Memory & Cognition
2003, 31 (5), 728-737

Word length and articulatory suppression affect short-term and long-term recall tasks

RICCARDO RUSSO and NICOLETTA GRAMMATOPOULOU
University of Essex, Colchester, England

In six experiments, subjects were required to recall (either serially or freely) lists of short and long words either immediately or after a study–test filled delay ranging from 30 to 60 sec. In three of these experiments, we investigated the effect that articulatory suppression had in modulating the word length effect in both immediate and delayed free recall. The results suggested that the effects of word length and articulatory suppression were comparable in the immediate and delayed recall tasks. The theoretical relevance of these findings is discussed.

Since the early 1960s, several researchers have interpreted the outcomes of empirical investigations of short- and long-term memory tasks as congruent with views suggesting the existence of dedicated short- and long-term memory storage devices (e.g., Atkinson & Shiffrin, 1968; Gillund & Shiffrin, 1984; Waugh & Norman, 1965). Short-term memory tasks are usually defined as consisting of the immediate retrieval of a limited number of stimuli that have been presented for a relatively short period of time during learning. Long-term memory tasks, on the other hand, have longer lists of stimuli and/or a relatively long delay imposed between learning and retrieval.

An influential model based on the preceding structural view of memory is the working memory model originally proposed by Baddeley and Hitch (1974; Baddeley, 1986). This model incorporates a limited-capacity controlling attentional system (i.e., the central executive) that supervises and coordinates the activities of dedicated limited-capacity slave systems used for temporarily storing small amounts of information. One of these slave systems in particular, the phonological loop, supports performance in short-term speech-based memory tasks such as the immediate serial recall of words or digits. This phonological loop comprises a temporary storage system (i.e., the phonological store) and an active system for rehearsing phonological information. Rehearsal refreshes the traces in the phonological store that are deemed to decay within approximately 2 sec (Baddeley, Thomson, & Buchanan, 1975).

Baddeley (1997) has summarized the main evidence in support of the existence of the phonological loop as comprising (1) the phonological similarity effect; (2) the irrelevant speech effect; (3) the word length effect; and (4) the effect of concurrent articulatory suppression during study. The phonological similarity effect consists in poorer short-term memory performance for words that are similar than for words that are dissimilar in sound or articulatory characteristics (see, e.g., Conrad, 1964). The irrelevant speech effect consists in poorer short-term memory performance for visually presented verbal information when irrelevant-speech–based material is played in the background (e.g., Salamé & Baddeley, 1982; but see, e.g., Jones & Macken, 1995). These results have been taken to support the view that the phonological loop holds phonological information.

The word length effect consists in poorer performance in short-term memory for long words than for short words (e.g., Baddeley et al., 1975; Cowan, Woods, & Borne, 1994). This effect has been explained by suggesting that more short words than long words can be rehearsed per unit time. Hence, the decaying traces of short words in the phonological store can be refreshed more efficiently than the decaying traces of long words (for a review on alternative interpretations of the word length effect, see Lovatt & Avons, 2001).

The final signature of the phonological loop is the effect of concurrent articulatory suppression at study. Articulatory suppression is achieved by asking people to repeat the same verbal information aloud continuously (e.g., the word *the*). This repetition is expected to affect the rehearsal of verbal information, hence refreshing the decaying phonological representation in the phonological store. As a consequence, the working memory model predicts that, when articulatory suppression is imposed during learning in a short-term memory task, fewer items should be recalled than are recalled in a standard learning condition. Congruent with this prediction, articulatory suppression affects immediate recall of verbal material (e.g., Baddeley et al., 1975).

Although the variables listed above have been used extensively in short-term memory tasks, there has been little investigation of their effect in long-term memory

We thank Nadia Gamboz, Lydia Tan, Laura Petrini, and Barbara Bacci for their help in data collection, and the members of the Memory Club of the Department of Psychology of the University of Essex for the useful discussions about the theoretical issues raised in the paper. Correspondence should be addressed to R. Russo, Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, England (e-mail: rrusso@essex.ac.uk).

tasks. The available empirical evidence seems to indicate that these variables may affect performance not only in short-term memory tasks, but also in delayed recall tasks. Longoni, Richardson, and Aiello (1993) showed that both immediate and 10-sec–delayed serial recall of lists of six words were affected when targets were phonemically confusable. Tehan, Hendry, and Kocinski (2001), using lists of four words, showed that short words were recalled better than longer words in both immediate and 12-sec–delayed serial recall tasks. Using a free recall task with no study–test delay, Watkins (1972) and Salthouse (1980) showed that more items were recalled from lists of 12 one-syllable words than from lists of either three- or four-syllable words. Finally, Richardson and Baddeley (1975) showed that articulatory suppression during learning impaired immediate and 20-sec–delayed free recall of lists 10 and 16 words long.

Overall, the available empirical evidence seems to suggest that some of the effects considered to be signatures of dedicated short-term storage devices to hold speech-based information occur in both short- and long-term memory recall tasks. However, given the relatively small number of studies showing the effect of these variables in long-term recall tasks, and given the relatively small range of study–test delays employed (either no delay or delays up to 20 sec), we decided to investigate more thoroughly the effect of word length on comparable immediate and delayed recall tasks. We report six experiments in which subjects were tested in either serial or free recall of lists of short and long words. Recall occurred either immediately or after a study–test filled delay ranging from 30 to 60 sec. In the last three of these six experiments, we investigated the effect of articulatory suppression in modulating the word length effect in both immediate and delayed free recall. To anticipate some of the main results obtained: It appeared that the effects of word length and of articulatory suppression were detected in both immediate and delayed recall tasks. The theoretical relevance of these findings will be discussed in the General Discussion section.

EXPERIMENT 1

The aim of this experiment was to assess the presence of the word length effect in immediate and delayed free and serial recall of lists of one- and four- to five-syllable long words. We used lists of six words, but with new words on each trial. The study–test delay lasted 30 sec. During this interval, subjects had to count aloud forward by 3 for 30 sec at a speed of about one number per second. At test, subjects were asked to recall the targets verbally.

