SHORT TERM MEMORY

Dissertation submitted for the
Degree of Doctor of Philosophy
in the University of Stirling

by

Neil Thomson

University of York
November, 1979

ABSTRACT

The eight experiments reported in this thesis are designed to investigate the idea that in verbal short-term memory (STM) material decays over time and this decay is prevented by rehearsal. It follows that the capacity of STM when measured in words should be inversely proportional to the time taken to rehearse the words. Consequently, subjects should be able to recall more short duration words than long duration words. In contrast to this hypothesis is the idea that the capacity of STM is a fixed number of chunks, where chunks are a structural characteristic of the material.

The first four experiments are designed so that these alternative hypotheses produce conflicting predictions and, in all cases, the hypotheses derived from decay theory are supported. It is shown that serial recall performance is very well predicted by the time taken to say the words and that the relationship between word duration and recall is of the type predicted by decay theory.

The second set of experiments are based on the assumption that both STM and long-term memory (LTM) contribute to performance in serial recall tasks. The purpose of the experiments is to determine whether it is the STM or LTM component that is sensitive to word duration. It is predicted, in line with a decay theory of forgetting in STM, that the STM component is sensitive to word duration. The experiments are designed to produce sizable contributions from both stores in order to test this hypothesis. The results support the hypothesis in showing that variables known to affect STM, such as acoustic similarity, interact with word duration, while variables known to affect LTM, such as repeated presentations of the same list, show no such interaction.

The results are interpreted in terms of decay theory and the different versions of this theory that have been proposed are considered. It is concluded that while no version of the theory is completely adequate, there is no evidence that invalidates the central assumptions, viz. that in STM items are forgotten by decay and that one of the functions of rehearsal is to prevent this decay.

PREFACE

The experiments reported here are original and were carried out by me and unless it is stated otherwise the ideas expressed are original.

I would like to thank Dr. Alan Baddeley, who supervised this work while he was at Stirling University, for his support and encouragement. I am grateful to Dr. Peter Hamilton and Dr. Philip Smith for taking over the supervision of this work when Dr. Baddeley left Stirling.

I should also like to thank Mrs. Marilyn Glicker who typed this thesis.

Neither this dissertation nor any part of it has been submitted to any other university.

CONTENTS

INTRODUCTION

The variety of meaning in human communication is mainly achieved by combining simple elements into different sequences. This ability to use serial order has long been recognized as a question of central importance in Psychology (Lashley, 1951). It is still, however, not well understood despite the abundance of research into serial recall and the mechanisms underlying it. The latter have been most commonly conceptualized as short- and long-term memory (STM and LTM) and STM in particular has conventionally been regarded as an input system which receives and holds information prior to its encoding in LTM (Waugh & Norman, 1965; Atkinson and Shiffrin, 1968). An alternative proposal has been that STM is an output store involved in the production of speech (Morton, 1970; Ellis, 1976). Although placing STM at opposite ends of the information processing system, both proposals require that the store be capable of dealing with serially ordered information. Speech as a form of human communication is serial by nature while the permanent storage of information cannot be. Thus, given that people decode serial information into a permanent message and vice-versa, the constraints of serial ordering will be felt most strongly at the points of perception and production. At input, the store would be responsible for holding serially organized information until it could be encoded into a permanent message, whereas at output, the store would hold material translated from a permanent message while it was being produced serially. In neither case would STM be holding material in the code used for permanent storage and consequently the meaning of the message should not interact with the properties of the store. Further, it might be expected that the store would have a relatively small capacity as the "perceptual units" of speech are fairly short whether taken as syllables, words, tone groups or phrases. By this account, then, STM is a peripheral store with a limited capacity, employing some form of speech code and insensitive to the meaning of a message. This description of STM is very much in agreement with those derived from experimental work as will become apparent from the research review below. However, there is one very important experimental result which would seem to contradict directly the assumption that STM is insensitive to the meaning of a message and it is this contradiction that provides the starting point for the thesis.

Miller (1956) showed that the amount people could remember in a serial recall task was 7±2 chunks where a chunk was defined as a subjective meaningful unit, e.g. HPD might be three chunks while PHD would be one. In so far as chunks are "meaningful units" they are units of permanent storage rather than units of some peripheral code and consequently, Miller's result would not be predicted by the view that STM is a peripheral store. Thus, there is a contradiction in the literature between the experimentally supported idea that STM is a peripheral store employing a speech code and the fact that the capacity of STM is constant when measured in units of permanent storage. The general aim of this thesis is to resolve this contradiction by testing the chunking hypothesis. This is not as straightforward as it might seem, because the chunk is defined subjectively and hence the hypothesis is essentially irrefutable. 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. Where lists are composed of sequences of these items chosen to be unrelated one to another this assumption seems plausible and allows a weaker version of Miller's hypothesis to be tested: that memory span is a constant number of chunks where chunks are defined experimental units. It will be shown that for tasks which rely heavily on STM this hypothesis fails for a variety of definitions of a chunk and Miller's results will be explained by arguing that he used tasks in which there was a large LTM contribution to recall. It will be assumed that, because LTM is responsible for the storage of a permanent message, it employs a semantic code and consequently is responsible for chunking.

The experimental work is based on two main assumptions: first, that at least two stores mediate performance on memory span and serial recall tasks, viz. STM and LTM, and secondly, that STM is a peripheral store responsible for holding the units of perception or production in a prescribed order. Before introducing the experiments, these two assumptions will be evaluated in the context of previous research.

RESEARCH REVIEW

The idea that STM is a peripheral store and particularly one involved in speech perception or production implies that it should employ a speech code; acoustic if it is an input store and articulatory if it is an output store. Conrad (1964) showed that errors produced in immediate serial recall were similar to those produced when attempting to identify verbal material in noise and, in particular, items that sounded similar tended to be confused. He called this the acoustic similarity effect (AS) and it implies that STM is sensitive to the physical properties of a message as would be expected of a peripheral store. Further, he showed that the AS effect was not dependent on the material being presented auditorially, but also occurred with visually presented material. This latter result is particularly important as it suggests that the store mediating performance is not modality specific. Thus, if it is to be argued that STM is an input store, it must occur after the material has been encoded in a form that is independent of modality. This would mean that the store is not subject to the same constraints of serial ordering as a truly peripheral store responsible for receiving serially organized messages. However, the experimental evidence reviewed below suggests that STM is a peripheral output store responsible for holding serially organized articulatory commands while they are being produced and consequently would not be modality specific.

Baddeley (1966a; 1966b) extended Conrad's results and showed that, while AS was detrimental to immediate serial recall, semantic similarity had little effect on performance. However, when list length was increased and a rehearsal preventing task interpolated between presentation and recall, AS had little effect while semantic similarity consistently impaired performance. Baddeley concluded that, in the first experiment, subjects were using STM almost exclusively and that this store employed a speech code, while in the second task, subjects were forced to make more use of LTM which employs a semantic code. Confirmation of these results was obtained by Conrad (1967) who showed that a distractor task removes the AS effect. Both Baddeley and Conrad assumed that there are two components underlying serial recall and that the STM component is rapidly lost if rehearsal is prevented.

These results are good support for the idea that STM is a peripheral store in so far as it is sensitive to the physical properties of the message rather than its semantic properties. However, they do not allow the location of the store in the information processing system to be determined. The fact that AS effects occur with visually presented material does not rule out the possibility that the code could be acoustic; it is not inconceivable, although perhaps rather implausible, that visual material could be translated into an acoustic code. Also, as Wickelgren (1969) has argued, it is possible that an abstract phonological code is used implying, as mentioned above, that STM is not as peripheral as has been claimed. The next set of experiments to be reviewed are attempts to determine the precise nature of the code employed by STM and, in particular, whether it is acoustic or articulatory.

Hintzman (1967) attempted to determine whether the AS effect was due to the material sounding similar or because the articulatory movements (or commands) necessary to produce AS material were similar. His approach was to try to find errors which were articulatory but not auditory or vice-versa. This is difficult due to the high correlation between acoustic and articulatory variables but Hintzman claimed that "place of articulation" was such a variable. He based the claim on some results of Miller and Nicely (1955) who showed that, in a task where speech sounds had to be identified in noise, there were many errors between consonants which differed only in place of articulation. However, these errors were not systematic, thus "D" and "G" were equally likely to be confused with "B" although "D" is closer to "B" than "G" in articulatory terms. In serial recall, however, Hintzman found that errors between items that differed only in place of articulation (such as "B", "D" and "G") were highly systematic in a manner predicted by the closeness of the articulatory features. Thus, he concluded that STM employs an articulatory rather than an acoustic code, implying that STM is a peripheral store involved with the production of speech. Hintzman's results have been criticized by Wickelgren (1969) who points out that the masking noise used by Miller and Nicely might have differentially masked different articulatory features, e.g. fricatives, characterized by a short noiseburst, could be expected to be masked more effectively by broad-band noise than vowels which have energy concentrated around a narrow band of frequencies. However, differences in place of articulation (labial, dental and velar for "B", "D" and "G") cause the length of the silent interval before vowel onset to be changed and it seems unlikely that white noise would differentially affect these. Ultimately, it is an empirical question whether Wickelgren's arguments discredit Hintzman's results and, to date, the appropriate experiments have not been done.

Levy (1971) was interested in the same question as Hintzman but used a rather different method to try to answer it. She started with the idea that either or both acoustic and articulatory codes might be employed and attempted to manipulate the serial recall paradigm in ways which altered the availability of these codes. She made two important assumptions: (1) That an auditory code can only be used when the material is presented auditorially, and (2) that articulatory suppression (Murray, 1968), i.e. saying an irrelevant item repeatedly during list presentation, prevents the use of an articulatory code. Thus, she combined the conditions of auditory and visual presentation with those of no suppression and suppression. (In fact the design was slightly more complicated than this but these are the only conditions relevant to the present discussion). According to her assumptions, subjects were able to employ an auditory code (auditory presentation with suppression), an articulatory code (visual presentation with no suppression), both auditory and articulatory codes (auditory presentation with no suppression) or neither (visual presentation with suppression). The results showed that there was a decrement in performance only for the condition in which neither code was available, i.e. there was no difference in performance when auditory, articulatory or both of these codes were available.

Levy concluded that subjects can use either code but do so disjunctively; subjects use one or the other of the codes but not both. However, the assumptions Levy makes are open to question. There is no a priori reason why an acoustic code should not be generated from a visual representation and secondly, there is no independent evidence that suppression stops the formation of an articulatory code. Her reasons for making this latter assumption are based on the idea that it is impossible to say two things at once, but the formation of an articulatory code is not the same thing as the execution of that code. Indeed, it would be rather strange if it were not possible to construct an articulatory code while speaking. It would imply that speech could not be fluent as each segment would have to be translated into an articulatory code and executed before the next section could be dealt with. It is more plausible to suppose that encoding and execution can occur in parallel. If Levy's assumption that articulatory suppression stops the formation of an articulatory code is abandoned, then rather different assumptions have to be made to explain her results.

The alternative explanation is that the material does not enter STM when presented visually with articulatory suppression, and in support of this view an experiment by Estes (1973) showed that the AS effect disappeared when material was presented visually with suppression but not when presented auditorially with suppression. However, this explanation assumes that suppression does not stop entry to STM with auditory material but only with visual material and there is no independent evidence that this is the case. The idea is not implausible, however, in that auditory information is very close to an articulatory code while written material would require far more translation and it is conceivable that suppression interferes with this translation process. Such ideas must remain at the level of speculation until far more is known concerning the effects of suppression. In fairness to Levy, it should be pointed out that there is no evidence which directly contradicts her explanation; the experiment by Estes (1973) fits quite well with Levy's theory in that if neither acoustic or articulatory codes are available then no AS effect should be apparent. However, to accept Levy's view is to abandon the idea of an STM which employs a specific code and to move towards a "levels of processing" approach (Craik and Lockhart, 1972).

By a levels of processing view, material is encoded in a variety of manners but the code used will depend on the nature of the material and the task. The view does not assume the existence of independent memory stores, but rather of different "levels" of coding. It is not proposed to discuss the merits and problems with this approach but it is fair to point out that it has been very strongly criticized for failing to be empirically distinguishable from a multi-store approach (Glanzer, 1976) and for an inherent circularity (Nelson, 1977; Baddeley, 1979). The view taken here is that, at present, there is no compelling reason to abandon the multi-store approach in favour of levels of processing and all explanations of the data will be couched in terms of STM and LTM stores.

Returning to the question under discussion of whether STM employs an acoustic or articulatory code, a rather different approach has been taken by Conrad (1970). He found that deaf people only showed the AS effect when they had some ability to speak. Conrad interpreted these results by assuming that STM employs an articulatory code and that those who could speak took advantage of this system in performing the task and hence showed the AS effect. Those who could not speak were unable to form an articulatory code and thus relied on a visual or some other type of representation. This experiment would seem to provide strong evidence that STM does not employ an acoustic code as presumably congenitally deaf people have never formed such a representation. However, it remains possible that the STM code is a more abstract phonetic code rather than an articulatory code as the ability to speak may well require the development of such representations.

In complete contradiction to Conrad's results, a recent paper by Colle and Welsh (1976) contains a set of experiments which seem to show that STM employs an acoustic rather than an articulatory code. Their procedure was to show that serial recall performance can be impaired by an irrelevant auditory message played to the subject while he is trying to remember a list of visually presented items. So that it could not be claimed that the irrelevant message interfered with an articulatory code, Colle and Welsh used a message spoken in German (which none of the subjects were able to speak). They found that recall was reduced when the irrelevant message was played and interpreted this result as evidence for acoustic coding in STM. However, their experiment is open to a number of objections: first, German and English share many common phonemes and consequently it may not be claimed that subjects are incapable of producing an articulatory code of the irrelevant message. Secondly, the serial recall task they used was unusual as there was a ten second silent interval between presentation and recall of the list and I have shown, in a set of experiments which do not form part of this thesis, that the effect disappears if this silent interval is removed. This makes Colle and Welsh's results difficult to interpret, but it seems possible that the irrelevant message was interfering with the subjects' rehearsal during the silent interval. In any event, as they stand these results cannot be taken as strong evidence for an acoustic code being employed by STM. Further, an experiment by Salame and Wittersheim (1978) suggests that "pink noise" affects serial recall by interfering with the perception rather than storage of the list items. If STM employs an acoustic code it should be predicted that the noise would interfere with the items in store, but this did not seem to occur.

To summarize the results discussed so far, it seems that the evidence strongly supports the idea that STM employs a speech code of some form. While no single experiment is crucial the evidence as a whole favours an articulatory code and consequently the idea that STM plays some role in the production of speech. However, the possibility that STM uses a more abstract phonetic code (say) cannot be ruled out. In addition to the experiments already discussed there is a body of evidence which, it has been claimed, supports the idea that STM employs a semantic code. If this were true, it would be very damaging to the position that STM is a peripheral store but would accord with Miller's chunking hypothesis as chunks are semantic units. In reviewing this work the arguments advanced by Baddeley (1972) will be used, viz. that the semantic coding found in serial recall tasks is due to the LTM component present in this paradigm. There is always the danger of circularity in adopting this position defining LTM as the store which employs a semantic code and then ascribing any evidence of semantic coding to LTM. This circularity may be avoided if other criteria are used for deciding whether material is recalled from STM and/or LTM and an obvious candidate is the presence or absence of a distractor paradigm. It is widely accepted that a rehearsal preventing task removes the STM component from recall, the main support for this assertion being the loss of the AS effect with a distractor task (Baddeley, 1966b; Conrad, 1967).

