

This article was downloaded by: [Fondren Library, Rice University]

On: 24 August 2011, At: 12:53

Publisher: Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office:
Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pqje20>

Concreteness effects in different tasks: Implications for models of short-term memory

Cristina Romani^a, Sheila Mcalpine^b & Randi C. Martin^c

^a University of Aston, Birmingham, UK

^b University of Birmingham, Birmingham, UK

^c Rice University, Houston, TX, USA

Available online: 03 Jan 2008

To cite this article: Cristina Romani, Sheila Mcalpine & Randi C. Martin (2008): Concreteness effects in different tasks: Implications for models of short-term memory, *The Quarterly Journal of Experimental Psychology*, 61:2, 292-323

To link to this article: <http://dx.doi.org/10.1080/17470210601147747>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Concreteness effects in different tasks: Implications for models of short-term memory

Cristina Romani

University of Aston, Birmingham, UK

Sheila McAlpine

University of Birmingham, Birmingham, UK

Randi C. Martin

Rice University, Houston, TX, USA

This study investigates concreteness effects in tasks requiring short-term retention. Concreteness effects were assessed in serial recall, matching span, order reconstruction, and free recall. Each task was carried out both in a control condition and under articulatory suppression. Our results show no dissociation between tasks that do and do not require spoken output. This argues against the redintegration hypothesis according to which lexical-semantic effects in short-term memory arise only at the point of production. In contrast, concreteness effects were modulated by task demands that stressed retention of item versus order information. Concreteness effects were stronger in free recall than in serial recall. Suppression, which weakens phonological representations, enhanced the concreteness effect with item scoring. In a matching task, positive effects of concreteness occurred with open sets but not with closed sets of words. Finally, concreteness effects reversed when the task asked only for recall of word positions (as in the matching task), when phonological representations were weak (because of suppression), and when lexical semantic representations overactivated (because of closed sets). We interpret these results as consistent with a model where phonological representations are crucial for the retention of order, while lexical-semantic representations support maintenance of item identity in both input and output buffers.

Although phonological information plays a prevalent role in short-term memory (STM) for word lists (Baddeley, 1966a, 1966b; Romani, McAlpine, Olson, Tsouknida, & Martin, 2005; Schweickert, Guentert, & Hersberger, 1990), a number of lines of evidence indicate that lexical and semantic variables are also important.

Evidence supporting the role of lexical and semantic factors in short-term retention is reviewed below. Different approaches have been taken to explain empirical findings. In one, lexical/semantic effects are a contribution of long-term memory (LTM) representations to STM recall and for this reason are a function of the opportunity that

Correspondence should be addressed to Cristina Romani, School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, UK. E-mail: C.Romani@Aston.ac.uk

This paper is based on the PhD dissertation of the second author, and it has been partially supported by NIH Grant DC00218 to Prof. Martin. We would like to thank Andrew Olson for his comments on earlier versions of the manuscript.

information has had to be transferred to LTM. In a second approach, lexical/semantic effects are attributed to “redintegration”—that is, to reconstruction of the words in the list, at the time of output, on the basis of the match of the surviving phonological information to long-term memory representations for words. In a final approach, lexical and semantic features of words are activated and stored as the words are processed and retained. Thus, these features can serve to boost the recall of words beyond the level possible with only surviving phonological information. The purpose of the present study is to evaluate these different views of memory by assessing one type of semantic variable—concreteness—in different conditions that all involve short-term recall.

Lexical-semantic effects in STM: Neuropsychological evidence

Patients with semantic dementia show a progressive loss of knowledge of the meanings of words. Several studies have shown that for these patients, short-term memory span is better for words whose meanings are still available (Forde & Humphreys, 2002; Knott, Patterson, & Hodges, 2000; Patterson, Graham, & Hodges, 1994). It has also been shown that aphasic patients may be specifically impaired in the retention of lexical-semantic representations even though semantic knowledge is intact. R. C. Martin and collaborators reported two patients of this kind: AB (R. C. Martin, Shelton, & Yaffee, 1994) and ML (R. C. Martin & He, 2004; R. C. Martin & Lesch, 1996). These patients, differently from patients with a phonological STM impairment, showed normal phonological effects, but no advantage of words over nonwords in span tasks. They also performed better on STM input tasks probing phonological versus semantic information (e.g., better performance on a rhyme probe task than on a semantic probe task). Additionally, AB and ML were impaired in sentence comprehension when the task involved a heavy lexical semantic load (R. C. Martin & He, 2004; R. C. Martin & Romani, 1994), and AB was impaired in learning lists of real words, although he showed no deficit

in learning other kinds of visual or verbal information (e.g., short stories; Romani & Martin, 1999).

Finally, it has also been shown that different cognitive impairments may have different impact on recall at different serial positions. N. Martin and Saffran (1997) have shown that, in a serial recall task, aphasic patients with a lexical-semantic deficit consistently failed to report items at the beginning of the list. This contrasts with patients with a more standard impairment of phonological STM who do better with initial items (N. Martin & Saffran, 1990, 1997; Saffran & Martin, 1990). It has been assumed that early items are more susceptible to semantic effects because they have more time to activate the corresponding semantic representations in LTM.

Lexical-semantic effects in STM: Evidence from neurologically intact participants

Lexicality and frequency

In serial recall, words are remembered better than nonwords (e.g., Gathercole, Pickering, Hall, & Peaker, 2001; Hulme, Maughan, & Brown, 1991; Hulme, Roodenrys, Brown, & Mercer, 1995), pseudohomophones are remembered better than control nonwords (e.g., BRANE vs. SLINT; Besner & Davelaar, 1982), and high-frequency words are recalled better than low-frequency words (e.g., Roodenrys, Hulme, Alban, Ellis, & Brown, 1994; M. J. Watkins, 1977). These lexical effects cannot be attributed to differences in rehearsal rate as they persist after controlling for speech rate (Hulme et al., 1991; Hulme et al., 1997) and under articulatory suppression (Gregg, Freedman, & Smith, 1989; Tehan & Humphreys, 1988).

Semantic similarity

Baddeley (1986) reported that semantic similarity has only a small and inconsistent effect on span tasks. However, he used a paradigm that strongly encouraged phonological retention. The sequences to recall were repeatedly drawn from the same pool of eight words, which remained visible, written on cards, throughout the testing session. This made

the task one of pure order retention, for which—we argue—a phonological code is most useful. When words have been drawn from an open set, semantically homogeneous lists (where all items are from the same category) have consistently shown an advantage over semantically heterogeneous lists (Fradet, Gil, & Gaonach, 1996; Huttenlocher & Newcombe, 1976; Poirier & Saint-Aubin, 1995). This facilitatory effect parallels what is found in free recall (Glanzer, Koppenaal, & Nelson, 1972; Greene & Crowder, 1984).

Concreteness

There are good reasons to assume that effects of concreteness or imageability are a type of semantic effect (we use concreteness and imageability interchangeably since they are highly correlated and difficult to distinguish). Aphasic patients are often better at reading, writing, and repeating concrete words (e.g., Coltheart, 1980; Marcel & Patterson, 1978; Saffran, 1982; Shallice & Warrington, 1975), although examples of the opposite pattern have also been reported (Breedin, Saffran, & Coslett, 1994; Warrington, 1975, 1981). This advantage has been generally explained by assuming that concrete words have richer semantic representations (e.g., Jones, 1985; Plaut & Shallice, 1991, 1993; Schwanenflugel & Shoben, 1983). A related explanation is that concrete words benefit from having associated visual as well as verbal features (Paivio, 1986, 1991).

Concreteness effects have been studied extensively using LTM tasks, particularly in the context of Paivio's dual coding theory (see Paivio, 1986, 1991), but they have been less studied using tasks tapping short-term memory retention. Two older studies reported no effect (Brenner, 1940; Paivio & Csapo, 1969), but items may not have been properly matched between conditions. More recent studies have reported positive effects. Saffran and Martin (1990) described patients who, in serial recall, show effects of imageability on early items and effects of frequency on late items. Bourassa and Besner (1994)—replicating a previous result by Tehan and Humphreys (1988)—found better recall for content words than

function words with and without articulatory suppression. These effects, however, disappeared when the two classes of words were equated for imageability, indicating that the grammatical-class effect was a by-product of an imageability effect. Finally, but most importantly, Walker and Hulme (1999) found concreteness effects in a serial recall task. The effects were independent of speech rate and of modality of recall (written or spoken) and were found both in forward and backward recall. No effects, however, were found with a matching task that required no spoken or written output.

Lexical-semantic effects in STM: Interpretations according to alternative views of memory

The results reported above indicate clearly that short-term recall does not draw exclusively on peripheral phonological representations. These results, however, are susceptible to different interpretations, which have different implications for models of memory.

The articulatory loop

Traditional models of memory have maintained a sharp distinction between a phonological STM and a semantic LTM (e.g., Baddeley & Levy, 1971; Craik & Levy, 1976; Kintsch & Buschke, 1969; Shallice, 1975). Baddeley's articulatory loop is the most popular STM model within this framework (e.g., Baddeley, 1986). In the articulatory loop, lexical-semantic effects in STM can be seen as the contribution of lexical semantic representations held in LTM to phonological representations held in an input buffer. The distinction between a phonological STM and a semantic LTM is maintained because transfer of information to LTM is not instantaneous but a function of the time representations spend in STM and/or of rehearsal (e.g., Baddeley & Patterson, 1971; Baddeley, Gathercole, & Papagno, 1998). These assumptions explain the prevalence of phonological effects in STM and the fact that semantic effects, when reported, have been generally associated with the beginning

part of the list to be recalled (e.g., Kintsch & Buschke, 1969; Walker & Hulme, 1999; O. C. Watkins & Watkins, 1977; also see N. Martin & Saffran, 1990, 1997, for neuropsychological evidence). Items at the beginning of the list have spent more time in STM and, thus, have had more chance to be rehearsed and transferred to LTM.

Within the articulatory loop framework, one can make two predictions regarding the presence of concreteness effects in STM: (a) They should be found only, or prevalently, in the primacy region of the serial position curve, and they should decrease monotonically from the beginning to the end of the list; (b) concreteness effects, especially those associated with the primacy region of the serial position curve, should be eliminated, or at least weakened, by suppression, which abolishes rehearsal and thus reduces the chance of transfer to LTM.

Redintegration

The redintegration hypothesis by Hulme and collaborators also shares the view that lexical-semantic effects in STM arise as a contribution of (LTM) lexical-semantic representations to buffered phonological representations (see Hulme et al., 1991, 1995, 1997; and also Schweickert, 1993). There are however, two crucial differences from the articulatory loop hypothesis. First, the activation of lexical-semantic representations is seen as an instantaneous process, which is part of what normally happens during language comprehension. Second, lexical-semantic representations contribute to activate phonological representations held in an output rather than in an input buffer. The idea is that lexical/semantic representations are used to reconstruct or “redintegrate” the degraded phonology of the words about to be spoken when a list is recalled. This would not be a conscious guessing strategy, but a mechanism hard-wired into the production system with the function of “cleaning up” output representations

by recirculating them through the lexical system (Hulme et al., 1991, 1995, 1997).

The redintegration hypothesis explains a number of effects. Redintegration will benefit: (a) real words over nonwords since nonwords do not have stored representations; (b) high-frequency over low-frequency words since high-frequency words are more easily retrieved using partial phonological information as a cue; (c) semantically homogeneous lists over heterogeneous lists since knowing the category of the words reduces the scope of the search.¹ The superiority of concrete words over abstract words is less clearly predicted. However, one could still assume that the richer semantic representations of concrete words will lead to increased activation at the lexical level.

The redintegration hypothesis makes different predictions about the presence of concreteness effects in STM. There is no reason to expect that lexical/semantic effects will be a function of serial position since they will occur automatically at the point of word production. Instead, this hypothesis predicts that concreteness effects will be found only in tasks that require overt output. In support of this claim, Walker and Hulme (1999) found that a concreteness effect was present in serial recall, but not in a matching-span task, which did not involve verbal output (participants only had to report whether two lists of words were the same or different). Gathercole et al. (2001), however, provided only partial support for the redintegration hypothesis. They investigated the effects of another lexical variable—whether the stimuli are real or made-up words—across tasks that do or do not require spoken production. They found that lexicality effects were much reduced in a matching-span task compared to serial recall, but they were still significant.

