Journal of Experimental Psychology:
Learning, Memory, and Cognition
1999, Vol. 25, No. 2, 447–463

0278-7393/99/$3.00

Think Before You Speak: Pauses, Memory Search, and Trace Redintegration Processes in Verbal Memory Span

Charles Hulme and Philip Newton
University of York

Nelson Cowan
University of Missouri—Columbia

George Stuart
University of York

Gordon Brown
University of Warwick

Immediate memory span and speed of memory search were assessed for words and nonwords of short and long spoken duration. Memory span was substantially greater for words than for nonwords and for short than for long items, though speed of memory search was unaffected by either length or lexicality. An analysis of the temporal pattern of responses in the memory span task indicated that inter-item pauses were longer between nonwords than words but that these pause durations were unaffected by item length. A model of verbal short-term memory span is described in which trace selection from a short-term store and the redintegration (restoration) of degraded phonological traces both occur in the pauses between saying successive items. Both trace selection and trace redintegration appear to play important roles in accounting for individual differences in memory span.

Verbal memory span, the longest sequence of words a person can repeat in the correct order immediately after hearing them, is strictly limited. In adults, memory span is typically equal to six or seven monosyllabic words. The present study is concerned with understanding the reasons for this severe limitation to human information-processing ability.

An influential theoretical interpretation of the limits of memory span has come from trace decay with rehearsal models. According to these models, verbal short-term memory is considered to have a limited capacity, with items being represented by traces that decay within a short period of time (e.g., Baddeley, 1986; Broadbent, 1958; Schweickert & Boruff, 1986). However, decay can be overcome by rehearsal (subvocal articulation), which refreshes the decaying representation of items in memory. An influential model of this type is the articulatory loop (e.g., Baddeley & Hitch, 1974; Baddeley, Lewis, & Vallar, 1984). This model provides a parsimonious explanation for many short-term memory effects.

One of the key pieces of evidence for the concept of an articulatory loop is the word-length effect (Baddeley, Thomson, & Buchanan, 1975); the fact that participants can recall more short than long words in order. Recall of words varying in length from one to five syllables has been shown to vary directly as a function of how quickly the words can be articulated, and participants can recall as many words as they can say in just under 2 s. It has been claimed that the relation between memory span and speech rate can account for variations in memory span across different types of materials and different individuals (Schweickert & Boruff, 1986; Standing, Bond, Smith, & Isely, 1980). In the articulatory loop model it is assumed that the representation of an item decays to the point at which it can no longer be used within about 2 s unless it is refreshed before that time; short words can be refreshed at a more rapid rate because they can be articulated more quickly than long words.

Although there is a close relationship between speeded articulation rate and short-term memory span, not all short-term memory effects can be explained in terms of this relationship. Hulme, Maughan, and Brown (1991) compared memory span for words and nonwords. It was found that memory span for nonwords was lower than for words and in both cases a linear function related memory span to speech rate for items of differing spoken durations. The function for nonwords had an equivalent slope but a lower intercept.

These findings were interpreted in terms of two processing components that contribute to short-term memory span. One component, which is indexed by the difference in recall between short and long items, appears to reflect differences in the storage demands of long and short items. The other component, which is indexed by the differences in recall between familiar items (words) and unfamiliar items (nonwords), appears to reflect the operation of long-term memory mechanisms. To put this another way, unfamiliar items, such as nonwords, are remembered even more poorly than we

---

Charles Hulme, Philip Newton, and George Stuart, Department of Psychology, University of York, Heslington, York, United Kingdom; Nelson Cowan, Department of Psychology, University of Missouri—Columbia; Gordon Brown, Department of Psychology, University of Warwick, Coventry, United Kingdom.

This research was supported by Grant G9507279N from the Medical Research Council of the United Kingdom, by Grant R 000 236 216 from the Economic and Social Research Council of the United Kingdom, and by National Institute of Childhood Health and Diseases Grant HD-21338. We would like to thank Philip Quinlan for his help and advice and three anonymous reviewers for their incisive comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Charles Hulme, Department of Psychology, University of York, Heslington, York YO1 5DD United Kingdom. Electronic mail may be sent to ch1@york.ac.uk.

would expect from the speed at which they can be articulated. This appears to be because such items lack a long-term (lexical) representation of their phonological (spoken) form (for further details, see Brown & Hulme, 1995; Hulme et al., 1991; Hulme, Roodenrys, Brown, & Mercer, 1995). In subsequent work, Hulme, Roodenrys, Schweickert, Brown, Martin, and Stuart (1997) have developed the idea that familiarity facilitates a redintegration or "pattern completion" process that operates to reconstruct partially degraded item traces at retrieval (cf. Lewandowsky & Murdock, 1989; Nairne, 1990). We suggest that such redintegrative processing is necessary at retrieval (immediately before response production) in a short-term memory task and that such processing is separate from the response production processes (such as motor programming and articulation) that occur immediately after it. We think of redintegration as being akin to speech perception whereby a noisy input (the degraded memory trace) is recognized. One interesting, and perhaps nonobvious, corollary of this view is that redintegration is considered to be more useful for long than for short items because partial loss of information from the representation of a long item may have less serious consequences than an equivalent loss of information from the representation of a short item (consider, e.g., the effects of losing information specifying the identity of the phoneme /p/ in the word *hip* or *hippopotamus*). Brown and Hulme (1995) explored the implications of the idea that redintegration is a length sensitive process in a mathematical model of short-term memory. They showed that the assumptions that degradation of memory traces occurred on a segmental basis coupled with the idea that redintegration occurred more effectively for long than short items, and for words than nonwords, could explain a wide range of short-term memory phenomena.

The mechanisms underlying the word-length effect in serial recall remain contentious and the rehearsal based account now seems untenable, for both theoretical and empirical reasons. Theoretically, Brown and Hulme (1995) showed that a computational model of a simple decay-based memory system could produce effects of word length without the necessity of postulating a rehearsal mechanism. Empirically, it now seems that at least part of the word-length effect reflects output processes. Cowan et al. (1992) conducted an experiment in which the length of words varied across the two halves of each list. The results showed that performance was worse when the first half of the list comprised long words. A further experiment using forward and backward recall showed that performance was worse when the longer words were recalled first, regardless of position in the stimulus list. The participants were not instructed whether recall would be forward or backward for each list until the list had been presented, so presumably they would rehearse all lists in the same manner. These results imply that a major part of the word-length effect is due to memory loss during output rather than during rehearsal, at least for the lists of around span length used in this experiment.

Further evidence for the importance of output effects was advanced by Cowan (1992). This article reported a fine-grained temporal analysis of 4-year-old children's spoken responses in a memory span task. Cowan found that the overall time for which a participant spoke when recalling the lists correlated well with their short-term memory performance. However, when this time was analyzed in terms of periods of speaking and silences between items, the inter-item pause durations correlated with span, not the mean word durations. Cowan amassed evidence for a model of memory span in which items yet to be recalled are "reactivated" in the pauses between items being spoken. These pauses, however, are short and would not allow time for the items to be covertly rehearsed, so the method of reactivation would have to involve a more rapid search of memory. In fact, the pattern of results obtained by Cowan (short pauses between successive items in the memory response and with pause lengths tending to increase as the number of items held in memory increased) was similar to previous results that Sternberg, Monsell, Knoll, and Wright (1978) and Sternberg, Wright, Knoll, and Monsell (1980) obtained in studies of adults' memory search during the speeded pronunciation of lists shorter than the participant's span.

Following on from this study, Cowan et al. (1994) studied memory span, speech rate, and output timing relationships in 4- and 8-year-old children. They found that the older children remembered more words than younger children and that their inter-item pauses when recalling the lists were shorter than those of the younger children. The length of words in the children's spoken responses, however, did not differ between age groups. In contrast, although both age groups recalled more short than long words, word length did not affect the length of inter-item pauses in the responses (though it did, as expected, affect the spoken duration of words in the response). The theoretical interpretation offered was that age and word length affect different mechanisms. Word length affects how long it takes to say each word, and therefore how much time there is for subsequent words in the list to be lost from short-term memory before they can be pronounced. On the other hand, age affects how rapidly and efficiently the participant engages in covert memory processes during inter-item pauses. These processes may not only permit the pronunciation of the next item, but also help to refresh the short-term memory traces of subsequent items.

In summary, these previous experiments have led to the idea that there are important processes involved in the response phase of memory span tasks that are not well represented in the currently popular working memory model of Baddeley (1986). The work of Cowan (1992) and Cowan et al. (1994) indicates that some of the processing that occurs during spoken responses, and perhaps particularly in the pauses between saying successive items, is critically related to memory performance (see also Cowan, Wood, Wood, et al., 1998). It appears that longer silent periods occur in participants' responses to longer lists and that longer silent periods occur in the responses of individuals with less well developed memory skills. However, it is not yet clear what processes are involved in the response phase of memory span tasks, and how these processes operate to affect memory performance.

The aim of the present study was to improve our understanding of the processes taking place in the response phase of memory span tasks. To that end, our strategy was partly manipulative and partly correlational in nature. The manipulative aspect was to vary the materials used in the memory span task. Two key variables were word length, which has been used before to investigate the basis of memory span response timing in children (Cowan et al., 1994), and lexicality, which has not been used in this way before in either children or adults.

