JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR **14**, 575–589 (1975)

Word Length and the Structure of Short-Term Memory

ALAN D. BADDELEY

Medical Research Council, Applied Psychology Unit

AND

NEIL THOMSON AND MARY BUCHANAN

University of Stirling, Scotland

A number of experiments explored the hypothesis that immediate memory span is not constant, but varies with the length of the words to be recalled. Results showed: (1) Memory span is inversely related to word length across a wide range of materials; (2) When number of syllables and number of phonemes are held constant, words of short temporal duration are better recalled than words of long duration; (3) Span could be predicted on the basis of the number of words which the subject can read in approximately 2 sec; (4) When articulation is suppressed by requiring the subject to articulate an irrelevant sound, the word length effect disappears with visual presentation, but remains when presentation is auditory. The results are interpreted in terms of a phonemically-based store of limited temporal capacity, which may function as an output buffer for speech production, and as a supplement to a more central working memory system.

Miller (1956) has suggested that the capacity of short-term memory is constant when measured in terms of number of chunks, a chunk being a subjectively meaningful unit. Because of the subjective definition of a chunk, this hypothesis is essentially irrefutable unless an independent measure of the nature of a chunk is available. Typically this problem has been avoided by making the simplifying assumption that such experimenter-defined units as words, digits, and letters constitute chunks to the subject. Hence, although Miller's hypothesis is not refutable in the absence of an independent measure of a chunk, it is meaningful to test a weaker version, namely that the capacity of short-term memory is a constant number of items, where items are defined experimental units. Words represent one commonly accepted type of item, and in this case, the chunking hypothesis would predict that the capacity of short-term memory, as measured in words, should be constant regardless of the size or duration of the words used.

A number of studies testing this hypothesis have used the recency effect in free recall as an estimate of short-term memory capacity. Craik (1968) found no reliable effect of word length on performance in the free recall of separate groups of words comprising one to five syllables. This invariance held true whether performance was measured in terms of either raw scores, or estimates of primary memory and secondary memory components. This result was replicated and extended by Glanzer and Razel (1974) who observed a recency effect which was constant when measured in number of items, even when an item comprised a whole proverb rather than a single word. They concluded from their study that short-term or primary memory

We are grateful to the Social Science Research Council and the Medical Research Council for financial support. We thank Ronald Bidgood for running Experiment VII, Rob Baker for phonetic advice, and Graham Hitch for many useful discussions.

Dr. Baddeley's address is: Medical Research Council, Applied Psychology Unit, Cambridge, CB2 2EF, England.

has a capacity of two items regardless of item duration or complexity.

Miller's generalization, however, was based on the memory span paradigm, and it is questionable whether recency and span depend on the same memory mechanisms. There is indeed a growing body of evidence suggesting that the recency effect in free recall is basically unrelated to short-term memory as measured by memory span. Such evidence includes:

(1) Craik's (1970) observation that a subject's memory span correlates more highly with the secondary memory than the primary memory component of free recall.

(2) Memory span shows clear evidence of speech coding, being impaired by both phonemic similarity (Conrad, 1964; Baddeley, 1966) and articulatory suppression (Levy, 1971). This is not the case for the recency effect in free recall which is unaffected by either phonemic similarity (Craik & Levy, 1970; Glanzer, Koppenaal, & Nelson, 1972) or articulatory suppression (Richardson & Baddeley, 1975).

(3) Baddeley and Hitch (1974) have shown unimpaired recency in free recall for subjects performing a concurrent memory span task involving the retention of a sequence of six digits. Since the memory span task did not interfere with recency, it is difficult to maintain the view that the two tasks are based on the same limited-capacity system.

Studies investigating the effect of word length on memory span do not in general support the weak version of Miller's hypothesis. Thus, unpublished work by Laughery, Lachman, and Dansereau (Note 1) and by Standing, Bond, and Smith (Note 2) have reported poorer performance in a memory span task when longer words are used. Mackworth (1963) found a high correlation between reading rate and memory span for a wide range of materials, including pictures, letters, digits, shapes, and colors. This result could be interpreted in terms of word length as a determinant of memory span, with reading rate providing an indirect measure of word length. The situation is, however, complicated by the fact that subjects in some cases were asked to label pictures, and in others to read words so that it is not clear whether the result is due to articulation time or to difficulty in retrieving the correct verbal label. Watkins and Watkins (1973) present the clearest published evidence for an effect of word length on memory span, in a study primarily concerned with the modality effect. They found evidence for a word length effect on earlier serial positions, but observed that the modality effect (the enhanced recall of auditorily presented items) did not interact with word length. They suggest that the word length effect observed may have been due to the greater difficulty of perceiving their four-syllable words which were presented at a 1/sec rate.

These studies do not support the hypothesis that memory span capacity is a constant number of items. However, it is always possible to save the item-based hypothesis by questioning the assumption that words constitute items. Given evidence that short-term memory is a speech-based system, it could be reasonably argued that its capacity should be measured in more basic speech units such as syllables or phonemes. The experiments that follow aim first to study the influence of word-length on memory span, secondly to explore the relative importance of number of syllables and temporal duration of a word as determinants of span, and thirdly to explore the implications of this for the question of whether the underlying memory system is time-based or item-based.