As already reported Baddeley (1966a, 1966b) showed that with relatively long lists and a distractor task, semantic similarity between list items was detrimental to recall. The distractor task was used to remove the STM component of recall and consequently this result shows quite clearly that LTM is sensitive to semantic structure. The additional result, that acoustic but not semantic similarity affects immediate recall, has been used as evidence that STM employs a speech code, but it also demonstrates that STM does not employ a semantic code. However, Shulman (1971) and Craik and Lockhart (1972) have argued that a semantic code takes longer to establish than an acoustic code and that this is why there were no semantic similarity effects in immediate recall. While Wetherick (1975) and Huttenlocher and Newcombe (1976) have showed that serial recall is improved by blocking list items into semantic categories, the latter authors additionally demonstrated that the advantage of blocked over random lists did not change as presentation rate increased. If a semantic code takes longer to establish than a speech code, then the advantage of a semantically organized list should only occur when the subject has sufficient time to complete the semantic coding and consequently should be reduced at fast presentation rates. The fact that this does not occur suggests that Schulman and Craik et al are wrong in their explanation of the data, although it is always possible for them to claim that the presentation rates used were not fast enough. Further evidence against their ideas comes from the work on semantic priming and subliminal perception. This work (e.g. Worthington, 1964; Wickens, 1970; Marcel, 1979) suggests that the meaning of a word can be processed at least as quickly as the word's physical and lexical properties. Marcel, using a semantic priming paradigm, presented letter strings which were either words or non-words and the subject's task was to make this word/non-word decision as quickly as possible. Before each letter string appeared, a priming word was presented which was followed by a mask so that the subject could not report what he had seen. Even so, the word/non-word decision was faster if the priming word was semantically related to the word in the letter string than if it was unrelated. The interesting fact is that subjects had evidently processed the meaning of the priming word but could not report its physical properties and one explanation of this would be that the meaning of a word is abstracted faster than the word's physical properties. While other explanations are possible, the experiment does underline the point that it is by no means certain that semantic information takes longer to encode than physical information as Schulman and Craik et al. have asserted.

A number of studies which claim to show semantic coding in STM have used tasks other than serial recall. In particular, the necessity to remember the order of the items has been removed and, when this is done, the task becomes free- rather than serial recall. Some confusion has arisen in this area because of the different terms used to describe paradigms and stores. Some have adopted the convention of calling tasks which involve immediate recall (or recall after a relatively short delay) "short-term memory tasks" and have used the terms primary and secondary memory to refer to the proposed stores mediating performance. This convention has not been universally accepted, however, and the terms short- and long-term memory are commonly used (as here) to refer to memory stores and the different paradigms are described by names such as delayed recall, immediate recall, etc. There is now a good deal of evidence that the important distinction is not between immediate and delayed recall tasks but rather between serial and free recall, and a useful convention is to use the words primary and secondary memory for the stores underlying free recall and short-term and long-term memory for the stores underlying serial recall. While it is almost certainly the case that LTM and secondary memory are the same thing, it is far less clear that STM and primary memory are the same store. In other words, the store responsible for the recency effect in free recall (primary memory) and the store partly responsible for performance in a serial recall paradigm (short-term memory) may well be different mechanisms. This assertion is in direct contradiction of early models of memory (e.g. Atkinson and Shiffrin, 1968) which proposed that recency and memory span were mediated by the same store. However, Baddeley and Hitch (1977) review a considerable amount of data which they argue supports the idea that recency is the product of a retrieval strategy rather than of short-term memory. Thus, the recency effect in free recall is unaffected by AS (Craik and Levy, 1970; Glanzer, Koppenaal and Nelson, 1972) or by articulatory suppression (Richardson and Baddeley, 1975) unlike serial recall which, as discussed above, is affected by both these variables. Further evidence comes from Craik (1970) who observed that memory span correlates more highly with the secondary memory component of free recall than with the primary memory component. Finally, Baddeley and Hitch (1974) have shown that recency is unimpaired in a free recall task for subjects performing concurrent memory span tasks. Since the memory span task did not interfere with recency, it is difficult to maintain the view that the two tasks are mediated by the same limited capacity store.

This conclusion allows rejection of the claim that STM employs a semantic code when the basis for the claim is experiments using free rather than serial recall. Thus, papers by Schulman (1971), Klein and Klein (1974) and Goldstein (1975) all used probe tasks in which the subject does not have to remember the order of the items and consequently may not have been mediated by STM. These results have been used to argue against a multi-store view and for a "levels of processing" approach (e.g. Cermak, 1972; Craik and Lockhart, 1972; Nelson, Wheeler, Bordon and Brooks, 1974) but in the context of the present arguments it would seem unnecessary to discuss them any further here.

The next set of experiments to be considered in this section are more directly concerned with Miller's chunking hypothesis. In claiming that subjects' STM capacity is 7±2 chunks, where a chunk is a subjective meaningful unit, Miller evidently conceived of only one mechanism underlying span; one that was capable of utilizing the semantic structure of the material. The arguments presented here have all been based on the idea that both LTM and STM mediate performance on memory span and serial recall tasks and indeed it was necessary to assume this to explain the

AS and semantic similarity effects. Thus, to be consistent, it should be the case that in so far as subjects employ chunking, they do so by using LTM. Miller and Selfridge (1950) measured subjects' span for different approximations to English and found that the two measures were related; the higher the approximation to English of the material the greater the subjects' span. Although consistent with the chunking hypothesis, it is difficult to use this data as strong support, for without an independent measure of a chunk, the number of chunks recalled for the different types of material could not be estimated. Tulving and Patkau (1962) and McNulty (1966) attempted to avoid this problem by operationally defining chunks as a sequence of two or more words recalled in the correct order. They found that using this measure there was no change in performance over different approximations to English; about seven chunks were recalled in each condition.

These data certainly suggest that where possible subjects employ a chunking strategy in serial recall tasks but, assuming that STM and LTM underlie serial recall, it is not possible to determine which of the stores is sensitive to semantic structure. An important point is that all these experiments used material which was suitable for semantic coding, while experiments showing AS effects have typically employed short lists of unrelated words. It seems reasonable to argue that if subjects use both LTM and STM in these tasks they will rely most strongly on the store which is most suitable for the task. Thus, material with little or no semantic structure produces strong AS effects, while approximations to English should cause LTM to be used to a greater extent and hence evidence of chunking to become apparent. This is not to say that STM will not contribute to recall in tasks where the material has semantic structure, but rather that the relative contribution of LTM will increase and tend to swamp any STM contribution. Evidence for these arguments is provided by Baddeley and Levy (1971) who showed that subjects had great difficulty in semantically coding meaningless paired associates presented for a single trial and that under these conditions, subjects relied heavily on rote rehearsal and, by implication, STM. The implication of these arguments is, that while LTM may well employ a chunking strategy, there is no reason to believe that STM does and thus the question of the capacity of STM remains open.

One final piece of evidence in support of this assertion comes from Tulving (1966). He found that something very like chunking occurs in free recall where subjects, when repeatedly presented with the same list, tend to become consistent in the order of recall of the items. This phenomenon Tulving called "subjective organization", but it is very similar to the operational definitions of chunking used by Tulving and Patkau (1962) and McNulty (1966), and the fact that it occurs over the secondary memory region of the serial position curve could be taken as support for the view that maybe grouping or chunking strategies are mediated by LTM. Other forms of grouping effect, especially those produced by the list being temporally grouped at input (Ryan 1969) are more difficult to ascribe to LTM. These will be considered in the discussion chapter, but it is worth noting that Glanzer (1976), using a free recall paradigm, found that temporal grouping at input affected the secondary memory (LTM) component. However, in serial recall paradigms the grouping effect occurs under conditions which should produce large STM and small LTM components and it is this data that will be discussed later.

The above arguments leave the question of the capacity of STM unanswered and this question is the main concern of this thesis. In considering the question of capacity, an important distinction should be made between the "physical capacity" of a store and its "practical capacity". Sperling (1960) showed by the ingenious technique of partial report, that the physical capacity of iconic memory was greater than the estimate obtained when conventional recall tasks were used. The conventional recall tasks underestimated the capacity of the store because there was a very fast rate of forgetting and consequently, given the relative slowness of recall, material was being lost while the subject was recalling. Thus, it would seem worth making a distinction between capacity as a measure of what a store can physically hold (physical capacity) and capacity when measured by recall (practical capacity), the latter being affected by such things as rate of forgetting, inter-item interference, etc. In fact, it is probably impossible to measure accurately a store's physical capacity as any measure of performance will take time during which forgetting could occur. Different theories of forgetting make different assumptions about the physical and practical capacities of STM. For example, decay theory assumes no limit on the store's physical capacity, but that its practical capacity will be determined by the rate at which information is received and the rate at which it decays. Displacement theory on the other hand, assumes that the store's physical capacity is fixed and that forgetting only occurs when this is exceeded. Thus, a sensible way to begin an inquiry into the capacity of STM is to consider the nature of forgetting in this store.

Brown (1958) and Peterson and Peterson (1959) independently showed that if after presenting a short sequence of items to a subject he was required to perform a task which prevented him from rehearsing (a distractor task), recall of the items was very poor. The amount of forgetting increased with the time spent performing the distractor task up to about 18 seconds after which there was little change. Brown argued that this result supported decay theory as the amount of forgetting was not changed by the similarity between the items in the list and those in the distractor task. Further confirmation of this last point was provided by Posner and Rossman (1965) who showed that it was not the similarity between the list items and the distractor material that determined recall but rather the amount of information that had to be processed in the distractor task. Melton (1963) argued that while Brown's results might rule out the idea that forgetting occurs by retroactive interference (RI), it is possible that rather than forgetting occurring by decay, it is caused by interference from previous list items (PI) and by interference between items in the same list. Support for this argument came from work by Keppel and Underwood (1962) who showed that forgetting on the first trial of the distractor task was almost zero.

More recently, a great deal of work has been done on this problem by Wickens and his colleagues (Wickens, Bjorn and Allen, 1963; Wickens, 1972). He has demonstrated that if list items are used which belong to the same conceptual category (e.g. the names of dogs) for a number of trials, forgetting increases to a relatively high level. However, if at this point a different class of list items is used (e.g. the names of plants), there is a substantial improvement in recall on this trial. Wickens explains this phenomenon by arguing that PI builds up over trials and that interference is increased by using semantically similar items. Switching to a new category of items causes a reduction in interference and hence an improvement in recall. This phenomenon he calls "release from PI" and it seems to show quite clearly that much of the forgetting in the distractor paradigm is due to interference from previous list items. The important question for the present purposes is whether this forgetting occurs in STM or LTM. The fact that a release from PI occurs with a change in semantic category would imply, in the context of the present arguments, that the forgetting is occurring in LTM. Consistent with this view is the fact that the release from PI when items are made similar along physical dimensions is extremely small compared to semantically similar items (Wickens, 1972). In support of this claim it will be remembered that Baddeley (1966b) only found a semantic similarity effect in serial recall when a distractor task was used. With immediate recall he found no semantic similarity effect but rather an AS effect. This result was interpreted by assuming that with immediate recall subjects rely heavily on STM and do not attempt to encode the material into STM. With delayed recall, however, it was assumed that subjects attempt to encode the material into LTM hence producing semantic similarity effects. For this argument to be coherent it must be explained why subjects use LTM when a distractor task is present and the most plausible reason is that distraction causes forgetting from STM (as well as from LTM) to a degree where its contribution to recall is virtually zero. This explains why there is no AS effect with a distractor task, but, so far, all the evidence has suggested that the distractor task causes forgetting from LTM through PI. What is required is an experiment which shows that there is also forgetting from STM with this paradigm. Baddeley and Scott (1971) used data from a number of sources to show, that if only the first trial was considered, forgetting did occur but reached a maximum after only five seconds of distraction rather than the eighteen seconds found when all trials were considered. As there can be no PI on the first trial, this could be interpreted as forgetting from STM and would suggest that material decays very quickly in this store if it is not rehearsed. More evidence is needed before this interpretation can be made with confidence, but it is at least consistent with the arguments presented so far.

A further implication of this interpretation is that interference does not seem a likely explanation of forgetting in STM. As mentioned above, there is no change in the amount of forgetting when the similarity of the list items to the distractor items is varied and further, the forgetting that is being ascribed to STM could not be caused by PI as it occurred on the first trial. The main alternatives that have been proposed to explain forgetting in STM are decay theory (Brown, 1958; Conrad and Hille, 1958) and displacement theory (Waugh and Norman, 1965) and these will now be considered.

Forgetting by displacement is closely tied to the idea that STM has a limited physical capacity. Forgetting is due the the store being full and new items displacing existing ones in competition for space. Advocates of this view argue that rehearsal stops other items from entering STM and it is by this means that it prevents material being forgotten. Thus, it is assumed that material gains automatic access to STM unless it is actively prevented from doing so, a view allied with the idea that STM is a peripheral input store (Waugh and Norman, 1965). The main problem with this theory in its "pure" form is an inability to cope with AS effects. Unless other assumptions are made, it cannot explain why less AS items can be stored than acoustically different items. A possible extension of the theory that allows it to cope with the AS effect is the assumption that items are stored at locations determined by their physical properties and, therefore, that AS items will be stored at similar or overlapping locations. If displacement acts at the points of item location then AS items will be competing with one another for the same space. This assumption alters the theory quite radically as there is now no limit on the store's physical capacity as was previously assumed, but rather, its capacity will be completely determined by the nature of the items to be stored; the more similar the items the less can be stored.

However, this form of displacement theory predicts that increasing item similarity should produce a loss of item information whereas, in fact, it causes a loss of order information (Wickelgren, 1965) and enhances item recall (Craik, 1968). Alternative modifications of the theory which allow a satisfactory explanation of AS effects have not been proposed and Crowder (1976) concludes that displacement cannot be a satisfactory theory of forgetting in STM until it is able to explain similarity effects.

Decay theory assumes that material is forgotten over time and unlike the other theories discussed does not state the cause of forgetting. For this reason it has never been felt to be a wholly adequate explanation of forgetting and has often formed the "null hypothesis" in experimental designs. However, it does predict that forgetting will occur over an empty period of time if the subject can be prevented from rehearsing. This prediction was tested by Reitman (1971) whose subjects performed a difficult tone detection task during the retention interval in serial recall. She found that no forgetting occurred and because she had taken precautions to prevent the subjects rehearsing inferred that decay did not occur and that forgetting was due to displacement. However, in a later study where she took extreme pains to determine what her subjects were doing in the retention interval she found that they had, in fact, been rehearsing (Reitman, 1974). Further, when she divided the subjects according to whether or not they rehearsed, forgetting only occurred in those that had not rehearsed during the retention interval. She concluded that her results supported decay theory over displacement.