Method

Subjects. Thirty-two undergraduate students from the University of Essex took part in this experiment. All were native English speakers. Sixteen were allocated to the free recall condition and 16 to the serial recall condition.

Materials and Design. A pool of 144 words was used to construct 24 lists of 6 words each. Twelve lists comprised short (i.e., one-syllable) words, and 12 lists comprised long (i.e., four- and five-syllable) words. The average frequency of occurrence of the words in these lists ranged from 80.3 to 88.7 according to the Kučera and Francis (1967) norms ($F < 1$).

Word length (long vs. short), study–test delay (no delay vs. 30 sec), and type of recall (free vs. serial recall) were the manipulated variables. Type of recall was manipulated between subjects; both word length and study–test delay were manipulated within subjects. Half of the subjects in each recall group were tested first in the no-delay condition and then in the 30-sec-delay condition, with this order reversed for the remaining subjects. The order of presentation of short- and long-word lists was pseudo-randomly arranged for each subject, with the constraint that no more than three lists of the same type were consecutively presented. The lists used for the immediate and delay conditions were randomly selected for each subject from the available pool. The order in which items were displayed within each list varied randomly for each subject.

Procedure. Each subject was presented with a total of 24 lists of six words each and told that some lists were made up of short words and other lists of longer words. Each word was displayed on a computer screen for 1.5 sec, with an interstimulus interval of 0.5 sec. Twelve lists were used in the no-delay condition and 12 lists in the 30-sec study–test delay condition. At the start of each delay condition, subjects were given a practice list of six two-syllable words.

In the no study–test delay condition, each list terminated with a beep. The subjects were told that as soon as they heard the beep, their task was to verbally recall all the items they could remember from the list just displayed. The subjects in the free recall condition were asked to recall the items verbally in any order they wanted. The subjects in the serial recall condition were asked to recall the items in the same order these were presented. In the serial recall condition only, if subjects did not remember a word, they were asked to say "blank" and to go on to the next word until the last item in the list. The subjects were given 30 sec to complete their recall. At the end of each recall phase, when the subjects were ready, a new list was presented.

The instructions for the 30-sec study–test delay condition were identical to those for the no-delay condition. The only difference was that, immediately after the last word in each study list was shown, a three-digit number was displayed. The subjects were told to count aloud forward by 3 for 30 sec (e.g., from 456) at a speed of about one number per second. At the end of this period, the subjects heard a beep and then had to recall the items as required (either freely or in serial order).

Results and Discussion

The free recall and serial recall data were analyzed separately.

Free recall. A 2 (word length: short vs. long) × 2 (study–test delay: no delay vs. 30-sec delay) within-subjects analysis of variance (ANOVA) was performed on the proportions of items correctly recalled by each subject in each condition (see Table 1). There was a significant main effect of word length [$F(1,15) = 39.77$, $MS_e = 0.012$, $p < .01$], indicating that more short than long words were recalled (.73 vs. .55). The significant main effect of study–test delay [$F(1,15) = 35.41$, $MS_e = 0.022$, $p < .01$] indicated that more items were recalled in the no-delay than in the 30-sec–delay condition (.75 vs. .53). Finally, the interaction was significant [$F(1,15) = 10.39$, $MS_e = 0.006$, $p < .01$], indicating that the word length effect was larger in the no-delay condition. Planned comparisons showed that the word length effect was significant for both the no-delay [$t(15) = 12.54$, $p < .01$] and the 30-sec–delay [$t(15) = 2.56$, $p < .025$] conditions.

Table 1
Experiment 1: Proportions of Words Recalled as a Function of Word Length (Short vs. Long), Study–Test Delay (Immediate vs. 30 sec), and Recall Instructions (Free Recall vs. Serial Recall)

	Free Recall				Serial Recall			
	Short		Long		Short		Long	
Study–Test Delay	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Immediate	.87	.08	.63	.09	.63	.16	.37	.14
30 sec	.59	.20	.47	.15	.42	.16	.29	.20

Serial recall. A 2 (word length: short vs. long) × 2 (study–test delay: no delay vs. 30-sec delay) within-subjects ANOVA was performed on the proportions of items recalled in the correct order by each subject in each condition (see Table 1). There was a significant main effect of word length [$F(1,15) = 36.60$, $MS_e = 0.017$, $p < .01$], indicating that more short words than long words were recalled (.53 vs. .33). The significant main effect of study–test delay [$F(1,15) = 33.13$, $MS_e = 0.001$, $p < .01$] indicated that more items were recalled in the no-delay than in the 30-sec–delay condition (.50 vs. .36). Finally, the interaction was significant [$F(1,15) = 6.06$, $MS_e = 0.012$, $p < .05$], indicating that the word length effect was larger in the no-delay condition. Planned comparisons showed that the word length effect was significant for both the no-delay [$t(15) = 6.33$, $p < .01$] and 30-sec–delay [$t(15) = 3.11$, $p < .01$] conditions.

Overall it appeared that a reliable word length effect was detected in both free and serial recall at both study–test delay conditions, albeit the size of the word length effect was slightly reduced in the 30-sec–delay condition. It is important to notice that the word length effect was reliable also after a filled interval of 30 sec, and that this study–test delay should have been sufficient to allow either the decay or the displacement of any target words from any hypothesized short-term memory storage device.

EXPERIMENT 2

It could be argued that subjects relied more on semantic information during retrieval of target words in the delayed than in the immediate recall condition. This strategy could be one of the possible causes of the reduced word length effect in the delayed condition of Experiment 1. In this immediate recall condition, subjects might have relied more on the phonological characteristics of the items in order to recall the words. Phonological codes may be more transient than semantic codes, possibly because of the larger interfering effects from subsequently presented stimuli; hence semantic codes may be more usefully employed in long-term memory tasks (e.g., Baddeley, 1966). If this is the case, it should be possible to mitigate the difference in the size of the word length effect between immediate and delayed recall conditions by using items that are unlikely to be semantically encoded (i.e., nonwords). With nonwords, subjects should rely more on the use of phonological information to recall target items in both delay conditions. Hence, more comparable word length effects may be detected at both delays. Experiment 2 was nearly identical to Experiment 1, the only difference being that nonwords were used as targets instead of words. Moreover, to minimize the risk of floor effects, lists of five items were used.

