Phonological and semantic strategies in immediate serial recall

Guillermo Campoy
University of Murcia, Spain

Alan Baddeley
University of York, UK

It has been suggested that certain theoretically important anomalous results in the area of verbal short-term memory could be attributable to differences in strategy. However there are relatively few studies that investigate strategy directly. We describe four experiments, each involving the immediate serial recall of word sequences under baseline control conditions, or preceded by instruction to use a phonological or semantic strategy. Two experiments varied phonological similarity at a presentation rate of one item every 1 or 2 seconds. Both the control and the phonologically instructed group showed clear effects of similarity at both presentation rates, whereas these were largely absent under semantic encoding conditions. Two further experiments manipulated word length at the same two rates. The phonologically instructed groups showed clear effects at both rates, the control group showed a clear effect at the rapid rate which diminished with the slower presentation, while the semantically instructed group showed a relatively weak effect at the rate of one item per second, and a significant reverse effect with slower presentation. The latter finding is interpreted in terms of fortuitous differences in inter-item rated associability between the two otherwise matched word pools, reinforcing our conclusion that the semantically instructed group were indeed encoding semantically. Implications for controlling strategy by instruction are discussed.

Verbal short-term memory (STM) has traditionally been conceived as relying on the maintenance of phonological representations. Evidence supporting this view emerged first from the observation by Conrad (1964) that errors in immediate recall of consonants tended to be similar in sound to the correct item (e.g., misrecalling V instead of P), even when the stimuli had been presented visually. Congruently with this, Conrad and Hull (1964) showed that lists of letters similar in sound (e.g., V, C, P, D, T) were harder to recall than lists of dissimilar letters (e.g., K, Z, W, R, Y), which was explained in terms of the greater difficulty of discriminating among similar phonological traces. On the basis of these and similar results, various memory models have included the idea that verbal material is phonologically coded in STM. According to the working memory model, for example, the short-term maintenance of verbal material is mainly accomplished by the phonological loop, which comprises a phonological store where phonological traces are maintained and a control process of subvocal rehearsal that refreshes traces and counteracts decay (Baddeley, 2007; Burgess & Hitch, 1999).
Even though it seems clear that phonological coding plays a major role in verbal STM tasks involving the immediate serial recall of digits or unrelated words, it is also the case that immediate recall of verbal material can benefit from the short-term maintenance of other kinds of information, especially semantic. Studies of the immediate recall of sentences, for example, have shown that participants can recall many more words from sentences than from lists of unrelated words (Brenner, 1940). There are a number of factors that could contribute to this phenomenon, including syntactic constraints and lexically based sequential redundancy. However, it seems likely that sentence superiority is in part a consequence of participants being able to maintain the overall meaning of the sentence and use this information at recall. In keeping with this idea, Jefferies, Lambon Ralph, and Baddeley (2004) showed that, whereas phonological errors (e.g., misrecalling *rut* instead of *rug*) were more common than semantic errors in word-list recall, semantic errors (e.g., misrecalling *carpet* instead of *rug*) were more frequent in sentence recall.

Although sentence recall is probably especially propitious for the participation of semantic information, there are a number of findings that suggest that the maintenance of semantic information may also play a role in the immediate serial recall of unrelated words. Some such evidence comes from studies showing poorer recall for words with a low semantic content (e.g., articles, prepositions, conjunctions) than for more meaningful words such as adjectives and nouns (Caza & Belleville, 1999; Tehan & Humphreys, 1988; see, however, Bourassa & Besner, 1994). This finding could be attributed to the fact that, although recall relies mainly on phonological traces, participants are able to encode and use semantic information to improve performance, with semantically richer words gaining a greater advantage. Further evidence comes from studies that manipulate semantic similarity. In an early study, Baddeley (1966) examined the immediate serial recall of five-word lists drawn from a set of eight phonological similar words, from a set of eight adjectives with similar meanings, or from a set of eight control words. Results showed a clear effect of phonological similarity, but also a small but significant effect of semantic similarity (worse recall of semantically similar items). This result suggests a contribution to recall from both phonological and semantic encoding, with similar words resulting in less distinctive traces and therefore poorer recall. A subsequent study by Baddeley and Levy (1971) showed that semantic coding could be induced in immediate serial recall of six-word sequences comprising three noun–adjective pairs, provided the pairs were semantically compatible (e.g., *priest–devout*), an effect that was eliminated when the pairs were made semantically similar to other pairs within the list (e.g., *bishop–pious*). A further study (Baddeley & Ecob, 1970) using semantically compatible or incompatible triplets (e.g., *my fine wine* vs *wine my fine*) found effects of both phonological and semantic similarity, with acoustic similarity impairing performance on immediate test, and effects of semantic similarity predominating after a delay—results that they interpreted as suggesting simultaneous encoding of both phonological and semantic features.

