Neuropsychologia 47 (2009) 747–760

Contents lists available at ScienceDirect

Neuropsychologia

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

Semantic memory is key to binding phonology: Converging evidence from immediate serial recall in semantic dementia and healthy participants

Paul Hoffman[a,*], Elizabeth Jefferies[b], Sheeba Ehsan[a], Roy W. Jones[c], Matthew A. Lambon Ralph[a]

[a] Neuroscience and Aphasia Research Unit, School of Psychological Sciences, Zochonis Building, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[b] University of York, UK
[c] RICE – Research Institute for the Care of Older People, Royal United Hospital, Bath, UK

ARTICLE INFO

Article history:
Received 3 September 2008
Received in revised form 17 November 2008
Accepted 2 December 2008
Available online 9 December 2008

Keywords:
Semantic binding
Interactive activation
Immediate serial recall
Serial position effects

ABSTRACT

Patients with semantic dementia (SD) make numerous phoneme migration errors when recalling lists of words they no longer fully understand, suggesting that word meaning makes a critical contribution to phoneme binding in verbal short-term memory. Healthy individuals make errors that appear similar when recalling lists of nonwords, which also lack semantic support. Although previous studies have assumed that the errors in these two groups stem from the same underlying cause, they have never been directly compared. We tackled this issue by examining immediate serial recall for SD patients and controls on "pure" word lists and "mixed" lists that contained a mixture of words and nonwords. SD patients were equally poor at pure and mixed lists and made numerous phoneme migration errors in both conditions. In contrast, controls recalled pure lists better than mixed lists and only produced phoneme migrations for mixed lists. We also examined the claim that semantic activation is critical for words in the primacy portion of the list. In fact, the effect of mixed lists was greatest for later serial positions in the control group and in the SD group recall was poorest towards the ends of lists. These results suggest that mixing nonwords with words in healthy participants closely mimics the impact of semantic degradation in SD on word list recall. The study provides converging evidence for the idea that lexical/semantic knowledge is an important source of constraint on phonological coherence, ensuring that phonemes in familiar words are bound to each other and emerge together in recall.

1. Introduction

Influential models of verbal short-term memory (STM) hold that the meanings of words make a critical contribution to short-term retention (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Martin & Gupta, 2004; Martin, Lesch, & Bartha, 1999; Patterson, Graham, & Hodges, 1994). Though the details of these models vary, they share the view that the interaction of long-term semantic representations with temporary phonological activation is key to our recall of word sequences. Support for this proposal comes from two main sources. In studies of healthy subjects, recall is sensitive to the semantic properties of the material to be remembered (e.g., Poirier & Saint-Aubin, 1995; Walker & Hulme, 1999) and in neuropsychological studies, impairments to semantic knowledge are associated with particular deficits in word list recall (e.g., Martin & Saffran, 1997; Patterson et al., 1994). It is generally held that these two approaches – variation in the semantic properties of the *stimuli* in healthy participants vs. variation in the semantic abilities of the *patients* in neuropsychological studies – examine the same underlying influence of semantics on verbal STM. However, in the absence of detailed comparisons between patients and controls, it is difficult to assess whether this is the case. In this study, we addressed this issue by directly comparing the phonological errors predicted by one theory of the semantic contribution to STM (semantic binding; Patterson et al., 1994) in the recall of healthy subjects and patients with semantic dementia.

Semantic dementia (SD) is a neurodegenerative disorder in which bilateral atrophy centred on the anterior temporal lobes is associated with a progressive and pervasive deterioration of semantic memory (Hodges, Patterson, Oxbury, & Funnell, 1992; Snowden, Goulding, & Neary, 1989). The semantic impairment affects a wide range of verbal and non-verbal tasks (Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000; Bozeat, Lambon Ralph, Patterson, & Hodges, 2002; Warrington, 1975), suggesting that SD involves the degradation of an amodal store of semantic knowledge (Rogers et al., 2004). Perhaps the most striking feature of the disease is the degree to which other cognitive functions are preserved: visuospatial abilities, episodic memory, non-verbal reasoning, syntax and phonology all remain largely intact (Hodges et al., 1992). SD patients perform well on phonological discrimination tasks such

* Corresponding author. Tel.: +44 161 275 7336; fax: +44 161 275 2873.
E-mail address: paul.hoffman@manchester.ac.uk (P. Hoffman).

0028-3932/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2008.12.001

as minimal pairs or rhyming judgements, show normal effects of phonological similarity in immediate serial recall (ISR) and their span for digits and nonwords is preserved (Jefferies, Patterson, Jones, Bateman & Lambon Ralph, 2004; Jefferies, Jones, Bateman, & Lambon Ralph, 2005; Majerus, Norris, & Patterson, 2007).

Despite this preservation of the phonological aspects of STM, word recall is markedly impaired in SD patients. A number of studies have divided words into "known" and "degraded" categories for individual SD patients, based on performance in semantic tasks. Patients recall words they understand ("known" words) more accurately than words whose semantic representations are now degraded, provided that a large enough set of words is sampled from to construct the word lists (Funnell, 1996; Jefferies, Jones, Bateman, & Lambon Ralph, 2004; Jefferies, Patterson, et al., 2004; Jefferies et al., 2005; Knott, Patterson & Hodges, 1997, 2000; McCarthy & Warrington, 1987; Patterson et al., 1994; Warrington, 1975). In addition to this known-degraded difference, studies have focused on the breakdown of phonological coherence that occurs for poorly understood words (Jefferies, Hoffman, Jones & Lambon Ralph, 2008; Knott et al., 1997; Majerus et al., 2007; McCarthy & Warrington, 1987, 2001; Patterson et al., 1994). Segments of phonology from different words are often recombined in patients' responses to form new words or nonwords (e.g., MINT, RUG → "*r*int, *m*ug"). These errors have been variously termed phoneme migrations, phoneme recombinations or "blend" errors and are characteristic of list recall in SD. Similar errors have been uncovered in patients with other forms of brain damage who show a reliable distinction between known and semantically degraded words (Caza, Belleville & Gilbert, 2002; Forde & Humphreys, 2002).

This particular pattern of errors led Patterson et al. (1994) to propose that semantic representations have a stabilising effect on phonological activation during language and verbal STM tasks. Termed the "semantic binding" account, this states that there are two sources of constraint on phonological processing that support verbal STM. First, phonological activation occurs whenever a word is spoken or comprehended. The phonological elements of any given word are, therefore, strongly associated with one another because they are co-activated whenever that word is encountered. This makes them likely to be produced together in recall. Second, semantic representations are also activated whenever words are encountered. The spread of activation between semantics and phonology (required for word comprehension and production) serves to constrain patterns of phonological activation, helping to ensure that the correct configuration of phonological elements remains active. The feedback from semantic to phonological representations is particularly important for STM tasks which involve the retention of large amounts of phonology. This approach draws on connectionist models of language that hold that a range of linguistic tasks can be accomplished through the operation of the basic underlying systems of semantics, phonology and orthography (Patterson & Lambon Ralph, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996).

N. Martin and colleagues have taken a similar approach, adapting the interactive activation model of single word production of Dell and O'Seaghda (1992) to account for repetition in semantically impaired stroke patients (Martin & Gupta, 2004; Martin & Saffran, 1997). In their model, activation flows bidirectionally between semantics and phonology, via an intermediate layer of lexical units. In repetition, initial activation of phonological units spreads to lexical and semantic nodes and feedback from the lexical and semantic levels helps to maintain activation of the phonological nodes. In this sense, the approach is similar to the semantic binding account. The interactive activation approach also holds that a number of cycles of activation spread between levels are necessary before sufficient semantic activation occurs. As a consequence, semantic activation is thought to benefit words presented earlier in lists more than those presented later. The assumption that early list items depend on semantic activation to a greater extent than later ones accounts for the finding that stroke patients with semantic impairments typically show reductions in the primacy effect in ISR (Jefferies et al., 2008; Martin & Saffran, 1990, 1997). It is not clear whether similar effects are typical in SD. Although one study has reported reduced primacy effects in two patients (Reilly, Martin, & Grossman, 2005), others have found that SD patients show normal serial position effects or, in a few cases, robust primacy but poor accuracy for later positions (Jefferies, Jones, et al., 2004; Jefferies et al., 2008; Knott et al., 1997, see also Forde and Humphreys, 2002).

There is also evidence from healthy individuals for the role of lexical-semantic knowledge in STM. People recall lists of words more accurately than nonwords (Brener, 1940; Hulme, Maughan, & Brown, 1991), high frequency words more accurately than low frequency words (Gregg, Freedman, & Smith, 1989; Hulme et al., 1997) and highly imageable words more accurately than more abstract words (Romani, McAlpine, & Martin, 2008; Walker & Hulme, 1999). In particular, healthy subjects show effects of semantic knowledge on phonological coherence. When healthy adults repeat word lists, phoneme migration errors are rare and item order errors predominate (Aaronson, 1968; Bjork & Healy, 1974). However, they occur frequently in repetition of lists of nonwords, which lack semantic representations (Ellis, 1980; Jefferies, Frankish, & Lambon Ralph, 2006; Treiman & Danis, 1988) and similar results have been reported in normally developing children (Archibald & Gathercole, 2007).

