



Lexical and semantic binding in verbal short-term memory

Elizabeth Jefferies^{a,*}, Clive R. Frankish^b, Matthew A. Lambon Ralph^a

^a University of Manchester, UK

^b University of Bristol, UK

Received 8 February 2005; revision received 30 August 2005

Available online 15 November 2005

Abstract

Semantic dementia patients make numerous phoneme migration errors in their immediate serial recall of poorly comprehended words. In this study, similar errors were induced in the word recall of healthy participants by presenting unpredictable mixed lists of words and nonwords. This technique revealed that lexicality, word frequency, imageability, and the ratio of words to nonwords all influence the stability of the phonological trace. These factors affected phoneme migrations and phoneme identity errors for both the words themselves and the nonwords they were presented with. Therefore, lexical/semantic knowledge encourages the phonological segments of familiar words to emerge together in immediate serial recall. In the absence of such knowledge, the elements of a particular item are more likely to recombine with the phonemes of other list items. These findings demonstrate the importance of lexical and semantic binding in verbal short-term memory.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Redintegration; Semantic binding; Semantic dementia; Verbal short-term memory; Immediate serial recall

It is well established that verbal short-term memory (STM) draws heavily on a phonological code, accounting for the occurrence of phonological errors in immediate serial recall tasks (Conrad, 1964) and the poorer recall of lists of phonologically similar than dissimilar items (Conrad & Hull, 1964). In line with these findings, the highly influential working memory model proposes that verbal STM is underpinned by a phonological loop subsystem that operates independently of long-term memory (Baddeley, 1986; Baddeley & Hitch, 1974). However, research has indicated that stable lexical and semantic representations also make an important contribution to immediate serial recall. Immediate serial recall

is better for: (1) words vs. nonwords (e.g., Brener, 1940; Hulme, Maughan, & Brown, 1991); (2) high compared with low frequency words (Gregg, Freedman, & Smith, 1989; Hulme et al., 1997; Roodenrys & Quinlan, 2000; Watkins & Watkins, 1977); (3) imageable/concrete words vs. abstract words (Bourassa & Besner, 1994; Walker & Hulme, 1999); (4) semantically grouped vs. ungrouped words (Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999).

Neuropsychological evidence also points to the role of lexical–semantic representations in immediate recall. Semantically impaired aphasic patients show reduced effects of frequency and imageability in immediate serial recall (Martin & Saffran, 1997). In addition, semantic dementia patients, who show a specific and progressive decline in semantic memory, have better immediate recall of words that they understand relatively well compared with words whose meanings have become

* Corresponding author. Fax: +44 161 275 2588.

E-mail address: beth.jefferies@manchester.ac.uk (E. Jefferies).

degraded over the course of the disease (Jefferies, Jones, Bateman, & Lambon Ralph, 2004; Jefferies, Jones, Bateman, & Lambon Ralph, 2005; Knott, Patterson, & Hodges, 1997, 2000; Patterson, Graham, & Hodges, 1994). Immediate serial recall differences between relatively well-known and semantically degraded words have also been reported in semantically impaired patients following cerebrovascular accident (Forde & Humphreys, 2002) and herpes simplex encephalitis (Caza, Belleville, & Gilbert, 2002).

Although it is widely accepted that lexical/semantic factors contribute to verbal STM, there is still considerable controversy about the mechanisms involved. The recall errors of patients with semantic dementia appear to offer an intriguing insight into this issue. Semantic dementia patients make numerous phoneme migration errors in immediate serial recall, particularly for words that they understand poorly (Patterson et al., 1994). This finding suggests that in healthy individuals, stable semantic representations help to maintain the correct configuration of phonological elements in STM. Patterson et al. (1994) found that immediate serial recall errors in semantic dementia typically consisted of incorrect combinations of phonemes from different words that largely preserve onset/rime syllable structure (e.g., the onsets in the words ‘mint, rug’ were exchanged to produce the response ‘rint, mug’). Treiman and Danis (1988) observed a similar pattern of phoneme migration errors in normal participants’ recall of lists of nonwords (e.g., ‘gir, vang, kus’ recalled as ‘gir, kang, vus’). In contrast, healthy participants rarely make phoneme migration errors in word recall. Instead, whole-item order errors constitute the majority of immediate serial recall errors in many studies (e.g., Henson, Norris, Page, & Baddeley, 1996; Pickering, Gathercole, & Peaker, 1998), particularly when a small pool of repeating items is used to construct the lists for recall (Coltheart, 1993; Gathercole, Pickering, Hall, & Peaker, 2001).

This striking association between a lack of semantic support and the emergence of phoneme migration errors led Patterson et al. (1994) to propose that semantic memory makes a major contribution to the coherence of phonological representations in STM. According to this “semantic binding hypothesis,” there are two sources of constraint that contribute to the phonological coherence of words. First, the constituents of words become associated in the phonological system because they are always activated together during speech production/comprehension; as a result, they are more likely to emerge together in immediate serial recall. Secondly, because specific semantic activation frequently co-occurs with the phonology of words, semantic constraints encourage word phonemes to emerge together in the correct order in immediate serial recall. As immediate serial recall tasks require subjects to maintain a phonological trace of an entire list of items, a loosening of

these constraints might allow phonemes to migrate between items. This hypothesis posits distinct semantic and phonological representations but no separate lexical level, in line with the parallel distributed processing (PDP) models of, for example, Seidenberg and McClelland (1989). ‘Lexical’ constraints on phonology result from a combination of long-term learning of frequently co-occurring phonemes and stable associations between phonological and semantic representations.

The semantic binding account suggests that verbal STM emerges from the language system; therefore, lexical and semantic constraints are integral to it. Several other researchers have also advocated this view: for example, Martin and Saffran (1997) suggested that verbal STM results from interactive activation between phonological, lexical, and semantic nodes following Dell and O’Searghda’s (1992) model of speech production. In addition, Gathercole and Martin (1996) likened the operation of phonological STM to McClelland and Elman’s (1986) TRACE model of speech perception.

A rather different account of the long-term memory contribution to verbal STM proposes that two separate mechanisms underlie immediate serial recall performance (Baddeley, Gathercole, & Papagno, 1998; Hulme et al., 1991, 1997; Schweickert, 1993). (1) There is a rapidly decaying phonological STM store, which is initially inert to the effects of lexical and semantic factors. (2) There is a later reconstructive process that compares the degraded STM trace with separate long-term phonological–lexical representations during recall to reinstate the correct phonological activation for words. Although this reconstructive process is underpinned by phonological–lexical representations, the model can account for semantic effects in immediate serial recall by assuming that semantic activation contributes to the selection of lexical candidates for reconstruction (Poirier & Saint-Aubin, 1995). This view, which has largely emerged from research on normal participants, suggests that the long-term memory contribution to verbal STM is late-stage and item specific. As the reintegration of each item is predicted to occur at a late stage, there is no mechanism by which lexical/semantic factors pertaining to one item can influence the phonological decay and reconstruction of other list items (see Schweickert, 1993). In line with this proposal, several computational models of immediate serial recall suggest (or, if extended, would suggest) that lexical-level representations restore the appropriate phonological activation for individual words at an output stage, after order and partial item information has been recovered (Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998). It is important to acknowledge, however, that not all reintegrative theories predict that late-stage reconstruction is specific to individual items. For example, the feature model (Nairne, 1990) proposes that degraded representations in primary memory are

matched to possible responses in secondary memory: this ‘search set’ is thought to consist principally of the items to be recalled and, consequently, the likelihood of correct redintegration for a particular item will depend on the nature of the other list items. In addition, recent research has suggested that the strength of inter-item associative links might influence the availability of lexical representations during late-stage redintegration (Hulme, Stuart, Brown, & Morin, 2003; Stuart & Hulme, 2000). Some aspects of the data presented here challenge the view that redintegration is purely item specific but might be accommodated by models that suggest redintegration is influenced by list composition.

The semantic binding and redintegration accounts are rather difficult to distinguish empirically. Both predict that variables thought to tap lexical/semantic processes will influence recall. In addition, the two theories are consistent with the observed influence of stable linguistic knowledge on item errors but not whole-item transpositions (Gathercole et al., 2001; Hulme et al., 1997; Poirier & Saint-Aubin, 1995, 1996; Saint-Aubin & Poirier, 1999, 2000; Walker & Hulme, 1999). According to the semantic binding hypothesis (Patterson et al., 1994), lexical/semantic factors have a major influence on the frequency of phoneme migrations that alter item identity: thus, when scoring recall at the level of whole items, the primary influence of stable linguistic knowledge is on item rather than order errors. Likewise, if the phonological trace of an item is degraded, item-specific redintegration should increase the probability of recalling the whole item and its constituent phonemes correctly but should not increase the probability of recalling the item in its correct serial position (Gathercole et al., 2001; Poirier & Saint-Aubin, 1995, 1996; Saint-Aubin & Poirier, 2000).