Experiment I

This study compared the memory span of subjects for sets of long and short words of comparable frequency of occurrence in English. One set comprised eight monosyllables, namely, *sum*, *hate*, *harm*, *wit*, *bond*, *yield*, *worst*, and *twice*. The other set comprised eight five-syllable words, namely *association, opportunity, representative, organization, considerable, immediately, university,* and *individual.*

Method

Five list lengths were used, comprising sequences of four, five, six, seven, and eight words. Eight sequences of each length were made up from the pool of short words, and eight from the pool of long words. In both cases, sequences were generated by sampling at random without replacement from the appropriate pool of words. All subjects were tested on both long and short words, and all received the sequences in ascending order of list length, beginning with sequences of four words and proceeding up to the point at which they failed on all eight sequences, whereupon testing on the pool of words in question was discontinued. Half the subjects began with the pool of long words, and half with the short words.

The words were read to the subject at a 1.5-sec rate, with each list being preceded by the spoken warning "Ready." Subjects were allowed 15 sec to recall the words verbally in the order presented. Subjects were allowed to familiarize themselves with the two pools of words at the beginning of the experiment, and these two pools remained visible to the subjects on prompt cards throughout the experiment. Several different prompt cards with the words in differing orders were used in this and subsequent experiments so as to prevent the subjects from using location on the card as a cue. The subjects were eight undergraduate or postgraduate students from the University of Stirling.

Results and Discussion

Performance was scored in terms of number of sequences recalled completely correctly (i.e., all the items correct and in the correct order). Figure 1 shows the level of performance at each sequence length for the long and the short words. There is a very clear advantage to the short word set which occurs at all sequence lengths and is characteristic of all eight subjects tested.

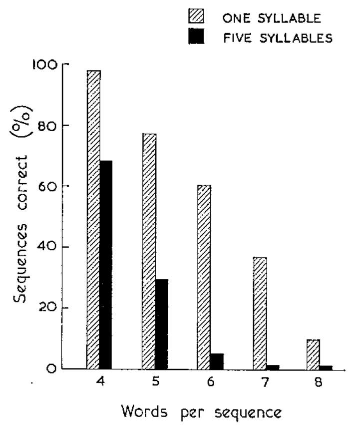

FIG. 1. Effect of word length on memory span. Mean percentage recall of long and short words as a function of sequence length.

There is little doubt that the sample of short words used results in better memory span performance than the sample of long words. However, it is arguable that polysyllabic words tend to be linguistically different from monosyllables. In particular, our polysyllables tended to be of Latin origin, compared to the monosyllables which seemed to comprise simpler words of Anglo–Saxon origin. Experiment II attempted to avoid this problem by using words from a single category, country names, a sample of material unlikely to come from any single language source.

EXPERIMENT II

Method

Sequences of five words were constructed by sampling without replacement from each of two pools. The pool of short words comprised the country names *Chad, Chile, Greece, Tonga, Kenya, Burma, Cuba, Malta,* while the long names were *Somaliland*, *Afghanistan*, *Venezuela*, *Czechoslovakia*, *Yugoslavia*, *Ethiopia*, *Nicaragua*, and *Australia*. The names were selected on the basis of their probable familiarity to the subjects, and because they had a similar frequency of repetition of initial and final letters within the pool. Subjects were tested on a total of eight sequences of five short names and eight sequences of five long names. Eight undergraduate subjects were tested using the same presentation procedure as Experiment I.

Results and Discussion

Table 1 shows the mean number of sequences recalled completely correctly, and the mean number of items recalled in the appropriate serial position, for long and short names. On both these scores all eight subjects showed a clear word length effect. Since the material in this study was very different at a linguistic level from the material used in the previous study, and since the effect is very large in both cases, it is clear that the word length effect is a robust phenomenon of some generality.

TABLE 1

MEAN NUMBER OF SEQUENCES AND ITEMS CORRECTLY RECALLED AS A FUNCTION OF WORD LENGTH IN EXPERIMENT II

	Short names		Long names	
	Mean	*SD*	Mean	*SD*
Sequences correct Max = 8	4.50	2.00	.88	1.27
Items correct Max = 5	4.17	.71	2.80	.24

However, in these and all previous experiments investigating the effect of word length, two major variables are confounded, namely a word's spoken duration and the number of syllables it contains. The results could therefore indicate either that memory span is limited in the number of items it can hold, with the item being the syllable, or that the temporal duration of the words determines the size of memory span. The latter possibility might be predicted by decay theory (Broadbent, 1958) which assumes that forgetting occurs as a function of time. Many studies have attempted to test the theory by measuring performance as a function of presentation rate, and while some studies report enhanced performance with rapid presentation as predicted by decay theory (Conrad & Hille, 1958), others have found the opposite (Sperling & Speelman, 1970). However, in none of these studies was the subject prevented from rehearsing, and this makes interpretation of the results difficult as the subject is effectively re-presenting the list to himself at a rate of his own choosing. This problem can be avoided by allowing the subject to rehearse while using lists of long- and short-duration words. As less long words than short words can be rehearsed in a given period of time, a word duration effect will be predicted by decay theory (Sperling, 1963). On the other hand, a simple displacement or interference model would predict an effect of number of items, but not duration. Thus, the hypothesis that short-term memory capacity is a constant number of items, where the syllable is the item, predicts no word length effect for words matched for syllable number, but differing in spoken duration. Decay theory, on the other hand, predicts that the amount recalled will be a function of word duration. The next experiment tests these predictions.