Wingfield and Byrnes (1972) attempted to test decay theory using the dichotic listening technique developed by Broadbent (1954). In this paradigm, two streams of information are presented simultaneously to both ears and it is known that subjects typically report all the items presented to one ear and then those presented to the other ear when the presentation rate is fast. Wingfield and Byrnes compared recall when subjects used this strategy with a condition in which subjects had to recall the items in their presentation order, i.e. alternating between ears. They found that this latter method produced poorer recall and further, that if the interval between presentation and recall of an item was measured, it was greater on average when subjects recalled in the order presented than when they recalled "ear by ear". When the probability of recall was plotted as a function of this interval (the time an item was stored) they found that the same simple function fitted both types of recall, i.e. the data points from both conditions lay on the same curve. The experiment is elegant in that the same number of items are presented and recalled in each condition so that differences in recall cannot be ascribed to interference or displacement. They interpreted their results as supporting a decay theory of forgetting. However, a problem with this experiment is that it is not clear what the subjects were doing. For example, if the items are stored in an order such that all the items presented to one ear are placed before all items presented to the other ear, then when asked to recall items in the order they were presented, subjects would have to use fairly complicated strategies to retrieve the appropriate items. The fact that the probability of recall is predicted by the time spent in store could then reflect the difficulty the subjects encountered in trying to retrieve correct items.

Although there is no direct support for decay theory, there are no experiments which contradict it, unlike interference and displacement theory. Further, decay theory is capable of explaining AS effects as demonstrated by Conrad (1965) whose detailed theory will be considered in the discussion section of this thesis. An important consequence of decay theory is that the capacity of STM will be equal to the decay time of an item. The theory assumes that items begin decaying as soon as they are registered in STM and that rehearsal "refreshes" the trace. It follows that the amount that can be recalled (the practical capacity of STM) will be equal to the number of items that can be rehearsed within the decay time. This, in turn, will be a function of the item's length and the subject's rehearsal rate. This is an important prediction as it is contrary to the prediction of the chunking hypothesis which assumes that capacity is a function of the number of "units" of some description. If it is accepted that STM employs a speech code, then these units are likely to be words or some subdivision of words such as syllables or phonemes. For decay theory, the code employed by STM is irrelevant as far as the capacity of the store is concerned; the only determinant of capacity being the decay time of items. It is clear then that chunking theory and decay theory produce different predictions concerning the capacity of STM; the former implying that capacity will be a fixed number of coded units while the latter predicts that it will be a fixed interval of time.

It has been argued above that Miller's experiments produced results supporting the chunking hypothesis because he used tasks in which subjects relied heavily on LTM. However, it remains to be shown that chunking does not occur in STM and the first set of experiments are designed to test this hypothesis and to contrast it with the predictions made by decay theory. The basic rationale is that words will be equated for number of chunks for a given definition of a chunk, but they will be chosen to differ in their spoken length. If chunking theory is correct, recall should be unaffected by this manipulation, while if decay theory is correct, the longer words should take longer to rehearse and hence fewer should be recalled. To anticipate the results, it will be shown that, for various definitions of a chunk, the chunking hypothesis fails in every case while decay theory is strongly supported.

THE CHUNKING HYPOTHESIS

The experiments reported in this chapter are designed to test the hypothesis that the capacity of short-term memory is a constant number of chunks. As stated in the introduction, this hypothesis cannot be tested unless a chunk is defined independently of the experimental results. The approach taken here is based on the idea that STM employs a speech code from which it follows that chunks should be some structural unit of speech such as words, syllables or phonemes. The different experiments test the hypothesis for different definitions of a chunk and in all cases the chunking hypothesis predicts that recall should be constant when measured in chunks. In contrast, decay theory predicts that the time taken to rehearse the items will be the crucial determinant of recall and consequently, that short words should be recalled better than long ones. In all the experiments, the predictions of decay theory and chunking theory will be opposed, i.e. the material will be chosen such that the words in the two conditions contain the same number of chunks but take different amounts of time to rehearse. Thus, chunking theory will always predict that recall will be the same in the two conditions while decay theory will predict that the shorter words should be better recalled.

The first two experiments designed along these lines were performed by Dr. A. D. Baddeley and myself (Baddeley, Thomson and Buchanan, 1975) and only the important details will be mentioned here. In both experiments a chunk was defined as a word and the two conditions comprised words of different lengths. In experiment one, subjects' memory span was estimated by presenting them with eight lists of 4, 5, 6, 7 and 8 words, always in ascending order of list length. In condition one the lists were constructed by sampling at random without replacement from a pool of eight five-syllable (long) words. In condition two the word pool contained eight one-syllable (short) words with the two pools of words matched for word frequency. All the subjects did both conditions, four of them doing the short words first and four the long words first. The lists were read to the subjects at a 1.5 second per word rate and recall was spoken.

The results showed that at each list length all the subjects performed better on the short words than on the long words when scored in terms of the number of lists completely correct. It was inferred that the capacity of STM is not a constant number of chunks when chunks are defined as words, but that the results supported decay theory given that the longer words took longer to rehearse than the short words.

The second experiment was a replication of the first with one important change. In the introduction it was argued that both LTM and STM are involved in serial recall and that LTM is sensitive to the semantic structure of the material. Thus, if for some reason the short words were more meaningful than the long ones, the better performance might be due to an increased LTM contribution. In fact, the long words tended to be more Latinate in origin and, therefore, possibly more complex semantically than the short words. In the second experiment then, an attempt was made to equate the meaningfulness of the two pools of words by using short and long names of countries. The short words were names such as Chad, Chile, Tonga, etc. and the long words were country names such as Czechoslovakia, Somaliland, Venezuela, etc. There were eight country names in each word pool and a serial recall paradigm was used in preference to memory span with list length constant at five words. As before the lists were constructed by sampling from a pool without replacement and each subject did eight lists in each condition. The lists were read to the subjects at a 1.5 second rate and recall was spoken. Eight subjects were tested.

Whether scored by words or sequences correct all the subjects did better with the short words. The means and standard deviations are shown in table 1.

This result together with that of the first experiment allows rejection of the chunking hypothesis where a chunk is defined as a word. Both experiments were designed to produce high STM and low LTM components by using immediate recall and choosing words which, when placed in a list had minimal semantic structure. Thus, it seems reasonable to conclude that STM is sensitive to word length as decay theory would predict. These results, however, do not rule out the possibility that the chunking hypothesis is true for definitions of a chunk which are smaller than the word and the next set of experiments were designed to test this possibility.

TABLE 1

	Short Names		Long Names	
	Mean	SD	Mean	SD
Sequences	4.50	2.00	0.88	1.27
Items	4.17	0.71	2.80	0.24

If it is the case that STM is involved with the production of speech as suggested in the introduction then it would seem plausible that the units of storage are smaller than the word. In particular, both the syllable (Huggins, 1964) and the phoneme (Wickelgren, 1969) have been proposed as candidates for units of analysis and/or production. The first two experiments made no attempt to equate the number of syllables or phonemes in the two conditions and it is possible that this caused the word length effect. The next experiment examines whether STM capacity is a constant number of syllables or, as decay theory would predict, is a function of the time taken to rehearse the material.

EXPERIMENT I

The hypothesis under test is that STM capacity is a constant number of syllables. Thus, if two sets of words are equated for syllable number, no difference in serial recall scores should be expected. In contrast to this view, decay theory predicts that STM capacity is determined by the spoken duration of the words, on the assumption that spoken duration will correlate highly with the time taken to rehearse the words. These predictions can be evaluated by matching two sets of words for syllable number but ensuring that one set is of longer spoken duration than the other.

METHOD

Two pools of disyllabic words, matched for frequency of occurrence in the language were compiled such that one set tended to have a longer duration when spoken normally. The words used are shown in table 2. Beside each word is

TABLE 2

Long Words	Frequency	Duration	Short Words	Frequency	Duration
Friday	40	0.7	Bishop	40	0.275
Coerce	1	0.8	Pectin	1	0.5
Humane	5	0.65	Ember	5	0.45
Harpoon	1	0.75	Wicket	1	0.5
Nitrate	2	0.9	Wiggle	2	0.5
Cyclone	3	0.875	Pewter	3	0.4
Morphine	1	0.75	Tipple	1	0.45
Tycoon	6 per 4 x 10^6	0.725	Hockle	6 per 4 x 10^6	0.45
Voodoo	4 per 18 x 10^6	0.65	Decor	4 per 18 x 10^6	0.5
Zygote	-	0.9	Phallic	-	0.425

shown its frequency and its spoken duration. The latter was measured by recording the words on to magnetic tape and playing this through an oscillograph. This machine plots the waveform of the signal against time and produces the result on light sensitive paper. Thus, the spoken duration of the words can be measured by finding the beginning and end of the word on the waveform trace and measuring the distance between them. The mean duration of the short words is 0.44 seconds and that of the long words 0.77 seconds. Even allowing for differences in pronunciation, it would seem reasonable to assume that the long words took subjects longer to rehearse than the short words.

From each pool of words ten lists of five words were compiled by sampling at random without replacement. These twenty lists were divided into four blocks of five, two blocks containing only short word lists and two containing only long word lists. A Williams latin square was then used to present these four blocks in counterbalanced order to each of twelve subjects as shown in figure 1.

The lists were read to the subjects at a 2 second rate and they 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 output was comparable for long and short words (Conrad & Hille, 1958). Subjects were instructed to begin recalling the material as soon as the list had been presented and to recall the words in order. They were told to say "blank" if they could not recall a word so that subsequently recalled words were not assigned the wrong ordinal position. It was stressed that recall

FIGURE 1

Subjects

1, 5, 9	2, 6, 10	3, 7, 11	4, 8, 12
Block 1	Block 2	Block 3	Block 4
Block 2	Block 4	Block 1	Block 3
Block 3	Block 1	Block 4	Block 2
Block 4	Block 3	Block 2	Block 1

must be serial; they should not recall the last items first even if they could say which position they were in the list. Before the task subjects were allowed to read through the words to familiarize themselves with the material. Twelve undergraduates from the University of Stirling served as subjects and were not paid for their services.

RESULTS

Figure 2 shows the mean number of words correctly recalled as a function of serial position. The results of a two way analysis of variance are shown in table 3. The dependent variable is the number correctly recalled and the independent variables are word length and serial position.

The results show that the short words were better recalled than the long words and that this difference was largest over early serial positions. As no a priori predictions were made concerning the interaction of the word length effect with serial position, it is not possible to perform planned comparisons on this data. Ad hoc tests such as "Tu keys HSD test" would add little to the results as the nature of the interaction is very clear from figure 2.

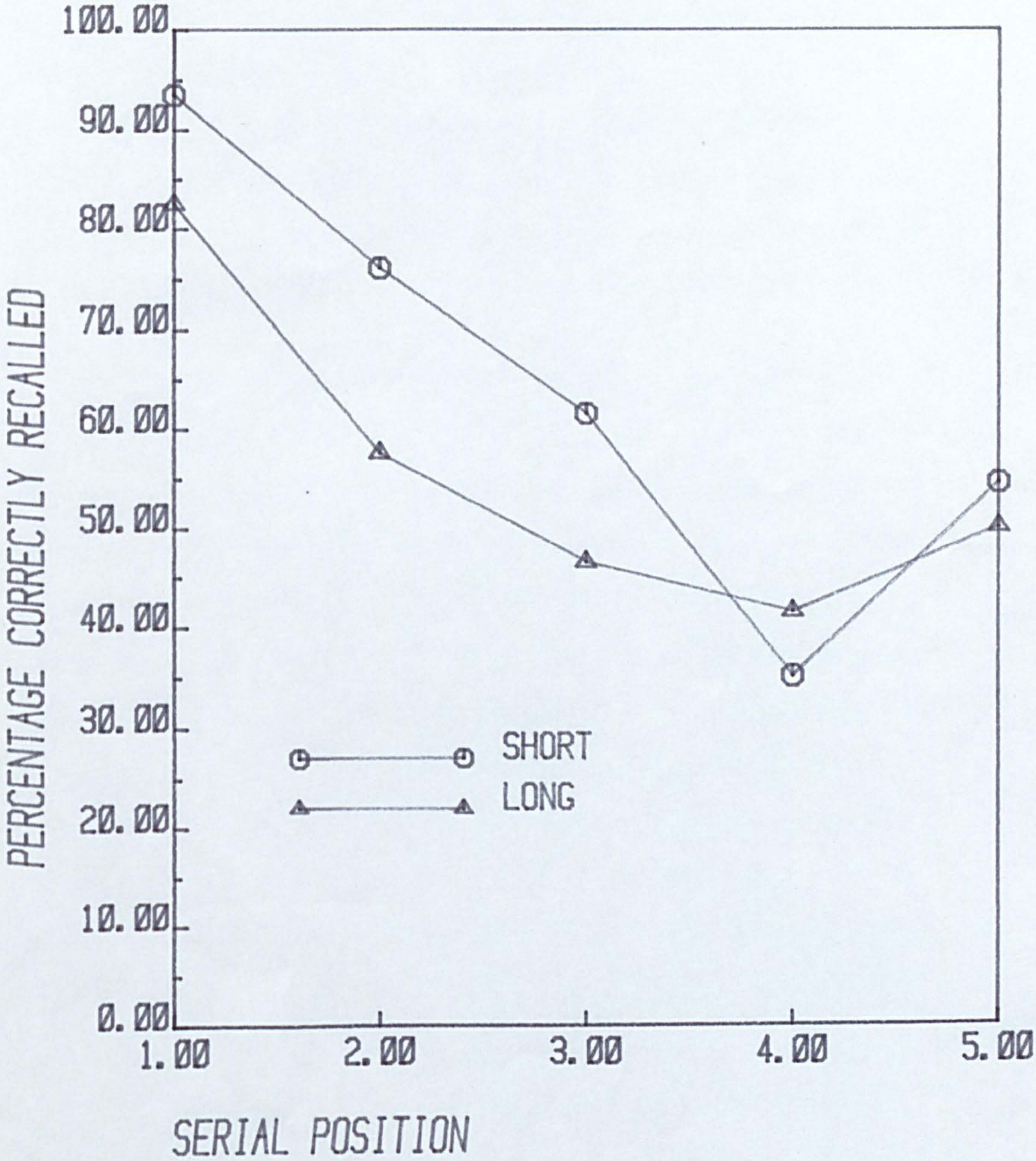
DISCUSSION

The results do not support the hypothesis that STM capacity is a constant number of chunks where a chunk is defined as a syllable. The results would be predicted by decay theory on the other hand and from subjects' informal reports it was clear that the long words were more difficult to remember because they took longer to rehearse; subjects

TABLE 3

Source	SS	df	MS	F	Probability
Word length (WL)	22.53	1	22.53	11.335	<.01
WL x subjects	21.87	11	1.98		
Serial Position (SP)	336.78	4	84.20	36.816	<.001
SP x subjects	104.42	44	2.33		
WL x SP	22.88	4	5.72	3.280	<.05
WL x SP x subjects	76.72	44	1.74		

FIGURE 2
SERIAL POSITION CURVES
FOR EACH WORD LENGTH



frequently stated that they could not rehearse the long words fast enough to remember the whole list.