The rationale for examining word length and lexicality effects can be summarized as follows. First, the absence of word length effects on the durations of silent periods in children's immediate memory responses (Cowan et al., 1994) suggests that covert rehearsal is not a critical process during those periods. We wished to determine if the same conclusion can be drawn for adult participants. Second, effects of lexicality on immediate memory can be attributed to the greater difficulty of a redintegration process for nonwords than for words. If redintegration is an important process taking place during the silent periods in the immediate memory response, then the silent periods should be longer between nonwords than words if redintegration is taking place during these pauses: This prediction has never been examined before. An additional prediction derived from the model described by Brown and Hulme (1995) is that there may be shorter pauses between long than short words. According to this model redintegration should operate more effectively (and therefore perhaps more quickly) for long than for short words.

Memory search, redintegration, and articulatory processing all are proposed to be independent mechanisms in span tasks (Brown & Hulme, 1995; Cowan et al., 1998). Because the span task offers no direct evidence regarding memory search, we explored the idea that memory search is a critical process for immediate serial recall using a version of Sternberg's (1966) search task in which the slope of the reaction time (RT)/set size function serves as a measure of the search speed per item in the set. If memory search is a mechanism that is independent from redintegration or articulatory processing, there may be no effect of lexicality or word length on search processes.

Clifton and Tash (1973) and Chase (1977) found no effect of word length on memory search speed, which is similar to what Cowan et al. (1994) found for the duration of interword intervals in memory span responses. We reexamined their findings here within our search task. We also examined memory search for both words and nonwords to determine the effects of lexicality on search rate. Previous studies have indicated that there may be effects of lexicality on memory search rate, but that they may be rather small. Brown and Kirsner (1980), using a set of 10 high-frequency monosyllabic words and 10 nonsense syllables, found slope values of 40 and 51 ms, respectively. In contrast, Puckett and Kausler (1984), using much larger item pools, found slope values of 41 ms for one- and two-syllable words and 81 ms for consonant–vowel–consonant nonwords. Both of these studies used visual presentation however, and we are not aware of any studies that have examined the effects of lexicality on memory search rate using auditory presentation as in the present study. This may be important because auditory presentation might result in modality-specific representations that carry phonetic information but decay gradually (Cowan, 1984; Cowan & Saults, 1995) and therefore go through intermediate points at which they can benefit from processes of redintegration (Hulme et al., 1997).

The correlational aspect of the present study involved relations between various timing measures and memory span. (This aspect was of only secondary importance, given that the data analysis process was extremely time consuming and therefore we were limited to 24 participants in the experiment.) First, there were the measures of preparatory intervals (the time between the end of a memory list and participants' beginning to repeat it) and inter-item intervals (the length of pauses between successive words in each participant's response). We wished to examine whether the length of preparatory intervals, or inter-item intervals, or both correlated with memory span performance. Cowan et al. (1994) found that older children recalled lists at their span length with shorter preparatory intervals than younger children. There was no such difference for inter-item pauses. However, when both age groups recalled lists of the same length, rather than span-length lists, older children recalled the lists with shorter inter-item pauses. Similarly, we expected that participants with longer memory spans could repeat lists of a particular length with shorter silent periods in their responses, resulting in correlations between those silent periods and memory span.

Second, we examined the correlations between memory span and memory search rate in the Sternberg (1966) task to investigate the proposal that memory search is an important process taking place in memory span. Cavanagh (1972) first proposed a relationship between search slope and span, and found a high correlation across a number of studies between mean memory search rate and mean memory span for different materials. Although Brown and Kirsner (1980) subsequently questioned whether such a relationship could be found on an individual-participant basis, Puckett and Kausler (1984) did find some evidence that it could. We reexamined this issue, and extended it to the auditory domain, using both words and nonwords of different lengths here.

Finally, in addition to reexamining the relation between search slope in this task and span, we examined the relation between the search task results and the other timing measures. If search is the critical process that is largely responsible for differences in the silent times between words in the response, there should be a correlation between search rate and inter-item silent intervals in the memory task.

The pattern of findings from previous studies of memory span, speech rate and memory scanning, relevant to the present study are summarized in Table 1. From these previous findings we confidently expected memory span to be higher for shorter items, and to be higher for words than for nonwords. The effects of length and lexicality on memory search speed have been less thoroughly investigated but the available studies suggest that item length does not affect search speed and that lexicality has only weak and possibly not reliable effects. No previous studies have investigated the effects of lexicality on memory response times, but tentatively we might expect that the covert memory processes involved in inter-item pauses would operate less efficiently for nonwords that are unfamiliar to participants than for familiar words. Finally, the findings summarized in Table 1 were culled from a range of studies using samples of both children and adults. No previous study has systematically investigated the within-subject relationships between measures of memory search, memory span

Table 1

Summary of Previous Studies of the Effects of Lexicality and Item Length on Measures of Memory Span, Speech Rate, Memory Search, and Memory Response Timing

Measure	Lexicality	Item length
Memory span	Words better than nonwords (Hulme et al., 1991)	Short better than long (Hulme et al., 1991)
Speech rate	No effect (Hulme et al., 1991)	Short faster than long (Hulme et al., 1991)
Memory search slope	Possible weak effect (Brown & Kirsner, 1980; Puckett & Kausler, 1984)	No effect (Chase, 1977; Clifton & Tash, 1973)
Memory response timing		
Word durations	Not known	Long greater than short (Cowan et al., 1994)
Inter-item intervals	Not known	No effect (Cowan et al., 1994)

and the timing of memory span responses, and how these relationships vary according to the lexical status or length of items that are to be remembered. This was a major aim of the present study.

Method

Each participant took part in three distinct procedures: memory span, speech rate, and memory search. For each of these procedures, the experimental materials were presented in a counterbalanced order. The order of testing for both variables, words versus nonwords and length, was counterbalanced across subjects in the same way for each procedure.

Participants

Twenty-four undergraduate and graduate students at the University of York participated in the experiment. Their ages ranged from 18 to 26 years, with a mean age of 18 years 9 months; 23 participants were female and 1 was male.

Materials

Two sets of eight words were used: eight one-syllable words (*Greece, maths, mumps, school, switch, stoat, scroll, zinc*), and eight five-syllable words (*Yugoslavia, physiology, tuberculosis, university, refrigerator, hippopotamus, periodical, aluminium*). These words were a subset of those originally used by Baddeley et al. (1975) and were matched for frequency and conceptual class.

Two sets of eight phonotactically legal nonwords were also used (Hulme et al., 1991): eight one-syllable items[1] (*bim, dof, fot, gug, mab, pid, sep, zog*), and eight three-syllable items (*arellum, bepavit, gossikos, jodazum, monosip, muttasek, tushebon, zegglepim*).

Procedure

The participants were each tested in two separate sessions, each lasting approximately 50 min. All of the tasks in the experiment were controlled by a Macintosh 7200/75 computer using an external amplified speaker to present the previously recorded and digitized items (see Cox, Hulme, & Brown, 1992).

Memory span. In the memory span procedure, before testing began, participants were initially asked to listen to and repeat each word and each nonword once to check their audibility.

To measure memory span, we presented participants with lists of items drawn randomly without replacement from each pool of eight items, at a rate of one item per second. Testing began with three-item lists for words and two-item lists for nonwords. Participants were required to listen to the lists and repeat them back in the order of presentation.

Participants were presented with four lists at each sequence length. If participants recalled any of the lists of a given length correctly, the length of the lists was increased by one item. Testing was discontinued when participants made errors on all four lists of a given length. Memory span was calculated as the greatest list length at which the participant could recall all lists correctly, plus .25 of a point for each subsequent list recalled correctly (as there were four lists presented at any length at which the participant had erred). Participants were instructed that they should say "pass" for any items they could not recall.

Speech rate. Following the memory span tasks, the participants' speech rate was measured. For each condition, the participants were presented with the eight items from that condition, in four pairs. The participants were instructed to repeat each pair 10 times as quickly as possible until told to stop by the experimenter, and the time taken to do this was recorded. The mean of these four times was then transformed to items articulated per second.

The entire memory-span and speech-rate measurement procedures were recorded on audiotape.

Memory search. Following the speech-rate measurement, participants undertook the memory search procedure. Participants were presented, by means of a computer, with an auditory warning signal consisting of a 150-ms tone. After an interval of 1 s, participants were presented with a spoken list of the stimulus words or nonwords of varying list lengths at a rate of one item per second. Lists of each stimulus type were blocked. Following a pause of 2.5 s at the end of each list, participants were then presented with a probe item and were required to decide whether it had been present in the preceding list. For each participant, the length of the lists varied randomly from trial to trial from one to four items. For each of the four stimulus types and each of the four list lengths, participants were presented with 15 positive trials (where the target had been presented) and five negative trials (where it had not); giving a total of 80 trials for each stimulus type. Participants were required to press the *Z* key on the computer keyboard if the probe item had been present in the list, or the period key if it had not. If participants failed to respond within 10 s, the computer recorded an error response and proceeded to the next trial. There was an intertrial interval of 1 s. RTs for correct responses were stored by the computer for analysis.

Results

The memory span task is discussed first, along with measurements of the timing of spoken responses in that task. This is followed by discussions of the speech rate task, the memory search task, and interrelations between the various measures. An alpha level of .05 was used for all statistical tests.