EXPERIMENT III

Method

Two pools of disyllabic words, matched for frequency, were produced such that one set tended to have a longer duration when spoken normally. The long word set comprised: *Friday*, *coerce*, *humane*, *harpoon*, *nitrate*, *cyclone*, *morphine*, *tycoon*, *voodoo*, and *zygote*, and the short words were *bishop*, *pectin*, *ember*, *wicket*, *wiggle*, *pewter*, *tipple*, *hackle*, *decor*, and *phallic*. The words were recorded by a female experimenter onto magnetic tape, which was then played through an oscillograph. This plots the wave-form of the signal against time, allowing the duration of the utterance to be measured. The mean duration of the long words was 0.77 sec, and of the short words, 0.46 sec.

From each pool of words, 10 lists of five words were constructed by sampling at random without replacement. The twenty lists were divided into four blocks of five, two comprising lists of short duration words and two of long duration words. A Latin square design was then used to present the blocks in counterbalanced order to each of the 12 subjects. Words were read at a 2-sec rate, and subjects were required to recall verbally at the same rate, paced by a metronome. Recall was paced so as to ensure that the mean delay between input and recall was comparable for long and short words (Conrad & Hille, 1958). Subjects were familiarized with the set of words and with the procedure, and were instructed to commence recall as soon as the last item in each list had been presented. Twelve undergraduates from the University of Stirling served as subjects.

Results and Discussion

Figure 2 shows the mean number of words correctly recalled as a function of serial position. A three-way analysis of variance involving subjects, word length, and serial position showed significant effects of word length, $F(1, 11) = 11.33$, $p < .01$, serial position, $F(4, 44) = 36.82$, $p < .001$, and a significant interaction between word length and serial position, $F(4, 44) = 3.28$, $p < .05$. Analysis by t test showed that the word length effect was significant for serial positions 1, 2, and 3, but not for positions 4 and 5.

These results are very similar to those of Watkins and Watkins (1973) showing a word length effect only for the earlier serial positions; this could reflect the masking of an underlying word length effect by the modality effect. However, the experiment

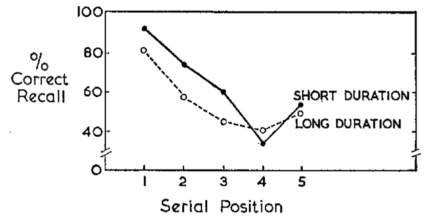

FIG. 2. Mean recall of disyllabic words of long and short temporal duration.

differs from the Watkins and Watkins study in using words which are matched for number of syllables, but differ in spoken duration. As such, the results are consistent with decay theory, and are inconsistent with the hypothesis that short-term memory holds a constant number of syllables.

The last version of Miller's weakened hypothesis to be investigated is that short-term memory holds a constant number of phonemes. In the last experiment, there was a clear tendency for the long words to have more constituent phonemes, thus the result is open to the interpretation that the word length effect represents a limit to the number of phonemes that can be held. Experiment IV compares performance on sets of words which are matched for number of constituent phonemes, but which differ in duration. Decay theory again predicts a difference in performance in favor of the short duration words.

EXPERIMENT IV

Two sets of words were generated with the following constraints: They differed in spoken duration; they were equal in number of syllables; they were matched for word frequency; and they were equal in number of phonemes (with Scottish pronunciation). Given all these constraints, the previous sets of words reduced from 10 to five; details are given in Table 2. Sequences of five words were produced, and the experiment performed using a procedure identical to that used in Experiment III, except that the presentation and paced recall rate was increased to 1 sec per word. Eight Scottish undergraduates served as subjects.

TABLE 2

Details of Words Used in Experiment IV

	Words	Frequency	Number of phonemes	Duration (sec)
Long	Coerce	1	5	.80
	Harpoon	1	6	.75
	Friday	40	5	.70
	Cyclone	3	6	.88
	Zygote	—	5	.90
Short	Wicket	1	5	.50
	Pectin	1	6	.60
	Bishop	40	5	.28
	Pewter	3	6	.40
	Phallic	—	5	.42

Results and Discussion

Subjects recalled a mean of 61.6% of the long words and 72.2% of the short. A three-way analysis of variance showed a significant effect of word length, $F(1, 7) = 18.9$, $p < .01$, and of serial position, $F(4, 28) = 38.06$, $p < .001$, but no interaction between serial position and word length, $F(4, 28) = 0.55$, $p > .05$.

It is clear then that word duration may influence span when the number of both syllables and phonemes is held constant. The absence of an interaction between word length and serial position is puzzling in view of the previous result, it may be due to either the change in material, or more likely, the change in presentation and recall rate. However, despite this minor discrepancy between experiments III and IV, both seem to concur in suggesting that the temporal duration of items is a powerful determinant of memory span. Before finally dismissing the hypothesis that short-term memory capacity is a constant number of items, a procedural point that could have distorted the results should be mentioned. In both experiments the same experimenter read out the words and it is possible that some incidental feature of her mode of delivery produced the observed effect. To avoid this possibility, the experiment was repeated using visual presentation at a 2-sec rate and to our dismay, a statistically reliable word length effect was not observed.