The interaction between word length and serial position is somewhat puzzling as neither of the theories being considered would predict such a result. One possible explanation is that the paced recall procedure caused the effect. Subjects found it difficult to recall words at a 2 second rate, often claiming that it was too slow. It is possible then that pacing recall at a faster rate would remove this interaction and this prediction is tested in the next experiment.

The final form of Miller's chunking hypothesis to be tested is where a chunk is defined as a phoneme. In the previous experiment the longer words tended to have more phonemes than the shorter words and so it does not rule out the possibility that STM capacity is a constant number of phonemes. The idea that STM employs a phonetic code has been advanced by Wickelgren (1969) and if forgetting in the store is not through decay but rather by displacement or interference, then it would be expected that the store's capacity would be a constant number of phonemes. It is important to notice that decay theory says nothing about the nature of the code employed, only that it is irrelevant as a determinant of the capacity of the store. Thus, none of the experiments reported here rule out the possibility that STM employs lexical, syllabic or phonemic codes, only that the capacity of the store is not a constant number of these structural items.

EXPERIMENT II

Two pools of words were constructed, matched for frequency, syllable number, and phoneme number but differing in spoken duration. As it happened these words were a subset of those used in the previous experiment and are shown in table 4. The phoneme count for each word was made by a linguist for a Scottish pronunciation of the words.

METHOD

Exactly the same design was used as for the previous experiment with the same number of lists blocked in the same manner. As there are only five words in each pool, subjects were always presented with the same words in any one condition but the order of the words always differed between lists. The lists were constructed such that no list had more than two adjacent words in the same order as any other list. For reasons discussed previously, recall was paced at a one second rate in this experiment and, because it is difficult to receive words at one rate and recall them at a different rate, the presentation rate was also increased to one word per second. Eight Scottish undergraduates served as subjects and were not paid for their services.

RESULTS

Figure 3 shows the mean number of words recalled in each condition plotted as a function of serial position. The results of a two way analysis of variance are shown in table 5. The dependent variable is the number correctly recalled and the independent variables are word length and serial position. The results show that there is a significant effect of word length with the short words being better recalled and that this effect does not

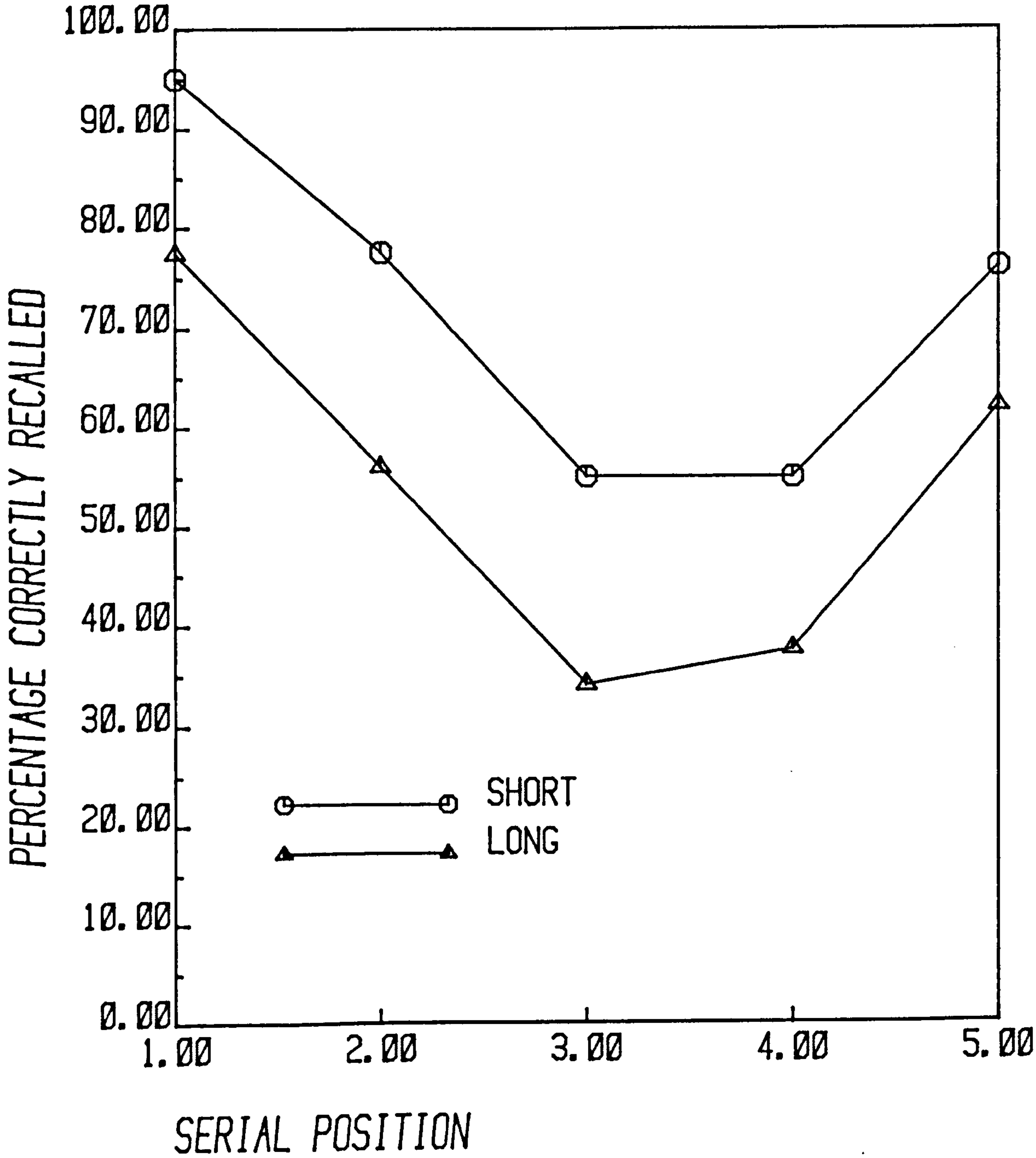
TABLE 4

Long Words	Frequency	Phoneme Number	Duration
Coerce	1	5	0.8 secs
Harpoon	1	6	0.75 secs
Friday	40	5	0.7 secs
Cyclone	3	6	0.875 secs
Zygote	-	5	0.9 secs
Short Words			
Wicket	1	5	0.5 secs
Pectin	1	6	0.6 secs
Bishop	40	5	0.275 secs
Pewter	3	6	0.4 secs
Phallic	-	5	0.425 secs

TABLE 5

Source	SS	df	MS	F	Probability
Word Length (WL)	59.51	1	59.51	18.95	>.001
WL x Subjects	21.99	7	3.141		
Serial Position (SP)	191.08	4	47.77	38.06	>.001
SP x Subjects	35.13	28	1.255		
WL x SP	4.18	4	1.044	0.553	ns
WL x SP x Subjects	52.83	28	1.887		

FIGURE 3
SERIAL POSITION CURVES
FOR EACH WORD LENGTH



interact with serial position.

DISCUSSION

The hypothesis that STM capacity is a constant number of phonemes is not supported but again decay theory is able to explain the results. The disappearence of the interaction of word length and serial position supports the view that it was caused by the subjects having to recall the material at a slow rate. However, this particular issue will not be pursued further as it would result in the research being sidetracked from the main purpose which is to determine the capacity of STM.

It would now seem that the chunking hypothesis has failed for all plausible definitions of a chunk while decay theory has consistently produced correct predictions. Thus, it would now seem more useful to accept decay theory as a working hypothesis and to examine in more detail the relationship between word length and recall. However, it should be noted that the design of the previous experiments does not permit the inference that the number of phonemes or syllables has no effect on recall, only that there is an effect of word length independent of these variables.

The next experiment was designed to enable the relationship between word length and recall to be examined more closely than had been possible with the previous experiments. To achieve this it was felt necessary to have a better measure of word duration. The spoken length of a word will be determined by two factors: the physical properties of the word (e.g. how many syllables it has and how long the vowels are), and secondly, by the way the subject says the word and how quickly he can speak. In the previous experiments the experimenter's pronunciation of a word was used to measure its length but as there is a large amount of variability in speech rate between subjects, although very little within a subject (Goldman-Eisler, 1961), a better estimate of the subject's rehearsal rate would be provided by measuring the subject's pronunciation of the words and this measure is used in the next experiment. The purpose of the experiment is to provide a "first look" at the relationship between rehearsal rate and recall before embarking on a large scale experiment.

EXPERIMENT III

Apart from some details mentioned below, the experiment was essentially the same design as experiment 2. The two important changes are the different measure of word length and the use of visual instead of auditory presentation. The reasons for using a different measure of word length have already been discussed but the use of visually presented material requires explanation. In the previous experiments the lists had been read to the subjects and it is possible that some aspect of the experimenter's presentation could have affected the results. Presenting the material visually removes this possibility but in some pilot experiments using this technique the word length effect disappeared. On interrogating the subjects it was found that those who showed the word length effect had used a rehearsal strategy while those who showed no effect had used imagery strategies. In a second pilot experiment it was found that the word length effect re-appeared when subjects were instructed to remember the items by rehearsing them. It is clear that imagery and rehearsal strategies would involve different processes and possibly different systems and as rehearsal is widely assumed to be a process employed by STM (e.g. Atkinson & Shiffrin, 1968; Waugh & Norman, 1965) and the prediction that word length would determine the amount recalled was made on the assumption that rehearsal is used to prevent decay, it would seem both necessary and reasonable to instruct subjects to use a rehearsal strategy and this was done in this experiment.

METHOD

The same material and design as experiment 2 was used. The words were typed on cards which were presented one at a time to subjects at a 2 second rate. Each card was put down on top of the previous one so that subjects could only see one word at a time. For the last list item a blank card was placed over it two seconds after it had been presented. The subjects were instructed to remember the material by "saying it over to themselves". They were reminded of this instruction approximately half way through the experiment. Recall was again spoken but was unpaced as it was felt that it was not clear what subjects did during paced recall and that if allowed to recall at their own rate the relationship between word length and recall might be clearer.

Rehearsal rate was estimated by timing the subjects' pronunciation of the words and this was done in two slightly different ways. In the first of these subjects were timed for reading aloud as fast as possible the ten five word lists in each condition. The lists were typed as two continuous columns of twenty-five words in upper-case on a sheet of paper. The subjects were instructed to read the words as fast as possible consistent with a correct pronunciation of each word. Subjects read each set of words four times, each reading being timed with a stop-watch. The mean of the four times was then calculated for each condition. Because adults can read faster than they can speak, the task of reading the material should not directly slow down their speech rate. However, reading could well interfere with speech in an indirect manner and for this reason a second method of estimating rehearsal rate was employed. This consisted of measuring the time subjects took to repeat ten times over three of the words from a particular pool. The subjects were given three words and after they had repeated them in sequence a few times were instructed to repeat them continuously as fast as possible consistent with correct pronunciation. They did this four times for each set of words, always with a different set of three words. Again the four times were averaged for each subject in each condition.

Both measures of rehearsal rate were taken after subjects had completed the serial recall task, half the subjects doing the "reading rate" test first and half the "speech rate" test first. The subjects were eight members of the Applied Psychology Unit (APU) subject panel and were paid for their services.

RESULTS

A two way analysis of variance was performed on the data with number recalled as the dependent variable and word length and serial position as the independent variables. There was a significant effect of word length, $F = 15.14$, $df(1,7)$, $p<.01$; of serial position, $F = 14.79$, $df(4,28)$, $p<.001$; but no interaction between word length and serial position, $F = 2.43$, $df(4,28)$, $p>.05$.

DISCUSSION

Before discussing the main results it is worth noting that the interaction of word length and serial position is again absent and would seem to confirm the idea that the results of experiment 1 were due to pacing recall at a slow rate.

Assuming that there is very little LTM component to recall, decay theory predicts that subjects will recall as much as they can rehearse in a fixed time interval, viz. the decay time of an item. This would imply that the ratio of number recalled (NR) to rehearsal rate (RsR) should be constant. This prediction can be best understood by using an example. If items decay in two seconds then the subject will be able to remember as much as he can rehearse within this interval. Thus if he can rehearse 2 words per second, he will be able to rehearse and so recall 4 words in two seconds. If the words were shorter and he could rehearse at three words per second, he would be able to rehearse 6 words within the decay time of an item. The amount recalled is equal to RsR multiplied by the decay time. This may be expressed as the equation:-

$$\text{Decay Time} \times \text{RsR} = \text{NR}$$

and thus: $\text{NR}/\text{RsR} = k$ where k is a constant equal to the decay time of an item

Substituting the measures of rehearsal rate that were used, viz. reading rate (RR) and speech rate (SR) both in units of words/sec. produces:

$$NR/RR = k_1$$

and $$NR/SR = k_2$$

The constants "k_1" and "k_2" will differ from the true decay time of an item by an amount proportional to the degree to which RR and SR differ from RsR. Assuming decay time to be the same for different lengths of words it can now be seen that the value of NR/RR should be the same for both word length conditions and similarly for the ratio NR/SR. These two ratios were calculated for each subject and are shown in table 6. A Wilcoxon matched pairs test revealed that there was no effect of word length for either the NR/RR ratio, $T = 9$, $N = 8$, $p > .05$, or for the NR/SR ratio, $T = 10$, $N = 7$, $p > .05$. Thus the results show that the subjects were recalling as much as they could read in 1.6 seconds or articulate in 1.3 seconds. This constitutes strong support for decay theory and suggests that the capacity of STM is determined by the decay time of items rather than by the number of any structural units such as phonemes, syllables, etc. It would now seem appropriate to design a large scale experiment to confirm this result for a wider range of word lengths and the next experiment to be reported was an attempt to do this. Before introducing the experiment two points need further discussion; the first is the question of the relative merits of the different measures of rehearsal rate, and the second the question of the effect on the results of an LTM component to recall.

TABLE 6

Subject	NR/RR (short words)	NR/RR (long words)
1	1.78	1.70
2	1.72	1.42
3	1.55	1.80
4	1.43	1.83
5	1.30	1.46
6	1.68	1.95
7	1.38	1.59
8	2.15	1.63
Mean	1.62	1.67

Subject	NR/SR(short words)	NR/SR (long words)
1	1.48	1.43
2	0.93	0.93
3	1.15	1.14
4	1.34	1.48
5	1.14	1.32
6	1.40	1.79
7	1.24	1.00
8	1.95	1.36
Mean	1.33	1.31

Some possible objections to using RR as an estimate of rehearsal rate have already been discussed, the main one being that the task of reading may interfere with and, therefore, slow down the rate of articulation. This is not too serious an objection so long as the interference is similar for different word lengths and as the words were matched for frequency, this seems a plausible assumption. While SR was felt initially to be a somewhat "purer" measure, it was found in practice that it was less easy to control what the subjects were doing. In particular they tended to pronounce the words very unclearly, in spite of the instructions, and it was often difficult to be sure that they had said the word at all. Further, because of time constraints, only a small subset of the word sequences could be used and it was very noticeable that some sequences could be said much faster than others. The RR measure had the advantage of using all the words in the pools a number of times and subjects were much more careful in their pronunciation with this technique. Finally, both measures produced the predicted result and so, given the greater control possible with the RR measure, this one was used in the next experiment.

