Psychological Review
1999, Vol. 106, No. 3, 551–581

0033-295X/99/$3.00

Memory for Serial Order: A Network Model of the Phonological Loop and Its Timing

Neil Burgess
University College London

Graham J. Hitch
University of Lancaster

A connectionist model of human short-term memory is presented that extends the "phonological loop" (A. D. Baddeley, 1986) to encompass serial order and learning. Psychological and neuropsychological data motivate separate layers of lexical, timing, and input and output phonemic information. Connection weights between layers show Hebbian learning and decay over short and long time scales. At recall, the timing signal is rerun, phonemic information feeds back from output to input, and lexical nodes compete to be selected. The selected node then receives decaying inhibition. The model provides an explanatory mechanism for the phonological loop and for the effects of serial position, presentation modality, lexicality, grouping, and Hebb repetition. It makes new psychological and neuropsychological predictions and is a starting point for understanding the role of the phonological loop in vocabulary acquisition and for interpreting data from functional neuroimaging.

Temporally modulated information-processing and sequential behavior are essential to people's everyday functioning, such as in interpreting auditory stimuli and in using language. In a classic article, Lashley (1951) argued that serially ordered behavior cannot be explained in terms of "associative chaining," whereby each element in a sequence is associatively linked to its neighbors. However, subsequent attempts to model serial order computationally have continued to use associative chaining (e.g., Lewandowsky & Murdock, 1989), and there remains a paucity of adequate alternative accounts of people's ability to recognize, remember, reproduce, and become familiar with intricately timed and rhythmic sequences such as those in speech.

In contrast, there has been substantial progress in identifying the cognitive subsystems mediating the immediate serial recall of sequences of verbal items. There is a large body of evidence from psychological experiments suggesting that this form of memory is mediated by a "phonological loop" comprising a speech input store and a control process of subvocal articulation (Baddeley, 1986; Baddeley & Hitch, 1974). However, this simple model is incomplete. For example, it does not specify the processing of serial order information. This is a serious omission given the inherent seriality of verbal information. Yet another limitation of the phonological loop concept is that it ignores long-term memory (LTM). It therefore cannot address the effects of item familiarity (Hulme, Maughan, & Brown, 1991) or learning phenomena such as the Hebb repetition effect (Hebb, 1961). This further omission has been brought into sharp focus by recent evidence that the phonological loop is involved in vocabulary acquisition (Baddeley, Gathercole, & Papagno, 1998). In this article, we develop a connectionist model of the phonological loop that builds on the insights of the conceptual model and extends it to encompass serial ordering and learning.

According to the conceptual model of the loop, the speech input store acts as a temporary phonological buffer with rapidly decaying contents. External speech enters the store automatically, and subvocal rehearsal can be optionally used to refresh its contents. This decay and rehearsal system is likened to a closed "tape loop" of inner speech (Baddeley, 1986). The control process of subvocalization is also necessary for visual stimuli to gain access to the phonological store. This simple conceptual model accounts for a range of well-established features of immediate serial recall such as limited memory span, effects of the articulation rate and phonemic similarity of items, and interference caused by irrelevant articulation (see Baddeley, 1986, for a review). The main findings and their interpretations within the model are as follows:

> Word length: Recall declines as word length is increased, such that $s \approx 2r + c$, where s is span (the maximum number of items that can be correctly recalled), r measures speech rate, and c is a constant (Baddeley, Thomson, & Buchanan, 1975; Schweickert & Boruff, 1986). This is consistent with the idea that articulation speed limits the rate at which decaying phonological traces can be refreshed through rehearsal.
>
> Articulatory suppression and word length: Articulation of irrelevant items (i.e., "articulatory suppression") during list presentation impairs recall (Murray, 1968) and removes the effect of word length for visual lists (Baddeley et al., 1975). Suppression also removes the word length effect for auditory lists when it is continued throughout recall as well as presentation (Baddeley, Lewis, & Vallar, 1984). These

---

Neil Burgess, Institute of Cognitive Neuroscience, and Department of Anatomy and Developmental Biology, University College London, London, England; Graham J. Hitch, Department of Psychology, University of Lancaster, Lancaster, England.

This research was supported by the Biotechnology and Biological Sciences Research Council of the United Kingdom. We thank Rik Henson, Tim Shallice, George Houghton, Giuseppe Vallar, Tom Hartley, John O'Keefe, and Chris Yeo for useful discussions; Ian Neath and Stephan Lewandowsky for helpful suggestions during the review; and Rik Henson and George Houghton for reading an earlier version of this article. Neil Burgess is a Royal Society University Research Fellow.

Correspondence concerning this article should be addressed to Neil Burgess, Institute of Cognitive Neuroscience and Department of Anatomy and Developmental Biology, University College London, London WC1E 6BT, England. Electronic mail may be sent to n.burgess@ucl.ac.uk.

effects are as expected given the reasonable assumption that articulatory suppression prevents rehearsal.

> Phonemic similarity: Lists of phonemically similar items are less well recalled than lists of phonemically dissimilar items (Conrad & Hull, 1964). This is explained in terms of interference between the contents of the phonological store.
>
> Articulatory suppression and phonemic similarity: Articulatory suppression during presentation and recall removes the phonemic similarity effect for visual but not auditory lists (Baddeley et al., 1984). This is explained by making the assumption that rehearsal is necessary for visual stimuli to access the phonological store, whereas auditory stimuli access it directly.

The phonological loop has also proved useful in interpreting neuropsychological evidence that auditory–verbal short-term memory (STM) can be selectively impaired by neurological damage (Warrington & Shallice, 1969). Such patients typically show abnormally low memory span for auditory–verbal materials, combined with normal learning and LTM, normal span for auditory, nonverbal materials and relatively unimpaired memory span for visual–verbal materials (see Shallice & Vallar, 1990). A particularly pure case is illustrated by Patient P.V. (Vallar & Baddeley, 1984b; see also Case J.B. in Shallice & Butterworth, 1977, and other cases reviewed by Vallar & Papagno, 1995). P.V. showed no effects of phonemic similarity, word length, or articulatory suppression in the recall of visual lists (Vallar & Baddeley, 1984a), so, according to the phonological loop model, she was not using subvocal rehearsal. However, P.V.'s auditory–verbal span of only two items was far below normal performance under articulatory suppression. Her auditory span was affected by the phonemic similarity of the items but not by word length, suggesting that her deficit corresponded to a marked reduction in the capacity of the phonological store and an absence of rehearsal. However, as her articulation skills were normal, her failure to rehearse probably reflected a strategy choice, as little would be gained from rehearsal given a deficient phonological store (Baddeley, 1990). Subsequent cases have given further support to the idea of separable systems for rehearsal and phonological storage (Vallar, Di Betta, & Silveri, 1997). Thus, the simple two-component model of the phonological loop can provide an account of relatively subtle patterns of performance associated with neurological damage.

The two major components of the phonological loop have recently been localized in a functional neuroimaging study (Paulesu, Frith, & Frackowiak, 1993). Paulesu et al. differentiated between the subvocal rehearsal system and the phonological store by comparing the activation associated with making rhyming judgments with that associated with remembering lists of visually presented letters. Rhyming judgments are assumed to involve articulatory speech output systems, but not the phonological store: They are impaired by articulatory suppression but not by chewing (Burani, Vallar, & Bottini, 1991) and could be performed normally by Patient P.V. (Vallar & Baddeley, 1984b). Paulesu et al. identified the articulatory speech output component (i.e., "subvocal rehearsal") with Broca's area, and the "phonological store" with the left supramarginal gyrus (close, but slightly posterior to, Wernicke's area). This is roughly consistent with prior neuropsychological data (see, e.g., Shallice, 1988) that identifies Broca's area with speech output processes and Wernicke's area with speech input processes (note that the data attributing such complex functions to locations are less clear than this simple picture; see, e.g., Blumstein, 1995). Interestingly, a review of lesions in patients assumed to have deficits in the phonological store (including P.V. and J.B.) suggested its localization to the region of the left inferior parietal lobule and left superior temporal gyrus (including the supramarginal gyrus and Wernicke's area, respectively; Shallice & Vallar, 1990).

In summary, it can be seen that the phonological loop is capable of explaining not only various aspects of verbal STM performance in normal individuals but also neuropsychological evidence and the results of functional neuroimaging studies. We note that the model has also proved successful in explaining new findings, such as the observation that variations in memory span across languages and across stages of development can be predicted from differences in speech rates (Ellis & Hennely, 1980; Hitch & Halliday, 1983; Hulme, Thompson, Muir, & Lawrence, 1984; Stigler, Lee, & Stevenson, 1986). Although aspects of the phonological loop account can and have been questioned, such as whether the word length effect can occur in the absence of rehearsal (Brown & Hulme, 1995; Cowan et al., 1992) or temporal decay of information (Neath & Nairne, 1995), or whether interference from irrelevant speech affects a specifically phonological record,[1] it retains its strength as a general theoretical framework.

The concept of the phonological loop is thus ripe for translation to a more detailed implementation going beyond the level of the simple analogy of a tape loop of inner speech used by Baddeley (1986) to provide a more realistic serial order mechanism and to accommodate long-term learning effects. There is a vast amount of detailed information gathered from STM experiments with which to constrain modeling choices (see below). Given its recent popularity, and beguiling hints at analogy with the operation of brain processes, connectionist modeling seems to be a fruitful framework for such an implementation. In previous work, Burgess and Hitch (1992) began the process of extending the ideas of the phonological loop into a fully determined connectionist model of short-term serial processing of verbal items. This model proved successful as an explanatory account of performance in immediate serial recall as a function of list length, word length, and items' phonemic similarity. It also explained serial position effects and some of the more detailed aspects of the types and proportions of errors including serial order errors and phonemic substitution errors. The main features of this model have also provided a useful framework for a number of subsequent models of serial recall from STM (Brown, Preece, & Hulme, in press; Glasspool, 1995; Gupta & MacWhinney, 1997; Hartley & Houghton, 1996; Neumann, 1994; Page & Norris, 1998).

However, Burgess and Hitch's (1992) model could not show recency and phonemic similarity effects simultaneously, nor could it reproduce the characteristic sawtooth serial position curve for lists of alternately phonemically similar and dissimilar items (Baddeley, 1968; see Figure 11). These failings suggest that phonemic and positional information interact inappropriately within the

---

[1] The impairment of recall by exposure to irrelevant speech was explained by Salamé and Baddeley (1982) in terms of speech inputs gaining automatic access to the phonological store. However, given subsequent controversy (specifically, whether this effect is sensitive to phonemic similarity between the list and distractor items; Macken & Jones, 1995), we chose not to build it into the model. However, note that it would be straightforward to implement Salamé and Baddeley's account within our model.

model (for solutions to this problem, see Burgess, 1995; Burgess & Hitch, 1996; Henson, Norris, Page, & Baddeley, 1996). Furthermore, interactions involving presentation modality and articulatory suppression were not fully modeled, despite the fact that they were critical in developing the concept of the phonological loop (see Baddeley et al., 1984). No mechanism for learning over repeated trials was included in the model, so it could not address the effects of prior experience on immediate serial recall.

Here we present a connectionist model of immediate serial recall that maintains the basic structure of our 1992 model but significantly extends its range of application to address the relevant neuropsychological data, differences between auditory and visual presentation, the effects of articulatory suppression and temporally grouped presentation, as well as effects of learning such as item familiarity, Hebb repetition, rehearsal, and multiple recalls of a list. The model has also been extended to include the recognition memory task of judging whether a probe item has been recently presented. Our immediate aim is to present an integrated, mechanistic account of these numerous and diverse verbal STM phenomena. Our longer term aim is to provide a starting point for an analysis of related processes such as vocabulary acquisition.

The worth of a model of this sort depends on the ability to interpret its proposed mechanisms in terms of both the psychological processes and the brain regions that might implement them (see also Gupta & MacWhinney, 1997). To this end, we attempt to be consistent with data from neuropsychology and recent functional neuroimaging regarding the modularity of the systems proposed in the model. We implement the model within the connectionist framework of functional units that are consistent with the computational properties that could be expected of groups of neurons. The detailed predictions made by the model with regard to performance are then examined by numerical simulation. Analysis of an earlier, simplified version of this model has been presented elsewhere (Burgess, 1995; see also Burgess & Hitch, 1996, and Hitch, Burgess, Towse, & Culpin, 1996). We hope that the model may be used not only to understand, but also to predict, experimental observations from psychology, neuropsychology, and functional neuroimaging.

The Phonological Loop: Ideas and Data

Our starting point is that some aspects of STM for serially ordered lists of pronounceable stimuli are well described, at a crude level, by the conceptual model of the phonological loop (Baddeley, 1986). As described earlier, these include the effects of speech variables (i.e., word length, phonemic similarity, and articulatory suppression) and their interactions under auditory and visual presentation as well as selective neuropsychological impairments of auditory–verbal STM. Accordingly, the first goal of our computational model was to simulate the cluster of "speech-based" effects already explained by the conceptual model of the phonological loop. To extend the model to tackle the problems of serial order and learning, it is necessary to decide which phenomena are to be explained because different investigators place different values on the relative importance of the experimental findings. The following list describes what we consider to be generally accepted key findings on serial order and learning in the context of verbal STM. We have included recency and suffix effects because these index major differences between auditory and visual presentation.

Serial Ordering Effects

Serial position curve: The probability of correctly recalling an item as a function of its serial position in the list (the "serial position curve") has a bowed shape exhibiting both primacy and recency (see, e.g., Crowder, 1972).

Presentation modality, suffixes, and phonemic similarity: Auditory presentation of stimuli increases the recency effect (Crowder, 1972; Penney & Blackwood, 1989). An auditory suffix removes this effect (Crowder & Morton, 1969), but a visual (graphemic) suffix has little effect on lists presented in either modality (Hitch, 1975; Morton & Holloway, 1970). The modality effect is attenuated by using phonemically similar items (Watkins, Watkins, & Crowder, 1974).

Temporal grouping: Temporal grouping by the insertion of pauses during presentation typically improves performance by reducing order errors (Ryan, 1969). Grouping leads to "scalloped" serial position curves (showing primacy and recency within each group) and has a greater effect for auditory than visual lists, with extra recency within each group (Frankish, 1985). Grouping has the greatest effect when lists are presented in subgroups of three items (Ryan, 1969).

Errors: The majority of errors in immediate serial recall of familiar items are "order errors" rather than "item errors" (Aaronson, 1968; Bjork & Healy, 1974) and tend to involve transpositions of neighboring items (Healy, 1974) and phonemically similar items (Conrad, 1964).

Errors and phonemic similarity: There is an interaction between phonemic similarity and serial position in lists of alternating similar and dissimilar items, in which the extra phonemic similarity of half of the items affects only recall of those items (producing the sawtooth serial position curves described earlier; Baddeley, 1968).

Influences of Longer Term Processes, Familiarity, and Learning Effects

Item familiarity: Span increases with the familiarity of the items used. Specifically, item familiarity affects the c in $s \approx 2r + c$ (Hulme et al., 1991).

Nonword lexicality: Span for visually presented nonwords that sound like words (e.g., BRANE) is higher than span for those that do not (e.g., SLINT; Besner & Davelaar, 1982). This effect remains under articulatory suppression.

List repetition: Recall performance improves if a list has recently been previously presented and recalled (the Hebb repetition effect; Hebb, 1961). The Hebb effect is not obtained when the repeated list is presented in a different temporal grouping (Bower & Winzenz, 1969). If a list is rehearsed aloud, the number of extra errors committed on each rehearsal decreases with the number of repeated recalls (see Heffernan, 1991).

Serial order intrusions: Serial order intrusions occur in which an item from a previous list is recalled as an error at the same position in the current list (Conrad, 1960).

Summary

We have reviewed the subsystems involved in verbal STM, as captured by the concept of the phonological loop (Baddeley, 1986). This model explains a cluster of phenomena associated with the effects of speech variables, together with data on the selective neuropsychological impairment of auditory–verbal STM. In contrast, there has been a relative lack of progress in describing how the verbal STM system deals with the encoding and reproduction of serial order or how its operation is modified by long-term learning. A rationale was given for implementing a fully determined connectionist model of the phonological loop that would tackle all these issues in a single functional architecture. An initial such model (Burgess & Hitch, 1992) was fruitful in extending the conceptual model to handle serial order effects, but it failed to reproduce some human data and contained no mechanism for long-term learning or modality effects. The key empirical observations relating to these effects were identified and proposed as constraints on the architecture of the new model. In the next section we describe the rationale for the architecture of the new model and outline its manner of operation.

Rationale for the Functional Architecture of the Model

Modeling Background

The selection of items in serial order raises many questions in its own right, such as whether the items considered are subcomponents of a physical movement, words in a sentence, or merely digits in a telephone number. Until fairly recently, detailed modeling of short-term recall has been primarily mathematical. Some of these models do not address the problem of serial order (e.g., Nairne, 1990), or they do so in a limited way, falling short of a full model of serial recall (Drewnowski, 1980; Schweickert, 1993). Models that tackle the problem of order differ as to whether they make use of associative chaining between successive items. For example, Estes (1972) and Lee and Estes (1977) followed Lashley (1951) and rejected chaining. They proposed models in which serial order is encoded by item–position associations, with positional information represented in a multilevel hierarchy (see also Lee & Estes, 1981). In another model, Shiffrin and Cook (1978) proposed that associations are formed between items and a set of nodes with chaining between nodes and not the items. At the other extreme, Lewandowsky and Murdock (1989) proposed the theory of distributed associative memory (TODAM), a holographic model that relies entirely on interitem associative chaining. However, none of these models addresses the cluster of effects so successfully explained by the phonological loop concept, nor do they address the related neuropsychological data.

The subsequent development of connectionism has produced many models that exhibit various sorts of serial behavior. However, the majority of these have tended to rely on some form of chaining mechanism that associates previous states to successive states. Such models typically rely on delay lines to propagate activation in output or hidden layers at previous times back to the hidden or input layers (Elman, 1990; Jordan, 1986) or to the same layer (Kleinfeld, 1986; Sompolinsky & Kanter, 1986). However, as noted earlier, Lashley (1951) showed that associative chaining is a poor model for order errors. One of his arguments was that anticipation errors in skills such as speech and typing suggest that order is encoded by a superordinate schema in which parts of the sequence become active before they are output. Recent studies of immediate serial recall confirm that the chaining of item or phoneme representations would generate errors that are incompatible with the observed data (Henson et al., 1996). The same point has also been demonstrated by simulation (Burgess & Hitch, 1992). An alternative to chaining as a means of item selection is therefore indicated.

Item Selection in Serial Recall

A line of connectionist models has been inspired by Lashley's (1951) suggestion that items in a sequence are partially activated by some scanning mechanism before output and by evidence of response competition from various fields (see, e.g., Eriksen, Coles, Morris, & O'Hara, 1985). Such models assume that serial order is imposed on a set of items by a mechanism referred to as competitive queuing (CQ) by Houghton (1990). These models typically posit a set of nodes, each representing an item, with variable activation values, and a competitive interaction between nodes such that the most active is selected for output. This is followed by suppression of the selected node, which enables selection of the next item in the list (see also Grossberg, 1987; Rumelhart & Norman, 1982). Houghton and Hartley (1996) provided a much fuller discussion of the background to the computational issues behind this solution to the problem of serial order. Biologically plausible mechanisms that might underlie the suppression of the activation of the winning item were spelled out by Houghton, Glasspool, and Shallice (1994).

With specific regard to immediate serial recall of lists of items, we note that when lists contain no repeated items, repeats at recall are rare and tend to be separated by many intervening serial positions (Henson et al., 1996). On the other hand, when the presented lists include repeats, there is a tendency for the second item to be overlooked (the Ranschburg effect; Jahnke, 1969). Taken together, these effects point to a CQ-like selection process in which items compete to be chosen, are suppressed once chosen, and slowly recover from suppression. It is interesting to note that the serial behavior of TODAM owes much to the way that the calculation of the probability of correct recall implicitly suppresses recalled items (Lewandowsky & Murdock, 1989, p. 33; see also Lewandowsky & Li, 1994; Mewhort, Popham, & James, 1994; Nairne & Neath, 1994).