Despite these similarities, the two viewpoints have different expectations about the effect of lexical/semantic factors on phoneme migrations. A central prediction of the semantic binding hypothesis is that phoneme migration errors should occur less frequently for words than for nonwords, as semantic memory is thought to provide a major source of constraint on the order/clustering of phonemes. The redintegration theory, in contrast, does not offer compelling predictions about the effect of lexical/semantic knowledge on phoneme recombinations. In standard experiments of immediate word recall, few phoneme migrations are observed, presumably because redintegration can reinstate phonemes that have intruded into words. This process is not expected to restore phonemes for nonwords in the same way, and the frequent phoneme migrations that are observed might be a typical consequence of trace decay/interference (see Treiman & Danis, 1988). In the absence of redintegration, equal numbers of word and nonword components might be expected to migrate because phonological decay is not thought to be influenced by lexical/semantic

factors (e.g., Schweickert, 1993). Few studies have directly compared phoneme migration errors for words and nonwords. However, Gathercole et al. (2001) found that for 9-year-old children, lexicality had a small but significant effect on the occurrence of phoneme order errors and a much larger effect on identity errors, both at the level of whole items and individual phonemes.

Phonological errors in word recall, while rare in studies involving normal participants, are of particular interest because they make it possible to study the effect of semantic and lexical factors on the breakdown of phonological coherence in STM. Knott and Monsell (unpublished manuscript) found that healthy participants could be induced to make more frequent phoneme migration errors in their recall of words if they were presented with lists that contained an unpredictable mixture of words and nonwords. The nonword phonemes, which were presumably not tightly bound together as coherent items, appeared to damage the phonological integrity of the words. The majority of recall errors were recombinations of phonemes from different list items, resembling the pattern for semantic dementia. In these circumstances, clear effects of lexicality, frequency, and imageability on the degree of phonological breakdown were observed.

In this study, we replicated and extended these findings. First, we used the mixed lists methodology to evaluate a central tenet of the semantic binding account—the claim that lexical and semantic knowledge has a stabilising influence on the phonological trace. By this view, there should be fewer phoneme migration errors for words than for nonwords. The semantic binding account also predicts that the strength of binding will be modulated by factors such as frequency and imageability: there should be fewer phoneme migration errors for lists containing highly frequent and imageable words compared with low frequency/imageability words. Phonemes were not repeated within a list, allowing the sources of all phoneme migration errors to be traced.

Second, we examined whether the lexical/semantic characteristics of the words in mixed lists had any impact on *nonword* recall. This is an important issue because the various accounts make different predictions. The semantic binding hypothesis suggests that lexical/semantic factors pertaining to a particular item will influence phoneme migrations for other items in the list. There should be fewer opportunities for nonword phonemes to migrate in lists containing particularly frequent/imageable words, as these are more phonologically stable. In contrast, if late-stage redintegration operates in an item-specific manner, lexical/semantic factors should only influence the reconstruction of the words they pertain to and not the nonwords they are mixed with (see, for example, Schweickert, 1993).

We also compared the recall of words and nonwords in mixed lists (Experiment 1) with the same

items in pure lists (Experiment 2), to explore the effects of list composition on immediate serial recall. The semantic binding account might anticipate that words and nonwords will be recalled at an intermediate level in mixed lists because the presence of highly stable words should reduce the opportunity for nonword phonemes to migrate. In addition, nonword phonemes, which are not tightly bound together, might forcibly recombine with the elements of words, impairing their recall. These two factors together determine the likelihood of a particular phoneme migration: our hypothesis is that the elements of a specific word are encouraged to migrate when lexical/semantic knowledge is lacking for an item and also when lexical/semantic binding is weak for the other items in the list. The item-specific redintegration account, in contrast, predicts no effect of list composition on phoneme migrations or recall accuracy, as lexical and semantic factors should only affect the probability of successfully reconstructing each specific item. A recent study reported that lexicality and word frequency effects were reduced in mixed compared with pure lists, contrary to the predictions of the item-specific redintegration account (Hulme et al., 2003). Hulme and colleagues found that nonwords were more likely to be recalled when they were mixed with words (although word recall was unaffected by list composition: a point we will return to later). However, this study did not examine the influence of list composition on phoneme migrations, making it difficult to test the specific predictions of the semantic binding hypothesis with this data set.

Experiment 1: Mixed lists of words and nonwords

Normal participants were presented with five-item lists that contained an unpredictable mixture of words and nonwords. We anticipated that the nonword phonemes would damage the phonological integrity of the words, in line with the findings of Knott and Monsell (unpublished manuscript): consequently, we expected to see frequent item errors but only occasional whole-item order errors in the experiment as a whole. In addition, following the semantic binding hypothesis (Patterson et al., 1994), we predicted that phoneme identity and migration errors should be less common for frequent and imageable words relative to nonwords. Another prediction was that lexical/semantic influences on phonological stability would not be restricted to specific items but instead would be observed across the entire phonological trace. Consequently, both word and nonword phonemes were expected to migrate less frequently for lists containing high frequency/imageability words and a higher ratio of words to nonwords.

Method

Participants

The participants were 30 undergraduates, aged between 18 and 32, who spoke English as a first language and had normal hearing. They were tested individually and took part for course credit.

Design and materials

The experiment examined immediate serial recall for auditorily presented lists of five consonant–vowel–consonant (CVC) stimuli (see Appendix A). Every list contained a mixture of words and nonwords. The ratio of words to nonwords was varied as a within-subjects factor. Lists contained one word and four nonwords, two words and three nonwords or three words and two nonwords. Word frequency and imageability were also included as within-subjects factors. Words were assigned to four frequency by imageability groups on the basis of estimates of written word frequency and imageability taken from the Celex database (Baayen, Piepenbrock, & van Rijn, 1993) and the MRC psycholinguistic database (Coltheart, 1981). Mean frequency was 179 counts per million for the high frequency (HF) words (range = 51–656) and 6 for the low frequency (LF) words (range = 1–13). Mean imageability was 602 for the high imageability (HI) words (range = 573–659) and 442 for the low imageability (LI) words (range = 340–501). There were no significant imageability differences between groups that varied in frequency and no frequency differences between imageability groups (all $t(58) < 1$). Words with homophones were excluded if the frequency or imageability of the homophone was higher than that of the target word. The words in each list were drawn from a single frequency and imageability group.

Each participant was tested on 60 lists. There were 15 lists in each of the frequency by imageability categories. For each word type, there were five lists for the three ratios of words to nonwords. The words/nonwords occurred in different serial positions across these five lists, to prevent the participants from anticipating which items would be words/nonwords in advance. For lists containing one word (w) and four nonwords (n), the word occurred once in every serial position in the five lists (wnnnn, nwnnn, nwnn, nnnwn, nnnnw). There were 10 possible arrangements of words and nonwords for lists containing two or three words. Five were selected that minimised the degree to which the words and nonwords were clustered. In lists containing two words and three nonwords, the two words were never adjacent (wnwnn, wnnwn, nwnwn, nwnnw, nnwnw). In lists containing three words and two nonwords, the two nonwords were never adjacent (nwnnw, nwnnw, wnwnw, wnwnw, wnwnw).

The nonwords were constructed from the words by recombining the initial consonants, vowels, and final consonants to form new items. All the nonwords were legal and pronounceable, and were randomly assigned to the lists. Phonemes were not repeated within a list. Items were not repeated in the course of the experiment.

Procedure

The items were recorded individually in a flat intonation by a female speaker and were digitised using a computer. Sound editing software (Cool Edit, Syntrillium) was used to position the items in the lists so that they occurred at a rate of one item per second. Presentation of the lists was controlled using SuperLab software (Cedrus). The order of the trials was re-randomised for each participant. A red exclamation mark appeared on the computer screen just prior to the start of each trial. A blue question mark prompted participants to recall the list aloud. Participants' responses were recorded on tape and were later transcribed. The participants were told in advance that the lists would contain a mixture of both words and nonwords, and were given four practice trials. They were asked to recall the items in serial order and to make an attempt at each target, even if they were not sure they were correct. They did not say "pass" if they could not remember an item: instead, when fewer than five responses were produced, the omissions of items were positioned in the response transcripts in a way that minimised the numbers of phonemes occurring at the wrong serial positions. In a small number of trials (.6%), participants produced six rather than five items and the final response was discarded from the analysis.

Results

Table 1 shows the percentage of words and nonwords recalled in the correct serial position in each condition. Table 2 shows the outcome of ANOVAs examining recall accuracy by subjects and items. We will focus below on the most theoretically informative aspects of the data; namely the effects of lexical/semantic factors on identity and order errors at the whole-item and phoneme levels. First, however, the interactive influences of these factors on recall accuracy can be succinctly summarised. Recall in position was affected by the full range of lexical/semantic factors (lexicality, frequency, imageability, and the ratio of words to nonwords). Word frequency and imageability impacted on recall accuracy for both the words themselves and the nonwords they were mixed with. Lexicality and frequency had a greater influence on recall when there were more words in the list. In contrast, imageability effects were more substantial for lists containing more nonwords (see Table 2). This semantic variable may have played a greater role when the coherence of the phonological trace was jeopardised by the presence of nonwords.

The influence of lexical and semantic factors (lexicality, frequency, imageability, and the ratio of words to nonwords) on the number of order and identity errors was examined at both the level of whole items and individual phonemes. Item order errors were target items recalled in the wrong serial position. Item identity errors included both omission and commission errors, and occurred when the target item was not recalled correctly in any position. At the phoneme level, the same criteria were applied: phoneme order errors were target phonemes that migrated to the wrong list position and phoneme identity errors were target phonemes not recalled in any position. In the example "hush, woam" → "lush, home", h → l is a phoneme identity error and w → h is a phoneme order error (both responses are also item identity errors). Contrasting these two levels of analysis established whether the phonemes of a target migrated together or independently. In Fig. 1, identity errors at the level of whole items are shown as a percentage of the number of presented items. Similarly, identity errors at the phoneme level are given as a percentage of the number of phonemes presented Figs. 1–6. We present a conditionalised measure of order memory in the figures and analyses. As the rate of order errors was dependent on total recall (items could not be recalled in the wrong position if they were not recalled at all), order errors at the level of items are expressed as a percentage of total item recall. Similarly, phoneme migration errors are shown as a percentage of total phoneme recall.