The multiple-codes hypothesis

An alternative proposal—which eliminates the distinction between a phonological STM and a

¹ However, Poirier and Saint-Aubin (1995) explicitly argued against this explanation because in their study wrong responses were rarely extraneous members of the same category.

semantic LTM—is that both phonological and lexical-semantic representations are involved in the retention of information at different retention intervals. This hypothesis has also taken different forms. Walker and Hulme (1999) have talked of the redintegration of temporary semantic representations, but they have left the role of buffered semantic representations underspecified. R.C. Martin and collaborators have put forward a more detailed model of the relation between language-processing and STM resources. According to this model, each language-processing component—including the semantic system—has an associated buffer component, which is responsible for the short-term retention of the corresponding representations. Buffered lexical-semantic representations would play a role in sentence comprehension and production, especially when individual word meanings need to be retained over time (e.g., R. C. Martin, Lesch, & Bartha, 1999; R. C. Martin & Romani, 1994, 1995; Romani & Martin, 1999). Finally, N. Martin and Saffran (1992, 1997) have also talked of a direct contribution of activated lexical-semantic representations to STM recall, although they have refrained from hypothesizing a separate semantic STM component in parallel with a phonological STM component.

In spite of their differences, all of these models assume that lexical semantic representations contribute directly to short-term memory recall as soon as a word is heard. As a result concreteness effects in STM should: (a) be present whether recall draws on an input or output buffer (i.e., they should be present in tasks with no overt output) and (b) not diminish with articulatory suppression nor decrease monotonically with serial position (at least in normal participants where access to lexical-semantic representations occurs in a matter of milliseconds).

Hypotheses that assume that different codes contribute directly to STM need to explain why phonological and not semantic effects are prevalent in STM experiments. However, the lack of semantic effects could be, at least in part, the result of the particular tasks and kinds of stimuli that have been used. Most studies have stressed

recall of *serial order* and have used a restricted set of items. These could be exactly the conditions that favour retention through a phonological code (for evidence that phonological coding is optimal for order retention see Baddeley & Lewis, 1981; Beaman & Jones, 1997; Gathercole et al., 2001; Watkins, Watkins, & Crowder, 1974; Wickelgren, 1965).

Plan of study

The purpose of the present study is to assess the predictions outlined above by assessing concreteness effects in a number of different conditions involving short-term recall.

We contrast tasks that do and do not involve overt output (serial recall and free recall versus a matching-span task and an order reconstruction task). If concreteness effects are found in all tasks this will argue against Walker and Hulme's (1999) contention that concreteness effects arise as a function of redintegration in an output buffer and, thus, occur only if the task requires overt output.

We also examine concreteness effects across serial positions both under control conditions and under articulatory suppression. If a concreteness effect occurs even under suppression, this will mean that it cannot be attributed to differences in speech rate. In addition, the lack of an interaction between concreteness, serial position, and suppression would suggest that concreteness effects are not due to a rehearsal-mediated process of transfer to LTM (which should mainly affect early items).

Finally, we examine whether phonological representations are particularly important to the recall of order information. We test two sets of predictions. The first is that conditions that disrupt phonological encoding make retention of order more difficult and favour semantic encoding. Suppression is one such condition (see Gupta & MacWhinney, 1995). Thus, semantic errors—which indicate lexical-semantic encoding—should be more frequent under suppression. Moreover, in free recall, suppression should abolish or reduce a recall strategy based on serial order. The second set of predictions is that

lexical semantic effects will be stronger when recall of items is more important. Thus, the concreteness effect should be stronger in free recall than in an order reconstruction task. Moreover, in a matching-span task, concreteness effects may be present only when words are drawn from an open set. When items are drawn from closed sets, lexical-semantic representations are less useful because the identity of the items is determined by the set.

EXPERIMENT 1

Concreteness and articulatory suppression in serial recall

The purpose of Experiment 1 was to assess the reliability and generality of concreteness effects in STM and to examine a possible interaction between concreteness and articulatory suppression. This experiment used a somewhat different methodology from that of Walker and Hulme (1999). First, we drew the stimuli from an open set instead of from a closed set. This should offer a better guarantee that stimuli in the different conditions are matched, since uncontrolled variation should cancel out across a larger set of stimuli. Secondly, we have used a more extended set of list lengths (Walker & Hulme used lists of 7 words; we use lists of 5, 6, and 7 words). Thirdly, we have used both a strict output procedure, as used by Walker and Hulme, and a less strict procedure, where participants had to write recalled words in the appropriate slots on the answering sheet, but not necessarily from left to right.

Method

Materials

Participants were asked to recall lists of five, six, and seven auditorily presented words. For each list length, 10 lists of abstract words and 10 lists of concrete words were presented. All words were bisyllabic nouns. Concrete words referred to items that could be experienced through the senses and, in particular, were easily pictured

(e.g., giraffe, custard, lettuce, pocket). Abstract words referred to qualities, feelings, and abstract concepts (e.g., magic, caution, symbol, anger). The experimental words were rated for concreteness on a 5-point scale (with 5 corresponding to the most concrete) by a group of 25 participants who did not take part in the memory experiments. These participants gave abstract words an average rating of 2.3 (range 1.3–3.9; $SD = 0.5$) and concrete words an average rating of 4.4 (range = 3.9–5.0; $SD = 0.2$). Abstract and concrete words were matched one to one for frequency, letter length, and phoneme length. There were no repetitions of the same words within or between lists. A complete listing of the experimental words is reported in Appendix 1.

Procedure

Participants were tested individually in a quiet room. Words were read aloud by the experimenter at a rate of about one every 1.5 s. As soon as the list was finished, participants had to write down all the words that they remembered in the appropriate ordinal slots on an answer sheet, leaving a blank where an item could not be recalled. In Experiment 1A, they were free to start with any word in the list, provided the final result reflected the order as they remembered it. For example, they could write down Word 5 first as long as they entered it in Slot 5. They were allowed to make alterations to the list they were working on, but not to lists already completed. In Experiment 1B, participants were told that they had to recall the items strictly from left to right. Thus, they could not regress to skipped items when filling out the response sheet. Whenever participants violated this instruction, they were reminded of it, and the last response was not counted. Only lists of six words were used in this condition. In all conditions, there was no time limit to complete recall. Concrete and abstract lists were presented blocked, and the order of presentation was counterbalanced across participants.

In the articulatory suppression condition participants had to repeat “Coca-cola” continuously, both while the words were being read out and in the recall phase while they were writing them down. They

Table 1. Serial recall of five-, six-, and seven-word lists

List		Control condition				Suppression condition					
		Concrete		Abstract		Diff	Concrete		Abstract		Diff
		Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Five words	Items/order	88.6	8.7	87.2	10.1	1.4	64.4	16.0	58.6	13.6	5.8
	Items	93.6	6.6	92.3	8.2	1.3	73.8	12.3	68.0	12.6	5.8
Six words	Items/order	79.1	12.5	75.2	12.8	3.9	57.0	11.3	54.9	12.3	2.1
	Items	87.8	8.4	85.4	8.4	2.4	66.0	9.8	61.5	11.0	4.8
Seven words	Items/order	62.5	11.3	57.1	9.9	5.4	44.7	10.3	40.2	9.2	4.5
	Items	72.9	9.3	68.9	8.0	4.0	52.5	10.7	47.7	9.4	4.8
Six words strict output	Items/order	79.5	14.1	69.9	13.9	9.6	47.3	15.6	40.1	10.4	7.2
	Items	86.0	9.7	80.7	10.3	5.3	56.9	14.3	49.2	9.6	7.8

Note: Scores expressed as a percentage of items correct and items in the correct order.

were not allowed to add any more words after they had stopped saying "Coca-cola". Participants were very good at complying with the instructions. Control and suppression conditions were presented to the same participants, and the order of presentation was counterbalanced. The same stimuli were used for the control condition and the articulatory suppression condition, but with a different randomization. Participants did two practice lists before starting the experimental lists.

Participants

All participants were native speakers of English. Most of them were undergraduate or postgraduate students who participated in order to gain research credits. Experiment 1A was carried out at the University of Birmingham, Experiment 1B at Rice University. In Experiment 1A, 36 participants recalled five-word lists, 25 recalled six-word lists, and 25 recalled seven-word lists. A total of 32 participants recalled six-word lists in Experiment 1B.

Scoring

We report two ways of scoring lists: (a) number of items correct in the correct order (henceforth items/order); (b) number of items correct even if they were reported in the wrong order (henceforth items).

Results

Results are reported in Table 1. The two types of scores for each list length were analysed using a two-way repeated measures analysis of variance (ANOVA) with articulatory suppression and concreteness as factors.

As expected, words in the suppression condition were always recalled worse than in the control condition. The effects of suppression were significant in all conditions: five words: items/order, $F(1, 35) = 142.4$; $MSE = 44.1$; $p < .001$; items, $F(1, 35) = 161.4$; $MSE = 27.1$; $p < .001$; six words: items/order, $F(1, 24) = 89.2$; $MSE = 45.2$; $p < .001$; items, $F(1, 24) = 234.5$; $MSE = 19.7$; $p < .001$; seven words: items/order, $F(1, 24) = 110.5$; $MSE = 33.4$; $p < .001$; items, $F(1, 24) = 187.9$; $MSE = 28.3$; $p < .001$; six words strict output: items/order, $F(1, 31) = 284.6$; $MSE = 38.9$; $p < .001$; items, $F(1, 31) = 356.3$; $MSE = 29.7$; $p < .001$.

Concrete words were always recalled better than abstract words. Concreteness effects were significant in all conditions for both scoring methods except for the six-word lists scored by items/order: five words: items/order, $F(1, 35) = 8.8$; $MSE = 13.6$; $p < .005$; items, $F(1, 35) = 17.6$; $MSE = 6.5$; $p < .001$; six words: items/order, $F(1, 24) = 2.8$; $MSE = 29.4$; $p = .11$; items, $F(1, 24) = 7.3$; $MSE = 15.8$; $p < .01$; seven words: items/order, $F(1, 24) = 6.7$; MSE

= 44.1; $p < .05$; items, $F(1, 24) = 6.7$; $MSE = 44.1$; $p < .05$; six words strict output: items/order, $F(1, 15) = 25.5$; $MSE = 32.1$; $p < .001$; items, $F(1, 31) = 23.7$; $MSE = 20.6$; $p < .001$.

The interaction between concreteness and suppression was significant only for the five-word lists: items/order, $F(1, 35) = 3.9$; $MSE = 11.2$; $p = .06$; items, $F(1, 35) = 8.1$; $MSE = 5.5$; $p < .01$. For this list length, concreteness was highly significant under suppression: items, $F(1, 35) = 16.1$; $MSE = 9.3$; $p < .001$; items/order, $F(1, 35) = 9.1$; $MSE = 16.7$; $p < .005$, but was not significant in the quiet condition: items, $F(1, 35) = 1.2$; $MSE = 8.0$; *ns*; items/order, $F(1, 35) = 3.0$; $MSE = 2.6$; *ns*. There were no significant interactions for the six- and seven-word lists: six words: items/order, $F(1, 24) = 0.5$; $MSE = 14.7$; $p = .49$; items, $F(1, 24) = 1.0$; $MSE = 13.1$; $p = .32$; seven words: items/order, $F(1, 24) = 0.1$; $MSE = 26.7$; $p = .76$; items, $F(1, 24) = 0.08$; $MSE = 21.3$; $p = .78$; six words strict output: items/order, $F(1, 31) = 1.0$; $MSE = 17.3$; $p = .34$; items, $F(1, 31) = 1.4$; $MSE = 13.1$; $p = .25$. These results suggest that the interaction found with five-word lists is the result of a ceiling effect in the control condition.

Comparison across conditions

We have also analysed effects of concreteness and type of scoring considering all four conditions together (lists of five, six, and seven words and lists of six words with strict output). We have used differences between concrete and abstract words (the concreteness effect) as the dependent measure. List length was a between-subjects factor, and suppression and type of scoring were within-subjects factors. There was no significant main effect of type of scoring, $F(1, 114) = 0.3$; $MSE = 20.3$; $p = .56$, but there was a significant interaction between type of scoring and suppression, $F(1, 114) = 5.2$; $MSE = 19.8$; $p = .3$. When scored by items, the concreteness effect was smaller in the control than in the suppression condition (3.5% vs. 5.8%). This difference was marginally significant, $F(1, 114) = 3.6$; $MSE = 71.3$; $p = .06$. Instead, the concreteness effect did not differ across conditions where the results

were scored by items/order. It was 4.7% in the control condition and 5.2% in the suppression condition, $F(1, 14) = 0.03$; $MSE = 89.1$; $p = .85$.

Discussion

Our results replicate and extend those of Walker and Hulme (1999) by showing concreteness effects at all list lengths in a task of immediate serial recall. The fact that results were replicated in spite of a number of procedural differences strengthens our confidence that concreteness effects in STM are robust and reliable. Thus, concreteness should be controlled, among other variables, in tasks assessing verbal short-term memory in different conditions.

Clearly concreteness effects are not due to a confounding with speech rate. Walker and Hulme (1999) did control for speech rate and still found an effect of concreteness. In our experiment, words were closely matched for syllable length and phoneme length, and recent studies have shown no effect of speech rates in words matched for length (e.g., Lovatt, Avons, & Masterson, 2000; Service, 1998). Moreover, similar effects of concreteness were found under articulatory suppression, which should abolish rehearsal and, with it, any differences linked to speech rate.