A Minimal Connectionist Model

The CQ mechanism alone would be sufficient to provide a minimal model of immediate serial recall simply by having CQ among items at presentation and once more during recall. As each item was presented, that item would receive the greatest activation, thus being selected by CQ at that time step and subsequently suppressed. The gradual recovery from suppression of these items would ensure that at the beginning of recall, the item that had been presented first would be the most active (it would have had the greatest time to recover), the second item in the list would be the second most active, and so on. Thus, selection, by CQ, of the most active item at each time step in recall would result in the selection of items in the correct order. Of course, there must also be errors to model the human data, and the particular characteristics of these error data indicate the presence of both phonological and temporal or contextual influences on item activations in addition to CQ among them. However, before discussing these influences, we consider the nature of item-level representations.

Item Representations

We assume that the selection of items in serial order takes place at a lexical level and involves the vocabulary from which list items are drawn. This is necessary to account for order errors in which an item from the experimental vocabulary is recalled in the wrong position (Conrad, 1965). The assumption that item representations are at a lexical level also provides a natural locus for modeling the effects of variables such as word frequency (Watkins, 1977) and item familiarity (Hulme et al., 1991) on recall.

Although the inclusion of lexical representations in a model of STM is relatively uncontroversial, the idea that serial ordering operates at this level raises a number of issues. First, encoding and outputting individual items also involves selection of their constituent elements in serial order. We assume that a separate selection process takes place at this sublexical level, but, for simplicity, we do not model these processes here. Hartley and Houghton (1996) showed how CQ could operate to select phonemes in serial order in a way that is compatible with the present type of model. Second, an alternative to CQ among item representations would be competition among place markers for the serial positions within a list. This would have some advantages in modeling memory for nonwords. However, given that place markers would have to be strongly associated to items in order to reproduce many of the phenomena of interest, parsimony favors selection among items (a discussion of how such a system handles nonwords is given below). A third issue is whether serial selection might also operate at a supralexical level. There is evidence that incoming sequences may be encoded in the form of familiar, learned "chunks" (Miller, 1956) that can vary in size to include units larger than individual words (Simon, 1974). Moreover, extensive training in digit span has been shown to increase span through increasing the extent to which sequences are recoded into higher order units that are themselves made up of chunks (Ericsson & Chase, 1982). A complete model of STM must include these important processes. However, for simplicity, consideration of such higher level chunking can be deferred while still allowing memory for such familiar lexical items as letters, digits, or words to be modeled. A separate consideration of chunking is supported by evidence that effects on chunking and phonological variables on immediate serial recall are independent of one another (Zhang & Simon, 1995).

Phonological Information: Neuropsychology and Phonemic Feedback

The main components of our model of the phonological loop, beyond CQ of items, are its subsystems for speech input and speech output. The assumption of separate input and output systems for speech is supported by independent evidence from studies of dual-task interference (Shallice, McLeod, & Lewis, 1985). Thus, there is minimal interference when a task involving phonological output is combined with one involving phonological input, but combining two input tasks or two output tasks is difficult. As noted earlier, neuropsychological data identifies speech output processes with Broca's area and speech input processes with Wernicke's area. Further arguments in favor of separate input and output systems were provided by Morton and Patterson (1980), Shallice and Vallar (1990), and Shallice (1988, chap. 7.1).

To model input and output systems, it is necessary to specify the way they are connected. Again, neuropsychological data are instructive. Along with Wernicke's and Broca's aphasias, a third category (conduction aphasia) has been identified and related to a disruption of the pathway between speech input and output systems. Conduction aphasic people can be divided into at least two subclasses: those unable to repeat lists of short words in the correct order (e.g., the classic STM patients mentioned earlier, J.B. and P.V.) and those unable to correctly repeat a single, long, low-frequency word (see, e.g., Goodglass & Kaplan, 1972). Note that these symptoms are not ascribed to simple deficits in speech comprehension or production (indeed, J.B. worked as a secretary), indicating unimpaired input and output phoneme layers. Thus, we are led to implement two pathways from input to output phonology, one pathway via an item representation and ordering mechanism, and a second direct pathway (see Shallice, 1988, and Shallice & Vallar, 1990, for further discussion).

Further evidence supporting the direct input–output connection comes from studies showing that speech inputs can be immediately repeated with little or no interference from various distractor tasks, suggesting a "privileged loop" of direct connections from input to output phonemic buffers (McLeod & Posner, 1984). The developmental ability to learn to associate a word with the same input and output phonology indicates that these direct connections should be reciprocal. Neuropsychological evidence supports this suggestion; for example, Patient M.K. (Howard & Franklin, 1987) did not appear to be able to convert output phonological representations into input phonological representations to access semantic form (thus, he could not give the meaning of a written pseudohomophone despite being able to pronounce it and being able to give the meaning of the corresponding written words). Thus, we model the connections between input and output phoneme layers by a unidirectional pathway via item representations and an ordering mechanism and by a bidirectional direct pathway (see Figure 1).

The pathway from output phonemes to input phonemes described earlier and the phonemic feedback it allows are important components of the model. This feedback enables visual material to activate input phoneme representations via subvocalization and enables subvocal rehearsal to reactivate input phoneme representations (thus allowing relearning of decaying connection weights; see below). It also allows selection of an item during output, initiated by activation feeding to item nodes from the context signal, to be influenced by phonemic feedback. This has important consequences in terms of accidental selection of phonemically similar items and the effect of an item's familiarity (with familiar items provoking stronger phonemic feedback). The model's phonemic feedback may be thought of as the mechanism by which one hears one's own voice when reading or rehearsing a list item.

The phonemic feedback can be thought of in terms of recurrent autoassociative networks or reverberatory loops, in that the

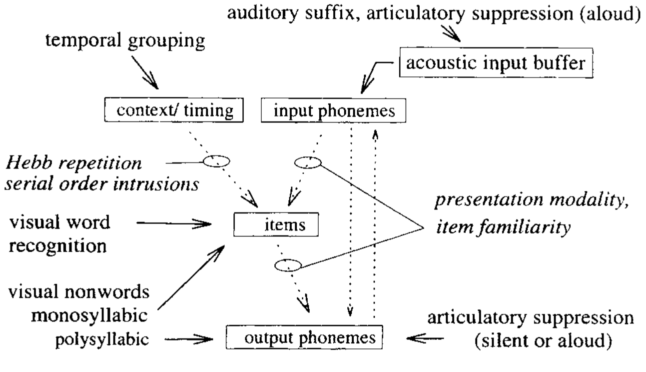

Figure 1. The structure of the model, indicating the routes by which visual and acoustic inputs reach the model and the locus at which articulatory suppression affects the performance.

output–input phoneme connections create a feedback loop involving the item and phoneme layers. Note, however, that our simulation of recall is such that activation passes around this loop only once before item selection occurs (see the section on implementation). Many models of memory have used autoassociative recurrent feedback (see the introduction and the section on modeling background). Its use in our model is related to the ideas of "clean-up" units (e.g., Plaut & Shallice, 1993) or "redintegration" (e.g., Hulme et al., 1997; Schweickert, 1993) in recall. A model of verbal STM that stresses reciprocal connections between corresponding input and output phoneme units was outlined by Campbell (1990). Some speech production models such as the one by Dell, Juliano, and Govindjee (1993) also make use of internal feedback from the output representation (but also see Levelt et al., 1991).

Positional Context, Timing, and Temporal Grouping

The idea that positional information is represented in STM is controversial. Inexact memory for position is demonstrated by the observation that incorrectly recalled items tend to be recalled near to their serial position during presentation. This is so regardless of whether they are order errors from the current list (Healy, 1974) or intrusions from the previous list in a block of trials (Conrad, 1960). However, the distribution of order errors is not a strong constraint because it can be simulated by models that do not represent positional information explicitly, as with the ordinal representation of the primacy model (Page & Norris, 1998). Conrad's evidence for serial order intrusions is more telling, as it is difficult to see how intrusions from a prior list can preserve positional information if such information is not represented in some way. Other evidence for the representation of positional information is the change in the pattern of order errors when a list is temporally grouped during presentation (Ryan, 1969). Under these conditions, incorrectly recalled items still tend to be recalled near their correct serial position, but superimposed on this is a tendency for errors to preserve within-group position, as in 387 154 → 357 184. Such a pattern is most directly explained by assuming representations of both between- and within-group positional information.

Burgess and Hitch (1992) modeled the coding of positional information as a "context" signal that functions similarly to attaching serial position "tags" to each item. To capture the idea that positional information is coded imprecisely, positional context was modeled as a moving window, such that there was similarity between the signals for successive positions. This view of context and context–item associations is similar to Lee and Estes's (1981) model, although in their case order errors were caused by perturbations to the positional associations of items rather than overlap between the representations of nearby positions themselves. Subsequently, Treisman, Cook, McNaish, and Crone (1994) proposed the existence of a central timing system composed of several sets of similar temporal oscillators, with each set giving rise to a distinct distribution of frequencies (see also Church & Broadbent, 1990). Treisman et al. found that exposing participants to a regular series of auditory clicks disturbed both their subjective estimation of time and electroencephalogram (EEG) results. They assumed that the clicks entrained subsets of oscillators whose natural frequency was close to the click rate, thus disturbing their usual contribution to both time estimation and the EEG. The rhythm of a regular series of verbal items may similarly entrain subsets of internal oscillators. If so, the outputs of such oscillators would overlap for successive items, thus providing a possible mechanism for our original idea of a moving window.

In recent work, Hitch et al. (1996) interpreted the context signal in terms of the oscillatory processes envisaged by Treisman et al. (1994) and showed how temporal grouping might have its effects by entraining multiple rhythms (see also Henson and Burgess, 1997). Modeling temporal grouping by modifying serial position tags has been proposed previously (see McNicol & Heathcote, 1986, pp. 76–77, for a brief review). However, those authors noted the need for a more complete model of order storage and retrieval in which to examine such mechanisms. McNicol and Heathcote (1986) suggested that the representation of serial position in grouped lists corresponds to points in a plane where one dimension reflects position within a list and the other dimension reflects position within a group. We adopted this approach previously (Hitch et al., 1996) and in the current model.

The characteristic shape of the serial position curve relating the probability of correct recall to serial position does not itself provide an argument for the representation of some correlate of serial position. Thus, many models that do not store positional information generate serial position curves that simulate human performance (e.g., TODAM; Lewandowsky & Murdock, 1989). However, positional models do generate a simple account of primacy and recency. That is, if item t is close to the beginning (or end) of the list, there will be fewer items preceding (or following) it and therefore fewer items associated to similar positional representations with which to be confused. For further discussion of the use of temporal, positional, and ordinal representations in models of STM, see Henson (1998b) and Henson and Burgess (1997).

Effects of Presentation Modality

So far, we have given priority to the architecture for dealing with the serial order of a sequence of verbal items with little regard for how they are presented. The most obvious effects of presentation modality concern recency (see, e.g., Crowder, 1972), so it would be reasonable to assume that auditory presentation induces a sharper positional coding in the context layer (see Henson, 1998b) or perhaps a strengthened association to the final context state. The former corresponds to the view that the time of occurrence is encoded more distinctively for auditory than visual stimuli (Gardiner, 1983; Glenberg & Swanson, 1986). However, this hypothesis has been applied mainly to LTM and has received somewhat mixed experimental support (Neath & Crowder, 1990; however, see Marks & Crowder, 1997). It is also possible to account for presentation modality at the level of coding by explicitly adding a set of modality-dependent features to the encoding scheme (see Nairne's, 1990, feature model). This approach has been successfully applied to the framework of our 1992 model by Neumann (1994). However, we are reluctant to add arbitrary extra features to the phonological coding of only auditory items, as there seems to be no a priori reason why one modality should benefit over the other in this way.

Given that the conceptual model assumes that auditory inputs access the phonological store directly, while visual inputs have first to be articulated (Salamé & Baddeley, 1982), we chose to focus on the routes by which different types of input reach the core part of our model. Campbell (1990) showed that a model with different access routes can explain modality differences such as recency and suffix effects and comparisons among spoken, lipread, and graphemic inputs. In Campbell's model, auditory, mouthed, or lipread items activate the input phoneme layer via a smaller capacity store specialized to speech stimuli. This store is similar to the precategorical acoustic store (Crowder & Morton, 1969) and holds a decaying trace of the item unless overwritten by a subsequent input. In contrast, visual (graphemic) items activate the output phoneme layer via a process of grapheme-to-phoneme conversion. Reciprocal connections between the two phoneme layers ensure that visual items also activate the input phoneme layer, but they do this without accessing the speech input store. We make assumptions similar to those of Campbell (1990), although in our model visual items activate item nodes directly, as we think that grapheme-to-phoneme conversion is probably not necessary in adult readers (with the exception of polysyllabic nonwords). Coltheart, Curtis, Atkins, and Haller (1993) and Zorzi, Houghton, and Butterworth (1998) provided further discussion of dual route models of reading, and Seidenberg and McClelland (1989) gave the alternative single-route view.

Learning

Tasks of immediate serial recall require short-term learning of the order of familiar items and of the sound of a novel item or nonword. Human performance in these tasks also shows longer term learning effects, such as improved performance on familiar sounding items (i.e., frequency effects among words and lexicality effects among nonwords) and on repeated lists (the Hebb repetition effect). Hebbian strengthening of connections over different time scales has biological plausibility from experiments showing simultaneous short- and long-term potentiation of synaptic pathways (see, e.g., Bliss & Collingridge, 1993; McNaughon, 1982), and related learning rules have been used in previous models of memory (Gardner-Medwin, 1989; Plaut & Shallice, 1993). We therefore incorporated "fast" and "slow" weights in the same connections to show long- and short-term learning effects, without altering the structure of the model. The decay of fast learning weights is also a simple way of modeling the temporal decay of information in the phonological loop, whereas the slowly varying weights are used to model the effects of item and list familiarity. The sounds of words and their pronunciations can be learned by strengthening connections between phonological and item representations, and serial order can be learned by strengthening connections between items and context–timing representations.

We make the simplifying assumptions that phonological representations are a primitive of the system (given that we are concerned with adult performance) and that semantic information can be ignored (because variables such as semantic similarity have relatively little effect on verbal STM; Baddeley, 1966; Saint-Aubin & Poirier, 1999). Indeed phonological and semantic memory appear to be dissociable, as evidenced by, for example, the tendency for J.B. and P.V. to remember the gist of a sentence rather than its surface structure.

A classic demonstration of long-term serial learning has been provided by the Hebb repetition effect, in which memory for a list presented for immediate serial recall improves with successive repetitions of the same list, provided that the repetitions are not too widely spaced (Hebb, 1961; Melton, 1963). Experimental evidence suggests that this effect and the effects of temporal grouping are strongly related as changing the grouping of the repeated list removes the Hebb effect (Bower & Winzenz, 1969). Modeling both effects in terms of context–item associations is consistent with this observation. Note that active recalls are necessary in producing the Hebb effect rather than repeated presentations alone (Cunningham, Healy, & Williams, 1984), so that learning in long-term weights (at least for context–item connections) should occur only at recall. However, this makes little practical difference to the simulation results we present below.

Errors

The individual errors committed by a participant are not deterministically controlled by experimental parameters, so we examine how the model breaks down in the face of "noise" or unreliability in its operation, as in our 1992 model (Burgess & Hitch, 1992). We model errors as occurring during recall following evidence that errors associated with the phonemic characteristics of the items do not occur during encoding (Conrad, 1964) but are associated with retrieval (Baddeley, 1968). That errors should occur primarily at the final output stage of the recall process (when items receive both context and phoneme inputs) is also indicated by the interaction between serial position and phonemic similarity on the likelihood of an error occurring. This is obvious from an inspection of serial position curves (see Drewnowski, 1980, or our Figure 11), in which the number of extra errors in lists of phonemically similar items, compared with lists of dissimilar items, depends on serial position in a way similar to the total number of errors. Because serial ordering is modeled at the item level, output is scored at the item level and within-items errors (such as misordering of an item's constituent phonemes) are not addressed.

The Core of the Model

The model that results from the considerations discussed earlier is summarized in this section. Its overall organization is shown in Figure 1, and detail of the architecture of the core component of the model is shown in Figure 2. Its main features are as follows: (a) Separate layers of nodes are used to represent items, input phonemes, output phonemes, and the context–timing signal. (b) The representation of phonemes and items is local and is such that one

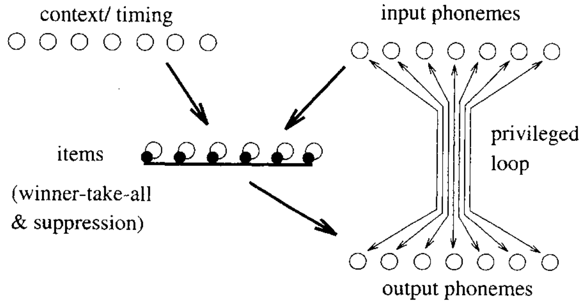

Figure 2. The basic architecture of the model. A single arrow between two layers represents connections from all units in one layer to all units in the other. Thick arrows are excitatory connections with short- and long-term plasticity; the full line with filled circles is a winner-take-all inhibitory interaction, with subsequent decaying inhibition of the winner; thin double-ended arrows are fixed one-to-one connections between corresponding input and output phoneme nodes.

node represents one item or phoneme. (c) Connections between the item layer and each of the phoneme layers are modifiable and represent the system's knowledge about the phonemic content of items. These connections enable phonemic inputs to activate item representations, as in list presentation, and allow item representations to drive speech output, as in recall. (d) Connections between the context and item layers are also modifiable. (e) Modifiable connections have a component that decays with time. (f) In addition, there are hard-wired (i.e., nonmodifiable) one-to-one connections between corresponding input and output phoneme nodes such that activation in one is automatically reproduced in the other. (g) CQ is implemented in the item layer. (h) The activation of output phoneme nodes in vocal or subvocal rehearsal is the rate limiting process, as items with more phonemes take longer to output.

The basic operation of the model common to acoustic or visual presentation can be briefly summarized as follows: Presentation of an item causes item nodes and input and output phoneme nodes to be temporarily activated via their interconnections. Competition between item nodes ensures that only one remains active (the "winner"). Simultaneously, a subset of nodes in the context layer is activated (these nodes correspond to the current state of the moving window; see Figure 3A and Burgess & Hitch, 1992). Modifiable connections linking currently active nodes in adjacent layers are strengthened by Hebbian learning. This entails strengthening associations between the context signal and the item selected by CQ and between this item and its input and output phonology. Before presentation of the next item, the winning item is inhibited, the strength of modifiable connections and inhibitions to previous items undergo decay, and the next context state is activated.

During recall, the slowly varying context–timing signal is repeated and evolves through the same successive states as during presentation. For each such state, item nodes receive activation from the context layer, and the most active "wins" and provides phonemic feedback via the output and input phoneme layers. These inputs combine additively, and the most strongly activated item node at this stage is output. The final output selection process is taken to be unreliable, modeled by adding zero mean Gaussian noise to the input to item nodes. Thus, the fundamental reason for errors at the item level is noisy selection of items receiving input from both the context–timing signal and the input phoneme nodes, as in our 1992 model (Burgess & Hitch, 1992). For example, serial order effects result from CQ of the item nodes, such that the activation of a selected node is subjected to decaying inhibition. The model incorporates a basic form of rehearsal that corresponds to repeating the recall phase. Note that errors occur solely during the final selection of output: The other processes in the model are assumed to be error free (e.g., the repetition of the context–timing signal,[2] and the suppression of items and their recovery from suppression).

There is a simple mapping onto the conceptual model of the phonological loop insofar as the input phoneme layer corresponds to the phonological store and the output selection mechanism corresponds to part of the speech production (articulation) system. The idea of subvocal rehearsal or "inner speech" corresponds to the use of output phoneme activations to reactivate input phoneme nodes, providing phonemic feedback to the items; this process is supported by the loop of connections from input phonemes to items to output phonemes and back to input phonemes. Temporal decay of the short-term connection weights learned at presentation corresponds to the decay of information in the phonological store and provides word length effects when noise is introduced into the system. Decay can be offset by relearning the short-term connection weights every rehearsal.

Presentation modality is assumed to affect the routes by which external stimuli gain access to the system. Auditory stimuli do so by activating the input phoneme layer first, whereas visual (graphemic) stimuli activate the item layer first. In addition, there is an acoustic input buffer for auditory stimuli.

As well as providing a quantitative mechanism for the operation of a phonological loop, the model contains several additional features. The most notable are (a) the context–timing signal, which plays a key role in generating effects of serial position, Hebb repetition, temporal grouping, and serial order intrusions; (b) the CQ mechanism, critical for generating serial output and giving the correct behavior for repeated items in recall; (c) separation of input and output phonology as consistent with neuropsychological and experimental data; and (d) a mechanism for including the effects of long-term learning.