It has been noted previously that the phoneme migration errors made by healthy individuals when recalling nonwords appear similar to those made by SD patients for words they no longer understand (e.g., Patterson et al., 1994). In both cases the errors appear to reflect the breakdown of phonological coherence in the absence of sufficient constraining semantic activation. There are also apparent qualitative similarities in the form of the phonological errors. When healthy subjects recall nonword lists, vowels are less likely to migrate than consonants (Ellis, 1980) and recombinations often involve the separation of an onset from its rime, with the vowel and consonant of the rime recalled as a unit (Treiman & Danis, 1988). These effects have also been observed in the word list recall of SD patients (Patterson et al., 1994). If it were the case that errors in both circumstances had a similar cause, this convergence of evidence from normal and impaired populations would provide strong support for the semantic binding hypothesis and related theoretical positions. However, to our knowledge the phonological errors made by SD patients in word list recall have never been directly compared to those made by healthy people in nonword recall. One reason for this is that the stimuli used in the two populations has differed. Studies of nonword recall in normal subjects have employed highly constrained experimental stimuli, typically using only consonant–vowel–consonant (CVC) nonwords, which allowed detailed quantitative analysis of the rate and type of phoneme migration errors (Jefferies et al., 2006; Treiman & Danis, 1988). In addition, these studies have avoided repeating the same phoneme within a list, avoiding ambiguity in the nature of phoneme movements. Conversely, studies of SD patients have often used tailor-made lists that contrast known and semantically degraded sets of words that differ for each individual patient (e.g., Jefferies, Jones, et al., 2004; Knott et al., 1997; McCarthy & Warrington, 2001; Patterson et al., 1994). Whilst this approach has been instrumental in revealing the conditions under which recall is impaired, it has resulted in small, restricted pools of words from which to concoct lists. As a consequence, lists presented to SD patients have featured words with varying phonological structures and containing repeating phonemes and the recall of SD patients has not been subjected to the same level of quantitative analysis seen in the studies of normal subjects.

The purpose of the present study was to directly compare word recall in a group of SD patients with recall in healthy subjects under conditions of reduced lexical-semantic support. Rather than simply giving healthy subjects lists of nonwords to repeat, we made use of a methodology that produces phoneme migrations for words *and* nonwords in healthy subjects, allowing SD patients and healthy volunteers to be directly compared for the same items. In a recent study, Jefferies et al. (2006) presented young, healthy individuals with mixed lists containing an unpredictable mixture of words and nonwords. The words in these mixed lists were recalled less accurately than those in pure lists containing only real words, suggesting that the presence of nonwords disrupted the integrity of the words they were presented alongside. There are two possible explanations for this effect. First, it may reflect disruption to a "redintegration" process that uses lexical knowledge to reconstruct decayed phonological traces (Hulme et al., 1991, 1997; Schweickert, 1993). On this view, the contribution of lexical/semantic knowledge occurs only at the point of recall, when degraded phonological information is "cleaned up" by matching it to entries in the lexical system (e.g., the degraded trace *e_eph_nt* could be cleaned up to produce *elephant*). Presenting words and nonwords in an unpredictable order would disrupt this process as the system would no longer know which degraded items should be reconstructed. Redintegration of degraded nonwords must be avoided as it would incorrectly convert them into words.

The second possibility is that the presence of nonwords disrupted the semantic binding process. On this view, the status of individual items affects the phonological stability of the entire list: the probability of a particular phoneme being recalled correctly depends both on (i) how strongly it is bound to the other phonemes in that item and (ii) on how strongly phonemes from other list items are bound to each other. Both of these factors have a strong influence on the degree of uncertainty about a target phoneme's position. Unfamiliar nonwords do not substantially benefit from lexical/semantic binding. As a consequence the positions of their phonemes are weakly represented in the phonological system and they are particularly prone to being recalled in incorrect positions. Word phonemes, by virtue of semantic activation, are more tightly bound to each other and less likely to migrate. However, when the phonological system is stressed by the presentation of a series of items, weakly bound nonword phonemes may intrude into word positions, disrupting the binding of words.

According to this proposal, the difference in the strength of binding for word and nonword phonemes is one of degree: although phoneme migrations are much more frequent for nonwords, word phonemes can migrate between items, even in pure word lists, especially when low frequency target words are presented (Jefferies et al., 2006). Consequently, words are not entirely immune from the destabilising effects of nonword phonemes. In line with this view, when presented with mixed lists, participants make some phoneme migration errors for words (although nonword phonemes still migrate more often). Therefore, mixed lists appear to disrupt the semantic binding process, making this paradigm a useful one in which to compare semantic binding effects in controls and individuals in SD.

In addition, by inducing phoneme migrations to words as well as nonwords, the mixed list method allows the effects of lexical frequency and imageability on phonological stability to be studied in healthy subjects. Jefferies et al. (2006) found fewer phoneme migrations in lists containing highly frequent and imageable words, providing further evidence that phonological coherence is dependent on the lexical and semantic status of items. Mixed lists might also provide a closer approximation to word list recall in SD because they mimic the varied nature of semantic degradation in the disease. Patients lose semantic knowledge gradually and often retain general knowledge of words despite losing more specific information about their meaning (Crutch & Warrington, 2006; Warrington, 1975). Therefore, lists presented to SD patients will contain some words that elicit partial semantic activation as well as some that do not – in contrast to nonword lists that rely purely on phonology.

In this study, we used the mixed list method to compare ISR in SD patients and age-matched controls directly. We employed pure word lists and mixed word/nonword lists, with the expectation that this second type of list would approximate the effects of semantic degradation in SD. All of the stimuli followed a CVC structure and phoneme repetitions within lists were avoided, allowing all phoneme migrations to be traced. To permit this design, we did not manipulate known vs. degraded status of the word targets for the SD patients: all of the patients received exactly the same lists as controls. We did, however, manipulate two variables known to influence comprehension and ISR in SD: frequency and imageability. As noted earlier, the ISR of healthy subjects is sensitive to both the frequency and imageability of the words presented. In SD, an exaggerated effect of frequency is found in ISR (Jefferies et al., 2008; Knott et al., 1997, 2000; Majerus et al., 2007; McCarthy & Warrington, 2001), which reflects the susceptibility of low frequency words to the disease. Semantic deterioration in SD is strongly graded by frequency, perhaps because less frequent words are less strongly represented in the semantic system to begin with, as well as being encountered less often in the course of the disease (Bozeat et al., 2000; Funnell, 1995; Jefferies, Patterson, Jones & Lambon Ralph, in press-b; Lambon Ralph, Graham, Ellis, & Hodges, 1998). Since the meanings of low frequency words are more likely to be degraded, these items are less likely to benefit from semantic binding. Similarly, highly imageable words may be more robust to semantic degradation than abstract items because they are thought to possess more detailed semantic representations (Jones, 1985; Plaut & Shallice, 1993). The majority of SD patients show marked positive effects of imageability on comprehension (Jefferies et al., in press-b) and better immediate recall of imageable than abstract words (Jefferies et al., 2008; Majerus et al., 2007; although this was not the case in Reilly et al., 2005).

An additional aim of the study was to test the prediction of Martin and Gupta (2004) that semantic factors have a larger effect on early items in list repetition. As discussed earlier, previous studies of SD have provided mixed support for the idea that semantic degradation in the disease selectively affects earlier serial positions. Evidence from healthy subjects is similarly equivocal. There is some evidence that repetition of nonword lists yields a serial position curve with a normal primacy effect but a reduced *recency* effect (Gupta, Lipinski, Abbs, & Lin, 2005; Hulme, Stuart, Brown, & Morin, 2003), which appears in direct contradiction to the idea that semantic information supports early items and that later items are retained on the basis of their phonology. Hulme et al. (1997) also found that frequency effects were largest in words presented later in lists and smallest for the primacy portion of the curve, which is also difficult to account for if it is assumed that high frequency words possess more robust semantic representations. However, interpretation of these effects is difficult because high and low frequency words, as well as nonwords, differ in the strength of their lexical and phonological representations in addition to their semantic properties. Regarding the effects of imageability, a variable more closely linked to semantic knowledge, several studies have found that the advantage for concrete words is absent for the final one or two words in lists but stable across all other positions (Romani et al., 2008; Allen & Hulme, 2006; Walker & Hulme, 1999). This provides some support for the notion that semantic factors are most critical for the primacy portion of the curve, although Romani et al. have noted that this theory predicts a monotonic decrease in imageability effects from the first position onwards, which was not consistent with their data.

In the present study, we compared the effects of semantic degradation in SD with that of disruption to semantic support through the presentation of mixed lists. ISR data were analysed at a number of levels to determine to what extent recall performance, phonological errors and serial position effects were affected by the presence of SD and by the presentation of nonword stimuli thought to disrupt semantic binding.

1. Item-level accuracy was analysed to determine the extent to which list composition (i.e., mixed vs. pure lists) affected recall in each group. We expected SD patients to be affected less strongly than healthy controls by list composition because the mixed list technique is thought to lessen lexical-semantic contributions to verbal STM for words (Jefferies et al., 2006).
2. Error rates at the whole item level were also analysed, with the prediction that SD patients would make more item intrusion errors for pure word lists and that the error pattern on mixed lists for controls would be similar to that of SD patients for pure lists.
3. Serial position curves were examined, to determine whether the primacy portion of the curve was selectively affected, either by the effects of SD or by the presentation of mixed lists, in line with the Martin and Gupta's (2004) hypothesis.
4. At the level of individual phonemes, we calculated the rate of phoneme migration errors. We expected patients to make frequent migrations for both list types, due to the breakdown of constraining semantic activation, but that controls would only make these errors for mixed lists in which phonological coherence was disrupted by the presence of poorly bound nonwords. We also examined the fate of phonemes from different syllabic positions, to determine whether the errors made by controls and patients were similar at this fine-grained level.

2. Method

2.1. Case descriptions

Six patients with a clinical diagnosis of SD were recruited from Bath and Liverpool, UK. Details of the cases and their scores on neuropsychological tests are given in Table 1. They fulfilled all of the published criteria for SD (Hodges et al., 1992) and MRI revealed focal atrophy of the inferior and lateral aspects of the anterior temporal lobes in every case (except KI for whom scanning was unavailable). Patients were tested on the following four semantic tasks, with every patient showing abnormal performance on every task: (a) Pyramids and Palm Trees test (Howard & Patterson, 1992). This test of associative semantic knowledge comprised picture-picture matching and word-word matching versions. (b) Picture naming for 64 items (Bozeat et al., 2000). Black-and-white line drawings from the Snodgrass and Vanderwart set (1980) were presented individually for naming. There were an equal number of living and man-made objects. (c) Word–picture matching for the same 64 items (Bozeat et al., 2000). The name of each item was spoken aloud and the patients attempted to select the corresponding object from an array of ten pictures. The nine distractor pictures belonged to the same category as the target. (d) Category fluency: patients generated as many exemplars from a specified semantic category as possible in 1 min. There were four living and four man-made categories (animals, fruits, birds, dog breeds, household items, tools, vehicles and types of boat).