Across the four experiments, concreteness effects were stronger under suppression and when results were scored in terms of item correct only. These results suggest that semantic variables are more important when recalling identity information and when the contribution of phonological representations is weaker, as is the case under suppression. Differences in the recall of item and order information are further explored in Experiment 3.

EXPERIMENT 2

Concreteness and articulatory suppression in a matching-span task

As outlined in the Introduction, Walker and Hulme (1999) found no concreteness effect in a

matching-span task and argued that concreteness effects are only found in tasks involving spoken output. Another possibility, however, is that Walker and Hulme drew items from a closed set rather than from an open set and that this methodological choice weakened concreteness effects. Roodenrys and Quinlan (2000) have demonstrated that frequency effects are stronger when open sets of words are used. The same may be true for concreteness effects. The purpose of Experiment 2 was to assess whether we could find a concreteness effect even in a matching task when open rather than closed sets of stimuli were used.

Method

Experiment 2A

Materials

The control condition involved 32 trials with concrete words and 32 trials with abstract words. In each case, 16 were “same” trials where the first list of five words was followed by an identical list; the other 16 were “different” trials where the second list included the same words but in a different order. Order permutations always involved adjacent words. There were four instances of each of the four possible permutations (involving Positions 1–2, 2–3, 3–4, 4–5). The 80 words involved in each task came from the original pool of words used in Experiment 1 plus 10 new pairs (see Appendix 2). Each word was used twice. Overall, concrete and abstract words were well matched for frequency (concrete = 26.5; $SD = 31.9$; abstract = 26.8; $SD = 31.4$), letter length (concrete = 6.1; $SD = 0.7$; abstract = 6.1; $SD = 0.7$), phoneme length (concrete = 5.2; $SD = 0.8$; abstract = 5.3; $SD = 0.7$), and syllable length (concrete = 2.01; $SD = 0.11$; abstract = 2.01; $SD = 0.19$).

To avoid possible floor effects, we used lists of four words under suppression. Thus, the suppression condition, involved 30 trials with concrete words and 30 trials with abstract words (15 “same” and 15 “different”). There were five instances of each of the three possible position permutations (1–2, 2–3, 3–4). A total of 60 words

were used twice each, for each of the concrete and abstract probe tasks. They were a subset of the 80 words used in the control conditions. They were, again, closely matched for frequency (concrete = 27.4; $SD = 29.2$; abstract = 27.9; $SD = 29.2$), letter length (concrete = 6.0; $SD = 0.7$; abstract = 6.1; $SD = 0.7$), phoneme length (concrete = 5.2; $SD = 0.8$; abstract = 5.3; $SD = 0.6$), and syllable length (concrete = 2.00; $SD = 0.00$; abstract = 1.98; $SD = 0.13$).

In both conditions, repetitions were avoided, both within lists and between adjacent lists, and care was taken to avoid lists where phonological or semantic similarity could serve as a clue for recall.

Procedure

The lists were presented via a minidisc recording at the rate of one word every second. A longer pause (about two seconds) was made between the target list and the probe list. Participants were asked to record their response by circling either “same” or “different” on an answer sheet next to the appropriate trial number. Each participant carried out the task with concrete and abstract words, in control and suppression conditions in a single testing session. The order of presentation of the four conditions was counterbalanced across participants.

At the end of the session, participants were asked to rate the words for concreteness on a scale from 1 to 5 (1 very abstract; 5 very concrete) and familiarity, again on a scale from 1 to 5 (1 very unfamiliar; 5 very familiar). Half of the participants did the familiarity rating first; the other half did the concreteness rating first. Familiarity ratings were 4.6 ($SD = 0.3$) for concrete words and 4.4 ($SD = 0.5$) for abstract words. Concreteness ratings were 4.9 ($SD = 0.4$) for concrete words and 2.2 ($SD = 0.6$) for abstract words.

Participants

A total of 20 students at the University of Aston participated in Experiment 2A. They were all native English speakers. They were all tested individually in a quiet room. Compliance with the instructions to rehearse was good.

Experiment 2B

Materials

The purpose of Experiment 2B was to use the same stimuli as those in Experiment 2A, but to directly compare an open-set with a closed-set condition. For this last condition, the sets of 80 abstract and concrete words were subdivided into eight pairs of matched sets each containing 10 abstract and 10 concrete words. The abstract and concrete words in each pair of sets were matched for frequency and length. Only one pair of sets was administered to each participant. The words in the sets were repeatedly sampled to produce 32 series of 5 words in the control condition and the 30 series of 4 words for the suppression condition. Thus, each word was used 16 times in the control condition and 12 times in the suppression condition. There were no word repetitions within sets. Each pair of sets was administered to 4 participants.

Procedure

The lists were recorded on a Macintosh computer using the program Sound Edit 16. These digitized lists were presented to participants using PsychoScope 1.2.5. The words were presented at the rate of one word per second. A 1,500-ms pause was made between the target list and the probe list. Participants were asked whether or not the lists were "identical" and indicated their response with a yes or no key press. Each participant was presented with concrete and abstract word lists under both the suppression and control conditions within the same experimental session. The order of presentation of the four conditions was counter-balanced across participants and lists.

Participants

A total of 32 participants carried out the open-set condition and a further 32 the closed-set condition. They were all students at Rice University who received research credit for their participation. They were all native English speakers. They were all tested individually in a quiet room. Compliance with the instructions to rehearse was good.

Experiment 2C

In Experiment 2A, we found, unexpectedly, that abstract rather than concrete words were recalled better under suppression. To check the reliability of this result, we ran the closed set of stimuli a second time with a second group of 32 Rice students after some minor modifications to the stimuli. We were worried that in a few sets the concrete words were more confusable because of pairs of similarly sounding words (e.g., button/butter in one set; garbage/garden in another; tablet/planet in another). Moreover, we checked for differences between concrete and abstract words in the frequency with which words within the same set started with the same letter. In everything else (including number of participants) Experiment 2C was identical to Experiment 2B.

Results

Results from all conditions are shown in Table 2.

An ANOVA with concreteness and suppression as within-subjects factors and open/closed set as a between-subjects factor showed a significant main effect of suppression, $F(1, 114) = 81.1$; $MSE = 144.5$; $p < .001$, no main effect of concreteness, $F(1, 114) = 0.8$; $MSE = 111.5$; $p = .36$, no main effect of type of set, $F(1, 114) = 0.1$; $MSE = 380.0$; $p = .73$, but significant interactions between suppression and concreteness, $F(1, 114) = 8.8$; $MSE = 69.5$; $p = .004$, and set and concreteness, $F(1, 114) = 8.5$; $MSE = 111.5$; $p = .004$. We, therefore, ran separate analyses for open and closed sets.

For the open set, there was a significant main effect of suppression, $F(1, 51) = 30.5$; $MSE = 150.5$; $p < .001$, no main effect of concreteness, $F(1, 51) = 1.7$; $MSE = 111.3$; $p < .19$, but an interaction between suppression and concreteness, $F(1, 51) = 5.5$; $MSE = 71.8$; $p < .02$. There was no effect of concreteness under suppression, $F(1, 51) = 0.1$; $MSE = 125.5$; $p < .72$, but concrete words were recalled better than abstract words in the control condition, $F(1, 51) = 10.0$; $MSE = 57.6$; $p < .003$.

Table 2. Matching-span task: same/different recognition of the order of series of five or four words for the control and suppression conditions

Set	Experiment	N	Control condition					Suppression condition				
			Concrete		Abstract		Diff	Concrete		Abstract		Diff
			Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Open	2A	20	75.9	12.8	70.5	15.0	5.4	64.7	16.1	63.7	15.1	1.0
	2B	32	81.3	10.6	77.1	12.4	4.2	68.6	14.5	70.5	12.3	-1.9
	Total	52	79.3		74.5		4.8	67.1		67.8		-0.7
Closed	2B	32	73.2	13.5	76.7	12.6	-3.5	63.9	13.5	69.0	14.8	-5.1
	2C	32	78.7	11.6	79.1	11.6	-0.4	62.7	12.8	68.9	13.7	-6.2
	Total	64	75.9		77.9		-2.0	63.3		69.0		-5.7

Note: Scores expressed as a percentage of correct responses.

For the closed set, there was a main effect of suppression, $F(1, 63) = 53.6$; $MSE = 139.6$; $p < .001$, a main effect of concreteness in the direction opposite of what was expected, $F(1, 63) = 8.2$; $MSE = 111.7$; $p < .006$, and a significant interaction between suppression and concreteness, $F(1, 63) = 5.5$; $MSE = 71.8$; $p = .02$. There was no effect of concreteness in the control conditions, $F(1, 63) = 1.5$; $MSE = 78.0$; $p < .23$, but abstract words were recalled better than concrete words under suppression, $F(1, 63) = 10.1$; $MSE = 101.3$; $p = .002$. This advantage was significant in both Experiment 2B, $F(1, 31) = 4.1$; $MSE = 100.7$; $p = .05$, and Experiment 2C, $F(1, 31) = 5.9$; $MSE = 104.9$; $p = .02$.

Discussion

Results in the control conditions were what we expected. There was a significant and positive concreteness effect with open sets, but no effect with the closed sets. Our results indicate that concreteness effects can be obtained even in a matching-probe task when an open set rather than a closed set is used. This suggests that concreteness effects are not contingent on whether or not the task requires overt production of the words to recall, contrary to Walker and Hulme's (1999) hypothesis. What is important, instead, is that the task encourages the use of lexical-semantic representations. The use of these representations is much more limited when the same words are

presented over and over again, as with closed sets. Here, the identity of the words becomes constrained by the stimulus set, and lexical-semantic effects have a reduced scope to emerge.

Results under suppression were unexpected. We found no concreteness effect with open sets and a significant, reversed effect with closed sets. This pattern, however, can be explained if, under suppression, the phonological representations supporting order became so weak that they no longer support the link with their corresponding lexical-semantic representations. This is consistent with a large number of participants performing at or near chance (18/30 correct or less) in the suppression condition, especially with the closed sets (participants near chance: concrete words/closed sets, $31/64 = 48.4\%$; concrete words/open sets, $16/52 = 30.8\%$). With this in mind, one can propose the following explanation.

In the control condition, a relatively good peripheral phonological record keeps lexical/semantic representations in their relative order. With the open sets of words, semantic activation helps to retain identity information. Abstract words—which have less rich semantic representations—lose identity information more often than concrete words. This leads to more errors and to a concreteness advantage. When the words come from closed sets, the role of semantic information in preserving identity information is more limited since information about the set already constrains the identity of the lexical items. In this condition,

errors rarely stem from loss of word identity information. This results in no difference between concrete and abstract words.

In the suppression condition, a much poorer peripheral phonological record often fails to keep semantic representations in the right order. Without order information, the identity advantage for concrete words is of no use. Thus, there is little or no concreteness effect with open sets. With closed sets, lexical semantic representations for the whole set become overactivated due to repeated exposure. Since phonological representations are very degraded, the semantic input to the buffer will not reintegrate existing representations, but, instead, it will replace them without regard to list order. This will have negative effects on the ability to remember the order of the items. Concrete words will be more susceptible to these negative effects, and this will lead to an inverse concreteness effect.

To give an example, let us suppose a participant has heard the words “carrot” and “giraffe” in this order. Capacity limitations and suppression have reduced the phonological representations of these words to a very faint record in the phonological buffer. However, in the closed-set condition, both of these words will receive strong lexical-semantic activation. This activation will sometimes completely overwrite the phonological record so that the participant will be aware that both “carrot” and “giraffe” have been presented, but will not remember in which order. Abstract words will receive less strong lexical-semantic activation. Thus, more often the original phonological record will be able to assert its predominance over lexical-semantic activation, which will more often support rather than displace the phonological record. This will result in the observed advantage for abstract words.

In the General Discussion, we present a model where retention of order is accomplished mainly through phonological representations, while semantic representations would help to maintain and reconstruct the identity of the phonological representations. In normal control conditions, semantic representations would be linked to the corresponding phonological representations, and this will be helpful. However, the matching task carried out

under suppression and with closed sets of words represents an extreme situation where the phonological representations may be too weak to command the ordering of the corresponding lexical-semantic representations. In this situation high semantic activation may become a hindrance. The results presented here fit well with this model and with the overall theme emerging from our study: phonology is important for the retention of order, while semantics supports the retention of identity.

EXPERIMENT 3

Concreteness and articulatory suppression in order reconstruction and free-recall tasks

The purpose of Experiment 3 was twofold. We wanted to assess concreteness effects in a second input task, and, in addition, we wanted to provide further evidence that phonological and semantic coding are preferentially used for the recall order and identity information. For these purposes, we contrast an order reconstruction task (Experiment 3A) with a free-recall task (Experiment 3B). Both tasks involve memory for all the stimuli in the list (thus, reducing the problem of chance performance). However, only order information has to be recalled in the order reconstruction task, while only item information has to be recalled in the free-recall task.