In the remainder of this section, we describe the core aspects of the model in detail. In the section on extending the core model, we describe how the model simulates the effects of presentation modality, articulatory suppression, temporally grouped presentation, and recognition memory judgments. The effects of lesioning the model are described in the section on performance.

Learning: Long- and Short-Term Connection Weights and Item and List Familiarity

All nodes are either active (activation value 1) or inactive (activation value 0). Simulations proceed in steps of the presentation, recall, or rehearsal of one item, corresponding to *wl* seconds (the item's spoken word length; see the section on implementation). Aside from the fixed one-to-one connections between input and output phoneme nodes, all of the connections in the model have a weight that shows both long and short-term Hebbian plasticity. The long-term weight is denoted by $W^l_{ij}(t)$, where t is the time since the start of a list and i and j denote the nodes at either end of the connection (which goes from node j in the layer above to node i in the layer below). $W^l_{ij}(t)$ is slowly incremental and saturates (i.e., it cannot exceed a maximum or minimum value in magnitude). The short-term weight, $W^s_{ij}(t)$, is capable of large increases to a maximum value given by $W^s_{\max}$ but decays by a factor Δ per second or Δ^{wl} per simulation step; that is,

$$W^s_{ij}(t + wl) = \Delta^{wl} W^s_{ij}(t),$$

where $\Delta < 1$. The net weight of a connection is the sum of the short- and long-term components:

$$W_{ij}(t) = W^l_{ij}(t) + W^s_{ij}(t).$$

As well as decay of the short-term connection weights, the model's modifiable connections change from one step to the next as a result of the following learning rules:

---

[2] An explicit mechanism for the generation of the context–timing signal might involve a set of oscillators in which patterns of activation reoccur naturally (see Henson & Burgess, 1997; Hitch et al., 1996; Brown, Preece, & Hulme, in press).

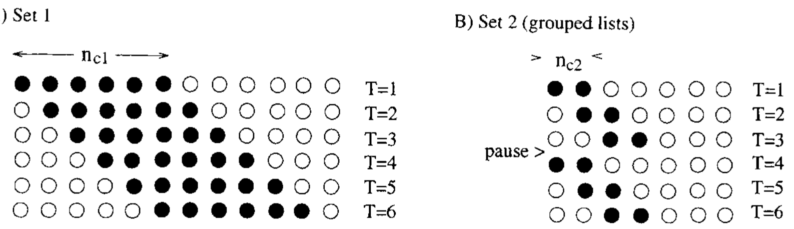

Figure 3. The context–timing signal: Filled circles are active nodes; empty circles are inactive nodes, the index *T* represents serial position. A: The representation for ungrouped lists shows a moving window of activation that is reset at the start of recall (adapted from Burgess, 1995; see also Burgess & Hitch, 1992). B: The second set of context nodes (Set 2) recruited by temporally structured presentation of a list. Temporal grouping of presentation by insertion of a pause between $T = 3$ and $T = 4$ causes the moving window of activation of these nodes to be reset. Adapted from Hitch et al., 1996.

$$W^s_{ij}(t + wl) = \begin{cases} W^s_{\max}[a_j(t) - \theta]a_i(t) & \text{if } a_i(t) > 0; \\ W^s_{ij}(t) & \text{otherwise,} \end{cases} \tag{1}$$

$$W^l_{ij}(t + wl) = \begin{cases} W^l_{ij}(t) + \beta[a_j(t) - \theta]a_i(t) & \text{if } a_i(t) > 0; \\ W^l_{ij}(t) & \text{otherwise,} \end{cases} \tag{2}$$

where $a_i(t)$ are the activation values of nodes *i* and *j*, respectively, and ϵ is a small positive constant determining the size of changes to long-term connection weights. The maximum and minimum values for long-term connection weights are $(1 - \theta)W^l_{\max}$ and $-\theta W^l_{\max}$, so that $W^l_{\max}$ is the range of possible values. For all but the connections between the input phoneme and item layers, the value of θ is zero, so that both short- and long-term connection weights are always positive or zero. For the input phoneme–item connections, $\theta = 0.5$ to allow balanced positive and negative connection weight values. These are necessary to ensure the specificity of an item's association to its phonemes: Learning brings not only a stronger positive association with its own phonemes but also a stronger negative association with other phonemes. As we show, this is important in modeling the effects of presentation modality and recall of lists of mixed words and nonwords. The above type of learning rule, in which the level of presynaptic activity determines whether the synaptic weight increases or decreases, has been proposed as a model for synaptic plasticity (see, e.g., Bienenstock, Cooper, & Munro, 1982).

All connection weights must be normalized by the number of active nodes in the layer from which they project, so that information represented by many active nodes (e.g., a word containing many phonemes) does not cause greater activation in the next layer than one represented by few active nodes. Thus, the maximum size of short-term connection weights is controlled by $W^s_{\max} = 1/n_c$ for context–item connections, $W^s_{\max} = 1/n_p$ for input phoneme–item connections, and $W^s_{\max} = 1$ for item–output phoneme connections, where n_p is the number of phonemes in the item and n_c is the number of active context nodes in the moving window (see n_{c1} in Figure 3A). The size of long-term connection weights is scaled similarly, with

$$\beta = \kappa\epsilon, \tag{3}$$

where $\kappa = 1/n_c$ for context–item connections, $\kappa = 1/n_p$ for input phoneme–item connections, $\kappa = 1.0$ for item–output phoneme connections, leaving ϵ to determine the size of long-term weight changes independently of the number of active nodes in a representation.

Item Selection

The "winner-take-all" component of CQ is implemented directly on the items (i.e., there is no extra "competitive filter" layer as in Burgess & Hitch, 1992). All the item nodes are compared and the one with the greatest input is given an activation of 1.0. All other item nodes are given an activation of zero before any modification of connection weights or spreading of activation to other layers occurs. Thus, only the winning item is associated to the currently active context and phoneme nodes. At the next time step, the activation of the winning item is suppressed and gradually recovers thereafter because of a decaying inhibition process.

The suppression of item node *i* (at time *t*) is modeled as a decaying inhibition, $I_i(t) < 0$, and forms one of the inputs to that node. $I_i(t)$ is set to an initial value of –2.0 and decreases by the decay factor Δ over each time step. Accordingly, the net input to item *i* at time *t* is

$$h_i(t) = E_i(t) + I_i(t) + \eta_i, \tag{4}$$

where $I_i(t) < 0$ is the decaying inhibition imposed following an item's selection at presentation or output, η_i is a noise term added at output only (η_i is a Gaussian random variable with zero mean), and $E_i(t)$ is the excitatory input to the item. During auditory presentation excitation comes from the input phoneme layer, during visual presentation it comes directly from visual recognition of the item, and during the output stage it comes from the context and phoneme layers together. Thus,

$$E_i(t) = \begin{cases} \sum_j W_{ij}(t)b_j(t) & \text{during auditory presentation,} \\ 1.0 & \text{during visual presentation,} \\ \sum_j W_{ij}(t)b_j(t) + \sum_k W'_{ik}(t)c_k(t) & \text{during output.} \end{cases} \tag{5}$$

where W_{ij} is the connection weight from input phoneme node *j* to item node *i*, which has activation b_j, and W'_{ik} is the connection weight from context node *k*, which has activation c_k. Note that the net input may be negative because of inhibition, I_i, and negative connection weights from nonconstituent phonemes.

Implementation

The algorithm used to implement recall of a list of *n* items follows. For simplicity, the model does not address differences in presentation and recall rates because each time step has a duration equal to the time taken to present or recall an item (i.e., the spoken word length; *wl*). This restriction corresponds to presentation being at the same rate as the "inner speech" controlling rehearsal and recall (i.e., at the participant's speech rate). The index T increases by 1 per time step and corresponds to *wl* seconds.

Presentation. Steps that apply specifically to acoustic or visual presentation are shown in parentheses and preceded by A or V, respectively:

0. Set activations, inhibitions, and short-term weights to zero, T = 1.
1. Set the context layer to the state for time T.
2. Suppress the last selected item (i.e., set the inhibitory input to it to –2).
3. Input item T (A = activate corresponding input phoneme nodes and spread activation to output phoneme nodes; V = activate corresponding item node and spread activation to output and input phoneme nodes).
4. Select the item node with the greatest input.
5. Learning: Increment all modifiable connection weights according to Equations 1 and 2.
6. Decay: Multiply short-term connection weights and inhibitions by a factor Δ^{wl}.
7. T → T + 1; if T < *n*, go to 1; otherwise go to recall (A = further increment short-term item–phoneme and phoneme–item connections).

Recall. The steps are the same for recall and rehearsal and are independent of the modality of presentation:

0. Set T = 1.
1. Set the context layer to state T.
2. Inhibit the last selected item.
3. Set all phoneme activations to zero (there is no external input).
4. Select the item node with the greatest net input.
5. Output: Activate output phonemes via item–phoneme connections, activate the corresponding input phonemes, and select the item node with the greatest net input in the presence of noise.
6. Learning, as seen earlier.
7. Decay, as seen earlier.
8. T → T + 1; if T < *n*, go to 1; otherwise, stop.

Thus, both presentation and recall involve updating the context layer, inhibiting the most recently selected item node and selecting the most active item node. In both cases, modifiable connections undergo learning and short-term connection weights and inhibitions undergo decay. The major difference between presentation and recall is that the activation of item and phoneme nodes is initiated by external inputs during presentation and by the context signal during recall.

Parameter Values for Simulations

The parameter values below were used in all of the simulations that follow. Better fits to individual experiments could be provided by fine tuning of parameters, such as the level of noise or the level of phonemic similarity. However, given the variability of the experimental data itself and the many minor differences of procedure that are not reflected in the model, our aim was simply to demonstrate the overall compatibility of our model with the psychological effects of interest. All results shown below are the average of 5,000 simulated trials, except that simulations of transposition gradients and confusion matrices used 10,000 trials to better sample the relatively low incidence of transposition errors.

Parameterization of the model. The occurrence of errors depends on the rate of decay in the model (parameterized by Δ) and the level of noise (parameterized by σ, the standard deviation of the Gaussian random variable, η_i, in Equation 4). These are the two main parameters of the model and were chosen to be $\Delta = 0.75$ and $\sigma = 0.3$, so that digit span is approximately 7 (see Table 1). The size of the inhibition $I_i(t)$ immediately following item selection is not critical, and the value 2.0 was chosen to be enough to prevent successive recalls of an item. The precise number of context nodes active at any one time of (n_c) does not have a great effect on the appearance of serial position curves, with little difference for values between 3 and 9. For consistency, we took $n_c = 6$ (as shown in Figure 3A).

At the start of a simulation, the values of long-term connection weights are set to model the effect of any learning prior to the simulated trial. For this we introduce two parameters: contextual familiarity (F_c) and phonemic familiarity (F_p). Simulation of an item having recently been presented or recalled in a given serial position prior to the simulated trial is achieved by activating the relevant item and context nodes and making one application of the learning rule with $\epsilon = F_c$ (Equations 2 and 3). The prior association of an item node with its constituent phonemes is achieved by activating the relevant item and phoneme nodes and making one application of the learning rule with $\epsilon = F_p$ (Equations 2 and 3). In simulations of the Hebb repetition effect, and in examining position-specific intrusions from previous lists, an item's contextual familiarity is $F_c = 0.3$. In simulations of four successive rehearsals of a list of unfamiliar items (see Figure 17), $\epsilon = 0.1$ (see Equations 2 and 3) for context–item connections (so that an item's contextual familiarity increases from 0.0 to 0.4; i.e., slightly more than for items recalled in previous lists). Otherwise, $F_c = 0.0$ and $\epsilon = 0.0$ for context–item connections.

Parameterization of the stimuli. The spoken word length for letters of the alphabet was taken to be *wl* = 0.4 s and slightly less (*wl* = 0.35 s) for digits. "Words" are composed of varying numbers of phonemes and have varying spoken duration as indicated. Phonemically dissimilar items have no phonemes in common, as in the list of letters, *q, r, n, o, h,* and *y*. Phonemically similar items in a list all have one common phoneme as in the list of letters *b, c, d, g, p,* and *t* or more than one common phoneme. Simulations use either letters or digits or artificial stimuli (with parameters as indicated). When simulations involve letters of the alphabet or digits, the actual phonemic composition of each item is used.

Unfamiliar words or nonwords are modeled by phoneme–item familiarity, $F_p = 0.05$, whereas familiar items (known words or pseudohomophones) are modeled by $F_p = 0.2$. The values of F_p were chosen to provide the magnitude of the experimentally observed effect of item familiarity (see Figure 5). For simulation of four successive rehearsals of a list of unfamiliar items (see Figure 17), $\epsilon = 0.05$, so that an item's phonemic familiarity increases from 0.05 to a maximum of 0.25 which is slightly more than for familiar words that have not just been repeated many times.

Extension of the Core Model

In this section we consider extending the core model to include effects of presentation modality, articulatory suppression, and temporal grouping of presentation. Figure 1 indicates where these effects are located within the model. We also present a model of recognition memory in which a probe item is judged according to whether it was present in a recently presented list.

Effects of Presentation Modality

As discussed earlier, we suppose that speech inputs (acoustic or lipread) and visually presented inputs arrive at the input phoneme layer via different routes. Thus, spoken items activate input phoneme nodes directly and therefore item and output phoneme nodes via the two sets of connections from the input phoneme layer. In contrast, visually presented words activate item nodes directly, then output phoneme nodes via the connections from item nodes and then input phoneme nodes via the connections from output phoneme nodes.

Acoustic inputs arrive via an auditory input buffer and are maintained there until the next item arrives. This causes continued activation of the input phoneme nodes (and hence also of the output phoneme nodes) at the end of a list or during pauses when the rhythm of the preceding items would indicate that the next item was due. The effect of this is to strengthen the short-term association between the last presented item and its phonemes. This is modeled by doubling the magnitude of the short-term phoneme–item and item–phoneme connections of items at the end of a list or the end of a group in a grouped list, so that $W^s_{\max} = 2.0/n_p$ and 2.0, respectively, in Equation 1.

The effect of an auditory suffix is to clear the buffered phonemic information from the previous item, resulting in performance identical to visual presentation (i.e., there is no extra increase in the connections to and from the previous item node). A visual suffix would not clear the auditory input buffer in this way. However, the exact effect of a visual suffix depends on processes such as attention that are beyond the scope of this model. The model suggests ways of handling lipread items (e.g., these excite input phonemes directly but without activating the acoustic input buffer) and the effects of delaying an auditory suffix (e.g., extra learning may occur before the buffer is cleared). However, because these are not central to our current concerns, we have not chosen to examine them in simulations.

Articulatory Suppression

Articulatory suppression is modeled as occupying the output phoneme layer. To model the effect of suppression in detail would require a more detailed simulation of the output phoneme layer than we have so far developed (e.g., including a mechanism for the ordering of phonemes within an item). However, within the scope of our simulation, articulatory suppression has two consequences.

Because the time taken for the activation of the output phoneme layer is the rate-limiting step in recall, it represents the time taken to subvocally rehearse or output an item. The effect of suppression during recall is to break the relationship between the rate of recall and the spoken word length of the items in the list. This is simply modeled by taking the time taken for the rehearsal or output of an item to be equal to a constant. We take this constant time to be 1.0 s to illustrate the effect in simulations. This will give performance independent of word length, as observed in experimental data (Baddeley et al., 1984).

The irrelevant activation of the output phoneme layer during suppression also causes phonemic feedback to the input phoneme layer. Visual items also use this pathway to activate the input phoneme nodes. For visual items, the interference caused by the extra activation of irrelevant phonemes in this layer impairs the learning of the input phoneme–item associations corresponding to the list item. This occurs naturally in the model, as the strength of these connections is normalized by the number of active phonemes in the input phoneme layer (see Equation 1 and the next paragraph). To simulate this, learning is simply blocked in the connections from input phonemes to items. By contrast, acoustic presentation activates the input phonemes directly from the auditory buffer. This input is assumed to have automatic control of the input phoneme activations, so the learning of input phoneme–item connections for acoustic items is unimpaired by suppression in the model.

Note that even for auditory items, relearning of the input phoneme–item association during rehearsal is blocked by suppression, as rehearsal relies on phonemic feedback to activate the input phoneme layer. Finally, suppressing aloud has a third consequence: filling the auditory input buffer and so removing the auditory modality effect in the same way that an auditory suffix would.

Our model of suppression differs slightly from the conceptual model of the phonological loop, in which visual inputs are said to be entirely prevented from entering the loop, whereas acoustic stimuli have automatic access (see, e.g., Baddeley, 1986). This view seems to overstate the power of articulatory suppression, as, despite interfering, reading or talking to oneself subvocally is possible under suppression (try it). In our model the only process terminally impaired by suppression is the reading of polysyllabic nonwords (which have to enter the system via the output phoneme nodes; see above).

Temporal Grouping

The effect of temporal grouping at presentation is modeled as recruiting a second set of context nodes (Set 2) that reset after every pause and at the start of recall, as opposed to the first set of context nodes (Set 1), which reset at the start of recall only (see Figure 3B and Hitch et al., 1996). During presentation, associations to items are formed from both sets of context nodes. At recall, the input from the second set of nodes multiplicatively modulates the signal from the first. Ignoring decay, the size of the short-term weights between an item node and the context nodes active at the time of the item's presentation is $1/n_c$, where n_c is the number of active nodes in the particular set of context nodes. Thus, at time t during recall, the strength of the input to item node i from a set of context nodes is equal to the fraction of currently active context nodes that were active during the item's presentation. We denote this fraction by $F^1_i(t)$ for the context nodes in Set 1 and $F^2_i(t)$ for those in Set 2. The compound timing signal reaching item i from the two sets is the product of the two signals, $F^1_i(t) \times F^2_i(t)$. If the list contains n items of word length wl, then the value of this input is reduced by a factor $\Delta^{n \times wl}$, where $n \times wl$ is the time between the presentation and recall of an item, to give the usual effect of decay of the short-term connection weights.

The effect of the second set of context nodes is to reduce the input to an item from the first set of context nodes at all serial positions during recall except those with the same within-group position as the item's presentation. This is because the pattern of activation in the second set of context nodes advances with each time step and resets at the end of each group (i.e., the pattern of activation depends only on the within-group position). If an item i is presented at a certain within-group position, then whenever recall reaches the same within-group position (at time t' say) the activation of context nodes in Set 2 will be the same as at presentation of item i, so $F_i^2(t') = 1$ and the input to the item will again be identical to that received in the ungrouped case: $F_i^1(t')$. However, at other within-group positions during recall (at time t'' say), the item's input $F_i^1(t) \times F_i^2(t)$ will be much reduced as $F_i^2(t'') < 1$. Thus, this model of temporal grouping will produce a decrease in the likelihood of an item being incorrectly selected at any recall position other than one with the correct within-group position (i.e., there will be an increase in overall performance and a relative increase in the number of confusions between items with the same within-group position). We also expect to see "scalloping" of the serial position curve with primacy and recency within each group for the same reason that primacy and recency are seen in ungrouped lists.

Recognition Memory

The recency of presentation (or recall) of an item is encoded in the extent to which the short-term connection weights associated with it have decayed. Thus, recognition memory judgments, such as deciding whether a probe item was present in a recently presented list (Sternberg, 1969), can be modeled without recall of the entire list in order by testing the extent of this decay. We model this as follows: When a probe item is presented, the corresponding item node is activated (activation 1.0); all other item nodes and the context nodes are set to zero activation, and the usual noisy choice of the most active item is performed following the activation of items by the phonemic feedback from the probe item. If the chosen item matches the probe item, then it is accepted as having been recently presented. The more recently an item has been presented, the stronger its short-term phoneme–item and item–phoneme connection weights will be and the greater will be the activation caused by phonemic feedback. If phonemic feedback is too weak, noise will dominate item activations and the probe is unlikely to be accepted as having been recently presented.

Note that we could model recognition memory judgments in terms of serial recall of the presented list and a comparison between each presented item and the probe item in accordance with Sternberg's (1969) interpretation of how the task is solved. The results of such a model are easy to infer from the simulations in the rest of this article. However, more recent analysis indicates that recognition memory judgments tend to be solved by reference to the decaying strength of item representations (see McElree & Dosher, 1989; Monsell, 1978), consistent with the finding that error rates and reaction times decrease as a function of the probe's serial position.

