Journal of Memory and Language 61 (2009) 97–111

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Slave to the rhythm: Experimental tests of a model for verbal short-term memory and long-term sequence learning

Graham J. Hitch [a,*], Brenda Flude [b], Neil Burgess [c]

[a] Department of Psychology, University of York, Heslington, York, YO10 5DD, United Kingdom
[b] Lancaster University, Lancaster, LA1 4YF, United Kingdom
[c] Institute of Cognitive Neuroscience and Department of Anatomy, University College London, 17, Queen Square, London, WC1N 3AR, United Kingdom

ARTICLE INFO

Article history:
Received 15 July 2005
revision received 13 February 2009
Available online 19 April 2009

Keywords:
Long-term memory
Short-term memory
Learning
Phonological loop
Working memory
Hebb Effect
Computational modelling
Vocabulary acquisition
Serial order

ABSTRACT

Three experiments tested predictions of a neural network model of phonological short-term memory that assumes separate representations for order and item information, order being coded via a context-timing signal [Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551–581005D]. Predictions were generated for long-term sequence learning and tested using the Hebb Effect, the improvement in immediate serial recall when a list is repeated. Results confirmed predictions that the Hebb Effect would be (1) insensitive to phonemic similarity and articulatory suppression, variables that impair immediate recall without affecting the context-timing signal and (2) reduced if the context-timing signal is altered by varying the temporal grouping pattern of the repeated list. Results highlighted an interesting shortcoming of the model in that participants were able to learn more than one sequence simultaneously. However, this problem was addressed by extending the model to include multiple context representations and a sequence-recognition process.

Introduction

According to the model of working memory proposed by Baddeley and Hitch (1974) short-term memory (STM) for a sequence of verbal items depends on a speech-based storage system known as the phonological loop (see also Baddeley, 1986, 2007). A wide range of empirical evidence suggests that an important function of this store is to support vocabulary acquisition (Baddeley, Gathercole, & Papagno, 1998), the idea being that a new word form must be held in short-term phonological storage in order for long-term learning to take place. However, the theoretical concept of the phonological loop requires elaboration if it is to offer an explanation of how such long-term learning takes place. This is primarily because it does not address the twin problems of how order information is represented in STM and how STM and long-term memory (LTM) interact. Burgess and Hitch (1999) described an implementation of the phonological loop as a neural network that included explicit mechanisms for both serial order and long-term learning. This was a development of an earlier neural network model that addressed serial order but not long-term learning (Burgess & Hitch, 1992). In the present article we report experiments that test qualitative predictions of the Burgess and Hitch (1999) model for sequence learning. The results broadly confirm the predictions but draw attention to the need to revise the model so as to be capable of learning multiple sequences without massive interference.

We begin by briefly describing the concept of the phonological loop. We then explain how the network model of Burgess and Hitch (1999) addresses serial order and long-term learning and show how the model can be used to predict the sensitivity of sequence learning to variables known to influence short-term recall. These predictions depend critically on whether the variable affects item or order information. We then introduce an experimental

* Corresponding author. Fax: +44 (0) 1904 433181.
E-mail address: gjh3@york.ac.uk (G.J. Hitch).

0749-596X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2009.02.004

procedure due to Hebb (1961) as a suitable vehicle for testing these predictions.

Phonological loop

As initially proposed, the phonological loop consists of two components: a phonological store that decays over time and a subvocal rehearsal mechanism capable of refreshing the contents of the store (Baddeley, 1986). This simplistic conceptual model was used to explain the effects of a cluster of related variables on verbal STM and their interactions, principally word length (Baddeley, Thomson, & Buchanan, 1975), phonemic similarity (Baddeley, 1966; Conrad & Hull, 1964) and articulatory suppression (Murray, 1967). The tendency for lists of long words to be harder to recall than short words was attributed to the extra time taken to rehearse long words and refresh their decaying memory traces. The tendency for lists of similar-sounding items to be less well recalled than dissimilar items was explained by the extra difficulty of discriminating partially decayed traces of similar items in the phonological store. Articulatory suppression involves repeating an irrelevant utterance and is used as a secondary task to disrupt subvocal rehearsal (Murray, 1967). Suppression impairs immediate serial recall and removes the word length and phonemic similarity effects (Baddeley et al., 1975; Murray, 1968; Peterson & Johnson, 1971), consistent with it disrupting the phonological loop. The precise pattern of interactions depends on presentation modality. Thus, suppression removes the word length effect for visual and auditory lists but removes the phonemic similarity effect only for visual lists (Baddeley, Lewis, & Vallar, 1984). This effect of presentation modality was explained by a modest elaboration of the model, such that auditory items access the phonological store automatically whereas visual items have first to be verbally recoded, suppression being assumed to block the recoding process in addition to preventing subvocal rehearsal.

The concept of the phonological loop has proved influential and has been successfully applied in areas beyond its initial remit, including child development, developmental disorders, neuropsychology and neuroimaging (see Baddeley, 2007). The theory has nevertheless attracted numerous challenges. These include the explanation of the word length effect (e.g., Caplan, Rochon, & Waters, 1992; Caplan & Waters, 1994; Lovatt, Avons, & Masterson, 2000; Service, 1998), the phonemic similarity effect (Jones, Macken, & Nicholls, 2004), and the idea of short-term forgetting as time-based decay (e.g., Lewandowsky, Duncan, & Brown, 2004). There has also been controversy about the use of the phonological loop to explain the way unattended irrelevant sounds affect immediate recall (e.g., Jones, Hughes, & Macken, 2006) and there are competing theoretical accounts (e.g., Nairne, 2002; Neath, 2000). The phonological loop theory has been defended against these challenges (Baddeley, 2007; see also Mueller, Seymour, Kieras, & Meyer, 2003), but these arguments are tangential to our present focus. We are concerned here with two entirely uncontroversial shortcomings of the theory, namely its omission of mechanisms for serial order and long-term learning.

The absence of a mechanism for serial order prevents the phonological loop giving an adequate explanation for errors where an item is recalled correctly but in the wrong position in the list (Conrad, 1965). Such order errors are highly characteristic of immediate serial recall. They typically involve items migrating to adjacent positions, and are a principal determinant of the bow-shaped serial position curve (e.g., Henson, 1998). Order errors typically increase when items are phonemically similar (Conrad, 1965) and decrease when items are presented in rhythmic temporal groups, for example by adding an extra pause after every third item (Frankish, 1985; Ryan, 1969). Temporally grouping item presentation results in multiply-bowed serial position curves and a change in the distribution of order errors suggesting that order is coded at separate levels, between and within groups (McNicol & Heathcote, 1986).

The absence of a mechanism for long-term learning prevents the phonological loop from explaining numerous well-established effects of long-term knowledge in immediate serial recall, such as differences between words and nonwords and more subtle linguistic frequency effects (e.g., Gathercole, 1995; Gathercole, Pickering, Hall, & Peaker, 2001; Hulme, Maughan, & Brown, 1991). We have already noted that theoretical elaboration is also required to explain how short-term phonological storage contributes to vocabulary acquisition (Baddeley et al., 1998). In a recent revision of the working memory model, Baddeley (2000a) included an explicit link between the phonological loop and linguistic knowledge. However, the revision is pitched broadly and does not specify how the proposed link operates, making it difficult to generate testable predictions.

Recently, a number of attempts have been made to develop computational models of verbal short-term memory that go beyond the phonological loop theory by specifying mechanisms in more detail (see e.g., Botvinick & Plaut, 2006; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1992, 1999; Farrell & Lewandowsky, 2002; Henson, 1998; Page & Norris, 1998). In general, these models have concentrated on addressing the problem of serial order, paying less attention to interactions between STM and LTM. Two exceptions that deal with both serial order and long-term learning are the connectionist models proposed by Botvinick and Plaut (2006) and Burgess and Hitch (1999). We focus here on our own model (Burgess & Hitch, 1999) showing first how a general understanding of the model can be used to derive novel predictions about sequence learning and then evaluating the model by testing these predictions experimentally. First we outline the way the model operates.

The Burgess and Hitch (1999) model implements the phonological loop as a localist neural network with two main components, a phonological/lexical store for item information and a context-timing signal that encodes the serial order of items. Phonological, lexical and timing information are represented in separate layers of nodes (see Fig. 1). Each node can transmit activation to nodes in adjacent layers according to the strengths of connections between them. Learning and forgetting occur through increases and decreases in the strengths of modifiable

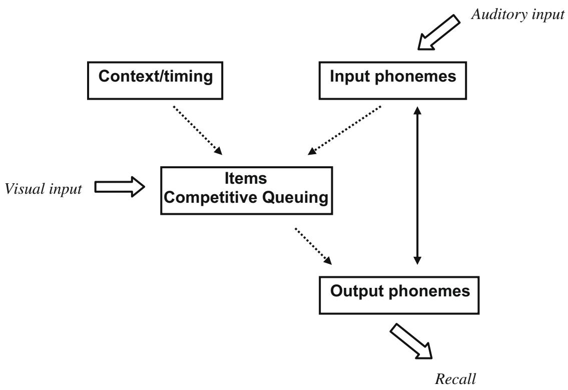

Fig. 1. Outline of the Burgess and Hitch (1999) model. Boxes denote layers of nodes, text in boxes denotes what the nodes represent. Nodes in adjacent layers are multiply interconnected. Modifiable connections are shown as dashed lines. Pre-wired connections are shown by a solid line. External arrows denote where visual and auditory inputs access the system and where recall is emitted.

connections (shown as dotted lines in Fig. 1). Presentation of an item activates its node in the item layer and triggers a learning process whereby connections between simultaneously active nodes in adjacent layers are strengthened. The process of strengthening connections has two components, one large-amplitude and short-lived, the other small-amplitude and long-lived. The short-lived process decays rapidly and is responsible for short-term forgetting. The long-lived process decays slowly and allows long-term learning to accumulate when an input is repeated, provided the repetitions are not too far apart.