More recently, Poirier and Saint-Aubin (1995); Saint-Aubin & Poirier, 1999; Saint-Aubin, Ouellette, & Poirier, 2005) have found effects of semantic similarity in a number of experiments in which similarity was manipulated by presenting words from the same semantic category (e.g., *horse, bull, goat, mare, pig, cow*) or from different categories. Overall, semantic similarity led to a better recall in these studies,1 which was explained by an extended version of the redintegration hypothesis. According to the redintegration hypothesis (Hulme, Maughan, & Brown, 1991; Schweickert, 1993), short-term recall involves the participation of reconstructive mechanisms by which partially degraded phonological traces are redintegrated on the basis of long-term knowledge about the constituent words. In their extended version of the hypothesis, Poirier and Saint-Aubin suggested that semantic information improves redintegration accuracy by providing an additional cue that delimits the number of suitable candidates in long-term memory (Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999).

---

1 This advantage for lists of semantically similar lists could be seen as conflicting with the results obtained by Baddeley (1966), who found worse recall for lists of similar words. However, this discrepancy is easily explained by relevant procedural differences. Semantic similarity in the studies by Saint-Aubin and Poirier was found to have a large beneficial effect on item recall, but also a slight detrimental effect on order recall (Saint-Aubin et al., 2005). The negative effect found by Baddeley (1966) can be explained by the fact that, in contrast to Saint-Aubin and Poirier, Baddeley used closed sets of stimuli, which were, moreover, displayed during recall. It seems clear that performance in this situation was a product of order rather than item recall, with the consequence of worse recall for similar lists.
Consider, for example, that the phonological representation of the word *cat* has lost the last phoneme. In this situation, the redintegration process could lead to error because the remaining trace would fit with a number of incorrect words (*cab*, *can*, *cache*, *cad*, *cash*, *cap*, etc.). Stored semantic information would be useful then to pick the candidate whose meaning is consistent with this semantic information. For Poirier and Saint-Aubin, the advantage of similar word lists would be a consequence of the fact that semantic similarity results in a more recallable and efficient semantic cue (Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999). Again, findings support the idea that participants in STM tasks encode semantic information and employ this information at recall.

Whereas there is broad agreement that semantic coding may play a role in immediate serial verbal recall, there is less agreement on how to interpret it. One possibility is to attribute it to the influence of long-term memory, the view initially taken by Baddeley (1972). Within the multi-component working memory framework, Baddeley (2000, 2007) has subsequently proposed that an additional component, the episodic buffer, provides a link between long-term and working memory, accounting for the short-term effects of semantic coding. Another approach is to propose a separate semantic STM (e.g., Martin, 2005), a view derived from studies of brain-damaged patients (Freedman & Martin, 2001; Hanten & Martin, 2000; Majerus, Van der Linden, Poncelet, & Metz-Lutz, 2004; Martin & He, 2004; Martin & Romani, 1994; Martin, Shelton, & Yaffee, 1994; Romani & Martin, 1999), from neuroimaging dissociations (Martin, Wu, Jackson, Freedman, & Lesch, 2003; Shivde & Thompson-Schill, 2004), and from behavioural data (Haarman, Ashling, Davelaar, & Usher, 2005; Haarman, Davelaar, & Usher, 2003; Haarman & Usher, 2001).

Regardless of these different conceptions and their theoretical consequences, the fact that verbal STM can be partially supported by stored semantic information is a potential source of important interpretative problems in verbal STM experiments, with different investigators obtaining conflicting results. The reason for this is that, although the use of a phonological strategy for recalling short sequences of verbal items may predominate, such a strategy may lose ground in favour of other strategies under certain conditions (Logie, Della Sala, Laiacona, Chalmers, & Wynn, 1996).

There are a number of instances of theoretically important results in which major inconsistencies appear to have resulted from switching between phonological encoding and another strategy, probably semantic. One instance concerns the claim that poor beginning readers fail to show the characteristic phonological similarity effect in verbal STM, which Liberman and colleagues (Liberman, Mann, Shankweiler, & Werfelman, 1982; Mann, Liberman, & Shankweiler, 1980) suggest might reflect a general failure of poor readers to encode visual material phonologically, providing a possible cause of their reading deficit. Subsequent research suggested that poor readers do encode phonologically at shorter list lengths, only failing to show a phonological similarity effect when list length greatly exceeded their memory span, which is typically shorter than that of good readers (Hall, Wilson, Humphries, Tinizmann, & Bowyer, 1983; Johnston, 1982). This result is generally attributed to the abandonment of a phonological strategy (Baddeley, 2007; Johnston, 1982). Such an interpretation has also been offered for data from adults, whereby phonological variables that are present at short list lengths may not be found with longer lists, particularly when other variables such as irrelevant speech or articulatory suppression are involved. For example, Salamé and Baddeley (1986) studied the combined effects of phonological similarity and irrelevant speech across a range of list lengths, finding additive effects of both variables that disappeared as list length increased, suggesting the absence of phonological coding with longer lists. A similar interaction between phonological coding and list length occurs when auditory presentation is used under articulatory suppression (Baddeley & Larsen, 2007; Jones, Macken, & Nicholls, 2004), resulting in a claim by Jones et al. that their data are inconsistent with the concept of a short-term phonological store. However, Baddeley and Larsen (2007) suggest that the results of Jones et al. simply reflect the abandonment of phonological coding as level of difficulty increases.

