

Articulatory loop explanations of memory span and pronunciation rate correspondences: A cautionary note

GERALD TEHAN and MICHAEL S. HUMPHREYS
University of Queensland, St. Lucia, Queensland, Australia

In recent years, a number of memory span findings have been attributed to the operation of an articulatory loop (Baddeley & Hitch, 1974). These attributions have been made on the basis of finding a correspondence between span differences and pronunciation rate differences. This experiment explored articulatory loop explanations for two material effects in memory span: the word-frequency effect (span for high-frequency words is larger than span for low-frequency words) and the word-class effect (span for function words is smaller than span for either nouns or adjectives). The results indicate that it is possible to obtain span differences without finding corresponding pronunciation rate differences. Moreover, span differences were as pronounced under articulatory suppression conditions as they were under rehearsal conditions. Both of these results limit the generality of articulatory loop explanations of memory span.

Explanations for performance on the memory span task have varied greatly over the years; however, recent evidence indicates that one of the best predictors of memory span is the time taken to pronounce the to-be-remembered items. Data from a number of studies suggest that span is equivalent to the number of items that can be pronounced in about 2 sec (Baddeley, Thomson, & Buchanan, 1975; Schweickert & Boruff, 1986; Standing, Bond, Smith, & Isley, 1980).

The relationship between span and pronunciation rate was first noted by Baddeley et al. (1975). In a series of experiments, Baddeley et al. found that short words were better remembered than long words in the memory span task. Word length was measured in a couple of different ways: (1) the syllables in the word were counted, and (2) in the case where words had been matched for frequency, number of syllables, and number of phonemes, the length of pronunciation was estimated by measuring the waveform produced by the pronunciation of each to-be-remembered word. Baddeley et al. argued that this *word-length effect* was consistent with the operation of an articulatory loop. The articulatory loop had previously been proposed as the mechanism in working memory that underlies performance on the span task (Baddeley & Hitch, 1974). It was seen as a phonological store that was driven by an articulatory control process (Baddeley, 1986). In the span task, for visual presentation, items are registered on the loop by way of subvocalization. The items in the loop decay rapidly, but can be refreshed by rehearsal. Thus span recall is a function of rehearsal and decay rates, such that the more that can be rehearsed in a given period of time, the more will be remembered. The

word-length effect is due, then, to the fact that more short words than long words can be rehearsed in any given unit of time. Additional confirming evidence for the operation of a rehearsal-driven articulatory loop is obtained when rehearsal is prevented by means of articulatory suppression. Under suppression conditions, at least for visual presentation, the word-length effect disappears. Furthermore, the concept of a rehearsal-driven articulatory loop accounted for individual differences in span, in that measures of rehearsal speed correlated quite highly with span scores. In the present experiment two measures of rehearsal speed were used: (1) an articulation measure, in which 3 words were rehearsed 10 times as quickly as possible, and (2) a reading measure, in which 50 words were read as quickly as possible (Baddeley et al., 1975).

In addition to the word-length effect, there is evidence that differences in span for a variety of materials covaries with rehearsal rates (pronunciation rates) for those materials. Watkins (1977) demonstrated a difference in span for high- and low-frequency words, whereas Wright (1979) showed a corresponding difference in the time taken to read 25-item lists of high- and low-frequency words. Schweickert and Boruff (1986) and Standing et al. (1980) showed that memory span for a number of different materials (e.g., digits vs. letters vs. words) covaries with pronunciation rates for these materials. Differences in pronunciation rates also have accounted for cross-cultural differences in memory span. Ellis and Hennelly (1980) explained the difference in digit span between bilingual Welsh and English speakers in terms of differential pronunciation rates for the two types of material. Stigler, Lee, and Stevenson (1986) found the same result in comparing bilingual Chinese and English speakers. Between-subject differences in memory span for digits in English, Spanish, Hebrew, and Arabic have been shown to be correlated with the average number of syllables in

Reprint requests should be sent to Gerald Tehan, Department of Psychology, University of Queensland, St. Lucia, Queensland, Australia 4067.

the number words of the respective languages (Naveh-Benjamin & Ayers, 1986).