---

[1] It should be noted that the nonword *dof* (*doff*) is actually a low-frequency English word, though its inclusion in the stimuli only operates against our finding of lexicality effects in memory. The nonword *fot,* pronounced to rhyme with *lot,* is not a word in British English.

Memory Span

In Figure 1, the mean memory span scores for the long and short words and nonwords are plotted as a function of speech rate. As would be expected from previous studies (e.g., Hulme et al., 1991), it is clear that increased item length had a detrimental effect on memory span and on speech rate and that memory span was greater for words than for nonwords. The memory span advantage for words was clearly independent of any differences in speech rate, because in this experiment speech rate for the nonwords (which were recalled more poorly than the words) was greater than for the words.

The memory span scores were subjected to a two-way analysis of variance (ANOVA) with repeated measures on both variables, lexicality (word or nonword) and length (short or long). This revealed significant main effects of lexicality, $F(1, 23) = 57.12$, $MSE = 0.33$, and length, $F(1, 23) = 112.36$, $MSE = 0.19$, but no significant interaction between these variables, $F(1, 23) = 3.48$, $MSE = 0.18$.

Measures of the Timing of Spoken Responses in the Span Task

The audiotaped recordings of each trial in the memory span procedure were digitized and the spoken duration of items and intervals between items were measured using a waveform editor on a microcomputer. For trials on which a list was recalled correctly, three measures were taken: the preparatory interval (defined as the pause between the termination of the last list item and the start of the spoken response), the item durations (the spoken duration of each item in the response), and the inter-item intervals (defined as the length of pause between the termination of one item in the response and the beginning of the next item). These temporal measures were taken for every stimulus type in every trial and for every participant when data were available.

All participants had usable data for an analysis of the mean of these response measurements "at span" (i.e., when

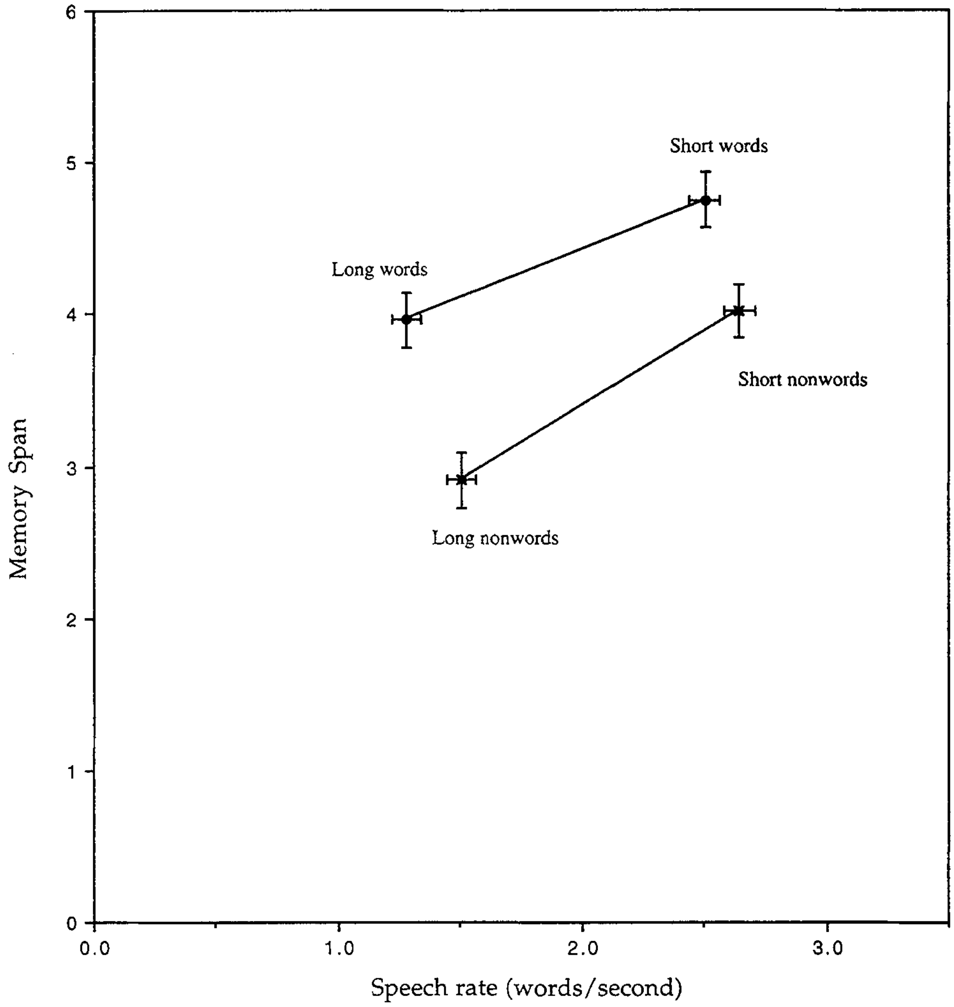

Figure 1. Mean memory span scores as a function of speech rate (and 95% confidence intervals, based on the interaction error term [Loftus, 1995]) for long and short words and nonwords.

the participant was entirely correct on all four trials at the highest list length given), though for 4 participants, in some conditions, perfect performance on four consecutive trials of a given list length was not achieved. In these few cases the data from correct responses on the shortest lists tested were taken as the closest approximation to at-span performance. (Equivalent analyses were also conducted on subspan and supraspan performance as well: These analyses yielded identical patterns to those for at-span performance and are not reported here, although some of the subspan measurements are later used in the correlational analyses.)

The means of the at-span preparatory intervals, item durations, and inter-item intervals from each stimulus type are shown in Table 2. The scores for each of these measures were subjected to a two-variable ANOVA with repeated measures on both variables, lexicality and length.

For preparatory intervals there was no significant main effect of lexicality, $F(1, 23) = 0.02$, $MSE = 60{,}117.51$; $\eta^2 = .0008$, nor length, $F(1, 23) = 1.24$, $MSE = 57{,}168.99$; $\eta^2 = .051$, and no significant interaction between these variables, $F(1, 23) = 0.50$, $MSE = 35{,}964.26$. Thus, the time taken to begin to repeat a span length list did not vary according to the length of the items contained in the list, nor according to whether the list items were words or nonwords.

For item durations there was a significant main effect of lexicality, $F(1, 23) = 260.97$, $MSE = 1{,}882.51$, and, as would be expected, a significant main effect of item length, $F(1, 23) = 823.10$, $MSE = 2{,}618.57$, but no significant interaction between these variables, $F(1, 23) = 0.07$, $MSE = 1{,}589.97$. The results of this analysis merely confirm that longer items (whether words or nonwords) took longer to say in the memory response than shorter items. The main effect of lexicality reflected the fact that the nonwords used in this experiment were, on average, of shorter articulatory duration than the words.

For inter-item intervals there was a significant main effect of lexicality, $F(1, 23) = 19.56$, $MSE = 22{,}236.25$, but no significant main effect of item length, $F(1, 23) = 1.24$, $MSE = 13{,}667.66$, $\eta^2 = .051$, and no significant interaction between these variables, $F(1, 23) = 1.46$, $MSE = 11{,}673.48$, $\eta^2 = .059$. Thus, participants paused for significantly longer between recalling successive nonwords than they did between successive words, a new finding; but the length of items in the list did not affect the duration of these inter-item pauses, consistent with previous research (Cowan et al., 1994). The model described by Brown and Hulme (1995) suggested that long words may be redintegrated more easily than short words and this might have been expected to lead to shorter pauses between long than short words. The absence of such an effect may reflect the fact that response preparation processes were also operating during the inter-item pauses and that these processes counteracted differences in the speed of redintegration between long and short words. In line with this suggestion, there is evidence to suggest that response preparation time is greater when participants repeat lists of long than lists of short words (Monsell, 1986). It is noteworthy that the pattern of inter-item intervals contrasted with the pattern for item durations. Participants paused longer between successive nonwords even though the nonwords used in this experiment were actually of shorter articulatory duration.

Table 2
Mean Preparatory Intervals, Item Durations, and Inter-item Intervals (and Standard Errors) for the Four Stimulus Types When Repeating Lists of Span Length

	Stimulus type			
Response measure	Short words	Long words	Short nonwords	Long nonwords
Preparatory intervals				
M	792	766	827	745
SE	71	54	48	46
Item durations				
M	540	837	395	697
SE	14	15	9	15
Inter-item intervals				
M	170	170	332	278
SE	23	26	28	39

Note. All numbers are in milliseconds. Standard errors were computed from the data for individual cells.

One other way of looking at the data from our response timing measures is in terms of the duration of the memory responses at span (Cowan, 1992; Cowan et al., 1994; Dosher & Ma, 1998). Doing so is important because we do not yet have a satisfactory estimate for theoretical purposes. Baddeley's (1986) working memory model predicts that the duration of spoken responses for correctly repeated span-length lists will be equal to the decay time of the phonological store (which has been estimated to be roughly 2 s). This proposition was suggested explicitly by Schweickert and Boruff (1986) and Schweickert, Guentert, and Hersberger (1990). However, actual measures of output times are few and have varied. Stigler, Lee, and Stevenson (1986) found a range of 1.3–4.6 s in output time for span-length lists, roughly consistent with Baddeley's theory. Dosher and Ma (1998) found a longer response period of 4–6 s, but they defined span as the length at which the participant was 50% correct, a definition that cannot be mapped easily onto Baddeley's model. The empirical question is, what is the duration of speech that corresponds to a participant's recall of a list at span?