In more detail, the activation of nodes in the context-timing layer changes continuously over time such that patterns of activation at adjacent time-steps overlap (see also Burgess, 1995; Henson & Burgess, 1997). As a result, the order of items is encoded in the form of broadly-tuned position-item associations. Presentation of an item also strengthens item-phoneme and phoneme-item connections. However, for familiar items such as digits, letters or words, only the short-term component of item-phoneme and phoneme-item connections changes: the long-term component is already saturated through pre-experimental learning.

When starting to recall a sequence, the context layer is given the pattern of activation it received during presentation of the first item in the list. Activation spreads to the item layer and from there to the output and then input phoneme layers, finally feeding back to the item layer again. At this point the item node with the strongest resultant activation is selected for recall. Immediately after its recall this item is strongly inhibited so as to prevent it from dominating subsequent steps in the recall process. This way of driving sequential output from a parallel set of activations is a feature of several models of serial order and is known as competitive queuing (see Grossberg, 1987; Henson, 1998; Houghton, 1990; Page & Norris, 1998). At the next step in recall, the context layer is given the pattern of activation it received during presentation of the second list item and the process is repeated, cycling on in this way until the end of the list is reached. Errors in recall result from the combined effects of decay of the connection weights and noise in activations at the competitive queuing output stage. A feature of the model is that recall of each item is a two-stage process in which the first stage involves processing serial order information and the second stage involves processing phonological information. The Primacy Model of Page and Norris (1998) also involves two such stages, but differs principally in assuming that serial order is cued by a primacy gradient of activation over items rather than position-item associations.

In summary, the Burgess and Hitch (1999) model embodies the concept of the phonological loop in the phonological-lexical part of a neural network whose connection weights undergo decay in time. A separate context-timing signal encodes serial order information by strengthening broadly-tuned position-item associations that undergo both short-term and long-term decay. These features allow the model to address serial order effects in both long-term learning and short-term recall.

Burgess and Hitch (1999) report simulations of immediate serial recall that reproduce effects of word length, phonemic similarity and articulatory suppression much as in Baddeley's (1986) conceptual model of the phonological loop. Thus, the word length effect arises because longer words take longer to recall, allowing more time for short-term connection weights to decay. The phonemic similarity effect arises from cross-talk in phonemic feedback during the second stage of the item-retrieval process producing errors where one item is substituted by another, as in human data (Conrad, 1965). It is important to appreciate that according to the model, these errors reflect the loss of phonemic and not order information, even though they appear empirically as order errors. Articulatory suppression is approximated by adding noise to nodes in the output phoneme layer. As a result, articulatory suppression impairs recall and removes its dependence on word length, as in human data. Modality effects are captured by assuming that auditory stimuli activate the input phoneme layer directly via an acoustic input buffer whereas visual stimuli access this layer indirectly via a route that involves the item layer and then the output layer (see Fig. 1). Thus, articulatory suppression prevents visual items from activating both the input and output phoneme layers but only prevents auditory items from activating the output phoneme layer. In this way, suppression abolishes the effect of phonological similarity for visual but not auditory presentation, as in human data (see simulations in Burgess & Hitch, 1999, Table 1).

Further simulations confirm that the network model generates effects of serial order and LTM that were not addressed by the phonological loop concept. For example, serial position curves and effects of temporal grouping in immediate serial recall are generated by the way the context-timing signal operates. Previous authors have suggested that serial position effects arise from differences in the distinctiveness of positional cues (e.g., Lee & Estes, 1981; Murdoch, 1960). The network model works in a similar way in that the context-timing signals are more distinctive for items at the ends of a list. Temporally grouped presentation of list items is assumed to recruit an additional set of units in the context layer whose pattern of activation tracks position within group (Henson & Burgess, 1997; Hitch, Burgess, Towse, & Culpin, 1996). The second component modulates the first, thus providing a more distinctive, two-dimensional coding of position relative to ungrouped presentation. The result is a large reduction in order errors accompanied by a small increase in interposition errors, where an item migrates from one group to a corresponding position in another group. In addition, the serial position curve has a scalloped shape that reflects the pattern of grouping. These are all characteristic features of human performance (Ryan, 1969). Temporal grouping effects have been found to be insensitive to phonemic similarity, word length and articulatory suppression, consistent with the model's major assumption that item (phonological) and order (timing) information are represented separately (Hitch et al., 1996).

Turning to long-term sequence learning, a basic requirement is to explain why prior-list intrusion errors in immediate serial recall tend to come from the corresponding serial position in the preceding list (Conrad, 1960; Henson, 1999). According to the Burgess and Hitch (1999) model a position-item association strengthened on a previous trial will influence recall at the same position in a later list if the slow component of its connection weight has not fully decayed. When an entire list of familiar items is repeated several times, long-term learning is assumed to result from cumulative changes to the slow weights of context-item connections. The present experiments sought to test this account using a procedure developed by Hebb (1961), consisting of an immediate serial recall task in which the same list was repeated every third trial. The intervening trials involved random 'Filler' lists. Even though participants were not informed about the experimental manipulation, recall of the critical list improved with repetitions, a phenomenon now known as the Hebb Effect. Given that Hebb used lists of digits, participants learned primarily order rather than item information. For this reason the Hebb Effect is a useful vehicle for examining theoretical assumptions concerning how serial order is encoded and learned.

Network model predictions for the Hebb Effect

As noted above, the Burgess and Hitch (1999) model assumes that learning a list of familiar items involves strengthening context-item connections through cumulative changes in their slow weights. The assumption that slow weights decay provides a natural account of the observation that the Hebb Effect disappears when the gap separating list repetitions is made sufficiently long (Melton, 1963). Of more interest here is the model's prediction that the Hebb Effect will be sensitive to experimental manipulations affecting the context-timing signal. For example, altering the timing of items in each presentation of a repeated list would be expected to impair sequence learning. In contrast, the model predicts that manipulations affecting the phoneme nodes will have little or no influence on the Hebb Effect for lists of familiar items. This is because the learning in question is of order and not items, and because item-phoneme and phoneme-item connections are saturated for familiar items. Thus, variables such as phonemic similarity and articulatory suppression are not expected to affect the rate of long-term learning, despite having their usual disruptive effects on immediate recall. The present experiments tested these predictions.

Experiment 1

We began by exploring the influence of articulatory suppression on the Hebb Effect. As argued above, the Burgess and Hitch (1999) model predicts that suppression will impair immediate recall by disturbing phonemic representations, but will have no effect on sequence learning because it does not affect context-item associations. There are few competing accounts of the Hebb Effect with which to compare these predictions. One pertinent suggestion is that long-term learning depends on 'active rehearsal' of the repeated sequence (Cunningham, Healy, & Williams, 1984), though the precise nature of this rehearsal process was left unclear. Cunningham et al. based their suggestion on their finding that presentations of a sequence did not lead to long-term learning when not accompanied by attempts at recall (see also Cohen & Johansson, 1967). In the present experiment articulatory suppression was required during both presentation and recall of sequences. If long-term sequence learning depends on subvocal rehearsal it should be reduced or even disappear under these conditions.

Method

Participants

Eighteen members of Lancaster University (14 females and four males) took part in the experiment in return for a small payment.

Stimuli and apparatus

Stimuli were 12-item lists made up from the digits 1–9. This list length was chosen to match a previous study of the Hebb Effect by Bower and Winzenz (1969, Experiment 3). Lists were constructed by randomly selecting items subject to the constraints that consecutive digits were never the same and ascending or descending runs of more than two digits were not allowed. In each list there was no more than one run and no digit appeared more than twice. Sixty four lists were presented. Each block of eight lists had an *aebecede* structure, where *a,b,c,d* are non-repeated ('Filler') lists and *e* is the repeated list. A Sony Walkman Professional was used to record the lists onto audio tapes and presentation was by a Rotel amplifier connected to a Technics speaker. Each list was preceded by the word 'ready' and items were presented at a rate of 1 item per second. There was a silent interval of about 12 s between lists.

