Available online at www.sciencedirect.com
ScienceDirect
Journal of Memory and Language 55 (2006) 587–600
Journal of Memory and Language
www.elsevier.com/locate/jml

Mixed-list phonological similarity effects in delayed serial recall ☆

Simon Farrell

Department of Experimental Psychology, University of Bristol, 12a Priory Road, Clifton, Bristol BS8 1TU, UK

Received 28 February 2006; revision received 26 May 2006
Available online 9 August 2006

Abstract

Recent experiments have shown that placing dissimilar items on lists of phonologically similar items enhances accuracy of ordered recall of the dissimilar items [Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 838–849.]. Two explanations have been offered for this effect: an encoding explanation, in which items similar to current memory contents are given less encoding weight and offer less competition for recall; and a retrieval explanation, which suggests that the long-term similarity structure of the items leads to dissimilar items being more distinct on mixed lists. These theories are compared in an experiment in which a filled delay was introduced between study and test. Simulations show the prominent enhancing effects of similarity after a delay are captured by a model that assumes encoding is sensitive to the similarity of items to other list items [Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9*, 59–79.], but are not handled by a retrieval model [the Start–End Model; Henson, R. N. A. (1998). Short-term memory for serial order: the Start–End Model. *Cognitive Psychology, 36*, 73–137.].

Keywords: Serial recall; Memory; Computational modelling; Start–End Model; Serial-Order-in-a-Box; Phonological similarity

Introduction

The finding of a detrimental effect of phonological similarity on short-term serial recall is a classic finding in the short-term memory literature (e.g., Baddeley, 1966, 1968; Conrad, 1964; Wickelgren, 1965a, 1965b), and has been pivotal in the development of theories of short-term memory (Baddeley, 1986; Burgess & Hitch, 1999). Recently, emphasis has also been given to phonological similarity due to its apparent lack of effect on the recall of phonologically distinct items on the same list. That is, mixing together phonologically similar (e.g., *B*, *P*, *V*) and dissimilar (e.g., *H*, *Q*, *R*) items on the same list leads to the standard phonological similarity effect (similar items are confused more than dissimilar items), but does not influence recall of the dissimilar items compared to a control list containing only dissimilar items (Baddeley, 1968; Henson, Norris, Page, & Baddeley, 1996). This "dissimilar immunity" finding has been taken as a benchmark in the development of recent serial recall models,

☆ Preliminary data and simulation results arising from this research were presented at the 2005 meeting of the European Society of Cognitive Psychology.

E-mail address: Simon.Farrell@bristol.ac.uk
URL: http://seis.bris.ac.uk/~pssaf.

0749-596X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2006.06.002

motivating a theoretical framework in which the ordering of items in a serial recall task is assumed to be stored in a mechanism separate and upstream from that responsible for phonological confusions (Burgess & Hitch, 1999; Henson, 1998; Henson et al., 1996; Page & Henson, 2001; Page & Norris, 1998).

However, more recent experiments have called the dissimilar immunity effect into question. Farrell and Lewandowsky (2003) investigated serial recall for phonologically mixed lists in more detail, examining the effects of phonological similarity on different types of errors. Farrell and Lewandowsky found that mixing phonologically similar and dissimilar items had heterogeneous effects on recall errors for dissimilar items: mixing resulted in enhanced ordered recall for dissimilar items (reflected in a reduction in transposition errors) and additionally caused a decrement in item memory for the same items (reflected in an increase in intrusion and omission errors). Farrell and Lewandowsky (2003) suggested that the increase in item errors on mixed lists, which generalised to the similar items on the same lists, was due to an increase in set size at recall, since the number of possible items that could have appeared on those lists was essentially double that of "pure" lists containing only one class of items (only dissimilar, or only similar). When limiting item errors using a forward serial reconstruction task (Experiment 2), or controlling for this confound by using different dissimilar items on mixed and pure lists (Experiment 3), Farrell and Lewandowsky found a consistent recall advantage for dissimilar items on mixed lists.

This *mixed-list advantage* is important theory-wise as it distinguishes between two classes of models accounting for serial recall performance. In the one class of models already identified, phonological similarity effects are assumed to occur in a "phonological confusion" stage separate to, and downstream from, a primary stage responsible for ordering items. Thus, phonological similarity is assumed not to affect the representation of order of items, and any phonological similarity effects that occur are solely related to the discriminability of items in this second stage (Burgess & Hitch, 1999; Henson, 1998; Henson et al., 1996; Page & Norris, 1998). These models are here referred to as *retrieval-based* models, to indicate that phonological similarity only has effects at retrieval in these models. Notably, under standard assumptions (see below) these models do not predict a mixed-list advantage, since dissimilar items are assumed to be as similar to each other as they are to similar items, meaning the list context makes no difference to the recall of dissimilar items. In contrast, a second class of models exists in which similarity can have effects on order encoding. In such models, it is assumed that, additional to the standard detrimental effects of similarity at retrieval, the encoding of items is sensitive to their similarity to other items on the list. For example, a principle incorporated in the Serial-Recall-in-a-Box (SOB) model (Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2006) is that of *similarity-sensitive encoding*, in which the similarity of items to current memory contents is used to determine their encoding. One consequence is that these items will offer less competition for retrieval, meaning that the competition for recall of a dissimilar item amongst a list of similar items will be less than that for recall of the same item amongst a list of other similar items, due to attenuated encoding of the similar items in the first case (Farrell, 2001; Farrell & Lewandowsky, 2003). Such an encoding-based explanation thus naturally leads to a prediction of a mixed-list advantage.

This paper provides further evidence against retrieval-based models by examining the effects of delaying recall on the mixed-list effect empirically and in encoding and retrieval-based models. A possible explanation for the mixed-list advantage in retrieval-based models is considered, and an experiment designed to distinguish this explanation from encoding-based models is presented. Finally, simulations are presented that further call into question the validity of dual-stage retrieval-based models.

Mixed-list effects in retrieval-based models

In modelling phonological similarity effects in serial recall, modellers have generally made simplifying assumptions about the representation of similarity. One important assumption has been that, while similar items bear some relation to one another, dissimilar items are assumed to be completely dissimilar to each other and to the similar items. For example, in his Start–End Model (SEM) Henson (1998) assumed that the phonological match between a similar item and any other similar item was one, while the match between a dissimilar item and any other item was zero (items were assumed to have a match of one with themselves). An important assumption implicit in this is that the mean pair-wise similarity between any two dissimilar items is equal to the mean pair-wise similarity between any dissimilar item and any similar item. Under these assumptions, retrieval-based models have been shown to predict a dissimilar immunity effect, where recall of dissimilar items is unaffected by the similarity of surrounding items (Henson, 1998; Page & Norris, 1998).