In our data the means (and standard errors of the mean) for this measure were as follows: short words = 3.321 s (0.160), long words = 3.939 s (0.187), short nonwords = 2.785 s (0.119), long nonwords = 2.928 s (0.198). It is apparent that the duration of speech corresponding to memory span was greater for words (3.63 s) than nonwords (2.86 s), and for long (3.43 s) than short (3.05 s) items. These durations were clearly somewhat longer than the estimate of 2 s derived from Baddeley's (1986) model.

These scores were subjected to a within-subjects ANOVA in which the variables were lexicality (word or nonword) and word length (short or long). This analysis revealed significant effects of lexicality, $F(1, 23) = 16.85$, $MSE = 851{,}823.78$, and length, $F(1, 23) = 4.82$, $MSE = 721{,}776.11$, but no significant interaction between these variables, $F(1, 23) = 2.65$, $MSE = 511{,}234.00$. The findings match those obtained with interword pause measures but not the other timing measures described above, suggesting that pauses were the main type of response segment contributing to the overall response length.

Speech Rate

A two-way ANOVA with repeated measures on both variables, lexicality and length, was also carried out on the speeded speech-rate data. This analysis revealed significant main effects of lexicality, $F(1, 23) = 35.02$, $MSE = 0.02$, and length, $F(1, 23) = 526.49$, $MSE = 0.06$, but no significant interaction between these variables, $F(1, 23) = 1.95$, $MSE = 0.02$. Thus, the nonwords in this experiment were repeated more quickly than the words, and within both classes of items, long items were repeated more slowly than short items (see Figure 1).

Memory Search

Participants' error rates in the search procedure were low, with 2.2% errors for short words, 3.4% errors for long words, 2.1% errors for short nonwords, and 3.7% errors for long nonwords.

An arcsin transformation was performed on the proportions of errors in each condition. These transformed error values were then subjected to a within-subjects ANOVA with two levels for the variable of lexicality (word and nonword) and two levels for the variable of length (short and long). The results revealed no significant effect of lexicality on errors, $F(1, 23) = 0.03$, $MSE = 0.00$, but a significant effect of length, $F(1, 23) = 8.39$; $MSE = 0.001$, confirming that participants made more errors with long items. There was no significant interaction between the variables of lexicality and length, $F(1, 23) = 0.17$, $MSE = 0.00$.

RTs for correct responses only were analyzed. The set size/median RT functions for positive ("yes") responses to words and nonwords are shown in Figures 2 and 3, respectively. (Median RTs were used to compute slopes to minimize the effects of skew.) These functions took the expected form, with linear increases in RT as set size increased (cf. Sternberg, 1966). The functions for words were strikingly similar to those for nonwords. In both cases RT was quicker for short items, and, perhaps surprisingly,

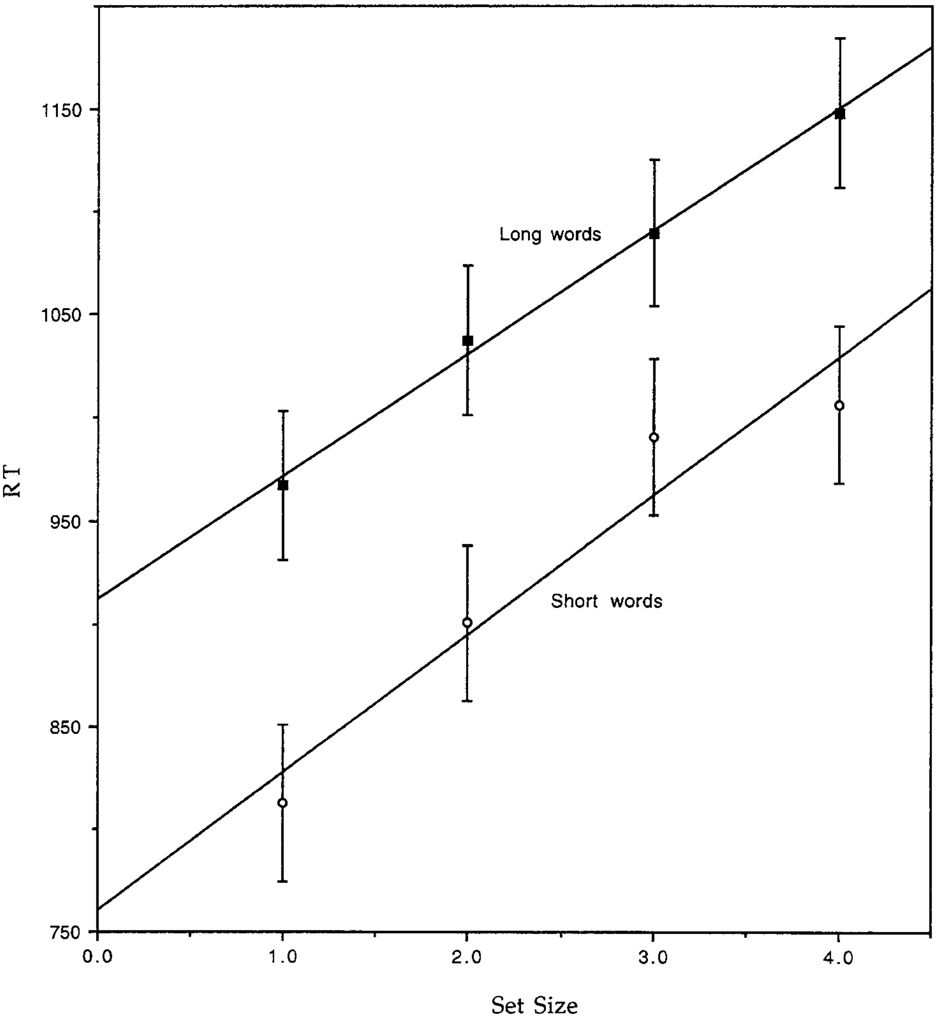

Figure 2. Median reaction time (RT; in milliseconds) as a function of set size for short and long words (and 95% confidence intervals, based on the error term for set size for each stimulus type [Loftus, 1995]).

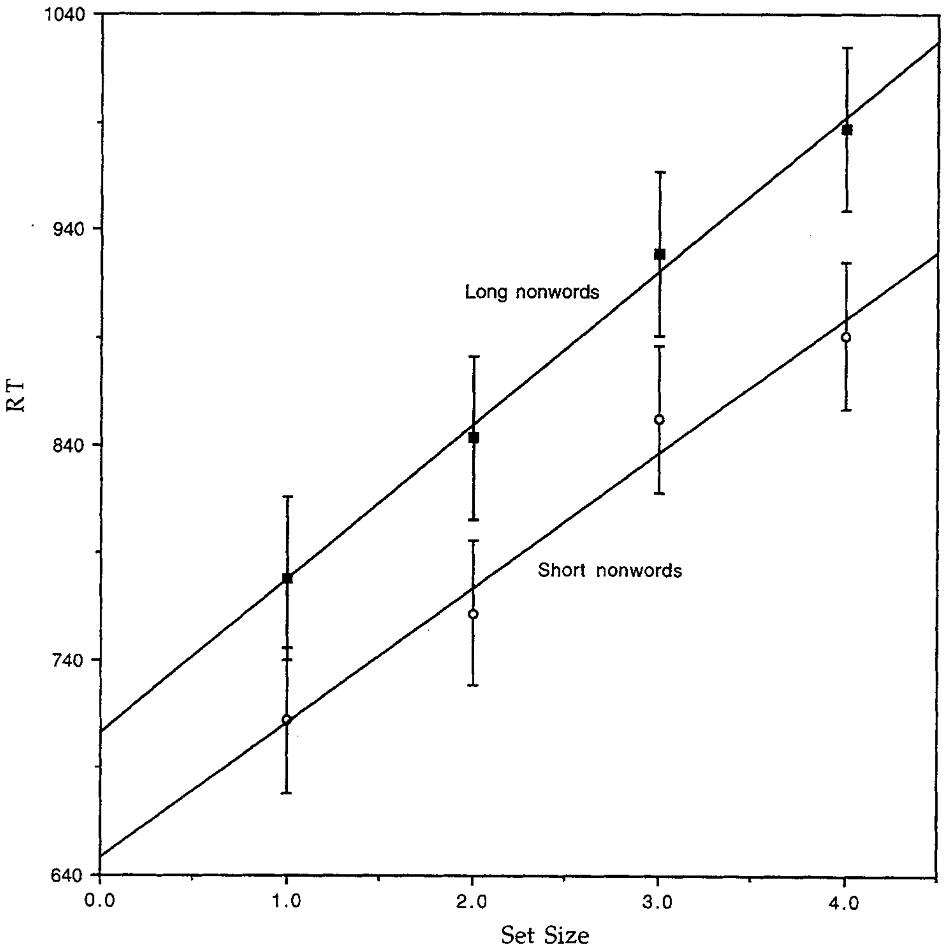

Figure 3. Median reaction time (RT; in milliseconds) as a function of set size for short and long nonwords (and 95% confidence intervals, based on the error term for set size for each stimulus type [Loftus, 1995]).

RT was slightly quicker for nonwords than words; however, it should be noted that the words used in the present experiment were longer than the nonwords and that this may have been responsible for differences in RT between the stimulus types.