However, a closer examination of the data revealed that most subjects did show the predicted effect, but that two out of eight did substantially better on the long words. On testing a further set of subjects and asking them how they remembered the material, it was found that those who did best on the short words reported using a rehearsal strategy, whilst those who did better on the long words reported using an imagery strategy. Use of this latter strategy was facilitated by the fact that the presentation rate had been reduced to 2 sec/word in order to obviate perceptual difficulties. As the subject of investigation is the articulatory short-term memory system, it is reasonable to instruct subjects to use a rehearsal strategy in order to avoid this difficulty. The next experiment, then, is a replication of the previous one, but using visual presentation with an instruction to the subjects to rehearse.

Experiment V

The same material and design were used as in the previous experiment, except that the material was presented visually on flash cards at a 2-sec rate, and recall was unpaced. The duration of the words was also measured in a different way. The duration of a word is determined by two sets of variables, the acoustic nature of the word and the subjects' articulatory rate. The latter variable has been shown to be very stable over a wide range of conditions within a subject, but to vary considerably between subjects (Goldman–Eisler, 1961), and, as decay theory assumes rehearsal rate determines performance, the subject's rather than the experimenter's pronounciation of the words was used.

Two different estimates of rehearsal rate were made. In the first of these, subjects were timed for reading the 10, five-word lists in each condition, as quickly as they could out loud, the 50 words being typed out in two columns. This was done four times for each word length after the memory task, times being recorded by stopwatch. The times so obtained were transformed into reading rate (RR) scores in units of words per second. The second estimate of rehearsal rate involved requiring the subject to repeat continuously three of the words from one of the pools out loud. Subjects did this as quickly as they could, and were timed by stopwatch for 10 repetitions of the three words. For each condition, they did this four times, always with a different set of three words, and always after the memory task. These times were transformed into articulatory rate scores (AR) in units of words per second. Half the subjects did the reading rate test first, and half the articulatory rate test first. The subjects, who were instructed to remember the lists by repeating the words to themselves, were eight members of the Applied Psychology Unit subject panel who were paid for their services.

Results and Discussion

Subjects recalled a mean of 53.4% of the long words correctly and in the right order, and 71.7% of the short words. Analysis of variance showed that there was a significant effect of word length, $F(1, 7) = 15.14, p < .01$, indicating that the word duration effect is not dependent on auditory presentation. There was again a significant effect of serial position, $F(4, 28) = 14.79$, $p < .001$, but the interaction between word length and serial position failed to reach significance, $F(4, 28) = 2.43$, $.05 < p < .1$. These results are again inconsistent with the hypothesis that short-term memory capacity is a constant number of items. An alternative view, that short-term memory is a time-based system, will next be explored and the adequacy of decay theory in this context empirically investigated.

Let us assume that the memory system underlying the word length effect exhibits trace decay, but that rehearsal may revive a decaying trace. It then follows that the amount recalled will be a function of rehearsal rate. Thus, if it can be assumed that reading rate (RR) and articulation rate (AR) are good estimates of rehearsal rate, then it should be possible to use them as predictors of memory span. Table 3 shows the ratio of memory span to reading rate and to articulation rate across conditions. A Wilcoxon matched pairs test showed that there was no effect of conditions for either the memory span–reading rate ratio, $T = 9, N = 8, p > .05$, or for the memory span–articulation rate ratio, $T = 10$, $N = 6$, $p > .05$. In short, Table 3 indicates that a subject can recall as many words as he can read in 1.6 sec, or can articulate in 1.3 sec. The next experiment explored this relationship in more detail using five different word lengths rather than two.

TABLE 3

RATIO OF MEMORY SCORE TO READING RATE AND TO ARTICULATION RATE FOR SUBJECTS IN EXPERIMENT V.

Subject	Memory score / reading rate K_L[a]	Memory score / reading rate K_S[b]	Memory score / articulation rate K_L[a]	Memory score / articulation rate K_S[b]
1	1.78	1.70	1.48	1.43
2	1.72	1.42	.93	.93
3	1.55	1.80	1.15	1.14
4	1.43	1.83	1.34	1.48
5	1.30	1.46	1.14	1.32
6	1.68	1.95	1.40	1.79
7	1.38	1.59	1.24	1.00
8	2.15	1.63	1.95	1.36
Mean	1.62	1.67	1.33	1.31

[a] K_L = Constant for long words.
[b] K_S = Constant for short words.

EXPERIMENT VI

Method

Five pools of 10 words were constructed.

TABLE 4

POOLS OF WORDS, MATCHED FOR CONCEPTUAL CLASS; USED IN EXPERIMENT VI

Number of syllables				
1	2	3	4	5
Stoat	Puma	Gorilla	Rhinoceros	Hippopotamus
Mumps	Measles	Leprosy	Diphtheria	Tuberculosis
School	College	Nursery	Academy	University
Greece	Peru	Mexico	Australia	Yugoslavia
Crewe	Blackpool	Exeter	Wolverhampton	Weston-Super-Mare
Switch	Kettle	Radio	Television	Refrigerator
Maths	Physics	Botany	Biology	Physiology
Maine	Utah	Wyoming	Alabama	Louisiana
Scroll	Essay	Bulletin	Dictionary	Periodical
Zinc	Carbon	Calcium	Uranium	Aluminium

Each pool comprised one word from each of 10 semantic categories, the items being matched as closely as possible for familiarity to the subjects. The sets differed in comprising words of either one, two, three, four, or five syllables, as may be seen in Table 4.