Clearly, item information is important even in a task that asks for recall of order (see our previous experiment and also Neath, 1997), and it has been demonstrated that a recall strategy based on serial order is used even in free recall (e.g., DeLosh & McDaniel, 1996; Kahana, 1996). Still, the order reconstruction task clearly emphasizes memory for order more than does the free-recall task. On this basis, we predict that a concreteness effect—which is linked to the recall of identity information—will be stronger in free recall than in the order reconstruction task. In addition, we expect that a recall strategy based on serial order should be less evident when free recall is carried out under suppression, which disrupts phonological coding.

Method

Experiment 3A: Order reconstruction

Neath (1997) previously reported concreteness effects in an order reconstruction task. Participants were presented with lists of six items and then were asked to assign numbers to the words to indicate the original order of presentation. However, a distraction task was inserted between presentation of each word and the following one, which made the task one of long-term memory. We used a similar task, but without an intervening distractor so that only temporary retention was required, in line with other STM tasks.

Materials

Participants were asked to sort lists of seven or eight words. The stimuli were the same as those for Experiment 1 with the addition of 10 new pairs (see Appendix 2). All added words were bisyllabic nouns like the original set. Concrete and abstract words were well matched for phoneme length (concrete: mean = 6.1, $SD = 0.7$; abstract: mean = 6.1, $SD = 0.7$), frequency (concrete: mean = 24.5, $SD = 27.9$; abstract: mean = 24.2, $SD = 27.9$) and letter length (concrete: mean = 5.3, $SD = 0.8$; abstract: mean = 5.4; $SD = 0.8$). For each length, 10 lists of concrete words and 10 lists of abstract words were presented. There were no repetitions of the same words either within or between series.

Procedure

Each stimulus was presented typed on a card. Cards were placed in front of participants one after another, with each card covering the previous one, at a rate of approximately one every three seconds (we used a slower rate to guarantee that participants had enough time to read the words). The cards were then placed back on the table in random order. Participants had to rearrange the cards into the order of presentation and copy the words down on an answer sheet. In the articulatory suppression condition, the participants had to continuously produce the word "Coca-cola" during

both presentation and recall of the stimuli. They were not allowed to add more words to the list once they had stopped saying "Coca-cola". The same participants carried out the control and suppression conditions with a one-week gap. The order of the two sessions was counterbalanced. The same stimuli were used for the control and the suppression conditions, but with a different randomization.

Scoring

Each item in the correct order scored 1 point. No credit was given for a run of items in the correct relative order but not in the correct absolute order. Thus, the series 1345672 would score only 1 point since the misplaced second item puts all the others out of place. This example is given for clarification; extreme cases of this sort did not occur.

Experiment 3B: Free recall

Materials and procedure

The stimuli were the same as those in Experiment 3A (although with different randomizations), and they were presented in the same way. At the end of each series, participants had to write down all the words they could remember, irrespective of order. All other details of the procedure were identical to Experiment 3A. Items were scored correct if they were part of the series regardless of order.

Participants

To facilitate comparison, the same participants carried out the two tasks. A total of 18 participants carried out the task with lists of seven words and a further 18 with lists of eight words. They were all native speakers of English. Most of them were undergraduates at the University of Birmingham who participated in order to gain research credits.

Results

Results are shown in Table 3. By chance, the group which was given eight words to remember performed slightly better than the group that was given seven words.

Table 3. Percentages of concrete and abstract words recalled correctly for lists of seven and eight words in an order reconstruction task and in free recall

Task	Words	Control condition				Suppression condition					
		Concrete		Abstract		Diff	Concrete		Abstract		
		Mean	SD	Mean	SD		Mean	SD	Mean	SD	Diff
Order reconstruction	Seven	82.3	12.3	80.4	10.4	1.9	72.9	16.3	64.0	14.4	8.9
	Eight	86.2	12.8	79.4	14.8	6.7	75.2	14.5	64.9	15.4	10.3
	Mean	84.2		79.9		4.3	74.3		64.3		10.0
Free recall	Seven	78.9	11.7	70.3	7.6	8.6	59.9	15.5	51.5	15.1	8.4
	Eight	83.3	13.7	70.3	11.0	13.0	67.6	16.5	51.7	9.9	15.9
	Mean	81.1		70.3		10.8	63.7		51.6		12.2

The scores in the order reconstruction task were analysed using a three-way mixed ANOVA with suppression and concreteness as within-subjects factors and list length (seven vs. eight) as a between-subjects factor. List length did not interact significantly with the other variables: List \times Concreteness, $F(1, 34) = 1.4$; $p = .25$; List \times Suppression, $F(1, 34) = 0.2$; $p = .66$. Performance was significantly better in the control condition than under suppression, $F(1, 34) = 49.7$; $MSE = 66.7$; $p < .001$, and was significantly better for concrete than abstract words, $F(1, 34) = 16.6$; $MSE = 60.8$; $p < .001$. Although the concreteness effect was larger under suppression, the Suppression \times Concreteness interaction failed to reach significance, $F(1, 35) = 2.4$; $MSE = 54.8$; $p = .13$.

Scores in free recall were similarly analysed using a three-way mixed ANOVA with suppression and concreteness as within-subjects factors and list length as a between-subjects factor. List length did not interact significantly with the other variables: List \times Concreteness, $F(1, 35) = 3.4$; $MSE = 28.4$; $p = .08$; List \times Suppression, $F(1, 35) = 0.001$; $p = .98$. Recall was significantly better in the control condition than under suppression, $F(1, 35) = 152.2$; $MSE = 44.7$; $p < .001$, and concrete words were recalled better than abstract words, $F(1, 35) = 42.7$; $MSE = 68.6$; $p < .001$. Again, there was no Suppression \times Concreteness interaction, $F(1, 35) = 0.1$; $MSE = 28.3$; $p = .75$.

Effects of concreteness in the two tasks were compared using a three-way repeated measures ANOVA with type of task (free recall vs. order reconstruction), concreteness, and suppression as within-participants factors. As predicted, there was a significant interaction between type of task and concreteness: The concreteness effect was significantly larger in free recall than in order reconstruction, $F(1, 34) = 10.6$; $MSE = 23.6$; $p < .005$. There was no three-way interaction, $F(1, 34) = 1.4$; $MSE = 35.9$; $p = .25$.

Order of output in free recall. Following previous studies (e.g., Burns, 1992; DeLosh & McDaniel, 1996; Nairne, Riegler, & Serra, 1991) we used the Asch–Ebenholtz (1962) measure of seriation to investigate the extent to which a recall strategy based on serial order was used even in free recall. This is a ratio of the pairs of words produced in the correct relative order out of the total number of pairs produced (e.g., an output like 1, 3, 2, 7; composed of three pairs, of which two are in the correct order; therefore $2/3 = .67$). A measure of .50 indicates no serial order. Results are presented in Table 4. They have been analysed using a mixed ANOVA with suppression and concreteness as within-subjects factors and number of items (seven or eight) as a between-subjects factor. The dependent variable was the proportion of pairs recalled in the proper order.

There were no significant effects of list length or concreteness, $F(1, 33) = 0.03$; $MSE = 0.015$;

Table 4. Proportion of items recalled in the order of presentation for abstract and concrete words in free recall

Words		Control condition		Suppression condition	
		%	SD	%	SD
Seven	Concrete	.69	.25	.45	.33
	Abstract	.69	.27	.45	.35
	Total	.69	.26	.45	.34
Eight	Concrete	.61	.26	.47	.32
	Abstract	.63	.26	.47	.26
	Total	.62	.26	.47	.29

$p = .85$, and no interaction between suppression and concreteness, $F(1, 33) = 0.07$; $MSE = 0.012$; $p = .79$. Instead, as predicted, there was a main effect of suppression. A strategy based on serial order was used significantly more in the control than in the suppression condition, $F(1, 33) = 53.3$; $MSE = 0.026$; $p < .001$. Follow-up analyses indicated that, in the control condition, significantly more items were recalled in the original order than would be expected by chance: for seven words, $\chi^2(1) = 109.1$; $p < .001$; for eight words, $\chi^2(1) = 45.1$; $p < .001$. This is consistent with previous evidence that a serial output is used even in free recall. There was, instead, no evidence that such a strategy was used under suppression. Here the number of pairs in the original order did not differ from chance: for seven words, $\chi^2(1) = 0.43$; $p = .51$; for eight words, $\chi^2(1) = 0.59$; $p = .44$.

Discussion

Experiment 3 supports the findings of Experiment 2 in showing that concreteness effects are found in STM tasks whether or not overt output of the words is required by the task. Taken together, these results suggest that lexical-semantic representations contribute to STM recall from the moment the words are presented and not only at the time of output. Moreover, the results of Experiment 3 support the findings of the previous two experiments in showing that significant concreteness effects are also obtained in the suppression conditions. These results indicate that

concreteness effects do not depend on transfer to LTM as a function to rehearsal (this point is explored further in the section examining serial position curves).

Experiment 3 also shows that the kinds of representations used in STM are a function of task demands. The effect of concreteness was stronger in free recall than in order reconstruction. Furthermore, a recall strategy based on serial order was not used in free recall under suppression. These results support the hypothesis that phonological representations are most relied upon when the task stresses recall of order information and/or when task conditions allow a good peripheral phonological record to be used. Under suppression, phonological coding is disrupted and so is the use of a retrieval strategy based on serial order.

ERROR ANALYSES

So far we have analysed level of performance, but not the types of error made in different conditions. Our hypothesis that the relation between phonological and lexical semantic coding in STM depends on the nature of the task makes two predictions regarding the errors. It is well established that phonological errors are prevalent in serial recall (e.g., Baddeley, 1966b; Conrad & Hull, 1964). In addition, however, our hypothesis predicts that more semantic errors will be made under suppression than in the control condition. This is because suppression should disrupt phonological coding and, thus, encourage reliance on alternative (lexical/semantic) representations. Moreover, more semantic errors should occur in free recall than in serial recall because lexical/semantic representations should be used more when the task does not stress the retention of order. Rates of different kinds of lexical substitutions in serial recall and free recall are reported in Table 5. Given the small number of semantic errors, results have been collapsed across lists and modality of output.

As expected, among lexical errors, phonological substitutions (e.g., cycle \rightarrow psycho; butter \rightarrow button) were by far the most common error type.

Table 5. *Types of lexical substitution in serial recall and immediate free recall*

		<i>Control</i>		<i>Suppression</i>		<i>Overall</i>	
Serial recall	Phonological	221	52.4	190	40.5	411	46.1
	Semantic	11	2.6	35	7.5	46	5.2
	Mixed (phon. + sem)	13	3.1	13	2.8	26	2.9
	Other	177	41.9	231	49.3	408	45.8
	Total	422		469		891	
Free recall	Phonological	89	49.4	65	43.9	154	47.1
	Semantic	7	3.9	15	10.1	22	6.7
	Mixed (phon. + sem)	20	11.1	21	14.2	41	12.5
	Other	64	35.6	46	31.1	110	33.6
	Total	180		147		327	

Only a few semantic substitutions (e.g., elbow → hand; cabbage → lettuce) were made, limiting the scope for analyses. Results, however, were in the expected direction. More semantic substitutions were made under suppression than in the control condition in both tasks: serial recall, $\chi^2(1) = 14.8$; $p < .001$; free recall, $\chi^2(1) = 5.1$; $p = .02$. There was no difference in the rate of semantic and phonological substitutions in serial and free recall: $\chi^2(1) = 0.8$; $p = .38$. However, significantly more mixed phonological/semantic substitutions (e.g., ideal → idea; section → sector; salmon → lemon) were made in free recall than in serial recall: $\chi^2(1) = 32.3$; $p < .001$.

SERIAL POSITION CURVES

One can look at recall of different serial positions to investigate whether the concreteness effect results from transfer to LTM. In immediate serial recall of lists that slightly exceed participants' capacity—like those used here—recall is better for the first and last items in the list and poorer in the middle. Generally, primacy effects have been linked to the use of semantic codes, while recency effects have been linked to phonological coding. This interpretation stems from a number of results.

O. C. Watkins and Watkins (1977) found that in the recall of both auditorily and visually presented lists, frequency had a stronger influence on items at the beginning of the list than on

those at the end. Walker and Hulme (1999) found that the concreteness effect was generally stable across serial positions, with the exception of the very last positions where it was absent. In contrast, phonological variables have been found to affect items at the end of the list most strongly. For example, Murray (1968) found that the advantage of auditory over visual stimuli was larger for items at the end of the list; Brooks and Watkins (1990) found that the detrimental effect of rhyming words was more pronounced when they were placed in the second (versus the first) part of the list. These results are consistent with those found with longer lists and free recall, which also indicate that items early in the list are most susceptible to lexical/semantic effects, while late items are most susceptible to phonological effects (e.g., Kintsch & Buschke, 1969).