However, analysis of the actual confusability of items reveals that this assumption is an oversimplification. Fig. 1 shows a 2-dimensional projection of a 3-dimensional multidimensional scaling (MDS) solution for phonological confusion data gathered by Hull (1973), details of which are presented later. At first glance, this perceptual space is consistent with basic assumptions in modelling of phonological similarity: rhyming items (*B*, *D*, *G*, *P*, *T*, *V*) are more tightly clustered than dissimilar

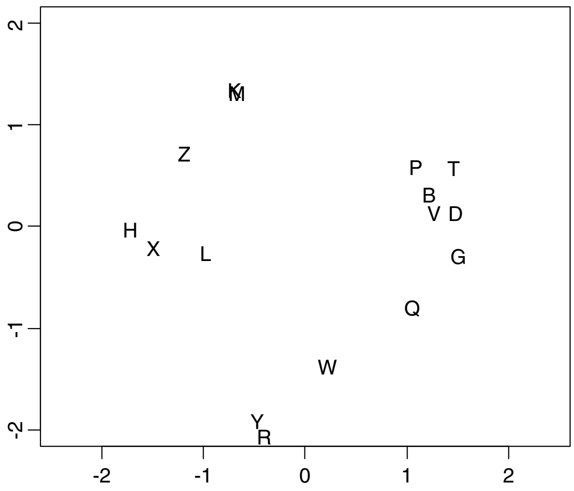

Fig. 1. A 2-dimensional projection of a 3-dimensional multidimensional scaling solution for acoustic confusions (data from Hull, 1973). Only letters from the experiments of Farrell and Lewandowsky (2003) are presented.

items. However, closer inspection shows that, although dissimilar items are more spread out than similar items, they also cluster separately from the dissimilar items. That is, Fig. 1 suggests that a dissimilar item will, on average, be more phonologically distinct on a list of similar items than on a list of dissimilar items. This is confirmed by examining average pair-wise Euclidian distances in the MDS space: the average pair-wise distance between any two dissimilar items (2.20) is smaller than that between any dissimilar item and any similar item (2.73).

Critically, incorporating this knowledge into retrieval-based models allows these models to capture the enhancing effects of mixed lists on dissimilar item recall, since dissimilar items are effectively made more distinct by placing them on mixed lists. Lewandowsky and Farrell (2006) modelled the mixed-list advantage in several models of serial recall. Critically, they found that the primacy model (Page & Norris, 1998), a retrieval-based model assumed to predict a null mixed-list effect, produced a small mixed-list advantage when phonological confusions were implemented according to an obtained MDS solution. Thus, retrieval-based models can in principle account for the mixed-list advantage when incorporating the phonological similarity structure of items generally used in serial recall experiments.

As noted in the preceding discussion of Fig. 1, the representational space that gives rise to mixed-list benefits in retrieval-based models is also responsible for the classic detrimental effects of phonological similarity on the recall of similar items. That is, the phonological space in which the distance between similar and dissimilar items drives the mixed-list advantage is the same space in which the average distance within the class of similar items, compared to that within the class of dissimilar items, drives the phonological similarity effect. This leads to a specific prediction from retrieval-based models: the mixed-list advantage should scale with the standard phonological similarity effect. Since models such as SEM (Henson, 1998) and the primacy model (Page & Norris, 1998) hold that the only effect of similarity is on confusions in a second stage, elimination of the phonological similarity effect will require a global change to the distinctiveness of all items (since these models have no mechanism for selective manipulation of classes of items).[1] This means that reducing or eliminating the standard phonological similarity effect should lead to a proportional reduction in the mixed-list advantage.

One factor that is well known to reduce the phonological similarity effect is the introduction of a filled delay between list presentation and recall (see, e.g., Page & Henson, 2001). Several studies indicate that delaying recall eliminates (Baddeley, 1968, Experiment 1; Page & Henson, 2001) or even reverses (Nairne & Kelley, 1999) the phonological similarity effect. Page and Henson (2001) analysed the results of Norris, Baddeley, and Page (2004) and found that the phonological similarity effect disappeared after 3 s of interpolated activity prior to recall. They argued that the rapid disappearance of the phonological similarity effect with delay indicated a rapid decay of phonological information as is assumed in SEM (Henson, 1998) and the primacy model (Page & Norris, 1998), prototypical retrieval-based models. Fig. 2 shows the effects of this phonological decay on the retrieval distinctiveness of items in SEM, obtained by simulating delay in SEM, and performing MDS on the resulting confusions predicted by the model. The change in the similarity structure from the left to right panel is solely due to decreasing phonological activations in the face of a constant amount of noise in the phonological confusion stage. As can be seen, a significant delay greatly reduces the extent to which the similar items are more tightly clustered (the standard phonological similarity effect), and also reduces the extent to which dissimilar items tend to cluster separately from

[1] In reviewing this paper, Gordon Brown pointed out that this prediction does not necessarily follow from models such as SIMPLE (Brown, Neath, & Chater, 2002) in which weights might be selectively applied to individual (phonological) dimensions. In SIMPLE it might be that manipulations such as delay have a greater effect on the dimension scaling within-class similarity (the vertical dimension in Fig. 1) than on the dimension representing between-class similarity (the horizontal dimension in Fig. 1). Although such a possibility is not discounted by the present data, such a change may render the model overly flexible (see, e.g., Navarro, Pitt, & Myung, 2004, for a discussion of flexibility in computational models).

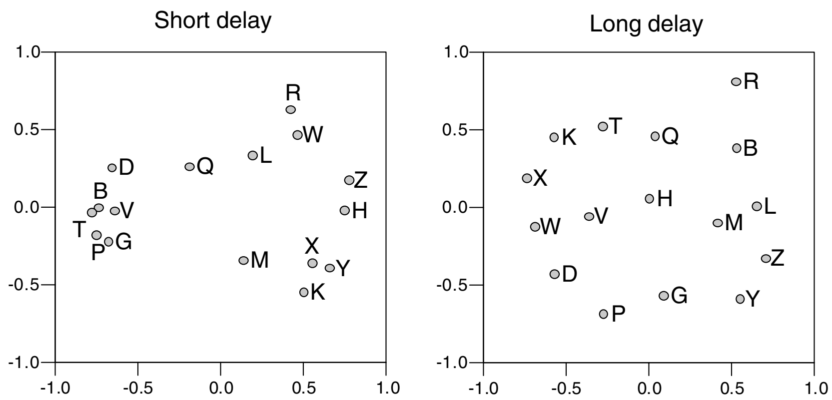

Fig. 2. Effects of delay on the phonological similarity structure in SEM (Henson, 1998). Both panels depict 2-dimensional MDS solutions for phonological confusions predicted by SEM. Predictions were obtained by running only the phonological confusion stage in SEM (by assuming the correct item was always passed from previous stages, and items were not suppressed), assuming equal initial activation for all phonological representations, and with decay applied to all items. The left panel shows the MDS solution for a short delay, while the right panel shows the MDS solution for an extended delay.

similar items. This stresses the extent to which the different effects of phonological similarity are tied together, and leads to the prediction that delaying recall will cause a scaled decrease in both the classical phonological similarity effect and in the mixed-list advantage.

In contrast, the SOB model, incorporating the mechanism of similarity-sensitive encoding, predicts that the mixed-list advantage should not scale with the standard phonological similarity effect, since similarity will have heterogeneous effects on ordered recall. Specifically, recall will be sensitive to similarity in two ways: (1) the similarity structure of the items at retrieval, producing a large detrimental effect of phonological similarity for similar items, and a small mixed-list advantage as per Fig. 1, and as in SEM; (2) an added advantage for dissimilar items on mixed lists due to decreased competition offered by similar items due to their reduced encoding strength. This means that even if the actions of delay are sufficient to reduce or eliminate the phonological similarity effect by ‘‘squashing’’ the similarity structure that causes similarity-based confusions in the model, a mixed-list advantage will still be evident after a delay given the added effects of similarity-sensitive encoding.