The model of using the strength of the item's short-term association with phonemic information is consistent with the observation of a phonemic similarity effect and an effect of articulatory suppression during these tasks but no similarity effect under suppression (Henson, Hartley, Hitch, & Burgess, 1999). This model also allows possible false acceptance of probe items from a recently presented previous list (see Monsell, 1978). It predicts a greater acceptance rate for words than nonwords, which have weaker phonemic feedback (due to smaller long-term phoneme–item associations).

Performance

This section is divided into two subsections concerning evaluation of the model at successive levels of detail. We initially consider performance in terms of overall capacity (principally item span), including simulation of articulatory suppression, lesions to the network, and the various experimental conditions reviewed earlier. In the second subsection we consider performance in more detail, such as the effects of serial position and the distributions of errors. Simulations relating to a particular experimental condition may be found in one or both of these subsections according to the level of detail that best shows the effect.

Capacity

"Memory span" or "item span" is measured as the list length for which the probability of correct recall would be 0.5. When this is not an integer, the probability of correct recall of the list lengths above and below span is linearly interpolated to give a measure of span. The model's item span, and the effect on it of articulatory suppression, is shown in Table 1 as a function of presentation modality, phonemic similarity, word length, item familiarity, prior list presentations (Hebb repetition), and temporal grouping. Item span was chosen as a measure of performance here and for assessing the effects of lesions (see the section on lesions and neuropsychological data) because it is not affected by floor or ceiling effects.

Table 1

Simulated Effect of Articulatory Suppression on Item Span

		List items						
		Familiar					Unfamiliar	
Condition	Modality	Dis	Sim	Dis & long	Hebb	Grouped	Dis	Sim
Control	Acoustic	7.1	4.8	4.9	7.3	9.1	5.7	4.0
Control	Visual	6.8	4.6	4.6	7.1	8.4	5.4	3.7
Suppression	Acoustic	4.0	3.2	4.0	4.1	4.7	3.3	2.9
Suppression	Visual	2.9	2.6	2.9	2.9	4.1	2.4	2.4

Note. Interaction of articulatory suppression with presentation modality, item familiarity, phonemic similarity, word length, Hebb repetition, and temporal grouping. Familiar list items were phonemically dissimilar (Dis) letters; phonemically similar (Sim) letters; dissimilar items twice as long as letters (Dis & long; duration = 0.8 s); items from lists of dissimilar letters previously repeated three times ($F_c = 0.3$ Hebb); and dissimilar letters temporally grouped in threes in presentation (Grouped). Unfamiliar items had $F_p = 0.05$ and resembled letters in other respects.

Table 2
Experimental and Simulated Effect of Articulatory Suppression on Percentage of Error

		List items							
		Familiar						Unfamiliar	
Condition	Modality	Dis	Sim	Short	Long	Ungrouped	Grouped	Dis	Sim
Control	Acoustic	14^{a} (10)	26^{a} (32)	22^{c} (19)	39^{c} (44)	36^{d} (33)	11^{d} (14)		
Control	Visual	10^{b} (12)	25^{b} (35)			64^{d} (35)	45^{d} (21)	29^{b} (27)	40^{b} (45)
Suppression	Acoustic	29^{a} (56)	47^{a} (64)	53^{c} (60)	58^{c} (61)	56^{d} (78)	32^{d} (73)		
Suppression	Visual	53^{b} (63)	52^{b} (67)			76^{d} (80)	66^{d} (75)	62^{b} (68)	63^{b} (69)

Note. Interaction of articulatory suppression with presentation modality, item familiarity, phonemic similarity, word length, and temporal grouping. Experimental data are shown as percentage of items in error in immediate recall of a list, and data from corresponding simulations are shown in parentheses. Data are organized according to whether list items were familiar or unfamiliar, phonemically similar (Sim) or dissimilar (Dis), short or long, and temporally grouped or ungrouped.

[a] Lists of six letters, adapted from Baddeley, Lewis, and Vallar (1984). [b] Lists of four-letter strings (phonemically similar and dissimilar pseudohomophones and nonwords, simulated as familiar and unfamiliar four-phoneme items of length 0.4 s with two or zero phonemes in common), adapted from Besner and Davelaar (1982). [c] Lists of five short or long words (simulated as four-phoneme familiar items of length 0.5 or 0.8 s with one phoneme in common). Data from "Exploring the Articulatory Loop," by A. D. Baddeley, V. J. Lewis, & G. Vallar, 1984, *Quarterly Journal of Experimental Psychology, 36A,* p. 239. Copyright 1984 by the Experimental Psychology Society. Adapted with permission. [d] Lists of nine dissimilar letters ungrouped or grouped in 3s, adapted from Hitch et al. (1996).

The simulations show the correct qualitative effects of phonemic similarity, word length, item familiarity, Hebb repetition, and temporal grouping. Performance is slightly better for acoustic items than visual ones across all conditions and significantly better for temporally grouped acoustic items than temporally grouped visual ones. The simulations also show the correct pattern of effects for articulatory suppression. Articulatory suppression reduces performance across all conditions, and for both acoustic and visual items the effects of item familiarity, Hebb repetition, and temporal grouping remain under suppression, whereas the effect of word length does not. There is a significant effect of phonemic similarity under suppression for acoustic items, but a much reduced one for visual items. This pattern of results closely matches the experimental data, when it exists (see Table 2). Table 2 shows experimental data expressed as the average percentage of error in the recall of an item in immediate serial recall of lists of a fixed length. Data from the corresponding simulations are shown in parentheses for comparison and show good quantitative agreement, with the exception of a slightly too strong effect of suppression overall and too good recall of long visual lists, grouped or otherwise (see the section on limitations and future work for more on this problem).

Simulations showing improved recall of familiar sounding items match the data from Besner and Davelaar (1982) and also demonstrate the effect of "support," in which nonwords that are phonemically similar to known words are better recalled than those that are not (Gathercole, Willis, Emslie, & Baddeley, 1991). In the model, the recall of items that are phonemically similar to familiar items is supported by making use of the long-term phoneme–item and item–phoneme connections of the node representing the familiar item.

That the Hebb repetition effect (although small) should remain under suppression is a prediction. The only noticeable difference in the patterns of simulated and experimental results is the slight effect of phonemic similarity for familiar visual items under suppression. This is due to the long-term component of item–phoneme connections, as can be seen from the absence of this effect for unfamiliar visual items under suppression. We discuss this further in the section on limitations and future work.

List length, word length, and articulation rate. The model's capacity can also be examined in terms of how performance declines as lists of more items, or items of longer spoken duration, are used. The probability of correctly recalling a list of acoustically

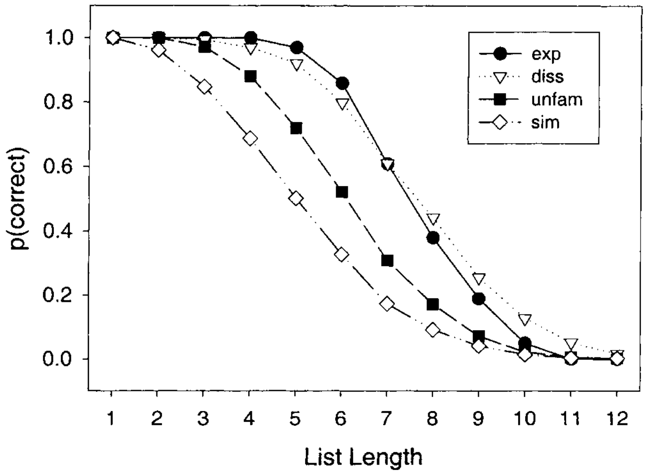

Figure 4. The probability of correctly recalling an entire list in order as a function of the number of items in the list. Curves show recall for the following simulated acoustic items (of the same duration as simulated digits, i.e., 0.35 s): phonemically dissimilar items (diss); phonemically similar items (sim; having one phoneme out of two in common); and unfamiliar items (unfam; having F_p = 0.05). Also shown are experimental data from the recall of acoustically presented lists of digits (exp). Adapted from Guildford and Dallenbach (1925).

presented digits versus list length is shown in Figure 4 and shows good agreement with the experimental data (Guildford & Dallenbach, 1925), with performance decreasing as a sigmoidal function of list length. This effect occurs for two reasons. First and most important, for longer lists more time elapses between the presentation and recall of an item, allowing the short-term connection weights to decay further, thus propagating less activation to the correct item node during output and leaving it more vulnerable to the effects of noise during the selection process. Second, there are more items competing for selection at each serial position (although only items from nearby serial position stand much chance of selection) and more serial positions at which an error can occur. Figure 4 also shows the probability of correctly recalling lists of unfamiliar and phonemically similar items (with the same spoken duration as the simulated digits). Again, the curves show the correct sigmoidal decrease, with marked phonemic similarity and item familiarity effects.

The relationship between performance and word length involves participants' articulation rate. The model assumes that presentation and recall occur at the same rate, equal to the speeded articulation rate of the participant on the list items. Within this assumption, performance decreases with the overall time taken to present and recall the list for the same reason that performance decreases for lists of more items (there is more time in which short-term connection weights decay).

Figure 5 shows item span as a function of the articulation rate for the list items (i.e., the inverse of the average spoken word length *wl* of the items in the list). Each item is composed of 5 (dissimilar) phonemes and has a spoken word length ranging from 0.6 to 1.5 s. The figure shows the correct, approximately linear, relationship between span s and articulation rate r for lists of familiar words, such that $s \approx \alpha r + c$, where the slope α and intercept c are both approximately 2.0. Experimental data from Hulme et al. (1991) are also shown for comparison.

Figure 5 also shows the item span versus articulation rate for lists of unfamiliar items. As described in the section on learning and long- and short-term connection weights, the familiarity of an item's phonemic composition is reflected in the long-term weights between items and output phonemes and between input phonemes and items. Experimental data from Hulme et al. (1991) are also shown for Italian words (after the participants have been told their meaning). The simulation models the beneficial effect of familiarity reasonably well (i.e., the curve for unfamiliar items is lower and approximately parallel to that for familiar items). However, the simulated unfamiliar item curve has a slightly shallower gradient. This may be due to our simulations not modeling the occurrence of errors because of misorderings of the phonemes within an item or exchanges of phonemes between items. This type of error, rare with familiar items, is a significant source of error in the recall of unfamiliar items (e.g., Trieman & Danis, 1988) and becomes increasingly important in short lists of long items for which other types of error (e.g., order errors or repetitions) are rare.

The effect of phonemic similarity can also be seen in Figure 5, where it lowers the slope of the span versus rate curve without changing the intercept. The simulations are of familiar items, so

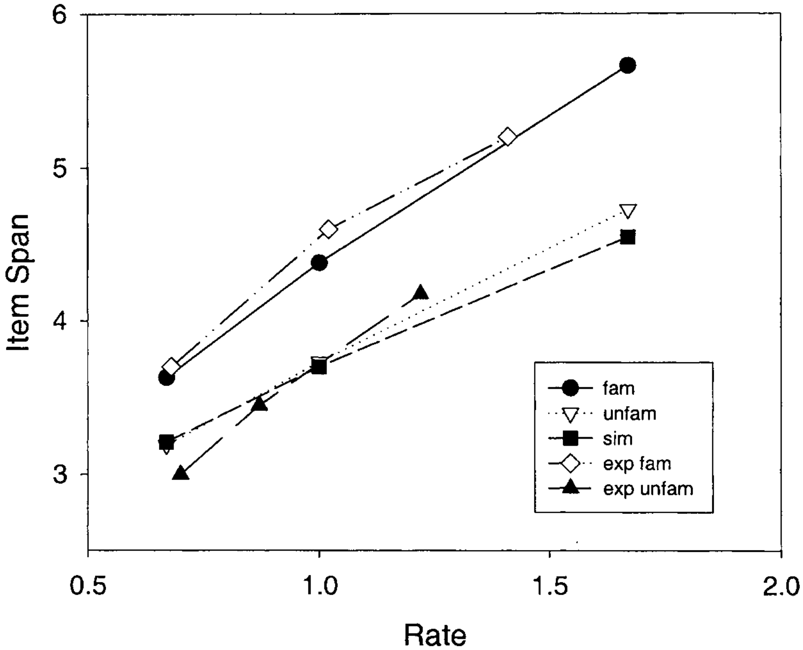

Figure 5. Item span versus speeded articulation rate (1/*wl*, or 1/word length). All simulations refer to recall of auditory presentation of five-phoneme items whose spoken duration varied from 0.6 to 1.5. Lists of three types of item were simulated: lists of familiar (F_p = 0.2) and phonemically dissimilar items (fam); lists of unfamiliar (F_p = 0.05) and phonemically dissimilar items (unfam); and lists of familiar (F_p = 0.2) and phonemically similar items, all having two phonemes in common (sim). Also shown are experimental data on recall by English speakers of known English words (exp fam) and Italian words whose meaning is known (exp unfam). Adapted from Hulme et al. (1991).

this effect is unlikely to be an artifact of ignoring within-items misorderings of phonemes. Preliminary support for this effect of phonemic similarity was provided by Schweikert, Guentert, and Hersberger (1990).

Lesions and neuropsychological data. Selective neuropsychological impairments can be addressed by the effects of damage to specific layers and connections within the model (labeled in Figure 6). In the case of Wernicke's aphasia, one of its characteristic symptoms, that of being unable to find the meaning of a pseudohomophone despite being able to pronounce it, would follow naturally from loss of the input phoneme layer (see C in Figure 6), its connections to the item layer (see D in Figure 6), or both. In Broca's aphasia, the impaired articulatory agility with unfamiliar words or phrases could be interpreted as damage to the output phoneme layer (see G in Figure 6). Such damage might particularly involve the mechanism for ordering phonemes in speech output, which we have not incuded in our model (however, see Hartley & Houghton, 1996).

Table 3 illustrates the effects of removing various subsets of layers and connections from the model on various aspects of STM performance. One of the major successes of the conceptual model of the phonological loop in neuropsychology is its ability to interpret the type of conduction aphasia exhibited by Patients P.V. and J.B. (Vallar & Baddeley, 1984a; Shallice & Butterworth, 1977) in terms of damage to the phonological store (see the introduction). Table 3 shows that removing the input phoneme layer (C; see Figure 6) or its connections to the item layer (D; see Figure 6) abolishes the ability to recall auditory lists while leaving a residual ability to recall visual lists, which is relatively insensitive to phonemic variables. However, of these two lesions, only lesion D preserves the immediate repetition of auditory words. Thus, lesion D best approximates the overall pattern of these patients' difficulties. However, this lesion overestimates the patients' impairment with auditory materials because it involves total removal of the means by which input phoneme activation could influence the item layer. The effects of partial lesions of the type hypothesized for the patients have yet to be explored.

A second type of conduction aphasia corresponds to impaired immediate repetition of a single auditory item but also to a relatively spared span for short familiar items (see the section on phonological information and Shallice, 1988, p. 164). This is modeled by removal of the connections from the input to output phoneme layers, (see H in Figure 6) which impairs immediate repetition of auditory nonwords because of a lack of direct activation of output phoneme nodes. This also causes an inability to create strong short-term associations from items to output phoneme nodes during presentation, leading to somewhat impaired recall of lists of auditory words and strongly impaired recall of lists of auditory nonwords. However, the repetition (reading) of visual words or nonwords, or lists of visual words and nonwords, is unimpaired because this does not rely on the direct input phoneme to output phoneme connections (it relies on the reverse, output phoneme to input phoneme, connections; see I in Figure 6). This pattern of inability to repeat auditory nonwords but relatively spared reading is characteristic of Patients M.K. (Howard & Franklin, 1990) and R. (Michel, 1979). Note however, that these patients have a more extensive and complicated pattern of impairment than P.V. and J.B. and may also have an inability to use phonological recoding of written words to access semantic information. That is, while they are able to pronounce written pseudohomophones (e.g., bote, tode, tial) they are not able to say what they would mean (e.g., Howard & Franklin, 1990).

Table 3
The Effect of Location-Specific Lesions on Item Span

Lesion	List items (acoustic) Dis	Sim	Long	Unfam	Recognition memory (%)	Word repetition
None	7.1	4.8	4.9	5.7	87	√
A or B	6.6	4.5	4.9	5.4	87	√
E or D	×	×	×	×	×	√
C, J, or K	×	×	×	×	×	×
G	2.1	2.1	2.3	2.1	×	×
F or I	2.1	2.1	2.3	2.1	×	√
H	3.4	2.6	3.3	2.3	35	×

Lesion	List items (visual) Dis	Sim	Long	Unfam	Recognition memory (%)
None	6.8	4.6	4.6	5.4	86
A or B	6.3	4.1	4.6	4.9	86
E	×	×	×	×	×
C, D, G, or I	2.1	2.1	2.3	2.1†	×
F	2.1	2.1	2.3	2.1	×
H, J, or K	√	√	√	√	√

Note. The interaction of lesions of the model with presentation modality, familiarity, phonemic similarity, and word length are shown in terms of item span. Also shown is performance on a recognition memory task, measured as the percentage of items correctly recognized as being from a recently presented list of six dissimilar letters. Lesions refer to loss of an entire layer or set of connections as labeled in Figure 6. List items were phonemically dissimilar letters (Dis); phonemically similar letters (Sim); items comparable to dissimilar letters but twice as long (duration = 0.8 s); and unfamiliar items comparable to dissimilar letters (Unfam; $F_p = 0.05$). × indicates that the model failed to perform at all; √ indicates normal performance; and † indicates that span would be unobtainable for polysyllabic nonwords, which can activate item nodes only during presentation via activation of the output and input phoneme layers.

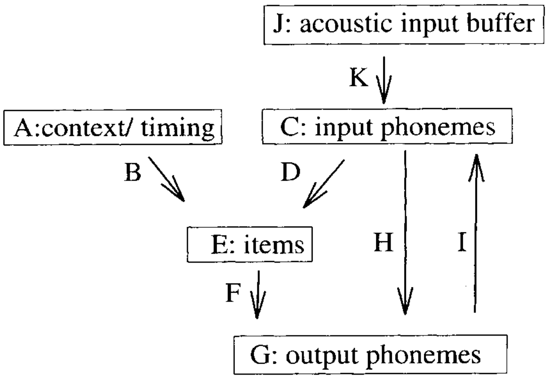

Figure 6. The layers and connections lesioned in Table 3.

Interestingly, removal of the context layer reduces span, but not greatly, because recovery of items from inhibition following presentation ensures that the first item is most active and so on (as in the CQ model outlined in the section on the minimalist connectionist model). Again, this effect is independent of presentation modality, and immediate repetition of a single word is spared. The effects of word length, phonemic similarity, and item familiarity, and the ability to make recognition memory judgments, also survive this lesion. The improvement in performance following Hebb repetition, temporal grouping, or rehearsal of a list of familiar items do not survive this lesion, and position-specific intrusions from previous lists would also not be expected. The existence of this cluster of symptoms is a prediction of the model.

Removal of the item layer (see E in Figure 6), which prevents any serially ordered recall of items independent of the modality of presentation, completely destroys span performance, with the only function spared being immediate repetition of a single word. The existence of such patients is a prediction of the model.

Loss of the layers or connections in the path providing phonemic feedback (items–output phonemes–input phonemes–items) drastically reduces item span and removes the effects of phonemic similarity, item familiarity, and word length (which can even be slightly reversed; see the section on predictions and neuropsychological data). Recognition memory judgments are also impaired by these lesions, consistent with the analysis of Starr, Barrett, Pratt, Michalewski, and Patterson (1990), who found that impairments in both tasks tended to go together (see the section on recognition memory judgments for more details of recognition memory performance). Loss of the item–output phoneme connections also produce this pattern of deficits as well as deficits in immediate repetition. Loss of the other two sets of connections within this loop has modality-specific implications as follows. Loss of the input phoneme–item connections prevents any activation of item nodes during acoustic (but not visual) presentation, reducing span to zero. Loss of the connections from output phonemes to input phonemes prevents visual information from activating item nodes and reduces visual, but not auditory, span to zero.

In summary, our model predicts many potential patterns of neuropsychological impairment. At this stage the model's strongest prediction is that of different clusters of symptoms as a result of major lesions to separate parts of its architecture. We singled out the predicted effects of damage to the context–timing system here. Assessing the significance of the subtly different patterns of memory impairment predicted as a consequence of this and other lesions will depend on progress in the ability to characterize neural damage at this level of specificity and in mapping the components of the model onto the brain.