Lexicality

Figs. 1A and B show whole-item order and identity errors for words and nonwords. The data from mixed lists of words and nonwords (Experiment 1) are shown in conjunction with the data examining pure lists (Experiment 2), although only the Experiment 1 data are discussed here. Item errors were common throughout the experiment whereas whole-item transpositions occurred infrequently because of the large number of phonological errors (accounting for the very different scales of the Figs. 1A and B). Lexicality had opposite effects on identity and order errors at the level of whole items. There was a highly significant effect of lexicality on the number of item identity errors ($t(29) = 17.34$, $p < .0001$); these errors occurred much more frequently for nonwords than for words. In contrast, whole-item transpositions occurred more commonly for words than for nonwords, as the phonological elements of nonwords were unlikely to be recalled together ($t(29) = 3.64$, $p = .001$).

Figs. 2A and B show order and identity errors at the level of individual phonemes for words and nonwords. Identity errors at the phoneme level, like those at the whole-item level, occurred less frequently for words than for nonwords ($t(29) = 9.99$, $p < .0001$). Phoneme order errors also occurred less frequently for words than for

Table 1

Percentage of words and nonwords recalled in mixed lists (Experiment 1) as a function of frequency, imageability, and the ratio of words to nonwords

	Words:nonwords in list	Words		Nonwords	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
High frequency, high imageability	3:2	77.8	12.7	48.0	19.9
	2:3	66.7	18.1	51.1	22.0
	1:4	66.7	23.7	45.5	20.9
	Mean	70.4	14.5	48.2	17.9
High frequency, low imageability	3:2	75.1	14.2	48.7	24.5
	2:3	57.3	19.5	36.7	17.4
	1:4	58.0	20.6	40.7	16.6
	Mean	63.5	14.3	42.0	16.4
Low frequency, high imageability	3:2	55.3	17.7	31.7	19.1
	2:3	58.3	16.4	38.2	18.5
	1:4	66.7	25.4	39.7	19.0
	Mean	60.1	15.0	36.5	14.8
Low frequency, low imageability	3:2	58.2	19.1	31.3	18.7
	2:3	41.3	19.3	31.3	16.0
	1:4	42.0	20.6	31.2	15.6
	Mean	47.2	15.3	31.3	14.2
Mean	3:2	66.6	11.9	39.9	15.9
	2:3	55.9	13.4	39.3	15.4
	1:4	58.3	16.0	39.3	15.4
	Mean	60.3	12.7	39.5	14.7

Note. Figures show percentage of items recalled in correct serial position.

Table 2

Analysis of recall accuracy in Experiment 1

Effect	By subjects	By items	Explanatory notes
Lexicality	$F(1, 29) = 244.1, p < .0001$	$F(1, 276) = 64.6, p < .0001$	Words > nonwords
Frequency	$F(1, 29) = 125.3, p < .0001$	$F(1, 276) = 22.4, p < .0001$	High freq > low freq
Imageability	$F(1, 29) = 43.1, p < .0001$	$F(1, 276) = 9.1, p < .01$	High imag > low imag
No. words in list	$F(1, 29) = 10.6, p < .001$	$F(2, 276) = 2.0, p = .14$	3 words > 2 words = 1 word
Lex × Freq	$F(1, 29) < 1$	$F(1, 276) < 1$	Frequency affects recall of both words and nonwords
Lex × Imag	$F(1, 29) = 7.9, p < .001$	$F(1, 276) < 1$	Imag effect larger for words than nonwords (significant for both)
Lex × no. words	$F(2, 58) = 11.5, p < .0001$	$F(2, 276) = 1.6, p = .21$	More words in list → words recalled more accurately
Freq × no. words	$F(2, 58) = 6.6, p < .01$	$F(2, 276) = 1.5, p = .23$	More words in list → larger frequency effect
Imag × no. words	$F(2, 58) = 11.7, p < .0001$	$F(2, 276) = 2.6, p = .08$	More nonwords in list → larger imageability effect
4-way	$F(2, 58) = 3.3, p < .05$	$F(2, 276) < 1$	

All significant effects are reported above, alongside non-significant effects of particular interest. The explanatory notes are based on post hoc tests by subjects that are not reported in full here. This analysis considered the number of items recalled in the correct serial position. Abbreviations used: Lex, lexicality; Freq, frequency; Imag, imageability.

nonwords ($t(29) = 7.50, p < .0001$), in contrast with order errors at the whole-item level. Nonwords were therefore more likely than words to fragment in verbal STM.

Frequency

Frequency affected the occurrence of order and identity errors in a similar way to lexicality. Identity errors at the whole-item level were more common for low

compared with high frequency words (high frequency words, mean item identity errors = 27.1% of items presented, $SD = 10.7$; low frequency = 42.9%, $SD = 12.9$; $t(29) = 10.11, p < .0001$). In contrast, there was no significant difference in the rate of item order errors between high and low frequency words (high frequency, mean = 6.4% of items presented, $SD = 5.5$; low frequency mean = 6.2%, $SD = 5.5$; $t(29) < 1$).

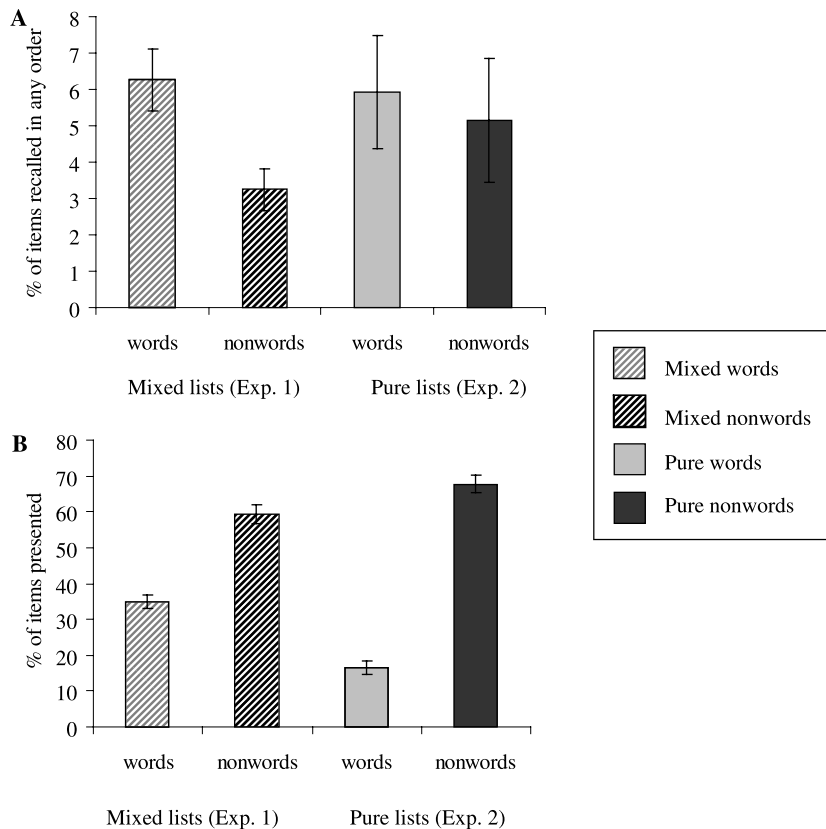


Fig. 1. Order and identity errors at the level of whole items, as a function of lexicality (Experiments 1 and 2). (A) Item order errors. (B) Item identity errors. Error bars show standard error of mean.

Figs. 3A and B show order and identity errors at the level of individual phonemes for high and low frequency words and the nonwords they were mixed with. Identity errors at the phoneme level, like those at the whole-item level, occurred less frequently for high than low frequency words ($t(29) = 7.63, p < .0001$). Phoneme order errors also occurred less frequently for high than low frequency words ($t(29) = 3.43, p = .002$), in contrast with order errors at the whole-item level. As the larger number of phoneme order errors for low frequency words could not be accounted for by whole-item migrations, this finding suggests that frequency affected the extent to which the phonemes of words were recalled together as a coherent item.

Word frequency affected the occurrence of phoneme order and identity errors for nonwords as well as for words in mixed lists (see Figs. 3A and B). Phoneme order errors were less common for nonwords that had been presented with high compared with low frequency words ($t(29) = 5.22, p < .0001$), as were phoneme identity errors ($t(29) = 6.10, p < .0001$). These results are consistent with the suggestion that the phonemes of high frequency words were more likely to remain together

in verbal STM, reducing the opportunity for nonword phonemes to migrate in mixed lists. Word frequency appeared to impinge on the stability of the whole phonological trace and not just on the phonological representation of the word items.

Imageability

Imageability had a rather smaller effect on recall overall. Nevertheless, several of the word frequency findings were reproduced for imageability. Figs. 4A and B show identity and order errors at the level of individual phonemes for high and low imageability words and the nonwords they were mixed with. Phoneme identity errors were less common for high than low imageability words ($t(29) = 5.52, p < .0001$), although imageability did not affect the number of phoneme migration errors for words ($t(29) < 1$). Again, the imageability of the words impinged on the stability of the phonological trace for nonwords. Phoneme order errors were less common for nonwords presented with high compared with low imageability words ($t(29) = 3.03, p = .005$), as were phoneme identity errors ($t(29) = 5.18, p < .0001$).