The previous experiment attempted to minimize the contribution of LTM as has already been discussed, but there is no independent evidence that the techniques used were successful. There is good reason to assume that LTM is not sensitive to word length in that Craik (1968) showed that word length did not affect free recall performance and, assuming that the secondary memory component of free recall is the same as the LTM component of serial recall, it may be inferred that word length does not affect LTM. Thus, any contribution from LTM should, all things being equal, add the same amount to recall of both long and short words. Using RR as the measure of rehearsal rate, the equation would then become:

$$NR = k_1.RR + c$$: where "c" is the LTM component in words.

This is the equation of a straight line with a non-zero intercept where the slope of the line is the decay time of an item and the intercept is the LTM contribution. This line should result if NR is plotted against RR for different word lengths and this prediction is tested in the next experiment. In order to increase the LTM contribution, the rehearsal instruction was omitted and this should cause the line to have a positive intercept value.

EXPERIMENT IV

The aim of this experiment is to enable the relationship between word length and recall to be examined in detail. RR will be used as the measure of word length and will be plotted against NR. If the above arguments are correct this should produce a straight line with a positive intercept value.

In order to plot the NR vs RR function with some accuracy, five different word lengths were used. With this range of word length it is not possible to control for number of syllables and phonemes so these were allowed to covary with word length. This does not alter the prediction as decay theory assumes that it is only word length that is important and not the structural characteristics of the words. If there is any effect of syllable number etc. then this could be expected to distort the predicted relationship and thus will not increase the chance of wrongly rejecting the null hypothesis. For each word length a pool of ten words was compiled, matched for word frequency and chosen such that each pool contained one instance from each of ten semantic categories. The words are shown in table 7. The reason for equating the words in the different pools for the semantic category from which they are derived is to try to ensure that the LTM component is similar for each word length.

The experiment used a within subjects design with all subjects tested on all the five different word lengths. For the purposes of analysis the experiment has one condition (word length) with five levels. It was decided not to block lists of one word length together as was done for the previous experiments but rather to present all the lists in a random order. It was felt that if the lists were blocked subjects might find strategies specific to a particular word length and hence confound comparisons between them. For each word length ten lists were constructed making fifty experimental lists in total.

METHOD

From each pool of words ten five-word lists were constructed by sampling from the pool at random without replacement. The words were then written on the back of IBM computer cards and placed in an IBM card puncher. This machine holds a stack of cards and, on pressing a button, will push a card into a window through which the word could

TABLE 7

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

be clearly seen. On pushing the button again the next card in the stack is placed in the window and thus, if the button is pushed in time with a metronome, a word appears in the window at a regular interval. The fifty lists were placed in random order and before the first word of each list was placed a card with three asterisks on it. After the last card in the list there was placed a blank card. These acted as a warning stimulus for the list and a mask for the last list item respectively. A television camera connected to a VTR was then focused on the card punch window and the words were changed at a 2 second rate. Thus a video recording was made of the window display. This recording was played to a 12 inch monitor to present the material to the subjects. At the end of each list the VTR was halted for about twelve seconds to allow subjects to recall. Recall was spoken and subjects were instructed as for the previous experiments, viz. to recall the words in order and to say "blank" if they could not remember a particular word. Subjects were given a short rest half way through the lists and no subjects found the task too tiring.

Reading rate was measured by requiring the subject to read lists of fifty words, each list being composed of five occurrences of each word from one of the pools. The words were typed in a completely random order in upper case letters in two columns on a sheet of A4 paper. Subjects were instructed to read the words aloud as quickly as possible consistent with a correct pronunciation of each word. Their reading times were measured with a stop-watch. Before the memory task they read five of these lists, one for each word length, twice through and each reading was timed. This procedure was then repeated after the memory task resulting in four measures for each word length. The reading task was split up in this manner as subjects found it tiring to do all the reading in a continuous session and their reading rates slowed considerably if they became tired.

Fourteen subjects were tested, all members of the APU subject panel and they were paid for their services.

RESULTS

The mean memory scores and mean reading rates in each condition are shown in figure 4. Separate analyses of variance were performed on these measures and the results are shown in table 8.

Thus the main effects of word length on memory performance and reading rate were significant and the word length effect has been replicated. The next set of analyses were concerned with the nature of the relationship between NR and RR and in particular, to determine whether the predicted straight line relationship exists. NR is plotted as a function of RR in figure 5 and it can be seen that the points fall on a straight line. Regression analysis showed that the slope of the best fitting line was 1.87 seconds, the intercept on the ordinate was 0.17 words and the standard error of the estimate is 0.10. The value of the intercept was shown by a "t" test to differ significantly from zero, $t = 3.71$, df(3), $p<.05$. It may be concluded, therefore, that the results are very well described by the equation $NR = RR.k_1 + c$. It seemed desirable, however, to check that individual subjects were

FIGURE 4
PERCENTAGE CORRECTLY RECALLED
AND READING RATE FOR EACH WORD LENGTH

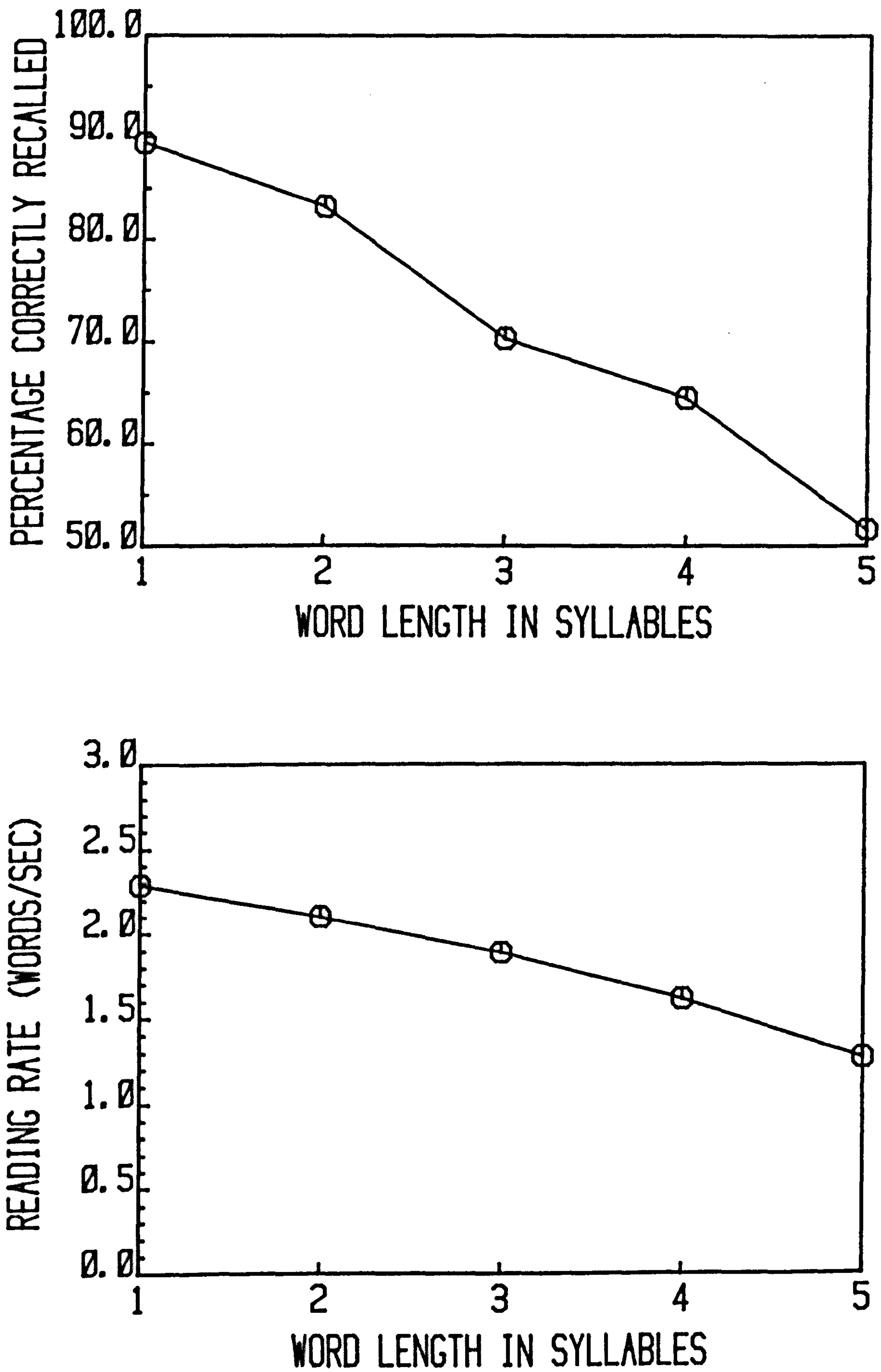


TABLE 8

Analysis of Memory Scores

Source	SS	df	MS	F	Probability
Word Length	3179.812	4	794.953	37.695	<.001
Subjects	3245.875	13	249.683	11.840	
Residual	1096.625	52	21.089		

Analysis of Reading Rate Scores

Source	SS	df	MS	F	Probability
Word Length	2823.199	4	705.800	244.023	<.001
Subjects	1191.672	13	91.667	31.693	
Residual	150.402	52	2.892		

behaving similarly to the averaged data and the regression lines for each subject were calculated and plotted in figure 6. Additionally, the raw data for this experiment has been included and are listed in appendix 1. The individual regression lines show a fair degree of spread but there is no sign of there being separable groups of subjects or of any subjects showing very different types of relationship between NR and RR.

The fact that some subjects show negative intercept values might seem disturbing as it should not be possible to have a negative LTM component. However, consider the range of values of word length over which the equation NR = RR.k + c could be expected to be true. If words took longer to say than the decay time of an item in STM then the equation would break down. For example if both long and short word conditions contained items that were all of longer spoken duration than the decay time of an item, a word length effect would not be found as, for both conditions, a single word would be too long to be stored in STM, and any recall would have to be from LTM. Consequently, the equation can only be true for words whose duration is less than or equal to the decay time of an item. As has been argued, the slope is a measure of this decay time and thus an "ordinate" should be drawn at a point corresponding to the length of a word equal to the slope. To the left of this ordinate will be words whose spoken duration is longer than the decay time of an item and, thus, the curve of the function should not be extrapolated beyond this point. This has to be done for each subject and it was found that no subject's data crosses the "ordinate" below zero; the worst case (the regression line furthest to the right) has a slope of 1.236 secs. and if the ordinate is drawn at this point, the regression line would intercept it above zero. This is important as, if some subjects had shown "true" negative intercepts, then the idea that the intercept was a measure of the LTM component would have been discredited.

The next analysis was performed on the memory score data and used analysis of variance with both word length and serial position as the independent variables. The