In contrast to their semantic impairments, the patients scored normally on the Rey figure copy (Rey, 1941), indicating preserved visuo-spatial skills. Four of the six cases displayed intact non-verbal reasoning on Raven's (1962) coloured progressive matrices, although performance was impaired in the remaining two. Digit spans in the group were good (with just one patient, NH, falling outside the normal range by one digit), suggesting preserved STM for numbers which are well-understood by SD patients (Cappelletti, Butterworth, & Kopelman, 2001; Jefferies, Patterson, et al., 2004). Since STM depends heavily on phonological processing, we assessed phonological skills in order to determine to what extent our patients' STM deficits could be attributed to poor phonological processing. There were four tasks. (a) Minimal pairs test from PALPA (Kay, Lesser, & Coltheart, 1992). Patients made same/different judgement to pairs of auditory items that were identical (e.g., miv–miv) or that differed by one phoneme (e.g., miv–niv). The remaining three tasks were taken from Patterson and Marcel (1992). (b) Phoneme manipulation. There were two versions of this task. In the phoneme subtraction version, the patients were asked to delete the first sound of an item and say what remained (e.g., vale → ale). In the phoneme addition version, patients were asked to add a phoneme to the rhyme of an item (e.g., ale → vale). All of the items were monosyllabic and the same 48 items were used in the two versions of the task. Words and nonwords featured as stimuli and target responses equally often. (c) Rhyme judgement. This task required patients to judge whether or not two spoken words rhymed (e.g., white–kite). There were 48 trials. Half of the 24 non-rhyming trials were composed of two phonologically similar words (e.g., tick–tin) (d) Rhyme production. A word was presented and the patient produced a rhyming word. Since our patients' word knowledge was impoverished, rhyming nonword responses were classed as correct in this task.

These tasks suggested mild phonological weakness in some patients (since controls tend to perform them at ceiling). This was most evident on the rhyme judgement task, which does place some demands on semantics. Patients needed to understand the concept of "rhyming" in a sufficiently detailed manner to distinguish rhyming pairs from other phonologically similar pairs. In any case, the phonological weakness revealed by these tasks was mild relative to the consistent and pervasive deficits apparent on the semantic tasks.

2.2. Control participants

In addition to the SD cases described above, 11 healthy controls were tested. They were aged between 51 and 72 (mean = 64.9) and were matched to the patients for age ($t(15) = 1.17$, *ns*) and educational level ($t(15) = 1.25$, *ns*).

2.3. Design and materials

Two sets of lists taken from Jefferies et al. (2006) were presented auditorily for ISR. All lists contained five CVC items. All of the items in pure lists were real words but mixed lists contained an unpredictable mixture of real words and nonwords (see Appendix A). There were 60 mixed lists. The ratio of words to nonwords in these lists varied: 20 lists contained one word and four nonwords (1W:4N lists), 20 contained two words and three nonwords (2W:3N lists) and 20 contained three words and two nonwords (3W:2N lists). Words and nonwords appeared in each serial position an equal number of times. Words in the lists were also varied for frequency and imageability in an orthogonal manipulation. Mean frequency from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993) was 179 counts per million for the high frequency words (range = 51–656) and 6 counts per million for the low frequency words (range = 1–13). Mean imageability from the MRC database (Coltheart, 1981) was 602 for high imageability words (range = 573–659) and 442 for low imageability words (range = 340–501). All of the words within a list were taken from a single frequency by imageability condition.

Participants were presented with 40 pure lists. These were created by taking the 2W:3N and 3W:2N mixed lists and replacing the nonwords in these lists with real words. Therefore, 100 words from the pure lists also appeared in the mixed lists; all that differed between the conditions was the list context (whether the items were presented amongst nonwords or real words). Words in these lists obeyed the same frequency and imageability constraints as the mixed lists: ten lists belonged to each frequency by imageability condition. For each set of lists, there was no repetition of phonemes within a list, and no repetition of items within the set.

2.4. Procedure

Items were recorded by a female speaker and were presented at a rate of one item per second. Presentation was controlled using SuperLab software (Cedrus). A red exclamation mark on the screen preceded each trial. After the list had been presented, this was replaced with a blue question mark, prompting recall of the list. Participants were told in advance that mixed lists would contain both words and nonwords, and pure lists only words. They received the mixed and pure lists in separate sessions, with an intervening period of at least 1 week. Half were presented with the pure lists in the first session and half with the mixed lists. Each session was preceded by four practice trials. Participants were instructed to recall the items in the same serial order as they were heard and were encouraged to try to recall all of the items, even if unsure. As our primary interest was in the phonological stability of items and not their serial order, we scored items as correct irrespective of whether they were in the correct serial position.

3. Results

Results were analysed first at the item level and then at the phoneme level. At the item level, we present (1) accuracy for the pure and mixed lists separately; (2) a direct comparison of recall for the words presented in both pure and mixed lists; (3) the effect of serial position and (4) item-level errors. Following this, we examine (5) rates of phoneme migration errors, (6) other phonemic errors and (7) phoneme recall as a function of syllabic position. We focus on group mean data throughout; however, a summary of performance for individual patients is given in Table 2.

Table 1
Patient details and background neuropsychology.

		EK	GE	JT	KI	NH	SJ	Controls	
	Max							Mean	s.d.
Age		60	52	66	65	66	60	64.9	5.9
Sex		F	M	M	M	F	F		
School leaving age		15	16	16	14	16	16	17.2	1.4
Years post onset		5	5	4	4	2	3		
Semantic									
Pyramids and Palm Trees[a]									
Words	52	35[b]	34[b]	31[b]	35[b]	36[b]	42[b]	51.2	1.4
Pictures	52	30[b]	35[b]	35[b]	31[b]	36[b]	48[b]	51.1	1.1
Picture naming	64	18[b]	14[b]	6[b]	15[b]	26[b]	30[b]	62.3	1.6
Word–picture matching	64	39[b]	34[b]	34[b]	36[b]	37[b]	59[b]	63.7	0.5
Category fluency (8 categories)	–	27[b]	7[b]	9[b]	27[b]	34[b]	31[b]	113.9	12.3
General									
Coloured progressive matrices[c]	36	33	33	36	21	20	34	–	–
Rey figure copy[d]	36	36	30	34	35	26[b]	33	34.0	2.9
Digit span[e]									
Forwards	–	7	7	8	8	4[b]	5	6.8	0.9
Backwards	–	4	4	4	5	3	3	4.7	1.2
Phonological									
Minimal pairs[f]									
Words	72	71	72	70	69	NT	72	70.1	3.4
Nonwords	72	71	65[b]	70	71	NT	70	70.9	2.9
Phoneme segmentation[g]									
Addition	48	37	44	44	38	NT	45	–	–
Subtraction	48	44	43	46	45	NT	47	–	–
Rhyme judgement[g]	48	42	47	46	40	NT	43	–	–
Rhyme production[g]	24	23	24	20	23	NT	22	–	–

[a] Howard and Patterson (1992).
[b] Denotes abnormal performance.
[c] Raven (1962).
[d] Rey (1941).
[e] Wechsler (1987).
[f] From the PALPA battery (Kay et al., 1992).
[g] Patterson and Marcel (1992).

3.1. Item accuracy

3.1.1. Pure lists

Fig. 1 shows response accuracy for pure lists and mixed lists, with words and nonwords within the mixed lists shown separately. Results for the pure lists were analysed using a $2 \times 2 \times 2$ mixed ANOVA which included frequency and imageability as within-group factors and participant group as a between-groups measure. The outcome of this analysis is given in Table 3. As expected, controls were much more accurate than SD patients. While both groups were influenced by frequency and imageability, SD patients were more sensitive than controls to frequency reflecting the more severe degradation of low frequency concepts in the disease.

Table 2
Summary of individual patient performance.

	EK	GE	JT	KI	NH	SJ	Controls	
							Mean	s.d.
Items recalled (%)								
Pure – words	33	33	35	44	27	24	79	11
Mixed – words	26	22	28	49	32	28	65	12
Mixed – nonwords	14	7	25	14	12	14	37	11
Phoneme migrations (% of all phonemes recalled)								
Pure – words	28	19	21	13	19	30	2	1
Mixed – words	28	30	26	9	14	29	6	3
Mixed – nonwords	38	38	25	20	29	36	13	6
Items recalled by serial position (pure lists)								
First item	45	65	50	50	55	50	90	5
Middle items[a]	23	27	33	43	24	18	76	14
Final item	53	18	23	43	8	18	77	15
Items recalled by serial position (mixed lists)								
First item	28	35	43	35	43	45	66	14
Middle items[a]	8	8	24	27	19	15	44	11
Final item	40	5	17	27	0	8	42	14

[a] Mean of positions two, three and four.

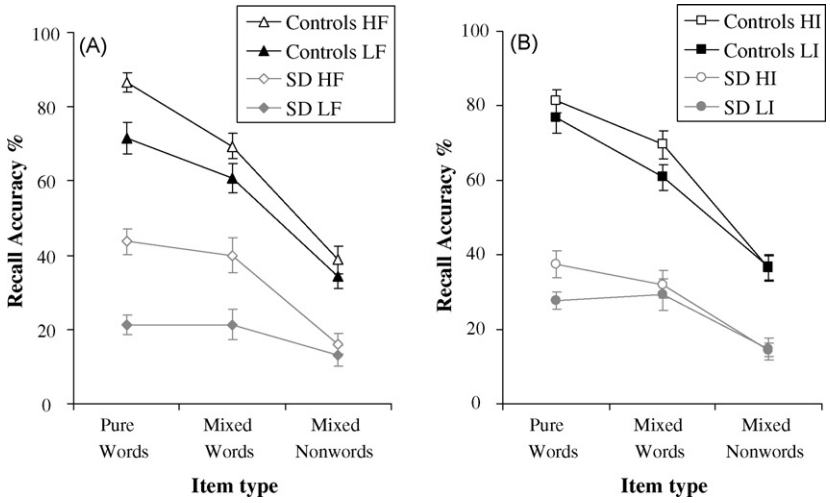

Fig. 1. Item accuracy for pure and mixed list recall. Item accuracy for (A) high vs. low frequency lists and (B) high vs. low imageability lists. HF = high frequency; LF = low frequency; HI = high imageability; LI = low imageability. Bars indicate one standard error of mean.