Preliminary empirical evidence for the potential importance of strategy comes from Logie et al. (1996). In this study, participants were interviewed after a series of verbal STM tasks in order to determine the strategies that they had employed. Results revealed that participants classified as using a semantic strategy showed weaker phonological effects (phonological similarity and word length) than the great majority of
participants who reported adopting a phonological strategy. A more direct approach to strategy was taken by Hanley and Bakopoulou (2003), Experiment 2), in a study involving the immediate serial recall of phonologically similar or dissimilar letter sequences. They left one group of participants free to select any strategy, instructed a second group to encode phonologically, while a third group was asked to use each item as the initial letter of a word, instructing them to link the words to form a sentence. They obtained the standard phonological similarity effect in the baseline group, a more marked phonological similarity effect when phonological coding was encouraged, and an absence of a phonological similarity effect under their lexical-semantic instructions.

The experiments that follow aim to build on this observation in the hope of developing methods within which strategy can be controlled and manipulated, so as to facilitate a resolution of controversies based on the possibility of strategy switching. We chose to study the consequences of attempting to induce a phonological or semantic strategy on the presence or absence in serial verbal STM of two well-established effects, the previously described phonological similarity effect, and the word-length effect, the finding that immediate recall is worse for lists of long words than for short words (Baddeley, Thomson, & Buchanan, 1975).

The precise mechanism underlying the word length effect remains a matter of controversy. According to the working memory model, the effect is a consequence of a faster rehearsal rate for short words, so that a larger number of short words can be reactivated before they are lost because of decay (Baddeley, 1986; Burgess & Hitch, 1999; Mueller, Seymour, Kieras & Mier, 2003). Other authors, however, have suggested that the word-length effect is a result of differences in the phonological complexity of the items (Campoy, 2008; Caplan, Rochon, & Waters, 1992; Hulme, Suprenant, Bireta, Stuart, & Neath, 2004; Romani, McAlpine, Olsen, Tsoukni, & Martin, 2005; Service, 1998), with more complex words having been proposed to result in less distinctive traces (Hulme et al., 2004; Hulme et al., 2006). Despite the differences, these explanations attribute the effect to phonological factors rather than to semantic coding, a view that is consistent with the self-report data collected by Logie et al. (1996). It could be predicted, therefore, that the induction of a semantic strategy would result in the reduction of the word length effect.

**EXPERIMENT 1**

In Experiment 1 we presented lists of five words for immediate serial recall. Three conditions were studied: a control condition in which no strategy was suggested, a second condition in which participants were instructed to employ a semantic strategy ("Try to think of the word meanings and link them"), and a third phonological strategy condition ("Try to maintain and rehearse the word sounds"). Experiment 1 investigated the effect of these strategies on the phonological similarity effect.

**Method**

**Participants.** A total of 60 undergraduates from the University of Murcia were tested, 20 in each strategy condition.

**Stimuli.** Two sets of Spanish words were used, one comprising eight phonologically similar nouns and the other containing eight phonologically dissimilar nouns (see Appendix A). The word sets were individually matched for word frequency, familiarity, imageability, and concreteness using the LEXESP database (Sebastián, Martí, Carreiras, & Cuetos, 2000). All the words comprised two consonant–vowel syllables (two phonemes per syllable) with the stress on the penultimate syllable. Words in the similar set shared at least the two vocalic sounds. Phonological similarity in the dissimilar set was minimised by ensuring that no words shared more than one phoneme in the same position. A computer program generated by E-Prime (Schneider, Eschman, & Zuccolotto, 2002) was used for stimulus presentation and response storage.

**Procedure.** Participants were tested individually in a sound-attenuated booth. The experiment consisted of three blocks of 16 trials, 8 trials with phonologically similar words and another 8 trials with dissimilar words. The order of presentation of similar and dissimilar lists within a block was determined at random. Lists were constructed by selecting words from the specific set at random, with the constraint that every word appeared in five trials per block, once in each of the five possible positions of the lists. Four practice trials were presented before the experimental trials.
Instructions about the strategy to be employed were given after the second practice trial, once participants were familiar with the procedure. Participants were told that the objective of the experiment was to test different memory strategies and that it was therefore crucial that they used the specified strategy even though they thought that a different strategy could lead to better performance. They were reminded of strategy instructions before each of the three trial blocks.

The sequence of events in each trial was the following. First, a row of dashes appeared indicating that a mouse button had to be pressed to initiate the trial. One second after the mouse click, a plus sign was presented in the centre of the screen for 1 second, followed by a five-word list. Items were presented at a rate of one item per second (950 ms on, 50 ms off). After the list presentation, the 16 experimental stimuli appeared on the computer screen and participants employed the mouse to click on the five list words in their presentation order (clicked words appeared in bold). Participants were instructed to click on a question mark in place of any word they did not remember. A restart button labelled ‘’ were used in case of error. When they had finished list reconstruction, they were required to press an unlabelled black button. The basic screen configuration during list reconstruction is illustrated in Figure 1. The arrangement of the 16 experimental stimuli within this configuration was established at random at the beginning of the experiment for each participant.