In the experiment now reported, the predictions of SEM[2] and SOB were tested by introducing a delay between presentation and recall.

[2] Henson’s (1998) SEM is taken as representative of dual-stage models. Page and Norris (1998) very clearly emphasize that their primacy model is intended as a model of immediate serial recall, and does not generalise to recall after ‘‘significant delays’’ (p. 762) such as those employed here.

Experiment

The purpose of the experiment was to test for the presence of a mixed-list advantage after a delay. A recall delay of 8 s was chosen for the delayed condition here as this has been previously found to greatly reduce or eliminate the phonological similarity effect for various materials (Baddeley, 1968; Bjork & Healy, 1974; Nairne & Kelley, 1999; Page & Henson, 2001).

Methods

Participants and design

Twenty-two undergraduates and employees of the University of Bristol voluntarily participated in the study. Undergraduates gained course credit, while others were reimbursed £7 for their time.

The experimental variables of delay, list type, and serial position were manipulated within participants. There were five different types of six-item lists: lists containing only similar items (*SSSSSS*), lists containing only dissimilar items (*DDDDD*), and three mixed lists containing a single dissimilar item (an *isolate*, following the terminology of Farrell & Lewandowsky, 2003) in position 2, 4 or 6 (*SDSSSS*, *SSSDSS*, and *SSSSSD*, respectively; *S* indicates a similar item; *D* indicates a dissimilar item). At each level of delay (2 s or 8 s), 10 lists of each list type were constructed, giving 100 lists in total for each participant.

Materials

Lists of six letters were constructed from two pools of consonants: a dissimilar pool, consisting of the letters *H*, *K*, *M*, *Q*, *R*, and *Y*; and a similar pool, consisting of the phonologically similar letters *D*, *P*, *T*, *B*, *V*, and *G*. Each list was constructed by randomly sampling from the two pools according to the list structure for each type of list. The sampling was constrained such that a letter could only appear once on each list, and common familiar sequences (e.g., 'TV') were disallowed. Lists were presented in a random order to participants (in contrast to the blocked presentation of Farrell & Lewandowsky, 2003, & Henson et al., 1996), to minimize list predictability (see also Lewandowsky & Farrell, 2006).

A random sequence of the digits 0–9 was also constructed for each trial. Digits were randomly sampled with replacement (with the constraint that an item should not be immediately repeated) to make sequences of four (delay = 2 s) or 16 (delay = 8 s) digits.

Procedure

The procedure was similar to that of Experiment 2 of Farrell and Lewandowsky (2003). Participants were informed that they would be presented with lists of six letters not containing any repeated items, which would be followed by sequences of digits. Participants were told that they should read aloud all the letters and numbers that appeared, and that they would be required to recall the letters from each sequence in order.

Each trial began with the presentation of a fixation cross in the middle of the screen for 1000 ms. After removal of the cross, a 500 ms was followed by presentation of the letters from the list, one by one, in upper case in the middle of the screen, each letter replacing the preceding one. Letters appeared for 450 ms, with a 50 ms inter-stimulus interval (ISI). After the final letter, there was a 500 ms, after which a sequence of four (delay = 2 s) or 16 (delay = 8 s) digits appeared one by one on the screen at a rate of 2/s, each digit was presented for 450 ms and followed by 50 ms delay. This digit shadowing task served to prevent rehearsal (cf, e.g., Nairne & Kelley, 1999; Page & Henson, 2001). The experimenter remained in the room for the duration of the experiment to ensure that participants complied with instructions. A delay period was included for the short delay condition to equalise the two conditions as much as possible except for the length of delay.

Following presentation of the list, a recall screen was presented using which participants were to reconstruct the order of the list items. This consisted of a box in the middle of the screen, above which the letters from the list were arrayed in a random order. Participants typed out the letters in their remembered sequence. As each letter was typed it briefly appeared in the box (300 ms), following which its representation above the box was dimmed to indicate it had been recalled. Thus, participants performed a serial reconstruction task, in that the identity of items was provided at recall. Nevertheless, participants were able to recall letters not appearing on the list, and could also repeat items, to allow generalisation to other experiments on mixed-list phonological similarity effects (Baddeley, 1968; Henson et al., 1996).

Results

Serial position curves for all levels of delay and list type are presented in Fig. 3. An ANOVA revealed all main and interaction effects to be significant, except for the three-way interaction (delay × list type × serial position). The results of this analysis are not reported as they do not directly address the mixed-list advantage.

To focus directly on mixed-list effects, Fig. 4 shows recall performance on similar, dissimilar, and isolate items at serial positions 2, 4, and 6 (the serial positions corresponding to the possible positions of isolates on mixed lists). Specifically, the serial position curves for the pure dissimilar and similar lists show performance on those lists only at positions 2, 4, and 6. The curve for the isolate lists shows performance on isolate items on mixed lists; this is a composite function where the first point depicts mean performance on the isolates appearing at serial position 2 on *SDSSSS* lists, the second point shows performance on the isolates appearing at serial position 4 on *SSSDSS* lists, and the third point shows mean performance on the isolates on *SSSSSD* lists. Fig. 4 suggests that, in general, isolates were recalled more accurately than their controls on pure dissimilar lists at both delay durations. To confirm this, a two (delay: short or long) × 3 (item type: similar, dissimilar, or isolate) × 3 (serial position: 2, 4, or 6) repeated measures ANOVA was conducted.[3] The ANOVA revealed significant main effects of delay [$F(1, 21) = 31.95$, $MSE = 0.058$, $p < .001$], item type [$F(1.84, 38.65) = 46.23$, $MSE = 0.041$, $p < .001$], and serial position [$F(1.54, 32.41) = 24.98$, $MSE = 0.13$, $p < .001$]. The interaction effects delay × item type [$F(2.00, 41.96) = 6.45$, $MSE = 0.021$, $p < .01$] and item type × serial position [$F(2.80, 58.80) = 4.01$, $MSE = 0.038$, $p < .05$] were significant; the delay × serial position interaction was marginally significant [$F(1.89, 39.74) = 3.16$, $MSE = 0.031$, $p = .056$]. The delay × item type × serial position interaction was not significant [$F(3.24, 68.01) = 1.53$, $MSE = 0.026$, $p = .21$]. Critically, post-hoc testing revealed that isolates were recalled more accurately than control items on dissimilar lists across delay [$t(21) = 2.47$, $p < .05$].

Further tests were run to examine the effects of delay on the phonological similarity and isolation effects. For each person, two scores were calculated at each level of delay: a mixed-vs.-pure score (*MVP*), being the mean difference in recall accuracy between isolate items and

[3] Mauchy's test of sphericity revealed a violation of the assumption of sphericity for some effects; where appropriate, effects are reported with the Greenhouse–Geisser correction.

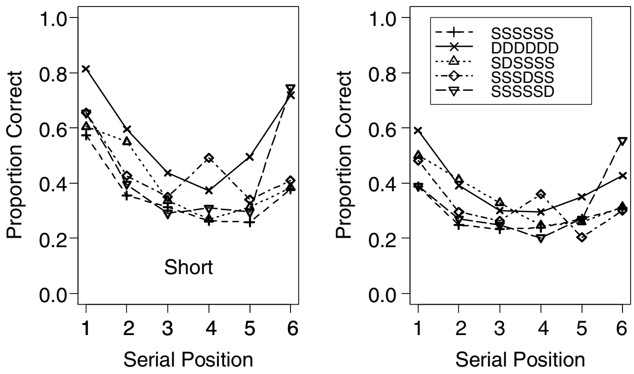

Fig. 3. Serial position functions for mean proportion correct. The left panel shows performance after a short delay (2 s); the right panel shows performance after a long delay (8 s).