Serial Position Effects

Here we consider the effects of various experimental manipulations on the probability of recalling the correct item at each serial position: the serial position curve.

Presentation modality and temporal grouping. Figure 7 shows the effects of presentation modality on serial position curves for lists of items of varying phonemic similarity, familiarity, and duration. Acoustic presentation improves performance on the final item over that for visual presentation in all conditions, with an improvement also being noticeable for the penultimate item. This effect is caused by the stronger phoneme–item associations of the final item, which strengthens the positive phonemic feedback it receives during recall. It also strengthens the negative feedback received during the recall of prior items (reducing errors caused by premature recall of the final item). Comparing the curves for long or unfamiliar items with those for phonemically similar items, we see that phonemic similarity weakens the auditory recency effect: reducing the improvement on the final item and removing any improvement on earlier items. This is due to the reduction in negative phonemic feedback between phonemically similar items and is consistent with the data from Watkins et al. (1974).

Presentation modality also plays an important role in the effect

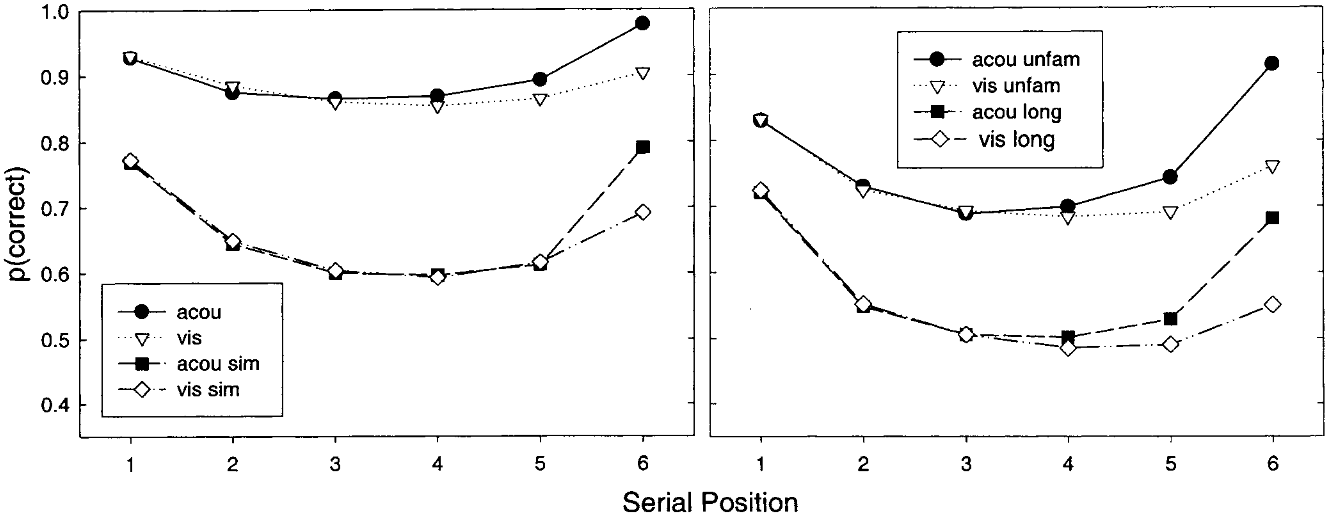

Figure 7. Simulation of the effect of presentation modality on serial position curves. Acoustic presentation (acou) leads to extra recency compared with visual presentation (vis) in all conditions. Left panel: the effect of using phonemically similar items (sim). Right panel: the effect of using unfamiliar items (unfam) and long items (long). Note that the size and extent of the extra recency for acoustic items were reduced by using phonemically similar items but not by using unfamiliar or long items. The items used were acoustically presented phonemically dissimilar letters (acou); visually presented phonemically dissimilar letters (vis); phonemically similar letters (acou sim and vis sim); unfamiliar items resembling dissimilar letters (acou unfam and vis unfam); and longer items (duration = 0.8 s) resembling dissimilar letters (acou long and vis long).

of temporal grouping. Figure 8 shows the effect of temporal grouping of visual and acoustic stimuli in the model. The model demonstrates both improved recall and scalloped serial position curves. The effect of grouping is greater for acoustically presented items than for visually presented ones and produces a more marked recency effect. Comparison with experimental data in Figure 9 shows that the magnitudes of these effects are qualitatively well modeled, including the relative effects of modality. The exception to this is performance on the final visual group, for which the model's performance is too good. This is due to the general problem that the model's serial position curves do not show as steep a primacy effect as the experimental data, particularly for long visual lists. Simulations (not shown) also show that the size of the temporal grouping effect is largely independent of the spoken duration and phonemic similarity of items, in agreement with experimental data (see Hitch et al., 1996).

Transposition gradients (i.e., the probabilities for items to be recalled at given serial positions) are shown in Figure 10. This shows the correct tendency for an item to be recalled near to its correct serial position, if not actually at it (see Healy, 1974, and Figure 14). Figure 10 also shows the probability of each item being recalled at each serial position for a temporally grouped list. This shows that although the overall number of transpositions have decreased, the relative number of transpositions between corresponding within-group positions (−3 and 3) have increased. This is consistent with the experimental data (of Ryan, 1969).

The model shows primacy and recency effects within a list (or within a group in a temporally grouped list) due to positional cuing from the context–timing signal: If an item is close to the beginning (or end) of the list, there are fewer cued items preceding (or following) it that compete for selection. There is also a slight knock-on effect of earlier errors that contributes to the primacy effect. A stronger effect of primacy could be obtained by changing the timing with which connection weights decay (see the section on limitations and future work and Figure 21). Similarly, positional cuing contributes to the transposition gradients being sharply tuned to the position of correct recall. Another important factor here is that items recover from suppression following selection at presentation in the order in which they were presented. Thus, an item presented in a certain serial position is likely to be the most active item at the same serial position during recall (with those before it having been selected and suppressed and those after it having had less time to recover from suppression at presentation; see the section on the minimalist connectionist model).

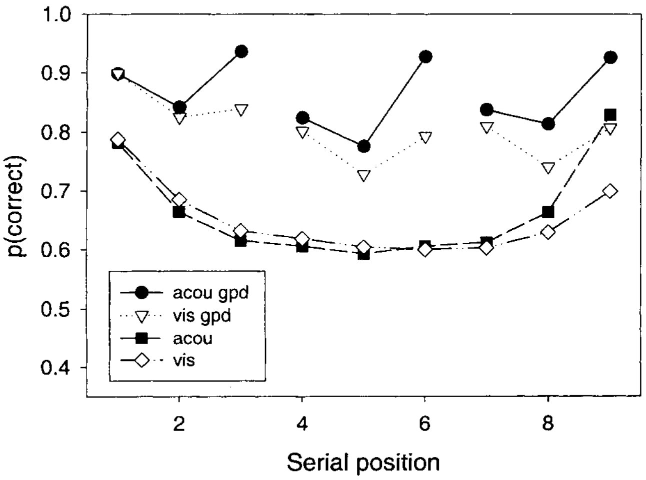

Figure 8. Simulated serial position curves for digits as a function of temporal grouping and modality at presentation. The curves show recall of lists of nine acoustically presented digits (acou); visually presented digits (vis); acoustic digits grouped in threes (acou grp); and visual digits grouped in threes (vis grp).

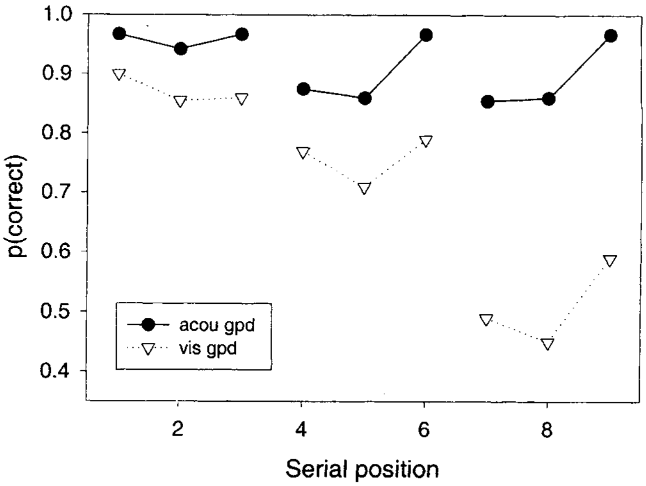

Figure 9. Experimental serial position curves for lists of visual (vis grd) and acoustic (acou gpd) digits grouped in threes. Adapted from Frankish (1985), Experiment 2.

Phonemic similarity. Phonemic similarity among list items has an effect on the serial position curve: producing a more deeply bowed shape (i.e., there is an interaction with serial position; see

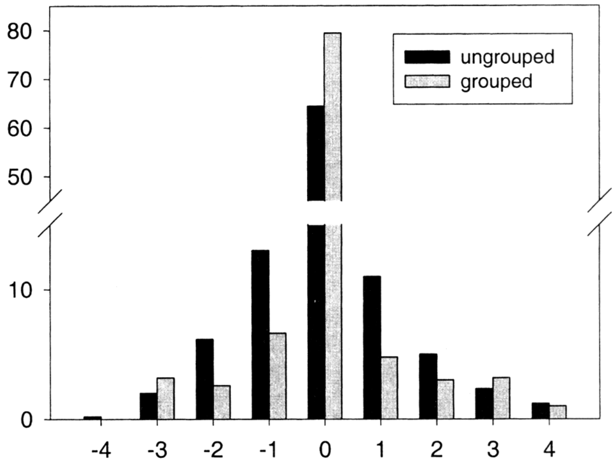

Figure 10. Simulation of the effect of temporal grouping on transposition gradients. Lists of nine visual items were used, either ungrouped or grouped 3 × 3. The figure shows the percentage of trials in which each item is recalled at a given position relative to its correct position (the correct position is marked by 0, earlier positions by −1, −2, etc., and later positions by 1, 2, etc.). Note that, overall, grouping produces a decrease in the chance of recalling an item in an incorrect position but that there is a relative increase in transpositions to the same relative position within nearby groups (i.e., Positions 3 and −3).

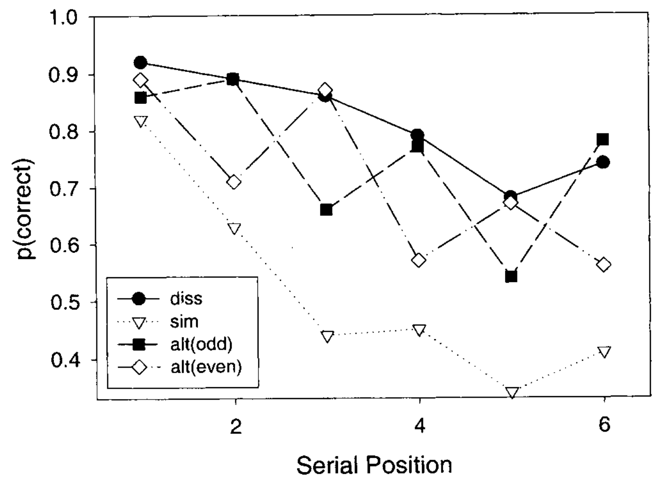

Figure 11. Effect of phonemic similarity: Experimental data adapted from Baddeley (1968, Experiment V). Serial position curves for visual lists of phonemically dissimilar letters (diss); lists of phonemically similar letters (sim); and lists of alternating similar and dissimilar letters with the similar ones in the even, alt(even), and odd, alt(odd), positions.

the upper and lower curves in Figure 11). We aimed to model this and to model the effect of phonemic similarity in lists of alternate phonemically similar and dissimilar items (e.g., the list of letters *y, b, h, d, q,* and *g* in which every second letter contains the phoneme/ee/). This produces zig-zagged serial position curves in which recall of the dissimilar items is unimpaired with respect to the recall of pure lists of dissimilar items (see Figure 11).

Figure 12 shows that the model qualitatively reproduces both the interaction between phonemic similarity and serial position as well as the approximate coincidence of the curves showing performance on dissimilar items whether they occur in an alternating list or a pure list of dissimilar items. However, the simulations show a better match to the size of the experimental phonemic similarity effect when the similar items are simulated as having three of four phonemes in common rather than the veridical one of two. This indicates a limitation of simply modeling all phonemes as equal because this underestimates the actual similarity of some items (see also Table 4 and the discussion of featural representations that follows). We note again that the simulated serial position curves do not show as steep a primacy gradient as the experimental data.

Figure 14 shows the transposition gradients for the simulated recall of lists of phonemically dissimilar items, phonemically highly similar items, and items of alternating similarity shown in Figure 12 (right). The simulations show the correct wider spread of transpositions for phonemically similar items than for phonemically dissimilar ones and the correct increased number of transpositions between the positions of the phonemically similar items in lists of alternating similarity. The corresponding experimental data from the study of Henson et al. (1996) are shown for comparison.

Figure 13 shows the corresponding simulation for unfamiliar items resembling the letters in all other ways. Overall performance is reduced, but the serial position curves show the same pattern. Thus, we predict that the effects of similarity and familiarity do not interact at this more detailed level of analysis, just as they had separate effects on overall performance (see Table 2 and Besner & Davelaar, 1982).

Our model for representing an item's spoken form as simply the list of constituent phonemes can be further tested by looking at the occurrence of transpositions between specific items as a function of their phonemic similarity. Table 4 shows a simulation of Conrad's (1964) confusion matrix for recalling visually presented lists of six letters. For most pairs of letters, the simulated rates of confusions match the experimental ones fairly well and show the main pattern of errors in which most confusions occur between the letters *B, C, P, T,* and *V* or between the letters *F, M, N, S,* and *X.* The model also shows the increased rate of errors between *C* and

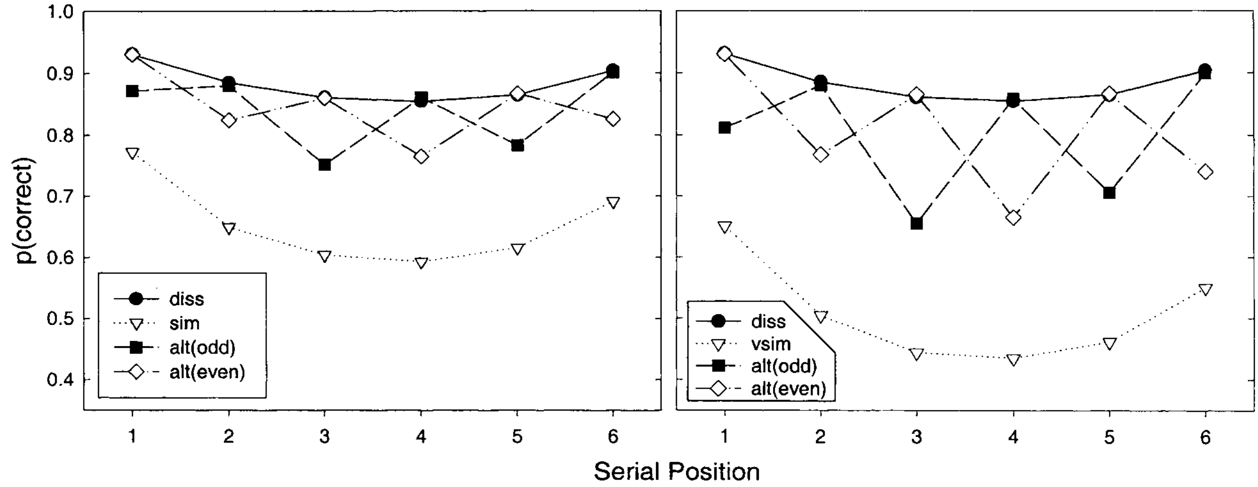

Figure 12. Effect of phonemic similarity. Left panel: simulated serial position curves for visual lists of phonemically dissimilar letters (diss); lists of phonemically similar letters (sim; having one of two phonemes in common); and lists of alternating similar and dissimilar letters with the similar ones in the even, alt(even), and odd, alt(odd), positions. Right panel: simulated serial position curves using dissimilar letters and lists of phonemically similar items (vsim; having three of four phonemes in common).

Table 4
Simulated and Experimental Confusion Matrix for Visual Letters

Presentation item	Recalled item									
	B	C	P	T	V	F	M	N	S	X
B	3.4 (—)	0.1 (0.1)	0.2 (0.7)	0.1 (0.2)	0.2 (0.4)	0.0 (0.0)	0.0 (0.1)	0.0 (0.1)	0.0 (0.0)	0.0 (0.0)
C	0.1 (0.1)	3.0 (—)	0.1 (0.1)	0.1 (0.3)	0.2 (0.2)	0.0 (0.1)	0.0 (0.0)	0.0 (0.0)	0.1 (0.1)	0.0 (0.0)
P	0.1 (0.6)	0.1 (0.2)	3.0 (—)	0.2 (0.5)	0.2 (0.2)	0.0 (0.1)	0.0 (0.1)	0.0 (0.0)	0.0 (0.1)	0.0 (0.0)
T	0.1 (0.0)	0.1 (0.1)	0.2 (0.2)	3.0 (—)	0.1 (0.1)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
V	0.1 (0.5)	0.1 (0.4)	0.1 (0.3)	0.2 (0.2)	3.0 (—)	0.0 (0.2)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)
F	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)	3.2 (—)	0.2 (0.1)	0.1 (0.2)	0.1 (0.2)	0.0 (0.2)
M	0.0 (0.1)	0.0 (0.0)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)	0.1 (0.1)	3.2 (—)	0.1 (1.1)	0.2 (0.0)	0.0 (0.1)
N	0.0 (0.0)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)	0.1 (0.1)	0.2 (0.9)	3.3 (—)	0.1 (0.1)	0.0 (0.1)
S	0.0 (0.0)	0.1 (0.2)	0.0 (0.0)	0.0 (0.1)	0.0 (0.1)	0.1 (0.8)	0.2 (0.1)	0.2 (0.2)	3.1 (—)	0.2 (0.4)
X	0.0 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.1)	0.0 (0.0)	0.2 (0.1)	0.2 (0.0)	0.1 (0.0)	0.5 (0.1)	3.4 (—)

Note. The percentage of trials in which a specific recalled item replaced a specific presented item to produce a single error in the recalled list. The stimuli were visually presented lists of six letters randomly selected (without repetition) from *B, C, P, T, V, F, M, N, S,* and *X*. Also shown in parentheses are the corresponding experimental data, adapted from Conrad (1964).

S. However, confusions between some pairs of letters are underestimated by the model, most noticeably *P* and *B* and *N* and *M*. These shortcomings indicate the inclusion of more detailed features than phonemes, as some phonemes are clearly more similar than others. For example, in a more detailed featural description the letters *N* and *M* are more similar than, say, *N* and *F* (equally, *B* and *P* are more similar than *B* and *C*). Spencer (1996, pp. 116–117; see also Wicklegren, 1969) specified a set of features by which consonant phonemes may be distinguished, such as the /n/ in the letter *N* (/e/ /n/). In this description, /n/ and /m/ differ by two features, /p/ and /b/ by one, whereas both /n/ and /f/ and /b/ and /s/ differ by five features.

The phonemic similarity effect occurs during the output of an item because of the extra phonemic feedback to other similar items (if present in the list). This increases the likelihood of a (nearby) similar item being recalled in its place, but it does not impair the recall of any intervening dissimilar item (see Figure 14). Our previous model did not show a zig-zag serial position curve because items phonemically similar to the presented item were also associated to its context state. This impaired recall of the dissimilar items, as well as the similar ones, due to the overlap between neighboring context states (see Burgess & Hitch, 1992, p. 451).

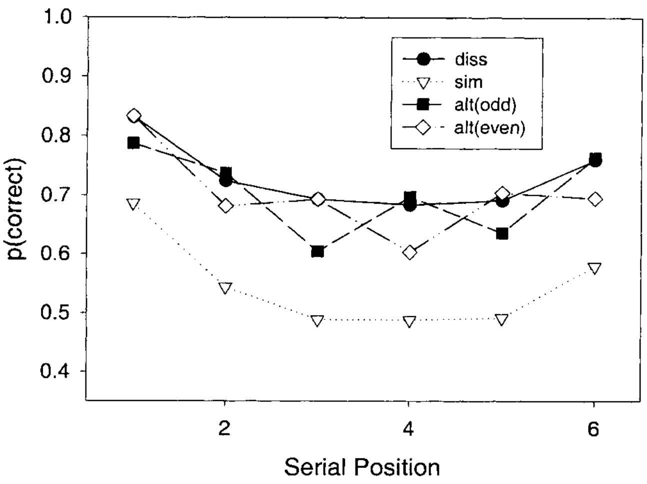

Figure 13. Effect of phonemic similarity on unfamiliar items: simulated serial position curves for visual lists of phonemically dissimilar unfamiliar items (comparable to letters but with $F_p = 0.05$; diss); lists of phonemically similar unfamiliar items (sim); and lists of alternating similar and dissimilar unfamiliar items with the similar ones in the even, alt(even), and odd, alt(odd), positions.