Design and procedure

Secondary Task (Suppression, No Suppression), Type of List (Repeated, Filler) and Position in Block (1,2,3,4) were manipulated in a $2 \times 2 \times 4$ within-subjects factorial design. Position in Block refers to successive pairs of Filler and Repeated lists. Thus position 1 refers to data from the first Filler and Repeated lists in each block, position 2 to the second, and so on.

Participants were tested individually. They were told they would hear lists of 12 items containing the digits 1–9. Immediately after list presentation they were to try and write down the digits in the same order on a response sheet containing eight rows of 12 blank squares. Responses were written from left to right without retracing. Participants were instructed to guess or leave a blank for any item they could not recall. There was normally a fixed interval of 12 s between trials but an extra pause was allowed on a few occasions when necessary. Previous responses were covered with a piece of card before presentation of the next list. In the articulatory suppression condition, participants were instructed to repeat "the, the, the,..." audibly during both presentation and recall of the lists at a rate of 3–4 utterances per s.

The 64 lists were divided into two halves. Order of presenting these halves and assigning them to the Suppression and No Suppression conditions was counterbalanced. Two practice trials were given before the start of each condition. At the end of the experiment participants were asked if they noticed anything unusual about the lists presented in each half.

Results

Recall was scored by totalling the number of errors made at each serial position for each list. Table 1 and Fig. 2 summarise performance. The main results can be seen by inspection: (i) errors on Repeated lists decreased with Position in Block; (ii) errors on Filler lists increased slightly with Position in Block and (iii) articulatory suppression increased errors overall but did not seem to have any other effect.

Errors in list recall were entered into a $2 \times 2 \times 4$ analysis of variance with factors Secondary Task (No Suppression, Suppression), Type of List (Repeated, Filler) and Position in Block (1–4). There were significant main effects of Type of List, $F(1, 17) = 19.1$, MSE = 15.9, $p < .001$, and Position in Block, $F(3, 51) = 6.07$, MSE = .69, $p < .001$, together with a significant interaction between these factors, $F(3, 51) = 11.1$, MSE = 1.23, $p < .001$. The interaction reflects the learning of Repeated lists relative to Filler lists. There was a significant main effect of Secondary Task, $F(1, 17) = 55.4$, MSE = 2.59, $p < .001$, reflecting more errors with suppression ($M = 7.39$) than without ($M = 5.98$).

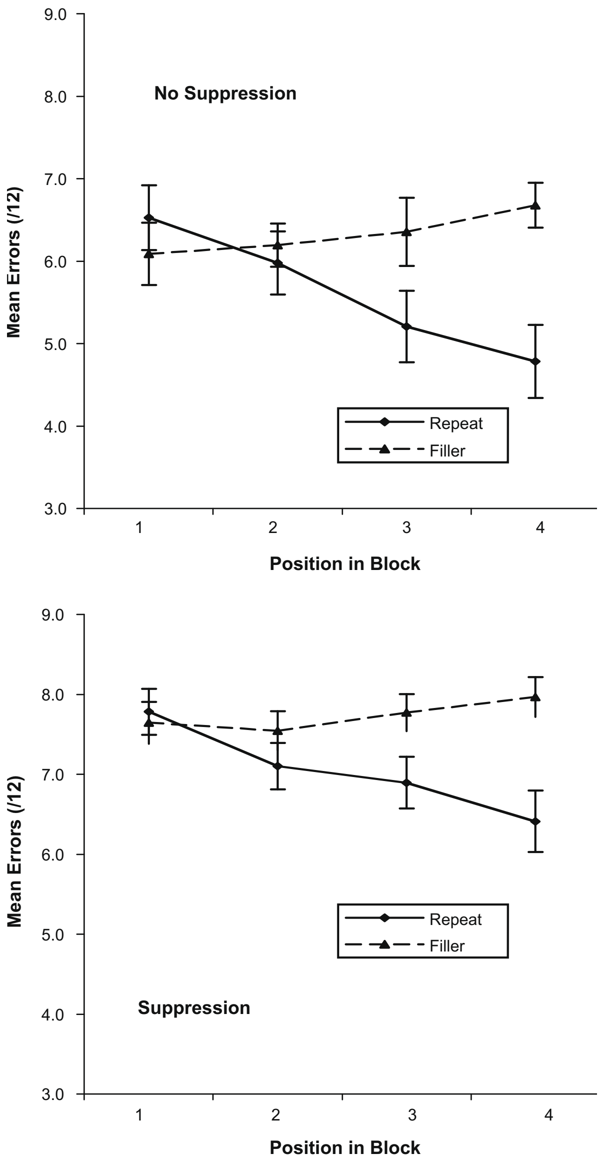

Fig. 2. Experiment 1: immediate serial recall of an auditory sequence of 12 digits with or without articulatory suppression as a function of trial Position in Block, for Repeated (Repeat) and control (Filler) lists. See text for details.

However, none of the interactions involving Secondary Task was significant, $F < 1$ in each case. Separate analyses of variance for the two suppression conditions confirmed that the 2-way interaction between Type of List and Position in Block was significant in each case. With no suppression, the simple main effect of Position in Block was significant for Repeated lists, $F(3, 51) = 11.7$, MSE = .94, $p < .001$, but not for Filler lists, $F(3, 51) < 1$. Similarly, with suppression the simple main effect of Position in Block was significant for Repeated lists, $F(3, 51) = 7.16$, MSE = .82, $p < .001$, and not for Filler lists, $F(3, 51) = 1.11$, MSE = .54.

In further analyses the errors on each list were subdivided into omissions, where no response was made, and

Table 1
Experiment 1: mean (*SD*) errors (max = 12) for Repeated and Filler lists as a function of position in trial-block and articulatory suppression.

	Position in Block				
	1	2	3	4	Mean
No Suppression					
Repeat	6.5 (1.7)	6.0 (1.6)	5.2 (1.8)	4.8 (1.9)	5.6 (1.9)
Filler	6.1 (1.6)	6.2 (1.1)	6.4 (1.8)	6.7 (1.2)	6.3 (1.4)
Suppression					
Repeat	7.8 (1.2)	7.1 (1.2)	6.9 (1.4)	6.4 (1.6)	7.1 (1.4)
Filler	7.6 (1.1)	7.5 (1.1)	7.8 (1.0)	8.0 (11)	7.7 (1.0)

substitutions, where an incorrect digit was recalled. Substitution errors were virtually all order errors as participants seldom responded with an item outside the set of digits 1–9. The two types of error were approximately equally frequent with omissions accounting for 48% of total errors. The totals for each type of error were entered into separate $2 \times 2 \times 4$ analyses with factors Secondary Task (No Suppression, Suppression), Type of List (Repeated, Filler) and Position in Block (1–4). The experimental effects on total errors were almost entirely due to omissions, the only significant effect to emerge in the analysis of substitution errors being the two-way interaction between List Type and Position in Block, $F(3, 51) = 4.49$, MSE = 19.54, $p < .01$. This interaction was due to a small non-significant decrease in substitution errors for Repeated lists (13%) coupled with a small non-significant increase (10%) for Filler lists.

Further analyses examined serial position effects (see Fig. 3). Correct-in-position responses were subjected to a $2 \times 2 \times 4 \times 12$ analysis of variance with factors Secondary Task (No Suppression, Suppression), Type of List (Repeated, Filler), Position in Block (1–4) and Serial Position (1–12). Effects associated with the first three factors have already been described in the analyses of total errors and we focus here on additional effects associated with serial position.

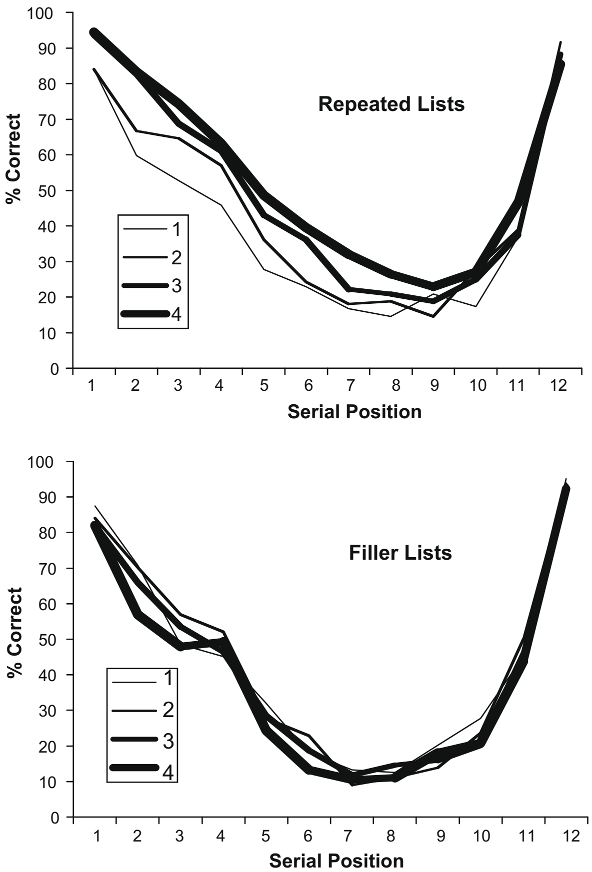

Fig. 3. Experiment 1: serial position curves for Repeated lists and Filler lists as a function of trial Position in Block (1,2,3,4).