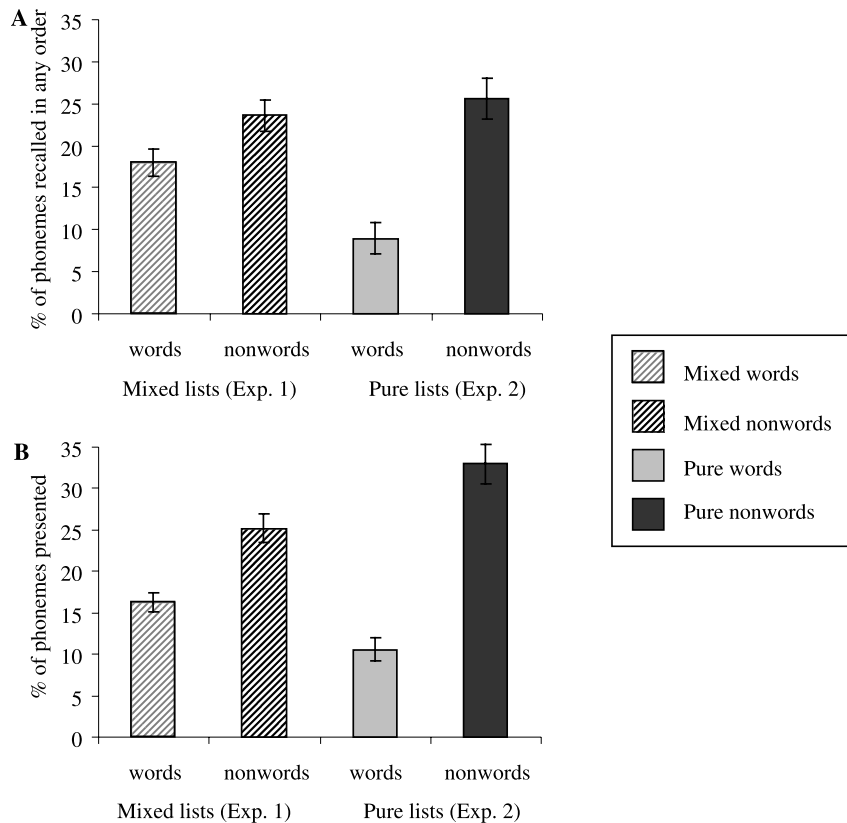


Fig. 2. Order and identity errors at the level of individual phonemes, as a function of lexicality (Experiments 1 and 2). (A) Phoneme order errors. (B) Phoneme identity errors. Error bars show standard error of mean.

Ratio of words to nonwords

The effect of the ratio of words to nonwords on the number of order and identity errors was also examined at the level of individual phonemes (see Figs. 5A and B). There were fewer phoneme migration errors when the ratio of words to nonwords was higher, and this effect occurred for both words (1 vs. 2 words in list; $t(29) < 1$; 2 vs. 3 words in list; $t(29) = 3.96$, $p = .0004$) and nonwords (1 vs. 2 words in list; $t(29) = 1.03$, ns ; 2 vs. 3 words in list; $t(29) = 2.56$, $p = .02$). In contrast, the ratio of words to nonwords in the lists only influenced the number of phoneme identity errors for words (1 vs. 2 words in list; $t(29) < 1$; 2 vs. 3 words; $t(29) = 3.49$, $p < .01$) and had no effect on nonwords (1 vs. 2 words in list; $t(29) < 1$; 2 vs. 3 words; $t(29) = 1.81$, ns). These results are again consistent with the suggestion that word phonemes are bound together more than nonword phonemes in verbal STM, reducing the opportunity for nonword phonemes to migrate in lists containing a higher proportion of words.

Phoneme type

Analyses examining the number of order and identity errors for C1, V, and C2 phonemes as a function of

lexicality were also conducted (see Figs. 6A and B). There was a main effect of phoneme type for both order errors ($F(2, 58) = 101.26$, $p < .0001$) and identity errors ($F(2, 58) = 45.41$, $p < .0001$), indicating that vowels were recalled more accurately than consonants. For phoneme migration errors, there was no interaction between phoneme type and lexicality ($F(2, 58) = 2.18$, ns), showing that the order memory advantage for words over nonwords was equal for vowels and consonants (collapsing across lexicality, V vs. C1—Bonferroni $t(29) = 10.94$, $p < .0001$; V vs. C2—Bonferroni $t(29) = 12.18$, $p < .0001$). In contrast, there was an interaction between phoneme type and lexicality for phoneme identity errors ($F(2, 58) = 31.08$, $p < .0001$). For phoneme identity errors, the recall advantage for words over nonwords was greater for consonants than vowels. Bonferroni t tests revealed that for nonwords, fewer identity errors involved V than C1 ($t(29) = 3.06$, $p < .01$) or C2 phonemes ($t(29) = 8.89$, $p < .0001$). For words, there was no difference between V and C1 phonemes ($t(29) = 1.09$, ns) and a less marked difference between V and C2 phonemes ($t(29) = 3.86$, $p < .01$). The acoustic energy of the vowels may have assisted their identification and maintenance, minimising the number of vowel identity errors. This

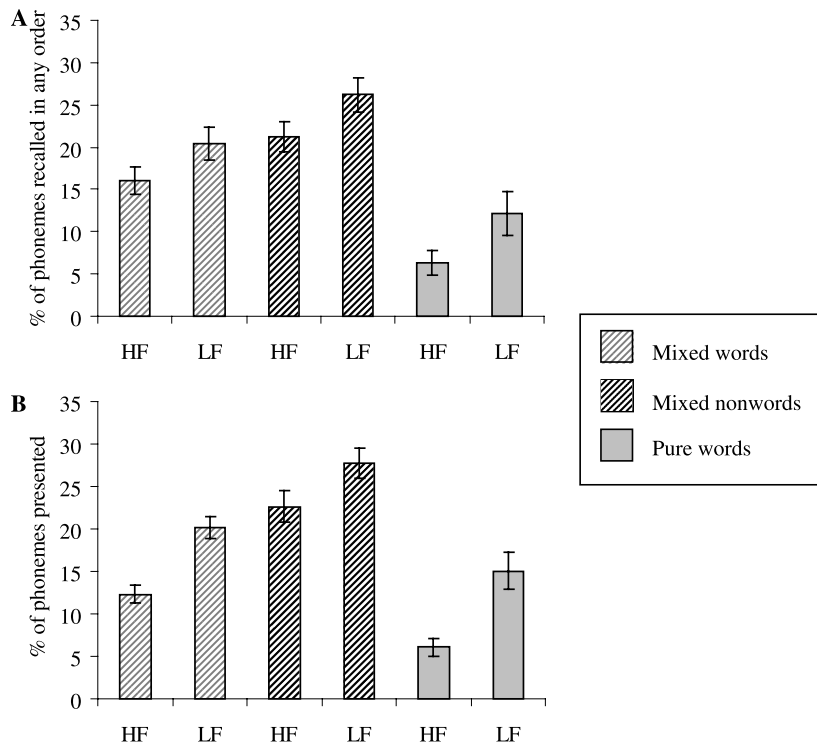


Fig. 3. Order and identity errors at the level of individual phonemes, as a function of word frequency (Experiments 1 and 2). (A) Phoneme order errors. (B) Phoneme identity errors. HF, high frequency; LF, low frequency. Error bars show standard error of mean.

effect would have been less critical for words than for nonwords however, since identification and memory of each word phoneme would have been supported by knowledge about the whole word.

Discussion

When participants attempted to recall lists composed of an unpredictable mixture of words and nonwords, a considerable number of phonological errors occurred for words as well as nonwords, mimicking immediate serial recall in semantic dementia. This was possibly because the nonword phonemes, which were not tightly bound together, were able to recombine with the elements of words, reducing their coherence. Some specific support for this suggestion was obtained in analyses that examined the impact of manipulating the ratio of words to nonwords in the lists. Word recall declined as the number of nonwords was increased because word phonemes were more likely to migrate between list items or be recalled incorrectly.

Phoneme migrations were influenced by lexical and semantic factors, in line with the predictions of the semantic binding theory. Word phonemes were less likely to migrate to new list positions than the elements of nonwords. Similarly, fewer phoneme migration errors

occurred for high compared with low frequency words.¹ In contrast, *whole-item* transpositions were more common for words than for nonwords (in line with the findings of Saint-Aubin & Poirier, 2000). Therefore, the phonemes of nonwords were more likely to fragment and migrate individually. Word phonemes, on the other hand, were more tightly bound together and often migrated as a coherent item. In addition, there were effects of lexicality, frequency and imageability on phoneme identity errors, indicating that lexical/semantic knowledge constrained phoneme identity as well as phoneme order.

Nonword phonemes were less likely to migrate in mixed lists containing a larger number of words and more frequent/imageable words. The phonemes of these words were apparently more tightly bound together, reducing the opportunity for nonword phonemes to migrate. These findings are inconsistent with the hypothesis that redintegration operates in an item-specific fashion (e.g.,

¹ Although imageability did not have a significant effect on migration errors for word phonemes, this variable did impinge on phoneme migrations for nonwords, suggesting that lexicality, frequency, and imageability all had an effect on phonological binding.