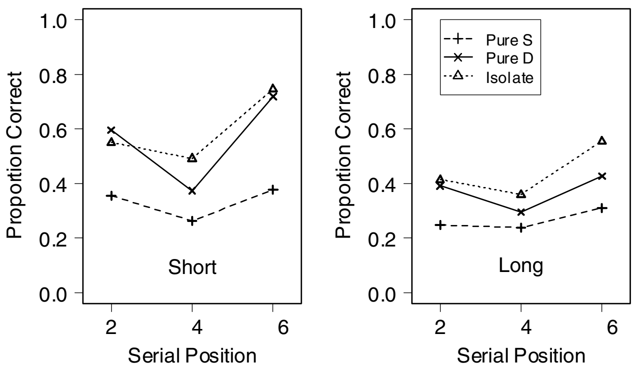

Fig. 4. Condensed serial position functions for mean proportion correct. The data presented are selected from Fig. 3, and show performance on isolates, dissimilar items on pure dissimilar lists, and similar items on pure similar lists, at serial positions 2, 4, and 6. The left panel shows performance after a short delay (2 s); the right panel shows performance after a long delay (8 s).

control items on dissimilar lists; and a phonological similarity effect (*PSE*) score, being the mean difference in recall accuracy between control dissimilar items and control similar items (the mean being calculated across the critical serial positions 2, 4, and 6 for both measures). Analysis using paired *t*-tests revealed a significant drop in the *PSE* scores from short (*mean* = 0.215) to long (*mean* = .111) delay [$t(21) = 3.47$, $p < .01$], and that the change in *MVP* scores across delay (mean *MVP* for short: .033; mean *MVP* for long: .071) was not significant [$t(21) = 1.12$, $p = .28$].

Discussion

Replicating other studies (Farrell & Lewandowsky, 2003; Lewandowsky & Farrell, 2006), dissimilar items on mixed lists were found to be more accurately recalled than dissimilar items on pure lists: a mixed-list advantage. Additionally, it was found that whereas the phonological similarity effect significantly decreased with delay, the change in the mixed-list advantage from short to long delay was not significant; if anything, there was a small tendency for these scores to increase with delay. The latter result suggests that the mixed-list advantage does not scale with the phonological similarity effect, contrary to the predictions of retrieval-based models.

One limit with using standard statistics here is that the prediction of retrieval-based models is a quantitative one: the mixed-list advantage is expected to scale with the classic phonological similarity effect. To quantitatively address this issue, some simulations are now reported in which two models of serial recall, SEM (Henson, 1998), a dual-stage model, and SOB (Farrell & Lewandowsky, 2002), a model incorporating similarity-sensitive encoding, were applied to the data from the experiment.

Simulations

Start–End Model (Henson, 1998)

SEM is a model of serial recall that stresses the role of positional markers in remembering order information. It is assumed that storage of an item consists of the formation of an episodic token containing a representation of the item along with information about its position. For simple, ungrouped lists, position is uniquely represented by the value of a start marker and an end marker, both of which vary across serial positions as an exponential function. Formally, the value of the start marker at serial position j, $s(j)$, is given by

$$S_0 S^{j-1}, \tag{1}$$

while the value of the end marker at position j, $e(j)$, is given by

$$E_0 S^{N-j}, \tag{2}$$

where S_0 and E_0 are the respective maximum values of the start and end markers, and S and E are the rate of change in start and end markers across serial positions (with N the number of items on the list). Each position j then has associated with it a vector containing a start element and an end element, $\mathbf{p}(j) = \{s(j), e(j)\}$.

Recalling items in order consists of matching a sequence of positional markers to the same information in the episodic tokens. The overlap, o, between any two positional markers $\mathbf{p}(i)$ and $\mathbf{p}(j)$ is given by

$$o[\mathbf{p}(i), \mathbf{p}(j)] = \sqrt{\mathbf{p}(i) \bullet \mathbf{p}(j)} \times \exp\left\{-\sqrt{\sum_k [p_k(i) - p_k(j)]^2}\right\}, \tag{3}$$

where k indexes the individual elements of the positional vector (representing start and end markers) and $\bullet$ is the dot product operator (see, e.g., Anderson, 1995). At recall, the positional marker corresponding to a position j is matched to all episodic tokens in parallel, with the competition for response at position j offered by item i being

$$c_c(i,j) = o[\mathbf{p}(i), \mathbf{p}(j)][1 - r(i)] + N(0, G_c). \tag{4}$$

In Eq. (4) the first term is the overlap between the positional marker for position j and the positional marker stored in the episodic token of the item presented at position i. The second term represents the application of response suppression to items. In SEM, it is assumed that recall of an item is followed by its suppression; this prevents perseverative recalling of items, and is consistent with the low probability of erroneous repetitions of items observed empirically (Brown, Preece, & Hulme, 2000; Henson, 1998). In the simulations reported here, as in Henson (1998), the response suppression applying to item i, $r(i)$, was set to 0 at the start of a list, and was then set to 1 following recall of that item. Following Henson, response suppression was assumed to wear off across recall as

$$r_t(i) = r_{t-1}(i) \exp(-R_s), \tag{5}$$

where t indexes output position, and R_s was set to a constant 0.5 following Henson (1998). The final term in Eq. (4) is normally distributed noise with mean 0 and standard deviation G_c, a free parameter.

Following application of Eq. (4), the item v offering the strongest competition is selected from the ordering stage, and matched against a set of phonological traces in a further phonological confusion stage. The competition offered at this second stage is similar in form to that for the ordering stage

$$c_p(u) = c_c(u) + p(u,v)a_p(v)[1 - r(u)] + N(0, G_p), \tag{6}$$

where $p(u,v)$ is the phonological overlap between items v and u, $a_p(v)$ is the activation of the phonological representation of item v; here, $a_p(v)$ was set to 0 for items not presented on a list, to reflect the provision of participants with the list items at recall in the experiment. In Eq. (6) $r(u)$ is the response suppression of item u, and G_p, which scales the addition of normally distributed noise, was a free parameter. Following application of Eq. (4), the item offering the strongest competition is selected for output.

Phonological similarity was implemented in SEM by incorporating a multidimensional scaling (MDS) solution. The MDS was performed on acoustic confusability data of Hull (1973), who recorded the perceptual identification errors of 135 participants listening to letters (excluding "O") and digits presented auditorily in noise (a Euclidian distance model was used). A 3-dimensional solution was selected on the basis of a scree plot of stress measures.

To incorporate this knowledge into SEM, pair-wise phonological overlaps $p(u, v)$ were calculated by scaling the Euclidian distance in MDS space between the items,

$$p(u,v) = \exp\left(-c\sum_k [m_k(u) - m_k(v)]^2\right), \tag{7}$$

where k indexes the dimensions of the MDS solution, and c was a free parameter.