Item familiarity. Figure 15 shows serial position curves as a function of item familiarity. Interestingly, simulated recall of lists of items of alternating familiarity does not show exactly the same pattern of errors as that for the recall of lists of items of alternating phonemic similarity. The recall of the familiar items in lists of alternating familiarity is impaired relative to the recall of the familiar items in pure lists of familiar items, in contrast to the unimpaired recall of phonemically dissimilar items in lists of alternating similarity. This is a prediction arising from the way item familiarity is modeled.

The presence of the unfamiliar items impairs recall of the neighboring familiar items because the relatively weak phonemic feedback to an unfamiliar item leaves it more susceptible to replacement by a familiar one. Thus, the extra errors caused by the unfamiliar items tend to involve familiar items in an alternating list (which is not the case with alternating phonemic similarity).

Word length. Figure 16 shows serial position curves as a function of word length. Curves are plotted for lists of short words, lists of long words, and lists in which short and long words are presented in alternation. Interestingly, simulated recall of lists of items of alternating spoken duration does not show exactly the same pattern of errors as that for the recall of lists of items of alternating phonemic similarity or alternating familiarity. Performance is worse for lists of long items than lists of short items and is roughly intermediate for alternating lists. The most important influence on performance is the total length of the list, as this is the time over which the short-term weights and inhibitions associated with each item have decayed; this explains the difference in overall performance. This length of time is the same for items in the same position in either of the alternating lists (even or odd). The slight reverse zig-zag nature of the curves for alternating lists arises from the length of time since presentation of the neighboring items competing for recall: This will be slightly longer for a short item

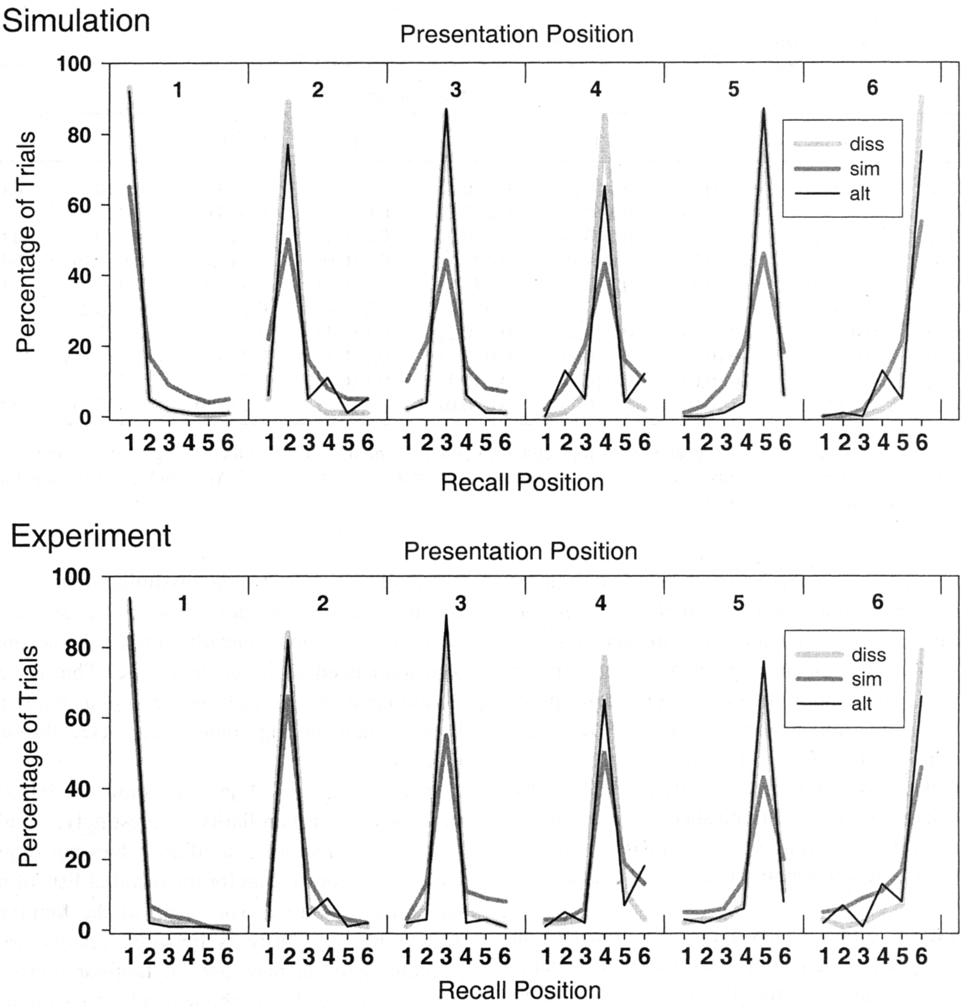

Figure 14. Transposition gradients as a function of phonemic similarity. The percentage of trials in which each item is recalled at each serial position is shown for visually presented lists of phonemically dissimilar letters (diss); phonemically similar letters (sim); and list of letters of alternating similarity (alt; with similar letters at the even positions). Top: data from the simulations shown in Figure 12 (right panel). Bottom: experimental data adapted from Henson et al. (1996). Note that for both the simulation and the experiment, the increase in transpositions in the alternating list compared with the dissimilar list is restricted to those involving similar items (i.e., even positions).

immediately following a long one than for a long one following a short one.

Rehearsal. The basic role postulated for the articulatory loop is a rehearsal mechanism for maintaining verbal items "on-line" for use by other systems. Therefore, it is important to examine this aspect of the model. A full model of rehearsal would include a process for selecting and implementing different rehearsal strategies. However, to achieve this would be difficult because it would involve specifying control systems external to the phonological loop (i.e., "executive processes"; Baddeley, 1996; Norman & Shallice, 1986), about which relatively little is known. We therefore opted for a simplistic approach in which rehearsal is implemented in the same way as serial recall, such that repeated rehearsal of a list is equivalent to repeated recall.

Figure 17 shows the effect of four successive rehearsals of a list of acoustically presented unfamiliar items. This demonstrates that the increase in the long-term component of context–item, phoneme–item, and item–phoneme connections during rehearsal brings stability after a small number of rehearsals, that is, fewer additional errors are committed as rehearsal progresses, similarly to the experimental data (Heffernan, 1991). This property of the model demonstrates that subvocal rehearsal could be used to reliably store a serially ordered list of unfamiliar items and could provide an important mechanism in at least the phonological aspects of learning new vocabulary. It also demonstrates the stability required to provide a rehearsal function that may be of use to other systems during the learning of language (see Baddeley et al., 1998).

Recognition memory judgments. Another way in which we assess the model's capacity is the accuracy of its recognition memory judgments: how well it can detect whether a single item was in the most recently presented list (based on the strength of

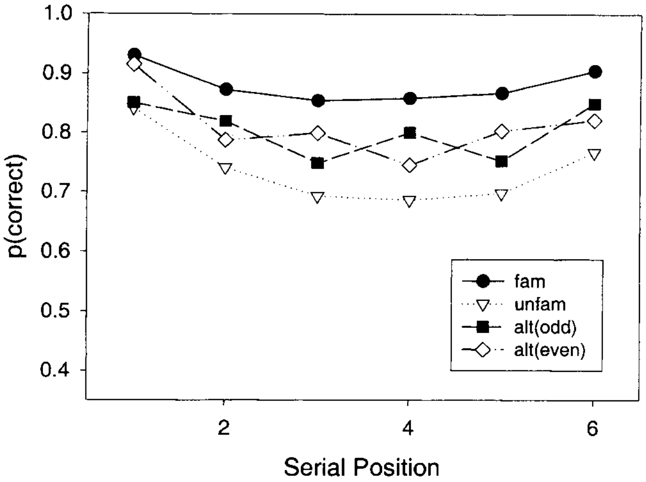

Figure 15. Effect of item familiarity: simulated serial position curves for lists of visual familiar items (dissimilar letters; fam); lists of comparable unfamiliar items (unfam; having $F_p = 0.05$); and lists of alternating familiar and unfamiliar items with the unfamiliar ones in the even, alt(even), and odd, alt(odd), positions.

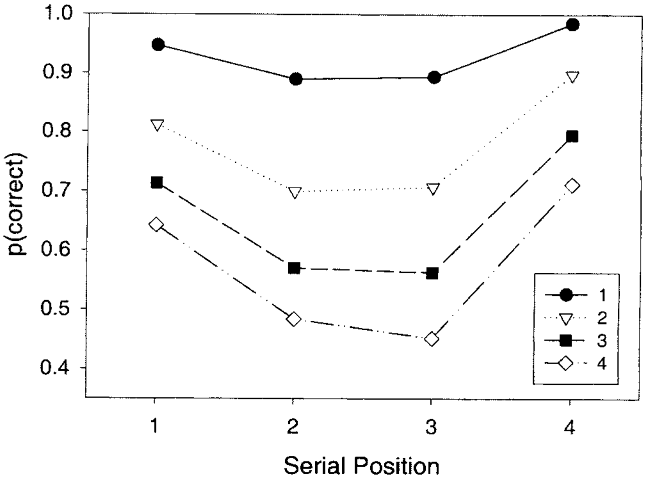

Figure 17. Effect of rehearsal for lists of unfamiliar items. Serial position curves for four consecutive recalls (1, 2, 3, and 4) of a list of four acoustic unfamiliar phonemically dissimilar items 0.4 s long. Note that the number of additional errors added by each recall decreases. Each recall increases contextual familiarity F_c by 0.10 (i.e., increases the long-term component of context–item connection weights from 0.0 to 0.4) and phonemic familiarity F_p by 0.05 (i.e., increases the long-term component of item–phoneme and phoneme–item connections from 0.05 to 0.25).

phonemic feedback). This is shown in Figure 18 as a function of the probe item's serial position and the phonemic similarity of the list items and in Figure 19 as a function of the probe item's serial position and the familiarity of the list items. The figures show the correct increase in accuracy with the serial position of the probe and that performance is worse for phonemically similar or unfamiliar items, consistent with experimental data (Henson, Hartley, Hitch, & Burgess, 1999; a primacy effect is observable for the first item in experiments, but not in the model).

The simulations also show effects that we are not aware of in the literature. These include that acoustic presentation gives a strong advantage over visual presentation for recognition memory judgments involving the last item (giving 100% correct) and that the acceptance rate is greater for familiar items. The curves showing accuracy of recency judgments on lists of items of alternating similarity or familiarity show a marked sawtooth pattern, with those for alternating familiarity reaching the curves for purely familiar or unfamiliar items at either extreme, whereas those for

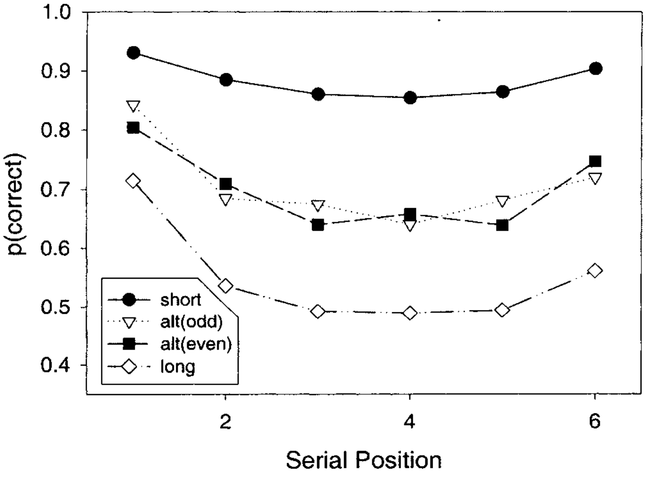

Figure 16. Effect of item length: simulated serial position curves for lists of short visual familiar items ($wl = 0.4$; short); lists of long items ($wl = 0.8$; long); and lists of alternating short and long items with the long ones in the even, alt(even), and odd, alt(odd), positions.

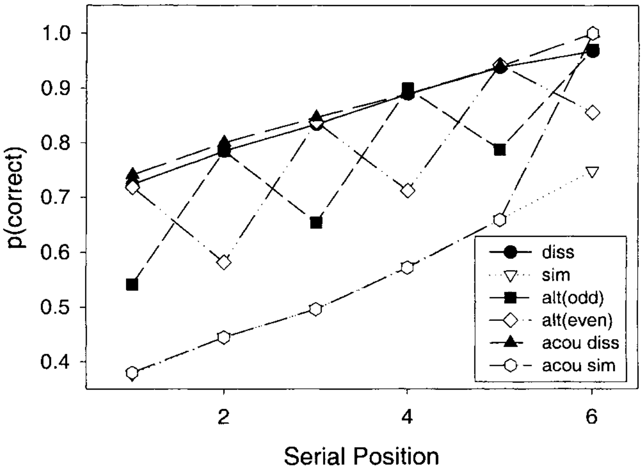

Figure 18. Recognition memory judgments as a function of the list position of the probe item and the modality of presentation and phonemic similarity of list items. Shown is the probability of correctly recognizing a letter from a recently visually presented six-item list of either phonemically similar letters (*b, d, g, p, v,* and *t;* sim); phonemically dissimilar letters (*q, r, n, o, h,* and *y;* diss); or lists of letters of alternating similarity with similar letters only at the even or odd positions, alt(even) and alt(odd), respectively. Also shown is the probability of correctly recognizing a letter as being from an acoustically presented list of phonemically dissimilar letters (acou diss) or similar letters (acou sim).

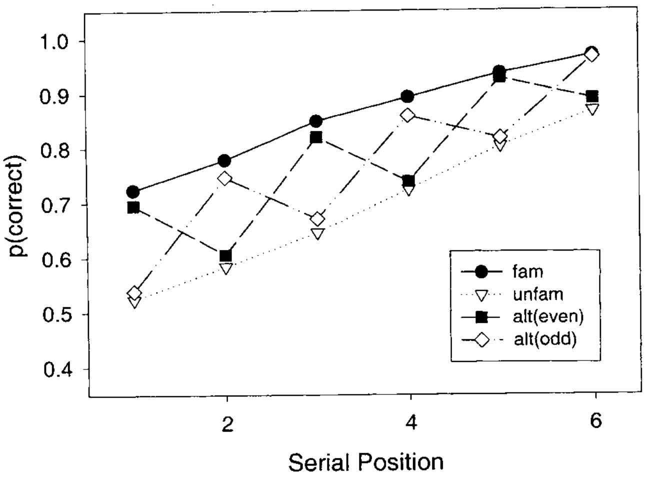

Figure 19. Recognition memory judgments as a function of the list position of the probe item and the modality of presentation and familiarity of list items. Shown is the probability of correctly recognizing an item from a recently visually presented six-item list of either familiar letters (*q, r, n, o, h,* and *y;* fam), comparable unfamiliar items (having $F_p = 0.05$; unfam), or lists of items of alternating familiar with unfamiliar items only at the even or odd positions: alt(even) and alt(odd), respectively.

alternating similarity coincide with the curve for purely dissimilar lists and reach midway to the curve for purely similar items. These curves are predictions.

Repetitions. Repeated recall of an item can occur once the suppression received after its first recall has decayed away. Figure 20 shows that repeats in recalling lists of six letters tend to be early list items that are recalled for a second time near to the end of the list, as found in the experimental data of Henson et al. (1996). In total, repeats accounted for 9.6% of all errors in dissimilar lists and 13.7% of all errors in similar lists, which compares well with experimental values (e.g., repeats formed 16% of transposition errors in Henson et al., 1996).

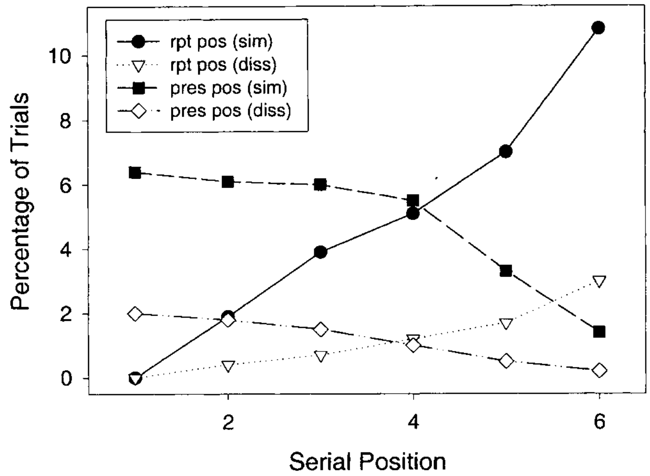

Figure 20. Erroneous repeats as a function of serial position. Curves show the percentage of trials in which the second occurrence of an incorrectly repeated letter occurs in a given position in lists of dissimilar letters, rpt pos (diss), or similar letters, rpt pos (sim). Also shown is the percentage of trials on which a letter that is repeated at recall was presented in a given position in lists of dissimilar letters, pres pos (diss), or similar letters, pres pos (sim).

Intrusion errors. Intrusion errors (i.e., recall of an item not in the presented list) are modeled as follows. Item nodes representing the vocabulary of possible responses are present in the model, each with the appropriate long-term phoneme–item and item–phoneme connections. What constitutes the set of possible responses is a complex issue, and, for the present, we do not want to try to model it in detail. Here we simply assume that the set of stimuli used throughout a block of trials define possible responses. Before being selected during presentation, all items have a baseline activation of −1.5, so that extralist items are available for recall but are less likely to be recalled than those selected at presentation. These items will be rarely recalled unless activated by long-term context–item associations due to inclusion in a previous list or long-term phoneme–item associations due to extreme phonemic similarity to a current list item.

The occurrence of serial order intrusions is illustrated in Table 5. This shows the probability of each item from a previously presented list being recalled at each serial position in the current list. The pattern of intrusions is generally similar to that for items from the current list, although the probabilities are much lower. However, the positional specificity of serial order intrusions is much lower than for recall of items from the current list. This is because the differences in the small inputs to possible intrusions from different serial positions have less effect on their probabilities of selection in the presence of noise of relatively larger amplitude.

Omission errors. In experimental data, the rate of omissions depends to some extent on the participant's instructions, (i.e., how likely he or she is to guess a response when unsure). The occurrence of omissions may be modeled, as in Burgess and Hitch's (1992) model, by assuming that the winning item activation must exceed a threshold value during output or an omission will occur. The value of the threshold corresponds to the participant's overall propensity to omit a response. Under this model, omissions become more common as the activation of items decreases (i.e., in long-duration lists, toward the end of a list, and in lists of nonwords rather than words).

Table 5
Serial Order Intrusions

Presentation position	Recall position 1	2	3	4	5	6
1	0.5 (92)	0.2 (5)	0.2 (2)	0.1 (1)	0.1 (0)	0.1 (1)
2	0.3 (5)	0.5 (87)	0.3 (5)	0.2 (2)	0.2 (1)	0.1 (1)
3	0.2 (2)	0.4 (5)	0.4 (85)	0.3 (5)	0.2 (2)	0.3 (1)
4	0.1 (0)	0.3 (1)	0.4 (5)	0.4 (85)	0.4 (5)	0.3 (2)
5	0.1 (0)	0.2 (0)	0.3 (2)	0.4 (5)	0.6 (85)	0.4 (5)
6	0.1 (0)	0.1 (0)	0.2 (0)	0.2 (1)	0.3 (5)	0.8 (89)

Note. The percentage of trials in which an item from a previous list was recalled in the current list as a function of serial position at presentation and recall. The percentages for recall of items in the current list are given in parentheses. Activation reaches the intruding item node from the long-term component of the context–item association ($F_c = 0.3$) and phoneme–item association ($F_p = 0.2$). Intruding items tend to appear near the serial position of their original list, but the tendency is less marked than for the recall of the current list items.