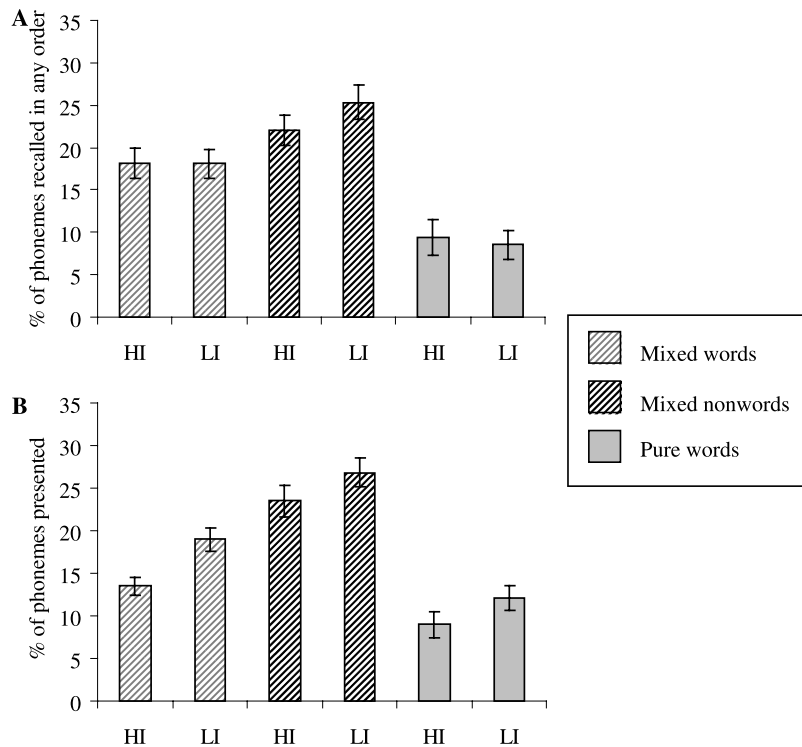


Fig. 4. Order and identity errors at the level of individual phonemes, as a function of imageability (Experiments 1 and 2). (A) Phoneme order errors. (B) Phoneme identity errors. HI, high imageability; LI, low imageability. Error bars show standard error of mean.

Schweickert, 1993). Our results are also not fully accommodated by recent modifications of the reintegration theory in which the availability of long-term lexical representations is influenced by the strength of associations between list items (Hulme et al., 2003; Stuart & Hulme, 2000). Although this viewpoint predicts that list composition can influence recall, the strength of associative links between nonwords and high frequency/imageability words might not be expected to differ substantially from those between nonwords and low frequency/imageability words (both would be close to zero, reflecting the absence of associative relationships between words and nonwords; see Hulme et al., 2003).

Experiment 2: Pure lists of words and nonwords

Participants were tested on pure lists of words and nonwords derived from the mixed lists described above, allowing a between-subjects investigation of the impact of mixing words with nonwords. The semantic binding account anticipates that both words and nonwords will be recalled at an intermediate level in mixed lists: nonword phonemes will recombine with the elements of words, impairing their recall, but the relative phonological stability of the words will also reduce nonword phoneme migrations.

Method

Participants

The participants were 20 undergraduates, aged between 18 and 23, who spoke English as a first language and had normal hearing. They did not take part in Experiment 1. They were tested individually and took part for course credit.

Design and materials

As in Experiment 1, participants were presented with lists of five CVC stimuli. Lexicality, frequency, and imageability were retained as within-subjects factors. However, each list was composed purely of words or nonwords, rather than a mixture of the two. The pure word lists were constructed by replacing the nonwords in the mixed lists with real words. Similarly, the pure nonword lists were constructed by replacing the words in the mixed lists with nonwords. Therefore, the participants were tested on the same stimuli as in Experiment 1, in the same serial positions, but in the context of pure rather than mixed lists. The new words, used to replace the nonwords in the mixed lists, had similar frequency counts and imageability ratings as the original Experiment 1 words (see Appendix B). The new nonwords were constructed from the words they replaced by changing the final consonants (e.g., gun to 'gudge'). Whenever possible, consonants were ex-

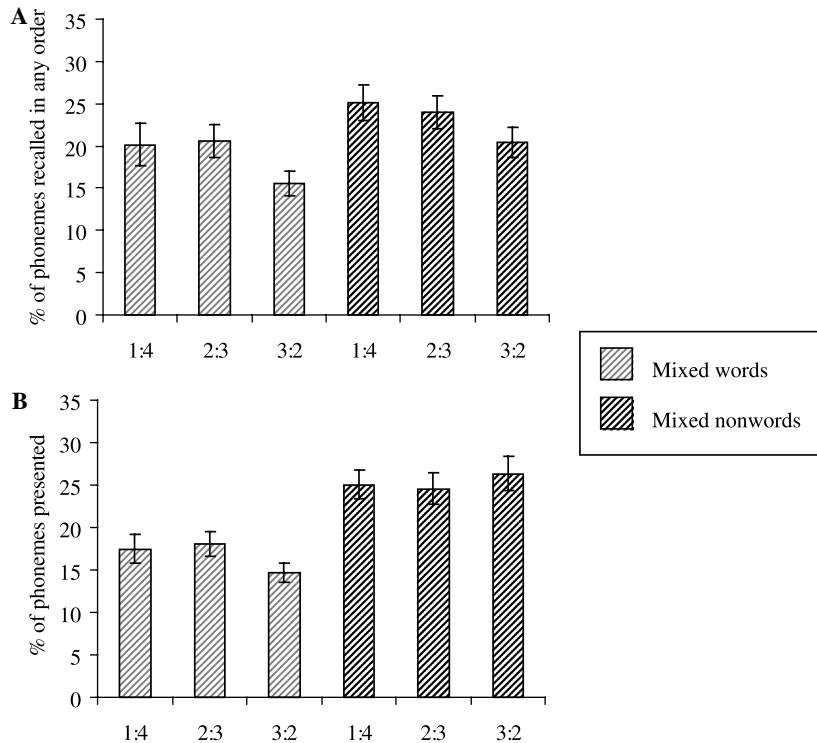


Fig. 5. Order and identity errors at the level of individual phonemes, as a function of the ratio of words to nonwords in mixed lists (Experiment 1). (A) Phoneme order errors. (B) Phoneme identity errors. 1:4 denotes list containing one word and four nonwords. 2:3 denotes two words and three nonwords. 3:2 denotes three words and two nonwords. Error bars show standard error of mean.

changed between items in order to construct the nonwords, although this did not prove feasible for a few common final consonants, given the requirement for every phoneme to be different within a list. All the nonwords were legal and pronounceable. Items were not repeated in the course of the experiment.

There were 60 nonword trials, corresponding to the complete set of lists used in Experiment 1. There were also 40 word trials, corresponding to the lists that contained two and three words in Experiment 1. In the word condition, participants were not tested on lists that originally contained four nonwords and one word because the overlap in items between the two experiments would have been low.

Procedure

A female speaker recorded the new words and nonwords individually in a flat intonation. Sound editing software (Cool Edit, Syntrillium) was used to replace either the words or nonwords from the original lists with these new items, in order to produce pure lists of words and nonwords. Presentation was at a rate of one item per second. The pure word and nonword lists were presented in separate blocks and participants were told in advance whether the block would contain words or nonwords. There were four practice trials at the start

of each block. The order of the blocks was counterbalanced across participants. The order of trials was the same as for the first 20 participants in Experiment 1. Other methodological details were as described for Experiment 1.

Results

As in Experiment 1, omissions were positioned in transcripts to minimise the number of errors. When participants recalled six rather than five items (.01% of trials for both words and nonwords), the final item was discarded. All of the analyses below include only those items that were presented in both experiments.² Table 3 shows the mean percentage of words and nonwords recalled in the correct serial position in pure and mixed lists. Table 4 shows an analysis of recall accuracy for mixed and pure lists by subjects and by items. In summary, we found that recall in position was better for nonwords and poorer for words in mixed relative to pure lists. The effects of frequency and imageability, however, did not interact with list composition.

² Additional analyses that included all of the items presented in the pure lists produced very similar outcomes.

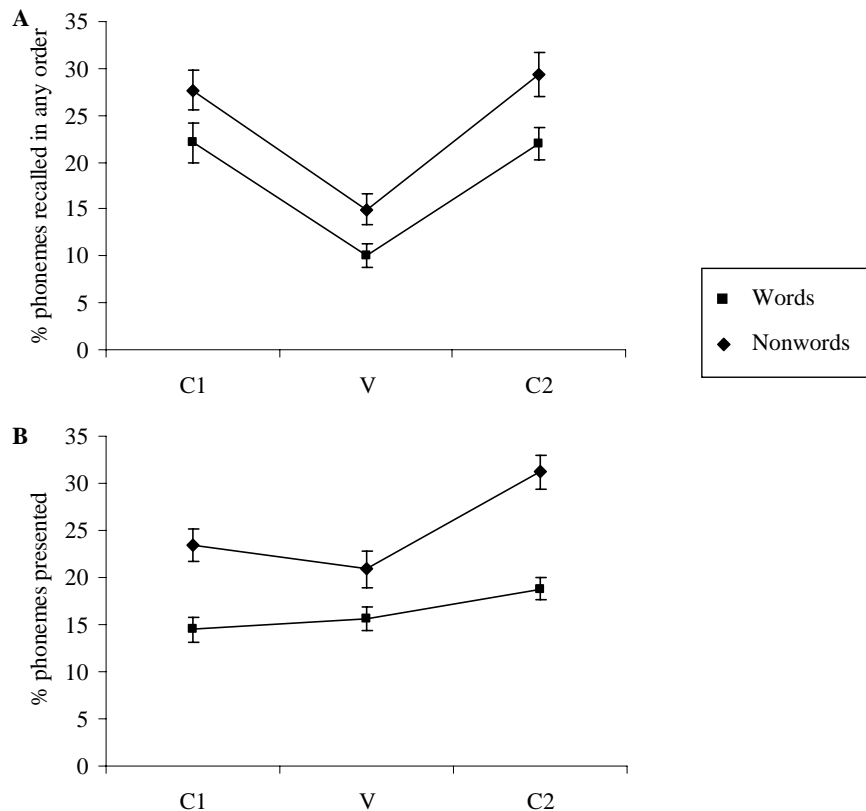


Fig. 6. Order and identity errors at the level of individual phonemes, as a function of phoneme type and lexicity in mixed lists (Experiment 1). (A) Phoneme order errors. (B) Phoneme identity errors. C1, initial consonant; V, vowel; C2, final consonant. Error bars show standard error of mean.