The main effect of Serial Position was highly significant, $F(11, 187) = 135.6$, MSE = 2.47, $p < .001$, reflecting extended primacy and recency limited to the final few serial positions. The improvement in recall due to list repetition declined towards later positions (see Fig. 3). This picture was supported by a cluster of significant interactions: Serial Position × List Type, $F(11, 187) = 7.35$, MSE = .67, $p < .001$; Serial Position × Position in Block, $F(33, 561) = 1.49$, MSE = .51, $p < .05$; and Serial Position × Position in Block × List Type, $F(33, 561) = 1.76$, MSE = .50, $p < .01$. The interactions involving List Type and Serial Position were not simply due to a ceiling effect at the final serial position as they remained significant when data from this position were excluded. The analysis revealed one further significant interaction, namely Serial Position × Articulatory Suppression, $F(11, 187) = 3.29$, MSE = 1.58, $p < .001$, reflecting a tendency for suppression to disrupt recall of items from the middle of lists more than the ends (see Fig. 4). As can be seen, suppression interacted with position in a different way from list repetition.

Virtually all participants (16/18) reported noticing that all or part of some lists were repeated in the control condition but fewer (11/18) did so in the suppression condition. Where partial repetition was noticed it was typically attributed to the beginnings and ends of the lists. Inspection of the data indicated that recall performance was unrelated to level of awareness. There were no significant effects associated with awareness in a 4-way analysis of variance on correct responses under suppression (Awareness × Type of List × Position in Block × Secondary Task).

Discussion

Articulatory suppression interfered with immediate serial recall, as expected, but did not interfere with the learning of repeated sequences. Furthermore, suppression had its biggest effect on items from middle positions whereas list repetition had its biggest effect at early positions. If one accepts that suppression disrupts rehearsal (Murray, 1967), these observations go against the idea that rehearsal is necessary for the Hebb Effect (Cunningham et al., 1984). The lack of effect of suppression on learning is somewhat counterintuitive in that any factor reducing the amount

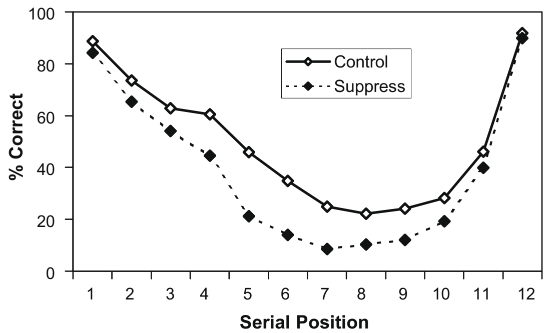

Fig. 4. Experiment 1: serial position curves for recall performance in the control and concurrent articulatory suppression conditions. Data pooled over Repeated and Filler lists.

of information recalled might be expected to reduce the amount of information gained. However, the Burgess and Hitch (1999) model successfully predicted the insensitivity of the Hebb Effect to articulatory suppression because it acts on phonological representations whereas sequence learning is mediated by context-item connections.

It is interesting to consider performance on filler lists in terms of the Burgess and Hitch (1999) model. Given that the context signal evolves with the timing of item presentation, all lists will involve strengthening connections to the same sequence of states of the context layer. The build-up of long-term strength in these position-item associations for a repeated list should interfere with immediate recall of non-repeated lists, as in the model's explanation of prior-list intrusions. There was little evidence for such an effect here (see Fig. 2). However, this observation should be treated with caution given that there is no control condition measuring performance on non-repeated lists in the absence of interpolated Hebb lists. The general question of interference between different lists is important and turned out to be the key reason for having to modify the Burgess and Hitch (1999) model in the light of the present experiments. We return to this issue towards the end of the paper.

That most participants reported becoming aware lists contained repeated information is not surprising given previous data on awareness in the Hebb paradigm (McKelvie, 1987; Sechler & Watkins, 1991). However, it is interesting that articulatory suppression tended to make repetition less noticeable whilst having no measurable effect on learning, consistent with the view that awareness and learning are not strongly associated (McKelvie, 1987) and the idea that the learning is implicit (see Seger, 1994). However, we note the limitations of retrospective reports of awareness and controversy surrounding the role of awareness in the Hebb Effect (see e.g., Sechler & Watkins, 1991).

Experiment 2

As described earlier, another prediction of the Burgess and Hitch (1999) model is that changing the timing of item presentation when a list is repeated should impair learning. Accordingly, we compared the effect of repeating a list using either the same or different patterns of temporal grouping. Both conditions involved repeating the same list of items, but in one the pauses between groups always occurred in the same places whereas in the other the locations of the pauses were different each time. According to the model, when the grouping pattern remains the same, context-timing nodes go through the same sequence of states during each presentation. However, when the grouping pattern varies, these nodes go through a different sequence of states each time the list is presented. It was expected that changing the pattern of grouping would reduce and not abolish learning because the rate of presenting items within groups always remained the same. Thus, according to the model, successive states of the context-timing signal for different grouping patterns would not be completely different from each other.

A previous study of the effect of temporal grouping in the Hebb procedure showed that changing the grouping pattern of a repeated list does have a large disruptive effect on learning (Bower & Winzenz, 1969; Experiment 3). However, this theoretically important result appears never to have been replicated independently. Accordingly, Experiment 2 set out to do this. Effects of articulatory suppression were also examined in order to confirm and extend the findings of Experiment 1. As before, suppression was predicted to disrupt immediate recall without affecting long-term learning because it interferes with phonological representations but not the context-timing signal. By the same token, suppression was not expected to alter the effect of changing the temporal grouping pattern of the repeated list.

Method

Participants

Forty eight participants (10 males and 38 females) were recruited from the student population at Lancaster University in return for a small payment.

Stimuli and apparatus

Twelve-item lists were generated in the same way as in Experiment 1. Five sets of five lists were constructed, each set being used to generate a block of eight trials with the form *aebecede*, where *a,b,c,d* were designated as non-repeated Filler lists and *e* was the Repeated list. A second set of master lists was obtained by reversing the within-list order of items in the first set. The second set was used to generate a further five blocks of eight trials with the same *aebecede* repetition pattern. The two sets of materials were assigned to either the 'same' or 'different' grouping conditions, and counterbalanced across participants.

Temporal grouping patterns were imposed to generate two versions of each set of materials. In the "Repeat Same" version, the four repeated lists in each block of eight trials were all given the same grouping pattern. In the "Repeat Different" version, the repeated lists in each trial-block were each given a different temporal grouping pattern. In both versions, the Filler lists were given different temporal grouping patterns from one another and from the Repeated lists.

Each list of 12 items contained four or five temporal groups of various sizes, the sets used being {4,4,3,1}, {4,4,2,2}, {4,4,2,1,1}, {4,3,3,2}, {4,3,3,1,1} and {4,3,2,2,1}. Permutations of the order of groups within each of these sets were used to generate a total of 130 different grouping patterns, i.e. a different pattern for every list apart from repeated lists in Repeat Same trial-blocks. The sequences of group-sizes used for repeated lists in the five Repeat Same trial-blocks were 3,2,1,4,2; 4,1,2,1,4; 1,3,4,1,3; 3,4,1,4 and 2,4,3,3. The sequences of group-sizes used for the four presentations of the repeated list in each Repeat Different block had their pauses in a different place on each presentation, so far as possible, as in the example, 1,4,2,3,2; 3,1,4,4; 2,4,3,2,1 and 4,2,4,2. Once the grouping patterns for the Repeated lists had been selected, the grouping patterns for the Filler lists were assigned haphazardly. Two practice lists with different grouping structures from any of the experimental lists were also created.

Lists were recorded on tape as in Experiment 1 but with a presentation rate of two items per second and a 1 s pause between groups.

Design and procedure

The experiment had a $2 \times 3 \times 4$ factorial design with one between-subjects factor, Secondary Task (Suppression, No Suppression), and two within-subjects factors, Type of List (Repeat Same, Repeat Different, Filler) and Position in Block (1,2,3,4). Position in Block refers to the position of successive pairs of Filler and Repeated lists within each block of eight trials, as explained in Experiment 1.

Instructions and procedure were as in Experiment 1 except that suppression was a between-subjects factor and participants were allowed as long as they needed to recall each list. Each participant received 10 blocks of eight lists. Half the blocks had Repeat Same grouping patterns, the other half Repeat Different. The blocks were presented in a different haphazard order for each participant.

Results

Recall was scored in terms of errors per list as in Experiment 1. Preliminary analysis showed that performance on Filler lists in Repeat Same and Repeat Different blocks did not differ, so these data were pooled. Table 2 summarises the results for Repeat Same, Repeat Different and Filler lists and Fig. 5 illustrates the trends across trials. The main findings are readily apparent: (i) lists were learned faster when they were repeated using the same grouping pattern; (ii) there was some learning when repeated lists had different grouping patterns and (iii) articulatory suppression disrupted immediate recall but did not appear to interact with learning, as in Experiment 1.