![Figure 1](image_url)  
*Figure 1. Example of screen configuration during list reconstruction.*

### Results and discussion

Raw data were scored according to a strict serial recall criterion, by which a response was counted as correct when the right item was recalled in the correct position. The mean percentage correct recall is presented in Table 1. A 2 (word similarity) × 3 (induced strategy) mixed analysis of variance revealed a main effect of word similarity, $F(1, 57) = 58.90; MSe = 58.87; p < .01; \eta^2_p = 0.51$, with the phonologically dissimilar words being better recalled. The main effect of strategy was not significant, $F(2, 57) = 1.88; MSe = 32.290; p = .16; \eta^2_p = 0.06$, but there was a significant strategy by similarity interaction, $F(2, 57) = 6.48; MSe = 58.87; p < .01; \eta^2_p = 0.00$. Further analysis indicated a significant effect of phonological similarity for the uninstructed control group, $F(1, 57) = 30.01; MSe = 177.74; p < .01; \eta^2 = 0.34$, and for the group instructed to use phonological coding, $F(1, 57) = 39.51; MSe = 177.74; p < .01; \eta^2 = 0.41$, but not for the group instructed to encode semantically, $F(1, 57) = 2.34; MSe = 177.74; p = .13; \eta^2 = 0.04$. When similar and different sequences were analysed separately to investigate the interaction further, there proved to be no difference between the three strategy groups on dissimilar items, $F(2, 57) < 1$; $\eta^2 = 0.00$, whereas the groups did differ on similar word performance, $F(2, 57) = 5.00; MSe = 193.04; p = .01; \eta^2 = 0.15$, with the semantic group showing significantly better performance than either the control or phonological groups: respectively, $F(1, 57) = 7.77$ and 7.20; $MSe = 193.04; p < .02; \eta^2 = 0.12$ and 0.11, which did not differ.

Some idea of the consistency of the phonological similarity effect can be obtained by the number of participants showing an overall effect (cf. Logie et al, 1996). This was 18 out of 20 for the control group, 20 out of 20 for the

<table>
<thead>
<tr>
<th>STRATEGIES AND IMMEDIATE SERIAL RECALL 333</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
</tr>
<tr>
<td>Experiment 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Induced strategy</th>
<th>Semantic</th>
<th>Phonological</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissimilar words</td>
<td>70.37 (14.02)</td>
<td>69.67 (13.54)</td>
<td>68.17 (13.93)</td>
</tr>
<tr>
<td>Similar words</td>
<td>66.67 (15.54)</td>
<td>54.42 (12.41)</td>
<td>54.88 (13.55)</td>
</tr>
<tr>
<td>Difference</td>
<td>3.71 (11.12)</td>
<td>15.25 (8.53)</td>
<td>13.29 (12.52)</td>
</tr>
</tbody>
</table>

Means (and standard deviations) of the percentages of correct responses as a function of phonological similarity and induced strategy in Experiment 1.
The results of Experiment 1 are straightforward. Both the control and phonological strategy groups showed the expected effect of phonological similarity, while the semantic strategy group did not. Strategy appeared to have no effect on retention of dissimilar words, but semantic coding appeared to protect participants from the effects of phonological similarity, leading to enhanced performance on these items. Experiment 2 extends our investigation of strategy effects to their influence on word length.

### EXPERIMENT 2

**Method**

**Participants.** Participants were 75 undergraduates from the University of Murcia, 25 in each strategy condition.

**Stimuli and procedure.** Two sets of Spanish words were used, one comprising eight disyllabic nouns and the other containing eight trisyllabic nouns (see Appendix B). The word sets were matched for word frequency, familiarity, imageability, and concreteness using the LEXESP database (Sebastián et al., 2000). All the words consisted of consonant-vowel syllables with the stress on the penultimate syllable. Phonological similarity was minimised by ensuring that no two words from a set shared a syllable or more than any two phonemes in the same position. Procedure was the same as in Experiment 1.

### Results and discussion

Table 2 shows the mean percentage of items correctly recalled. A 3 (strategy) × 2 (word length) ANOVA showed a main effect of length, $F(1, 72) = 41.15$; $MS_e = 27.92$; $p < .01$; $\eta^2_p = 0.36$, no main effect of strategy, $F(2, 72) < 1$; $\eta^2_p = 0.02$, and a marginally significant interaction, $F(2, 72) = 2.44$; $MS_e = 27.92$; $p = .09$; $\eta^2_p = 0.06$, with all three groups showing a significant overall word length effect—for the phonological, control, and semantic groups, $F(1, 72) = 29.14, 11.79$, and 5.19 respectively; $MS_e = 55.84$; $p < .04$; $\eta^2 = 0.29$, 0.14, and 0.07. The number of participants showing a word length effect was 23 out of 25 for the phonological strategy group, 20 out of 25 for the control group, and 16 out of 25 for the semantically instructed group.