FIGURE 5
MEAN NUMBER OF WORDS RECALLED
VS READING RATE

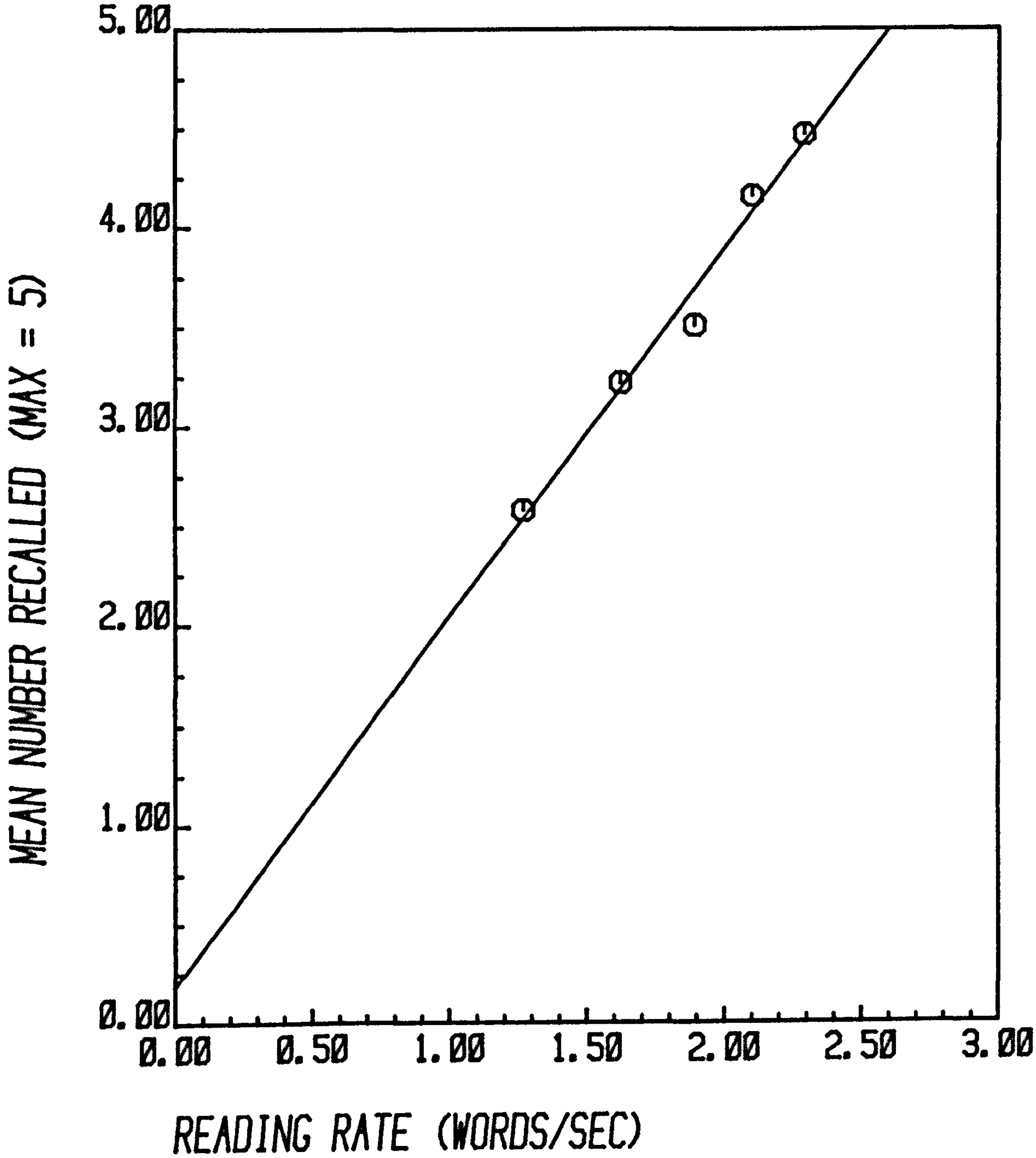
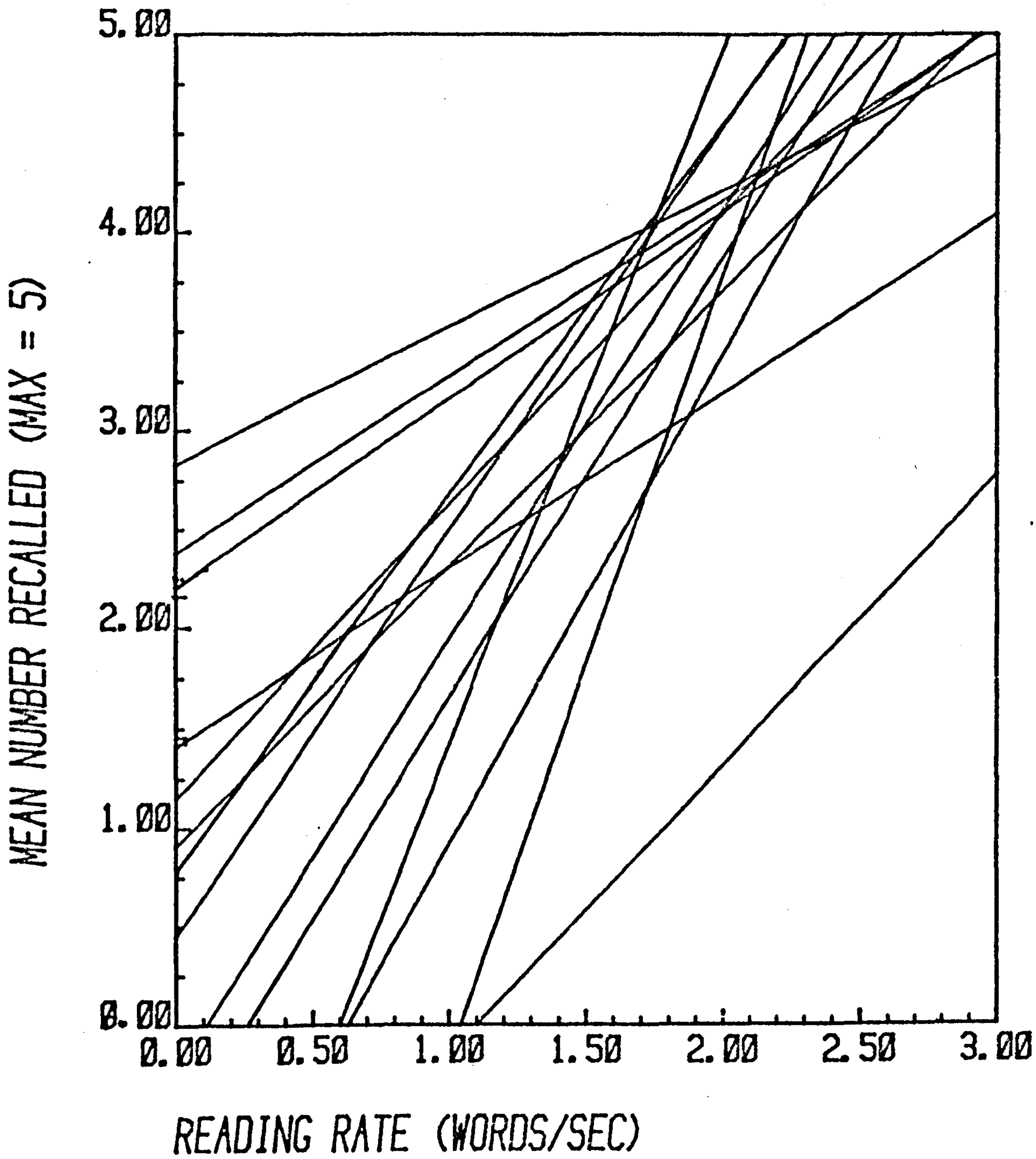


FIGURE 6
REGRESSION LINES
FOR INDIVIDUAL SUBJECTS



serial position curves are shown in figure 7 and the analysis of variance results in table 9.

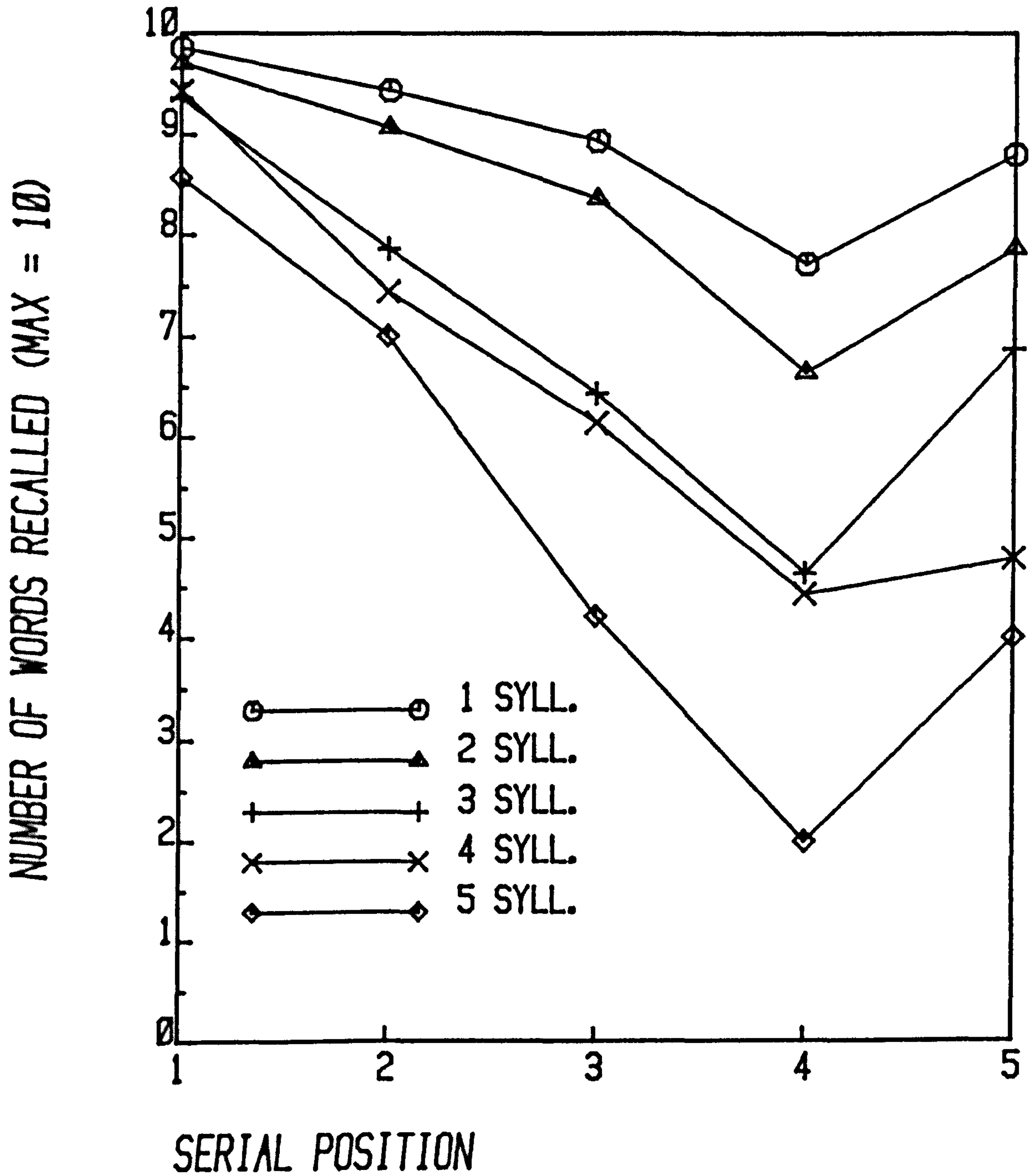
The very slight difference in the F ratio for this analysis from that of the previous analysis of the memory scores (37.710 and 37.695) is due to different analysis programmes being used, the second of which used single precision arithmetic while the first programme used double precision. Clearly the word length x serial position interaction has returned, but it is very unlike the interactions obtained in the previous experiments. In experiment 1 the word length effect was maximal at early serial positions and almost absent over the last two serial positions while, in this experiment, the smallest word length effect is to be found at the early serial positions. This is almost certainly due to a ceiling effect on the first list item; for the shortest list length eleven subjects scored the maximum possible on the first item (10, as there were ten lists for each word length) and the remaining 3 subjects only missed one item each. Thus, little can be made of this interaction and it will not be discussed further.

The final set of analyses were purely ad-hoc. Given that within subjects, reading rate predicts serial recall, it seemed interesting to ask whether this relationship held across subjects. In other words, are fast readers (or fast rehearsers) better at serial recall than slow readers? This might be the case if the decay time of an item was constant between subjects but there is no a priori reason for believing this to be so. However, analysis of the data showed a strong relationship across subjects between

<u>TABLE 9</u>

Source	SS	df	MS	F	Probability
Subjects	616.734	13			
Word Length (WL)	650.417	4	162.604	37.710	<.001
WL x Subjects	224.223	52	4.312		
Serial Position (SP)	794.246	4	198.561	46.261	<.001
SP x Subjects	223.194	52	4.292		
WL x SP	159.526	16	9.970	6.898	<.001
WL x SP x Subjects	300.634	208	1.445		

FIGURE 7
SERIAL POSITION CURVES FOR
EACH WORD LENGTH

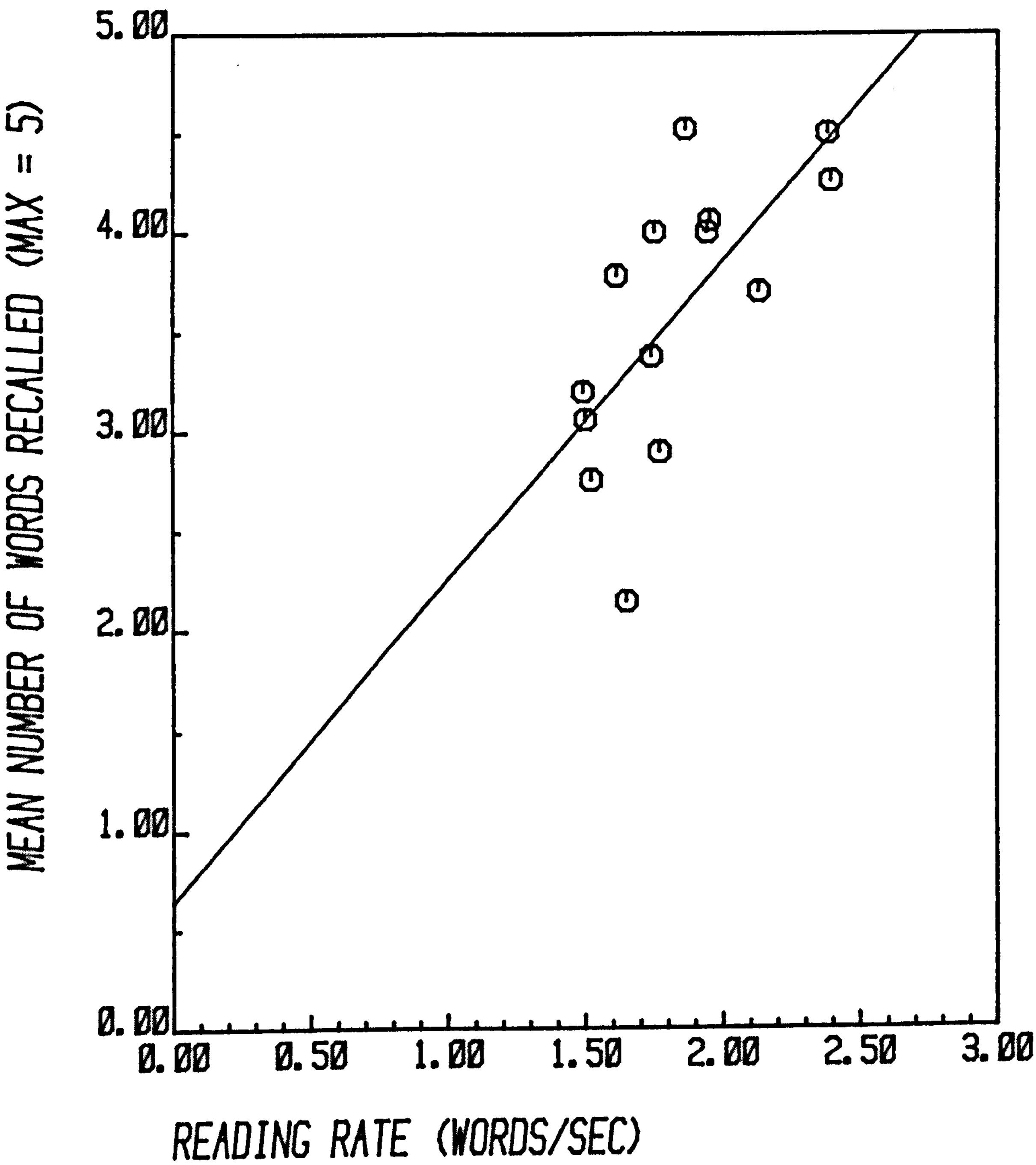


memory scores and reading rate. The mean reading rate and memory score was averaged for each subject across the five word length conditions and is plotted in figure 8. The correlation between the two variables across subjects was significant, $r = 0.69$, $df(13)$, $p<.005$. A regression analysis showed that the best fitting straight line had a slope of 1.632 seconds, an intercept value of 0.56 words and the standard error of the estimate was 0.56. This result is interesting as it suggests that decay time is fairly constant between subjects and adds support to the idea that decay might be mediated by neural events (Hebb, 1961). Further, it suggests that the physical capacity of people's STM does not differ much and the differences in actual capacity are due to differences in rehearsal rate. In so far as STM is involved in a number of cognitive tasks (Atkinson & Shiffrin, 1968; Baddeley & Hitch, 1974) it might be possible to improve performance on these tasks by increasing speech rate. Further, as digit span forms part of the most commonly used IQ tests (e.g. the WISC), it might be possible to improve children's performance on this test by concentrating on their fluency of speech rather than by attempting to increase their powers of reasoning. This must remain speculation at present but the results suggest some interesting lines of further research.

DISCUSSION

The main result of this experiment is the success of the prediction made by decay theory:- that the equation $NR = k_1.RR + c$ would describe performance. This result provides strong support for decay theory and suggests that

FIGURE 8
SCATTER GRAPH AND REGRESSION LINE
FOR SUBJECTS AVERAGED OVER CONDITIONS



the capacity of STM is constant in units of time rather than units of structure. The weakest point in the explanation is the idea that the intercept reflects the LTM component as no independent evidence of this assertion has been provided. In fact, to be very critical, there is no evidence that the slope of the function is a measure of the decay time in STM although it is rather implausible to suppose that LTM has such a small capacity. The next experiments are designed to supply this evidence. To date, the experiments have been designed so that subjects rely very heavily on STM while the LTM component is minimal. Even the last experiment only produced a very small intercept value and it is necessary to show that manipulations that are known to influence LTM change the intercept, while those that affect STM change the slope. Further discussion of the results will be postponed until after these experiments have been reported.