These correspondences between memory span and pronunciation rate have generally been interpreted as evidence for the involvement of the articulatory loop (for an exception, see Schweikert & Boruff, 1986). However, one piece of evidence was briefly reported that may not be consistent with articulatory loop explanations of memory span. Humphreys, Lynch, Revelle, and Hall (1983) reported a study in which it was found that children's span for function words (articles, prepositions, and conjunctions) was poor relative to span for other grammatical classes. This is odd, since these function words tend to be used very frequently in the language and thus should be particularly well recalled (Watkins, 1977). In addition, there seems to be little reason to suppose that function words should be rehearsed or read more slowly than the other word classes.

The present experiment had two aims. First, we sought to confirm that the word-class effect could be found with an adult population. Second, we sought to establish how consistent both the word-class effect and the word-frequency effect were with articulatory loop explanations of memory span.

The subjects in the experiment studied lists of high-frequency and low-frequency nouns, adjectives, and function words for immediate serial recall. These lists were studied under conditions in which the subject was free to rehearse and under conditions in which rehearsal was suppressed. Span estimates and rehearsal time measures were taken for the different materials.

In this experiment, the six different types of words were matched for spoken duration, using procedures similar to those used by Baddeley et al. (1975). If memory span is purely a matter of word length, matching the various materials for spoken duration should result in equivalent span estimates for the different materials, since the same number of items in each case should fit on the articulatory loop. However, Baddeley and Lewis (1984) suggested that the best determinant of span is the functional rate of rehearsal. Thus, even if words are matched for spoken duration, span differences could still be present if one set of material is rehearsed faster than another. Geffen and Luszcz (1983) showed that this is the case

for word frequency. Even when high-frequency and low-frequency words have the same spoken duration, high-frequency words can be read more quickly than low-frequency words.

One final prediction can be derived from the concept of the articulatory loop. If corresponding rehearsal and span differences emerge under conditions in which the subject is free to rehearse, the articulatory loop hypothesis would predict that these differences in span should disappear once rehearsal has been suppressed, provided that the stimuli were presented visually (Baddeley et al., 1975).

Failure to find rehearsal differences and corresponding span differences or the failure of span differences to disappear under suppression would severely limit the generality of the concept of the articulatory loop. It would also raise doubt as to the source of the other span differences mentioned.

METHOD

Subjects

Twenty-four introductory psychology students participated for course credit.

Materials

In both the span tasks and the rehearsal task, the subjects studied lists constructed from six 10-item pools. There was a 10-item pool for each of the six classes used: high-frequency nouns, high-frequency adjectives, high-frequency function words, low-frequency nouns, low-frequency adjectives, and low-frequency function words. The words in each pool were all two-syllable words and were matched for spoken length. The mean spoken lengths of the words in each pool are presented in Column 2 of Table 1. The spoken length was determined by use of a Fairlight computer. Each word was spoken by the first author into a microphone, the sound was digitized, and a waveform of the sound was displayed on the terminal, from which a measure of spoken duration could be derived. The numbers in Table 1 are the averages of 10 measurements for each word spoken in an Australian accent.

Column 3 of Table 1 presents the word-frequency data for the items used. High-frequency items were selected from the Kučera and Francis (1967) norms as having frequency counts of greater than 100. The criterion for low-frequency items was a frequency count of less than 15.

Procedure

The subjects viewed 20 trials in each of the six types of material used; each trial was formed by randomly permuting the 10-item pool of that material. All subjects saw 20 identical trial lists for each type of material. Each subject studied each material type twice, once under rehearsal con-

Table 1
Spoken Duration (in msec), Word Frequency, Rehearsal Rate (in sec),
and Span Estimates for the Six Materials Used

Word Class	Spoken Duration	Word Frequency	Rehearsal Rate	Span Articulation	Span Suppression
High-Frequency Words					
Adjective	750	210.5	19.62	4.78	3.89
Noun	745	205.7	19.67	4.85	3.91
Function	735	545.4	19.41	4.61	3.73
Low-Frequency Words					
Adjective	740	7.8	19.40	4.46	3.71
Noun	745	6.3	20.94	4.54	3.65
Function	745	7.1	20.23	4.52	3.36

ditions and once under suppression conditions, in which the subjects repeated the word "the" throughout the presentation of the items. These conditions were counterbalanced across subjects. The order of presentation of the six types of material was also counterbalanced. Thus each subject studied the six material types under rehearsal conditions and then the same six material types under suppression conditions, or vice versa. All items were presented on a computer-controlled video display unit.