Of particular interest are the effects of lexicality and item length on the slopes of the set size/RT functions. These slopes have traditionally been interpreted as a measure of the speed with which an internal memory search process operates. The mean slopes (±*SE*) across subjects of all four functions (in ms/item) were very similar: short words = 67.1 (±7.6), long words = 59.6 (±8.5), short nonwords = 62.6 (±8.4), long nonwords = 71.2 (±10.7), indicating that the speed of memory search did not appear to vary either as a function of lexicality or length.

The slope values for each participant from each stimulus type were subjected to a two-way ANOVA with repeated measures on both variables, lexicality and item length. Consistent with our conclusions about the similarity of the slopes, neither the main effect of lexicality, $F(1, 23) = 0.15$, $MSE = 1{,}994.87$, nor item length, $F(1, 23) = 0.004$, $MSE = 1{,}547.89$, was significant. In each case, the size of the effect was tiny ($\eta^2 = .006$ and .0001, respectively). The interaction between these variables also was not significant, $F(1, 23) = 1.76$, $MSE = 888.27$.

Summary of Results Based on Means and ANOVAs

To summarize the key evidence so far, memory spans revealed effects of both word length and lexicality, as expected. Also as expected from previous research (Cowan et al., 1994), there was no effect of word length on the durations of inter-item pauses in the responses. In contrast, the previously untested variable, lexicality, was found to have large effects on the inter-item pauses. Longer pauses for the nonwords occurred despite the finding that the speech rates in the speeded task were faster for the nonwords than the words in this study. Finally, the memory search results show that memory scanning rate did not measure the same processes as span or span response timing: Neither lexicality nor word length had any effect on the memory search slopes.

Interrelationships Between Measures of Memory Span, Speech Rate, Memory Search, and Memory Response Timing

A key issue in the light of theories in this area is the extent to which we can predict individual differences in memory span performance from the measures of speech rate and memory search rate (search slope) we have obtained. To this end, we examined the pattern of correlations between these measures. We also examined the extent to which individual differences in patterns of memory response timing (preparatory intervals, interword intervals, and item durations) predicted variations in memory span. In past studies, the silent periods between words in responses to correctly repeated lists of a fixed list length best indicated the participants' mnemonic capability (Cowan, 1992; Cowan et al., 1994, 1998). Therefore, for the response timing measures we selected subspan list lengths for which data were available for all participants. For this purpose we used List Length 4 for one-syllable words, List Length 3 for five-syllable words, List Length 3 for one-syllable nonwords, and List Length 2 for three-syllable nonwords.

Before considering the pattern of correlations obtained, it is necessary to consider the properties of the measures. The extent to which any variable can correlate with another is constrained by its reliability and variance (the greater the variance and the more reliable a measure, the higher its potential correlation with other variables will be). Table 3 presents the mean, standard deviation, and reliability (Cronbach's α) for each of the 28 measures included in our correlational analyses. It is clear that the measures had adequate variability and, in most cases, adequate reliability. It is notable that the measures of overall RT, word durations, speech rate, and span were highly reliable, whereas the memory search rate and the response interval measures were slightly less so. We consider issues of the relative reliabilities of the different measures further in the context of discussing the pattern of correlations obtained.

Correlations across subjects between the measures of memory span, speech rate, memory search, and memory response timing are shown in their entirety in the Appendix. Several generalizations are supported by these correlations, as follows.

Correlations between measures across conditions. First, and most simply, the groupings of related measures from the four conditions of the experiment tended to correlate quite highly and positively with each other (see the Appendix). For each measure there were six between-condition correlations (short words correlated with long words, short words correlated with short nonwords, etc.) and all six of these correlations were significant for speech rate, overall RT in the memory scanning task, and word durations. These high intercorrelations probably reflect in part the fact that these three sets of measures had the highest reliabilities of those included here. The between-condition correlations for memory span, preparatory intervals, and interword intervals were also mostly positive and significant.

Table 3
Means, Standard Deviations, and Reliabilities of the Measures Derived From the Study

Measure	*M*	*SD*	α
Memory span			
1-syllable words	4.74	0.64	.77
5-syllable words	3.96	0.63	.69
1-syllable nonwords	4.01	0.63	.80
3-syllable nonwords	2.91	0.63	.80
Speech rate[a]			
1-syllable words	2.52	0.35	.89
5-syllable words	1.28	0.12	.82
1-syllable nonwords	2.65	0.36	.92
3-syllable nonwords	1.51	0.16	.76
Memory scanning[b]			
RT			
1-syllable words	946	209	.98
5-syllable words	1,080	313	.98
1-syllable nonwords	805	208	.98
3-syllable nonwords	893	236	.97
Slope			
1-syllable words	76	58	.77
5-syllable words	56	45	.48
1-syllable nonwords	60	37	.60
3-syllable nonwords	76	53	.56
Memory response timing[b]			
Preparatory intervals			
1-syllable words	770	332	.50
5-syllable words	758	265	.79
1-syllable nonwords	885	216	.69
3-syllable nonwords	711	166	.63
Inter-item intervals			
1-syllable words	187	105	.76
5-syllable words	142	106	.69
1-syllable nonwords	316	125	.77
3-syllable nonwords	204	127	.80
Item durations			
1-syllable words	545	63	.89
5-syllable words	833	177	.87
1-syllable nonwords	404	49	.83
3-syllable nonwords	707	74	.88

Note. RT = reaction time. The means for preparatory intervals, inter-item intervals, and word durations in the memory span responses differ from those in Table 2 because the present means are for the subspan fixed list lengths used in the correlational analyses, not for the span-length lists reported in Table 2.
[a] Words per seconds. [b] Times (in milliseconds).

The one clear exception to this pattern of moderate to high between-condition correlations amongst measures comes for the measures of memory search slope. It is particularly notable that slope for one-syllable words hardly correlated at all with any of the other three measures of slope. This was not an artifact of poor reliability, however, because this measure had the highest reliability of any of the slope measures, and showed substantial correlations with other measures (notably measures of span). Furthermore, the two measures of memory search slope from the nonword conditions did correlate significantly and positively with each other in spite of the fact that their reliabilities were somewhat lower. It appears that the slope of the memory search function for one-syllable words was measuring something fairly reliably that was different from what was assessed by the slope for the other three sets of items (short and long nonwords and long words) used in this study. This difference may reflect the fact that only for short, familiar words were the phonological representations retrieved quickly and automatically along with each word's lexical identity.

Correlations between different timing measures. As might be expected, a number of the different timing measures correlated with each other (see the Appendix). In particular, among the measures of response timing in the memory-span task, the measures of preparatory intervals and interword pauses tended to correlate positively, indicating that there were probably some common mechanisms involved in determining the duration of these two types of pauses in the memory span task. It is particularly notable that within each condition, the correlation between preparatory intervals and inter-item intervals was always significant. In contrast, on the whole the correlations between word duration and preparatory intervals and word duration and inter-item intervals were lower and generally not significant.

We had three sets of speeded measures that were external to the memory span task: speech rate, overall RT, and slope from the Sternberg memory scanning task. The main thing to note here is that, perhaps surprisingly, measures of speech rate were quite highly correlated with measures of overall RT (see the Appendix). That is, participants who articulated words faster in our maximal speech rate task tended to respond faster overall in the Sternberg memory search task (which involved a nonverbal keypress response). Both measures may reflect a general speed of processing factor, as the research of Kail and Park (1994) would suggest (see also Kail and Salthouse, 1994). In contrast, measures of slope and overall RT in the Sternberg search task were very weakly intercorrelated. Hence, knowing how fast a participant responds in the Sternberg task gives no indication as to how great an increase in RT will be produced by including an additional item in the memory search set in that task.

The key patterns among the correlations from the Appendix between memory span and the timing measures are summarized in Table 4. First, consider the relationship between our measures of timing in the memory-span task and our measures of speeded responding from tasks external to the memory-span task. The duration of interword intervals in the span task correlated with speech rate. Cowan et al. (1994) found no such correlation, but their speech rate task involved only one pronunciation of each list per trial rather than cyclic repetition, and excluded preparatory intervals. It may be the cyclic nature of the speech rate task that results in its correlation with interword intervals in the present study.

The correlations between memory search slope and interword intervals were, on the whole, weak. However, we should note that the slope of the search function averaged across conditions did correlate with the inter-item intervals ($r = .41$) and the preparatory intervals ($r = .47$) when we considered data from correct responses to lists of span length, as opposed to the data from subspan lists that we have been considering so far. We took this as evidence that when participants are at their limits of performance in the memory span task, the time taken to perform a rapid search of the contents of short-term memory may place constraints on their ability to repeat the list correctly, as Cowan (1992) and Cowan et al. (1994) suggested.

Finally, the measures of speech rate and memory search slope also appeared more or less completely uncorrelated with each other, a finding that was obtained also by Cowan et al. (1998, Experiment 2) using less conventional procedures.

Correlations between memory span and timing measures. The issue of most theoretical interest is the extent to which the timing measures we have obtained predict variations in memory span. First, and contrary to some earlier studies (e.g., Baddeley et al., 1975), measures of speech rate correlated only weakly with variations in memory span. The only within-condition correlation between span and speech rate that was significant was for three-syllable nonwords, where the correlation was very substantial ($r = .67$).