THE WORD LENGTH EFFECT

The purpose of this set of experiments is to determine the nature of the processes underlying the function $NR = RR.k_1 + c$. The working hypothesis is that the slope of the function is a measure of the decay time in STM while the intercept on the ordinate is the LTM component. In the last experiment of the previous chapter it was shown that both slope and intercept were significant quantities; however, the intercept value was extremely small and possibly of little theoretical importance. If the working hypothesis is correct then it should be possible to increase the size of the intercept by changing the task so that the subject relies on LTM to a greater extent. The most obvious way of increasing the LTM component is to use a 'multi-trial' procedure in which the subject is presented with the same list for a number of trials with recall measured after each presentation. Hebb (1961) showed that repeated presentations of a list, even when other lists were interspersed, produced increases in recall. To explain this result in a way that is consistent with the arguments presented in the introduction, it must be assumed that the learning occurs in LTM as only this store is capable of holding the information for any length of time without continual rehearsal. Thus, in the next experiment a multi-trial procedure is used and it is predicted that over trials the LTM component (the intercept) will increase while the STM component (the slope) will remain constant.

A second aim of the next experiment was to test further the idea that decay rate is constant between subjects and hence that fast rehearsers are better at serial recall than slow rehearsers. As mentioned in the results of the previous experiment, this would be an important discovery as it would imply that decay is mediated by some rather basic neural processes. The most severe test of the hypothesis is to make word length a between subjects variable and determine whether the function $NR = RR.k_1 + c$ still holds. This was done in the next experiment where each subject is tested with only one word length.

EXPERIMENT V

Three levels of word length (1, 3 and 5 syllables) were combined with four levels of recall (after 1, 2, 3 and 4 presentations of the list) in a 3 x 4 split plot design. Necessarily the four levels of recall were a within subjects variable, but for reasons given above word length was made a between subjects variable. Each subject did ten lists of one word length with each list being presented four times consecutively. Recall was measured after each presentation. With word length as a between subjects variable the amount of error variance will be higher than for a within subjects design and, consequently, 45 subjects were tested, 15 in each word length.

METHOD

Three pools of sixteen words were constructed, matched for frequency and for semantic category, but differing in word length. The words are shown in table 10.

The lists, all eight words in length, were constructed by sampling at random without replacement eight times from a pool. Ten lists of each word length were compiled. For each word length the lists were placed in a random order

TABLE 10

One Syllable	Three Syllable	Five Syllable
Zinc	Calcium	Aluminium
Maine	Wyoming	Louisiana
Mumps	Leprosy	Tuberculosis
Gas	Parafin	Electricity
Fraud	Forgery	Impersonation
Judge	Alderman	Representative
Scroll	Bulletin	Periodical
Maths	Botany	Physiology
Monk	Methodist	Presbyterian
School	Nursery	University
Switch	Radio	Refrigerator
Stoat	Gorilla	Hippopotamus
Greece	Mexico	Yugoslavia
Crewe	Exeter	Weston-super-mare
Test	Audition	Examination
Rome	Amsterdam	Constantinople

and each subject was presented the lists in the same order. Subjects were first shown the three word pools and allowed to familiarize themselves with the words. The first list was then read to the subject at a 1.5 second rate and as soon as the list was finished the subject attempted to recall the words in the order presented. Subjects were instructed, as for previous experiments, to say 'blank' if they could not recall a particular word. This procedure was then repeated with the same list until it had been presented and recalled four times. There was no pause between list presentations; subjects were presented with the list as soon as they had finished recalling the previous presentation. A short interval was left between different lists and after five lists subjects were given a short break.

After this phase of the experiment subjects were given a copy of the lists (80 words in all) and asked to read them aloud as fast as possible consistent with the correct pronunciation of each word. This was repeated four times, each reading being timed with a stopwatch.

Forty-five subjects were tested, fifteen in each word length. They were assigned randomly to a particular word length with the constraint that no more than fifteen subjects were run in each word length. The subjects were members of the APU subject panel and were paid for their services.

RESULTS

An analysis of variance was performed on the raw memory scores with number recalled as the dependent variable and word length, presentation trials and serial position as the independent variables. The results are shown in table 11.

Thus the word length effect has been replicated for a between subjects design and the lack of interaction with trials implies that the slope of the NR vs RR function does not change over trials. This function is plotted in figure 9 where it can be seen that the points are well fitted by straight lines. A regression analysis done separately for each trial produced the results shown in table 12 and the straight lines drawn in figure 9 are derived from this analysis.

The main points to be noted are that the intercept increases over trials as predicted while the slope remains relatively constant. Further, it should be noted that the value of the slopes is considerably less than that obtained in the previous experiment. These results will be considered further in the discussion, but first the serial position curves will be considered. The significance of the three way interaction between word length, serial position and trials requires that all the serial position curves be illustrated for the purposes of comparison. Figures 10, 11, 12 and 13 show the set of three curves for each word length over trials one to four respectively. Figures 14, 15 and 16 show the set of four curves for each trial over word lengths one to three respectively. Both sets of graphs have been included for the sake of completion and to enable the reader to ensure that no possible explanation of the interaction has been overlooked.

TABLE 11

Source	df	SS	MS	F	Probability
Subjects (S)	44	3940.352			
Word length (WL)	2	767.239	383.620	5.078	<0.02
WL x S	42	3173.122	75.550		
Trials (T)	3	3146.369	1048.790	188.283	<0.001
T x WL	6	32.993	5.489	0.985	n.s.
T x WL x S	126	701.854	5.570		
Serial Position (SP)	7	4649.366	644.195	136.264	<0.001
SP x WL	14	123.861	8.847	1.815	<0.05
SP x WL x S	294	1433.054	4.874		
SP x T	21	293.603	13.981	11.648	<0.001
SP x T x WL	42	75.345	1.794	1.495	<0.05
SP x T x WL X S	882	1058.646	1.200		

TABLE 12

Trial	Slope	Intercept	Standard Error of Reg Coef
1	0.711	1.263	.139
2	1.161	1.720	.171
3	1.343	2.369	.073
4	1.111	3.625	.039