Evaluation

The model successfully reproduces the basic phenomena of the phonological loop, namely the sensitivity of serial recall to list length, word length, phonemic similarity, and articulatory suppression. It also successfully reproduces the interactions between articulatory suppression, phonemic similarity, and presentation modality as well as the linear relationship between the amount recalled and articulation rate. These were the main phenomena that the original concept of the phonological loop was designed to explain. The results considered in the rest of this section were not addressed by the phonological loop hypothesis and were modeled incompletely or not at all in Burgess and Hitch's (1992) model. These results constitute major advances over the range of application of the previous model and, interestingly, follow from relatively minor extensions or adjustments to it.

The structure of the model has been evaluated by the process of lesioning each component and comparing the resulting changes in performance with the relevant neuropsychological data. The model shows good agreement with the pattern of impairments shown by STM patients. The model's performance compares reasonably well with that of patients with the two types of conduction aphasia affecting (a) repetition of lists of short familiar items in order and (b) repetition of a single long low frequency word. It also compares well with the evidence relating impairments of recognition memory judgment with impairments in STM span tasks in most but not all cases (e.g., Starr et al., 1990).

The shape of the curve relating the probability of recalling a list to list length is modeled correctly, as are the effects of item and list familiarity (see the section on capacity). The critical feature enabling the model to explain learning and LTM effects are inclusion of long- as well as short-term Hebbian changes to connection weights (which were absent in Burgess & Hitch's, 1992, model).

The model succeeds in reproducing the bow-shaped curve relating the probability of recall to serial position. Serial position effects are attributable to the way discriminability of patterns of activation in the context nodes change for successive items. The effect of presentation modality on the shape of the serial position curve is successfully modeled by assuming that auditory stimuli enter a buffer from which they can activate input phoneme nodes during pauses at the end of a group or list, strengthening the short-term association between an item and its constituent phonemes. The model also explains the effect of temporal grouping on recall. It does so by extending the contextual input to include more than one dimension to enable it to reflect the rhythmic pattern of temporal grouping.

The model reproduces the zig-zag serial position curves characteristic of recall of sequences in which phonemically similar items are sandwiched between phonemically dissimilar items. It further predicts the effects on recall of alternating lists of words and nonwords and of short and long items. The use of rehearsal as a way to provide stable temporary storage of ordered items, including unfamiliar ones, for potential use in language learning is demonstrated. The model also reproduces the major types and distributions of errors in immediate serial recall. These include realistic transposition gradients, the effects of phonemic similarity on order errors, repetition errors, and serial order intrusions from prior lists.

The model maps onto the phonological loop idea of decaying phonological information being refreshed by rehearsal. The decaying short-term connection weights between item and input and output phoneme layers corresponds to the phonological store. The successive selection of items, reactivation of phoneme nodes, and relearning of short-term connections correspond to rehearsal. The feedback of phonemic information from output to input may be considered as hearing one's own inner speech. The model predicts that a third component: a context–timing signal is also present. Both the phoneme and context inputs to the item layer serve to increase span, and, the former underlie effects of phonemic similarity, word frequency and lexicality, and recency judgments, whereas the latter underlie visual recency, position-specific intrusions, Hebb repetition, and temporal grouping effects.

Limitations and Future Work

Although simulating a wide range of human data, the current model shows some limitations in terms of data that it does not address and in terms of its performance on the data that it does address. We note that modeling the effects of lesions (and also of articulatory suppression) cannot be complete, in that the residual behavior shown in experiments is often interpreted as being attributable to preserved processing outside the articulatory loop itself, such as in the visuospatial sketchpad (Baddeley, 1986; Salway & Logie, 1995). Thus, the effects of lesions or interference are likely to be greater in our simulations than in reality because we have not modeled the entire brain and all of the compensatory mechanisms that may be used when one particular process is impaired. Otherwise, the principal limitations, and suggestions for their rectification, are as follows:

Long-term item representations. There is a problem using the connections to an item node to form a long-term representation of the phonology of a novel item. Presentation of a new item recruits the most active item node (i.e., the one representing the most phonemically similar familiar item). This has the advantage that performance is better for nonwords that resemble a known word than for those that do not (the lexicality effect). In the long run, however, as the novel item becomes more familiar, the item node will develop strong connections to the phonemes in both the novel and the familiar items, resulting in confusion whenever they both appear in a list. The problem of representing a repeated item in a list raises similar questions (see Henson, 1998a). A related problem is that the long-term component of connections between phoneme and item nodes causes a slight phonemic similarity effect to remain in the recall of familiar visual items under articulatory suppression.

An appealing alternative to long term item–phoneme connections is that familiarity of recognition and production develop via long-term modification of horizontal connections within the input and output phoneme layers, with only short-term modification of the weights to and from item nodes. The phoneme layers themselves would have an autoassociative memory function in which phonemic representations of words existed as attractor states whose depth corresponded to the item's familiarity. A model of this sort would require only a relatively small number of item nodes (one for each serial position) with temporary association to the appropriate item in the current list (i.e., becoming purely place holders and resembling the nodes in the context layer). The model's item and phoneme layers could be considered as a recurrent autoassociative network if phonemic feedback were calculated several times during recall instead of once only (see the section on the role of phonemic feedback). This might be a first step toward such an attractor-based model.

The primacy effect shown by the model is weak. The simplest way to achieve steeper primacy is for all connection weights to decay from the start of recall, rather than the weights associated with an item decaying from the time at which it was last selected. This way, connection weights will have decayed by more when the end of list items are recalled than when recalling the items at the start of the list. Figure 21 shows the increased primacy in simulations of weight decay from the start of recall. This solution is analogous to that used to produce steep primacy in the primacy model (Page & Norris, 1998). However, it is difficult to motivate short-term connection weights (or activations in the primacy model) that decay throughout one recall (or presentation) of the list and are then reset to their original values at the start of the next recall before starting to decay again. For a single immediate recall, the primacy model can achieve this by assuming that error-free rehearsal of all prior list items occurs in the gaps between item presentations, and that decay during rehearsal is negligible relative to decay during recall of the list. By contrast, we have made the simpler assumption that presentation, recall, and rehearsal all occur at the same rate (in fact, rehearsal is identical to recall) and settle for serial position curves of slightly shallow slope.

Ordered activation of phonemes. The ordered activation of the phonemes corresponding to successive syllables should be modeled in more detail because the two phonemic buffers must also include order information, which we have omitted here for simplicity. Obviously, the input and output representations of, say, *cat* and *act* must differ. An extension to the model that incorporates order information for both words and nonwords, and constrains the output to be pronounceable, has been described by Hartley and Houghton (1996). In principal, this could be incorporated into our model (see Gupta & MacWhinney, 1997).

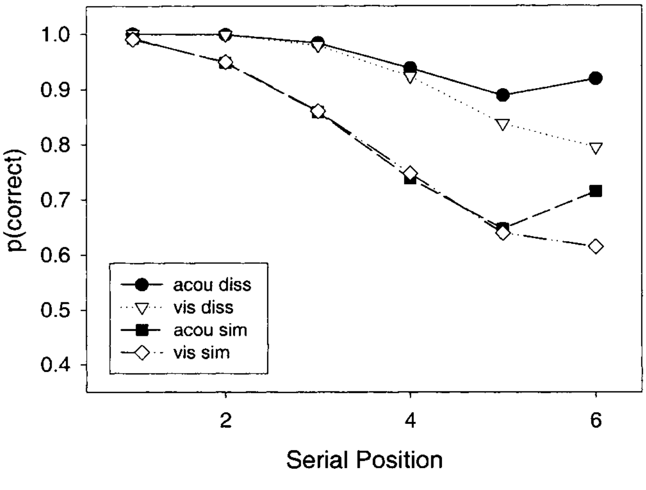

Figure 21. Serial position curves for simulations in which short-term weights connected to an item decay from the beginning of recall, rather than from the last time the item was selected. Curves show the recall of lists of six familiar items of a spoken duration of 0.8 s. Items were presented acoustically or visually and were phonemically dissimilar (acou diss and vis diss, respectively) or similar (having one of two phonemes in common; acou sim and vis sim, respectively). This figure may be compared with Figure 12 (left panel), which shows recall of visually presented letters (which have a spoken duration of 0.4 s, giving the same total time across items for connection weights to decay as in the simulation shown here). Note that here the primacy effect is more marked than in the simulation discussed in the rest of this article.

Time scales and long-term data. Relative to the time scale over which short-term connection weights decay, we have considered the long-term connection weights as static. However, it is clear that they, too, must decay over some longer time scale because, after a long gap without repetition, the pronunciation of novel items becomes more error prone, serial order intrusions from a previous list become less common, and the Hebb effect disappears when the delay between repetitions is long enough. The longer time constants associated with the long-term decay of phoneme–item–phoneme connections and of context–item connections need to be determined. A second problem concerns the rate of change of the "context" vector, which, in simulations, is equal to both the rate of presentation and recall of items and is interpreted as reflecting the rhythm of presentation (see the section on temporal grouping and Hitch et al., 1996), that is, automatically adjusting to variable rates of presentation. A possible solution to how this on-line adjustment could be achieved in practice was presented by Hartley (1999).

The question of time scales within the model also arises with regard to similarities found in the patterns of errors after short delays and delays of many hours (Nairne, 1992). These may indicate mechanisms common to STM and LTM for serial position. However, there are also major differences, such as that errors in STM tend to be phonemically related, whereas errors in LTM tend to be semantically related (see, e.g., Baddeley, 1986). Most persuasively, an almost total dissociation between impairments in STM and LTM in neuropsychological cases can be found (amnesic people are typically unimpaired in immediate recall, whereas STM patients such as J.B. and P.V. do not appear to have LTM deficits apart from those thought to be caused by their short-term deficit, such as learning new vocabulary; see, e.g., Shallice, 1988). Our model is aimed at immediate recall, the influences of long-term learning on it, and its role in building up long-term representations of an item's phonology. Both the model's overall performance and the phonemic similarity effect that it shows fall rapidly to floor levels when recall follows a filled delay of more than a few seconds, consistent with a contribution to STM but not LTM. Thus, although the contents of LTM might be read out into STM structures, how this might be done, and how or whether our model should relate to explicit recall from LTM, is a subject for future work.

Predictions

In this section we discuss specific predictions made by the model in terms of psychological, neuropsychological, and neurophysiological experiments. Before doing this, however, we note that, because of its wide scope, the model could be used to make numerous predictions beyond those related to the specific simulations we have chosen to perform. These are perhaps best expressed in terms of the clustering of effects that the model's architecture imposes. Thus, we would expect data from studies of dual-task interference, individual differences, development, neuropsychology, and functional neuroimaging to indicate separable components corresponding to the four main layers of the model (see Figures 1 and 2).

Psychological data. The most obvious predictions arise from the model's use of a separate context–timing signal (as proposed by Burgess & Hitch, 1992). This signal helps to correctly order familiar items in a list and underlies the temporal grouping and Hebb effects, but it is not needed for performing a recognition memory judgment. Thus, we predict that temporal grouping and Hebb repetition effects should be independent of articulatory suppression, word length, and phonemic similarity (that this is true for the temporal grouping effect has recently been supported; see Hitch et al., 1996). For the same reason, we predict that temporal grouping and Hebb-like repetition of lists should not affect recognition memory judgments, whereas we would predict an interaction between grouping and repetition, such that the Hebb repetition effect requires repetitions of the list to have the same temporal structure and so involve the same sets of context nodes (consistent with the findings of Bower & Winzenz, 1969). More generally, any temporal component of a concurrent task would be expected to interfere with immediate serial recall. Recent evidence offers some support for this prediction: Concurrent paced finger tapping was found to interfere with the recall of supraspan lists (Vandierendonck, De Vooght, & Van der Goten, 1998).

The recall of mixed lists of words and nonwords (see Figure 15) and the prediction that words are more likely than nonwords to form intrusion errors in these lists have not been investigated experimentally to our knowledge. The way in which item familiarity is encoded by long-term connection weights has a straightforward implication regarding overall performance: A filled pause before recall will be more damaging to lists of nonwords than lists of words; as, for nonwords, phonemic feedback during recall relies to a greater extent on short-term decaying connection weights.

The detailed characteristics of the errors committed in simulations of the model are emergent features. In many cases the simulated behavior of the model corresponds to experimental data known to us (see above), but in some the relevant data were not available, enabling us to make some further predictions. These include lower zig-zags for alternating familiar and unfamiliar items than for similar and dissimilar items; no zig-zags for short and long items; the span versus articulation rate curves for lists of items of varying phonemic similarity (however, see Schweickert et al., 1990); and the lower positional specificity of serial order intrusions compared with transposition errors within a list (see also Henson, in press).

The effects of presentation modality, phonemic similarity, and item familiarity on the performance of recognition memory judgments are simulated in the section on recognition memory judgments. These simulations include recognition memory judgments on items from lists of alternating familiarity or similarity, showing zig-zag curves of performance. As far as we know, the simulated effect of all of the above manipulations are predictions (but see Henson, Hartley, Hitch, & Burgess, 1999).

Our simulations of articulatory suppression highlight an intriguing corollary of the independence of performance from word length under articulatory suppression. If performance decreases linearly with word length (as in the phonological loop idea), serial recall of lists of long words (taking more than 1.0 s to say) might be improved by suppression. This should be checked experimentally, but participants may opt to use a system other than the phonological loop, if possible, or not to rehearse when presented with such long stimuli.

Neuropsychological data. Although most lesions to the model impair both serial recall and recognition memory judgments (consistent with the analysis of Starr et al., 1990), some predictions are possible. A lesion to the context–timing system alone would reduce item span for serially ordered recall and the effects of temporal grouping and Hebb repetition, but would not impair performance on tasks such as recognition memory judgments, rhyming judgments, or finding the meaning of a pseudohomophone. Similarly, we would predict the complementary pattern of impairment: patients who have some damage to part of the pathway that provides phonemic feedback (from items to output phonemes to input phonemes to items) and show impaired serially ordered recall and reduced phonemic similarity effects but who do show effects of temporal grouping and Hebb repetition.

An interesting question raised by our model is whether the component responsible for serial ordering in verbal STM plays a similar role in STM for nonverbal items. Neuropsychological evidence suggests a dissociation between verbal and nonverbal span tasks (De Renzi & Nichelli, 1975), and most of our model is specialized for phonological material. However, the context–timing signal is not necessarily as specialized. The same signal might also be involved in nonverbal immediate serial recall tasks such the Corsi blocks task. A common ordering mechanism should lead to similar patterns of order errors in the serial recall of Corsi blocks and verbal sequences, and recent work indicates that such similarities do exist (Jones, Farrand, Stuart, & Morris, 1995), although, as would be expected, the correspondence is not perfect (Smyth & Scholey, 1996). Similarities have also been found in STM for the serial order of olfactory stimuli (White & Treisman, 1997). Similarly, patients with a more general timing deficit might be expected to show deficits in STM for serial order as the context–timing system need not be specialized for use with verbal material, and the CQ mechanism might well apply more generally to the serial ordering of behavior (Houghton & Hartley, 1996).

The model's use of decaying inhibition causes items to become more active over time (similarly to a "primacy gradient"; Page & Norris, 1998), in contrast to the decay of connection weights. Note that using decaying inhibition (or growing activation) to provide such a gradient gives an advantage to long words over short ones. Thus, when a lesion removes the effect of some decaying connection weights, performance on short words can be impaired more than on long ones. This may produce slightly better overall performance on long words than short ones (see Table 3). This result could be translated into a prediction that neuropsychological patients should be found who show such counterintuitive behavior. However, it must be remembered that the model assumes that presentation and recall rates are both equal to the participants' rate of speeded articulation.

Neurophysiological data. It is generally difficult for models of verbal behavior to make contact with neurophysiology data because detailed single-unit data are generally not recorded in humans. However, as mentioned above, parts of our model relate to the serial ordering of behavior in general. In principle, the representation of serial order in the brain might take one (or several) of many forms, such as activations that increase with proximity to the start or end of the sequence (Henson, 1998b; Houghton, 1990); the start of the sequence (Page & Norris, 1998); or the positional representation of the context layer in our current and 1992 model. Which type of representation or representations are in fact used in the brain may be difficult to determine solely on the basis of behavior.

The simplest interpretation of our context signal predicts neurons that are active for specific serial positions in a sequence of items, irrespective of which items are recalled at each position. Interestingly, cells have recently been observed in the presupplementary motor area of monkeys that fire at specific serial positions within a learned sequence of three movements, independently of which movement is made at which position (J. Tanji, personal communication, March 17, 1998). These complement the cells in the supplementary motor area that fire before the performance of the three actions in a specific order (Tanji & Shima, 1994). Cells in primary motor cortex have also been found to encode the serial position of visual stimuli in a task requiring memory for their serial order (Carpenter, Georgopoulos, & Pellizzer, 1999). Both of these observations of the neural coding of serial order correspond closely to the way units represent serial order in our predicted context–timing layer.

Discussion

Locating Functional Modules Within the Brain

A simple mapping from the functional modules in the model to the likely brain areas supporting them would place the input phoneme layer in the supramarginal gyrus (just posterior to Wernicke's area) and the output phoneme layer in Broca's area. Beyond this, however, it is not clear whether to interpret the item layer as corresponding to locations close to Broca's area and the output system or close to Wernicke's area and the input system. In the former case, CQ of items would be providing a simple model of the planning of a speech output (or articulation), the task for which it was originally developed (Houghton, 1990). The cells found in the supplementary motor area of monkeys, whose activity is related to particular ordered sequences of forthcoming multiple movements (Tanji & Shima, 1994), indicate that this area might be expected to contribute to the premotor planning of the output of a list in humans. However, the identification of STM patients such as P.V. and J.B. with posterior lesions is most consistent with the item layer being located near Wernicke's area (and hence associated with speech input processes). This position is also more consistent with functional neuroimaging of lexical processing that has activated inferior parietal and middle and inferior temporal areas on the left (Demonet et al., 1992). In this case, items might simply be recognized as ordered sets of input phonemes (as in James & Miikkulainen, 1995) but could still be sequentially processed in a manner analogous to CQ despite their distributed representation (e.g., similarly to the model of olfactory processing in Granger, Ambrose-Ingerson, Staubli, & Lynch, 1990).

The location and nature of the proposed third component, the repeatable context–timing signal, is an interesting subject for speculation. The recent observation, discussed above, of cells that fire according to the serial position within a sequence of three movements makes the primary motor cortex and presupplementary motor area the most obvious candidates. Treisman et al. (1994) have suggested a central timing signal that can be detected in the EEG and in interval judgments. This could be related to our context–timing signal, although a location for it has not been proposed to our knowledge. Another possible location for at least part of this system is the cerebellum, which has been proposed to have a general timing function (Ivry & Keele, 1989) and has long been associated with smooth motor coordination (see Welsh, Lang, Suglhara, and Llinas, 1995, for a review). More specifically, timed finger-tapping tasks thought to involve the cerebellum, and the nonmotor aspect of subjective estimation of time intervals, have recently been shown to selectively activate parts of the cerebellum in a positron emission tomography study (Jueptner et al., 1995). Other recent work has pointed to a role for the cerebellum in human cognitive function, including language (Leiner, Leiner, & Dow, 1993), but remains highly speculative at this stage (see Daum & Ackermann, 1995, for a review). Interestingly, Paulesu et al. (1993) found that the cerebellum and the supplementary motor area were selectively activated by their list recall task rather than by rhyming judgments. They suggested that these areas "were automatically engaged as part of a more general neuronal network devoted to language planning and execution" (Paulesu et al., 1993, p. 342), a role that would also apply to the timing of serial recall of a list. The location of the context–timing signal is the subject of a current functional neuroimaging study (Henson, Burgess, & Frith, 1999).

Distributed Versus Localist Representations and Biologically Plausibility

The model presented here is localist, one unit in the model represents one piece of information (e.g., an item or phoneme), and does not respond to any other piece of information. Such a unit is intended to correspond to many neurons in the brain, but these neurons would also have to be localist for the correspondence to make sense. Within connectionist modeling, localist representations are sometimes viewed as being less biologically plausible than distributed ones. Is there any basis for this? Given that a particular part of a connectionist model processes a certain type of information, should the information be represented by the activity of individual units in a distributed or a localist fashion? To address this question, we consider the following (artificially polarized) positions: (a) a "distributed" position, where information about the state of the system is distributed across many units, such that each unit is active for a random subset of all the possible states of the system, and such that which states a unit contributes to is independent of what those states represent; (b) a "localist" position, where the activity of a single unit represents some identifiable information. The "grandmother cell hypothesis" (i.e., that the concept of one's grandmother is encoded by an individual cell) is a special case of this position, although, more generally, a particular stimulus might be represented by graded activity among many cells, each with well defined but less specific receptive fields (e.g., female relatives, people wearing glasses, etc.).