The neuropsychological literature has also provided supporting evidence. Aphasic patients with lexical/semantic deficits have been found to show reduced primacy effects, while patients with more prominent phonological deficits have been found to show reduced recency effects (N. Martin & Saffran, 1990, 1997). In addition, N. Martin and Saffran (1997) found a positive correlation between the presence of an imageability effect and the presence of a primacy effect in their group of patients.

The prevalent finding that lexical/semantic effects are associated with the primacy region of the serial position curve has been explained by assuming that lexical-semantic effects take some

time to be established, and, thus, they are more evident for items at the beginning of the list, which have spent more time in STM. In addition, in models that endorse a phonological STM/semantic LTM distinction—like Baddeley's articulatory loop—semantic effects are a function of transfer to LTM and, thus, potentially, a function of rehearsal. These models predict not only stronger concreteness effects in the primacy part of the serial position curve, but also an interaction with suppression. Stronger concreteness effects in the primacy region of the recall curve should be eliminated by suppression, which prevents rehearsal and, thus, weakens transfer to LTM.

Not all results from the literature, however, have been consistent. Some patients with lexical/semantic deficits show clear primacy effects (Forde & Humphreys, 2002; Knott, Patterson, & Hodges, 1997; R. C. Martin et al., 1999). Moreover, Hulme and collaborators, using a paradigm very similar to O. C. Watkins and Watkins (1977), found the opposite relation between frequency and serial position (Hulme et al., 1997; Hulme, Stuart, Brown, & Morin, 2003). Frequency effects increased, rather than decreased, from left to right.

In the light of these contradictory results, we examine the interaction between concreteness, serial position, and articulatory suppression in three different tasks: serial recall, free recall, and order reconstruction. Consistent with previous studies, we expect clear primacy and recency effects in the serial recall and free-recall tasks. What to expect in the order reconstruction task is less clear. In this task, items are dealt with according to the random order of the cards on the table. This may disrupt a retrieval strategy based on serial order.

Results

Figure 1 shows the results for serial recall, Figure 2 for free recall, and Figure 3 for the order reconstruction task. Each graph shows results for different lists and for the control and the suppression conditions.

In serial recall and in free recall results are similar. Generally, the data show U-shaped functions with clear recency and primacy effects (the exception is the control condition with seven words, which is likely to be a chance result). In addition, like Walker and Hulme (1999), the concreteness effect appears relatively constant for all positions with the exception of the last two where it is generally absent. The pattern is the same in the control and in the suppression conditions. In the order reconstruction task, instead, the curves are flatter, and there are no recency effects.

To assess the hypothesis that concreteness effects will be stronger in the initial part of the curve, we carried out a three-way repeated measures ANOVA with primacy/recency, suppression, and concreteness as factors. We have considered the first half of the serial curve to belong to the primacy region and the second half to belong to the recency region. In the case of lists with odd numbers of words, the word in the middle was not considered. Results are presented by task cumulating across list lengths, since the pattern did not change with length.

In serial recall with no enforced output order, there was a highly significant interaction of serial position (primacy/recency) and concreteness. The advantage for concrete words was 8.0% in the primacy region and -1.0% in the recency region, $F(1, 83) = 32.3$; $p < .001$; $MSE = 1.1$. Post hoc ANOVAs showed that concreteness was highly significant in the primacy region: control, $F(1, 85) = 13.8$; $p < .001$; suppression, $F(1, 85) = 42.0$; $MSE = 1.15$; $p < .001$, but not significant in the recency region: control, $F(1, 85) = 0.2$; $MSE = 0.64$; $p = .66$; suppression, $F(1, 85) = 1.7$; $MSE = 1.45$; $p = .20$. This pattern is unlikely to be related solely to the level of performance, which was similar in the two regions of the curve (cumulating across conditions: recency, 65.2% correct; primacy, 71.2% correct). There was also a significant three-way interaction between serial position, concreteness, and suppression, $F(1, 83) = 7.2$; $p = .009$; $MSE = 0.96$. The pattern was stronger in the suppression condition, however, post hoc ANOVAs showed significant Concreteness \times Serial Position interactions in

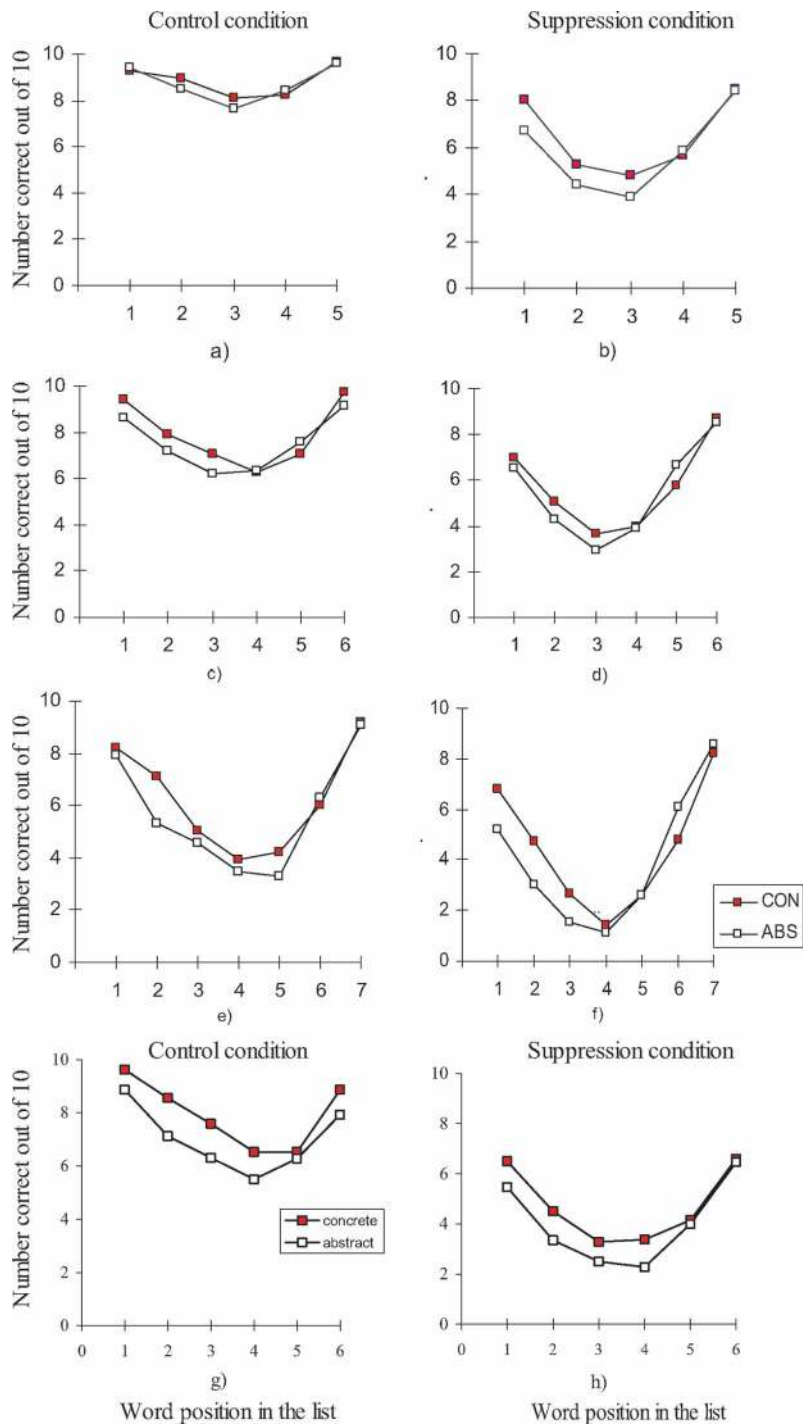


Figure 1. Serial recall: Mean correct by serial position in the control and suppression conditions; (a) and (b) are lists of five words; (b) and (c) are lists of six words; (e) and (f) are lists of seven words; and (g) and (h) are list of six words recalled in strict order of presentation.

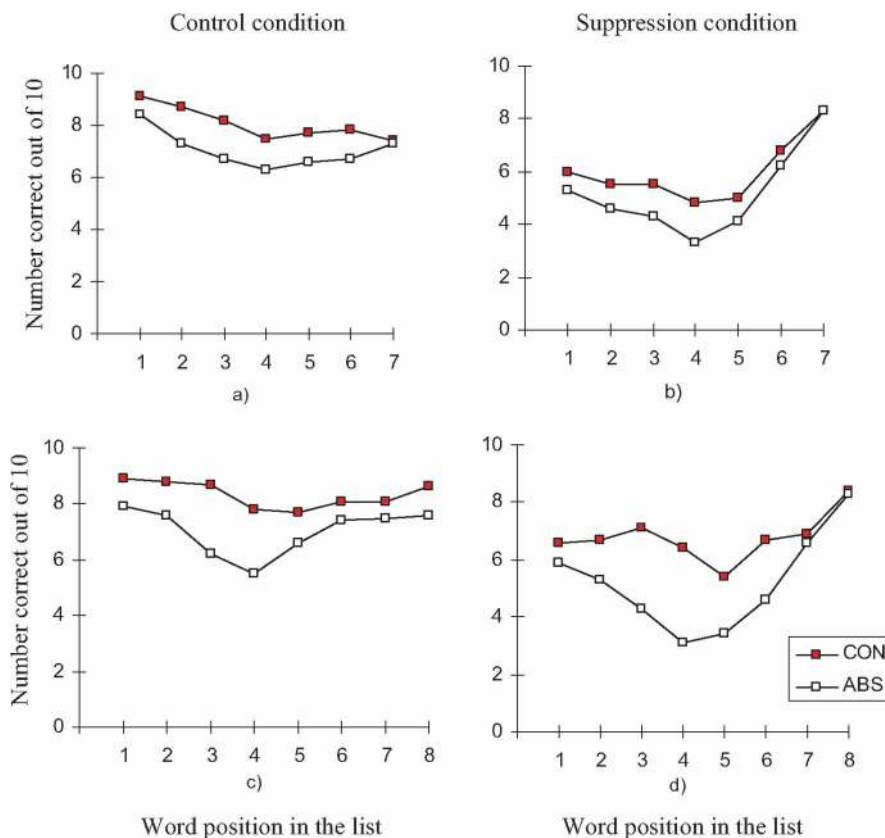


Figure 2. Free recall: Mean correct by serial position for concrete and abstract words in the control and suppression condition; (a) and (b) are lists of seven words; (c) and (d) are lists of eight words.

both conditions: control, $F(1, 83) = 7.0$; $MSE = 0.74$; $p = .009$; suppression, $F(1, 83) = 28.2$; $MSE = 1.27$; $p = .000$.

In serial recall with strict serial output, the interaction between concreteness and serial position was again significant, $F(1, 31) = 4.3$; $p = .046$; $MSE = 15.0$. The concreteness effect was stronger in the primacy than in the recency portion of the serial position curve (3.2% vs. 1.8%). There was no three-way interaction with suppression, $F(1, 31) = 0.8$; $p = .37$; $MSE = 5.7$.

In free recall, there was also a highly significant interaction of serial position and concreteness, $F(1, 35) = 8.6$; $MSE = 1.1$; $p = .006$, with a larger concreteness effect in the primacy region. Post hoc analyses, however, showed that concreteness was significant both in the primacy region, $F(1, 35) =$

45.6; $MSE = 1.7$; $p = .000$, and in the recency region, $F(1, 35) = 11.5$; $MSE = 1.7$; $p = .002$. There was no significant three-way interaction of position, concreteness, and suppression, $F(1, 35) = 0.2$; $MSE = 1.1$; $p = .67$. In the control condition, the concreteness effect was 14.9% in the primacy region and 6.4% in the recency region. Under suppression, it was 14.8% in the primacy region and 8.4% in the recency region.

In the order reconstruction task, there were no significant two-way or three-way interactions: Concreteness \times Position, $F(1, 35) = 2.6$; $p = .12$; $MSE = 5.9$; Concreteness \times Suppression \times Position, $F(1, 35) = 0.2$; $p = .70$; $MSE = 7.4$.