Delay was implemented in SEM as the decay of phonological representations. Henson (1998) assumed that delaying recall has two effects: the decaying of phonological activations of items, and the updating of positional markers to incorporate changes in context during the retention interval. To focus on phonological distinctiveness, especially given the emphasis on decay of phonological representations in Page and Henson (2001) and Henson (1998), and to maximise comparability with SOB, a sole locus of delay effects was specified here as the squashing of the phonological representations over time by changes in $a_p(v)$. Following Henson (1998), during the retention interval phonological representations were assumed to decay with each additional presented item, such that $\mathbf{a}_p = \exp(-R_s D)$, where R_s is a free parameter, and D was the number of distractor digits between list presentation and recall (4 for the short delay condition, and 16 for the long delay condition).

Serial-Order-in-a-Box model (Farrell, 2001; Farrell & Lewandowsky, 2002)

The Serial-Order-in-a-Box (SOB) model was originally introduced as a simple connectionist model of serial recall incorporating three basic principles argued to offer a self-contained, mechanistic account for basic serial recall phenomena (Farrell & Lewandowsky, 2002). The main principle of relevance here is that of similarity-sensitive encoding, introduced previously. This is the assumption that encoding of an item is sensitive to the similarity of that item to current memory contents. Farrell and Lewandowsky (2002) showed that this assumption naturally leads to a primacy gradient, and argued that similarity-sensitive encoding, as implemented in SOB, thus offers a process account for an assumption often made in serial recall models (Brown et al., 2000; Page & Norris, 1998).

The model employed here (see also Farrell, 2001) extends the basic model of Farrell and Lewandowsky (2002) by relaxing an assumption of orthogonality of input items made in that model. The model also incorporates positional markers as in most models of serial recall (Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998); this allows potential extension of the model to hierarchical phenomena (Hitch, Burgess, Towse, & Culpin, 1996) beyond the purview of the basic SOB model (Farrell & Lewandowsky, 2002), and is also necessitated by the incorporation of similarity between items (Lewandowsky & Farrell, 2006). Finally, since the focus here was on the probability of recall and not the dynamics of recall, the recurrent recall procedure implemented in SOB was replaced by a simpler matching mechanism that serves the same process of converting a "noisy" item representation into a recallable item (see Farrell, 2001; Farrell & Lewandowsky, 2002; Lewandowsky, 1999; Lewandowsky & Farrell, 2000).

The model consists of two layers of units; an input layer ($N = 16$) used to represent positional markers, and an output layer ($N = 150$) representing list items, where N refers to the number of units in each layer. The two layers are fully interconnected by a weight matrix $\mathbf{W}$.

As in SEM, it is assumed that each list position has associated with it a positional marker. For the simulations here, markers of 16 elements (consisting of real values) were constructed such that the similarity (as measured by cosine) between any two items was an exponential function of their distance; that is,

$$\cos(\mathbf{p}_i, \mathbf{p}_j) = t_c^{(|i-j|)}, \tag{8}$$

where $\mathbf{p}_i$ and $\mathbf{p}_j$ are the positional markers for position i and j, respectively. At list presentation, a pattern of activation corresponding to the position i is placed across the input units, and the pattern of activation corresponding to the item presented at position i is placed across the output units, and the weights between units updated to reflect learning. Formally, the item presented at position i, $\mathbf{v}_i$, is associated with the positional marker corresponding to its position, $\mathbf{p}_i$, by updating the input–output weights using Hebbian learning:

$$\Delta\mathbf{W}_i = \eta_e(i)\mathbf{v}_i\mathbf{p}_i^{\mathrm{T}}. \tag{9}$$

The weighting of the position–item associations is not constant but, embodying one of the key principles of SOB, is sensitive to the similarity of that item to items previously presented on the list. The measure of similarity used is called *energy*, and is given by

$$E_i = -\mathbf{v}_i^{\mathrm{T}}\mathbf{W}_{i-1}\mathbf{p}_i. \tag{10}$$

Energy is a measure of the overlap between a particular item–position pairing, and the item–position pairings previously stored. With the exception of the first item, which was given a constant weighting, the encoding weight of items, η_e, was inversely related to their energy:

$$\eta_e(i) = \begin{cases} 1, & i = 1, \\ \frac{-\phi_e}{E_i}, & i > 1, \end{cases} \tag{11}$$

where E_i is the energy of the item at position i, and ϕ_e is a model parameter. Because the positional markers are correlated, this means that storing an item will reduce the encoding rate of all following similar items, though the extent of this reduction will also depend on their proximity (see Appendix A).

Recall consists of stepping through the positional markers in order, and using them to cue for the item with which they were associated at encoding (see, e.g., Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998). First, the pattern of activation corresponding to the positional marker, $\mathbf{p}_i$, is placed across the input units, and the units in the item (output) layer update their activations according to the sum of the activations weighted by the connections $\mathbf{W}$

$$\mathbf{v}_i' = \mathbf{W}_i\mathbf{p}_i. \tag{12}$$

As in most associative connectionist models, the output is not an accurate representation of the item being cued for, being contaminated by cross-talk (Anderson, 1995). Following Brown et al. (2000), it was assumed that items were recalled by matching them to an experimental vocabulary. As in the implementation of SEM, to reflect the presentation of list items to participants at output, potentially recallable items on a particular trial were restricted to items presented on that trial's list. The noisy output $\mathbf{v}_i'$ was matched to items by calculating the similarity between the noisy output and each item

$$s(\mathbf{v}'_i, \mathbf{v}_k) = \exp[-cD(\mathbf{v}'_i, \mathbf{v}_k)^2], \tag{13}$$

where D is the Euclidian distance between the two vectors $\mathbf{v}'_i$ and $\mathbf{v}_k$, and c was a free parameter as in SEM. Before application of Eq. (13) the distances were normalized by subtracting the minimum distance across all the matched items from the distance for each item; this ensured that the obtained distances were close to 0, as otherwise, due to the amplification inherent in Eq. (12), the distances were all extremely large and did not allow computation of exponential scaling in Eq. (13). An item was then recalled by random sampling according to probabilities calculated using the Luce–Shepard choice rule (Luce, 1963; Shepard, 1957):

$$P(\mathbf{v}_v) = \frac{s(\mathbf{v}'_i, \mathbf{v}_v)}{\sum_{k=1}^{n} s(\mathbf{v}'_i, \mathbf{v}_k)}, \tag{14}$$

where $P(\mathbf{v}_v)$ is the probability of recalling item $\mathbf{v}_v$.

Following recall of an item in SOB, that item is suppressed. The response suppression occurring in SOB contrasts with that in SEM in two respects. One is that response suppression is implemented using the same mechanism as original learning (Farrell & Lewandowsky, 2002); that is, response suppression consisted of the unlearning of the association between the recalled item and the positional marker used as a cue for the item. Formally,

$$\Delta\mathbf{W}_j = \eta_s(j)\mathbf{v}_{o,j}\mathbf{p}_j^{\mathrm{T}}, \tag{15}$$

where $\mathbf{v}_{o,j}$ is the item recalled at position j, and $\mathbf{p}_j$ is the positional marker for position j. A second difference from SEM is that the weighting of response suppression, $\eta_s(j)$, is calculated from the energy of the output item in the context of the positional marker $\mathbf{p}_j$, $E_j = -\mathbf{v}_{o,j}^{\mathrm{T}}\mathbf{W}_{j-1}\mathbf{p}_j$, as

$$\eta_s(j) = \frac{-E_j}{\phi_s E_1}. \tag{16}$$

In Eq. (16) E_1 is the energy of the item first recalled, and ϕ_s was a free parameter weighting response suppression. Output interference was also assumed in SOB, and was implemented by adding Gaussian noise with standard deviation N_{O} (a free parameter) to each weight in $\mathbf{W}$ after each retrieval.