If we consider the whole range of correlations between span and our processing measures, it is clear that the patterns of correlation differed between the word and the nonword conditions. There were many more significant correlations between our processing measures and memory span in the two nonword conditions than in the word conditions (see Table 4 and the Appendix). A reasonable summary of this pattern of relationships is to say that memory span for words

Table 4

A Summary of Important Correlations Between Timing Measures and Memory Span, Speech Rate, and Search Slope

Timing measure	Correlation with memory span	Correlation with speech rate	Correlation with memory search slope
Preparatory intervals in the span task	3 signif.; words only, $-.5 < r < -.38$	*ns*	2 signif.; nonwords, $r = .47, .51$
Interword intervals in the span task	4 signif.; nonword span only, $-.53 < r < -.44$	5 signif.; $-.60 < r < -.45$	2 signif.; $r = .42$, $-.43$ (opp. sign)
Word durations in the span task	*ns*	1 signif.; 5-syl. words, $r = -.44$	*ns*
Speech rate in the speeded articulation task	1 signif.; for 3-syl. nonwords, $r = .67$	—	*ns*
Scanning slope	3 signif.; 1-syl. words only, $r = -.61$, $-.69$, $-.43$	*ns*	—

Note. There were 16 correlations possible for each cell of the table. These comprise the 1-syllable words, 5-syllable words, 1-syllable nonwords, and 3-syllable nonwords for each of the two measures defining that cell correlated with one another. Signif. = significant; syl. = syllable; opp. = opposite.

was not well predicted by our processing measures, with the notable exception of the very strong correlation between span for one-syllable words and memory scanning speed for one-syllable words. In contrast, memory span for nonwords was predicted not only by memory scanning rate, but also by a number of our other speed of processing measures (see Table 4). Span for long unfamiliar items (three-syllable nonwords) seemed particularly sensitive to variations in processing speed between participants.

It is clear that this pattern of correlations cannot be attributed to variations in the reliabilities of the measures. So for example, none of our measures of word duration in the responses predicted span, although these are highly reliable measures. Conversely, our measure of memory scanning slope for one-syllable words was the best overall predictor of variations in memory span performance, but it was far from the most reliable measure included here. Our finding that memory search for one-syllable words was a good predictor of memory span for all stimulus types (with the possible exception of memory span for five-syllable words) confirms and extends previous findings of relationships between memory search rate and memory span (e.g., Puckett & Kausler, 1984). It is worth noting that Puckett and Kausler also found that memory search rate with visual presentation for short words was the only measure of memory search that, across subjects, significantly predicted memory span performance.

Both inter-item intervals and preparatory intervals correlated significantly with span in the nonword data, though the corresponding correlations in the word data were weaker and in no case significant (see Table 4 and the Appendix). Thus, at least for nonwords, better memory span performance was associated with shorter silent intervals in the memory response to subspan lists. This finding matches our earlier finding that young children with poorer memory spans pause for longer than older children with better memory spans (Cowan et al., 1994; see also Cowan et al., 1998).

Finally, and importantly, there was no significant correlation between the duration of spoken items in the response and memory span for any of the item sets used here. Thus, there is no evidence that the speed with which participants actually articulate the individual items in the memory response is related to individual differences in memory span performance for those items. This again is consistent with Cowan (1992) and Cowan et al. (1994) and at odds with the original working memory model (Baddeley, 1986; Schweickert & Boruff, 1986).

Regression analyses. Given that the processing speed measures tended to correlate positively with each other to varying degrees and that they also predicted span (particularly for nonwords) it became important to assess the relative power of these different speeded measures as predictors of memory span. To do this, we conducted four simultaneous regression analyses using each of the memory span measures (one-syllable words, five-syllable words, one-syllable nonwords, and three-syllable nonwords) as the dependent variable. There were five predictors in each equation: memory scanning speed for one-syllable words (because of its uniformly high correlation with the different measures of span), speech rate for three-syllable words (because it was the strongest correlate of memory span amongst our speech rate measures), overall RT for that stimulus type in the memory scanning task, length of preparatory intervals for that stimulus type, and length of interword intervals for that stimulus type.

These analyses assessed the extent to which each of the five independent variables made a unique contribution to the prediction of memory span when the effects of all other variables in the equation were controlled for. A summary of the results of these analyses is presented in Table 5. The results are straightforward. For one-syllable words and one-syllable nonwords, the only unique predictor of memory span was the speed of memory search. However, for three-syllable nonwords, in addition to this, there were also unique predictions from speech rate and the length of inter-item intervals. For five-syllable words, however, as would be expected from the pattern of correlations shown in Table 4, none of the predictors accounted for unique variance in span. Overall, our predictors were only weakly related to memory span for five-syllable words (total variance accounted for = 28%), whereas in the other conditions the same variables were powerful predictors of memory span performance (total variance accounted for ranged from 55% to 71%).

This pattern, with both maximal speech rate and memory search measures making independent contributions to the

Table 5

Summary of Simultaneous Regression Analyses (Unique r^2 Values for Each Variable) Predicting Memory Span in Each Condition From Measures of Processing Speed

Predictor	1-syllable words	5-syllable words	1-syllable nonwords	3-syllable nonwords
Slope 1-syllable words	.40***	.02	.26**	.13**
Speech rate 3-syllable nonwords	.00	.05	.00	.11*
RT[a]	.01	.06	.00	.01
Preparatory interval[a]	.06	.12	.07	.00
Interword interval[a]	.01	.00	.02	.09*
Total r^2 unique	.48	.25	.35	.34
Total r^2	.55	.28	.55	.71

Note. RT = reaction time.
[a] Measure from relevant condition.
$*p < .05$. $**p < .01$. $***p < .001$.

prediction of memory span, confirmed effects obtained by Cowan et al. (1998), who with different stimuli and procedures found independent predictive relationships between short-term memory and both articulation rate and memory search rate in children and adults, even though the latter two measures were not correlated with one another. Given the numerous procedural differences between these two studies, it is pleasing to see a convergence of evidence to support the conclusion that both memory search speed and articulation rate measures were independent predictors of individual differences in memory span.

Discussion

The results of the present experiment clarify a number of central theoretical issues concerning the mechanisms of verbal short-term memory and the origins of individual differences in short-term memory performance. The major effects obtained in this study, summarized in Table 6, include the following: (a) memory span was greater for words than nonwords, even though the words used here were articulated more slowly than the nonwords; (b) memory span was greater for short than long items; (c) there were longer inter-item pauses in the span task responses for nonword stimuli than for word stimuli, a novel effect; (d) there were no effects of lexicality or word length in a memory search task; (e) the duration of speech corresponding to memory span was greater for words (3.63 s) than nonwords (2.86 s), and for long (3.43 s) than short (3.05 s) items. In the sections that follow, we outline some of the important conclusions that follow from the present findings.

Limitations of the Articulatory Loop Model

We began by considering the idea that verbal short-term memory depends on the operation of an articulatory loop (Baddeley, 1986). The articulatory loop theory is one explicit statement of a trace decay with rehearsal mechanism. According to this theory, a representation of speech is held in a store that is subject to passive decay so that if an item cannot be recalled or rehearsed within about 2 s it will be forgotten. Rehearsal in this model is measured by the rate at which the items to be remembered can be articulated, and the effects of word length are explained by the idea that fewer long than short words can be articulated and so refreshed within the decay time of the store.

Our results pose some serious problems for this theory. In the first place, the duration of speech that can be remembered in spoken recall appears to be in the region of 3.6 s for lists of words. Thus, it appears that the temporal limits of the mechanism holding speech are longer than suggested in the articulatory loop model (roughly 2 s). More seriously, memory span for words (3.6 s) consists of a sequence of much longer temporal duration than for nonwords (2.8 s). Thus, the limits of span are not constant in terms of the time taken to articulate the material that can be remembered as postulated by the articulatory loop model (e.g., Baddeley, 1986; Schweickert & Boruff, 1986).

The finding that spoken memory span responses are longer than the 2-s duration of an articulatory loop is similar to a previous finding in a study with children (Cowan et al., 1994). In this study, the average length of a span length response was actually in the region of 4 to 5 s. The finding of a rather shorter response duration in adults is at first surprising. However, the inter-item pauses and the preparatory intervals for the children in this earlier study were considerably longer than those for the adults in the present study: Thus, much more of the duration of the children's responses consisted of silence. We take this difference to support the idea that an important determinant of forgetting from short-term memory is output interference. The act of recalling early items in a list actually harms one's ability to recall later ones. Adults recall more items than children, but because these items are recalled more quickly, and because the act of recall leads to forgetting, this results in a memory response of shorter temporal duration.

Our findings concerning the total duration of recall are consistent with data collected by Dosher and Ma (1998). Dosher and Ma examined both spoken and nonspoken recall and found that the duration of the response in span-length lists was fixed at about 3.5 to 6 s for spoken recall and at about 6 to 7 s for nonspoken recall. The fact that our estimates of spoken span (2.7 to 3.9 s) are lower than those of Dosher and Ma probably reflects differences in the definition of span between the two studies. We defined span as the longest list duration at which the participant recalled all lists correctly, whereas Dosher and Ma defined span as a list length corresponding to 50% correct recall. Another difference is that whereas Dosher and Ma examined span for lists of "forgetting-matched" words, we examined span for words and nonwords, the latter being distinctly more forgettable in Dosher and Ma's terms. The finding that people can recall words for a longer period than they can recall nonwords shows that, as Dosher and Ma assumed, the rate of forgetting during recall is based on factors that include more than simple memory decay. In our view, recall may continue longer for words than nonwords because the redintegration process is much easier for words, allowing recall from a more degraded memory trace. Our measurements of inter-item pauses in the span task responses support this hypothesis of why words are easier to recall than nonwords (see below).