From each pool, 10 lists of five words were produced by sampling at random without replacement. The 50 lists were then presented visually on video tape in completely random order; hence subjects were unaware on any given trial what set would be used and so were unlikely to use a different strategy for words of different length. Half the subjects received Lists 1–25 first, and half Lists 26–50 first. Words were written on cards and presented at a 2-sec rate by a card changer which was viewed by a video camera and recorded. A card containing a row of asterisks served as a warning that the list was about to appear. Twelve seconds were allowed for spoken recall.

Reading rate was measured in this experiment by requiring the subjects to read lists of 50 words comprising five occurrences of each item in a given set. The words were typed in uppercase in random order in two columns on a sheet of paper. Subjects were instructed to read the lists aloud as quickly as they could, consistent with pronouncing each word correctly. Their reading times were measured by stopwatch. Subjects read each list a total of four times, twice before beginning the memory task and twice after completing it. Half the subjects began both tests by reading the one-syllable list and proceeding up to the five-syllable list, while the remainder of the subjects were tested in the reverse order. The subjects, who were tested individually, comprised 14 members of the Applied Psychology Unit's panel who were paid for their services.

Results

Figure 3 shows the effect of word length on mean percentage of words correctly recalled in the appropriate position, and mean reading rate.

Memory scores. Analysis of variance showed a significant effect of conditions, $F(4, 52) = 36.70$, $p < .001$, and of subjects, $F(13,52) = 11.84$, $p < .001$. A Newman–Keuls test between conditions showed that words of one or two syllables were better recalled than words of three or four, which in turn were better than five-syllable words ($p < .05$ in each case).

Reading rate. Analysis of variance showed a significant effect of conditions, $F(4, 52) =$

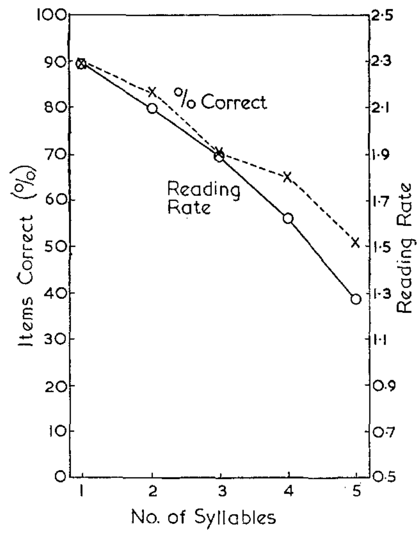

FIG. 3. Mean reading rate and percentage correct recall of sequences of five words as a function of word length.

244.02, $p < .001$. A Newman–Keuls test between conditions showed that each condition was significantly different from every other one ($p < .01$ in each case).

The next set of analyses tested the prediction made by decay theory, that the ratio memory span to reading rate is constant across conditions. Figure 4 shows memory span plotted as a function of reading rate, the line being fitted by the method of least squares. The slope of the line is 1.87, and the intercept on the ordinate 0.17. The standard error of the estimate is 0.10. The value of the intercept differs significantly from zero, $t(3) = 3.71$, $p < .05$. Thus the results are well described by the function $S = c + kR$, where S is the memory span, R is reading rate, and k and c are constants.

One final question of interest is whether such a relationship holds across subjects as well as across word samples, or in other words as to whether fast readers are also good memorizers. This proved to be the case; there was a substantial correlation between memory span and reading rate, $r(13) = .685$, $p < .005$.

Discussion

The results show that the manipulations were effective in producing sets of words of different spoken duration, and that memory score for these words was well predicted by their duration. It has also been shown that fast readers tend to be good memorizers. The relationship between reading rate and memory span thus appears to be remarkably straightforward. Again the ratio of reading rate to span is approximately constant, indicating in the present study that subjects are able to remember as much as they can read out in 1.8 sec. At this stage of research, however, it is probably imprudent to generalize this result too widely. There are many variables which change memory span, but which are unlikely to change reading rate (e.g., list length, word meaningfulness, interpolated delay). It would be of interest to know whether these variables have an effect on the slope, as predicted by decay theory, or on the intercept of the function. Only if the intercept stays consistently near to zero for a variety of conditions can the simple form of decay

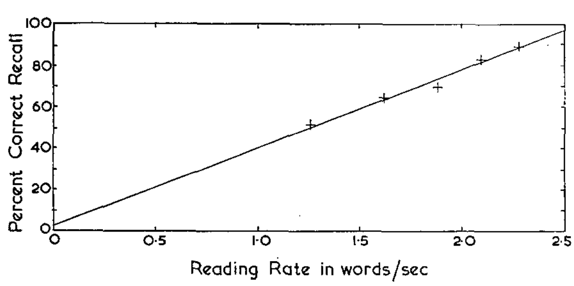

FIG. 4. The relationship between reading rate and recall observed in Experiment VI.

theory under discussion be accepted. The main result of the experiment however, when seen in conjunction with the previous studies, is that short-term memory capacity, as measured by memory span, is constant when measured in units of *time*, not in units of structure.

The time-based system, which presumably underlies the effects observed, is broadly consistent with a decay theory component of short-term forgetting. Decay theory ascribes to rehearsal the role of reviving a decaying trace, and it is this function of rehearsal that requires the prediction of a word length effect. It follows that if rehearsal could be prevented then, providing the presentation rate was the same for both long and short words, no word length effect should occur. The next experiment was designed to test this prediction. The technique used to stop rehearsal was that of articulatory suppression (Murray, 1968) in which the subject is required to articulate an irrelevant item during presentation of the list.