Method

Subjects. Thirty-two undergraduate students from University of Essex took part in this experiment. All were native English speakers. Sixteen subjects were allocated to the free recall condition and 16 to the serial recall condition. None had taken part in the previous experiment.

Materials, Design, and Procedure. A pool of 120 pronounceable nonwords was used to construct 24 lists of 5 items each. Twelve lists were made of short nonwords (i.e., one-syllable nonwords), and 12 lists of long nonwords (i.e., four- and five-syllable nonwords). The short nonwords were taken from McCann and Besner (1987). Each long nonword was created by changing the first three or four letters of an English word. Long nonwords received, by a set of three native English speakers, an average rating of pronounceability of at least 5 on a 7-point rating scale (1 = *very low*; 7 = *very high*). Moreover, the long nonwords did not appear to bring to mind real words easily (i.e., each word had an average rating of less than 4 on the type of scale above). The design and the procedure were those used in Experiment 1, the only difference being that subjects read the nonwords aloud during learning. Moreover, in order to familiarize the subjects with the pool of nonwords used in the experiment, they were asked to read these items aloud before the beginning of the experiment.

Results and Discussion

Free recall. A 2 (nonword length: short vs. long) × 2 (study–test delay: no delay vs. 30-sec delay) within-subjects ANOVA was performed on the proportions of items correctly recalled by each subject in each condition (see Table 2). There was a significant main effect of item length [$F(1,15) = 213.8$, $MS_e = 0.004$, $p < .01$], indicating that more short nonwords than long nonwords were recalled (.54 vs. .31). The significant main effect of study–test delay [$F(1,15) = 80.6$, $MS_e = 0.007$, $p < .01$] indicated that more items were recalled in the no-delay than in the 30-sec–delay condition (.52 vs. .33). Finally, the interaction was marginally significant [$F(1,15) = 4.19$, $MS_e = 0.004$, $p < .06$]. Planned comparisons showed that the nonword length effect was significant at both the no-delay [$t(15) = 12.11$, $p < .01$] and the 30-sec–delay [$t(15) = 8.99$, $p < .01$] conditions.

Serial recall. A 2 (word length: short vs. long) × 2 (study–test delay: no delay vs. 30-sec delay) within-subjects ANOVA was performed on the proportions of

Table 2
Experiment 2: Proportions of Nonwords Recalled as a Function of Item Length (Short vs. Long), Study–Test Delay (Immediate vs. 30 sec), and Recall Instructions (Free Recall vs. Serial Recall)

	Free Recall				Serial Recall			
	Short		Long		Short		Long	
Study–Test Delay	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Immediate	.65	.10	.39	.07	.65	.10	.42	.15
30 sec	.43	.12	.23	.09	.45	.15	.21	.13

items recalled in the correct order by each subject in each condition (see Table 2). There was a significant main effect of nonword length [$F(1,15) = 170.2, MS_e = 0.005, p < .01$], indicating that more short nonwords than long nonwords were recalled (.55 vs. .31). The significant main effect of study–test delay [$F(1,15) = 58.3, MS_e = 0.011, p < .01$] indicated that more items were recalled in the no-delay than in the 30-sec–delay condition (.53 vs. .38). Finally, the interaction was not significant ($F < 1$). Planned comparisons showed that the nonword length effect was significant at both the no-delay [$t(15) = 8.46, p < .01$] and 30-sec–delay [$t(15) = 11.14, p < .01$] conditions.

The results indicated that with nonword targets, whose recall is likely to be mainly supported by the retrieval of phonological information, a robust nonword length effect was observed both for immediate recall and at 30-sec filled study–test delays. Hence, these results replicate those obtained in Experiment 1. Moreover, it appeared that with targets unlikely to be semantically encoded, the "word length" effect did not differ significantly in the two study–test delay conditions used. This result strengthens those obtained in Experiment 1 indicating that a variable considered to be a signature of short-term memory has comparable effects in equivalent short- and long-term memory tasks.

EXPERIMENT 3

The previous experiments demonstrated the presence of a significant word length effect in delayed free and serial recall of relatively short lists. It would be interesting to assess the presence of a word length effect in more traditional long-term recall tasks. Previous studies had already provided some positive evidence on this issue. Watkins (1972) and Salthouse (1980) showed that more items were recalled from lists of 12 one-syllable words than from lists of either three- or four-syllable words. On the other hand, Craik (1968) did not find a word length effect in the free recall of lists whose length ranged from 9 to 18 items. However, it is important to notice that no word length effect was found when 6-item lists were used. Since the immediate free recall of 6-item lists can be considered a short-term memory task, it appears that in Craik's study, word length did not affect performance in both short- and long-term tasks. A further feature of the preceding studies was that subjects started the free recall task as soon as each study list ended.

To provide clearer evidence on the effect of word length on delayed recall, in Experiment 3 the subjects were given lists of 14 words (either short or long) to study. The subjects were then required to free recall any item they could after a 45-sec filled interval imposed at the end of each study list.

Method

Subjects. Sixteen undergraduate students from the University of Essex took part in this experiment. All were native English speakers. None had taken part in the previous experiments.

Materials and Procedure. A pool of 84 words was used to construct six lists of 14 items each. Three lists were made of one-syllable words, and three lists of four- and five-syllable words. The average frequency of the words in these lists according to the Kučera and Francis (1967) norms ranged from 72.6 to 83.6. Their average familiarity rating ranged from 507 to 542, and their imagery value ranged from 397 to 422, according to the ratings reported in the MRC psycholinguistic database (M. Coltheart, 1981). The order in which the lists were presented to each subject was pseudo-randomly arranged so that no more than two lists of each word type was presented in succession. The order in which the items were displayed within each list varied randomly for each subject. The procedure was the same as that employed in the longer study–test delay condition of Experiment 1, the only difference being that in Experiment 3 the delay was 45 sec instead of 30 sec. No immediate study–test condition was used in this experiment. The subjects were given 90 sec to recall the targets.