Despite this suggestion of reduced robustness of the word length effect under semantic encoding instructions, we failed to observe the clear interaction with strategy we found for phonological similarity. On the other hand, we clearly could not assert the absence of such an effect with any degree of conviction. We suspect that one problem with Experiment 2 might be that of processing the longer words at the rapid one per second rate employed. We therefore decided to replicate both studies at a rate of 2 seconds per word, which we felt would make it easier for participants to perform the somewhat complex recoding required for the semantic strategy (see Shulman, 1970, 1971). Experiment 3, therefore, examines the effect of phonological similarity as in Experiment 1, except for rate of presentation, which occurs at a slower rate of 2 seconds per word.

### EXPERIMENT 3

**Method**

**Participants.** Participants were 42 undergraduates from the University of Murcia, 14 in each strategy condition.

**Stimuli and procedure.** Stimuli were the phonologically similar and dissimilar words used in Experiment 1. The procedure was identical, except that lists were presented at a rate of one item every 2 seconds (950 ms on, 1050 ms off).

### Results and discussion

Means of percentage correct recall are shown in Table 3. A 2 (word similarity) × 3 (induced strategy) ANOVA showed a main effect of word similarity, $F(1, 65) = 27.33$; $MS_e = 38.21$; $p < .01$; $\eta^2_p = 0.32$, no main effect of strategy, $F(2, 65) < 1$; $\eta^2_p = 0.02$, and a marginally significant interaction, $F(2, 65) = 2.18$; $MS_e = 38.21$; $p = .11$; $\eta^2_p = 0.06$, with all three groups showing a significant overall word similarity effect—for the phonological, control, and semantic groups, $F(1, 65) = 14.38, 10.12$, and 2.97 respectively; $MS_e = 55.84$; $p < .01$. The number of participants showing a word similarity effect was 26 out of 25 for the phonological strategy group, 23 out of 25 for the control group, and 21 out of 25 for the semantically instructed group.
observed at a rapid rate, then we might expect to see the difficulty of processing longer words presented in Table 3. A mixed analysis of variance revealed a main effect of phonological similarity, $F(1, 39) = 30.23; MSe = 41.81; p < .01; \eta^2_p = 0.44$, no effect of strategy, $F(2, 39) = 2.29; MSe = 344.84; p = .11; \eta^2_p = 0.11$, and a significant interaction between these two factors, $F(2, 39) = 4.57; MSe = 41.81; p = .02; \eta^2_p = 0.19$. This interaction was investigated by tests of simple main effects. These showed significant effects of phonological similarity for the phonological and control groups, respectively, $F(1, 39) = 22.10$ and $16.74; MSe = 83.62; p < .01; \eta^2 = 0.36$ and $0.30$, but not for the semantic group, $F(1, 39) < 1; \eta^2 = 0.01$. As in Experiment 1, therefore, results showed that the induction of a semantic strategy resulted in the elimination of the phonological similarity effect. Again there was no effect of strategy with dissimilar words, $F(2, 39) < 1; \eta^2 = 0.03$, but there was a strategy effect for similar words, $F(2, 39) = 4.50; MSe = 203.70; p = .02; \eta^2 = 0.18$. This effect was due to the fact that similar words were better recalled in the semantic group than in the phonological and control groups, respectively, $F(1, 39) = 4.63$ and $7.86; MSe = 203.70; p < .05; \eta^2 = 0.11$ and $0.17$, revealing that the elimination of the phonologically similarity effect in the semantic group was mainly due to better recall of similar words.

A total of 11 of the 14 participants showed impaired recall for phonologically similar words for both the control and the phonological strategy groups compared to the 8 out of 14 under semantic coding instructions. Our results therefore suggest that both control and phonological groups continue to encode phonologically at this slower rate. Our final experiment is concerned with the word length effect. If, as we speculated, the inconclusive results of Experiment 2 reflect the difficulty of processing longer words presented at a rapid rate, then we might expect to observe clearer effects of strategy at the slower 2-second rate.

### EXPERIMENT 4

#### Method

Participants. Participants were 45 undergraduate volunteers from the University of Murcia, 15 in each strategy condition.

Stimuli and procedure. These were the same as in Experiment 2, except that the presentation rate was one item every 2 seconds (950 ms on, 1050 ms off).

#### Results and discussion

Means of the percentage of correct recall are presented in Table 4. A 2 (word length) × 3 (induced strategy) mixed analysis of variance showed a main effect of strategy, $F(2, 42) = 5.29; MSe = 327.12; p = .01; \eta^2_p = 0.20$, no overall effect of word length, $F(1, 42) = 1.76; MSe = 27.46; p = .19; \eta^2_p = 0.04$, and a significant interaction between strategy and word length, $F(2, 42) = 7.46; MSe = 27.46; p < .01; \eta^2_p = 0.26$. This interaction was studied by tests of main effects. These tests revealed that there was a main effect of word length in the phonological condition, $F(1, 42) = 9.66; MSe = 54.92; p < .01; \eta^2 = 0.19$, whereas the main effect of word length in the control group was not significant, $F(1, 42) = 2.02; MSe = 54.92; p = .16; \eta^2 = 0.05$. Unexpectedly, the main effect of word length was significant in the semantic condition, but in the direction of better performance on long words, $F(1, 42) = 5.00; MSe = 54.92; p = .03; \eta^2 = 0.11$. In contrast to Experiment 2, therefore, results showed that strategy induction had a marked impact on the effect of strategy comparisons.