EXPERIMENT VII

In this experiment, the recall of visually presented lists of long and short words was compared under two conditions: (1) with the subject remaining silent during list input and being free to rehearse, and (2) with the subject required to articulate an irrelevant sequence of items. The design thus involved four conditions, comprising two word lengths in each of two presentation conditions.

Method

Two pools of 10 words each were produced, one of one-syllable words and one of five-syllable words, matched for word frequency. From each pool, 16 five-word lists were constructed by sampling at random without replacement. Each set of 16 lists was divided into two equal blocks and a Latin square used to determine order of presentation of the blocks. All subjects did all conditions. The lists were presented at a 1.5-sec rate on a memory drum, and subjects were instructed to recall the items in the order presented. In the suppression conditions, subjects counted repeatedly from one to eight, keeping rate of articulation as constant as possible at about three digits per second. They began counting before the list appeared and stopped to recall as soon as the last item had been presented. In the no-suppression condition, subjects were simply told to try to remember the words. The subjects, 12 undergraduate students from the University of Stirling, were familiarized with the pools of words before being tested.

Results and Discussion

Figure 5 shows the mean percentage of words recalled in the correct serial position as a function of word length for the two presentation conditions. Analysis of variance showed a significant effect of word length, $F(1, 11) = 17.73$, $p < .005$, of suppression, $F(1, 11) = 67.89$, $p < .001$, and a significant interaction between word length and suppression, $F(1, 11) = 16.30$, $p < .005$. The data may be summarized by saying that the

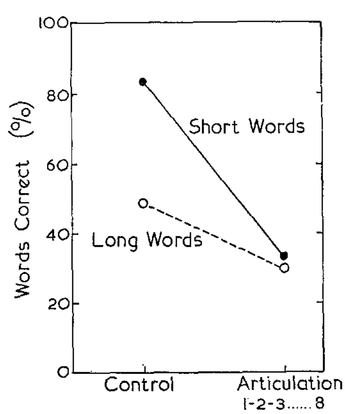

FIG. 5. Effect of articulatory suppression on the word length effect. The influence of articulatory suppression on the recall of auditorily presented long and short words.

word length effect disappears under suppression. Thus, these results are consistent with decay theory if it can be assumed that suppression stops rehearsal. Unfortunately, this latter assumption is open to dispute, since the effects of suppression seem to be dependent on presentation modality (Levy, 1971; Peterson & Johnson, 1971). In particular, suppression has been shown to have a large effect on visually presented material, but little effect on auditorily presented material. It would seem unlikely that suppression stops rehearsal with visual presentation, but not with auditory. An alternative explanation might be to assume that suppression stops the transformation of a visual stimulus into a phonemic code. Thus, given that the word length effect is mediated by a system employing a speech code, and that under suppression, visually presented material does not enter this system, we have an alternative explanation for the above results. Experiment VIII was designed to throw light on this issue.

Experiment VIII

This was essentially a replication of the previous experiment, with the addition of a condition involving auditory presentation. This expanded the design into a $2 \times 2 \times 2$ design, with two levels of word length (one and five syllables), two articulatory conditions (suppression and no suppression), and two presentation modes (auditory and visual). All subjects did all conditions, with the number of replications per condition being reduced to five. The experiment was run in two halves, with the four conditions of one modality in each half. Half the subjects did the visual conditions first, and half the auditory conditions first. Within a modality, the order of the conditions was determined by a Latin square. New pools of words were used, taken from the one- and five-syllable pools of Experiment VI. Presentation rate was slowed to 2 sec; in the auditory condition, the lists were read to the subject, whilst in the visual condition, the lists were presented on a memory drum. In all conditions, recall was verbal. In all other respects the procedure was as for the previous experiment; the subjects were 16 members of the Applied Psychology Unit panel who were paid for their services.

Results and Discussion

The mean percentage of words correctly recalled in the appropriate serial position is shown in Fig. 6. Analysis of variance showed significant effects of word length, $F(1, 15) = 14.02$, $p < .005$, suppression, $F(1, 15) = 85.68$, $p < .001$, and modality, $F(1, 15) = 39.66$, $p < .001$. The interaction between word length and modality was significant, $F(1, 15) = 8.81$, $p < .01$, as was the suppression × modality interaction, $F(1, 15) = 33.13$, $p < .001$. The remaining two-way interaction, word length × suppression, just failed to reach significance $F(1, 15) = 4.49$, $.05 < p < .10$. The three-way interaction did reach significance, $F(1, 15) = 6.23$, $p < .05$. This last result indicates that the change in the word length effect produced by suppression is different in the two presentation modalities.

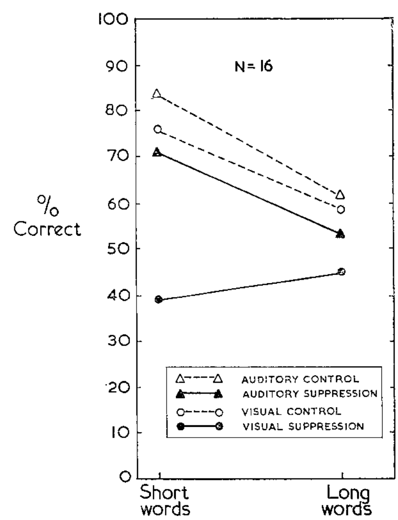

FIG. 6. The influence of articulatory suppression on the recall of long and short words as a function of modality of presentation.