There was no steady decrement of the concreteness effect across the primacy region of any tasks. In serial recall, averaging across tasks and

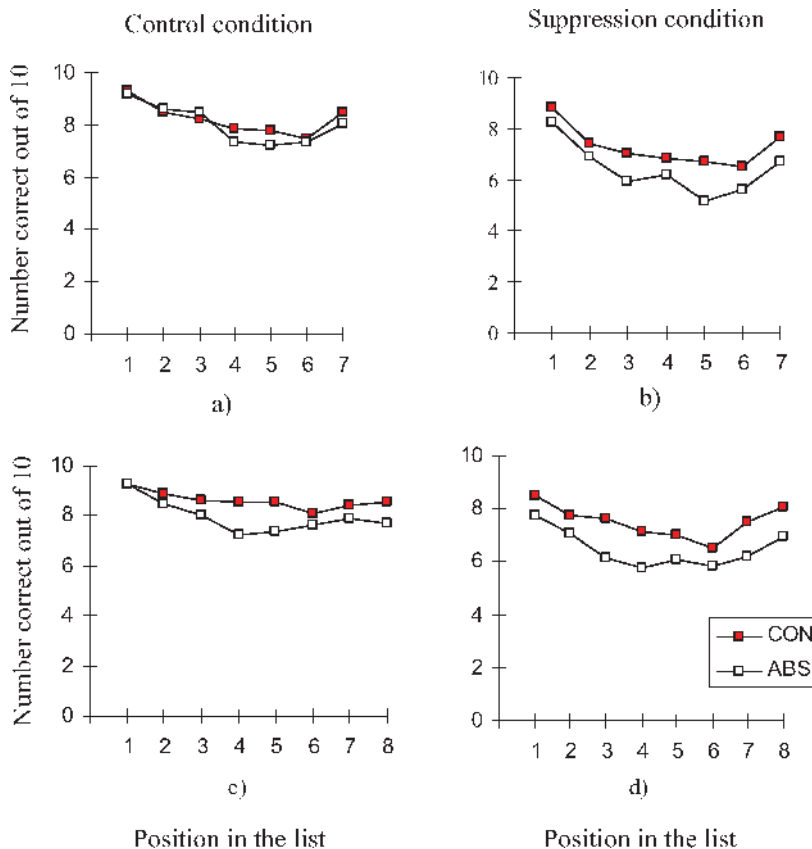


Figure 3. Order reconstruction: Mean correct by serial position for concrete and abstract words in the control and suppression conditions; (a) and (b) are lists of seven words; (c) and (d) are lists of eight words.

conditions, the concreteness effect for the first three positions was 9.6%, 11.5%, and 8.9%. In both free recall and order reconstruction, results, if anything, showed a trend in the opposite direction. For the first four positions, the concreteness effect was for free recall, 7.6%, 12.4%, 19.9%, and 21.0%; for order reconstruction, 3.5%, 3.8%, 7.4%, and 9.4%. These last patterns are probably linked to level of performance. The effects are stronger when performance is lower, and there is more room for the effects to emerge.

Discussion

The serial position curves that we obtained in serial and free recall are similar to those reported by

Walker and Hulme (1999). They are bow shaped and show stronger concreteness effects in the primacy region of the curve. These results are in agreement with those of other studies, which have found stronger lexical/semantic effects in the primacy regions (e.g., N. Martin & Saffran, 1997; Saffran & Martin, 1990; O. C. Watkins & Watkins, 1977). Results did not change when we enforced strict serial output.

We found no evidence, however, that stronger concreteness effects in the first part of the serial position curve are linked to rehearsal. Across tasks, the pattern was similar in control conditions and under suppression. If anything, the results in serial recall showed the opposite pattern, with stronger concreteness effects under suppression.

These results contrast with predictions made by models that assign an important role to rehearsal in the transfer of information from short- to long-term memory (e.g., Baddeley, 1986).

Stronger concreteness effects in the primacy region, independent of suppression, are consistent with the hypothesis that it takes time for lexical-semantic representations to be activated and contribute to short-term recall. This hypothesis, however, predicts a steady decrement of the concreteness effect across the serial position curve, which was not observed.

Taken together, our results suggest, instead, that concreteness effects are present all along, but their effect is masked for the very last positions by the presence of strong peripheral records: acoustic/phonological, in the case of auditory presentation, or orthographic, in the case of visual presentation (see Chincotta, Underwood, Ghani, Papadopoulou, & Wresinski, 1999; Cowan, Baddeley, Elliott, & Norris, 2003). This interpretation is supported by the results found with the order reconstruction task. In this task, there is no recency effect, and, at the same time, there is no interaction between concreteness and serial position.²

The serial position curves for concreteness are different from those for frequency (Hulme et al., 1997, 2003). One explanation suggested by Walker and Hulme (1999) is that both temporary semantic traces and temporary phonological traces are susceptible to the redintegration process, but that, by the end of the list, phonological traces need more reconstruction because presentation of further words in the list degrades them more (this assumption, however, is ad hoc). Another possibility suggested by Hulme et al. (2003) is that frequency and concreteness effects have different sources, and only concreteness effects depend on redintegration. Whatever the final explanation, one may note that, consistently, across experiments, lexical-semantic effects are reduced for the last two positions in the list indicating a

masking of these effects in the presence of a strong phonological record.

GENERAL DISCUSSION

The experiments reported in this study have shown consistent, positive effects of concreteness in tasks tapping immediate recall of items in the proper order (serial recall), recall of items independent of order (free recall), and recall of item positions (matching span and order reconstruction). In all of these tasks, concreteness effects were present in both the control and the suppression conditions, although suppression modulated the strength of the effects and, in some cases, changed their direction. We believe that our results have important implications for the nature and locus of semantic effects in short-term memory, for the relative contribution of different kinds of representations to recall, and, ultimately, for theoretical models of memory.

The locus and nature of concreteness effects

A first conclusion forced by our results is that both semantic and phonological representations contribute flexibly and independently to short-term recall as soon as they become available. This hypothesis is in contrast with earlier views that attributed semantic effects exclusively to LTM or with more recent views that link semantic effects to output processes.

Our results are inconsistent with the hypothesis that lexical-semantic variables affect STM tasks because recall of items in the primacy region draw from LTM while recall of items in the recency region draw from STM (see Cohen, 1970; M. J. Watkins, 1977; O. C. Watkins & Watkins, 1977). All traditional dual models of memory have assumed that transfer to LTM is not instantaneous, but that it occurs only after items have been held and rehearsed in STM for

² However, on average, the items presented first in the list will be those that have spent more time in STM (and been rehearsed more). Thus, the hypothesis that lexical-semantic effects are a function of transfer to LTM predicts stronger concreteness effects for these items.

some time. This hypothesis predicts that the concreteness effects should show both a steady reduction across positions and a decrease under suppression. Our results provided no evidence for these claims. In fact, in serial recall with item-only scoring, the concreteness effect was stronger, not weaker, under suppression. With the matching task and the closed set of words, the concreteness effect was again stronger under suppression, although it reversed direction.

Our results are also inconsistent with the redintegration hypothesis of Hulme and collaborators, at least to the extent that redintegration only occurs at output (Hulme et al., 1991, 1995, 1997; Walker & Hulme, 1999). According to this hypothesis, effects of lexical semantic variables in STM are strictly associated with the processes occurring at the time of language production. Phonological representations held in an output buffer will receive reinforcing activation from corresponding lexical and semantic representations, which will help to reintegrate lost information. Contrary to this hypothesis, however, we found concreteness effects even in matching-span and order reconstruction tasks, neither of which requires word production. We have argued that lack of significant findings may result from using a restricted pool of words. Consistent with this, in the control condition of the matching-span task, we found positive concreteness effects with open sets but not with closed sets of words.

Gathercole et al. (2001) found that lexicality effects (the advantage of words over nonwords) were much stronger in serial recall than in a matching-span task. In contrast, we found no clear differences for concreteness. Averaging across control conditions, the concreteness effect was: 5.1 in serial recall, 4.8 in the matching-span task with open sets, and 4.3 in the order reconstruction task. The different strength of the lexicality effect across tasks is not surprising. Redintegration plays a reduced role in a matching-span task since a yes/no response does not depend on complete phonological representations. Concrete and abstract words differ less in term of redintegration than do words and nonwords. Both can be reintegrated (although one may

argue that stronger semantic representations will facilitate redintegration of concrete words). Concrete and abstract words, instead, differ in the strength/richness of their semantic representation. Therefore, our results suggest that semantic representations contribute to the retention of word identity similarly across input/output tasks.

The prevalence of phonological representations in the recall of order

A number of separate results from our study converge to indicate that, while semantic representations are very important in the recall of identity information, phonological representations play a special role for the recall of order information. First, Experiment 3 was designed to directly contrast retention of items and retention of order. We found that concreteness effects were stronger in free recall than in order reconstruction. Secondly, a number of results indicated that suppression—which we assume impairs phonological coding—changes the way that phonological and semantic representations contribute to the task.

Suppression impairs retention of order information. We found that a recall strategy based on serial order was used in the control condition of free recall, but not under suppression. These results are consistent with those of Jones and collaborators who have argued that irrelevant speech—which interferes with phonological representations in a way similar to articulatory suppression (see Neath, 2000)—has a particularly disruptive effect on tasks involving retention of order (Beaman & Jones, 1997).

Suppression impairs phonological coding, but encourages semantic coding. In serial recall, when results were scored as item-correct only, concreteness effects were stronger under suppression. In addition, more semantic errors were made under suppression than in control conditions, both in serial recall and in free recall.

Finally, when the task emphasizes retention of order, but conditions both impair phonological retention and increase semantic activation, concreteness may begin to have negative, rather than positive, effects. The matching-span task, as close

to a pure order task as possible, fits this description when run under suppression and with a closed set of words. We have assumed that, in these conditions, semantic representations—which are highly activated due to repetition—become delinked from the corresponding phonological representations that maintain order information in the buffer. In these conditions, semantic information may actually contribute to disrupt the fragile representation of order at the phonological level.

Theoretical interpretations

As mentioned in the Introduction, several models from the literature have put at their core the idea that language-processing components and STM resources are tightly related. In these models, lexical semantic representations contribute to recall across time delays. These models, however, have assumed different numbers and kinds of buffers. The model proposed by R. C. Martin et al. (1999) is presented in Figure 4a. In this model, each storage component has its associated buffer. The representations in the buffers correspond to the activated representations in the storage components, and representations in the buffers and in the storage components mutually contribute to each other's activation. This model, therefore, assumes a distinction between storage components (or "knowledge structures") and components involved in the temporary buffering of representations. This is in contrast to the model by N. Martin and Saffran (1997), which largely dispenses with separate buffer components.

In the N. Martin and Saffran (1997) model, STM is equated to *activated* stored representations. This echoes earlier accounts. For example, in Shiffrin's (1976) "associative memory model", STM was viewed as the activated portion of LTM. In Craik and Lockhart's model (Craik, 1983; Craik & Lockhart, 1972), STM corresponded to a limited-capacity processor, which could operate at different levels and shift between them (e.g., visual, phonological, semantic). However, N. Martin and Saffran recognize the problem of representing order simply

through patterns of activation in the storage components (item repetitions are particularly difficult to represent). Thus, although their model has no semantic buffer, the representations in the lexical system are related to a sequence place holder with a role similar to that of a phonological buffer. The computational models of Gupta and MacWhinney (1997) and the feature model by Nairne (Nairne, 1990; Neath, 2000) have endorsed similar ideas. The idea that buffered phonological representations serve as place holders fits well with our results, which indicate a special role for these representations in the retention of order information.

In Figure 4b, we present a hybrid model that incorporates features of the R. C. Martin et al. (1999) model with the idea that buffered phonological representations serve as place holders during both speech comprehension and speech production. In this model, semantic effects in STM are due to bidirectional links between phonological lexical representations and semantic representations. Words with richer semantic representations will more strongly activate the corresponding phonological lexical representations, which will feed activation to the buffered phonological representations. To be parsimonious, this model implements a semantic buffer without hypothesizing a separate structural component. Instead, the semantic buffer is equated to the temporally activated representations in lexical-semantic memory. We assume that activated lexical-semantic representations will rely on buffered phonological representations for encoding order (see also Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005, for a computational model with an activation-based semantic buffer). This feature of the model is supported by the reversal of concreteness effects when the task emphasizes order and when the activation of semantic representations overpowers phonological representations (as in the matching-span task with closed set of words under suppression). This result indicates that, by themselves, semantic representations are unable to encode order, or may do this only in a rough/approximate fashion on the basis of relative degree of activation.