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Semantic</th>
<th>Phonological</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissimilar words</td>
<td>78.51 (10.89)</td>
<td>73.10 (16.43)</td>
<td>75.12 (12.67)</td>
</tr>
<tr>
<td>Similar words</td>
<td>76.73 (12.31)</td>
<td>61.61 (13.83)</td>
<td>65.12 (16.37)</td>
</tr>
<tr>
<td>Difference</td>
<td>1.79 (6.35)</td>
<td>11.49 (10.20)</td>
<td>10.00 (10.32)</td>
</tr>
</tbody>
</table>

Means (and standard deviations) of the percentages of correct responses as a function of phonological similarity and induced strategy in Experiment 3.

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Semantic</th>
<th>Phonological</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short words</td>
<td>74.22 (11.95)</td>
<td>73.50 (16.63)</td>
<td>86.95 (12.10)</td>
</tr>
<tr>
<td>Long words</td>
<td>78.50 (9.17)</td>
<td>67.55 (15.32)</td>
<td>84.22 (13.39)</td>
</tr>
<tr>
<td>Difference</td>
<td>−4.28 (8.18)</td>
<td>5.95 (6.40)</td>
<td>2.72 (7.54)</td>
</tr>
</tbody>
</table>

Means (and standard deviations) of the percentages of correct responses as a function of word length and induced strategy in Experiment 4.
word length, an effect that was significant for both short and long words; respectively, $F(2, 42) = 4.55$ and 6.48; $MSe = 188.57$ and 166.01; $p < .03$; $\eta^2 = 0.18$ and 0.24. Short words were better recalled in the control group than in the semantic or phonological groups, respectively, $F(1, 42) = 6.44$ and 7.19; $MSe = 188.57$; $p = .01$; $\eta^2 = 0.13$ and 0.15, whereas long words were worse recalled in the phonological group than in the semantic or control groups, respectively, $F(1, 542) = 5.41$ and 12.55; $MSe = 166.01$; $p < .03$; $\eta^2 = 0.11$ and 0.23. As predicted, a majority of phonologically instructed participants showed better performance for short words (12 out of 15). The opposite effect was shown in the semantic group with 12 of the 15 showing a long word advantage, while in the control group 8 of the 15 showed a short word advantage.

**GENERAL DISCUSSION**

Results throughout this study showed that strategy instructions had marked consequences on the phonological similarity and word-length effects. With respect to the effect of phonological similarity, Experiment 1 showed better recall of dissimilar words in both phonological and control groups, while in accordance with our initial expectations the induction of the semantic strategy resulted in the elimination of the phonological similarity effect. Results were equivalent in Experiment 3, in which the presentation rate was slowed from one item per second to one item every 2 seconds.

Results regarding the word-length effect were more complex. In Experiment 2, strategy had only a marginal effect, showing a tendency in the predicted direction of a smaller word-length effect in the group instructed to use a semantic strategy. With a slower presentation rate in Experiment 4, however, the effect in the semantic group was not only eliminated but reversed, whereas the standard word length effect was found in the phonological group. Although this result was unexpected, reverse word-length effects have been described previously in the literature by Romani et al. (2005). In that study a reverse effect was found when words were presented under articulatory suppression, a procedure in which participants were instructed to articulate an irrelevant word or phrase repeatedly in order to prevent subvocalisation. According to Romani et al. (2005) this reverse effect could be a consequence of the fact that, under suppression, redintegration processes become more relevant, given that participants tend to rely on lexical-semantic representations in this situation. They suggest that long words would benefit more from redintegration because, being longer, they could provide more fragments of phonological information, reducing the number of suitable candidates in long-term memory (Brown & Hulme, 1995). However, this account raises the question of why a number of previous articulatory suppression experiments resulted in the mere elimination of the length effect rather than in its inversion (Baddeley, Chincotta, Stafford, & Turk, 2002; Baddeley, Lewis, & Vallar, 1984; Neath & Surprenant, 1995; Neath, Surprenant, & LeCompte, 1998).

An alternative, and perhaps more parsimonious, explanation could be that the reverse effect in the present experiment, and possibly that of Romani et al. (2005), was a consequence of some uncontrolled characteristic of the stimuli. While long and short words were matched on word frequency, rated familiarity, imageability, and concreteness, further examination of the two sets suggested the possibility that our longer words might be easier to relate to each other semantically.