3.1.2. Mixed lists

Analysis of the mixed lists considered frequency, imageability and participant group (SD vs. normal participants), as above, and also the lexicality of the item (i.e., whether it was a word or nonword; see Table 3 for details). In the analyses presented below, we considered the effects of word frequency and imageability on *word* recall. We also assessed whether the frequency and imageability of words influenced recall of the *nonwords* they were presented amongst. Therefore, where effects of frequency and imageability on nonwords are reported, this refers not to the frequency and imageability of the nonwords themselves but rather of the words they were presented alongside.

Overall, words were recalled more accurately than nonwords and this lexicality effect was larger for the control group than the patients, consistent with the idea that SD patients show a reduction in semantic binding. Frequency and imageability affected the words to a greater extent than the nonwords they were presented with. Unlike the pure lists, in these lists the frequency effect was not larger in the patients than in controls.

We also considered the effect of the composition of the mixed lists (i.e., the ratio of words to nonwords in the list; see Fig. 2). This was examined using a $2 \times 2 \times 3$ ANOVA (results in Table 3) which included lexicality, participant group and list composition (1W:4NW, 2W:3NW, 3W:2NW). This revealed a main effect of list composition as well as interaction between composition and group. The healthy participants showed better performance for both word and nonword stimuli when recalling lists containing a greater number of words. Bonferroni post-hoc tests indicated that controls

Table 3
Analyses of item recall accuracy.

Effect	Statistic	Explanatory notes
Pure lists		
Group	$F(1,15) = 83.3$, $p < .0001$	Controls > SD
Frequency	$F(1,15) = 131$, $p < .0001$	High frequency > low frequency
Imageability	$F(1,15) = 14.1$, $p < .005$	High imageability > low imageability
Frequency × group	$F(1,15) = 5.1$, $p < .05$	Larger frequency effect in SD
Imageability × group	$F(1,15) = 1.9$, $p > .1$	Both groups equally sensitive to imageability
Mixed lists		
Group	$F(1,15) = 35.2$, $p < .0001$	Controls > SD
Lexicality	$F(1,15) = 103$, $p < .0001$	Words > nonwords
Frequency	$F(1,15) = 27.9$, $p < .0001$	High frequency > low frequency
Imageability	$F(1,15) = 6.6$, $p < .05$	High imageability > low imageability
Lexicality × group	$F(1,15) = 7.9$, $p < .02$	Larger lexicality effect in controls
Lexicality × frequency	$F(1,15) = 13.0$, $p < .005$	Frequency effect larger for words (but significant for words and nonwords)
Lexicality × imageability	$F(1,15) = 12.8$, $p < .005$	Imageability effect larger for words (no effect for nonwords)
Mixed lists – number of words		
List composition	$F(2,30) = 6.0$, $p < .01$	Higher accuracy for lists containing more words
Composition × group	$F(2,30) = 6.3$, $p = .005$	Controls affected by list composition; SD patients were not
Composition × lexicality	$F(2,30) = 3.4$, $p < .05$	List composition affected words more than nonwords
Pure vs. mixed lists		
Group	$F(1,15) = 408$, $p < .0001$	Controls > SD
Frequency	$F(1,15) = 131$, $p < .0001$	High frequency > low frequency
Imageability	$F(1,15) = 18.2$, $p = .001$	High imageability > low imageability
List type	$F(1,15) = 22.9$, $p < .001$	Pure > mixed
List type × group	$F(1,15) = 8.4$, $p < .05$	No list type effect in SD
List type × frequency	$F(1,15) = 5.4$, $p < .05$	Larger frequency effect in pure lists
List type × frequency × imageability	$F(1,15) = 5.3$, $p < .05$	Pure: larger imageability effect on high frequency lists. Mixed: larger imageability effect on low frequency lists

All main effects and significant interactions are reported. Explanatory notes are based on post-hoc tests not reported in full here. Analysis of pure vs. mixed lists focused on words presented in both conditions.

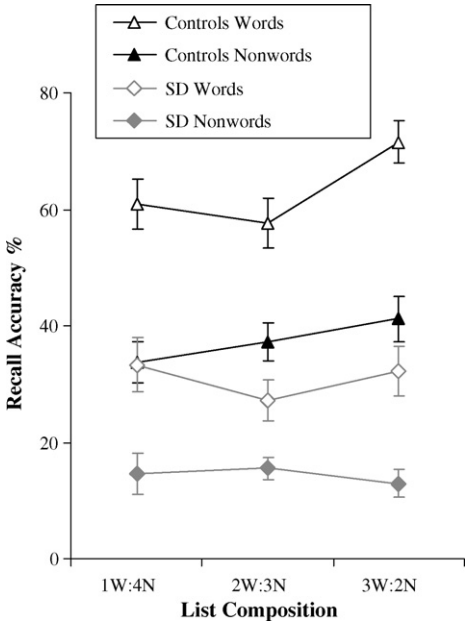

Fig. 2. Item accuracy for mixed lists. Bars indicate one standard error of mean.

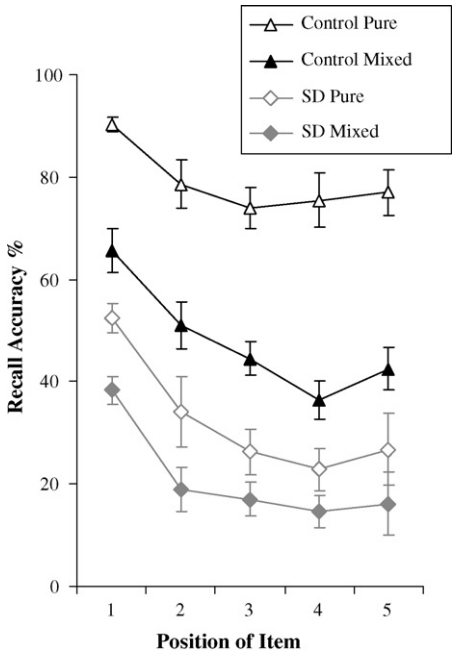

Fig. 3. Serial position effects. Bars indicate one standard error of mean.

recalled words in 3W:2N lists more accurately than those in lists containing fewer words (3W:2N vs. 2W:3N: $t(10) = 5.37$, $p < .001$; 3W:2N vs. 1W:4N: $t(10) = 2.86$, $p = .05$). Nonwords in 3W:2N lists were also more likely to be recalled than those in 1W:4N lists ($t(10) = 3.07$, $p < .05$). In contrast, list composition had no effect for the SD patients (all $t < 2.03$). It appeared therefore that for the controls, word targets benefited from lexical/semantic binding and the phonological stability of the entire list improved as a result. This effect was absent from the patients' recall.

3.1.3. Comparison of words in pure and mixed lists

This analysis only considered words that appeared in both mixed and pure lists, allowing exactly the same targets to be compared in different list contexts. Accuracy levels on these words resembled those in Fig. 1 and are not reproduced here. Table 3 shows an analysis of the effects of frequency, imageability, list type (pure vs. mixed) and participant group. The main effects of frequency, imageability and group from the previous analyses were replicated. There was also an effect of list type: words were more likely to be recalled accurately in pure lists than mixed lists. Most importantly, controls recalled words more accurately in pure lists than in mixed lists but list type had no effect on the recall of the SD patients.

3.2. Serial position effects

Fig. 3 shows serial position curves for the pure and mixed lists. A 5 × 2 × 2 (serial position × list type × group) ANOVA revealed significant main effects of list type ($F(1,15) = 261$, $p < .0001$) and serial position ($F(4,60) = 21.1$, $p < .0001$). While there was no interaction between serial position and group ($F(4,60) < 1$), there was a significant three-way interaction between group, serial position and list type ($F(4,60) = 2.63$, $p < .05$). Post-hoc tests confirmed that the serial position curve differed between pure and mixed lists for healthy participants ($F(4,40) = 3.12$, $p < .05$). For both types of list, controls showed a pronounced primacy effect and a smaller recency effect. However, performance declined more steeply across serial positions for the mixed lists, such that the difference between pure and mixed list recall was larger for later positions. Conversely, SD patients exhibited parallel serial position curves for the two types of list, showing no interaction between list type and serial position ($F(4,20) < 1$). Patients showed the same steep decline across serial positions for pure word lists that was evident when controls recalled mixed lists (a direct comparison of the patients' pure list performance with the controls' mixed list recall revealed no position by group interaction: $F < 1$), indicating that the recency portion of the curve was most susceptible to semantic degradation. When individual patients' data was considered, only one patient out of six showed better performance for the final word to be presented (EK). All of the other patients recalled the first item more accurately than any others in the list (see Table 2).

3.3. Errors

Errors are shown in Table 4, divided into item errors, which included any responses not present in the target list, omissions and serial order errors. Item errors largely accounted for the SD patients' poor performance on both the mixed and pure lists. A 2 × 2 (list type × group) ANOVA was conducted on each error type. SD patients made more item errors than controls overall ($F(1,15) = 81.8$, $p < .0001$), item errors were more common for mixed lists than pure lists ($F(1,15) = 68.9$, $p < .0001$) and there was an interaction

Table 4
Errors at the whole item level.

	Pure lists			Mixed lists		
	Item errors	Omissions	Order errors	Item errors	Omissions	Order errors
SD	57 (13)	11 (14)	17 (7)	69 (13)	10 (13)	12 (5)
Control	11 (5)	10 (9)	7 (6)	37 (10)	15 (12)	5 (4)

Item errors and omissions are expressed as a percentage of items presented. Order errors are expressed as a percentage of the total items recalled irrespective of position. Standard deviations in parentheses.

between group and list type ($F(1,15) = 9.51$, $p < .01$). For controls, there was a marked increase in the number of item errors to mixed lists compared with pure lists; in SD this increase was smaller. The prevalence of item errors in SD is suggestive of a lack of phonological coherence that also characterised healthy participants' recall of mixed lists. Indeed, many of the item errors made by the patients were "blends" of phonology from separate list items. These errors are analysed in more detail in the following section. The prevalence of omissions was similar across groups and list types: there was no significant effect of list type ($F(1,15) = 1.98$, *ns*) or group ($F(1,15) = .17$, *ns*) and no interaction ($F(1,15) = 2.53$, *ns*). Order errors were more common in the SD patients than in controls ($F(1,15) = 6.63$, $p < .05$) but these errors were not influenced by list type ($F(1,15) = 2.24$, *ns*) and there was no interaction ($F(1,15) = .72$, *ns*).