Specifically, the word length effect is abolished by suppression in the visual modality, but is unchanged in the auditory modality.

The results demonstrate very clearly how the effects of suppression change with presentation modality and provide support for the view that suppression stops the visual to auditory transformation. These results can still be fitted into the simple decay and rehearsal hypothesis, but only if the assumption is made that articulatory suppression does not prevent rehearsal, but simply inhibits the translation of visual material into a phonemic code.

General Discussion

The experiments described have shown: (1) That memory span is sensitive to word length across a range of verbal materials. (2) That when number of syllables and number of phonemes are held constant, the word length effect remains. (3) A systematic relationship between articulation time and memory span, such that memory span is equivalent to the number of words which can be read out in approximately 2 sec. (4) That memory span is correlated with reading rate across subjects. (5) That articulatory suppression abolishes the word length effect when material is presented visually.

We shall discuss the implications of these results for existing empirical generalizations about memory, and will then attempt to fit them into a conceptual framework.

The most obvious implication of these results is for Miller's (1956) suggestion that memory span is limited in terms of number of chunks of information, rather than their duration. It suggests a limit to the generality of the phenomenon which Miller discusses, but does not, of course, completely negate it. The question remains as to how much of the data subsumed under Miller's original generalization can be accounted for in terms of temporal rather than structural limitations. Consider, for example, the tendency for subjects' memory span for letter sequences to vary with order of approximation to English. McNulty (1966) has shown that higher orders of approximation to English lead to higher memory span performance when measured in terms of number of letters recalled, but not when measured in terms of number of adopted chunks. It seems highly probable that sequences which can be reduced to a relatively small number of chunks (e.g., THEMILEAKE) will be not only well remembered, but also spoken much more rapidly than sequences which cannot be reduced in this way (e.g., YVSPCWUECR). Such a view has, of course, been explored in some detail by Glanzer and Clark (1962) in connection with the recall of verbally recodable visual patterns. Both letter sequences and the types of pattern used by Glanzer and Clark are encodable into articulatory sequences which can be produced within the 1- to 2-sec limit implied by our results. It would clearly be desirable to explore the relationship between articulation time and memory for materials such as approximations to English prose, which are broadly consistent with Miller's chunking hypothesis (Tulving & Patkau, 1962), but which would seem likely to involve articulation times considerably in excess of 2 sec.

A second general point arises from the contrast between the recency effect in free recall, which apparently shows no word length effect (Craik, 1968; Glanzer & Razel, 1974), and the memory span task, which clearly does show such an effect. This fits in with the general pattern of results mentioned in the Introduction, suggesting that memory span relies on phonemic coding, whereas recency does not. This provides further evidence for the suggestion that the two reflect quite different underlying memory processes. Since most current views of the nature of short-term memory assumed a common primary memory system underlying both, it is unclear how they would interpret the word length effect.

One approach which does not have this drawback is the framework suggested by Baddeley and Hitch (1974), who explicitly postulate a working memory system which is responsible for performance on memory span tasks, but is not responsible for the recency effect in free recall. The formulation is based on a range of experiments in which subjects were required to perform reasoning, prose comprehension, or free recall learning tasks while simultaneously holding sequences of up to six random digits in short-term memory. In general, the results suggested that subjects could hold up to three items with virtually no effect on performance, but when required to remember six items, a decrement appeared. A tentative formulation was suggested in terms of a working memory system which acts as a central executive, and a supplementary articulatory rehearsal loop with a capacity of about three items.

Most of the experiments in the present series fit neatly into this broad framework, on the assumption that the word-length effect is the result of the limited capacity of the rehearsal loop. Looked at from this viewpoint, our data suggest that the articulatory loop system is time-based, and hence has a temporally limited capacity. When access to the loop is prevented by articulatory suppression, memory depends entirely on the capacity of the executive working memory system, which is not phonemically based, and does not have the same temporal limitation as the articulatory loop. The tendency for memory span to be impaired by phonemic similarity among the items to be remembered can also be attributed to the operation of the articulatory rehearsal system. As in the case of the word-length effect, the phonemic similarity effect disappears when articulatory suppression occurs with visually presented material (Levy, 1971). Finally, the existence of patients who have drastically impaired digit span, and yet who appear to show none of the general cognitive impairments that might be anticipated from most views of the role of short-term memory (Shallice & Warrington, 1970), can readily be accounted for within this framework if it is assumed that such patients are defective in the operation of the articulatory rehearsal system, while having the executive component of the working memory system intact.

One of our results however does present a problem for such a view. This is raised by the observation that the word-length effect does occur despite articulatory suppression, provided the material is presented auditorally. Levy (1971) has shown a similar pattern of results for phonemic similarity, with the similarity effect disappearing under suppression when visual presentation is used, but not when the presentation is auditory. On the straightforward assumption of an articulatory rehearsal loop which is entirely synonymous with subvocalization, it should follow that suppression effects could not be avoided by auditory presentation. Experiment VIII, however, suggests that although articulatory suppression produces an overall impairment in performance with auditory presentation, it does not influence the word-length effect. This would therefore seem to point to articulation as being a means of converting the visual stimulus into a phonemic code which may be accepted by some form of storage system. With auditory presentation, the material is presumably already encoded in an appropriate form, and can be fed into the supplementary system without the need for articulation. The fact that articulatory suppression still impairs performance, even with auditory presentation, may imply either that there is an additional advantage to be gained by articulation, or simply that the task of suppression provides a secondary task which takes up some of the general processing capacity which might otherwise be devoted to remembering the items presented.