We explored this possibility by asking a group of 48 undergraduates from the University of Murcia to rate, on a scale of 1 to 7, the level to which the words from each set could be semantically related to each other. The mean rating for long words was 4.15 ($SD = 1.64$) and for short words, 2.98 (1.36), a statistically significant difference ($M$ difference = 1.17; $SD$ difference = 1.78); $t(47) = 4.54; p < .01$. Although it was clearly undesirable that our long and short word sets differ in this way, it does in fact strengthen the evidence for the effectiveness of our two strategic instructions, leading to diametrically opposite word length effects based on phonological versus semantic strategy instructions. The absence of a clear effect in either direction in the control group, in contrast to a clear standard word length effect at the 1-second presentation rate, suggests that slower presentation leads to a wider range of strategies, with some participants still relying on phonological coding, while others were presumably using an alternative, possibly semantic code.

---

2 A reverse word-length effect has also been found by Caplan and collaborators (Caplan & Waters, 1994; Caplan et al., 1992), but this effect was probably a consequence of problems with the word sets, such as differences in phonological similarity (see Baddeley & Andrade, 1994; Mueller et al., 2003).
Table 5 shows the effect size of the phonological similarity and word length effects across the four experiments, expressed in terms of Cohen’s $d$. The effect size for both phonological similarity and word length was consistently greatest under phonological instruction, and least under semantic coding. Under control conditions we observed a substantial effect of phonological similarity at both rates of presentation, and a rather weaker word length effect at the 1-second rate, that was further reduced at the slower 2-second rate, probably because the reduced rate of presentation gave more scope for developing and using a wider range of strategies. Under semantic coding instructions, the effects of phonological similarity were weaker at both rates of presentation, while in the case of word length a modest decrement for long words at the faster rate reverses when more time is allowed. These results are consistent with the assumption that phonological coding is relatively easy at rapid rates of presentation, but may become less common when presentation rate decreases, allowing other strategies to develop.

Our results have implications at both a methodological and theoretical level. Methodologically, they indicate that the same material may be encoded semantically or phonologically, and that the relevant encoding strategies can be influenced by instruction. However, our results also indicate that there are limits to such strategic control, with verbal/phonological coding being particularly appropriate with rapid presentation, while effective semantic coding of unrelated words appears to require slower rates. Placing these results in the broader context of studies described in the introduction, phonological coding appears to be particularly appropriate when serial order is required under rapid presentation conditions. Provided the material broadly conforms to the speaker’s language habits, rehearsal is relatively easy and does not impose a demanding executive load. On the other hand, the phonological code is relatively impoverished compared to the richness of semantic encoding, leading to a more limited capacity and poorer long-term storage and retrieval (Craik & Lockhart, 1972). As such, it is optimal for short lists of items, where serial order is critical and semantic coding difficult.

In contrast, semantic coding is richer leading to better long-term encoding. As a result, it is advantageous in situations where the retention of item information is crucial, for example when large or open sets of words are used, rather than the small closed sets that characterise studies based on digits, letters, or limited word pools. It is, however, less easy to encode serial order semantically when arbitrary sequences of items are used, hence the need for more encoding time. However, this is not the case when sequences are able to take advantage of language habits, as in sentence recall (Baddeley & Ecob, 1970; Baddeley, Hitch, & Allen, 2008). Semantic coding is therefore likely to be helpful for sequences that exceed the optimal length for phonological encoding, and for tasks where item information is critical. Finally, one should note that, given appropriate material, both semantic and phonological encoding can proceed in parallel, with both enhancing performance (Baddeley & Ecob, 1970; Baddeley et al., 2008).

What of the theoretical implications of our results? They do not, and were not intended to, decide between theoretical options, but they do have implications for the way in which one might account for data of critical theoretical significance, namely a number of observations that phonological coding effects appear to disappear as list length increases (Hall et al., 1983; Hanley & Broadbent, 1987; Jones, Hughes & Macken, 2006; Salamé & Baddeley, 1986). One interpretation of this phenomenon is in terms of the abandonment of a phonological coding strategy (Baddeley & Larsen, 2007). The potential capacity to control strategy by instruction offers the possibility of testing this hypothesis directly.

### Table 5

Effect size (Cohen’s $d$) of phonological similarity and word length effects in Experiments 1–4

<table>
<thead>
<tr>
<th>Induced strategy</th>
<th>Experiment</th>
<th>Effect</th>
<th>Presentation rate</th>
<th>Semantic</th>
<th>Phonological</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>Similarity</td>
<td>1 s per item</td>
<td>0.33</td>
<td>1.79</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Word length</td>
<td>1 s per item</td>
<td>0.43</td>
<td>1.22</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Similarity</td>
<td>2 s per item</td>
<td>0.28</td>
<td>1.13</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Word length</td>
<td>2 s per item</td>
<td>−0.52</td>
<td>0.93</td>
<td>0.36</td>
</tr>
</tbody>
</table>
against other hypotheses; for example that longer list lengths simply overload the phonological system, rendering phonological coding unhelpful regardless of strategy. The results of such a study would be particularly relevant to the multi-component model of working memory (Baddeley, 2007), but would also inform other theoretical approaches to STM such as the feature hypothesis (Nairne, 1990) the SIMPLE model (Brown, Neath, & Chater, 2007) and Cowan’s (2005) embedded processes model, which do not necessarily assign a separate status to the process involved in short-term phonological storage.