Item errors were also classified according to whether the erroneous response was a word or a nonword. Controls produced words for 89% of their errors in pure lists but only 69% in mixed lists. This may reflect a strategic response monitoring process that prevented nonword responses in the pure list condition. Nevertheless, a strong lexical bias remained in the mixed condition. In the SD group, 75% of errors were words on the pure lists, compared with 68% in the mixed condition. A 2 × 2 (list type × group) ANOVA revealed main effects of group ($F(1,15) = 7.46$, $p < .02$) and list type ($F(1,15) = 36.5$, $p < .0001$) as well as a significant interaction ($F(1,15) = 8.84$, $p < .01$). This indicates that the SD group were more likely than controls to produce nonwords for the pure word lists. On mixed lists, controls and patients produced nonwords at a similar rate.

3.4. Item-level summary

There were striking similarities at the item-level between ISR under conditions of pathologically degraded semantic knowledge in patients with SD and artificial disruption to lexical/semantic binding through stimulus manipulation in healthy participants. Results from the control group replicated many of the findings of Jefferies et al. (2006), which used the same materials in a group of younger healthy individuals. The mixing of words with nonwords had a detrimental effect on word recall. This was evident in the effect of mixed list composition – recall accuracy improved as the ratio of words to nonwords shifted in favour of words – and is consistent with the idea that words benefit from additional lexical/semantic constraints not available to nonwords. SD patients produced a different pattern of performance. The presence or absence of nonwords had no effect on their ability to remember words, nor were they affected by the number of nonwords present in the mixed lists. In addition, healthy participants produced rather different serial position curves for the two list types, with the mixed lists promoting a steeper decline across serial positions and a particularly pronounced primacy effect. This contrast was absent from the SD group, who showed sharp declines across serial positions regardless of whether lists contained nonwords. These findings are not consistent with the notion that semantic activation is particularly crucial for recall of early list items (Martin & Gupta, 2004; Martin & Saffran, 1997).

SD patients showed strong positive effects of word frequency, presumably because highly familiar concepts are more resistant to semantic deterioration (Bozeat et al., 2000; Funnell, 1995; Lambon Ralph et al., 1998). Finally, controls were more likely to produce words when they made errors in the pure list condition, making nonword errors more frequently in mixed lists. This tendency was diminished in the patients, who made nonword errors at equivalent levels for both list types. However, the most striking feature of the SD cases' performance was their strong tendency to distort the phonology of list items and blend phonemes from different words together. These phoneme migrations are examined in more detail in the next section.

Table 5
Analyses of phoneme migration rates.

Effect	Statistic	Explanatory notes
Pure lists		
Group	$F(1,15) = 111$, $p < .0001$	SD > controls
Frequency	$F(1,15) = 68.2$, $p < .0001$	Low frequency > high frequency
Imageability	$F(1,15) = 14.4$, $p = .002$	Low imageability > high imageability
Frequency × group	$F(1,15) = 46.5$, $p < .0001$	Larger frequency effect in SD
Mixed lists		
Group	$F(1,15) = 146$, $p < .0001$	SD > controls
Lexicality	$F(1,15) = 42.7$, $p < .0001$	Nonwords > words
Frequency	$F(1,15) = 17.4$, $p = .001$	Low frequency > high frequency
Imageability	$F(1,15) < 1$	No imageability effect
Frequency × group	$F(1,15) = 3.7$, $p = .07$	Trend towards larger effect in SD
Mixed lists – number of words		
List composition	$F(2,30) = 2.9$, $p = .07$	Trend toward more migrations in lists containing fewer words
Composition × group	$F(2,30) = 3.6$, $p < .05$	Controls affected by list composition; SD patients were not
Pure vs. mixed lists		
Group	$F(1,15) = 82.9$, $p < .0001$	SD > controls
Frequency	$F(1,15) = 43.2$, $p < .0001$	Low frequency > high frequency
Imageability	$F(1,15) = 1.2$, *n.s.*	No imageability effect
List type	$F(1,15) = 3.5$, $p = .08$	Trend toward mixed > pure
Frequency × group	$F(1,15) = 16.6$, $p = .001$	Larger frequency effect in SD
List type × group	$F(1,15) = 2.3$, *n.s.*	Effect did not differ between groups
List type × frequency × group	$F(1,15) = 9.9$, $p < .01$	SD: Weak trend toward mixed > pure for high frequency lists. No difference in effect for controls

All main effects and significant interactions are reported. Explanatory notes are based on post-hoc tests not reported in full here. Analysis of pure vs. mixed lists focused on words presented in both conditions.

3.5. Phoneme migrations

Our classification of phonemic errors was as follows. For each phoneme in the target list, we checked whether the phoneme was part of the participant's response. If the phoneme was absent, it was not considered further. If it was present and was recalled in the correct position in the list, it was classed as correct, and if it was recalled in an incorrect list position as a result of a whole item migration (i.e., all three phonemes making up a word/nonword migrated together) it was also classed as correct, as this error preserved the phonological structure of the item. However, any other cases where the phoneme had been recalled in the wrong place in the list were counted as phoneme migrations. These errors always resulted in the formation of new word or nonword. Phoneme migrations were expressed as a percentage of the number of phonemes recalled in any position in the list.

Migrations on pure lists were considered with respect to stimulus frequency and imageability. For mixed lists, we considered frequency, imageability and lexicality and, in a separate analysis, list composition (ratio of words to nonwords). Migrations involving the words presented in both types of list were also directly compared. The results of these analyses are presented in Table 5.

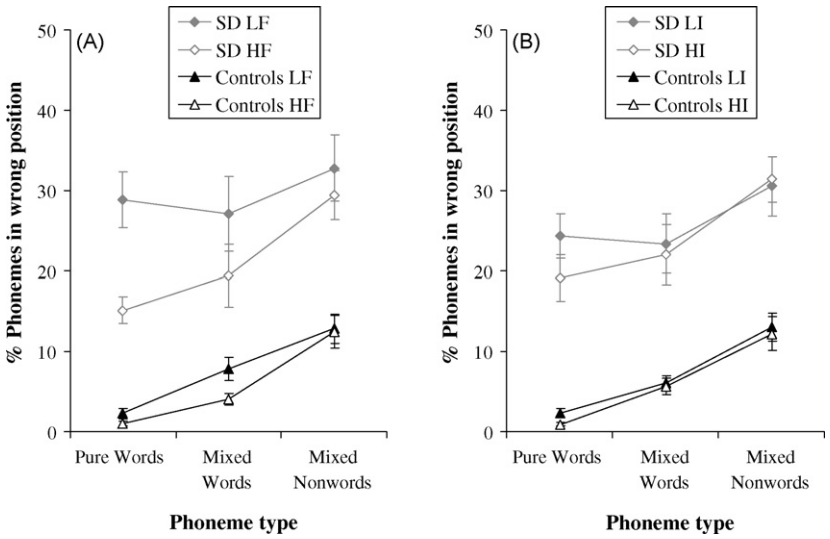

Fig. 4. Rate of phoneme migrations. Phoneme migrations for (A) high vs. low frequency lists and (B) high vs. low imageability lists. HF = high frequency; LF = low frequency; HI = high imageability; LI = low imageability. Bars indicate one standard error of mean.

3.5.1. Pure lists

The phoneme migration rates for each frequency and imageability condition are shown in Fig. 4. As expected, migrations for words in pure lists occurred much more frequently in the SD group than in controls: just 2% of phonemes were recalled in the wrong position in the control group compared with 22% in the patients. Phoneme migrations were also more likely to occur on low frequency and low imageability lists. This is consistent with the view that imageable items benefit from stronger semantic binding due to their more detailed semantic representations. More frequent words might also benefit from enhanced semantic binding as well as participants' greater familiarity with the phonological forms. However, the SD patients showed a much larger effect of frequency, consistent with a semantic contribution to the effect.

3.5.2. Mixed lists

The SD patients made more phoneme migration errors than the healthy participants on these lists. Phonemes originating in nonwords were more likely to migrate than the constituents of words. In the control group, the rate of migrations steadily decreased as the number of words in the list increased, indicating that the presence of more words improved the phonological stability of the list and prevented migrations for both words and nonwords (see Fig. 5). In the SD group, there was an apparent trend in the same direction for words while nonword phonemes appeared *more* likely to migrate in lists containing a higher proportion of words. However, neither of these effects approached statistical significance, reflecting the patients' insensitivity to the composition of lists.

3.5.3. Comparison of words in pure and mixed lists

This analysis considered only those phonemes originating in words that were present in both pure and mixed lists. As well as replicating the effects in the previous analyses, there was a tendency for migration errors to occur more frequently in mixed lists, although this effect did not differ between the two groups. There was also a frequency × group × list type interaction, perhaps because SD patients were more sensitive to frequency on the pure lists than the mixed lists, while controls showed a larger frequency effect on the mixed lists which elicited larger numbers of phoneme migrations.

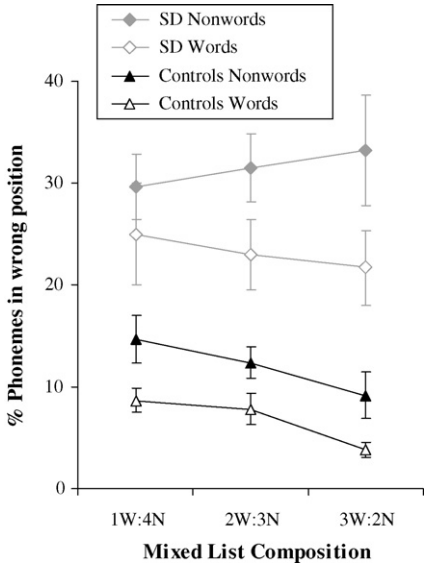

Fig. 5. Phoneme migrations in mixed lists. Bars indicate one standard error of mean.