Table 6
Summary of the Effects of Lexicality and Item Length on Measures of Memory Span, Speech Rate, Memory Search, and Memory Response Timing

Measure	Lexicality	Item length
Span	wd > nw	sh > lg
Speech rate	nw > wd	sh > lg
Search slope	*ns*	*ns*
Response timing		
Preparatory interval	*ns*	*ns*
Word durations	wd > nw	lg > sh
Interword intervals	nw > wd	*ns*
Duration of span response	wd > nw	lg > sh

Note. wd = word; nw = nonword; sh = short; lg = long.

Relationships Between Maximal Speech Rate and Memory Span

Within the articulatory loop model, speech rate was seen as a measure of the speed of rehearsal and rehearsal speed in turn was seen as a causal determinant of memory span (Baddeley, 1986). Contrary to this model, however, we have found that measures of maximal speech rate are weak predictors of individual differences in memory span. The only significant correlation between memory span and speech rate was for three-syllable nonwords. The absence of equivalent correlations in any of the other three conditions of the experiment (in the presence of other correlations involving span and speech rate in each of these conditions that were much higher) is strong evidence against the idea that speech rate is directly and causally related to memory span performance. In the past we have subscribed to such a theory as an explanation for developmental improvements in memory span with age (e.g., Hulme, Thomson, Muir, & Lawrence, 1984). In light of the present data and some other recent studies (cf. Cowan et al., 1994), such a simple causal theory no longer seems tenable. This is not to say that rehearsal plays no role in memory span tasks. Our claim is simply that it cannot bear the explanatory burden that is sometimes placed on it in theories such as the articulatory loop (cf. Brown & Hulme, 1995).

It remains possible that articulatory rehearsal does play a limited role in memory span under some conditions. The high ($r = .67$) correlation between speech rate and memory span for three-syllable nonwords found here is consistent with the view that, for those stimuli at least, articulatory rehearsal might be used. Alternatively, it might be that if participants find it difficult to articulate long unfamiliar items rapidly this contributes to greater output interference, and so more forgetting during the act of recalling these items.

The Timing of Spoken Memory Responses and Their Implications for Models of Memory Span

One major focus of the present study has been an analysis of the timing of spoken responses and the way in which these patterns may inform us about the mental processes occurring as participants recall a list. We have demonstrated that inter-item pauses in responses to correctly repeated span-length lists were substantially longer for nonwords than for words, though, in contrast and consistent with earlier findings (Cowan et al., 1994), these inter-item pauses did not differ between long and short items. The finding of substantial differences in inter-item pauses between words and nonwords is a new finding. This occurred even though the span-length responses included more items for word lists than for nonword lists and interword pauses are known to increase as a function of list length (Cowan, 1992; Cowan et al., 1994, 1997). In these previous studies, individual and developmental differences in interword pauses were not obtained when each individual was examined at a list length equal to span. The present study shows that the effects of lexicality on interword pauses are even more potent, showing up in a comparison of span-length lists of each type.

Our theoretical interpretation of inter-item pauses in spoken recall involves the idea that a number of covert processes relevant to recall occur during these pauses. It will be helpful to outline a general model of some of the hypothetical processes operating at recall in order to develop an argument relevant to our present findings.

We assume that after a spoken list of items has been presented in a short-term memory experiment, the participant encodes the list in a phonological code. This phonological code must represent the identity of items in the list and their order of occurrence. We argue that the participant must rely on this phonological representation to generate a spoken response corresponding to the list of items presented in the correct order. We assume that the phonological representation of the list of items is subject to degradation and an additional assumption would be that this degradation may, at least partly, be produced by output interference: That is to say, the act of recalling early items in a list harms the representation of later items in the list (Cowan, 1992; Cowan et al., 1994; Hulme et al., 1997).

For the participant to be able to perform this task, there must be some means of retrieving items from the phonological store. We assume for the moment that this involves some form of rapid search or inspection process similar to that first proposed by Sternberg (1966; though search need not be a serial process as Sternberg originally assumed; see, e.g., Ratcliff, 1978). It is necessary to assume that this inspection process has access to temporal or ordinal cues about item position within the store. Finally, we also assume that the retrieval of an item from the store is at least a two-stage process. At the first stage, we assume that a memory representation or trace of a candidate item is identified for a given list position that is to be recalled (trace selection). However, at this stage the trace that is identified may be incomplete due to degradation. At the second stage, we propose that incomplete traces must be restored to completion, or redintegrated, before they can be used as a response (trace redintegration).

To clarify these two stages, let us suppose that when retrieving the third item in a list the trace identified is *hip* ____ *tamus*. That is, the trace is incomplete, but specifies that it is a long item, and that the first syllable is *hip* and the last two syllables are *tamus*. This information may be quite sufficient for a redintegration process to recreate the complete trace of the item *hippopotamus*. Our argument is that redintegration may be akin to the process of speech perception and depends on knowledge of the phonological structure of spoken words (see, e.g., Hulme et al., 1991, 1995, 1997; Schweickert, 1993).

It should be clear from the foregoing that multiple processes are likely to be occurring in the pauses between recalling successive items from a list. However, from our perspective we would argue that the two retrieval processes (trace selection and trace redintegration) are likely to be major determinants of the duration of pauses. We therefore expect both factors that affect the speed of retrieving candidate items from a phonological store (e.g., list length) and factors that affect the ease of redintegration (e.g., lexicality) to affect the length of inter-item pauses. The present experiment has yielded support for both of these hypotheses.

We used the Sternberg (1966) memory search task as a measure of the speed of our first stage of retrieval from short-term memory (trace selection). We found that neither lexicality nor length affected the speed of memory search (measured by the slope of the RT/set-size function). These findings contradict the empirical generalization (based on between-subjects comparisons) suggested by Cavanagh (1972) that memory span is higher for materials with higher scanning rates. The very large differences in memory span between words and nonwords and between long and short items in our study were not associated with any difference in memory search rate for these same materials in the same participants. Our interpretation of this pattern is that trace selection at retrieval was not affected by the spoken length of the item nor by its familiarity.

We also found, however, that across subjects the speed of memory search correlated with the length of inter-item pauses (i.e., a participant's average memory search speed correlated with their average inter-item pause duration measured on lists of span length, $r = .41$, $p < .05$, though the equivalent correlations between memory search speed and inter-item pause durations measured on subspan lists were generally lower [see the Appendix]). This supports the idea that one determinant of the length of pauses in recall is how quickly participants can search the contents of short-term memory. We also found support for the idea that rapid memory search skills are related to good spoken memory span performance. In fact, as Table 5 shows, memory search rate for one-syllable words was the most powerful predictor of individual differences in memory span in the present study. It appears, therefore, that a major determinant of individual differences in memory span are variations in retrieval speed as indexed by variations in the rate of memory search.

It is probably important that in the present study it was memory search rate for one-syllable words that best correlated with memory span across conditions (an equivalent pattern was also evident in the data of Puckett & Kausler, 1984). Memory search for one-syllable words might plausibly be distinguished from search in the other conditions by the relative ease of the redintegration process; presumably monosyllabic words remain intact and can be searched relatively efficiently without the need for a lengthy redintegration process. In this view, search rate for one-syllable words may give the purest measure of the speed of a trace selection process that is unimpeded by variations in the ease of trace redintegration.

Our results also support the idea that trace redintegration is an important determinant of memory span performance and that this may be one of the major processes occurring in inter-item pauses. The comparison between words and nonwords gives us a manipulation to assess the role of redintegration: Nonwords lacking representations of their spoken form in lexical memory are difficult to redintegrate. One of the clearest findings of the present study was that participants pause for substantially longer between successive nonwords (roughly 300 ms) than between successive words (roughly 170 ms) when recalling a spoken list of span length. This suggests that in spoken recall a major determinant of inter-item pause duration is the difficulty of redintegrative processing, which is greater for nonwords than for words.

This difference in inter-item pauses is all the more striking given the fact that the response durations for span-length lists of words consist of considerably more speech activity (longer durations of speech with longer item durations) than the equivalent response durations for span-length lists of nonwords. This suggests that the words that are recalled successfully in span-length lists may have been subject to greater degradation (due to greater output interference) than the nonwords. The shorter pauses between consecutive words than nonwords at recall therefore occur despite the fact that the words used in the present experiment had longer spoken durations than the nonwords.

The further finding that, for nonwords, individual differences in inter-item pause durations correlated significantly with individual differences in memory span suggests that variations in participants' ability to redintegrate nonwords in the inter-item pauses may be a determinant of how well they can recall those items.

Summary and Conclusions

We have investigated the effects of length and lexicality on memory span and how these variables affect memory search and the pattern of pauses in responses to the memory span task. It is clear from our results that lexicality and length have very large effects on memory span, whereas neither of these variables has any substantial effect on the rate of memory search. Interestingly, we have shown for the first time that the pauses between successive items in the spoken memory span response are very much longer for nonwords than for words, whereas these same pauses are insensitive to the effects of item length.