In conclusion, our results reinforce the results of Hanley and Bakopoulou (2003) in suggesting that strategies can be reliably manipulated by instruction. Second, they provide preliminary support for the argument that differences in strategy may underpin some of the apparent differences in empirical findings within the field. They do not, of course, demonstrate that such explanations are in fact correct, but they do provide a way ahead, both through the direct manipulation of strategy under the relevant conditions and, in the longer term, through the development of indirect methods. For example, if strategy instruction can be used in neuroimaging studies to identify a clear signature for phonological and semantic strategies, then this would allow the monitoring of strategies used by participants in imaging studies, and as such should benefit both subsequent theoretical interpretation, and the standardisation of methods across laboratories.

Manuscript received 22 August 2007
Manuscript accepted 14 December 2007
First published online 30 March 2008

REFERENCES


 Neath, I., & Surprenant, A. M. (1995). The word-length effect and articulatory suppression. Unpublished manuscript, Purdue University, IN, USA.


### APPENDIX A

**Spanish nouns used in Experiments 1 and 3**

<table>
<thead>
<tr>
<th>Dissimilar word</th>
<th>Phonetic transcription</th>
<th>English translation</th>
<th>Similar word</th>
<th>Phonetic transcription</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOTE</td>
<td>/bo.te/</td>
<td>boat</td>
<td>LAGO</td>
<td>/la.xo/</td>
<td>lake</td>
</tr>
<tr>
<td>CIMA</td>
<td>/ki.ma/</td>
<td>summit</td>
<td>LAZO</td>
<td>/la.8o/</td>
<td>bow</td>
</tr>
<tr>
<td>FOCO</td>
<td>/fo.ko/</td>
<td>spotlight</td>
<td>MAGO</td>
<td>/ma.xo/</td>
<td>magician</td>
</tr>
<tr>
<td>MURO</td>
<td>/mu.ro/</td>
<td>wall</td>
<td>PALO</td>
<td>/pa.lo/</td>
<td>stick</td>
</tr>
<tr>
<td>PERA</td>
<td>/pe.ra/</td>
<td>pear</td>
<td>PATO</td>
<td>/pa.to/</td>
<td>duck</td>
</tr>
<tr>
<td>PINO</td>
<td>/pi.no/</td>
<td>pine</td>
<td>RAMO</td>
<td>/ra.mo/</td>
<td>bouquet</td>
</tr>
<tr>
<td>RUTA</td>
<td>/ru.ta/</td>
<td>route</td>
<td>SACO</td>
<td>/sa.ko/</td>
<td>sack</td>
</tr>
<tr>
<td>SOPA</td>
<td>/so.pa/</td>
<td>soup</td>
<td>VASO</td>
<td>/ba.so/</td>
<td>glass</td>
</tr>
</tbody>
</table>

### APPENDIX B

**Spanish nouns used in Experiments 2 and 4**

<table>
<thead>
<tr>
<th>Short word</th>
<th>Phonetic transcription</th>
<th>English translation</th>
<th>Long word</th>
<th>Phonetic transcription</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUBO</td>
<td>/ku.b3o/</td>
<td>bucket</td>
<td>GARAJE</td>
<td>/ga.ra.xe/</td>
<td>garage</td>
</tr>
<tr>
<td>FOCO</td>
<td>/fo.ko/</td>
<td>spotlight</td>
<td>JINETE</td>
<td>/xi.ne.te/</td>
<td>rider</td>
</tr>
<tr>
<td>GOMA</td>
<td>/go.ma/</td>
<td>rubber</td>
<td>MALETA</td>
<td>/ma.le.ta/</td>
<td>suitcase</td>
</tr>
<tr>
<td>LAGO?</td>
<td>/la.xo/</td>
<td>lake</td>
<td>NEVERA</td>
<td>/ne.be.r3/</td>
<td>fridge</td>
</tr>
<tr>
<td>NIDO</td>
<td>/ni.do/</td>
<td>nest</td>
<td>REGAZO</td>
<td>/re.ya.8o/</td>
<td>lap</td>
</tr>
<tr>
<td>PIPA</td>
<td>/pi.pa/</td>
<td>pipe</td>
<td>SOTANA</td>
<td>/so.ta.na/</td>
<td>cassock</td>
</tr>
<tr>
<td>TAZA</td>
<td>/ta.8a/</td>
<td>cup</td>
<td>TESORO</td>
<td>/te.so.ro/</td>
<td>treasure</td>
</tr>
<tr>
<td>VENA</td>
<td>/be.na/</td>
<td>vein</td>
<td>ZAPATO</td>
<td>/ba.pa.to/</td>
<td>shoe</td>
</tr>
</tbody>
</table>