Item representations (vectors corresponding to patterns of activation) were constructed to reflect the configuration of items in phonological MDS space. Each item vector was divided into three segments each containing 50 elements; each segment was taken to represent one of the three MDS dimensions. Two random vectors of 150 elements from the set $\{-1,+1\}$ were then constructed for each simulation replication; these vectors represented, respectively, the origin and the maximum possible value of all coordinates. MDS coordinates were rescaled to a value p within the range 0–1, and for any given stimulus, each feature was then sampled from the maximum-coordinate vector with probability p and from the origin with probability $1 - p$.

Delay was implemented in SOB following a precedent in the OSCAR model of Brown et al. (2000). In OSCAR, forgetting during delay periods is treated as additional output interference and implemented as the addition of Gaussian noise to the weights storing associations between temporal markers and items. A similar assumption was made here, with delay being modelled by the addition of random Gaussian noise (mean 0 and standard deviation G_{o}) for each distractor during the retention interval (4 for the short delay condition, and 16 for the long delay condition).

Simulation details

Each model had five free parameters that were used to fit both delay conditions in parallel. In SEM, the free parameters were the ratio of the maximum end marker to the maximum start marker (F_0, such that $E_0 = F_0S_0$; see Henson, 1998, for details); the standard deviation of noise added to the positional confusion stage (G_{c}); the standard deviation of noise added to the phonological confusion stage (G_{p}); the scaling of phonological distances in calculating phonological similarities (c); and the decay rate of phonological representations during the retention interval (R_{p}). The first four of these were left as free parameters as they were also freely varied in the demonstrations of Henson (1998; see his Table B1); R_{p} was left as a free parameter to allow SEM to quantitatively capture the effects of delay. In SOB, the free parameters were the positional similarity, t_{c}; the response suppression parameter, ϕ_s; a parameter scaling output interference, N_{O}; the scaling of matches in calculating response probabilities, c; and the noise added to weights for each distractor digit, G_{o}. To limit the number of free parameters, the parameter ϕ_{e}, left as a free parameter in other applications (Lewandowsky & Farrell, 2006), was set to be a fixed proportion of the positional similarity parameter t_{c} (see Appendix A).

Model fitting proceeded in two steps. First, each model was fit to the 60 data points in Fig. 3 by minimizing the root mean squared deviation (*RMSD*) between the model predictions and the data. This preliminary fit then provided starting points for a second phase of parameter estimation in which only the data for pure lists was fit, and predictions for the mixed-list advantage were examined. This procedure, a form of cross-validation, was used to examine the unconstrained predictions of the mixed-list advantage given the models' best account of the phonological similarity effect. Parameters were estimated at both stages using the Nelder–Mead algorithm (Nelder & Mead, 1965). At each step model predictions were obtained from 1000 simulation replications.

Simulation results

The overall predictions of SEM after being fit to the data for the pure lists are shown in Fig. 5, and the critical points replotted, as for the data, in Fig. 6 (the minimized *RMSD* for pure list fits was .011; the best-fitting parameter estimates were $F_0 = 0.439$; $G_c = .028$; $G_p = .409$; $c = .128$; $R_p = .081$). Inspection of these figures reveals SEM to be unable to quantitatively capture the mixed-list advantage seen in the data (Figs. 3 and 4) despite the model's ability to capture the standard phonological similarity effect. In particular, in both Figs. 5 and 6 it is apparent that although SEM can produce a slight mixed-list advantage for the short delay condition, this advantage is effectively eliminated after a long delay. The same sets of predictions from SOB (pure list $RMSD = .044$; parameter estimates were $t_c = 0.496$; $\phi_s = 1.209$; $N_O = 1.436$; $c = 0.298$; $G_o = 0.297$) are shown in Fig. 7 (overall predictions) and Fig. 8 (condensed plot of predictions). Comparison of these figures to Figs. 3 and 4 reveals the predictions of SOB to be strikingly similar to the data; in particular, the model clearly predicts the enhancing effects of mixing on dissimilar items for both short and long delay.

Table 1 presents two types of *RMSD*s between the models and the data: the *RMSD* between the isolate points in data and models, and the *RMSD* between the *MVP* scores in the model and the data. The *RMSD* values confirm SEM's inability to quantitatively capture the data, and the superiority of SOB in predicting the mixed-list advantage. In particular, it can be noted that although the models are equally able to capture the *MVP* effect at a short delay (*MVP RMSD*), at a long delay SOB clearly provides a quantitatively superior account of the data.

The two-stage fitting procedure outlined above, a simple version of cross-validation (e.g., Efron & Gong, 1983), was used to determine the extent to which the models predict the mixed-list advantage by virtue of their core principles, and not due to the action of the model in some restricted region of parameter space. In order to provide a more global index of the models' per-

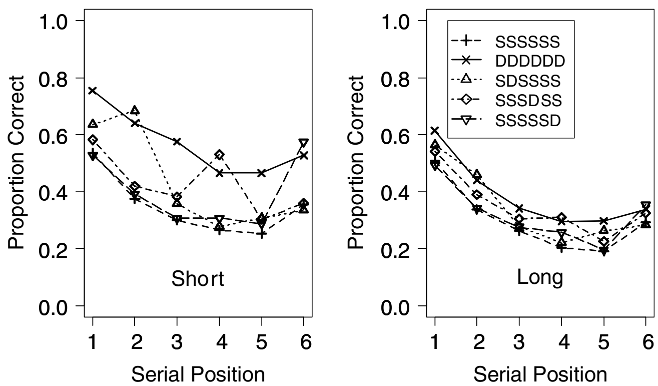

Fig. 5. Serial position functions for mean proportion correct predicted by SEM.

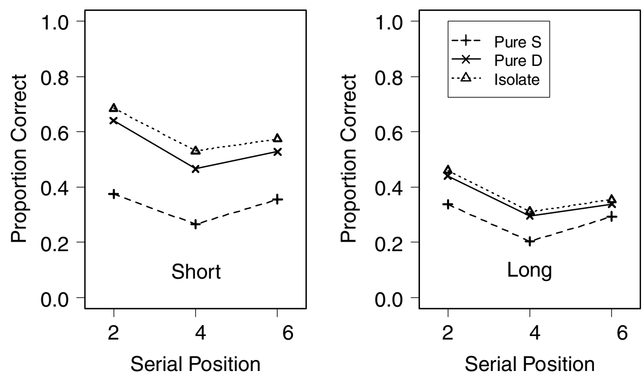

Fig. 6. Condensed serial position functions for mean proportion correct predicted by SEM.

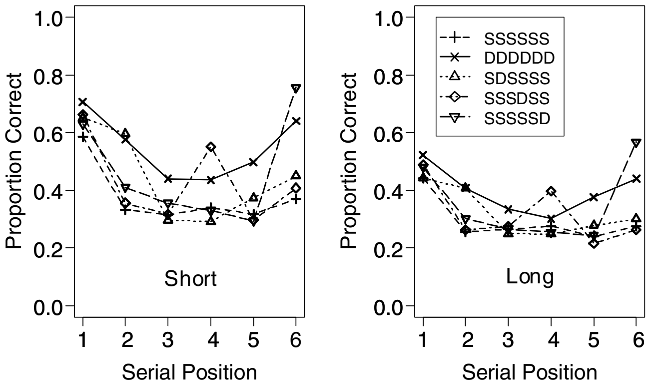

Fig. 7. Serial position functions for mean proportion correct predicted by SOB.