Data were entered into a mixed $2 \times 3 \times 4$ analysis of variance with factors Secondary Task (No Suppression, Suppression), Type of List (Repeat Same, Repeat Different, Filler) and Position in Block (1,2,3,4). There was a significant main effect of Secondary Task, $F(1, 46) = 21.6$, MSE = 24.5, $p < .001$, reflecting more errors with suppression ($M = 7.6$) than without ($M = 5.7$). There were no significant interactions involving Secondary Task (all $F_s < 1.5$). However, as the patterns of results with and without suppression were of interest, separate 3 (Type of List) $\times$ 4 (Position in Block) analyses of variance were carried out for each condition.

Table 2
Experiment 2: mean (*SD*) errors (max = 12) for lists repeated with the same or different grouping as a function of position in trial-block and articulatory suppression.

	Position in Block				
	1	2	3	4	Mean
No Suppression					
Repeat Same	6.4 (1.8)	5.4 (2.0)	5.2 (2.2)	4.3 (2.4)	5.3 (2.2)
Repeat Different	6.2 (1.3)	5.8 (1.6)	5.8 (1.8)	5.3 (1.9)	5.8 (1.7)
Filler	6.0 (1.4)	5.9 (1.6)	6.2 (1.4)	6.4 (1.4)	6.1 (1.5)
Suppression					
Repeat Same	8.0 (1.5)	7.6 (1.6)	7.2 (1.7)	6.8 (1.9)	7.4 (1.7)
Repeat Different	8.0 (1.7)	7.5 (1.7)	7.7 (1.4)	7.2 (1.7)	7.6 (1.6)
Filler	7.8 (1.6)	8.1 (1.2)	7.9 (1.5)	8.2 (1.2)	8.0 (1.4)

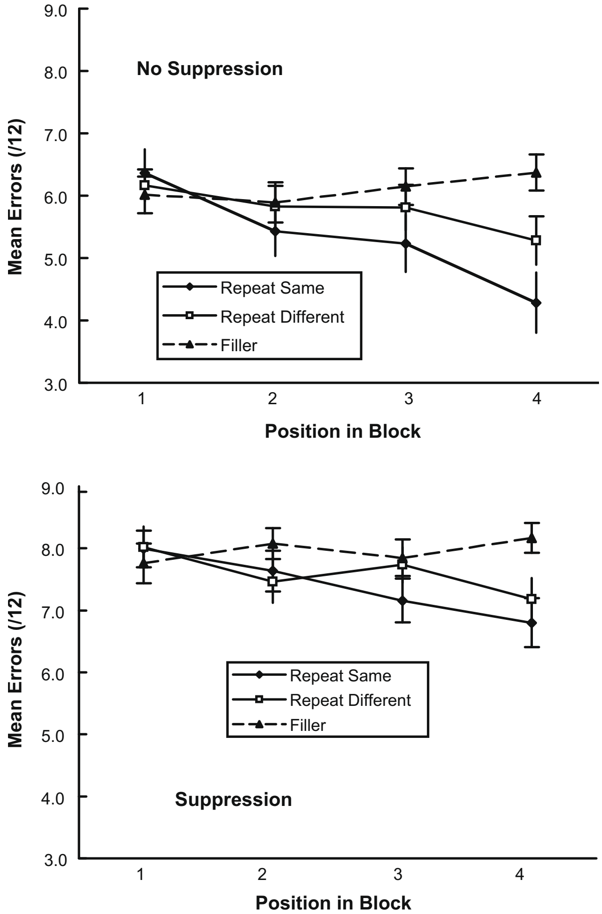

Fig. 5. Experiment 2: effect of repeating an auditory sequence of digits using either the same (Repeat Same) or a different grouping pattern (Repeat Different), with or without articulatory suppression. Non-repeated lists designated as Filler. See text for explanation of abscissa.

In the No Suppression condition, there were significant effects of Type of List, $F(2, 46) = 9.21$, MSE = 1.58, $p < .001$, Position in Block, $F(3, 69) = 7.48$, MSE = 1.22, $p < .001$, and the interaction, $F(6, 138) = 7.96$, MSE = .79, $p < .001$. Tukey tests indicated that error rates for the three types of list were significantly different from one other ($p < .05$ or better), with least errors for Repeat Same lists and most for Filler lists. Errors fell significantly over trials for Repeat Same and Repeat Different lists, $F(3, 69) = 12.2$, MSE = 1.44, $p < .001$ and $F(3, 69) = 2.90$, MSE = 1.10, $p < .05$, respectively, but rose significantly for Filler lists, $F(3, 69) = 3.77$, MSE = .27, $p < .05$. As regards the critical prediction that learning would be faster when repeated lists had the same grouping pattern, a separate 2 (Repeat Same vs. Repeat Different) $\times$ 4 (Position in Block) analysis revealed that the interaction was significant, $F(3, 69) = 2.91$, MSE = 1.02, $p < .05$.

The pattern of results for the Suppression condition was broadly similar, though overall performance was worse and differences tended to be smaller. Thus, there were significant effects of Type of List, $F(2, 46) = 11.0$, MSE = .71, $p < .001$, Position in Block, $F(3, 69) = 5.14$, MSE = .74, $p < .01$, and the interaction, $F(6, 138) = 7.85$, MSE = .43, $p < .001$. Tukey tests on mean errors summed over trials indicated that each type of repeated list was recalled significantly better than Filler lists ($p < .05$) but the two did not differ from one another on this measure. Analysis by simple effects showed that learning was significant for both Repeat Same Lists, $F(3, 69) = 10.1$, MSE = .65, $p < .001$, and Repeat Different lists, $F(3, 69) = 5.56$, MSE = .55, $p < .01$, whereas Filler lists showed a non-significant increase in errors with Position in Block, $F(3, 69) = 2.20$, MSE = .39, $p < .10$. A 2 (Repeat Same vs. Repeat Different) $\times$ 4 (Position in Block) analysis was used to test the prediction that changing the grouping pattern of the Repeated list would disrupt learning. This prediction was confirmed by a significant interaction, $F(3, 69)$ 3.07, MSE = .44, $p < .05$. Thus, articulatory suppression did not prevent Repeated lists from being learned, replicating Experiment 1, and did not alter the dependence of learning on whether lists were grouped the same way or differently on each presentation.

As our focus is on assessing the specific predictions of the network model generated in advance, and because little is added beyond what was reported in Experiment 1, we do not report data on errors and serial position effects here (or in Experiment 3).

Discussion

The results replicate and extend the previous observation that a repeated list is learned faster when its temporal grouping pattern remains the same (Bower & Winzenz, 1969). However, unlike the previous study, changing the grouping pattern of the repeated list did not abolish learning. It is not easy to explain this difference as the present study used broadly the same method, including the same list length and spacing between presentations of repeated lists. It is possible that the actual grouping patterns in the changed grouping condition were more similar in the present study but unfortunately, the original investigation does not give sufficient detail to assess this.

Overall, the results confirm the prediction of the Burgess and Hitch (1999) model that repeated presentation of a sequence leads to slower learning when its temporal grouping pattern is varied. This follows from the assumption that sequence learning involves strengthening associations to states of an internal context-timing signal that varies with the timing of item presentation. We note also that the model explains the residual Hebb Effect when the grouping pattern of the repeated list varied in terms of the similarity in the timing of items within groups.

Effects of articulatory suppression confirm and extend the observations made in Experiment 1. As before, suppression led to a substantial impairment in immediate recall but did not interfere with the learning of repeated lists. The new finding is that the effect on learning of changing the temporal grouping pattern of a repeated list was still detectable under suppression, despite low levels of performance. This is further evidence consistent with a general distinction between mechanisms for serial order and for phonemic information.

Experiment 3

So far, we have shown that learning a sequence of familiar items depends on the timing of item presentation but is unaffected by articulatory suppression. We interpret the absence of an effect of suppression as evidence that phonological coding is unimportant in this particular learning process. However, list presentation was auditory in Experiments 1 and 2 and, as already noted, articulatory suppression has different effects on immediate serial recall depending on the modality of list presentation. In particular, suppression removes the phonemic similarity effect for visual, but not auditory presentation (Baddeley et al., 1984). According to the phonological loop model, these effects are explained by assuming that visual stimuli have to be subvocalised in order to access the phonological store whereas auditory stimuli gain access to this store automatically. One could therefore interpret the persistence of the Hebb Effect under suppression in Experiments 1 and 2 as consistent with an effect that depends on the phonological store (or input phonology in the Burgess & Hitch, 1999 model). A stronger test of the hypothesis that the Hebb Effect does not involve changes at the level of phonological coding would be to show that it survives under articulatory suppression when items are presented *visually*. There is some evidence to support this prediction (Page, Cumming, Norris, Hitch, & McNeil, 2006, Experiment 1). Experiment 3 sought to test it further.