The influence of lexicity, frequency, and imageability on the number of order and identity errors in pure lists was examined using a series of *t* tests, as in Experiment 1. Order errors were again conditiona-

lised on the number of items recalled in any position. In addition, the effect of mixing words with nonwords on these error types was investigated for each factor using a series of repeated-measures ANOVAs, in which list type (i.e., mixed lists from Experiment 1 vs. pure lists from Experiment 2) was entered as a between-subjects variable.

Table 3

Percentage of words and nonwords recalled in pure and mixed lists

	Pure lists		Mixed lists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Words</i>				
High frequency, high imageability	87.8	11.5	73.3	11.3
High frequency, low imageability	85.8	8.7	68.0	13.9
Low frequency, high imageability	74.0	17.0	56.5	14.7
Low frequency, low imageability	68.6	18.9	51.5	17.2
Mean	79.1	11.6	62.3	11.8
<i>Nonwords</i>				
	31.1	11.5	39.4	14.8

Note. Figures show percentage of items recalled in correct serial position. Only items tested in both mixed and pure list conditions were included.

Lexicality

At the whole-item level, there was a main effect of lexicity across the two experiments for order errors: words were more likely to be recalled in the wrong serial position than nonwords ($F(1, 48) = 5.92, p = .02$) and this effect did not interact with list type ($F(1, 48) = 2.05, ns$). Item identity errors, on the other hand, showed a highly significant lexicity by list type interaction ($F(1, 48) = 1117.83, p < .0001$). Bonferroni *t* tests indicated that item identity errors were *more* frequent for words in mixed compared with pure lists ($t(48) = 10.44, p < .0001$) and *less* frequent for nonwords in mixed compared with pure lists ($t(48) = 11.97, p < .0001$). Consequently, the effect of lexicity on item identity errors was larger for the pure than mixed lists.

Table 4
Analysis of recall accuracy for pure lists (Experiment 2) vs. mixed lists (Experiment 1)

Effect	By subjects	By items	Explanatory notes
Lexicality	$F(1, 48) = 970.7, p < .0001$	$F(1, 128) = 259.3, p < .0001$	Words > nonwords
Lexicality × list type	$F(1, 48) = 121.2, p < .0001$	$F(1, 128) = 132.2, p < .0001$	Word recall: pure > mixed lists; nonwords: mixed > pure lists
Frequency	$F(1, 48) = 91.1, p < .0001$	$F(1, 96) = 29.0, p < .0001$	High freq > low freq
Freq × list type	$F(1, 48) < 1$	$F(1, 96) < 1$	Equivalent frequency effect for pure and mixed lists
Imageability	$F(1, 48) = 14.37, p < .001$	$F(1, 96) = 2.22, p = .14$	High imag > low imag
Imag × list type	$F(1, 48) < 1$	$F(1, 96) < 1$	Equivalent imageability effect for pure and mixed lists

All significant effects are reported above, alongside non-significant effects of particular interest. The explanatory notes are based on post hoc tests by subjects not reported in full here. This analysis considered the number of items recalled in the correct serial position. Only items presented in both the mixed and pure list experiments were included, although a separate analysis that examined the recall of the complete set of items yielded a very similar pattern of results. *Abbreviations:* Lex, lexicality; Freq, frequency; Imag, imageability.

As in Experiment 1, nonword phonemes were more likely than word phonemes to migrate to the wrong serial position in pure lists (see Figs. 2A and B). The substantial effect of lexicality on phoneme order errors ($F(1, 48) = 191.31, p < .0001$) was larger for pure compared with mixed lists ($F(1, 48) = 46.51, p < .0001$), primarily because the number of phoneme migration errors for words was greater (Bonferroni $t(48) = 3.57, p = .002$). In contrast, list type did not affect the number of phoneme migration errors for nonwords (Bonferroni $t(48) < 1$). There were also fewer phoneme identity errors for words than nonwords ($F(1, 48) = 357.14, p < .0001$) and this effect interacted with list type ($F(1, 48) = 66.74, p < .0001$). Bonferroni t tests indicated that the number of phoneme identity errors increased for words when they were mixed with nonwords ($t(48) = 3.18, p < .01$) and decreased for nonwords when they were mixed with words ($t(48) = 2.73, p < .01$).

Frequency and Imageability

The frequency of the words in the pure lists influenced the numbers of both phoneme order errors ($t(19) = 3.19, p = .005$) and phoneme identity errors ($t(19) = 5.03, p < .0001$), providing a replication of the results obtained with mixed lists (see Figs. 3A and B). There was no interaction between frequency and list type for either phoneme order or identity errors (both $F(1, 48) < 1$). In addition, the imageability of the words in the pure lists influenced the numbers of phoneme identity errors ($t(19) = 3.34, p < .01$) but not phoneme order errors ($t(19) < 1$), again providing a replication of the results obtained with mixed lists (see Figs. 4A and B). There was no interaction between imageability and list type for either phoneme order errors ($F(1, 48) < 1$) or phoneme identity errors ($F(1, 48) = 2.56, ns$).

Lexicality of errors in pure and mixed lists

We examined the proportion of errors that were real word responses for words and nonwords in mixed and pure lists, in order to determine if participants were sensitive to the lexical status of target items. This analysis was based on 6525 responses. There was no difference in the lexical status of errors to words and nonwords in mixed lists (58% and 60% words respectively; $\chi^2(1) = 1.44, ns$). In contrast, for the pure lists, there was a significant effect of lexicality on the proportion of errors that were real words (91% vs. 52% for words and nonwords, respectively; $\chi^2(1) = 162.99, p < .0001$), suggesting that participants were able to use knowledge of the lexical status of the targets to constrain their responses in pure but not mixed lists. Similarly, the ratio of words to nonwords in the mixed lists had no effect on the lexical status of errors for either words or nonwords (words—60, 52, and 58% for the ratios 1:4, 2:3, and 3:2; $\chi^2(2) < 1$; nonwords—60, 58, and 62%; $\chi^2(2) = 2.57, ns$). The nature of the word items in mixed lists did influence the lexicality of errors, however. Both word and nonword items in lists containing high imageability words were more likely to be recalled as words (words—61% vs. 55%; $\chi^2(1) = 3.46, p = .06$; nonwords—64% vs. 56%; $\chi^2(1) = 16.43, p < .0001$). Similarly, words and nonwords mixed with high frequency words were more likely to be recalled as real words (words—64% vs. 53%, $\chi^2(1) = 13.83, p < .001$; nonwords—62% vs. 58%; $\chi^2(1) = 3.06, p = .08$).

Discussion

Participants recalled pure lists of words and nonwords providing a comparison with the mixed lists examined in Experiment 1. Recall of both words and nonwords was intermediate in mixed lists: the nonwords harmed word recall whereas the words boosted nonword

recall. Likewise, in Experiment 1, words and nonwords were recalled more accurately in lists containing a higher ratio of words to nonwords. These results together indicate that the degree of ‘lexical/semantic binding’ for one list item impinges on the coherence of the entire phonological trace. Word phonemes that are tightly bound together reduce the opportunity for nonword phonemes to migrate and similarly, nonword phonemes act as “free radicals,” recombining with the elements of words and reducing their phonological stability. However, one potential caveat should be noted. Mixing words with nonwords (Experiment 1 vs. 2) increased the number of phoneme migrations for words but primarily improved nonword recall by decreasing phoneme *identity* errors: in contrast, the semantic binding hypothesis specifically predicts a change in phoneme order errors (indeed, the ratio of words to nonwords in Experiment 1 did affect the number of nonword phoneme migrations). Of course, the different error types are not independent of each other; list type may have had no effect on order errors for nonwords because identity errors (98% of which were omissions) prevented them from occurring.