Results and Discussion

The mean proportion of words correctly recalled from short-word lists was .45 ($SD = .14$); for long-word lists, this figure was .38 ($SD = .15$). A one-way repeated measures ANOVA showed that this difference was significant [$F(1,15) = 6.70, MS_e = 0.007, p < .025$].

The results obtained in this experiment showed the presence of a reliable word length effect in the free recall of lists of 14 words following a filled study–test delay of 45 sec. This result complements and extends those obtained in Experiments 1 and 2. Overall, it then appears that word length similarly affects performance in short- and long-term recall tasks. In the next three experiments, we investigated the concurrent effect of word length and articulatory suppression in short- and long-term recall tasks.

EXPERIMENT 4

Concurrent articulatory suppression during visual presentation at learning is known to significantly reduce the word length effect in immediate recall tasks (see, e.g., Baddeley et al., 1975; V. Coltheart, Avons, & Trollope, 1989; Longoni et al., 1993). However, LaPointe and Engle (1990) have shown that articulatory suppression removes the word length effect only when target items are selected from a limited pool of stimuli (e.g., the same words are used as targets in different study lists). However, when LaPointe and Engle used new items in every study–test trial, the word length effect was not affected by articulatory suppression. In Experiment 4, we attempted to investigate the effect that articulatory suppression during visual presentation of items at learning can have on the word length effect in a long-term memory task. The subjects were asked to freely recall lists of 14 words following a study–test filled delay of 60 sec. Lists were made of either short or long words, and the subjects were required either to study the lists silently or to continuously repeat aloud the word *Coca-Cola* during the study phase. As in Experiment 3, different targets were used in each study list.

Method

Subjects. Forty-two undergraduate students from the University of Essex took part in this experiment. All were native English speakers. None had taken part in the previous experiments.

Table 3
Experiment 4: Proportions of Words Recalled as a Function of Word Length (Short vs. Long) and Study Condition (Standard vs. Articulatory Suppression)

	Word Length			
	Short		Long	
Study Condition	*M*	*SD*	*M*	*SD*
Standard	.39	.10	.34	.11
Articulatory suppression	.33	.09	.30	.09

Materials, Design, and Procedure. The same set of words used in Experiment 3 and a new set of 24 words were used to construct eight lists of 14 items each. Four lists were made of one-syllable words, and four lists of four- and five-syllable words. The average frequency of the words in these lists according to the Kučera and Francis (1967) norms ranged from 72.6 to 83.6. Their average familiarity rating ranged from 507 to 555, and their imagery value ranged from 383 to 422, according to the rating reported in the MRC psycholinguistic database (M. Coltheart, 1981). We used a two-way within-subjects design. Each subject was presented with both short-word (i.e., S) and long-word (i.e., L) lists, and the presentation of these lists occurred under both standard and articulatory suppression conditions. The order in which the lists were presented to the subjects was either LSSL or SLLS. The lists used for the standard and articulatory suppression conditions were randomly selected for each subject from the available pool. The order in which items were displayed within each list varied randomly for each subject. For each subject, the same list order was used in both the standard and the articulatory suppression conditions. The order in which the standard and the articulatory suppression conditions were administered was counterbalanced across subjects. Because of an error in allocating subjects to the experimental conditions, the counterbalancing between list order and study conditions was not perfect (i.e., 23 subjects were tested in the articulatory suppression condition first, 8 with the LSSL and 15 with the SLLS list order; 19 subjects were tested in the standard condition first, 11 with the LSSL and 8 with the SLLS list order).

The procedure in the standard condition was the same as that employed in the longer study–test delay condition of Experiment 1, the only difference being that in Experiment 4 the delay was 60 sec. In the articulatory suppression condition, the subjects were told that during the presentation of the study list they had to repeat out aloud, with a rate of about one occurrence per second, the word *Coca-Cola*. The subjects were given 90 sec to free recall the targets.

Results and Discussion

A 2 (word length: short vs. long) × 2 (study condition: standard vs. articulatory suppression) within-subjects ANOVA was performed on the proportions of items correctly recalled by each subject in each condition (see Table 3). Both main effects were significant, indicating that more items were recalled from lists of short words than from lists of long words [.36 vs. .32; $F(1,41) = 14.5$, $MS_e = 0.05$, $p < .01$], and that more words were recalled in the standard than in the articulatory suppression condition [.36 vs. .32; $F(1,41) = 13.5$, $MS_e = 0.007$, $p < .01$]. The interaction was not significant ($F < 1$).[1] Planned comparisons showed that the word length effect was significant in the standard condition [$t(41) = 2.95$, $p < .01$] and was marginally significant in the articulatory suppression condition [$t(41) = 1.94$, $p < .06$].

The results obtained in this experiment indicated that both word length and articulatory suppression significantly affected performance in the delayed free recall of lists of 12 words. Thus Experiment 4 replicated the results obtained by Richardson and Baddeley (1975) and extended them; both effects were now obtained following a longer retention interval. Moreover, from the results obtained in Experiment 4, it appeared that word length and articulatory suppression affected delayed recall performance in an additive way. With this finding, the results obtained by LaPointe and Engle (1990) in an immediate recall task were also obtained in a delayed recall task.

In the next experiment, we intended to further assess the effect that articulatory suppression during visual presentation of items at learning can have on the word length effect in a short-term memory task. Subjects studied lists of seven words either under articulatory suppression or silently. Immediately after the end of the learning phase, they were asked to free recall as many items as they could from the immediately preceding study list.

EXPERIMENT 5

Method

Subjects. Sixteen students participated in this experiment. All were native English speakers. None had taken part in the previous experiments.