A second aim of Experiment 3 was to provide further evidence that participants rely on phonological short-term memory in the Hebb procedure. Baddeley (2000b) has suggested that participants may tend to abandon phonological coding for long lists and it is possible that the use of sequences exceeding span in the typical Hebb procedure encourages such a strategy. If so, the insensitivity of sequence learning to articulatory suppression in Experiments 1 and 2 could conceivably reflect participants' use of non-phonological coding strategies in this paradigm. Experiment 3 examined the effect of manipulating the phonemic similarity of the list items. If effects of both phonemic similarity and Hebb repetition are observed in the same experiment it would provide strong evidence that the sequence of items undergoing long-term learning is held in phonological short-term memory.

The manipulation of phonemic similarity allows two further observations. First, consistent with previous studies of immediate serial recall of visual sequences (e.g., Baddeley et al., 1984;Murray, 1968), articulatory suppression should remove the phonemic similarity effect. A demonstration that articulatory suppression interacts with phonemic similarity but not the Hebb Effect would be further evidence that the mechanisms for sequence learning and for maintaining phonological representations are distinct. Second, sequence learning can be examined as a function of phonemic similarity. Given that phonemic similarity typically results in increased order errors in immediate serial recall (Conrad, 1965), it would be natural to expect learning through repeated immediate serial recall to be particularly difficult for a list of similar items. However, as explained earlier, the Burgess and Hitch (1999) model explains the phonemic similarity effect in terms of cross-talk amongst item-phoneme connections that are independent of the context-item connections involved in learning serial order. Thus, the model makes the counterintuitive prediction that phonemic similarity should not impair sequence learning, despite having its normal effect of disrupting STM for serial order.

To summarise, Experiment 3 examined the effects of phonemic similarity and articulatory suppression in the Hebb procedure for visually presented lists. If participants are relying on phonological short-term memory, they should show the standard phonemic similarity effect in immediate recall, and this effect should disappear when phonological coding is prevented by articulatory suppression. However, the Burgess and Hitch (1999) model predicts no effect of either phonemic similarity or articulatory suppression on the learning of repeated sequences. These predictions follow directly from the assumption that learning a sequence of familiar items involves changes to context-item connections that are distinct from item-phoneme and phoneme-item connections. An alternative, intuitive hypothesis would hold that because phonemic similarity disrupts memory for the order of items in immediate recall, it should also disrupt learning the order of items in a repeated sequence.

Method

Participants

Twenty four participants (14 males and 10 females) were recruited from the student population at Lancaster University. They were given a small payment for taking part.

Stimuli and apparatus

Three types of 8-item list were constructed: Phonologically Similar lists were permutations of the letters BCDGPTVK (where the non-rhyming K was included to make up the set); Phonologically Dissimilar lists were permutations of the letters JFHQRYZL; Filler lists consisted of four items from each of the above sets. There were six Phonologically Similar and six Phonologically Dissimilar lists, constructed so that no letter occurred in the same position more than twice. There were 24 different Filler lists.

Altogether there were 72 experimental trials. These were arranged in six blocks of 12 trials each. Within each block, one Phonologically Similar list and one Phonologically Dissimilar list was each presented four times and there were four different Filler lists. These three List Types were rotated over successive groups of three trials (e.g., SDF, SDF, SDF, SDF where S = Similar, D = Dissimilar and F = Filler). Over the six blocks, each of the possible orderings of the three List Types was used once. An additional three Filler lists were constructed to give as practice immediately before the experimental trials. Six lists of eight digits from 1 to 9 were also created for task familiarisation. A Hypercard programme written by S. Slavin was used to present stimuli visually using the display screen of a Macintosh computer.

Design and procedure

The experiment had a $2 \times 3 \times 4$ factorial design. Secondary Task (No Suppression, Suppression) was varied between subjects, the other factors were varied within subjects. These were List Type (Phonologically Similar, Phonologically Dissimilar, Filler) and Position within Block (1,2,3,4). List position 1 refers to the first list of each type in a block, list position 2 the second, and so on.

Participants were tested individually. Letters were presented centre-screen, in upper case, at a rate of two items per second (each item was visible for 0.33 s, and there was a blank ISI of 0.17 s). An asterisk signalled the start of the list and a question mark appeared as a cue for recall after presentation of the final list item. Recall was written on prepared sheets as in the previous experiments. Six lists of digits were given for general practice. The experimental lists were preceded by three Filler lists that were not scored. Trials were self-paced by means of the space bar and natural breaks occurred when participants had to start a new response sheet after every 12 lists. Half of the participants were required to suppress articulation using the same procedure as in the previous experiments.

Results

Scoring was as before. Table 3 gives means and standard deviations for errors per list as a function of its position within each trial-block (i.e. 1,2,3,4 – see Method). Fig. 6 plots the same data and shows that, without articulatory suppression, phonemic similarity impaired recall but had no effect on the learning of Repeated lists. With suppression, phonemic similarity no longer impaired recall and learning of repeated lists remained evident.

The error data were entered into a $2 \times 3 \times 4$ analysis of variance with factors Secondary Task (No Suppression, Suppression), Type of List (Phonemically Similar, Phonemically Dissimilar, Filler) and Position in Block (1,2,3,4). There was a significant main effect of Secondary Task, $F(1, 22) = 14.7$, MSE = 8.15, $p < .01$, reflecting poorer recall

Table 3
Experiment 3: mean (*SD*) errors (max = 8) for lists of phonemically similar and dissimilar items as a function of position in each trial-block and articulatory suppression.

	Position in Block				
	1	2	3	4	Mean
No Suppression					
Phonologically Similar	4.4 (.9)	4.1 (1.5)	3.8 (1.6)	3.5 (1.7)	3.9 (1.4)
Phonologically Dissimilar	3.6 (.8)	3.5 (1.2)	2.8 (.8)	2.8 (1.2)	3.2 (1.1)
Filler	4.0 (.9)	4.2 (1.3)	4.2 (.9)	4.6 (1.2)	4.3 (1.1)
Suppression					
Phonologically Similar	5.5 (.8)	5.0 (.8)	4.6 (1.2)	4.1 (1.4)	4.8 (1.1)
Phonologically Dissimilar	5.2 (.9)	5.2 (1.2)	5.0 (1.2)	4.8 (1.1)	5.0 (1.1)
Filler	5.2 (.9)	5.3 (.6)	5.7 (.6)	5.4 (.7)	5.4 (.7)

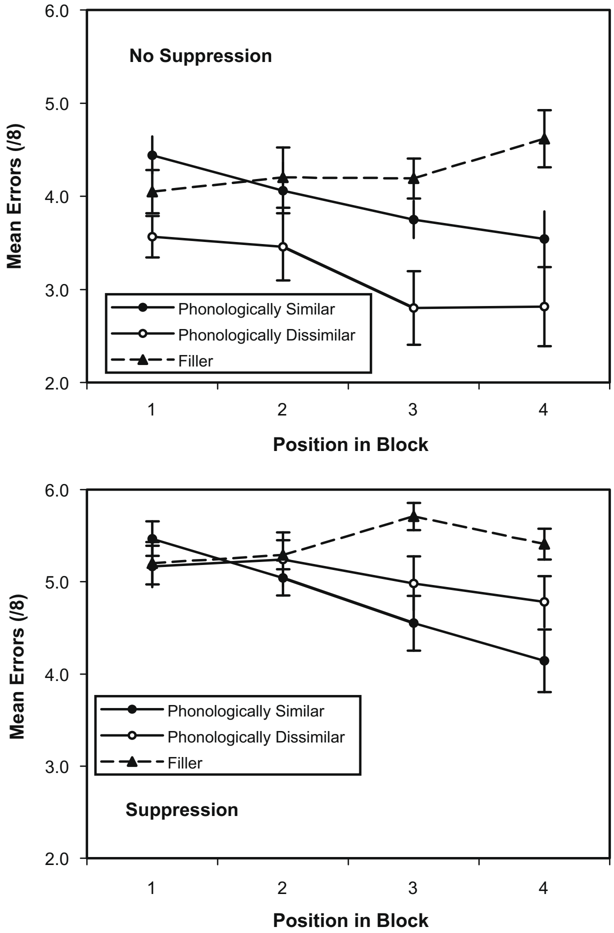

Fig. 6. Experiment 3: effect of repeating a visual sequence of phonemically similar or dissimilar letters, with or without articulatory suppression. Non-repeated lists designated as Filler. See text for explanation of abscissa.

with articulatory suppression ($M = 5.08$) than without ($M = 3.79$). The significant main effect of Type of List, $F(2, 44) = 9.91$, MSE = 1.32, $p < .01$, and the interaction between Secondary Task and Type of List, $F(2, 44) = 5.14$, MSE = 1.32, $p < .01$, reflect the disappearance of the phonemic similarity effect with suppression. There was also a significant main effect of Position in Block, $F(3, 66) = 6.76$, MSE = .41, $p < .01$, modified by an interaction with Type of List, $F(6, 132)$, = 8.00, MSE = .36, $p < .01$. Analysis of simple effects indicated that the drop in errors with repetitions was significant for both Phonologically Dissimilar lists, $F(3, 66) = 6.47$, MSE = .33, $p < .01$, and Phonologically Similar lists, $F(3, 66) = 14.1$, MSE = .40, $p < .01$. In contrast, there was a non-significant increase in errors in recalling Filler lists with their position in block, $F(3, 66) = 1.92$, MSE = .40. There were no other significant effects (Secondary Task × Position in Block, $F(3, 66) = 1.08$, MSE = .41; Type of List × Secondary Task × Position in Block, $F(6, 132) < 1$.