The up-and-down measure of span used by Watkins (1977) was adopted here. In the up-and-down procedure, if a trial was correctly recalled, the length of the next list was increased by one item; if an error was made, the next list was decreased by one item. Span was the average of the list lengths for the last 16 trials, the first 4 being discarded because of the arbitrary determination of the length of the first list. The first list for each type of material was always three items.

Between the two span sessions, an estimate of rehearsal speed for each type of material was obtained following the procedures used by Baddeley et al. (1975) and Wright (1979). The subjects read a list of 50 words, as quickly and as clearly as they could. Each list was constructed by randomly permuting the items in the pool five times. The items were presented on a page in two columns. The time taken to read the 50 words was measured by stopwatch. Two measures were taken for each type of material. The correlation between the two readings of each material was .89.

RESULTS

Rehearsal Speed

Column 4 of Table 1 provides the rehearsal speed data. This data was subjected to a 2×3 repeated measures analysis of variance (ANOVA) in which word frequency and word class were the two factors. The effects of frequency and class were reliable [$F(1,23) = 20.71$, $MSe = .675$, and $F(2,46) = 8.96$, $MSe = .853$, respectively]. (An alpha level of .05 was used in all analyses.) The interaction between frequency and class was also significant [$F(2,46) = 10.01$, $MSe = .692$]. A Newman-Keuls post hoc analysis confirmed that there were no differences in rehearsal speed between the different word classes when high-frequency members were used, but rehearsal differences emerged when low-frequency members of each class were used.

Memory Span

The memory span data are presented in Columns 5 and 6 of Table 1. These data were analyzed by means of $2 \times 2 \times 3$ repeated measures ANOVA, in which rehearsal/suppression, word frequency, and word class were the factors. Reliable effects were found for all three factors: suppression hurt performance [$F(1,23) = 76.04$, $MSe = .796$], span was higher for high-frequency words [$F(1,23) = 32.65$, $MSe = .146$], and there were differences among the word classes [$F(2,46) = 4.04$, $MSe = .212$]. A Newman-Keuls analysis suggested that span for the function words was reliably lower than span for the other two types of word class, which did not differ from each other. This finding replicates Humphreys et al.'s (1983) result.

The critical interactions between rehearsal condition and word frequency and between rehearsal condition and word class are presented in Figure 1. Neither interaction was significant; nor was the three-way interaction between rehearsal condition, frequency, and word class.

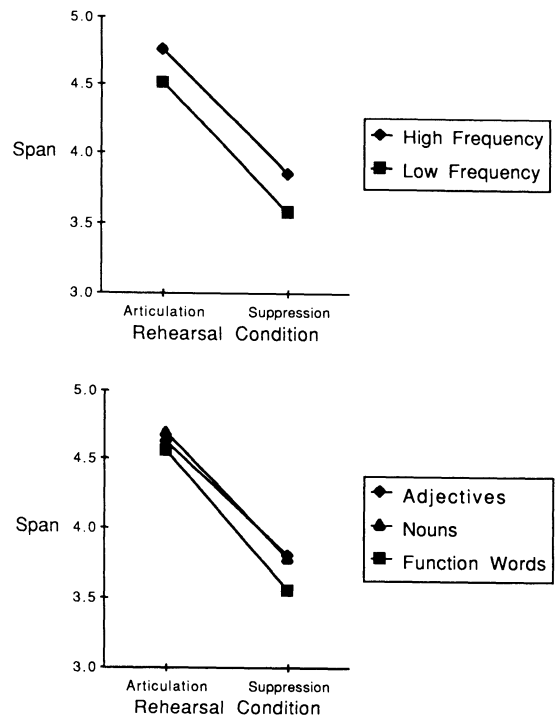


Figure 1. Memory span for high- and low-frequency words, and for nouns, adjectives, and function words, under rehearsal and articulatory suppression conditions.