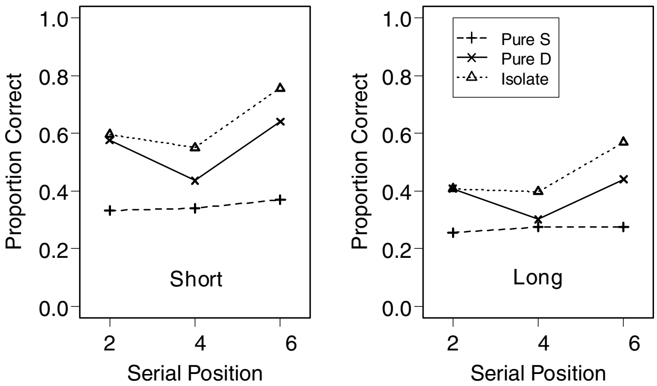

Fig. 8. Condensed serial position functions for mean proportion correct predicted by SOB.

Table 1
*RMSD*s between model predictions and data for SEM, SOB, and the control SOB

	Short		Long	
	Isolate	*MVP*	Isolate	*MVP*
SEM	.128	.061	.122	.071
SOB	.044	.064	.024	.023
SOB control	.092	.105	.093	.089

formance, the models were again fit to all the data points from the experiment (i.e., the first stage of the above procedure), with the exception that the goodness-of-fit measure used was χ^2, calculated using equations presented in Press, Teukolsky, Vetterling, and Flannery (1992, pp. 661–666). Importantly, this χ^2 measure weights each deviate by the standard error of that mean in the data; thus, data points with more accurate estimates (i.e., lower standard errors) will be given greater weighting, and we can be confident that any differences in model fits are not due to one model being better at fitting noise.[4]

The results of this simulation exercise revealed that for SEM, $\chi^2 = 63.83$, whereas for SOB, $\chi^2 = 39.02$; this equates to a likelihood ratio[5] of over 243,580, clearly favouring SOB. This further demonstrates SOB's better quantitative account of the data, and indicates that this superiority is not due a better ability of SOB to fit noise.

To confirm that the mixed-list advantage predicted by SOB was directly due to the similarity-sensitive encoding mechanism embodied in Eq. (11), a further simulation was run examining the predictions of a control version of SOB. This control model was identical to SOB in all respects except for the manner in which the encoding weight of items (η_e) was determined. Unlike SOB, these weights were insensitive to item similarity in the control model, and assumed to drop off exponentially across serial positions (Brown et al., 2000; Lewandowsky, 1999; Lewandowsky & Farrell, 2000),

[4] Thanks go to Dennis Norris for suggesting the use of χ^2.

[5] Likelihoods were obtained by $L = \exp(-0.5\chi^2)$, assuming data points were independently Gaussian distributed.

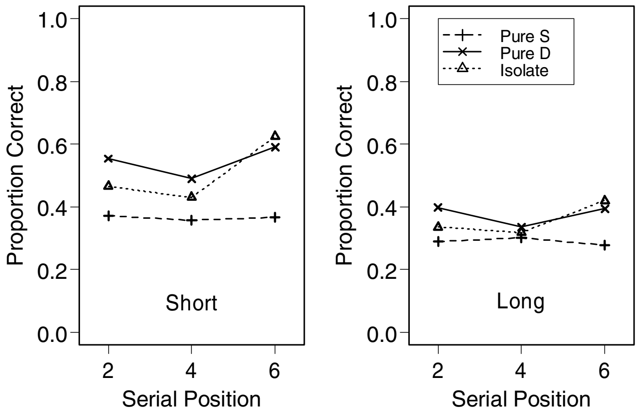

Fig. 9. Condensed serial position functions for mean proportion correct predicted by a control version of SOB.

$$\eta_e(i) = (i-1)^{-\lambda}, \quad (17)$$

where λ was a free parameter. Thus, phonological confusions in the model are dictated by the similarity structure in Fig. 1 as in SEM and SOB; unlike SOB, and like SEM, storage of items is not sensitive to item similarity. Following application of SEM and SOB, the control version of SOB was fit to the entire data set, and the used the resulting parameter values were then used as starting points for fitting of only the pure list data. Fig. 9 summarises the key predictions of the model, and Table 1 gives the *RMSD*s for the isolate points. The figure and table show that the control model without similarity-sensitive encoding is unable to produce a mixed-list advantage; indeed, the model predicts a mixed-list interference effect clearly inconsistent with the data (see similar suggestions in Henson et al., 1996 & Page & Norris, 1998, for chaining models and OSCAR, respectively). Not only does this confirm the tight link between the predicted mixed-list advantage and similarity-sensitive encoding, as discussed below this also suggests some adaptive benefit of similarity-sensitive encoding.

General discussion

Convergent with other studies (Farrell & Lewandowsky, 2003; Lewandowsky & Farrell, 2006), the present paper confirmed the existence of a mixed-list phonological similarity advantage in serial recall: placing phonologically dissimilar items on lists of similar items enhances their ordered recall. More importantly, it has been shown that one explanation for the mixed-list advantage as embodied in SEM (Henson, 1998), the increased distinctiveness of dissimilar items on mixed lists at retrieval, under-predicts the extent of the advantage in delayed recall. Since this model predicts that the detrimental effect of similarity on similar items and the mixed-list advantage for dissimilar items should scale together, it cannot account for the presence of a mixed-list advantage when the phonological similarity effect is minimized. In contrast, the SOB model, in which encoding is sensitive to the similarity of items predicts the prominent mixed-list advantage after a delay. In this model, the reduced encoding given to successive similar items will be present even after a delay, allowing the model to predict a mixed-list advantage in the face of a reduced or absent phonological similarity effect.

One point to note is that the mixed-list advantage was observed here even when randomizing the order of lists. Previous experiments examining mixed-list effects have tended to present lists in a blocked design, such that lists generated from a particular list structure (e.g., *SSSDSS*) are presented contiguously in an experiment (Baddeley, 1968; Farrell & Lewandowsky, 2003; Henson et al., 1996). One concern with such an experimental design is that any effects observed, particularly the mixed-list advantage, might be due to demand characteristics given the predictability of list contents when lists are blocked. The finding of a mixed-list advantage in the experiment presented here, where the predictability of list contents was limited by randomization, shows the enhancing effect of similarity mixing is not due to expectations on the part of participants (see also Lewandowsky & Farrell, 2006).

The results provide further support for models of serial order memory in which similarity has effects on the encoding of order, and call into question "dual-stage" models in which the encoding and representation of order is insensitive to the nature of items themselves (Brown et al., 2002; Henson, 1998; Page & Norris, 1998). Apart from SOB, one other encoding-based model that offers an account of the mixed-list advantage is the feature model (Nairne, 1990). Lewandowsky and Farrell (2006) showed that the feature model qualitatively predicts the mixed-list advantage, although quantitatively the effect was fairly small. Given the small effects in that simulation, it is unclear whether the mixed-list advantage would be apparent in the model after a delay; however, at a general level the model should predict that the mixed-list advantage will not scale with the classic phonological similarity effect as is assumed in phonological decay models (Page & Henson, 2001), since there is an added effect of similarity at encoding.