Further two-way analyses looked at performance on Repeated lists only. With no suppression, there were significant effects of Phonemic Similarity, $F(1, 11) = 7.44$, MSE = 2.00, $p < .01$, and Position in Block, $F(3, 33) = 8.16$, MSE = .45, $p < .01$, and no interaction, $F(3, 33) < 1$, MSE = .34. Thus phonemically similar lists were recalled worse but nevertheless learned at the same rate as dissimilar lists when participants were free to subvocalize. With suppression, there was a significant effect of Position in Block, $F(3, 33) = 10.8$, MSE = .33, $p < .01$, no effect of Phonemic Similarity, $F(1, 11) < 1$, MSE = 1.47, and an interaction of borderline significance, $F(3, 33) = 2.89$, MSE = .34, $p = .05$. The interaction reflects a tendency for learning to be slightly faster for phonemically similar lists when participants suppressed articulation.

Discussion

The main results of this study are straightforward. Articulatory suppression and phonemic similarity of list items had their standard effects on immediate serial recall. Thus, suppression and phonemic similarity each disrupted performance and suppression reduced the effect of phonemic similarity (Baddeley et al., 1984; Murray, 1968). However, the rate of learning a repeated list was unaffected by articulatory suppression, consistent with the results of Page et al. (2006). We noted earlier that the insensitivity of the Hebb Effect to articulatory suppression with auditory presentation in Experiments 1 and 2 left open the possibility that sequence learning involved changes in phonological storage. Confirmation that this result generalises to visually presented lists gives stronger support for the Burgess and Hitch (1999) model's proposal that changes involving phonological representations play no role in sequence learning for lists of familiar items. The insensitivity of Hebb repetition learning to phonemic similarity when participants were free to rehearse confirms a further prediction of the model derived from its assumption that repeating a sequence of familiar items strengthens context-item connections whilst having no effect phoneme-item connections. However, one unexpected result deserves comment. This is the tendency for lists of phonologically similar items to be learned slightly *faster* under articulatory suppression, resulting in the standard similarity effect tending to reverse over trials. This intriguing observation was not predicted by our model (see also Burgess & Hitch, 2006, Simulation 3). However, as this marginal interaction has not been replicated, any attempt at interpretation would be speculative.

It is interesting to note that the absence of any change in the phonemic similarity effect as learning proceeds seems inconsistent with the idea of separate short-term and long-term stores in the classic modal model of memory (e.g., Atkinson & Shiffrin, 1968). In this account learning involves transfer of information from the short-term to the long-term store. It seems reasonable to assume that the size of the phonemic similarity effect is an index of the contribution of the short-term store to recall, in which case it should become smaller as a list is learned.

A further aspect of the results requires comment. Because the experiment involved repeating two lists within each block of trials, the results imply that participants could learn two sequences simultaneously. This observation may conceivably reflect special circumstances in which the repeated lists were highly distinct from each other, and two out of every three trials contained a repeated list. However, in everyday life we clearly do have the ability to learn multiple sequences in parallel, suggesting that the phenomenon has some generality. This is theoretically significant because it highlights a difficulty for the Burgess and Hitch (1999) model according to which presentation of different lists sharing the same timing of item presentation generates the same sequence of activations in the context layer, leading to massive interference when two lists are repeated in parallel. That this does not occur strongly suggests the assumption of a single context layer is incorrect.

General discussion

We started by noting that the phonological loop theory of verbal STM (Baddeley, 1986) does not address how order information is maintained nor how STM interfaces with LTM. These are important general issues if one wishes to explain the role of STM in learning new word forms (Baddeley et al., 1998) and have been addressed in a network model of verbal STM capable of immediate serial recall and long-term sequence learning (Burgess & Hitch, 1999). The model makes the strong assumption that there are separate mechanisms for order and item information. Thus, order information is coded by connections between a context-timing signal and item representations whereas item information is coded by connections to phonological representations. These connections are modifiable, with both short-term and long-term plasticity. Rapid decay constrains immediate serial recall, whereas slow decay allows cumulative long-term learning when a previously-presented sequence is repeated. We showed how a general understanding of the model can be used to generate qualitative predictions about sequence learning. Specifically, we used the model to predict that long-term learning of a sequence of familiar items should be sensitive to factors affecting memory for order but not factors affecting memory for items. We tested these predictions experimentally using a procedure developed by Hebb (1961) that allows simultaneous assessment of short-term recall and long-term learning. The predictions were largely borne out in the form of an experimental dissociation between on the one hand effects of articulatory suppression and phonemic similarity (affecting mainly item information), and on the other effects of temporal grouping (affecting mainly order information). Thus, articulatory suppression and phonemic similarity had their well-established effects of disrupting immediate serial recall, but neither suppression nor phonemic similarity affected long-term sequence learning. In contrast, altering the timing of a repeated sequence in successive presentations impaired long-term learning of the sequence. The only exception to this neat pattern was the unpredicted tendency for long-term learning under articulatory suppression to be marginally faster when sequences contained phonemically similar items (see Experiment 3). On balance, we take confirmation of the general predictions of the network model as suggesting the way it deals with item and order information and the relationship between STM and LTM is broadly appropriate.

There was, however, one incidental observation that drew attention to an obvious limitation of the network model. This was that participants were capable of learning more than one sequence simultaneously (Experiment 3). If all sequences recruit a common context-timing signal, then repeating two sequences in parallel should generate a large amount of interference. People's ability to learn two different sequences at the same time, together with weak evidence for position-specific transfer effects in other studies of sequence learning (Cumming, Page, & Norris, 2003; Hitch, Fastame, & Flude, 2005) strongly suggests the need to revise our model. (Burgess & Hitch, 2006) describe a revision that introduces an added process whereby a sequence can be recognised as having been encountered previously. This is achieved by having multiple sets of context nodes. When a sequence is presented, the input is matched against all sets of context nodes in parallel. As successive items are presented a cumulative match is computed for each context-set and sets whose cumulative match falls below a threshold are discarded. In this way the cohort of active context-sets reduces until the sequence is either 'recognised' as similar to one that has been encountered before or perceived as 'new' and assigned a context-set with no significant long-term connections. This process of progressive matching is analogous to Marslen-Wilson's (1987) cohort model of auditory word recognition, but operating at the level of items within lists rather than the acoustic–phonetic components of individual words. Burgess and Hitch (2006) report simulations showing that the revised model is capable of reproducing the main findings of the present experiments.

To return to the present experimental findings, we note that they can be related to Baddeley's (2000a) proposal of a multi-modal episodic buffer in addition to the phonological loop, each store connected to a different region of long-term memory. The episodic buffer integrates information from different working memory stores (i.e. the visuo-spatial buffer store and the phonological loop) and links it with relevant information in long-term memory. One way of mapping the present data onto this account is to assume that the episodic buffer deals with serial order information whereas the phonological loop deals with (verbal) item information (see Burgess & Hitch, 2005). Interestingly, there are a number of empirical similarities between immediate serial recall for non-verbal and verbal stimuli (see e.g., Avons, 1998; Jones, Farrand, Stuart, & Morris, 1995; Smyth, Hay, Hitch, & Horton, 2005) and the Hebb repetition effect has been observed for non-verbal as well as verbal sequences (Gagnon, Foster, Turcotte, & Jongenelis, 2004). These similarities are consistent with the possibility of a common ordering mechanism linked to different subsystems. Alternatively, there may be specialised ordering systems for different storage systems and response streams. Jones et al. (1995) postulate a common ordering mechanism that depends on associative chaining. However, Henson, Norris, Page, and Baddeley (1996) discuss substantial difficulties with chaining as an account of order in immediate serial recall and it is difficult to see how this approach could explain the present data on the Hebb Effect.

The present results also have implications for alternative computational models of serial order and short-term memory. A number of models besides ours make a distinction between the process of selecting an item and retrieving its phonological features during recall (e.g., Brown et al., 2000; Henson, 1998; Page & Norris, 1998). As noted earlier, none of these models addresses long-term learning. However, they could in principle be elaborated by specifying learning mechanisms that accommodate the dissociation between phonological and timing variables in the Hebb Effect. Such changes could be readily incorporated into models in which item selection involves some form of context signal (Brown et al., 2000; Henson, 1998). However the primacy gradient in the Primacy Model (Page & Norris, 1998) would need substantial modification to account for effects of temporal grouping in short-term recall and long-term learning. The present results may also present a problem for the idea that item and order information are coded conjointly in a distributed network, as in the model of Botvinick and Plaut (2006). On the face of it, such a system of representation seems difficult to reconcile with the present evidence that, for familiar items at least, item and order information behave separately.