Materials, Design, and Procedure. The pool of words used in Experiment 4 was used to construct 12 lists of seven words each. We used lists of seven words instead of six words, as in Experiment 1, to reduce the risk of ceiling effects. Six lists were made of one-syllable words, and six lists of four- and five-syllable words. As in Experiment 4, each subject was presented with both short-word and long-word lists and the presentation of these lists occurred under both standard and articulatory suppression conditions. Half of the subjects were tested under articulatory suppression first; the remaining half was tested in the standard condition first. Within each study condition, half of the subjects received the short-word lists first; the remaining half received the long-word lists first. The procedure used in the standard condition was the same one employed in the no study–test delay condition of Experiment 1. In the articulatory suppression condition, the subjects were told that during the presentation of the study list they had to repeat out aloud, with a rate of about one occurrence per second, the word *Coca-Cola*. The subjects were given 90 sec to recall the targets.

Results and Discussion

A 2 (word length: short vs. long) × 2 (study condition: standard vs. articulatory suppression) within-subjects ANOVA was performed on the proportions of items correctly recalled by each subject in each condition (see Table 4). Both main effects were significant, indicating that more items were recalled from short- than from long-word lists [.64 vs. .53; $F(1,15) = 19.4$, $MS_e = 0.001$, $p < .01$] and that more words were recalled in the standard than in the articulatory suppression condition [.65 vs. .51; $F(1,15) = 38.1$, $MS_e = 0.008$, $p < .01$]. The interaction was not significant ($F < 1$). Planned comparisons showed that the word length effect was significant both in the standard condition [$t(15) = 3.81$, $p < .01$] and in the articulatory suppression condition [$t(15) = 3.69$, $p < .01$].

Table 4
Experiment 5: Proportions of Words Recalled as a Function of Word Length (Short vs. Long) and Study Condition (Standard vs. Articulatory Suppression)

	Word Length			
	Short		Long	
Study Condition	*M*	*SD*	*M*	*SD*
Standard	.71	.13	.59	.14
Articulatory suppression	.56	.14	.46	.10

The same pattern of results obtained in Experiment 4 was obtained in Experiment 5, where subjects performed the immediate free recall of lists of seven words. Overall it appeared that the effect of word length and articulatory suppression was equivalent in short-term and long-term memory tasks. Experiment 5 also replicated the results obtained by LaPointe and Engle (1990) in an immediate recall task. They showed that using an unlimited pool of items, as we did in the present study, articulatory suppression did not remove the word length effect. Hence, articulatory suppression seems to significantly reduce the word length effect only when targets are repeatedly selected from a fixed pool of stimuli (see, e.g., Baddeley et al., 1975; V. Coltheart et al., 1989; Longoni et al., 1993).

In the next experiment, we investigated the effect of articulatory suppression on the word length effect in immediate and delayed serial recall of short lists of words selected from a closed pool of items. This has been one of the most common methods of investigating the effect of variables such as word length and articulatory suppression on short-term memory (e.g., Baddeley et al., 1975; V. Coltheart et al., 1989; Longoni et al., 1993; Tehan et al., 2001). Therefore, in Experiment 6, we intended to extend the use of this method to a delayed serial recall task. The subjects were given lists of either the same short or the same long words to be studied silently or under articulatory suppression. The subjects were then asked to serially recall the items studied either immediately after their presentation or after a 45-sec filled interval.

EXPERIMENT 6

Method

Subjects. Forty-eight students participated in this experiment. All were native English speakers. None had taken part in the previous experiments.

Materials. One pool of 10 short words and one pool of 10 long words were selected from the larger pool of words used in Experiment 4. The short words were *Race*, *Edge*, *Flow*, *Sin*, *Draw*, *Mood*, *Cause*, *Myth*, *Cast*, and *Trust*. The long words were *Representative*, *Independence*, *Mechanical*, *Territory*, *Sympathetic*, *Comparison*, *Communication*, *Opportunity*, *Establishment*, and *Literature*. These two pools did not differ significantly with respect to word frequency, familiarity rating, or imagery value ($ps > .10$). Twelve different random samples of 5 short words and 12 different random samples of 5 long words were drawn from the two pools. Each sample was then used to randomly create 12 different study lists of short words and 12 different study lists of long words. Hence, the lists created from each set of words contained always the same words, but these were ordered in different ways. In summary, 12 sets of 12 study lists of short words and 12 sets of 12 study lists of long words were created. Each set was then used to test two different subjects. The same procedure was used to create study lists comprising 6 short words and study lists comprising 6 long words. To generalize the results obtained in this experiment, we used lists of two different lengths commonly used in immediate serial recall tasks.

Design and Procedure. We used a mixed factorial design, with list length (i.e., five words vs. six words) manipulated between subjects (24 subjects were randomly allocated to each list length condition). Word length (i.e., short words vs. long words), study condition (i.e., standard vs. articulatory suppression), and study–test delay (no delay vs. delay) were all manipulated within subjects.

Study conditions and study–test delay conditions were counterbalanced between subjects so that a random half of the subjects were tested in the no-delay study–test condition first and the remaining half in the delayed test condition first. Half of the subjects in each study–test order condition were first given the standard study condition and then the articulatory suppression condition; the remaining subjects were first tested in the articulatory suppression condition and then in the standard condition. In each of the four experimental conditions obtained by crossing the study–test delay and the study condition factors, subjects were given three lists of short words and three lists of long words to study and then to recall serially. Short- and long-word lists were randomly intermixed.

The procedures for the standard condition and for the articulatory suppression condition were the same as in Experiment 5. In the articulatory suppression condition, the subjects were told that during the presentation of the study list they had to repeat aloud, with a rate of about one occurrence per second, the word *Coca-Cola*. The study–test delay was either no delay or 45 sec. In this case, the subjects were asked to count backward by three. At the end of each study list, the subjects were given 90 sec to verbally recall the target words in the same order as that in which they had been presented. If the subjects did not remember a word they were asked to say "blank" and to go on to the next word until they reached the last item in the list.