The comparison between mixed and pure lists also highlights the importance of strategic factors in binding/redintegration. Although these processes are usually thought to be automatic (e.g., Hulme et al., 1991), strategic factors might also operate. In pure word lists, the participant (or the system) may use knowledge of lexical status to appropriately constrain output. In mixed lists, however, the items are far less predictable as words or nonwords: in our study, participants’ errors to words were predominantly other words for pure but not mixed lists, suggesting that knowledge of lexical status was used strategically to constrain output for the pure word lists only. This might account for discrepancies between our findings and those of a previous study. Hulme et al. (2003) compared recall of pure and alternating lists of words/nonwords. The alternating lists had a predictable structure, in contrast to our mixed lists that varied the serial position and the number of words/nonwords across trials. The recall of nonwords in Hulme et al.’s alternating lists was boosted by the presence of words (in line with our findings), although the recall of words was *not* substantially affected by the presence of nonwords. The poorer recall of words in mixed relative to pure lists that we observed could therefore have arisen due to a loss of knowledge about lexical status (as opposed to an impact of unbound nonword phonemes on the stability of the phonological trace for words). These strategic influences on redintegration are unlikely to be the whole story however. For the unpredictable mixed lists, participants apparently lacked knowledge about which items were words/nonwords (from their errors), suggesting that word recall was poorer when the number of nonwords was higher because the nonword phonemes directly interfered with the phonological

integrity of the words. In addition, for lists containing an unpredictable mixture of words and nonwords, it is unclear how an item-specific redintegration mechanism could discriminate between the degraded traces of words, which need to be lexically reconstructed, and the phonological traces of nonwords, which do not. Consequently, this theory appears to have some difficulty explaining how effects of lexicality persist in unpredictable mixed lists.

General discussion

Two experiments explored the influence of lexical and semantic factors on the integrity of representations in phonological STM. In the first experiment, participants recalled lists composed of an unpredictable mixture of words and nonwords. This methodology made it possible to study the effect of lexical/semantic factors on the coherence of the phonological trace for both word and nonword items. In the second experiment, the mixed lists were compared with pure lists of words and nonwords. There were three key findings across the experiments, which are summarised below:

- (1) Lexical/semantic knowledge influenced the rate of both phoneme identity and phoneme order errors. Word phonemes were more likely to be recalled in the correct configuration, whereas nonword elements were more likely to migrate to the wrong list position or be forgotten altogether. In line with previous studies, lexical/semantic knowledge chiefly reduced item identity as opposed to order errors (Gathercole et al., 2001; Hulme et al., 1997; Poirier & Saint-Aubin, 1995, 1996; Saint-Aubin & Poirier, 1999, 2000). It was unusual for a whole item to migrate to the wrong serial position and when this did occur, it was more likely to be a word than a nonword because the word phonemes were less likely to recombine with the elements of other items (see Saint-Aubin & Poirier, 2000; for a similar finding).
- (2) Phonological errors were more frequent for words when they were presented in mixed lists with nonwords, relative to pure word lists (Experiment 2). Similarly, phonological coherence for individual words in mixed lists was influenced by the ratio of words to nonwords in the list (Experiment 1). The nonwords appeared to jeopardise the phonological stability of the words by recombining with their constituent phonemes.
- (3) Frequency, imageability, and the ratio of words to nonwords affected the recall of nonwords as well as words in mixed lists (Experiment 1). Similarly, nonwords were recalled more accurately when they were mixed with words com-

pared with their recall in pure nonword lists (Experiment 2). The nonword phonemes may have had more limited opportunities to migrate when the phonological coherence of the other list items was greater.

The observation that lexical/semantic factors affect the rate of phoneme migration errors provides explicit support for the semantic binding hypothesis (Patterson et al., 1994). According to this viewpoint, two sources of constraint encourage the phonemes of a particular item to emerge together as a coherent item. First, the phonemes of familiar words become associated in the phonological system by virtue of the fact that they are produced and heard together; this facilitates their binding into coherent items. In addition, the semantic system can constrain the clustering of phonemes into meaningful items because the semantic and phonological patterns corresponding to particular words become associated during speech production/comprehension. In immediate serial recall tasks, in which participants attempt to maintain the phonology of an entire sequence of items, these constraints reduce opportunities for phonemes from a particular item to recombine with those of other list items. In fact, the likelihood of a phoneme transposition error for a given item is influenced by both the strength of lexical/semantic binding between the phonemes of that item and also the strength of binding for other list items, which affects the availability of unconstrained phonemes to intrude. This is consistent with our result that lexical/semantic knowledge pertaining to one list item had a knock-on effect on the phonological coherence of the rest of the list. These pattern completion properties for familiar words reduce phoneme identity as well as migration errors (in line with our data), because the elements of familiar words are strongly associated and bolster each other's activation levels.

The standard redintegration viewpoint (e.g., Hulme et al., 1991; Schweickert, 1993), in contrast, does not offer a specific explanation of the impact of lexical/semantic factors on phoneme order errors. Redintegration is purported to restore parts of words that have been maintained incorrectly, reducing the number of phoneme identity not order errors. In pure word lists, missing or incorrect phonemes can be reconstructed, explaining why few phonological errors are observed. In mixed lists, however, this process is not expected to correct the identity of word phonemes that intrude into nonwords, and yet clear effects of lexical/semantic variables on phoneme migrations can be seen. In addition, the original item-specific redintegration theory conflicts with our finding that lexical/semantic knowledge of an item affects the phonological coherence of the other items in the list. Other versions of the redintegration hypothesis which anticipate effects of list

composition on recall are able to accommodate some but not all of our findings. For example, Stuart and Hulme (2000) suggested that inter-item associations can influence the availability of lexical representations in redintegration: although associative links between words might be disrupted in mixed lists leading to intermediate levels of recall, this account does not predict an effect of word frequency/imageability on nonword recall because nonwords are expected to be minimally associated with all types of words. In addition, the feature model of Nairne (1990) suggests that redintegration is affected by the distinctiveness of each item relative to the retrieval set, which is typically composed of the other list items. Words and nonwords in mixed lists might be more distinctive than the same items in pure lists, providing an explanation of the superior recall of nonwords in mixed lists. However, it is more difficult to account for the *poorer* recall of words in mixed lists by this view.

Imageability had a weaker impact on phoneme migration errors than word frequency. In both experiments, the phonemes of high/low imageability words were equally likely to migrate, contrary to the predictions of the semantic binding hypothesis (however, the phonemes of nonwords mixed with imageable words were less likely to be recalled in the wrong position, so imageability did influence the stability of the phonological trace to some degree). Imageability effects were weaker than frequency effects throughout the experiment—this variable may have influenced phoneme identity errors but not phoneme migrations because lexical/semantic factors had a smaller impact on phoneme migrations in general. The fact that imageability had a significant impact on phoneme migrations for *nonwords* in mixed lists makes this explanation of the null result quite likely in our view. We acknowledge an alternative explanation, however: differences in the effects of frequency and imageability might have arisen from the operation of these variables in the theoretical framework that we favour. Patterson et al. (1994) suggested that constraints on phoneme order could be provided by: (1) a tendency for activation within the phonological system to settle on familiar patterns (i.e., lexical phonological knowledge) and (2) the strong associations between specific semantic and phonological patterns pertaining to the same word (i.e., lexical semantic knowledge). Imageability, a semantic variable, should only affect the strength of the second of these constraints. In contrast, word frequency is likely to affect both types of constraint: frequent words will have stronger pattern completion properties at the phonological level and also stronger connections between phonology and semantics. Frequency might therefore have made a greater contribution than imageability to phonological coherence for two reasons: first, because this variable affects both lex-

ical–phonological and lexical–semantic binding, and second, because this variable directly affects the phonological layer.

Taken together, our results support the notion that lexical/semantic factors influence the accuracy of immediate serial recall (at least in part) by altering the efficacy of the pattern completion processes that operate for familiar items. Activation-based accounts like the semantic binding hypothesis (Patterson et al., 1994), which suggest that these processes are integral to the operation of the phonological system underlying verbal STM, and are more consistent with several of our findings than the two-stage degradation plus late-stage reintegration theory. In particular, the semantic binding hypothesis is endorsed by the observed impact of lexical/semantic knowledge on the rate of phoneme migrations. This account is also consistent with the finding that the lexical/semantic characteristics of particular list items affected the stability of the entire phonological trace. In contrast, the standard reintegration viewpoint does not provide an account of the role of lexical/semantic knowledge in phoneme recombinations, and list composition effects are inconsistent with the concept of item-specific reintegration.

With modification, the reintegration theory could also accommodate lexical/semantic binding effects. For example, phonological traces could be reconstructed at output by assigning phonemes held in a temporary store to positions in the speech sequence. The selection of phonemes would be constrained by lexical/semantic knowledge so that elements would be more likely to be produced together if they formed a meaningful item. If, through a failure of lexical/semantic binding, a phoneme from one item migrates into another, this would have consequences for the stability of other list items—for example, target phonemes that were not produced might invade neighbouring items. This proposal departs from the traditional view that specific items are reconstructed from long-term lexical representations, and instead suggests there are multiple constraints on phoneme selection in immediate serial recall, some of which result from lexical/semantic binding. The issue of whether this binding only operates during recall, as envisaged by the reintegration account, or throughout verbal STM tasks is beyond the scope of this paper; however, there are marked parallels between immediate serial recall and matching span, a verbal STM paradigm that does not require overt recall, which suggest that lexical/semantic knowledge contributes to phonological stability *prior* to recall (Jefferies, Frankish, & Lambon Ralph, *in press*). In matching span, lexical/semantic knowledge aids the detection of phoneme transpositions, just as it prevented phoneme migrations during immediate serial recall in the current study. In addition, the detection of word phoneme transpositions in matching span was

impaired by presenting the words in unpredictable mixed lists—similarly, the current study found that word phonemes migrated more frequently in mixed lists relative to pure word lists. We therefore favour the view that lexical/semantic binding effects are an integral part of short-term maintenance within the language system.