We have argued that these findings are compatible with a multicomponent model of verbal memory span. According to the model we have described, a spoken list of verbal items is encoded into a phonological code (as proposed in many other models, most notably by the articulatory loop model of Baddeley, 1986). This phonological representation is subject to degradation and a major source of this degradation is output processes: The act of recalling early items in a list contributes substantially to the degradation of later items held in memory. We have emphasized that one major component of our model is a retrieval mechanism and that retrieval involves at least two separable components. At the first stage, we assume that the phonological memory representation of a candidate item is identified for a given list position that is to be recalled (trace selection) but that for many items a second stage involving the redintegration (restoration) of the retrieved trace is necessary before the trace can be used to generate a response. We believe this redintegration process depends on mechanisms analogous to speech perception and may involve the lexical identification of the degraded memory trace; such an identification process is difficult for nonwords, however, because they only have a newly formed, weak lexical representation.

References

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

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A., Lewis, V. J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 36*(A), 233–252.

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

Broadbent, D. E. (1958). *Perception and communication.* Oxford, England: Pergamon Press.

Brown, G. D. A., & Hulme, C. (1995). Modelling item length effects in memory span: No rehearsal needed? *Journal of Memory and Language, 34,* 594–621.

Brown, H. L., & Kirsner, K. (1980). A within-subjects analysis of the relationship between memory span and processing rate in short-term memory. *Cognitive Psychology, 12,* 177–187.

Cavanagh, J. P. (1972). Relation between the immediate memory span and the memory search rate. *Psychological Review, 79,* 525–530.

Chase, W. G. (1977). Does memory scanning involve implicit speech? In S. Dornic (Ed.), *Attention and performance VI* (pp. 607–628). Hillsdale, NJ: Erlbaum.

Clifton, C., Jr., & Tash, J. (1973). Effect of syllabic word length on memory-search rate. *Journal of Experimental Psychology, 99,* 231–235.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341–370.

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language, 31,* 668–684.

Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language, 31,* 1–17.

Cowan, N., Keller, T. A., Hulme, C., Roodenrys, S. J., McDougall, S., & Rack, J. (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory and Language, 33,* 234–250.

Cowan, N., & Saults, J. S. (1995). Memory for speech. In H. Winitz (Ed.), *Human communication and its disorders* (Vol. 4, pp. 81–170). Timonium, MD: York Press.

Cowan, N., Wood, N. L., Nugent, L. D., & Treisman, M. (1997). There are two word length effects in verbal short-term memory: Opposed effects of duration and complexity. *Psychological Science, 8,* 290–295.

Cowan, N., Wood, N. L., Wood, P. K., Keller, T. A., Nugent, L. D., & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127,* 141–160.

Cox, R., Hulme, C., & Brown, G. D. A. (1992). STM Experimenter: Using Hypercard and MacRecorder in short-term memory experiments. *Behavior Research Methods, Instruments and Computers, 24,* 575–579.

Dosher, B. A., & Ma, J.-J. (1998). Output loss or rehearsal loop? Output time versus pronunciation time limits in immediate recall for forgetting-matched materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1217–1232.

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

Hulme, C., Roodenrys, S., Brown, G. D. A., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of Psychology, 86,* 527–536.

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

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. L. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38,* 241–253.

Kail, R., & Park, Y.-S. (1994). Processing time, articulation time, and memory span. *Journal of Experimental Child Psychology, 57,* 281–291.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86,* 199–255.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review, 96,* 25–57.

Loftus, G. R. (1995). Data analysis as insight: Reply to Morrison and Weaver. *Behavior Research Methods, Instruments and Computers, 27,* 57–59.

Monsell, S. (1986). Programming of complex sequences: Evidence from the timing of rapid speech and other productions. In H. Heuer & C. Fromm (Eds.), *Experimental Brain Research,* Suppl., Series 15: *Generation and modulation of action patterns* (pp. 72–86). Berlin, Germany: Springer.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18,* 251–269.

Puckett, J. M., & Kausler, D. H. (1984). Individual differences and models of memory span: A role for memory search rate? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 72–82.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition, 21,* 167–175.

Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 419–425.

Schweickert, R., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science, 1,* 74–77.

Standing, L., Bond, B., Smith, P., & Isely, C. (1980). Is the immediate memory span determined by subvocalization rate? *British Journal of Psychology, 71,* 525–539.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652–654.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic Press.

Sternberg, S., Wright, C. E., Knoll, R. L., & Monsell, S. (1980). Motor programs in rapid speech: Additional evidence. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 507–534). Hillsdale, NJ: Erlbaum.

Stigler, J. W., Lee, S.-Y., & Stevenson, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition, 23,* 1–20.

(Appendix follows)

Appendix

Correlations of Memory Span, Speech Rate, Memory Scanning, and Response Times in the Memory Span Task

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. SPAN1SNW													
2. SPAN1SW	.31												
3. SPAN3SNW	.62**	.35											
4. SPAN5SW	.39	.49*	.46*										
5. SR1SNW	.05	−.19	.17	.32									
6. SR1SW	−.14	−.01	.20	.22	.74**								
7. SR3SNW	.21	.06	.67**	.28	.46*	.57**							
8. SR5SW	.05	.02	.25	.09	.54**	.75**	.41*						
9. RT1SNW	−.23	.19	−.43*	−.22	−.53**	−.46*	−.68**	−.42*					
10. RT1SW	−.10	−.03	−.30	−.08	−.28	−.39	−.52**	−.43*	.66**				
11. RT3SNW	−.08	.22	−.48*	−.12	−.32	−.31	−.64**	−.26	.81**	.57**			
12. RT5SW	.15	.15	−.04	.08	−.01	−.20	−.24	−.28	.51*	.79**	.56**		
13. SLOPE1SNW	−.30	−.04	−.06	.17	.09	.07	.11	.01	.02	−.14	−.03	.03	
14. SLOPE1SW	−.61**	−.69**	−.43*	−.29	.06	.03	−.10	−.14	.04	.17	−.04	−.16	.06
15. SLOPE3SNW	−.31	−.11	.01	−.14	.05	.32	.20	.12	.00	.04	−.14	.10	.51*
16. SLOPE5SW	−.16	−.14	.02	−.03	.01	.26	.26	.10	−.23	−.23	−.19	−.41*	.21
17. PI1SNW	−.47*	−.09	−.38	−.22	−.27	−.33	−.27	−.26	.40*	.26	.23	.17	.51*
18. PI1SW	−.20	−.38	−.25	−.37	−.09	−.24	−.25	−.22	.30	.22	.34	.24	.26
19. PI3SNW	−.50*	−.07	−.42*	−.17	−.08	.04	−.31	.17	.32	.06	.23	−.01	.47*
20. PI5SW	−.08	−.34	−.35	−.31	.00	−.18	−.26	−.14	.17	.43*	.29	.55**	.18
21. IWI1SNW	−.45*	−.05	−.53**	−.07	−.24	−.30	−.51*	−.35	.45*	.33	.36	.24	.13
22. IWI1SW	−.15	−.20	−.26	−.10	−.18	−.47*	−.45*	−.50*	.37	.43*	.36	.38	−.10
23. IWI3SNW	−.22	−.10	−.49*	−.24	−.34	−.16	−.25	−.16	.24	.06	.17	−.05	.42*
24. IWI5SW	−.32	−.20	−.44*	−.12	−.02	−.30	−.20	−.60**	.28	.48*	.21	.47*	.23
25. WD1SNW	−.17	.09	−.13	.39	.13	.08	−.08	−.16	.07	.12	.02	−.03	.31
26. WD1SW	−.27	.12	−.24	.08	−.32	−.25	−.25	−.38	.38	.19	.12	−.04	.25
27. WD3SNW	−.12	.06	−.20	.21	−.24	−.26	−.10	−.32	.24	.08	.17	−.06	.40
28. WD5SW	−.10	.17	−.19	.10	−.30	−.22	−.15	−.44*	.24	.04	.13	−.12	.22

Note. SPAN = memory span; SR = speech rate; RT = overall reaction time in Sternberg (1966) memory scanning task; SLOPE = slope (search rate) in Sternberg memory scanning task; PI = preparatory interval; IWI = interword interval; WD = word durations; 1SNW = 1-syllable nonword; 1SW = 1-syllable word; 3SNW = 3-syllable nonword; 5SW = 5-syllable word.
$^{*}p < .05$. $^{**}p < .01$.

Appendix (*continued*)

14	15	16	17	18	19	20	21	22	23	24	25	26	27
−.05													
.16	.37												
.14	.32	.12											
.25	.19	.21	.52**										
.23	.20	.25	.49*	.23									
.11	.34	−.06	.42*	.67**	.19								
.22	−.08	−.22	.51*	.35	.42*	.30							
.30	−.27	−.43*	.20	.49*	.04	.37	.64**						
.03	.37	.27	.46*	.20	.43*	.26	.29	−.03					
.31	.08	.00	.52**	.38	.04	.47*	.40	.36	.29				
.05	−.01	.21	.13	.02	.02	−.13	.04	−.09	−.01	.28			
.00	.03	.22	.41*	.13	.07	−.17	.24	−.02	.34	.44*	.72**		
−.01	.01	.37	.46*	.26	.21	−.03	.10	.00	.45*	.35	.63**	.75**	
−.11	.07	.28	.24	.08	.02	−.12	.12	.05	.34	.29	.72**	.76**	.72**

Received July 8, 1997
Revision received August 26, 1998
Accepted August 21, 1998 ■