Results and Discussion

A preliminary 2 (list length: five vs. six words) × 2 (study–test delay: no delay vs. delay) × 2 (word length: short vs. long) × 2 (study condition: standard vs. articulatory suppression) mixed ANOVA was performed on the proportions of items correctly recalled, in their serial order, by each subject in each condition (see Table 5). This analysis showed that the effect of the list length factor was significant [$F(1,46) = 7.85$, $MS_e = 0.11$, $p < .01$], indicating that a larger proportion of words was recalled from five-word than from six-word lists (.48 vs. .39, respectively). The list length × delay interaction was marginally significant [$F(1,46) = 3.73$, $MS_e = 0.04$, $p < .06$], indicating that the reduction in performance between the no-delay and the delay test conditions was larger in the five-word (i.e., .59 vs. .37) than in the six-word condition (i.e., .46 vs. .32). None of the remaining interactions involving the list length factor approached significance ($Fs < 2.4$, $ps > .10$). Therefore, the main analysis was performed after collapsing the levels of the list length factor. This analysis showed significant main effects of study–test delay [$F(1,47) = 78.7$, $MS_e = 0.04$, $p < .01$], word length [$F(1,47) = 95.0$, $MS_e = 0.02$, $p < .01$], and articulatory suppression [$F(1,47) = 75.8$, $MS_e = 0.02$, $p <$

Table 5
Experiment 6. Proportions of Words Recalled in Their Correct Order as a Function of Item Length (Short vs. Long), Study–Test Delay (Immediate vs. 45 sec), and Study Condition (Standard vs. Articulatory Suppression)

	Standard				Articulatory Suppression			
	Short		Long		Short		Long	
Study–Test Delay	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Six-Word Lists								
Immediate	.72	.19	.45	.19	.42	.17	.25	.17
45 sec	.38	.21	.31	.19	.32	.20	.25	.17
Five-Word Lists								
Immediate	.84	.14	.58	.17	.53	.19	.43	.17
45 sec	.40	.21	.36	.16	.38	.17	.34	.15
Pooled								
Immediate	.78	.17	.51	.19	.47	.18	.34	.18
45 sec	.39	.20	.33	.18	.35	.19	.30	.17

.01], indicating that a larger proportion of words was recalled in the immediate (no delay) than in the delayed recall (delay) condition (.52 vs. .34, respectively), in the short-word than in the long-word condition (.50 vs. .37), and in the standard than in the articulatory suppression condition (.50 vs. .36). The study–test delay × study condition interaction was significant [$F(1,47) = 41.7$, $MS_e = 0.02$, $p < .01$], indicating that the effect of suppression was greater in the no-delay than in the delay study–test condition. The study–test delay × word length interaction was significant [$F(1,47) = 33.5$, $MS_e = 0.02$, $p < .01$], indicating that the effect of word length was greater in the no-delay than in the delay study–test condition. The study condition × word length interaction was significant [$F(1,47) = 11.06$, $MS_e = 0.01$, $p < .01$], indicating that the word length effect was greater in the standard than in the articulatory suppression condition. Finally, the three-way interaction was significant [$F(1,47) = 6.13$, $MS_e = 0.02$, $p < .02$].

To further investigate the significant three way interaction two 2 (word length: short vs. long) × 2 (study condition: standard vs. articulatory suppression) within-subjects ANOVAs were conducted on the data obtained in the immediate and in the delayed recall conditions. In the immediate recall condition, both main effects were significant [Fs$(1,47) > 99.5$, $p < .01$]. The study condition × word length interaction was significant [$F(1,47) = 17.3$, $MS_e = 0.01$, $p < .01$], indicating that the word length effect was larger in the standard (.78 vs. .51) than in the articulatory suppression condition (.47 vs. .34). Planned comparisons showed that the word length effect was significant in both the standard and the articulatory suppression conditions [$t(23) = 10.64$ and $t(23) = 5.46$, ps $< .01$, respectively].

In the *delayed* recall condition, the main effect of word length was significant [$F(1,47) = 10.4$, $MS_e = 0.02$, $p < .01$], indicating that the proportion of short words recalled was larger than the proportion of long words recalled (.37 vs. .31). The main effect of study conditions was marginally significant [$F(1,47) = 3.7$, $MS_e = 0.01$, $p < .065$], indicating that the proportion of words recalled in the standard condition was larger than the proportion of words recalled under articulatory suppression (.36 vs. .32). The interaction was not significant ($F < 1$), indicating that the word length effect was equivalent in the standard (.39 vs. .33) and in the articulatory suppression (.35 vs. .30) conditions. Planned comparisons showed that the word length effect was significant in both the standard and the articulatory suppression conditions [$t(23) = 2.19$, $t(23) = 2.53$, ps $< .05$, respectively].

In summary, Experiment 6 showed that, when words were studied silently from a closed pool, the word length effect was significant in both immediate and delayed serial recall, therefore replicating and extending the standard findings of a significant word length effect in immediate serial recall to a comparable delayed serial recall task. Articulatory suppression reduced significantly the word length effect when recall was immediate. However, it is worth noticing that the word length effect was also significant under articulatory suppression when testing was immediate. In previous studies, more short than long words had been recalled under articulatory suppression; however, the relatively small sizes of the samples used may have prevented these differences from being significant (e.g., Baddeley et al., 1975; V. Coltheart et al., 1989). The word length effect was significant under both standard and articulatory suppression in the delay test condition. Moreover, articulatory suppression did not significantly reduce the size of the word length effect in the delay test condition. Overall, Experiment 6 indicated that word length and articulatory suppression affected in very similar ways immediate and delayed serial recall of words when these were selected from a closed pool of items.