Acknowledgments

We thank Karalyn Patterson for suggesting the mixed list methodology as a means of observing more frequent phoneme migration errors on word stimuli in healthy participants. Raymond Knott, Stephen Monsell, and Karalyn Patterson carried out the unpublished work that originally used this method. The first author was supported by an ESRC studentship and a grant from the NIMH (MH64445).

Appendix A. Stimuli used in Experiment 1

A.1. Words: High frequency, high imageability

ball, bed, board, boat, book, face, fish, foot, girl, gun, head, hill, home, horse, house, king, leg, male, men, neck, night, park, phone, rain, rock, seat, teeth, wall, wife, wine

A.2. Words: High frequency, low imageability

base, date, death, form, god, half, hope, jack, job, lead, line, loss, mass, name, part, race, role, rule, shape, sharp, shock, side, size, term, thing, thought, till, top, voice, week

A.3. Words: Low frequency, high imageability

bat, boot, cane, cart, coin, dime, dove, fan, foam, fog, geese, harp, hawk, hedge, hen, hoof, jeep, kite, limb, mouse, noose, rat, rib, surf, thorn, toad, web, weed, wig, yacht

A.4. Words: Low frequency, low imageability

bang, bet, bid, curse, dip, foul, germ, hurt, hush, jade, kale, knoll, lean, lodge, mall, mash, psalm, rack, raid, rhyme, sage, sap, thud, ton, verb, vice, wharf, whiff, wrath, zone

A.5. Nonwords

baf, bal, bam, barl, barss, beel, beng, beuffe, bick, bim, bol, bon, boof, bot, bowne, burge, berl, cowl, cun, dap, deaf, dem, dibe, dit, dop, dorth, fak, fal, feem, feen, fet, fid, fik, fing, foate, fod, fok, fon, forp, garl, gen, gid, girse, gis, goyt, haid, hal, han, harg, heem, heen, hees, heff, hess, het, hin, hoak, hoat, hoess, hol, hom, hon, houne, hoys, hud, hus, jarm, jid, jong, jook, jote, jud, jurn, jurz, kang, keem, keet, kep, kerm, kerze, ket, koese, korp, lan, laysh, leet, lep, lidge, lif, loate, lood, looth, lut, mal, med, mek, min, moess, morke, mort, mot, mun, nate, ned, neek, noid, nood, nooke, nop, paim, pid, poeth, rab,

rad, raig, raim, raish, raowl, rel, roak, roarss, rork, rorl, rorm, rorn, rorsch, rud, ruuge, sawg, saybe, sayde, saysh, seipe, seithe, sek, sharf, sharl, shart, sherb, sherp, siebe, siefe, sisle, sorl, taybe, tayde, tayje, tayne, tayse, tayve, tharj, tharss, thayte, thert, thit, thoape, tice, tiefe, tiege, vayze, vike, vipe, vite, voan, wais, wann, warthe, weem, weis, werp, wid, woam, woan, woash, wol, wole, wote, yourss, zime, zine

Appendix B. Additional stimuli used in Experiment 2

B.1. Words: High frequency, high imageability

book, cash, dark, dog, feet, gun, heart, king, love, mouth, nose, park, pool, red, road, roof, room, ship, shop, song, sun, teeth, wheel, white, wood

B.2. Words: High frequency, low imageability

call, cut, feel, fell, fine, firm, lead, long, lord, mean, miss, move, part, piece, rise, save, south, thick, turn, type, wait, warm, wide, wish, work,

B.3. Words: Low frequency, high imageability

cage, cart, cave, chalk, cheese, duck, geese, gem, gym, heel, juice, lamb, leaf, limb, morgue, nail, noose, peach, pearl, pet, pig, shed, thumb, wig, wool

B.4. Words: Low frequency, low imageability

bait, bang, cheat, chic, curse, dirge, hail, jerk, kale, keel, latch, lodge, loon, meek, nerve, nip, pawn, push, rap, rung, sod, tuck, verb, whack, whoop

Note. Some of these items also appear in Appendix A. The mixed lists that contained a single word were not tested as pure word lists, allowing their items to be reused.

B.5. Nonwords

bav, bayth, baz, bem, besh, buthe, bip, boash, boove, borch, borf, daysss, dess, dieje, dif, dutt, fam, fape, feuke, fiss, foade, foage, fodge, forsh, fout, geeth, girfe, goz, gudge, harn, harss, heg, hep, herch, heth, hidge, hoad, hobe, horg, horp, howke, huth, huthé, jad, jaim, jeese, jerss, jod, karch, kaych, kaysh, kerg, kidge, kienze, koite, leeb, leeth, lesh, liepe, ling, lom, lov, mav, mave, maz, mep, mord, mout, naze, ness, nide, nooth, noz, pard, parsh, rabe, radge, raining, ral, rayfe, riesh, rizz, roaje, rol, rosh, ruebe, saf, sape, sarg, seeve, sieje, sert, sharb, shayse, shoss, sime, teep, thib, thorb, thorm, thush, tiss, tope, tov, tud, turp, verp, vime, voig, weck, weeb, weef, weige, widge, wime, wiv, worg, worg, yod, zope

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database [CD-ROM]*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.
- Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop—A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin Review*, *1*, 122–125.
- Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology*, *26*, 467–482.
- 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.
- Caza, N., Belleville, S., & Gilbert, B. (2002). How loss of meaning with preservation of phonological word form affects immediate serial recall performance: A linguistic account. *Neurocase*, *8*, 255–273.
- Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology A*, *33*, 497–505.
- Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term memory recall of repeated and novel words lists. *Memory and Cognition*, *21*, 539–545.
- Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, *55*, 75–84.
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, *55*, 429–432.
- Dell, G. S., & O'Seaghda, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*, 287–314.
- Forde, E. M. E., & Humphreys, G. W. (2002). The role of semantic knowledge in short-term memory. *Neurocase*, *8*, 13–27.
- Gathercole, S. E., & Martin, A. J. (1996). Interactive processes in phonological memory. In S. E. Gathercole (Ed.), *Models of short-term memory*. Hove, UK: Psychology Press.
- Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology A*, *54*, 1–30.
- Gregg, V. H., Freedman, C. M., & Smith, D. K. (1989). Word frequency, articulatory suppression and memory span. *British Journal of Psychology*, *80*, 363–374.
- 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 in immediate serial recall. *The Quarterly Journal of Experimental Psychology A*, *49*, 80–115.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*, 685–701.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on

- short-term memory tasks: Evidence for a reintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, 23, 1217–1232.
- Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, 49, 500–518.
- Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (in press). Lexical and semantic influences on item and order memory in immediate serial recognition: Evidence from a novel task. *The Quarterly Journal of Experimental Psychology*.
- Jefferies, E., Jones, R., Bateman, D., & Lambon Ralph, M. A. (2004). When does word meaning affect immediate serial recall in semantic dementia? *Cognitive Affective and Behavioral Neuroscience*, 4, 20–42.
- Jefferies, E., Jones, R. W., Bateman, D., & Lambon Ralph, M. A. (2005). A semantic contribution to nonword recall. Evidence for intact phonological processes in semantic dementia. *Cognitive Neuropsychology*, 22, 183–212.
- Knott, R., & Monsell, S. (unpublished manuscript). Words in short-term memory—what preserves their phonological structure?
- Knott, R., Patterson, K., & Hodges, J. R. (1997). Lexical and semantic binding effects in short-term memory: Evidence from semantic dementia. *Cognitive Neuropsychology*, 14, 1165–1216.
- Knott, R., Patterson, K., & Hodges, J. R. (2000). The role of speech production in auditory–verbal short-term memory: Evidence from progressive fluent aphasia. *Neuropsychologia*, 38, 125–142.
- Martin, N., & Saffran, E. M. (1997). Language and auditory–verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, 14, 641–682.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1–86.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory and Cognition*, 16, 343–352.
- Page, M., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105, 761–781.
- Patterson, K., Graham, N., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, 6, 57–69.
- Pickering, S. J., Gathercole, S. E., & Peaker, S. M. (1998). Verbal and visuospatial short-term memory in children: Evidence for common and distinct mechanisms. *Memory and Cognition*, 26, 1117–1130.
- Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology A*, 48, 384–404.
- Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology*, 50, 408–412.
- Roodenrys, S., & Quinlan, P. T. (2000). The effects of stimulus set size and word frequency on verbal serial recall. *Memory*, 8, 71–78.
- Saint-Aubin, J., & Poirier, M. (1999). Semantic similarity and immediate serial recall: Is there a detrimental effect on order information. *The Quarterly Journal of Experimental Psychology A*, 52, 367–394.
- Saint-Aubin, J., & Poirier, M. (2000). Immediate serial recall of words and nonwords: Tests of the retrieval based hypothesis. *Psychonomic Bulletin Review*, 7, 332–340.
- Schweickert, R. (1993). A multinomial processing tree model for degradation and reintegration in immediate recall. *Memory and Cognition*, 21, 168–175.
- Seidenberg, M., & McClelland, J. (1989). A distributed developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Stuart, G., & Hulme, C. (2000). The effects of word co-occurrence on short-term memory: Associative links in long-term memory affect short-term memory performance. *Journal of Experimental Psychology: Learning Memory and Cognition*, 26, 796–802.
- Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning Memory and Cognition*, 14, 145–152.
- Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, 25, 1256–1271.
- Watkins, O. C., & Watkins, M. J. (1977). Serial recall and the modality effect: Effects of word frequency. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 712–718.