Participants' reports of awareness suggest an area for further study. Whilst we acknowledge that strong conclusions cannot be drawn from our data, the apparent absence of any role of awareness is consistent with Seger's (1994) classification of the learning seen in Hebb repetition as implicit. We noted earlier that our computational model contains no higher-level mechanism for awareness or any means of generating task-specific strategies. Thus, whilst the model can predict the effect of using a particular strategy (such as a given pattern of rehearsal), it cannot generate such strategies without further assumptions. This is a general limitation of current computational models. However, the apparent unimportance of awareness and rehearsal strategies in the Hebb Effect is encouraging for attempts to explain it in terms of simple underlying mechanisms.

Although our modelling approach receives general support from the present data, it remains deliberately simplistic. Thus there are many aspects of serial order and the interaction between verbal short-term and long-term memory that the model does not address. For example, patterns of errors in immediate serial recall suggest that context varies relative to the ends of sequences as well as the beginnings (Henson, 1998; Ng & Maybery, 2002). This shortcoming can be addressed without making major changes to the ordering mechanism (see Henson & Burgess, 1997). A more substantial limitation is that the model deals with serial order at a lexical level, ignoring order at lower levels of representation such as phonology (and higher levels such as syntax). We note however that serial ordering can be successfully modelled at the phonological level by a similar mechansim (Hartley & Houghton, 1996) and that Gupta and MacWhinney (1997) have presented a related model that combines serial ordering at the lexical and phonological levels. Modelling serial order at the phonological level is particularly important if one wishes to simulate the learning of new word forms, as opposed to sequences of familiar lexical items, as here. Any such attempt should be informed by empirical differences between sequence learning at the phonological and lexical levels. For example, we have shown here that the rate of learning sequences of familiar items (digits or letters) is not affected by phonemic similarity or articulatory suppression. In contrast, the rate of learning new word forms is substantially impaired by both phonemic similarity and articulatory suppression (Papagno, Valentine, & Baddeley, 1991; Papagno & Vallar, 1992), as one might expect given that the ordering in this case is of phonological representations.

Part of the success of the original theory of the phonological loop was almost certainly due to it offering a simple, understandable account of a robust set of human data that could be easily applied to a range of problems. By specifying mechanisms more precisely, computational models such as ours attempt to build on the original theoretical account. However, computational models are inevitably more complex than conceptual models, and it is important to ensure that they too can be translated into relatively simple understandings that can be applied by non-experts. By showing how a broad understanding of the principles underlying our own particular model can be used to generate qualitative, sometimes counterintuitive predictions, we hope to have contributed to making it more transparent and applicable.

To conclude, we have tested empirical predictions regarding the learning of verbal sequences derived from a qualitative analysis of a network model of memory that incorporates (a) short-term and long-term plasticity and (b) separate mechanisms for item and order information (Burgess & Hitch, 1999). One prediction was that phonemic similarity and articulatory suppression would not affect the learning of repeated sequences because they do not act on order information, even though both variables disrupt immediate serial recall Another was that sequence learning would be sensitive to the consistency of temporal grouping, a factor assumed to affect order information. These predictions were confirmed in a series of three experiments. At the same time an obvious limitation of the model became evident, namely its inability to learn more than one sequence at once. A modest but important revision has addressed this shortcoming by introducing a recognition process whereby an incoming sequence is compared with long-term memory for previously learned sequences (Burgess & Hitch, 2006). Three further areas seem ripe for exploration. One is the learning of novel word forms, where the critical ordering process is sublexical and involves sequential constraints. Another is immediate recall of natural lexical sequences such as sentences, where chunking plays an important role. A third concerns interactions between multiple sequences. These would go beyond the level of explanation so far achieved, and might shed light on the relative merits of localist (e.g., Burgess & Hitch, 2006) and distributed (e.g., Botvinick & Plaut, 2006) approaches to serial order.

Acknowledgments

We acknowledge useful discussions with many colleagues including Alan Baddeley, Dennis Norris, Mike Page, Tom Hartley, and Garry Pilmore, and referees who commented on an earlier version of this paper that included work that is was subsequently reported in a companion paper (Burgess & Hitch, 2006). We are also grateful to Simon Slavin for programming Experiment 3. Neil Burgess was supported by a Medical Research Council Senior Fellowship.

References

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

Avons, S. E. (1998). Serial report and item recognition of novel visual patterns. *British Journal of Psychology, 89*, 285–308.

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. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D. (2000a). The episodic buffer: A new component of working memory. *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A. D. (2000b). The phonological loop and the irrelevant speech effect: Some comments on Neath. *Psychonomic Bulletin & Review, 7*(3), 544–549.

Baddeley, A. D. (2007). *Working memory: Thought and action*. Oxford: Oxford University Press.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158–173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.). *Recent advances in learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

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.

Botvinick, M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review, 113*, 201–233.

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digit series. *Journal of Experimental Psychology Monograph, 80*(2, Pt. 2), 1–17.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107*, 127–181.

Burgess, N. (1995). A solvable connectionist model of immediate recall of ordered lists. In G. Tesauro, D. Touretzky, & T. K. Leen (Eds.), *Advances in neural information processing systems* (7th ed., pp. 51–58). Cambridge: MIT Press.

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. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551–581.

Burgess, N., & Hitch, G. J. (2005). Computational models of working memory: Putting long-term memory into context. *Trends in Cognitive Sciences, 9*, 535–541.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*, 627–652.

Caplan, D., Rochon, E., & Waters, G. S. (1992). Articulatory and phonological determinants of word-length effects in span tasks. *Quarterly Journal of Experimental Psychology, 45A*, 177–192.

Caplan, D., & Waters, G. S. (1994). Articulatory length and phonological similarity in span tasks: A reply to Baddeley and Andrade. *Quarterly Journal of Experimental Psychology, 47A*, 1055–1062.

Cohen, R. L., & Johansson, B. S. (1967). The activity trace in short-term memory: A re-evaluation. *Journal of Verbal Learning and Verbal Behavior, 6*, 139–143.

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

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

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

Cumming, N., Page, M. P. A., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory, 11*, 43–63.

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.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9*, 59–79.

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

Gagnon, S., Foster, J. K., Turcotte, J., & Jongenelis, S. (2004). Involvement of the hippocampus in implicit learning of supra-span sequences: The case of SJ. *Cognitive Neuropsychology, 21*, 867–882.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition, 23*, 83–94.

Gathercole, S. E., Pickering, S., Hall, M., & Peaker, S. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology, 54A*, 1–30.

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

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

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

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). New York: Oxford University Press.

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

Henson, R. N. A. (1999). Positional information in short-term memory: Relative or absolute? *Memory & Cognition, 27*, 915–927.

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., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology, 49A*(1), 80–115.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology, 49A*, 140–158.

Hitch, G. J., Fastame, C., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory, 3–4*, 247–258.

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.

Hulme, C., Maughan, S., & Brown, G. G. 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.

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

Jones, D. M., Hughes, R. W., & Macken, W. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language, 54*, 265–281.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store. *Journal of Experimental Psychology: Learning, Memory and Cognition, 30*(3), 656–674.

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.

Lewandowsky, S., Duncan, S., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review, 11*, 771–790.

Lovatt, P., Avons, S. E., & Masterson, J. (2000). The word length effect and disyllabic words. *Quarterly Journal of Experimental Psychology, 53A*(1), 1–22.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102.

McKelvie, J. S. (1987). Learning and awareness in the Hebb digits task. *Journal of General Psychology, 114*, 75–88.

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.

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*(6), 1353–1380.

Murdoch, B. B. Jr., (1960). The distinctiveness of stimuli. *Psychological Review, 67*, 1631.

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

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

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

Neath, I. (2000). Modelling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review, 7*, 403–423.

Ng, H. L. H., & Maybery, M. T. (2002). Grouping in short-term verbal memory: Is position coded temporally. *Quarterly Journal of Experimental Psychology, 55A*, 391–424.

Page, M. P. A., Cumming, N., Norris, D., Hitch, G. J., & McNeil, A. M. (2006). Repetition learning in the immediate serial recall of visual and auditory materials. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*, 716–733.

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

Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign language vocabulary learning. *Journal of Memory and Language, 30*, 331–347.

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology, 44A*, 47–67.

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

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

Sechler, E. S., & Watkins, M. J. (1991). Learning to reproduce a list and memory for the learning. *American Journal of Psychology, 104*, 367–394.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin, 115*, 163–196.

Service, E. (1998). The effect of word length on immediate serial recall depends on phonological complexity, not articulatory duration. *Quarterly Journal of Experimental Psychology, 51A*, 283–304.

Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visuo-spatial domain: Reconstructing sequences of unfamiliar faces. *Quarterly Journal of Experimental Psychology, 58A*, 909–930.