SOB, the encoding-based model presented here, provides a natural account of the mixed-list advantage in its mechanism of similarity-sensitive encoding. Importantly, similarity-sensitive encoding is not a mechanism specifically developed to account for the mixed-list advantage, but is a potentially general adaptive memory mechanism. One adaptive benefit of similarity-sensitive encoding can be seen in the demonstrations of Farrell and Lewandowsky (2002), who showed that incorporating similarity-sensitive encoding in a connectionist network naturally leads to a primacy gradient, an element necessary to the operation of a number of models of serial recall (Henson, 1998; Page & Norris, 1998). More generally, it may be seen as an adaptive mechanism to deal with the problem of cross-talk in the connectionist network we suggest underlies short-term memory (see also Brown et al., 2000); as shown in the preceding simulations, removing the similarity-sensitive encoding mechanism from SOB leads to excessive interference from similar items. Brown et al. (2000) motivated the encoding gradient in their OSCAR model by arguing that a rational organism, in deciding how to allocate attentional resources to predict changes in a dynamic environment, should pay increasingly less attention to a stream of homogeneous (i.e., similar) items. SOB's similarity-sensitive encoding can be seen as a computational instantiation of this principle in serial order memory, and sits well in an emerging framework in which some cognitive and neural mechanisms have developed to circumvent the limitations of connectionist networks whilst taking advantage of their powerful properties (e.g., McClelland, McNaughton, & O'Reilly, 1995; O'Reilly, 1998).

Appendix A

In the simulations of SOB it was desired that the encoding parameter ϕ_e, be fixed as a constant proportion of the positional similarity parameter t_c. To allow this scaling, it is first necessary to show that the energy measure to which ϕ_e is applied is directly proportional to contextual similarity t_c. Take the case where a single item-order association has been stored

$$\mathbf{W}_1 = \mathbf{v}_1\mathbf{p}_1^{\mathrm{T}} \tag{A.1}$$

The energy of a new association to be formed, $\mathbf{v}_2\mathbf{p}_2^{\mathrm{T}}$, will then be

$$\begin{aligned} E_2 &= -\mathbf{v}_2\mathbf{p}_2^{\mathrm{T}}\mathbf{W}_1\mathbf{p}_2 \\ &= -\mathbf{v}_2\mathbf{p}_2^{\mathrm{T}}\mathbf{v}_1\mathbf{p}_1^{\mathrm{T}}\mathbf{p}_2 \\ &= -(\mathbf{v}_2 \bullet \mathbf{v}_1)(\mathbf{p}_1 \bullet \mathbf{p}_2). \end{aligned} \tag{A.2}$$

Item vectors were of equal length $\sqrt{N}$ (N being the number of vector elements), and the algorithm for generating the positional markers produced vectors of constant length 4; given that

$$\cos(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a} \bullet \mathbf{b}}{\sqrt{(\mathbf{a} \bullet \mathbf{a})}\sqrt{(\mathbf{b} \bullet \mathbf{b})}} \tag{A.3}$$

(Anderson, 1995), the energy is then

$$\begin{aligned} E_2 = &-\left[\sqrt{(\mathbf{v}_1 \bullet \mathbf{v}_1)}\sqrt{(\mathbf{v}_2 \bullet \mathbf{v}_2)}\cos(\mathbf{v}_1,\mathbf{v}_2)\right] \\ &\times \left[\sqrt{(\mathbf{p}_1 \bullet \mathbf{p}_1)}\sqrt{(\mathbf{p}_2 \bullet \mathbf{p}_2)}\cos(\mathbf{p}_1,\mathbf{p}_2)\right] \\ = &-[N\cos(\mathbf{v}_1,\mathbf{v}_2)][16t_c] \end{aligned} \tag{A.4}$$

Thus, given two items $\mathbf{v}_1$ and $\mathbf{v}_2$ the energy between the associations $\mathbf{v}_1\mathbf{p}_1^T$ and $\mathbf{v}_2\mathbf{p}_2^T$ is directly proportional to the positional similarity parameter t_c. Given a fixed sequence of item vectors, obtaining a fixed set of learning rates for varying t_c is then accomplished via Eq. (11) by making ϕ_e directly proportional to t_c, $\phi_e = kt_c$. The constant k was estimated from the SOB model fits of Lewandowsky and Farrell (2006) by dividing the average energy of the second incoming association in their simulations by their $t_c = .25$, giving a k of 976.

References

Anderson, J. A. (1995). *An introduction to neural networks*. Cambridge, MA: MIT Press.

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, 20*, 249–263.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

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

Brown, G. D. A., Neath, I., & Chater, N. (2002). A ratio model of scale-invariant memory and identification. *Manuscript under revision*.

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

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.

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

Efron, B., & Gong, G. (1983). A leisurely look at the boostrap, the jackknife, and cross-validation. *The American Statistician, 37*, 36–38.

Farrell, S. (2001). *Similarity-sensitive encoding, redintegration, and response suppression in serial recall*. Unpublished PhD thesis, University of Western Australia, Perth.

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

Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 29, 838–849.

Henson, R. N. A. (1998). Short-term memory for serial order: the Start–End Model. *Cognitive Psychology, 36*, 73–137.

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

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: a working memory analysis. *Quarterly Journal of Experimental Psychology, 49A*, 116–139.

Hull, A. J. (1973). A letter–digit matrix of auditory confusions. *British Journal of Psychology, 64*, 579–585.

Lewandowsky, S. (1999). Redintegration and response suppression in serial recall: a dynamic network model. *International Journal of Psychology, 34*, 434–446.

Lewandowsky, S., & Farrell, S. (2000). A redintegration account of the effects of speech rate, lexicality, and word frequency in immediate serial recall. *Psychological Research, 63*, 163–173.

Lewandowsky, S., & Farrell, S. (2006). Phonological similarity in serial recall: Constraints on theories of memory. *Unpublished manuscript*.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.). *Handbook of mathematical psychology* (Vol. 1, pp. 103–189). New York: Wiley.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419–457.

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

Nairne, J. S., & Kelley, M. R. (1999). Reversing the phonological similarity effect. *Memory & Cognition, 27*, 45–53.

Navarro, D. J., Pitt, M. A., & Myung, I. J. (2004). Assessing the distinguishability of models and the informativeness of data. *Cognitive Psychology, 49*, 47–84.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7*, 308–313.

Norris, D., Baddeley, A. D., & Page, M. P. A. (2004). Retroactive effects of irrelevant speech on serial recall from short-term memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 30*, 1093–1105.

O'Reilly, R. C. (1998). Six principles for biologically based computational models of cortical cognition. *Trends in Cognitive Sciences, 2*(11), 455–462.

Page, M. P. A., & Henson, R. N. A. (2001). Computational models of short-term memory: modelling serial recall of verbal material. In J. Andrade (Ed.), *Working memory: A work in progress* (pp. 177–198). London: Routledge.

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

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in C: The art of scientific computing*. New York, NY: Cambridge University Press.

Shepard, R. N. (1957). Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. *Psychometrika, 22*, 325–345.

Wickelgren, W. A. (1965a). Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 4*, 53–61.

Wickelgren, W. A. (1965b). Short-term memory for phonemically similar lists. *American Journal of Psychology, 78*, 574–576.