GENERAL DISCUSSION

Summarizing present results, it appeared that in Experiment 1 a significant word length effect was found in both immediate and delayed (30 sec) free and serial recall of words. Experiment 2 replicated this finding and extended it to the use of target items more likely to be phonologically encoded and unlikely to be semantically processed (i.e., nonwords). In Experiment 3, a significant word length effect was found in a delayed (45 sec) free recall of lists of 14 words. Experiment 4 replicated this and extended it to the use of a study–test delay of 60 sec. Moreover, it appeared that articulatory suppression affected delayed free recall, and that recall performance was independently affected by articulatory suppression and word length. Experiment 5 showed that articulatory suppression and word length affected significantly and independently the immediate free recall of lists of 7 words. In all of these experiments, we used different items for each of the study lists. In Experiment 6, we used instead a closed pool of words (cf. Baddeley et al., 1975), and subjects were tested on a serial recall task. In this experiment, it appeared that the word length effect was significant at all study–test delay conditions and when recall followed both silent and articulatory suppression study conditions. Moreover, the word length effect was reduced by articulatory suppression, particularly at immediate test. Overall, the present study showed that word length and articulatory suppression, two variables considered to be signatures of a dedicated short-term storage device to hold speech-based information (i.e., the phonological loop) have comparable effects in both long-term and short-term memory tasks.

It could be argued, as an anonymous reviewer suggested, that all theories assuming separate short- and long-term memory storage devices propose that if information is not encoded in a short-term memory storage device, this memory cannot make it to a long-term storage device. Thus, all variables that negatively affect short-term memory should, as the present study demonstrates, also negatively affect long-term memory tasks. That is, short- and long-term storage devices should not be considered to operate independently, but rather to operate interdependently. Although it has been suggested that short- and long-term memory storage devices operate interdependently, not all theories provide a specification of the processes leading to the transfer of information from short- to long-term storage devices. For example, an efficient phonological loop is considered essential for long-term learning of phonological information (see, e.g., Baddeley, Papagno, & Vallar, 1988). However, no clear specification of the processes involved in the transfer of information from the phonological loop to a more permanent store has been provided either in qualitative accounts of the working memory model (e.g., Baddeley, 1986) or in quantitative models of the phonological loop (e.g., Page & Norris, 1998).

Although the present results may be handled by models in which short- and long-term storage devices operate interdependently, it could be argued that the present results can be more parsimoniously accounted for by theoretical approaches that do not postulate dedicated short- and long-term memory storage devices (cf. Nairne, 2002). As Robert Crowder (1993) has cogently expressed, "needing to retain information over brief intervals really means that humans beings require *memory*, not that it need be a dedicated subsystem with different properties from other subsystems" (p. 144). With respect to the present results, it could be argued that long words were not rehearsed to the same extent as short words were, and hence their recall was affected. Similarly, rehearsal might have been affected by articulatory suppression, so that recall suffered. Insufficient rehearsal of targets might have affected their recall in a number of ways. For example, less rehearsed items might have received less opportunity to be rehearsed in conjunction with other targets, so that poorly rehearsed items should have been less likely to be associated to other targets at learning. Hence, recall might have been impaired if this was based on interitem associations formed with other targets. Similarly, it is reasonable to assume that the phonological characteristics of poorly rehearsed items might not have been sufficiently encoded. Hence, if the retrieval of target items was based mainly on the recollection of their phonological traces, recall might then have been impaired.

This account of the present results does not rule out the possibility that the effect of word length and of articulatory suppression on recall could, in principle, be removed. This could happen, for example, when study conditions allow comparable encoding of either semantic/associative or phonological characteristics of short and long target words (a similar argument could be made for targets studied under articulatory suppression vs. standard learning conditions). For example, if targets were presented repeatedly during learning, this might induce comparable encoding of semantic/associative information for short and long words. If this hypothesis is correct, then the word length effect should be removed when target *words* are repeatedly presented during learning. However, repeated presentation of targets might have a limited influence on the word length effect when *nonwords* are used as targets. In the case of "novel words," there are no preexisting lexical–semantic representations. Without the support of preexisting lexical–semantic representations of targets, recall should mainly be based on their phonological characteristics. Since longer nonwords contain more phonological information than shorter nonwords do, then, given the same number of learning trials for both word and nonword targets, the word length effect might well be significant in the case of nonwords, whereas it could become nonsignificant for target words. Papagno and Vallar (1992) and Papagno, Valentine, and Baddeley (1991), using a paired-associate long-term learning task, obtained exactly this outcome. Subjects were asked to study either eight pairs of unrelated words (e.g., *fox–sign*) or eight word/nonword pairs (e.g., *oak–sumu*) in a series of study–test trials. At the end of each study trial, the subjects were given the first member of each pair as a cue to recall the second. There was a maximum of five study–test trials. The order in which items at learning and words at test were presented varied randomly in each study–test trial. Under these learning conditions, it appeared that although neither word length, nor phonological similarity, nor articulatory suppression affected words' learning, these variables did affect nonwords' learning. On the assumption that these variables interfere with an efficient functioning of the phonological loop, Papagno and colleagues interpreted their results as suggesting that efficient long-term learning of phonological information is disrupted by variables that negatively affect the functioning of the phonological loop. However, as indicated above, a more parsimonious account of their results would not necessarily need to postulate the existence of a dedicated phonological loop.

A nonstructural account can also be provided to explain similar results obtained by Baddeley et al. (1988), using the long-term paired-associate learning task described above, with a neuropsychological patient (P.V.) who showed a profound impairment in immediate recall tasks. They demonstrated that P.V. displayed an impairment in long-term learning, but only when word–nonword pairs were used. Baddeley and colleagues interpreted their finding as evidence that efficient phonological short-term storage is crucial for the acquisition of new phonological information. However, it is relevant to point out that this result does not necessarily prove that poor long-term learning of phonological information is the consequence of a defective phonological loop. It could well be the case that the brain damage sustained by P.V. might have led to the impairment of some crucial aspect of the processing of phonological information and/or of the storage of phonological information (not necessarily of some sort of phonological loop) essential to support performance in both immediate and delayed recall tasks.