Suppose one tentatively assumes a supplementary phonemically based store, what might its other characteristics be? Could it, for example, be equivalent to the precategorical acoustic store suggested by Crowder and Morton (1969)? This seems unlikely for two reasons: First because the word length effect occurs with visual presentation, provided suppression is avoided, whereas the precategorical acoustic store does not appear to be operative unless auditory presentation is used. Secondly, Watkins and Watkins (1973) have presented evidence suggesting that the precategorical acoustic store is not sensitive to the effect of word length; if this is so, it can clearly not be used to explain the word-length effect. An alternative is to suggest that the system is an output buffer of some type: A limited-capacity store for holding the motor program necessary for the verbal production of letter names has been suggested by Sperling (1963). It seems plausible to assume that some form of buffer store is necessary for the smooth production of speech; and indeed the existence of the eye–voice span in reading points to some such temporary storage process (Morton, 1964), since what the reader is saying when reading aloud lags consistently behind the point at which he is fixating. Such a buffer system would need to be separate from the act of articulation, since it is presumably necessary to set up new articulatory programs while existing programs are operating. On this interpretation, therefore, articulatory programs can be set up or at least primed, either by the act of overt or covert articulation, or indirectly through auditory stimulation. It is tentatively suggested that such a system may be necessary for fluent speech, and may have the supplementary advantage of providing an additional backup system for the immediate retention of phonemically codable material. Such a view is clearly very tentative and leaves unspecified the complex problem of how such a store might be interfaced with the other components of the system so as to account for even the basic phenomena of the memory span. It does, however, have the advantage of linking together the existing data in a way which is both internally consistent and also likely to generate testable hypotheses.

References

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

Baddeley, A. D., & Hitch, G. Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation.* New York: Academic Press, 1968, Vol. 8.

Broadbent, D. E. *Perception and communication.* London: Pergamon Press, 1958.

Conrad, R. Acoustic confusion in immediate memory. *British Journal of Psychology*, 1964, **55**, 75–84.

Conrad, R., & Hille, B. A. The decay theory of immediate memory and paced recall. *Canadian Journal of Psychology*, 1958, **12**, 1–6.

Craik, F. I. M. Two components in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 996–1004.

Craik, F. I. M. The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 143–148.

Craik, F. I. M., & Levy, B. A. Semantic and acoustic information in primary memory. *Journal of Experimental Psychology*, 1970, **86**, 77–82.

Crowder, R. G., & Morton, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, **5**, 365–373.

Glanzer, M., & Clark, W. H. Accuracy of perceptual recall: An analysis of organization. *Journal of Verbal Learning and Verbal Behavior*, 1962, **1**, 289–299.

Glanzer, M., Koppenaal, L., & Nelson, R. Effects of relations between words on short-term storage and long-term storage. *Journal of Verbal Learning and Verbal Behavior*, 1972, **8**, 435–447.

Glanzer, M., & Razel, M. The size of the unit in short-term storage. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 114–131.

Goldman-Eisler, F. The significance of changes in rate of articulation. *Language and Speech*, 1961, **4**, 171–174.

Levy, B. A. The role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 123–132.

Mackworth, J. F. The relation between the visual image and post-perceptual immediate memory. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 75–85.

McNulty, J. A. The measurement of "adopted chunks" in free recall learning. *Psychonomic Science*, 1966, **4**, 71–72.

Miller, G. A. The magical number seven, plus or minus two: Some limits to our capacity for processing information. *Psychological Review*, 1956, **63**, 81–97.

Morton, J. The effects of context upon speed of reading, eye movements and eye-voice span. *Quarterly Journal of Experimental Psychology*, 1964, **16**, 340–354.

Murray, D. J. Articulation and Acoustic confusability in short-term memory. *Journal of Experimental Psychology*, 1968, **78**, 679–684.

Murray, D. J. Articulation and Acoustic confusability in short-term memory. *Journal of Experimental Psychology*, 1968, **78**, 679–684.

Peterson, L. R., & Johnson, S. F. Some effects of minimising articulation on short-term retention. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 346–354.

Richardson, J. T. E., & Baddeley, A. D. The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 623–629.

Shallice, T. & Warrington, E. K. Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, 1970, **22**, 261–273.

Sperling, G. A model for visual memory tasks. *Human Factors*, 1963, **5**, 19–31.

Sperling, G., & Speelman, R. G. Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

Tulving, E., & Patkau, J. E. Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology*, 1962, **16**, 83–95.

Watkins, M. J., & Watkins, O. C. The postcategorical status of the modality effect in serial recall. *Journal of Experimental Psychology*, 1973, **99**, 226–230.

Reference Notes

1. Laughery, K. R., Lachman, R., & Dansereau, D. D. Short-term memory: Effects of item–pronunciation time. (Unpublished.)
2. Standing, L., Bond, B., & Smith, P. The memory span. (Unpublished.)

(Received July 1, 1975)