In terms of experimental data, *localist* is distinguished from *distributed* by the prediction that activity of a particular neuron can be identified with a single, well-circumscribed experimental event. This type of cellular activity has been routinely found in various brain regions of various animals. Some of the best known examples are cells maximally responding to specific types of visual stimuli (see, e.g., Zeki, 1993), including high-level stimuli such as a face looking in a particular direction (Perrett, Hietanen, Oram, & Benson, 1992); the occurrence of a particular syllabic or subsyllabic component to a bird's song (Yu & Margoliash, 1996); being in a particular spatial location (O'Keefe, 1976); heading in a particular direction (Taube, Muller, & Ranck, 1991); and reaching in a particular direction (Georgopoulos, Kettner, & Schwartz, 1988).

By contrast, we are not aware of direct evidence for distributed representations, although this type of representation would be hard to find because an apparently random neuronal response might be interpreted simply as a failure to identify the correct experimental variable. Part of the allure of distributed representations is that they are highly efficient: Even if neurons had only two types of responses (i.e., on or off), N neurons could encode 2^N bits of information in a distributed fashion. However, it is not clear how important a consideration this is in a brain of around 10^{11} neurons (see, e.g., Kandel, 1991).

In summary, we do not regard localist representations as being less plausible than distributed ones. Indeed, there is much direct evidence for their existence and none for purely distributed ones. Our model is localist, and the functional constraints imposed by the STM data make it hard to envisage an equivalent distributed model (e.g., an item must be suppressed in its entirety following selection, whereas similar items must not be suppressed at all). However, because local models inevitably encode information across several units and apparently random single-unit responses may one day be understood, the question of whether to use distributed or local representations may be one of interpretation: "a matter of whether we are looking at the forest or the trees" (McClelland & Rumelhart, 1986, p. 5).

Relation to Other Work

It is interesting to note the similarities between the present model and Estes's (1972; Lee & Estes, 1981) perturbation theory of serial ordering. Thus, in Estes's models, serial order information is stored via links between items and contextual "control elements," which shift as a function of time and represent positional information at different levels in a hierarchy (e.g., trial, group, and within-groups position). This has a clear parallel with the use of context–item associations in the present model, except that positional information in the present model is not fully hierarchical (i.e., group position is not encoded in the account of temporal grouping effects). Another similarity that highlights an interesting difference is the use of noise to cause errors. Thus, in Estes's model, noise perturbs the context–timing signal, whereas in our model memory for context is error free and noise disturbs item information only during output.

More recently, the primacy model (Page & Norris, 1998) and the oscillator-based associative recall model (OSCAR; Brown et al., in press) have been proposed, and both have strong similarities to our model. OSCAR can be thought of as an implementation of Burgess and Hitch's (1992) model using temporal oscillators to provide the context signal, extending its range of application in the temporal domain but also retaining its flaws. The first stage of the primacy model may be thought of as a minimal CQ model (see the section on the minimalist connectionist model) or in terms of Houghton's (1990) model modified to include a "start node" but without the corresponding "end node." The second stage of the primacy model roughly corresponds to the phonemic feedback in our model but involves a second set of item nodes (with attendant "primacy gradient" of activations) that serves purely to introduce extra errors when there is phonemic similarity among items. Our use of input and output phoneme representations, and the reciprocal connections between them, is necessary to provide a mechanistic explanation of the effects of articulatory suppression, presentation modality, and rehearsal and to account for neuropsychological data. They are also necessary to explain people's ability to learn to both recognize and produce the same sound for a given word (see *Phonological Information: Neuropsychology and Phonemic Feedback*). Such an explanation is not provided by the primacy model. The second difference with the primacy model concerns whether a separate context layer is necessary. We believe that it is, based on the Hebb repetition effect, position-specific intrusions, and temporal grouping effects. Again, these cannot be accounted for by the primacy model.

We note that the use of gradients relating to both the start and end of the list (Houghton, 1990) can be shown to support a model of immediate recall that is consistent with error patterns, including those related to temporal grouping and serial order intrusions (Henson, 1998b). Indeed, serial order intrusions and the pattern of errors in temporally grouped lists argue in favor of start and end node representations (see Henson, in press), although a drawback of end nodes is that they tend to require the system to know which is the last item in advance (however, see Houghton et al., 1994). Thus, some modification of our context signal may be indicated so as to encode serial position relative to the start and end of the list, rather than its absolute value (see Henson & Burgess, 1997).

Apart from the details of implementation, neuropsychological data may turn out to be a good way of deciding between different model structures. We believe these data force models of verbal STM to adopt the separate layers and connections between item, phonemic input, and phonemic output representations in our current model (see also Burgess and Hitch, 1992). We agree with Gupta and MacWhinney's (1997) emphasis on the importance of attempting to tie functional modules to specific brain regions and with their decision to base the item-level structure of their model on our 1992 model and the phoneme-level structure on that of Hartley and Houghton (1996). Here we have provided a different development of our 1992 model, associated its components with brain regions, and showed that it is capable of explaining a wider range of psychological data on verbal STM than any other current model of this type.

Conclusion

We have proposed a mechanistic account of how the brain organizes the storage and immediate serial recall of verbalizable items that is compatible with implementation by neurons and synapses. The functional architecture of the model is motivated by independent psychological and neuropsychological evidence, and our simulations demonstrate that it can account for a wider range of experimental data than any other model of this type. Simulated experiments include the probability of making an error in the immediate recall of a list of items as a function of serial position, list length, word length, phonemic similarity, presentation modality, articulatory suppression, rehearsal, temporal grouping, word frequency, and lexicality. It also models the effects of Hebb repetition and serial order intrusions from previous lists and can be related to the performance of neuropsychological patients. The model builds on our earlier model (Burgess & Hitch, 1992) and shows that the basic ideas can be usefully extended. It proposes an explicit mechanism for the conceptual model of a phonological loop (Baddeley, 1986) and retains that model's successes in explaining effects on overall performance levels while also extending it to explain the detailed error patterns in immediate serial recall. Our extended model predicts novel psychological experiments because many of the simulated error patterns have not yet been checked experimentally. We have also attempted to map the functional modules in the model onto brain regions in order to make contact with data from neuropsychology and functional neuroimaging. We hope that the model presented here is a starting point for understanding the role of verbal STM in language learning and will provide a framework for the prediction and interpretation of neuropsychological and functional neuroimaging data.

References

Aaronson, D. (1968). Temporal course of perception in an immediate recall task. *Journal of Experimental Psychology, 76,* 129–140.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362–365.

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 20*(A), 249–264.

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

Baddeley, A. D. (1990). The development of the concept of working memory: Implications and contributions of neuropsychology. In G. Vallar & T. Shallice, *Neuropsychological impairments of short-term memory* (pp. 54–73). Cambridge, England: Cambridge University Press.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 5–28.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105,* 158–173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *Recent Advances in the Psychology of Learning and Motivation, 1,* pp 47–90.

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

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.

Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology, 36,* 707–711.

Bienenstock, E. L., Cooper, L. N., & Munro, P. W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience, 2,* 32–48.

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

Bliss, T. V. P., & Collingridge, G. L. (1993). A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature, 361,* 31–39.

Blumstein, S. E. (1995). The neurobiology of the sound structure of language. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 915–929). Cambridge, MA: MIT Press.

Bower, G. H., & Winzenz, G. (1969). Group structure, coding and memory for digit series. *Journal of Experimental Psychology Monographs, 80*(2, Pt. 2).

Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed. *Journal of Memory and Language, 34,* 594–621.

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

Burani, C., Vallar, G., & Bottini, G. (1991). Articulatory coding and phonological judgements on written words and pictures: The role of the phonological output buffer. *European Journal of Cognitive Psychology, 3,* 379–398.

Burgess, N. (1995). A solvable connectionist model of immediate recall of ordered lists. *Advances in Neural Information Processing Systems, 7,* 51–58.

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

Burgess, N., & Hitch, G. J. (1996). A connectionist model of STM for serial order. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 51–72). Hillsdale, NJ: Erlbaum.

Campbell, R. (1990). Lipreading, neuropsychology, and immediate memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 268–286). Cambridge, England: Cambridge University Press.

Carpenter, A. F., Georgopoulos, A. P., & Pellizzer, G. (1999, March 12). Motor cortical encoding of serial order in a context-recall task. *Science, 283,* 1752–1757.

Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition, 37,* 55–81.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 568–608.

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

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

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

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

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

Crowder, R. G. (1972). Visual and auditory memory. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye* (pp. 251–276). Cambridge, MA: MIT Press.

Crowder, R. G., & Morton, J. (1969). Precatagorical acoustic storage (PAS). *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.

Daum, I., & Ackermann, H. (1995). Cerebellar contributions to cognition. *Behavioural Brain Research, 67,* 201–210.

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17,* 149–195.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115,* 1753–1768.

De Renzi, E., & Nichelli, P. (1975). Verbal and non-verbal short-term memory impair following hemispheric damage. *Cortex, 11,* 341–354.

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

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word length effect: Implications for intelligence testing and the ease of mental calculation in Welsh and English. *British Journal of Psychology, 71,* 43–51.

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

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. *American Scientist, 70,* 607–615.

Eriksen, C. W., Coles, M. G. H., Morris, C. L. R., & O'Hara, W. P. (1985). An electromyographic examination of response competition. *Bulletin of the Psychonomic Society, 23,* 165–168.

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

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

Gardiner, J. M. (1983). On recency and echoic memory. *Philosophical Transactions of the Royal Society of London, Series B, B302,* 267–282.

Gardner-Medwin, A. R. (1989). Doubly modifiable synapses: A model of short and long term auto-associative memory. *Proceedings of the Royal Society of London, Series B, B238,* 137–154.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12,* 349–367.

Georgopoulos, A. P., Kettner, R. E., & Schwartz, A. B. (1988). Primate motor cortex and free arm movements to visual targets in three-dimensional space: II. Coding of the direction of movement by a neuronal population. *Journal of Neuroscience, 8,* 2928–2937.

Glasspool, D. W. (1995). Competitive queuing and the articulatory loop. In J. Levy, D. Bairaktaris, J. Bullinaria, & P. Cairns (Eds.), *Connectionist models of memory and language* (pp. 5–30). London: UCL Press.

Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 16–31.

Goodglass, H., & Kaplan, E. (1972). The assessment of aphasia and related disorders. Philadelphia: Lea & Febiger.

Granger, R., Ambrose-Ingerson, J., Staubli, U., & Lynch, G. (1990). Memorial operation of multiple interacting simulated brain structures. In M. A. Gluck & D. E. Rumelhart (Eds.), *Neuroscience and connectionist theory* (pp. 95–129). Hillsdale, NJ: Erlbaum.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science, 11,* 23–63.

Guildford, J. P., & Dallenbach, K. M. (1925). The determination of memory span by the method of constant stimuli. *American Journal of Psychology, 36,* 621–628.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59,* 267–333.

Hartley, T. (1999). *Syllabic phase: Bottom-up representations of the structural properties of speech.* Manuscript in preparation.

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.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye, A. Fessard, R. W. Gerard, & J. Konorski, (Eds.) *Brain mechanisms and learning: A symposium* (pp. 37–46). Oxford: Blackwell Scientific Press.

Heffernan, T. M. (1991). *Measurement of working memory capacity in children.* Unpublished doctoral dissertation, Manchester University, Manchester, England.

Henson, R. N. A. (1998a). Item repetitions in short-term memory: Ranschburg repeated. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1162–1181.

Henson, R. N. A. (1998b). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36,* 73–137.

Henson, R. N. A. (in press). Positional information in short-term memory: Relative or absolute? *Memory & Cognition.*

Henson, R. N. A., & Burgess, N. (1997). Representations of serial order. In J. A. Bullinaria, D. W. Glasspool, & G. Houghton (Eds), *4th Neural Computation and Psychology Workshop, London 9–11 April 1997: Connectionist representations.* New York: Springer.

Henson, R. N. A., Burgess, N., & Frith, C. D. (1999). *Recoding, storage, rehearsal and grouping in phonological short-term memory: An fMRI study.* Manuscript submitted for publication.

Henson, R. N. A., Hartley, T., Hitch, G. J., & Burgess, N. (1999). *Selective interference with verbal short-term memory for item and order information.* Manuscript in preparation.

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

Hitch, G. J. (1975). The role of attention in visual and auditory suffix effects. *Memory & Cognition, 3,* 501–505.

Hitch, G., Burgess, N., Towse, J., & Culpin, V. (1996). Grouping and immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 116–139.

Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London, Series B, B302,* 325–340.

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., Glasspool, D. W., & Shallice, T. (1994). Spelling and serial recall: Insights from a competitive queuing model. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 365–404). New York: Wiley.

Houghton, G., & Hartley, T. (1996). Parallel models of serial behaviour: Lashley revisited. *Psyche, 2(25).*

Howard, D., & Franklin, S. (1987). Three ways for understanding written words, and their use in two contrasting cases of surface dyslexia. In D. A. Allport, D. MacKay, W. Prinz, & E. Scheere (Eds.), *Language perception and production: Relationships between listening, speaking, reading and writing* (pp. 340–366). London: Academic Press.

Howard, D., & Franklin, S. (1990). Memory without rehearsal. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 287–318). Cambridge, England: Cambridge University Press.

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., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1217–1232.

Hulme, C., Thompson, N., Muir, C., & Lawrence, A. L. (1984). Speech rate and the development of short-term memory. *Journal of Experimental Child Psychology, 38,* 241–253.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience, 1,* 136–152.

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

James, D. L., & Miikkulainen, R. (1995). SARDNET: A self-organizing feature map for sequences. *Advances in Neural Information Processing Systems, 5,* 577–584.

Jones, D. M., Farrand P., Stuart, G. P., & 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.

Jordan, M. I. (1986). Attractor dynamics and parallelism in a connectionist sequential machine. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (pp. 531–546). Hillsdale, NJ: Erlbaum.

Jueptner, M., Rijntjes, M., Weiller, C., Faiss, J. H., Timmann, D., Mueller, S. P., & Diener, H. C. (1995). Localization of a cerebellar timing process using PET. *Neurology, 45,* 1540–1545.

Kandel, E. R. (1991). Nerve cells and behavior. In E. R. Kandel, J. H.

Schwarz, & T. M. Jessell (Eds.), *Principles of neural science* (pp. 18–32). New York: Elsevier.

Kleinfeld, D. (1986). Sequential state generation by model neural networks. *Proceedings of the National Academy of Science, USA, 83,* 9469.

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

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

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1993). Cognitive and language functions of the human cerebellum. *Trends in Neurosciences, 16,* 444–447.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122–142.

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

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

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.

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

McClelland, J. L., & Rumelhart, D. E. (Eds.). (1986). *Parallel distributed processing* (Vol. 2). Cambridge, MA: MIT Press.

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.

McLeod, P., & Posner, M. I. (1984). Privileged loops from percept to act. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 55–66). Hillsdale, NJ: Erlbaum.

McNaughton, B. L. (1982). Long-term synaptic enhancement and short-term post-tetanic potentiation in the rat fascia dentata act through different mechanisms. *Journal of Physiology, 324,* 249–262.

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.

Michel, F. (1979). Préservation du language "ecrit malgré un déficit majeur du language oral. *Lyon Medical, 241,* 141–149.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

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

Morton, J., & Holloway, C. M. (1970). Absence of a cross-model "suffix effect" in short-term memory. *Quarterly Journal of Experimental Psychology, 22,* 167–176.

Morton, J., & Patterson, K. E. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 91–118). London: Routledge.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78,* 679–684.

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

Nairne, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science, 3,* 199–202.

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

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., & Nairne, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review, 2,* 429–441.

Neumann, C. (1994). *A feature-based approach to serial recall effects in immediate memory.* Unpublished master's thesis, Purdue University, West Lafayette, IN.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York: Plenum.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental Neurology, 51,* 78–109.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761–781.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature, 362,* 342–344.

Penney, C. G., & Blackwood, P. A. (1989). Recall mode and recency in immediate serial-recall: Computer users beware. *Bulletin of the Psychonomic Society, 27,* 545–547.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series, B, B335,* 23–30.

Plaut, D. C., & Shallice, T. (1993). Perseveration and semantic influences on visual object naming errors in optic aphasia: A connectionist account. *Journal of Cognitive Neuroscience, 5,* 89–117.

Rumelhart, D. E., & Norman, D. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science, 6,* 1–36.

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

Saint-Aubin, J., & Poirier, R. (1999). Semantic similarity and immediate serial recall: Is there a detrimental effect on order information? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 52*(A), 367–394.

Salamé, P., & Baddeley, A. D. (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.

Salway, A. F. S., & Logi, R. H. (1995). Visuospatial working-memory, movement control and executive demands. *British Journal of Psychology, 86,* 253–269.

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition, 21,* 168–175.

Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number of magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 419–425.

Schweickert, R., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science, 1,* 74–77.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge, England: Cambridge University Press.

Shallice, T., & Butterworth, B. (1977). Short-term memory impairment and spontaneous speech. *Neuropsychologia, 15,* 729–735.

Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech-perception and production separate processes? *Quarterly Journal of Experimental Psychology, 37*(A), 507–532.

Shallice, T., & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 11–53). Cambridge, England: Cambridge University Press.

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

Simon, H. A. (1974, February). How big is a chunk? *Science, 183,* 482–488.

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

Sompolinsky, H., & Kanter, I. (1986). Temporal association in asymmetric neural networks. *Physical Review Letters, 57,* 2861.

Spencer, A. (1996). *Phonology: Theory and description.* Oxford, England: Blackwell.

Starr, A., Barrett, G., Pratt, H., Michalewski, H. J., Patterson, J. V. (1990). Electrophysiological measures of short-term memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 94–110). Cambridge, England: Cambridge University Press.

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

Stigler, J. W., Lee, S., & Stevenson, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited tore. *Cognition, 23,* 1–20.

Tanji, J., & Shima, K. (1994). Role for supplementary motor area cells in planning several movements ahead. *Nature, 371,* 413–416.

Taube, J. S., Muller, R. U., & Ranck, J. B. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats: I. Description and quantitative analysis. *Journal of Neuroscience, 10,* 420–435.

Treisman, M., Cook, N., Naish, P. L. N., & McCrone, J. K. (1994). The internal clock: Electroencephalographic evidence for oscillatory processes underlying time perception. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47*(A), 241–289.

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

Vallar, G., & Baddeley, A. D. (1984a). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior, 23,* 151–161.

Vallar, G., & Baddeley, A. D. (1984b). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case-study. *Cognitive Neuropsychology, 1,* 121–141.

Vallar, G., & Papagno, C. (1995). Neuropsychological impairments of short-term memory. In A. D. Baddeley, B. A. Wilson & F. Watts (Eds.), *Handbook of memory disorders* (pp. 135–165). New York: Wiley.

Vallar, G., Di Betta, A. M., & Silveri, M. C. (1997). The phonological short-term store-rehearsal system: Patterns of impairment and neural correlates. *Neuropsychologia, 35,* 795–812.

Vandierendonck, A., De Vooght, F., & Van der Goten, K. (1998). Interfering with the central executive by means of a random interval repetition task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 51*(A), 197–218.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain, 92,* 885–896.

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

Watkins, M. J., Watkins, O. C., & Crowder, R. G. (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.

Welsh, J. P., Lang, E. J., Suglhara, I., & Llinas, R. (1995). Dynamic organization of motor control within the olivocerebellar system. *Nature, 374,* 453–457.

White, T. L., & Treisman, M. (1997). A comparison of the encoding of content and order in olfactory memory and in memory for visually presented verbal materials. *British Journal of Psychology, 88,* 459–472.

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

Yu, A. C., & Margoliash, D. (1996, September 27). Temporal hierarchical control of singing in birds. *Science, 273,* 1871–1875.

Zeki, S. (1993). *A vision of the brain.* Oxford: Blackwell Scientific Press.

Zhang, G., & Simon, H. A. (1995). STM capacity for Chinese words and idioms: Chunking and acoustic loop hypotheses. *Memory & Cognition, 13,* 193–202.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1131–1161.

Received February 28, 1996
Revision received December 7, 1998
Accepted December 7, 1998 ■