3.6. Other phoneme errors

In addition to phoneme migrations, two other errors of commission were analysed at the level of individual phonemes. If a participant reported a phoneme that was not present in the presented list, this was classed as an intrusion error. If the same phoneme was produced more than once within a list, a repetition error was recorded. Rates of phoneme intrusions and repetitions are given in Table 6. A 2 × 2 (list type × group) ANOVA confirmed

Table 6
Phoneme intrusions and repetitions.

	Pure lists		Mixed lists	
	Intrusions	Repetitions	Intrusions	Repetitions
SD	15.2 (6.1)	11.2 (1.6)	18.2 (7.2)	13.4 (2.5)
Control	2.4 (1.4)	2.5 (1.4)	7.2 (2.1)	6.7 (1.9)

Errors expressed as a percentage of phonemes reported. Standard deviations in parentheses.

that the SD patients made more intrusion errors than the controls ($F(1,15) = 39.8$, $p < .0001$) and that intrusions were more common in mixed lists ($F(1,15) = 18.2$, $p = .001$). The two variables did not interact ($F(1,15) < 1$, *n.s.*). A similar picture emerged when repetitions were considered. Again, SD patients made more errors than controls ($F(1,15) = 97.2$, $p < .0001$) and errors were more common for mixed lists ($F(1,15) = 39.7$, $p < .0001$). Here, the list type by group interaction also approached significance ($F(1,15) = 3.92$, $p = .06$), suggesting a tendency for SD patients to be less affected by the presence of nonwords than controls.

It is also worth noting that while SD patients (and controls on mixed lists) showed disruption of both the order and identity of list phonemes, they almost always reproduced the correct structure of the list items. All presented items had a CVC structure and only 6% of the errors made by both the SD cases and healthy participants deviated from this structure. These errors most often involved adding an additional consonant to form a cluster in the onset or coda. Participants never gave multi-syllabic responses.

3.7. Effect of phoneme position

Phoneme recall was analysed as a function of syllabic position within items. Vowels (V) were recalled more accurately than initial (C1) and final consonants (C2), with the poorest recall for C2 phonemes (see Fig. 6). We directly compared words in mixed and pure lists, using a position (C1, V, C2) × list type × group ANOVA. This indicated a main effect of position ($F(2,30) = 37.6$, $p < .0001$) along with a position × group interaction ($F(2,30) = 4.47$, $p < .05$). Crucially, the three-way interaction between position, group and list type was significant ($F(2,30) = 4.54$, $p < .02$). In the control group, C2 phonemes were recalled less accurately than C1 phonemes in mixed lists, but both consonants were recalled at similar levels in pure lists (see Fig. 6, Panel A). SD patients, on the hand, showed a clear disadvantage for C2 on both pure and mixed lists (Panel B). This disadvantage for the final consonant mirrored the pattern seen in *both* groups when recalling the nonwords.

We also investigated whether errors were more likely to preserve the rime (VC) of items as a complete unit compared to the preservation of onset and vowel (CV). Considering all item errors, we summed the number that contained either a CV or VC segment of a presented item. As can be seen in Table 7, errors were more likely to contain presented CV segments than VC. This difference was significant in controls for pure lists ($\chi^2 = 8.34$, $p < .005$) and mixed lists ($\chi^2 = 53.4$, $p < .0001$) but only for mixed lists in the patient group ($\chi^2 = 19.1$, $p < .001$; pure lists: $\chi^2 = 1.66$, *n.s.*).

3.8. Phoneme-level summary

Controls made a larger number of phoneme migrations when recalling mixed than pure word lists. They were also more likely to repeat phonemes and to include phonemes from outside the list when responding to mixed lists. SD cases also frequently incorporated extra-list phonemes or repeated phonemes within lists, but these tendencies were present both for pure and mixed lists. In all conditions, vowels were more likely than consonants to be recalled correctly and the final consonant of items was particularly vulnerable to the mixed list manipulation. This phoneme was also successfully recalled least often by SD patients.

4. General discussion

In the present study, verbal STM in SD patients and age-matched controls was directly compared. Previous studies have revealed that SD patients make numerous phoneme migration errors when recalling word lists, consistent with a breakdown in "semantic binding" – the way in which semantic activation helps to constrain sequences of phonemes into coherent units (Patterson et al., 1994; see also Jefferies et al., 2008; Knott et al., 1997; Majerus et al., 2007). These errors appeared to resemble those made by normal subjects when recalling lists that contain nonwords, suggesting a similar breakdown in phonological coherence in the absence of lexical-semantic constraints (Ellis, 1980; Jefferies et al., 2006; Treiman & Danis, 1988). In this study, we directly compared the effects of these two forms of disruption to lexical/semantic support for verbal STM. When SD patients and healthy participants were tested on pure word lists and mixed lists containing words and nonwords, we found that SD patients' errors in pure word lists were similar to those in mixed lists for controls. The main findings of the study can be summarised as follows:

1. SD-pure list recall resembled control-mixed list recall. In general, the same pattern of performance was seen for SD patients whether they were recalling pure or mixed lists. Controls responded in a similar way but *only* for the mixed lists. This was most obvious in the rate of phoneme migration errors, which SD patients made frequently for both list types but controls only made to mixed lists.
2. The effect of lexical/semantic disruption was largest in the recency portion of the serial position curve. In controls, the difference between pure and mixed list recall was greatest for final list items. In patients, the primacy effect was robust and recall of later items was most impaired.
3. SD patients were not sensitive to the composition of lists; their word recall was equally impaired irrespective of whether lists contained only real words or words and nonwords. Likewise, within the mixed lists the ratio of words to nonwords present had no effect on their accuracy (in contrast to controls, who improved when lists contained a higher proportion of words). A similar proportion of the patients' errors were nonwords for both list types; conversely, controls were more likely to produce real words in pure lists.
4. Accuracy and error rates were related to the lexical and semantic status of list items. Lexicality, frequency and imageability all influenced success of recall and rate of phoneme migrations (with the one exception that imageability did not affect migrations for the mixed lists). Although SD patients showed smaller lexicality effects than controls, they remained sensitive to the frequency and imageability of words, which may reflect the particular susceptibility of rarer, more abstract words to semantic degradation (see below).

These results are readily accommodated by models that view repetition and verbal STM as arising from an interaction between

Table 7
Errors preserving CV and VC sections of items.

	Pure lists			Mixed lists		
	Total errors	Errors containing list CV	Errors containing list VC	Total errors	Errors containing list CV	Errors containing list VC
SD	663	128 (19%)	110 (17%)	1192	271 (23%)	187 (16%)
Control	218	74 (34%)	47 (22%)	1159	427 (37%)	266 (23%)

Percentage of total errors given in parentheses.

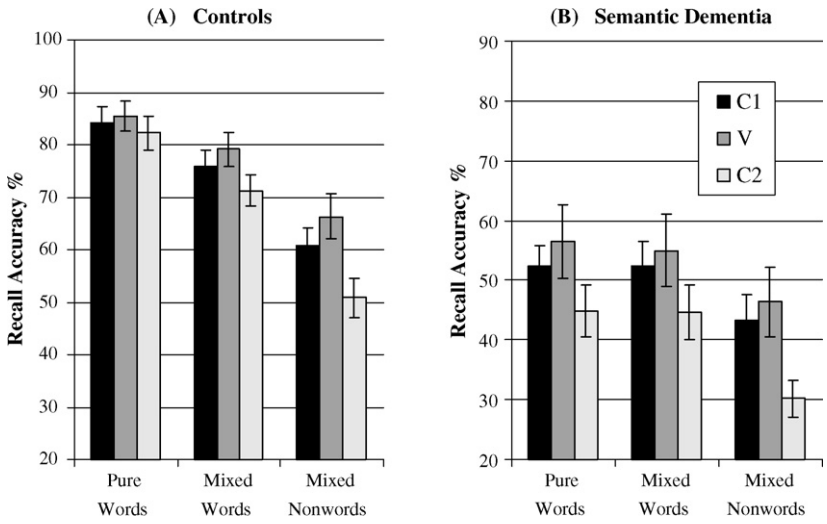

Fig. 6. Recall of phonemes by syllabic position. Phoneme recall accuracy as a function of syllabic position for (A) controls and (B) SD patients. C1 = first consonant; V = vowel; C2 = final consonant. Bars indicate one standard error of mean.

phonological and semantic representations (Jefferies et al., 2006; Martin & Saffran, 1997; Martin et al., 1999; Patterson et al., 1994). The semantic binding view, which specifically states that feedback from the activation of a word's meaning provides constraint over its phonological configuration, provides a particularly parsimonious account of the observed findings in patients and controls. Martin and Saffran's (1997) adaptation of the interactive activation model was also largely supported by the data. However, we found no evidence for one specific prediction made by this model, that the contribution of semantic activation is largest for the initially presented items. Instead, it was later list items that were most affected by the mixing of words and nonwords in controls and only one of the six SD patients showed any tendency towards recalling recent words most accurately.

One possible explanation for this disparity, discussed by Forde and Humphreys (2002), is that the temporal dynamics of the model differ for lists of different lengths. Martin and Saffran (1997) based their serial position account on the performance of stroke aphasics recalling single words and two-word lists. For the longer lists of five words used here (and typical of studies of ISR in SD) there was much more time between a word's presentation and its spoken recall. This extra time may have allowed the maximum possible semantic activation to be achieved for all words, eliminating the unusual serial position effect. However, we also found a subtle but statistically significant tendency for controls to show the largest decrements for mixed list presentation towards the ends of lists. This effect might have arisen as a result of output interference, whereby the phonological representations of later list items were affected by the recall of earlier ones. Consequently, later items might depend on semantic binding to a greater extent to maintain their phonological integrity. Hulme et al. (1997) proposed a similar explanation for their finding of larger frequency effects towards the ends of lists.