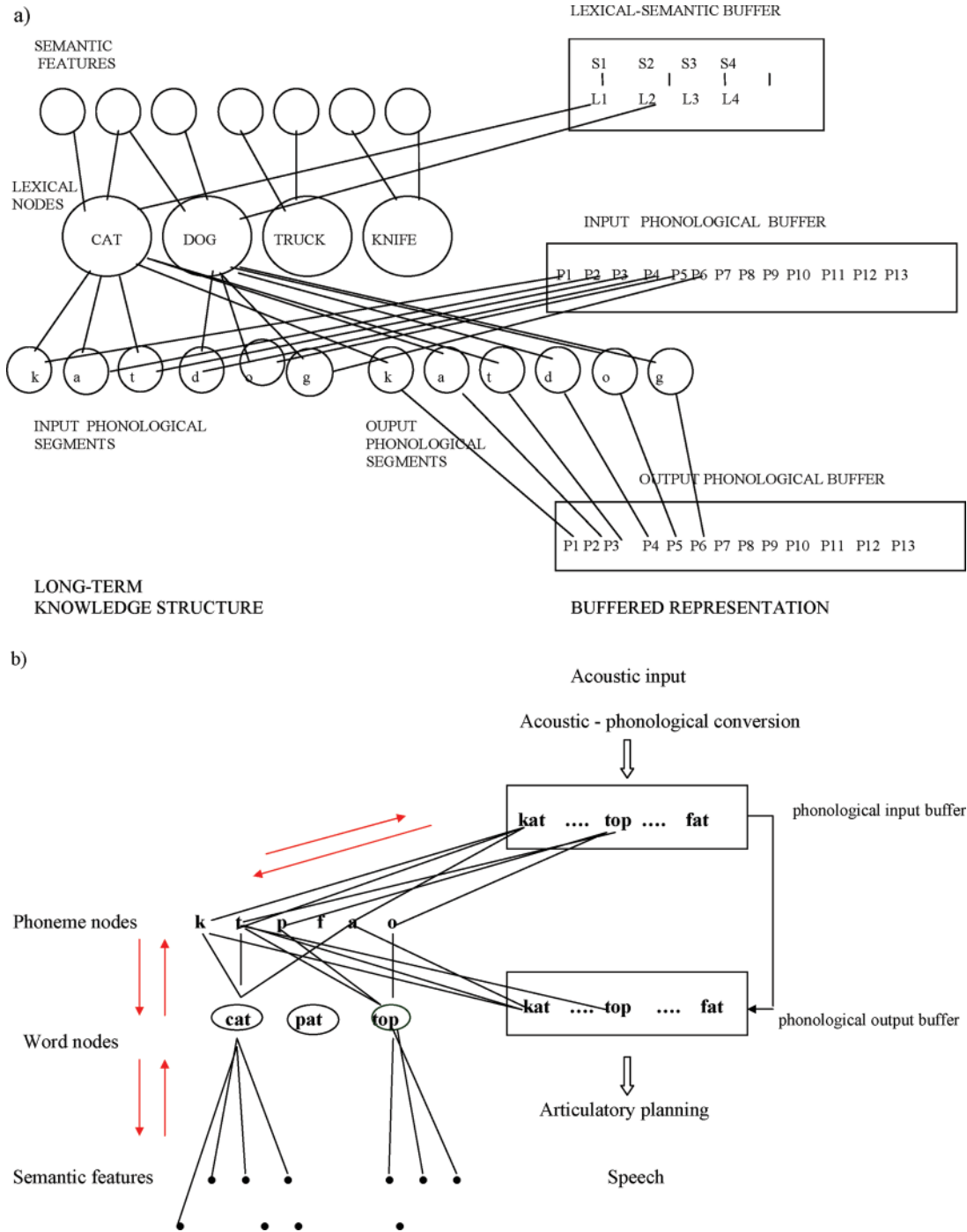


Figure 4. (a) The multiple-buffer model of short-term memory (from R. C. Martin et al., 1999). (b) A place holder model of short-term memory.

As in the R. C. Martin et al. (1999) model, we assume a separate syntactic buffer attached to the language component responsible for deriving syntactic structure and propositional meaning in comprehension and production of connected speech (not represented in Figure 4b; see R. C. Martin & Freedman, 2001; R. C. Martin & Romani, 1994). We also assume both input and output phonological buffers that feed into and are fed by a single set of lexical-phonological representations. There is no real evidence for separate input and output phonological representations (e.g., Hillis, 2001), whereas evidence for two buffers has been presented by a number of studies (see R. C. Martin et al., 1999; Romani, 1992; Shallice, Rumiaty, & Zadini, 2000). The two buffers have different roles since the input buffer converts acoustic into phonological representations while the output buffer converts phonological into articulatory representations.³

CONCLUSIONS

The results that we have presented suggest that lexical-semantic representations contribute to short-term recall from the moment they are heard without the need to invoke lengthy processes of encoding in LTM or the refreshing of representations in an output buffer. These results are broadly consistent with views of memory as a resource linked to specific kinds of representation rather than a system of independent and multi-purpose components (e.g., Barnard, 1985; Dell, 1986, 1988; Just & Carpenter, 1992; N. Martin & Saffran, 1997; R. C. Martin et al., 1999; Monsell, 1984). More specifically, our results are consistent with the existence of short-term

memory resources linked to the activation of lexical-semantic representations. Converging evidence for the existence of these resources, separate from phonological resources, comes from recent behavioural, neuropsychological, neuroimaging, and computational studies (we have reviewed behavioural and neuropsychological studies; for neuroimaging studies, see Fiez, 1997; Shivde & Thompson-Schill, 2004; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001; for computational studies, see Davelaar et al., 2005; Gupta & MacWhinney, 1997; Neath, 2000).

We have also presented evidence that lexical and phonological representations contribute differently to recall depending on task demands and that phonological representations are optimal for the retention of order. These results are well explained by a model where more peripheral (phonological) representations held in (input and output) buffers serve as place holders to which a number of other representations are attached (e.g., lexical, semantic). The ability to represent order may, in fact, be what distinguishes a structural component like the phonological buffers from a “virtual” component like the semantic buffer. Further studies should corroborate this conclusion.

It remains to ask why the competing view, that we use phonology to hold on to immediate memories and semantics to hold on long-term memories, has been so attractive. One answer is that this distinction embodies some important truth. Things are learned better the more they are meaningful, organized, and integrated with existing knowledge. However, this view conflates two kinds of semantic memory. Memory for individual items is part of the lexical-semantic system. Activation in this system may, in fact, decay at a rate quite similar to that of phonological representations. Instead, it is memory for integrated meanings like

³ Particularly compelling evidence for two buffers has been provided by the case of an amnesic patient—MS (R. C. Martin et al., 1999)—who performed very poorly in word serial recall, but very well in input STM tasks that did not require spoken responses (e.g., a rhyme probe task and an order recognition task). His performance can be explained by assuming that he is impaired in all tasks requiring word production—STM tasks included—because of damaged links between semantic and phonological lexical representations. The only way to reconcile MS’s pattern of performance with a single set of buffered representations would be to hypothesize that he performs well on input tasks because these tasks are normally not influenced by lexical-semantic representations. The results presented here, however, show the opposite.

propositions or, at an even higher level, for events, facts, and stories, which is characterized by a much slower rate of decay depending on the degree of integration with existing knowledge (see Romani & Martin, 1999, for related evidence).

The second reason for the widespread belief that only phonology is used in short-term retention is that most experimental evidence has been drawn from a restricted set of tasks, which require the serial recall of a number of unrelated words. To compound the problem, items have often been drawn from small closed sets. In these conditions, the task becomes mainly one of order retention, for which phonological coding is most efficient. Our results are well explained by a model where phonological representations play a crucial role in the retention of order as place holders, without forcing the conclusion that only phonological coding is available in STM.

Original manuscript received 11 May 2005

Accepted revision received 24 November 2006

First published online 10 May 2007

REFERENCES

- Asch, S. E., & Ebenholtz, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *Journal of Psychology*, *54*, 3–31.
- Baddeley, A. D. (1966a). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology*, *18*, 302–309.
- Baddeley, A. D. (1966b). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, *18*, 362–365.
- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173.
- Baddeley, A. D., & Levy, B. A. (1971). Semantic coding and short-term memory. *Journal of Experimental Psychology*, *89*, 132–136.
- Baddeley, A. D., & Lewis, V. J. (1981). Inner active processes in reading: The inner voice, the inner ear and the inner eye. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 108–129). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Baddeley, A. D., & Patterson, K. (1971). The relation between long-term and short-term memory. *British Medical Bulletin*, *27*, 237–242.
- Barnard, P. (1985). Interacting cognitive subsystems: A psycholinguistic approach to short-term memory. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2, pp. 197–285). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 459–471.
- Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology*, *36*, 701–711.
- Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin and Review*, *1*, 122–125.
- Breedin, S. D., Saffran, E. M., & Coslett, H. B. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology*, *11*, 617–660.
- Brener, L. R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology*, *26*, 467–482.
- Brooks, J. O., & Watkins, M. J. (1990). Further evidence of the intricacy of memory span. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 1134–1141.
- Burns, D. J. (1992). The consequences of generation. *Journal of Memory and Language*, *31*, 615–633.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency book*. Boston: Houghton Mifflin Company.
- Chincotta, D., Underwood, G., Ghani, K. A., Papadopoulou, E., & Wresinski, M. (1999). Memory span for Arabic numerals and digit words: Evidence for a limited-capacity, visuo-spatial storage system. *Quarterly Journal of Experimental Psychology*, *52A*, 325–351.
- Cohen, R. L. (1970). Recency effects in long-term recall and recognition. *Journal of Verbal Learning and Verbal Behaviour*, *9*, 672–678.
- Coltheart, M. (1980). Deep dyslexia: A right hemisphere hypothesis. In M. Coltheart, K. Patterson,

- & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 326–380). London: Routledge & Keegan Paul.
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, *55*, 429–432.
- Cowan, N., Baddeley, A. D., Elliott, E. M., & Norris, J. (2003). List composition and word length effect in immediate serial recall: A comparison of localist and globalist assumptions. *Psychonomic Bulletin and Review*, *10*, 74–79.
- Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London*, *B302*, 341–359.
- Craik, F. I. M., & Levy, B. A. (1976). The concept of primary memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 4. Attention and memory* (pp. 133–175). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
- Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revised: Empirical and computational investigations of recency effects. *Psychological Review*, *112*, 3–42.
- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*, 124–142.
- DeLosh, E. L., & McDaniel, M. A. (1996). The role of order information in free recall: Application of the word-frequency effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, 1136–1146.
- Fiez, J. (1997). Phonology, semantics, and the role of left inferior prefrontal cortex. *Human Brain Mapping*, *5*, 79–83.
- Forde, M. E., & Humphreys, G. W. (2002). The role of semantic knowledge in short-term memory. *Neurocase*, *8*, 13–27.
- Fradet, A., Gil, R., & Gaonach, D. (1996). Codages phonologique et sémantique en mémoire à court terme chez des sujets adultes jeunes ou âgés. [Phonological and semantic coding in short-term memory in young and ageing adults]. *Revue de Neuropsychologie*, *6*, 53–81.
- Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology*, *54*, 1–30.
- Glanzer, M., Koppelaar, L., & Nelson, R. (1972). Effects of relations between words on short-term storage and long-term storage. *Journal of Verbal Learning and Verbal Behavior*, *11*, 403–416.
- Greene, R. L., & Crowder, R. G. (1984). Effects of semantic similarity on long-term recency. *American Journal of Psychology*, *97*, 441–449.
- Gregg, V. H., Freedman, C. M., & Smith, D. K. (1989). Word frequency, articulatory suppression and memory span. *British Journal of Psychology*, *80*, 363–374.
- Gupta, P., & MacWhinney, B. (1995). Is the articulatory loop articulatory or auditory? Reexamining the effects of concurrent articulation on immediate serial recall. *Journal of Memory and Language*, *34*, 63–88.
- Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural basis. *Brain and Language*, *59*, 267–333.
- Hillis, A. E. (2001). The organization of the lexical system. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology* (pp. 185–210). Ann Arbor, MI: Edwards Brothers.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*, 685–701.
- Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of Psychology*, *86*, 527–536.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 1217–1232.
- Hulme, C., 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.
- Huttenlocher, J., & Newcombe, N. (1976). Semantic effects on ordered recall. *Journal of Verbal Learning and Verbal Behavior*, *15*, 387–399.
- Jones, G. V. (1985). Deep dyslexia, imageability and ease of predication. *Brain and Language*, *24*, 1–19.

- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.
- Kahana, M. L. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*, 103–109.
- Kintsch, W., & Buschke, H. (1969). Homophones and synonyms in short-term memory. *Journal of Experimental Psychology*, *80*, 403–407.
- Knott, R., Patterson, K., & Hodges, J. (1997). Lexical and semantic binding effects in short-term memory: Evidence from semantic dementia. *Cognitive Neuropsychology*, *14*, 1165–216.
- Knott, R., Patterson, K., & Hodges, J. (2000). The role of speech production in auditory-verbal short-term memory: Evidence from progressive fluent aphasia. *Neuropsychologia*, *38*, 125–142.
- Lovatt, P. J., Avons, S. E., & Masterson, J. (2000). Output decay in immediate serial recall: Speech time revisited. *Journal of Memory and Language*, *46*, 227–243.
- Marcel, A. J., & Patterson, K. E. (1978). Word recognition and production: Reciprocity in clinical and normal studies. In J. Requin (Ed.), *Attention and performance VII* (pp. 207–226). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Martin, N., & Saffran, E. M. (1990). Repetition and verbal STM in transcortical sensory aphasia: A case study. *Brain and Language*, *39*, 254–288.
- Martin, N., & Saffran, E. M. (1992). A computational account of deep dysphasia. *Brain and Language*, *43*, 240–274.
- Martin, N., & Saffran, E. M. (1997). Language and short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, *14*, 641–682.
- Martin, R. C., & Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory*, *9*, 261–280.
- Martin, R. C., & He, T. (2004). Semantic short-term memory and its role in sentence processing: A replication. *Brain and Language*, *89*, 76–82.
- Martin, R. C., & Lesch, M. (1996). Associations and dissociations between language impairment and list recall: Implications for models of STM. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 149–178). Hove, UK: Psychology Press.
- Martin, R. C., Lesch, M., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, *41*, 3–29.
- Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple components view. *Neuropsychology*, *8*, 506–523.
- Martin, R. C., & Romani, C. (1995). Remembering stories but not words. In R. Campbell and M. A. Conway (Eds.), *Broken memories: Case studies in memory impairment*. Oxford, UK: Blackwell.
- Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, *33*, 83–111.
- Monsell, S. (1984). Components of working memory underlying verbal skills: A distributed capacities view. A tutorial review. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 327–350). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, *4*, 679–684.
- Nairne, J. (1990). A feature model of immediate memory. *Memory and Cognition*, *18*, 251–269.
- Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *17*, 702–709.
- Neath, I. (1997). Modality, concreteness, and set-size effects in a free reconstruction of order task. *Memory and Cognition*, *25*, 256–263.
- Neath, I. (2000). Modelling the effects of irrelevant speech on memory. *Psychonomic Bulletin and Review*, *7*, 403–423.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.
- Paivio, A., & Csapo, K. (1969). Concrete-image and verbal memory codes. *Journal of Experimental Psychology*, *80*, 279–285.
- Patterson, K., Graham, N., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, *6*, 57–69.
- Plaut, D. C., & Shallice, T. (1991). Effects of word abstractness in a connectionist model of deep dyslexia. *Proceedings of the 13th Annual Meeting of the Cognitive Science Society* (pp. 73–78). Chicago.

- Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*, 377–500.
- Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *Quarterly Journal of Experimental Psychology*, *48A*, 384–404.
- Romani, C. (1992). Are there distinct input and output buffers? Evidence from a brain-damaged patient. *Language and Cognitive Processes*, *7*, 131–162.
- Romani, C., & Martin, R. C. (1999). A deficit in the short-term retention of lexical-semantic information: Forgetting words but remembering a story. *Journal of Experimental Psychology*, *128*, 56–77.
- Romani, C., McAlpine, S., Olson, A., Tsouknida, E., & Martin, N. (2005). Length, lexicality and articulatory suppression in immediate recall: Evidence against the articulatory loop. *Journal of Memory and Language*, *52*, 398–415.
- Roodenrys, S., Hulme, C., Alban, J., Ellis, A., & Brown, G. D. A. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory and Cognition*, *22*, 695–701.
- Roodenrys, S., & Quinlan, P. T. (2000). The effects of stimulus set size and word frequency on verbal serial recall. *Memory*, *9*, 71–78.
- Saffran, E. M. (1982). Neuropsychological approaches to the study of language. *British Journal of Psychology*, *73*, 317–337.
- Saffran, E. M., & Martin, N. (1990). Neuropsychological evidence for lexical involvement in short-term memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 428–447). Cambridge, UK: Cambridge University Press.
- Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *9*, 82–102.
- Schweickert, R. (1993). A multinomial processing tree model for degradation and reintegration in immediate recall. *Memory & Cognition*, *21*, 167–175.
- Schweickert, R., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science*, *1*, 74–77.
- Service, E. (1998). The effect of word length on immediate serial recall depends on phonological complexity, not articulatory duration. *Quarterly Journal of Experimental Psychology*, *51A*, 283–230.
- Shallice, T. (1975). On the contents of primary memory. In S. Dornic & P. Rabbitt (Eds.), *Attention and performance V* (pp. 269–280). New York: Academic Press.
- Shallice, T., Rumiati, R. I., & Zadini, A. (2000). The selective impairment of the phonological output buffer. *Cognitive Neuropsychology*, *17*, 517–546.
- Shallice, T., & Warrington, E. K. (1975). Word recognition in a phonemic dyslexic patient. *Quarterly Journal of Experimental Psychology*, *27*, 187–199.
- Shiffrin, R. M. (1976). Capacity limitations in information processing, attention, and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 4. Attention and memory* (pp. 177–236). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Shivde, G., & Thompson-Schill, S. (2004). Dissociating semantic and phonological maintenance using fMRI. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 10–19.
- Tehan, G., & Humphreys, M. S. (1988). Articulatory loop explanations of memory span and pronunciation rate correspondences: A cautionary note. *Bulletin of the Psychonomic Society*, *26*, 293–296.
- Wagner, A. D., Pare-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left frontal cortex guides controlled semantic retrieval. *Neuron*, *31*, 329–338.
- 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 & Cognition*, *25*, 1256–1271.
- Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635–657.
- Warrington, E. K. (1981). Concrete word dyslexia. *British Journal of Psychology*, *72*, 175–196.
- Watkins, M. J. (1977). The intricacy of memory span. *Memory and Cognition*, *5*, 529–534.
- Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior*, *13*, 430–447.
- 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.
- Wickelgren, W. A. (1965). Short-term memory for phonemically similar lists. *American Journal of Psychology*, *78*, 567–574.

APPENDIX 1:
STIMULI FOR EXPERIMENT 1

Freq. = frequency counts from Carroll, Davies, and Richman (1971)

<i>Abstract</i>	<i>Freq.</i>	<i>Letter length</i>	<i>Phoneme length</i>	<i>Concrete</i>	<i>Freq.</i>	<i>Letter length</i>	<i>Phoneme length</i>
EXPANSE	4	7	6	GIRAFFE	5	7	5
CARNAGE	0.2	7	6	CUSTARD	0.3	7	7
CAUTION	6	7	5	LETTUCE	6	7	5
TREASON	1	7	6	LEOPARD	5	7	6
DECADE	6	6	5	MISSILE	7	7	5
RUMOUR	1	6	5	TANKER	1	6	6
LEADER	54	6	5	FLOWER	51	6	4
LESSON	39	6	5	BASKET	39	6	6
COMBAT	5	6	6	PARCEL	5	6	6
PENSION	1	7	6	PENDANT	0.1	7	7
STANZA	7	6	6	SUBWAY	6	6	5
MISSION	23	7	5	HARBOUR	27	7	6
MAGIC	50	5	5	WAGON	58	5	5
VOLUME	43	6	6	POCKET	45	6	5
TALENT	6.4	6	6	ELBOW	8	5	4
METHOD	97	6	5	GARDEN	106	6	5
MASTER	57	6	6	HANDLE	51	6	5
RHYTHM	40	6	4	BUTTER	37	6	5
VOYAGE	25	6	4	SALMON	29	6	5
MIXTURE	25	7	6	PASTURE	25	7	6
FUTURE	61	6	6	VALLEY	62	6	4
SYMBOL	58	6	6	BOTTLE	56	6	4
TORTURE	2	7	6	SERPENT	2	7	7
ANGER	14	5	5	JELLY	14	5	4
HAZARD	4	6	6	PUDDLE	3	6	4
DRAMA	11	5	5	BACON	12	5	4
FACTOR	36	6	6	ROCKET	38	6	5
NATIVE	37	6	5	NEEDLE	31	6	4
AFFECT	30	6	5	SPIDER	30	6	6
FEATURE	12	7	5	CHIMNEY	12	7	5
VIGOUR	4	6	5	ZEBRA	3	5	5
SAMPLE	17	6	5	BUTTON	21	6	5
SECTION	94	7	6	KITCHEN	85	7	5
CYCLE	17	5	4	TIGER	18	5	5
ADVICE	24	6	5	GARAGE	24	6	5
ANGLE	51	5	4	APPLE	53	5	3
COURAGE	37	7	5	CHICKEN	37	7	5
SEASON	65	6	4	FINGER	64	6	6
DAMAGE	23	6	5	PALACE	24	6	5
SPIRIT	40	6	6	RABBIT	35	6	5
MOMENT	140	6	6	WINDOW	151	6	5
BUDGET	4	6	5	SOCKET	3	6	5
MERIT	2	5	5	RAZOR	2	5	5

(Continued overleaf)

Appendix 1: (Continued)

<i>Abstract</i>	<i>Freq.</i>	<i>Letter length</i>	<i>Phoneme length</i>	<i>Concrete</i>	<i>Freq.</i>	<i>Letter length</i>	<i>Phoneme length</i>
COWARD	4	6	5	VELVET	8	6	6
HABIT	22	5	5	DONKEY	24	6	5
FAVOUR	19	6	5	JACKET	21	6	5
VALUE	80	5	5	SUGAR	79	5	5
JUSTICE	13	7	6	CABBAGE	20	7	5
MINUTE	116	6	5	FOREST	115	6	6
SECRET	47	6	6	PLANET	46	6	6
PROMISE	19	7	6	HAMMER	21	6	6
GLORY	8	5	5	DEVIL	7	5	5
IDEAL	9	5	4	KITTEN	15	6	5
PORTION	19	7	6	DRAGON	21	6	6
INSULT	2	6	6	FALCON	1	6	6
SUPPORT	58	7	6	BALLOON	35	7	5
DEBATE	5	6	5	CRADLE	6	6	5
MORAL	8	5	5	CORAL	13	5	5
DOMAIN	3	6	5	WALRUS	3	6	6
DUTY	26	4	5	MIRROR	33	6	5
INDEX	19	5	6	BULLET	9	6	5
HONOUR	37	6	4	CANOE	26	5	4
FLAVOUR	7	7	6	LEMON	5	5	5
IMPACT	11	6	6	COTTAGE	13	7	5
FASHION	21	7	5	RIFLE	21	5	5
SORROW	7	6	4	SADDLE	26	6	5
FOCUS	9	5	4	BUBBLE	9	6	5
LUMBER	21	6	6	CANDLE	21	6	6
ESTEEM	0.8	6	5	LOBSTER	4	7	7
GOSSIP	2	6	5	BANDAGE	2	7	6
Sum	1866.4	422	369		1895.4	423	364
Mean	26.66	6.03	5.27		27.08	6.04	5.20
SD	28.8	0.7	0.7		28.9	0.7	0.8

APPENDIX 2:
ADDITIONAL STIMULI FOR EXPERIMENT 2 AND 3

Freq. = frequency counts from Carroll, Davies, and Richman (1971). Phon. length = phoneme length

<i>Abstract</i>	<i>Freq.</i>	<i>Letter length</i>	<i>Phon. length</i>	<i>Syllable length</i>	<i>Concrete</i>	<i>Freq.</i>	<i>Letter length</i>	<i>Phon. length</i>	<i>Syllable length</i>
Experiment 2									
DESTINY	22	7	7	3	REFEREE	1	7	6	3
GLAMOUR	5	7	6	2	CUSHION	8	7	5	2
MERCY	20	5	5	2	MARBLE	21	6	5	2
RANGE	160	5	5	1	RIVER	165	5	4	2
TEMPER	12	6	5	2	PHEASANT	7	7	7	2
BARGAIN	7	7	6	2	TABLET	3	6	6	2
DECEIT	2	6	5	2	WARRIOR	5	7	6	2
GESTURE	32	7	5	2	TICKET	16	6	5	2
HASSLE	0	6	5	2	YOGHURT	0	7	6	2
OFFENCE	0	7	5	2	SCOOTER	0	7	5	2
Sum	260.0	63	54	20		226.0	65	55	21
Mean	26.0	6.3	5.4	2.0		22.6	6.5	5.5	2.1
SD	48.3	0.8	0.7	0.5		50.5	0.7	0.8	0.3
Experiment 3									
MARGIN	9	6	5	2	MARBLE	8	6	5	2
BURDEN	6	6	5	2	PIGEON	6	6	5	2
MANDATE	0.03	7	6	2	BABOON	0.03	6	5	2
CALORIE	0.8	7	6	2	REVOLVER	1	8	7	2
MALICE	2	6	5	2	STEEPLE	2	7	6	2
CONTRAST	23	8	8	2	GLASSES	24	7	6	2
ANGUISH	1	7	6	2	JIGSAW	1	6	5	2
REVENGE	3	7	6	2	PANTHER	3	7	6	2
FORTUNE	16	7	7	2	FEATHER	17	7	6	2
ABSENCE	8	7	6	2	PUPPET	6	6	5	2
Sum	68.8	68	60	20		68.0	66	56	20
Mean	6.8	6.8	6	2		6.8	6.6	5.6	2
SD	7.5	0.6	0.9	0.0		7.8	0.7	0.7	0.0

Downloaded by [Fondren Library, Rice University] at 12:53 24 August 2011