The results of the present experiments may also be considered to be at variance with the model proposed by Cowan (1999). Cowan's model consists of three hierarchically embedded components: a central executive, a long-term memory store, and an automatic attention-orienting system. Short-term tasks are based on long-term memory information that has become temporarily activated. Central executive resources are allocated to activate the items to be attended. The focus of attention is capacity limited (i.e., about four items). Activated items that are not any more within the focus of attention are said to be in the "active memory." They rapidly decay to resting levels unless they are not brought back into the focus of attention. Reinstatement into the focus of attention can be achieved, for example, by rehearsing target items. The active memory component and the focus of attention in Cowan's model can be considered comparable to traditional conceptions of short-term storage devices.

According to Cowan's (1999) model, only the performance in immediate recall tasks depends on items within the focus of attention and in active memory. Thus, any condition that negatively affects the activation of targets in active memory should also affect immediate recall. Long words delay the output of the remaining targets in immediate recall tasks more than short words do. Since delayed output is associated with greater decay, then a word length effect should be observed in immediate recall. On the other hand, performance in long-term memory tasks does not depend on active memory, and hence a word length effect should not be observed in delayed recall tasks. Contrary to this prediction, we detected a significant word length effect in delayed recall tasks. Analogously, Cowan's model predicts that articulatory suppression, by limiting rehearsal opportunity, should be associated with a rapid decay of targets in active memory, and hence only immediate recall should be affected by articulatory suppression. Contrary to this prediction, we showed that articulatory suppression affected performance in both immediate and delayed recall tasks.

In summary, we have shown that signature effects of short-term storage devices are not confined to short-term recall tasks but are also present in long-term recall tasks. This finding should be coupled with other findings showing that variables that have been traditionally associated to long-term memory tasks also affect performance in immediate recall tasks. For example, Walker and Hulme (1999) showed that concrete words are more easily recalled than abstract words in immediate serial recall of short lists of words. Similarly, memory span appears to be affected by word frequency (e.g., Hulme et al., 1997) and by word co-occurrence (Stuart & Hulme, 2000). It then appears that variables that should have relatively selective effects on short-term memory also affect long-term memory tasks, and that variables that affect long-term memory tasks also affect immediate recall tasks. In light of these findings, we think that the present results can be more parsimoniously accounted for by memory models that do not distinguish between separate short- and long-term memory storage devices.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, **18**, 362-365.

Baddeley, A. {D.} (1986). *Working memory*. London: Oxford University Press.

Baddeley, A. {D.} (1997). *Human memory: Theory and practice* (rev. ed.). Hove, U.K.: Psychology Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In K.W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 8, pp. 67-89). New York: Academic Press.

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory & Language*, **27**, 586-595.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **14**, 575-589.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, **33A**, 497-505.

Coltheart, V., Avons, S. E., & Trollope, J. (1989). Articulatory suppression and phonological codes in reading for meaning. *Quarterly Journal of Experimental Psychology*, **42A**, 375-399.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, **55**, 75-84.
Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). New York: Cambridge University Press.
Cowan, N., Woods, N. L., & Borne, D. N. (1994). Reconfirmation of the short-term storage concept. *Psychological Science*, **5**, 103-106.
Craik, F. I. M. (1968). Two components in free recall. *Journal of Verbal Learning & Verbal Behavior*, **7**, 996-1004.
Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, **21**, 142-145.
Gillund, D., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.
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, & Cognition*, **23**, 1217-1232.
Jones, D. M., & Macken, W. J. (1995). Phonological similarity in the irrelevant speech effect: Within- or between-stream similarity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 103-115.
Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1118-1133.
Longoni, A. M., Richardson, J. T. E., & Aiello, A. (1993). Articulatory rehearsal and phonological storage in working memory. *Memory & Cognition*, **21**, 11-22.
Lovatt, P., & Avons, S. E. (2001). Re-evaluating the word-length effect. In J. Andrade (Ed.), *Working memory in perspective* (pp. 199-218). Philadelphia: Psychology Press.
McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in memory. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 14-24.
Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, **53**, 53-81.
Page, M. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, **105**, 761-781.
Papagno, C., Valentine, T., & Baddeley, A. {D.} (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory & Language*, **30**, 331-347.
Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology*, **44A**, 47-67.
Richardson, J. T. E., & Baddeley, A. D. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning & Verbal Behavior*, **14**, 623-629.
Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning & Verbal Behavior*, **21**, 150-164.
Salthouse, T. A. (1980). Age and memory: Strategies for localizing the loss. In L. W. Poon, J. L. Fozard, L. S. Cermack, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference* (pp. 47-65). Hillsdale, NJ: Erlbaum.
Stuart, G., & Hulme, C. (2000). The effects of word co-occurrence on short-term memory: Associative links in long-term memory affect short-term memory performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 796-802.
Tehan, G., Hendry, L., & Kocinski, D. (2001). Word length and phonological similarity effects in simple, complex and delayed serial recall tasks: Implications for working memory. *Memory*, **9**, 333-348.
Walker I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1256-1271.
Watkins, M. J. (1972). Locus of the modality effect in free recall. *Journal of Verbal Learning & Verbal Behavior*, **11**, 644-648.
Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, **72**, 89-104.

NOTE

1. In a supplementary analysis, we checked whether the unbalanced counterbalancing of word length order and of the order of study conditions might somehow have affected the results. It appeared that the main effects or interactions including these conditions were not significant [$F(1,38) < 2.81$, $p > .10$], with the following exceptions: (1) There was a significant interaction between the order in which subjects performed articulatory suppression and the effect of articulatory suppression on recall performance [$F(1,38) = 13.25$, $p < .01$], indicating that the effect of articulatory suppression was larger when subjects were tested in the standard condition first. (2) There was a significant interaction between the order in which word lists were presented (LSSL vs. SLLS) and the effect of articulatory suppression on recall performance [$F(1,38) = 4.38$, $p < .05$], indicating that the effect of articulatory suppression was larger if subjects were given word lists in the SLLS order.

Overall, these two significant interactions do not qualify the main outcome of Experiment 4 (i.e., that the effects of both word length and articulatory suppression were significant and that they affected recall performance in an additive way).

(Manuscript received August 1, 2002;
revision accepted for publication March 14, 2003.)