An alternative explanation proposed to account for lexical–semantic effects in healthy subject's recall is the action of a "redintegration" process that reconstructs degraded phonological traces at the point of recall (Hulme et al., 1991, 1997; Schweickert, 1993). On this view, decayed phonological representations are "cleaned up" by matching them to entries in the lexical system (e.g., the degraded trace *e_eph_nt* could be cleaned up to produce *elephant*). This proposal differs substantially from those described above, in that the retention of verbal information involves a purely phonological code and lexical-semantic effects arise only as a result of reconstruction at the point of recall. Even so, the poor word recall of SD patients is consistent with this hypothesis, since their impoverished lexical knowledge would affect their ability to reconstruct degraded traces. It is also consistent with the strong tendency for controls to produce real words in their errors, even to mixed lists where the majority of stimuli were nonwords. Presenting mixed lists would disrupt the redintegration process, unless the system had some mechanism for determining which degraded traces were words that should be reconstructed and which were nonwords that should be left alone.

However, a failure of lexical reconstruction does not appear to be a complete explanation of the effects of mixed list presentation. Jefferies, Frankish and Noble (in press-a) recently tested healthy participants with unpredictable mixed lists of the kind employed in this study and predictable mixed lists in which words and nonwords alternated. Participants produced word responses to word targets more often in the latter type of list, suggesting that knowledge of the lexical status of items was used strategically to ensure that words were produced in the correct positions in the list. However, this effect was relatively subtle: overall *accuracy* did not differ between the two types of mixed list and phoneme migrations were common in both cases. This suggests that much of the advantage for word recall over nonwords stems from an automatic binding process rather than strategic redintegration of items known to be words.

How does the semantic binding hypothesis account for the effect of nonwords on the phonological integrity of words? We believe that the binding of words in STM tasks depends not just on the strength of their phonological and semantic representations but also on the degree to which other list items are bound. Although lexical/semantic binding is usually sufficient to prevent phoneme migrations for real words, this is not always the case: phoneme transpositions occasionally occur in spontaneous speech (Vousden, Brown, & Harley, 2000) and occur more often in immediate serial recall when phonologically similar words are presented (Page, Madge, Cumming, & Norris, 2007). As the demands on the phonological system increase, the probability of phoneme order errors should also increase. For our five-item lists, when all of the items were words the phonemes of each item were sufficiently bound to ensure that phonemes were recalled together. However, when some of the items lacked semantic representations to bind their phonemes together (nonwords in controls; semantically degraded words in SD), these phonemes were liable to be recalled in incorrect list positions and to displace other phonemes in the list. Despite their stronger binding, this sometimes resulted in word phonemes being displaced into other list positions. We believe that this list-level approach is the most parsimonious way to

explain migration errors, which overwhelmingly involve phoneme exchanges between *different* items in a list rather than within a single item.

However, even assuming a list-level approach to phonological stability, there remain a number of ways in which effects of long-term linguistic knowledge could arise. The semantic binding account, along with approaches that apply the interactive activation model of word production to repetition (Martin & Gupta, 2004; Martin & Saffran, 1997; Martin et al., 1999), posit that verbal STM arises as a result of interactions between semantic and phonological levels of representation. Familiar words are associated with strongly instantiated patterns of activation at both of these levels and either could feasibly contribute to binding. Could our findings be explained entirely as an effect of long-term representations at the phonological level? We think not, for two reasons. First, patients with SD, who showed progressive deterioration of semantic knowledge, showed numerous phoneme migration errors for words despite having largely intact phonological processing (see also Jefferies et al., 2005). Second, both patients and controls were influenced by imageability. This variable relates to a word's meaning rather than its phonology.[1] Not only were concrete words recalled more accurately than more abstract words, their phonemes were also less likely to migrate, directly relating the phonological coherence of words to the status of their semantic representations. The greater stability of concrete words may arise because concrete words possess richer and more detailed semantic representations, thus providing a stronger input to the phonological system (Jones, 1985; Plaut & Shallice, 1993). Despite this, it is clear that long-term phonological knowledge also influences verbal STM. When recalling lists of nonwords, healthy individuals are more accurate for nonwords with high phonotactic probabilities (Gathercole, Frankish, Pickering, & Peaker, 1999; Thorn & Frankish, 2005) and those composed of high frequency syllables (Nimmo & Roodenrys, 2002). Such items benefit from stronger associative links within the phonological system itself, indicating an additional source of constraint over phonology that can operate independently of the semantic effects we have focused on here. SD patients are also sensitive to phonotactic probability (Majerus et al., 2007), suggesting that these effects can continue in the face of semantic degradation.

Frequency effects, observed in both patients and controls, could be attributed to phonological or semantic sources of binding. High frequency words are represented more robustly in the phonological system since their phonemes are co-activated more often but their common occurrence might also lead them to possess better-established semantic representations. While both of these sources of constraint might contribute to the frequency effect for healthy subjects, the effects of semantic knowledge should be eroded in SD. Why then did the SD patients show larger frequency effects than the controls? Although we did not formally assess the known–unknown status of the words used in this study, the degree of semantic degradation in SD is strongly predicted by frequency (Bozeat et al., 2000; Funnell, 1995; Jefferies et al., in press-b; Lambon Ralph et al., 1998) and it is likely that the patients possessed less intact semantic knowledge of low frequency targets. Consequently, the SD patients would have shown weakened semantic binding for the lower frequency target words. High frequency words were less likely to be degraded and so might have benefited from greater levels of semantic binding. The enhanced frequency effect therefore suggests that SD patients did derive *some* benefit from their remaining semantic knowledge.

Turning to form of the phonological errors uncovered in this investigation, in common with previous reports (Ellis, 1980; Jefferies et al., 2006; Patterson et al., 1994; Treiman & Danis, 1988) we found that patients and controls recalled vowels more accurately than consonants. Vowels appeared more resistant than consonants to reductions in semantic binding, possibly reflecting their greater acoustic intensity. However, we did not find that phoneme binding within the rime of items was more stable than the link between onset and vowel. The preservation of the rime has been reported as key characteristic of nonword recall in healthy subjects (Treiman & Danis, 1988) and computational models account for this finding by representing syllable onsets and rimes with separate nodes (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996). Nimmo and Roodenrys (2002) also failed to find an advantage for rimes in nonword ISR, and speculated that differences in the articulatory features of the phonemes they used in C1 vs. C2 positions could account for the anomaly. A similar explanation could hold here, as we did not match the phonemes presented in the two consonant positions. Of greater relevance, this tendency to preserve CV rather than VC segments was seen for both patients (in pure word lists) and healthy participants (in mixed lists).

In addition to phoneme order errors, errors of phoneme *identity* were common, with increased numbers of phoneme intrusions and repetitions for SD patients and in controls when repeating mixed lists (see also Jefferies et al., 2006). This points to a general role for semantic binding in preserving the integrity of phonological traces, rather than a specific function in preserving serial order. When semantic activation is disrupted, the links between phonemes are weakened. In some cases, this could result in phonological segments re-combining but, in other situations where a phoneme receives insufficient activation, it might be forgotten entirely. In these cases, in order to produce CVC responses participants must fill the gaps left by these lost phonemes. Such attempts could lead to repeated use of other list phonemes as well as intrusions from outside the list. Both phoneme order and item errors could reflect an underlying failure of semantic representations to provide sufficient constraint over temporary phonological activations.

Finally, there is one practical outcome of this study worth noting. By demonstrating that the mixed lists method closely mimics the effects of SD on list recall, we have highlighted the usefulness of this technique for investigating the effects of semantic disruption on phonological integrity in STM, something which has previously been investigated primarily in neuropsychological populations. There are potentially many advantages of being able to conduct such experiments in healthy subjects, not least that patients are often reluctant to perform ISR tasks which they find highly demanding. In conclusion, poor phonological coherence in verbal STM has been argued to reflect an underlying failure of semantic representations to provide sufficient constraint over temporary phonological activations. We have demonstrated that two rather different causes of this effect – the degradation of semantic representations in SD vs. use of unpredictable mixed lists in healthy participants that artificially disrupt semantic contributions to ISR – have strikingly similar consequences for list recall. These findings provide important converging evidence for the role of semantic representations in verbal STM.

[1] However, Reilly and Kean (2007) recently reported systematic differences in the surface properties of abstract and concrete nouns. Abstract nouns tend to be longer and more morphologically and phonologically complex. These differences are unable to account for imageability effects in the present study, in which all stimuli were monosyllabic CVCs.

Acknowledgements

We are indebted to the patients and their carers for their generous assistance with this study. We also thank Dr. David Bateman

(Royal United Hospital, Bath) and Dr. Mark Doran (Walton Centre for Neurology and Neurosurgery, Liverpool) for referring two of the patients to us. The work was supported by a grant from the NIMH (MH64445), an RCUK fellowship awarded to EJ, a studentship from the University of Manchester to PH and an MRC programme grant (G0501632).

Appendix A. Experimental stimuli

Words used in pure vs. mixed list comparison given in capitals.

Pure Lists

High frequency, high imageability

BALL, teeth, WIFE, nose, dark PHONE, mouth, book, SEAT, red wood, BOAT, sun, GIRL, roof gun, ROCK, ship, heart, MALE song, love, NIGHT, park, HEAD road, HORSE, white, LEG, FISH cash, HOME, FACE, pool, WINE NECK, room, FOOT, dog, HILL BOARD, shop, MEN, wheel, HOUSE WALL, RAIN, feet, BED, king.

High frequency, low imageability

THOUGHT, piece, JACK, wish, fell MASS, feel, thick, DATE, rise fine, SHAPE, south, TERM, lord call, SIDE, wait, move, JOB long, mean, HALF, wide, RACE type, THING, warm, SHOCK, VOICE lead, NAME, SIZE, cut, HOPE WEEK, miss, GOD, turn, SHARP TOP, save, RULE, FORM, work BASE, LINE, part, DEATH, firm.

Low frequency, high imageability

HAWK, gym, FAN, wool, geese CANE, pet, wig, DIME, cheese shed, COIN, leaf, BOOT, morgue cart, DOVE, noose, lamb, JEEP thumb, peach, YACHT, cave, HEN juice, THORN, heel, BAT, FOAM pearl, RIB, WEED, chalk, HOOF KITE, nail, MOUSE, pig, HEDGE RAT, gem, HARP, SURF, duck TOAD, WEB, limb, FOG, cage.