DISCUSSION

Two predictions were made on the basis of the concept of an articulatory loop. The first was that there should be a correspondence between rehearsal rate and span. For word frequency, this correspondence was found. High-frequency words were read (rehearsed) more quickly than low-frequency words, thus replicating Wright's (1979) and Geffen and Luszcz's (1983) results, and span for high-frequency words was larger than span for low-frequency words, replicating Watkins's (1977) results. In contrast to this, no correspondence was found between span and rehearsal rate for the three types of word class; that is, span differences were obtained when both high-frequency and low-frequency instances were used, but reading rate differences among word classes emerged only when low-frequency instances were used. Moreover, nouns that produced the highest memory span estimates took the longest time to read. Thus, on the basis of correspondence between span and rehearsal, the pattern of results for word frequency is consistent with the concept of an articulatory loop. The same cannot be said, however, of the word-class effect.

The second critical test for the articulatory loop hypothesis concerns span effects under articulatory suppression. The second prediction made was that once rehearsal was prevented, the articulatory loop would become inoperative, with the consequence that all span differences should disappear (Baddeley et al., 1975). In this experiment, word-frequency and word-class effects are as strong, if not stronger, under conditions of articulatory suppression as under conditions in which the subject is free to rehearse. Span differences obtained under suppression indicate that the word-frequency and word-class effects cannot simply be attributed to the operation of an articulatory loop in the same way that the word-length effect can be.

In summary, these results have two clear implications for the concept of the articulatory loop. First, finding span differences without the corresponding differences in pronunciation rate presents difficulties for an articulatory loop explanation of span or, for that matter, any theory

that views span as a simple race between pronunciation rate and trace decay. The data suggest that processes or mechanisms in addition to the articulatory loop are having an effect on span. In terms of the model of working memory proposed by Baddeley and Hitch (1974), it would seem that their notion of a central executive plays just as important a role in span as does the articulatory loop.

The second implication is more methodological in nature. Previously, the correspondence between span and pronunciation rate has been seen as sufficient evidence of the operation of the articulatory loop. However, the articulatory loop predicts both the correspondence between span and articulation rates and the elimination of span differences under suppression, for visually presented items. The word-frequency data suggest that obtaining one will not guarantee the emergence of the other. Finding a correspondence between span and pronunciation rate can no longer be seen as sufficient evidence for the articulatory loop; instead, future research will need to show consistent effects under both tasks associated with the articulatory loop (i.e., articulatory suppression and pronunciation rate/span correspondence). In light of this, one wonders how many of the other span effects mentioned in the introduction would remain under suppression conditions.

REFERENCES

- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- BADDELEY, A. D., & LEWIS, V. (1984). When does rapid presentation enhance digit span? *Bulletin of the Psychonomic Society*, **22**, 403-405.
- 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.
- ELLIS, N. C., & HENNELLY, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, **71**, 43-51.
- GEFFEN, G., & LUSZCZ, M. A. (1983). Are the spoken durations of rare words longer than those of common words? *Memory & Cognition*, **11**, 13-15.
- HUMPHREYS, M. S., LYNCH, M. J., REVELLE, W., & HALL, J. W. (1983). Individual differences in short-term memory. In R. F. Dillon & R. R. Schmeck (Eds.), *Individual differences in cognition* (pp. 35-64). New York: Academic Press.
- KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- NAVEH-BENJAMIN, M., & AYERS, T. J. (1986). Digit span, reading rate, and linguistic relativity. *Quarterly Journal of Experimental Psychology*, **38A**, 739-751.
- SCHWEICKERT, R., & BORUFF, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 419-425.
- STANDING, L., BOND, B., SMITH, P., & ISLEY, C. (1980). Is the immediate memory span determined by subvocalization rate? *British Journal of Psychology*, **71**, 525-539.
- STIGLER, J. W., LEE, S., & STEVENSON, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition*, **23**, 1-21.
- WATKINS, M. J. (1977). The intricacy of memory span. *Memory & Cognition*, **5**, 529-534.
- WRIGHT, C. E. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory & Cognition*, **7**, 411-419.

(Manuscript received for publication October 19, 1987.)