Low frequency, low imageability

MALL, nip, RACK, cheat, sod PSALM, push, latch, JADE, tuck kale, ZONE, rung, DIP, verb loon, WRATH, meek, bang, HUSH chic, hail, VICE, rap, GERM bait, SAP, dirge, RHYME, KNOLL lodge, RAID, WHIFF, curse, TON WHARF, nerve, SAGE, keel, BET HURT, pawn, FOUL, BID, whack MASH, LEAN, jerk, THUD, whoop.

Mixed Lists

High frequency, high imageability

Teeth, beng, sisle, woam, fik siefe, book, rorl, han, med fing, hees, park, bot, raim siebe, wote, rad, gun, kerm bol, heem, foate, saysh, king BALL, pid, WIFE, moess, shart PHONE, korp, bim, SEAT, looth wid, BOAT, jarm, GIRL, hoys jote, ROCK, ned, forp, MALE mal, jong, NIGHT, koese, HEAD taybe, HORSE, roak, LEG, FISH cow-t, HOME, FACE, bal, WINE NECK, raig, FOOT, dibe, HILL BOARD, tayve, MEN, ruuge, HOUSE WALL, RAIN, fok, BED, thoape.

High frequency, low imageability

Part, hes, neek, wole, burge tayse, role, vike, min, deef het, sharf, lead, saybe, voan hom, jurn, vipe, till, rud feen, thit, bam, haid, loss THOUGHT, seipe, JACK, raish, feem MASS, heen, rel, DATE, poeth houne, SHAPE, garl, TERM, fak cun, SIDE, lep, mort, JOB lood, mun, HALF, vite, RACE tayde, THING, jurz, SHOCK, VOICE gid, NAME, SIZE, ket, HOPE WEEK, hin, GOD, tayje, SHARP TOP, sayde, RULE, FORM, zine BASE, LINE, hoat, DEATH, rork.

Low frequency, high imageability

Cart, wol, gis, dem, rorn mot, geese, bick, sharl, wan nop, hus, limb, fet, vayze leet, bon, hud, wig, sherp lut, fod, kep, rorm, noose HAWK, jid, FAN, barl, tice CANE, thert, roarss, DIME, beel fid, COIN, laysh, BOOT, tharss fon, DOVE, rab, zime, JEEP sorl, weem, YACHT, kerze, HEN weis, THORN, harg, BAT, FOAM paim, RIB, WEED, sawg, HOOF KITE, fal, MOUSE, barss, HEDGE RAT, warthe, HARP, SURF, mek TOAD, WEB, hal, FOG, keem.

Low frequency, low imageability

Bang, lif, dop, nook, wais tiege, kale, boun, dap, tharj tayne, lidge, curse, rorsch, boof nood, hef, kang, verb, loate nate, dorth, sek, sherb, lodge MALL, beuffe, RACK, goyt, hoess PSALM, hoak, lan, JADE, tiefe keet, ZONE, seithe, DIP, burl woan, WRATH, morke, baf, HUSH dit, hon, VICE, woash, GERM thayte, SAP, jud, RHYME, KNOLL hol, RAID, WHIFF, girse, TON WHARF, noid, SAGE, raowl, BET HURT, yourss, FOUL, BID, gen MASH, LEAN, jook, THUD, werp.

References

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

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, *55*, 64–88.

Archibald, L. M. D., & Gathercole, S. E. (2007). Nonword repetition and serial recall: Equivalent measures of short-term memory. *Applied Psycholinguistics*, *28*, 587–606.

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

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

Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*, 1207–1215.

Bozeat, S., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2002). When objects lose their meaning: What happens to their use? *Cognitive, Affective and Behavioural Neuroscience*, *2*, 236–251.

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology*, *26*, 467–482.

Cappelletti, M., Butterworth, B., & Kopelman, M. (2001). Spared numerical abilities in a case of semantic dementia. *Neuropsychologia*, *39*, 1224–1239.

Caza, N., Belleville, S., & Gilbert, B. (2002). How loss of meaning with preservation of phonological word form affects immediate serial recall performance: A linguistic account. *Neurocase*, *8*, 255–273.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *33*, 497–505.

Crutch, S. J., & Warrington, E. K. (2006). Partial knowledge of abstract words in patients with cortical degenerative conditions. *Neuropsychology*, *20*, 482–489.

Dell, G. S., & O'Seaghda, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*, 287–314.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838.

Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior*, *19*, 624–634.

Forde, E. M. E., & Humphreys, G. W. (2002). The role of semantic knowledge in short-term memory. *Neurocase*, *8*, 13–27.

Funnell, E. (1995). Objects and properties: A study of the breakdown of semantic memory. *Memory*, *3*, 497–518.

Funnell, E. (1996). Response biases in oral reading: An account of the co-occurence of surface dyslexia and semantic dementia. *Quarterly Journal of Experimental Psychology*, *49A*, 417–446.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 84–95.

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., Lipinski, J., Abbs, B., & Lin, P. H. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language*, *53*, 141–162.

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

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

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*, 1783–1806.

Howard, D., & Patterson, K. (1992). *Pyramids and Palm Trees: A test of semantic access from pictures and words*. Bury, St. Edmunds, UK: Thames Valley Test Company.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*, 685–701.

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

Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, *49*, 500–518.

Jefferies, E., Frankish, C., & Lambon Ralph, M. A. (2006). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language*, *54*, 81–98.

Jefferies, E., Frankish, C. R., & Noble, K. (in press-a). Lexical coherence in short-term memory: Strategic reconstruction or "semantic glue"? *Quarterly Journal of Experimental Psychology*.

Jefferies, E., Hoffman, P., Jones, R., & Lambon Ralph, M. A. (2008). The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: A comparative study. *Journal of Memory and Language*, *58*, 66–87.

Jefferies, E., Jones, R., Bateman, D., & Lambon Ralph, M. A. (2004). When does word meaning affect immediate serial recall in semantic dementia? *Cognitive, Affective and Behavioral Neuroscience*, *4*, 20–42.

Jefferies, E., Jones, R. W., Bateman, D., & Lambon Ralph, M. A. (2005). A semantic contribution to nonword recall? Evidence for intact phonological processes in semantic dementia. *Cognitive Neuropsychology*, *22*, 183–212.

Jefferies, E., Patterson, K., Jones, R. W., Bateman, D., & Lambon Ralph, M. A. (2004). A category-specific advantage for numbers in verbal short-term memory: Evidence from semantic dementia. *Neuropsychologia*, *42*, 639–660.

Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (in press-b). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology*.

Jones, G. V. (1985). Deep dyslexia, imageability and ease of prediction. *Brain and Language*, *24*, 1–19.

Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)*. Hove, UK: Lawrence Erlbaum Associates.
Knott, R., Patterson, K., & Hodges, J. R. (1997). Lexical and semantic binding effects in short-term memory: Evidence from semantic dementia. *Cognitive Neuropsychology*, *14*, 1165–1218.
Knott, R., Patterson, K., & Hodges, J. R. (2000). The role of speech production in auditory–verbal short-term memory: Evidence from progressive fluent aphasia. *Neuropsychologia*, *38*, 125–142.
Lambon Ralph, M. A., Graham, K. S., Ellis, A. W., & Hodges, J. R. (1998). Naming in semantic dementia—What matters? *Neuropsychologia*, *36*, 775–784.
Majerus, S., Norris, D., & Patterson, K. (2007). What does a patient with semantic dementia remember in verbal short-term memory? Order and sound but not words. *Cognitive Neuropsychology*, *24*, 131–151.
Martin, N., & Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology*, *21*, 213–228.
Martin, N., & Saffran, E. (1990). Repetition and verbal STM in transcortical sensory aphasia: A case study. *Brain and Language*, *39*, 254–288.
Martin, N., & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, *14*, 641–682.
Martin, R. C., Lesch, M. F., & 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.
McCarthy, R. A., & Warrington, E. K. (1987). The double dissociation of short-term memory for lists and sentences: Evidence from aphasia. *Brain*, *110*, 1545–1563.
McCarthy, R. A., & Warrington, E. K. (2001). Repeating without semantics: Surface dysphasia? *Neurocase*, *7*, 77–87.
Nimmo, L. M., & Roodenrys, S. (2002). Syllable frequency effects on phonological short-term memory tasks. *Applied Psycholinguistics*, *23*, 643–659.
Page, M. P. A., Madge, A., Cumming, N., & Norris, D. G. (2007). Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus. *Journal of Memory and Language*, *56*, 49–64.
Patterson, K., Graham, N., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, *6*, 57–69.
Patterson, K., & Lambon Ralph, M. A. (1999). Selective disorders of reading? *Current Opinion in Neurobiology*, *9*, 235–239.
Patterson, K., & Marcel, A. J. (1992). Phonological ALEXIA or PHONOLOGICAL alexia? In J. Alegria, D. Holender, J. Junca de Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 259–274). Amsterdam: Elsevier.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.
Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study in 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. *The Quarterly Journal of Experimental Psychology*, *48A*, 384–404.
Raven, J. C. (1962). *Coloured progressive matrices sets A, AB, B*. London: H.K. Lewis.
Reilly, J., & Kean, J. (2007). Formal distinctiveness of high and low-imageability words: Analyses and theoretical implications. *Cognitive Science*, *31*, 157–168.
Reilly, J., Martin, N., & Grossman, M. (2005). Verbal learning in semantic dementia: Is repetition priming a useful strategy? *Aphasiology*, *19*, 329–339.
Rey, A. (1941). L'examen psychologique dans le cas d'encaphalopathie traumatique. *Arch Psychologie*, *28*, 286–340.
Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*, 205–235.
Romani, C., McAlpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology*, *61*, 292–323.
Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, *21*, 168–175.
Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.
Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioural Neurology*, *2*, 167–182.
Thorn, A. S. C., & Frankish, C. R. (2005). Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *31*, 729–735.
Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *14*, 145–152.
Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology*, *41*, 101–175.
Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 1256–1271.
Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635–657.
Wechsler, D. (1987). *Wechsler Memory Scale – Revised (WMS-R)*. New York: Psychological Corporation.