FIGURE 9
NUMBER RECALLED VS READING RATE
FOR EACH TRIAL

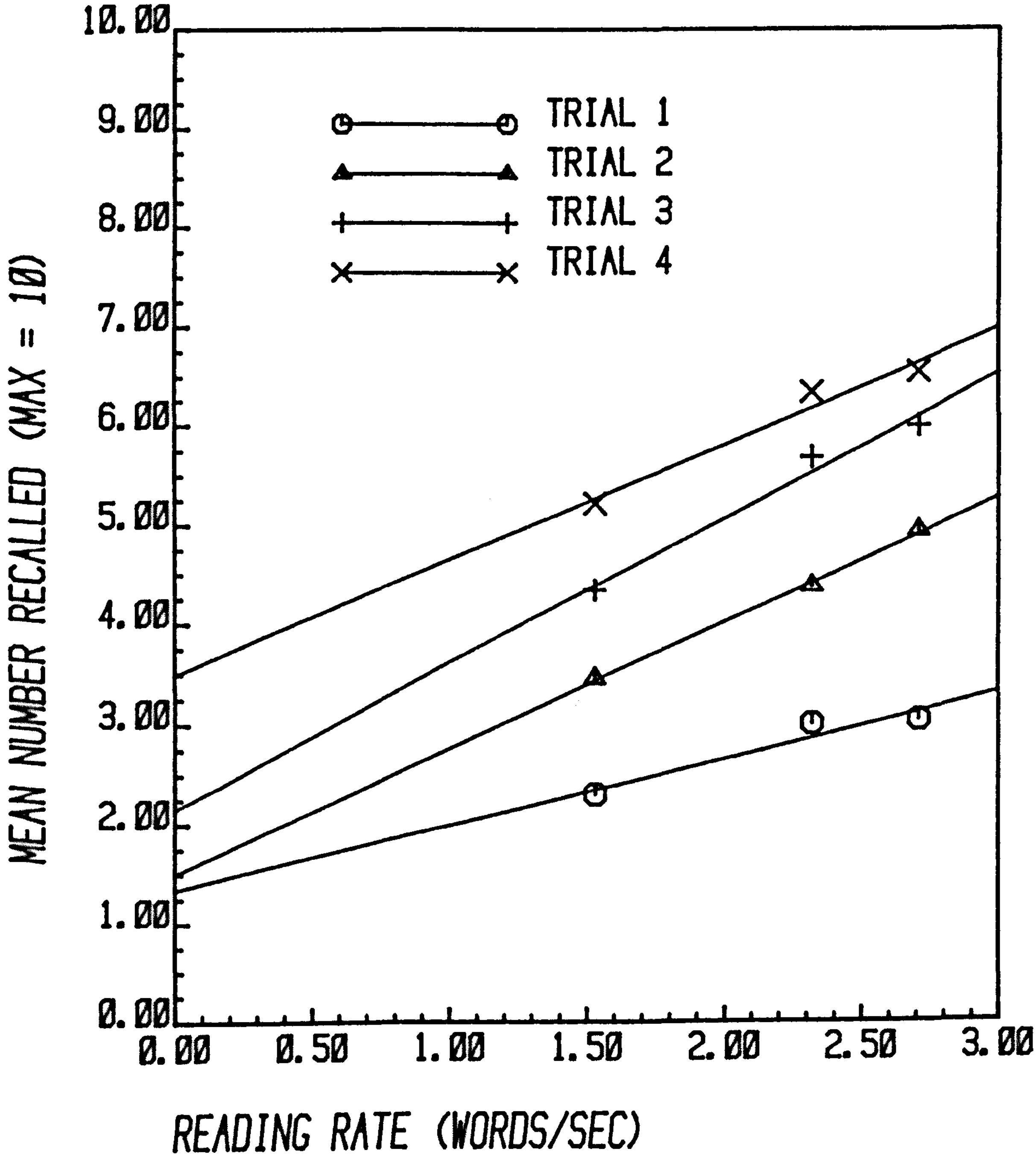


FIGURE 10
SERIAL POSITION CURVES FOR
EACH WORD LENGTH ON TRIAL 1

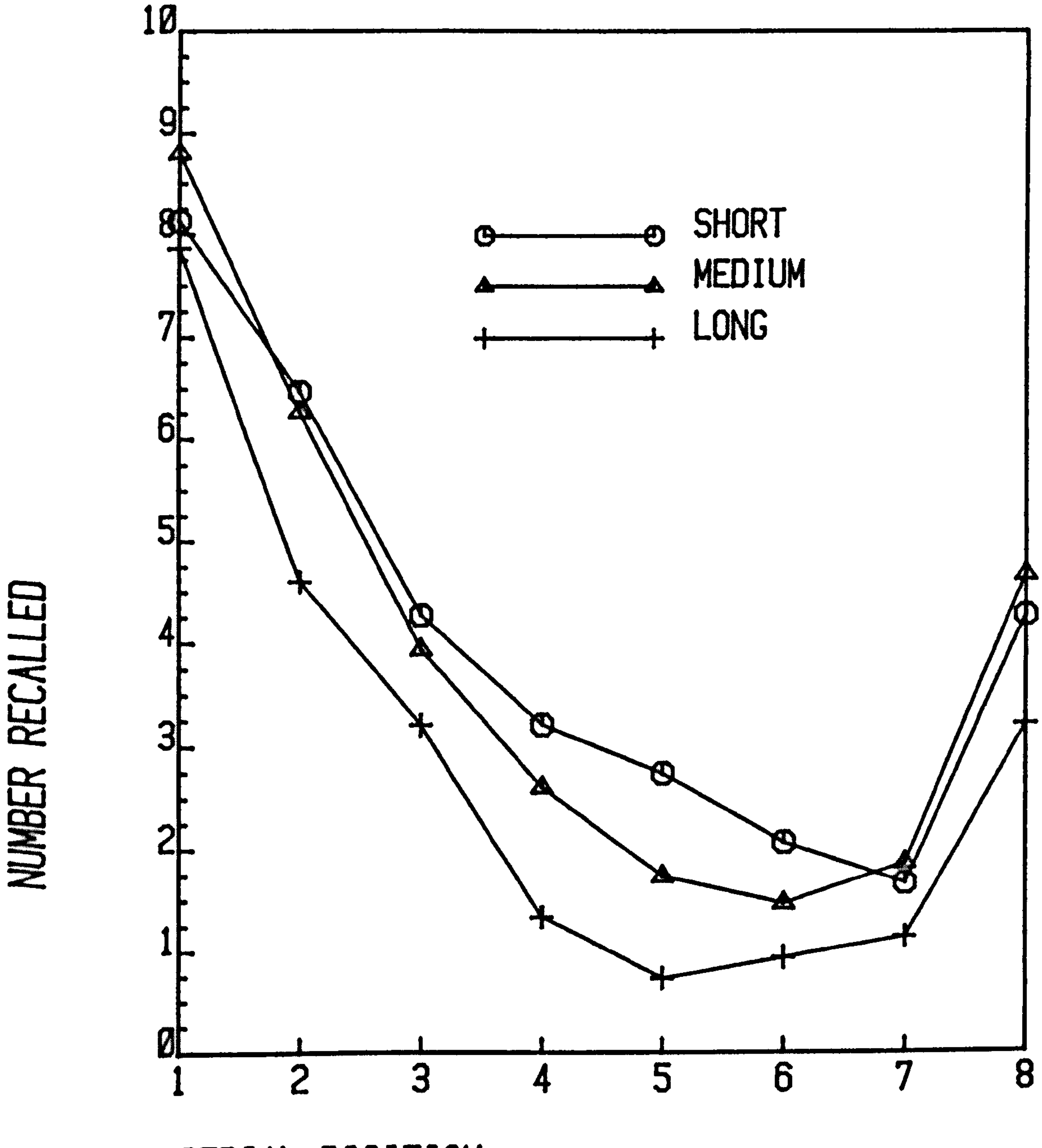


FIGURE 11
SERIAL POSITION CURVES FOR
EACH WORD LENGTH ON TRIAL 2

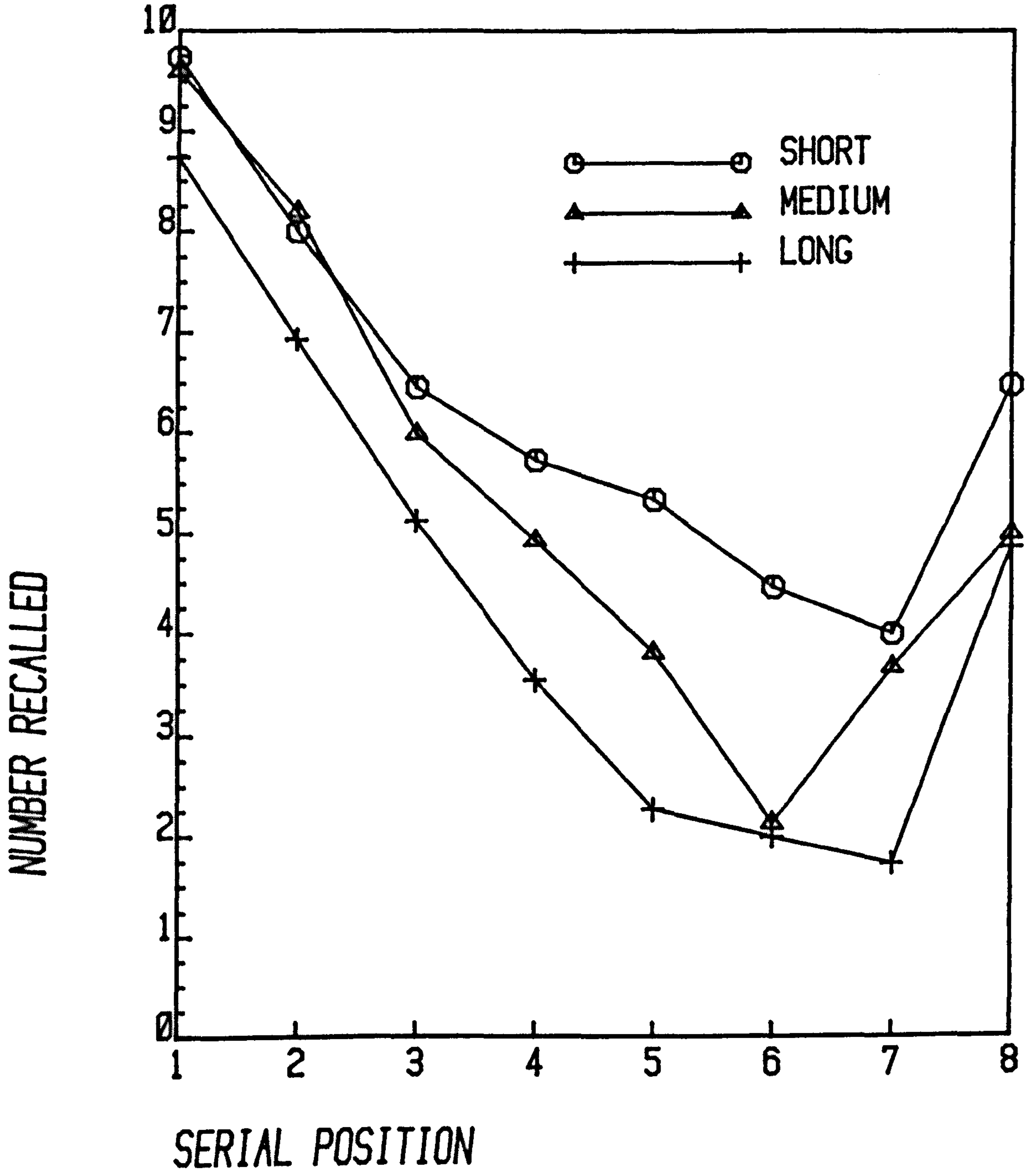


FIGURE 12
SERIAL POSITION CURVES FOR
EACH WORD LENGTH ON TRIAL 3

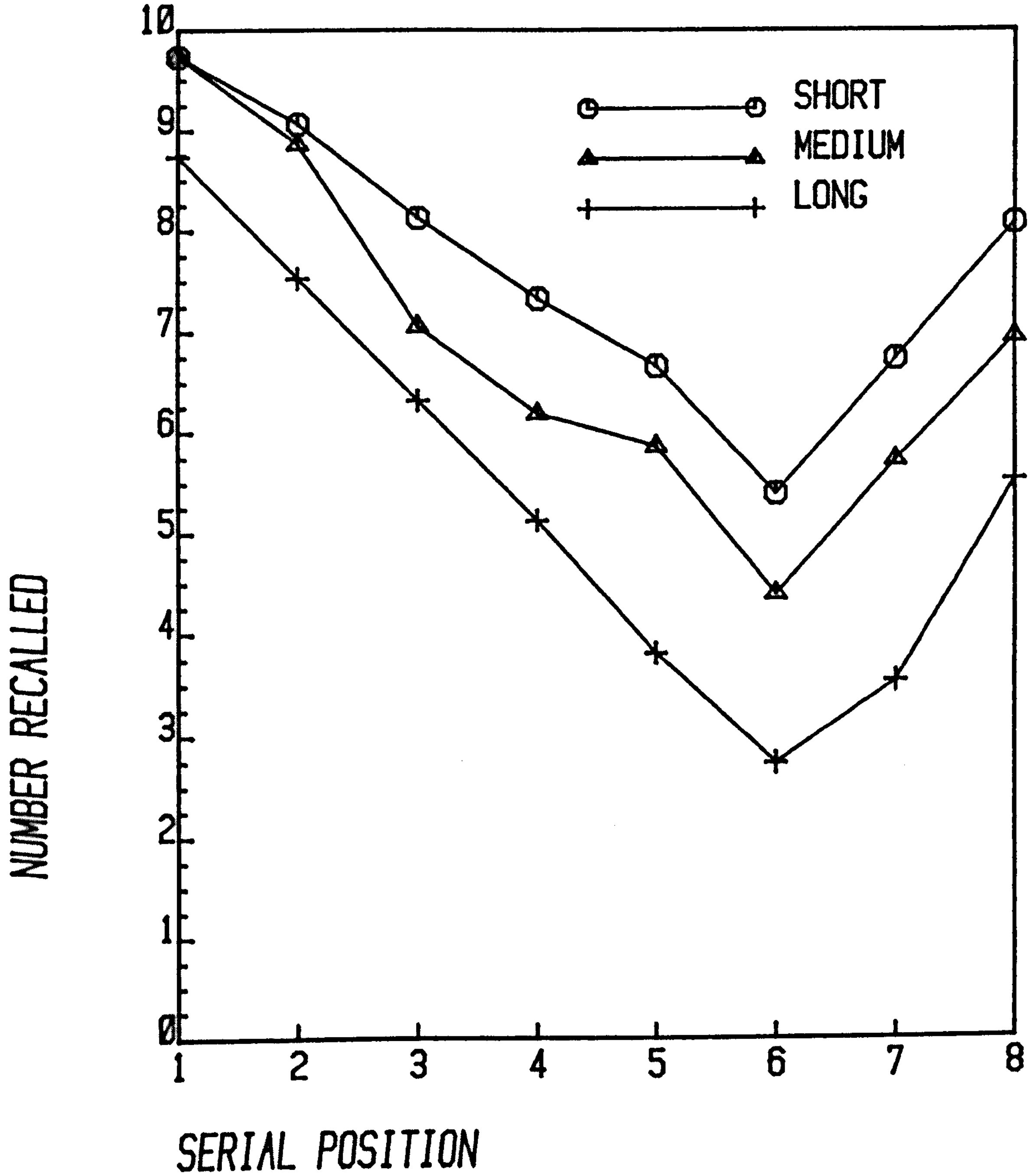


FIGURE 13
SERIAL POSITION CURVES FOR
EACH WORD LENGTH ON TRIAL 4

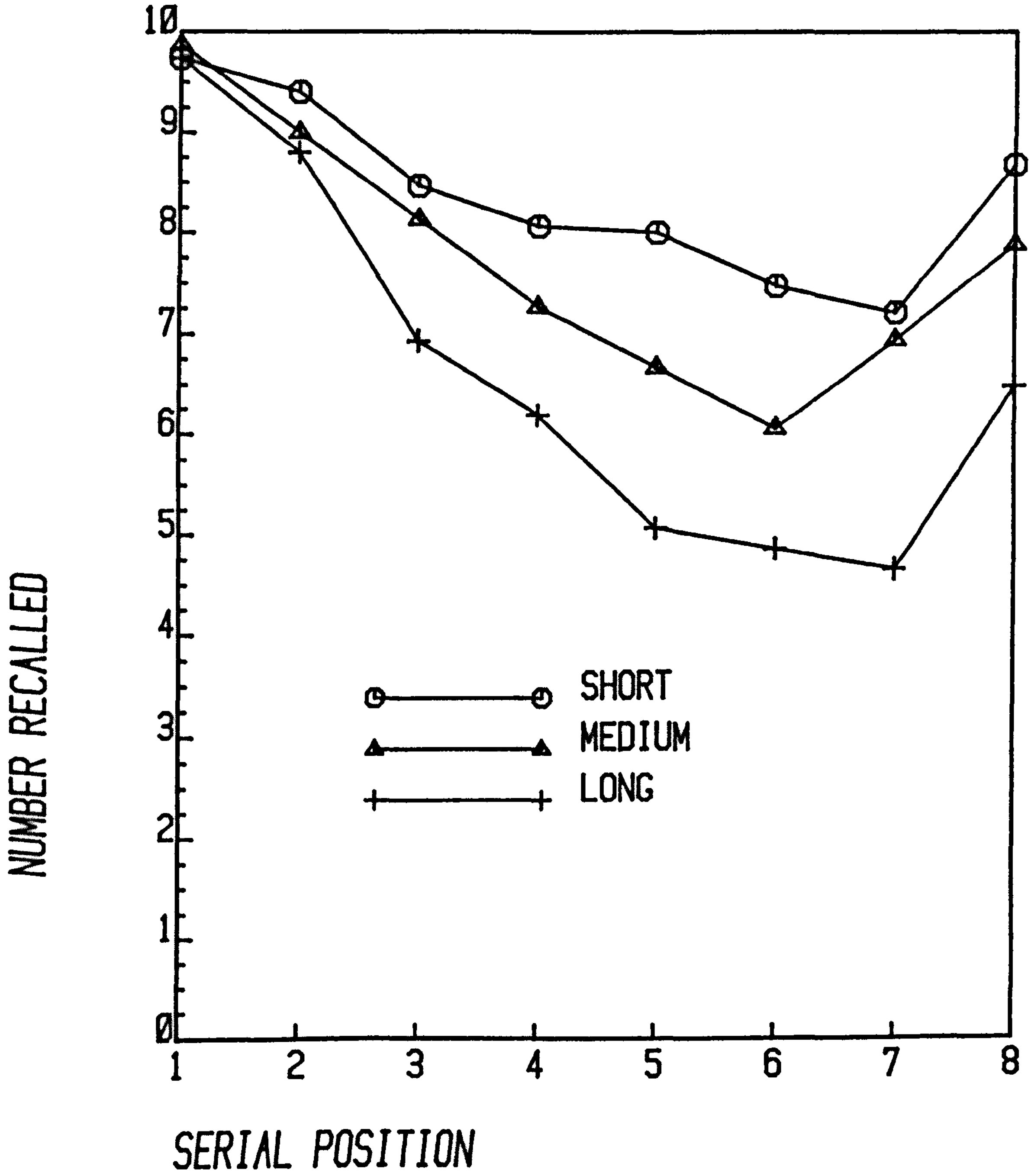


FIGURE 14
SERIAL POSITION CURVES FOR
SHORT WORDS OVER THE 4 TRIALS

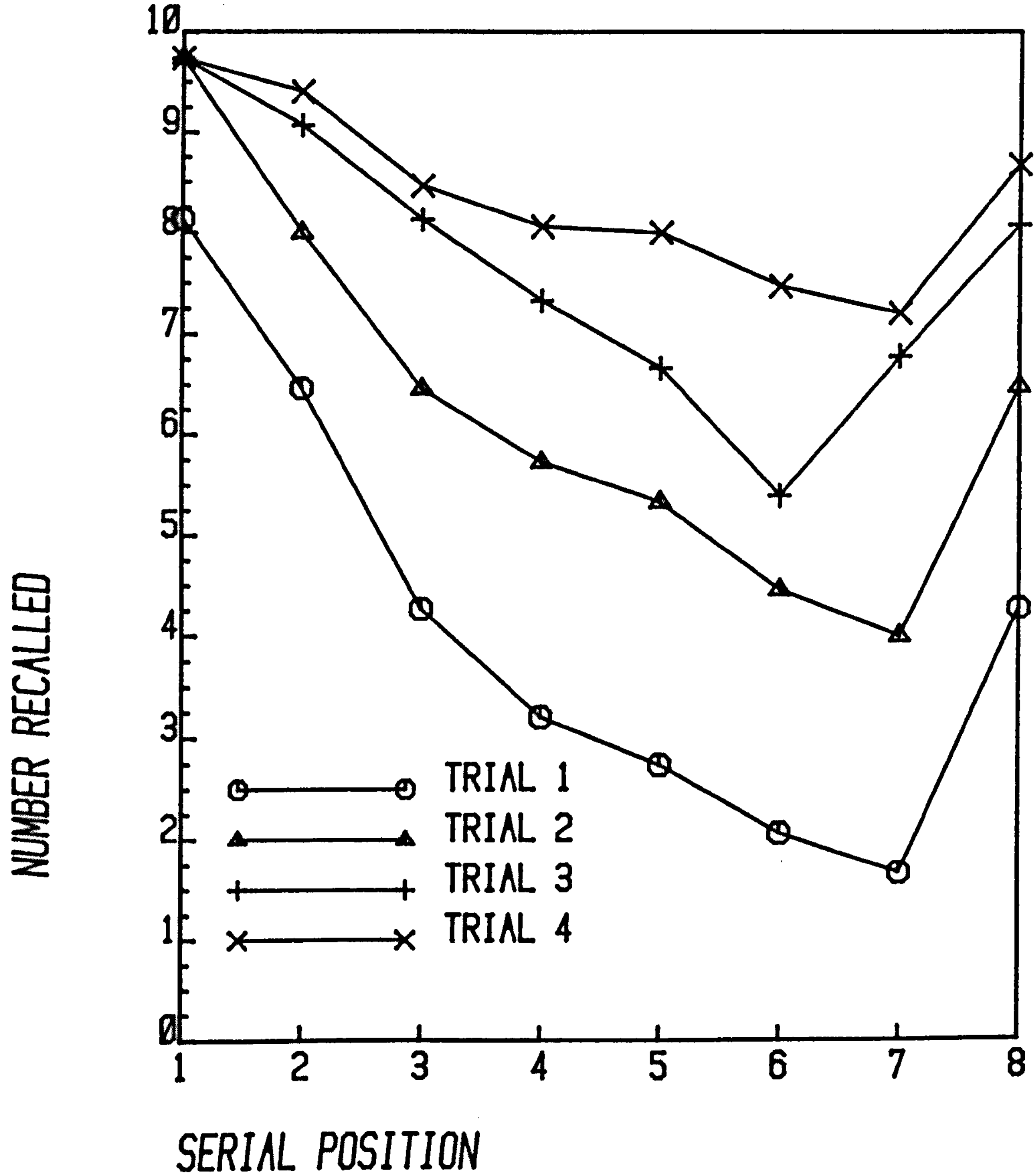


FIGURE 15
SERIAL POSITION CURVES FOR
MEDIUM WORDS OVER THE 4 TRIALS

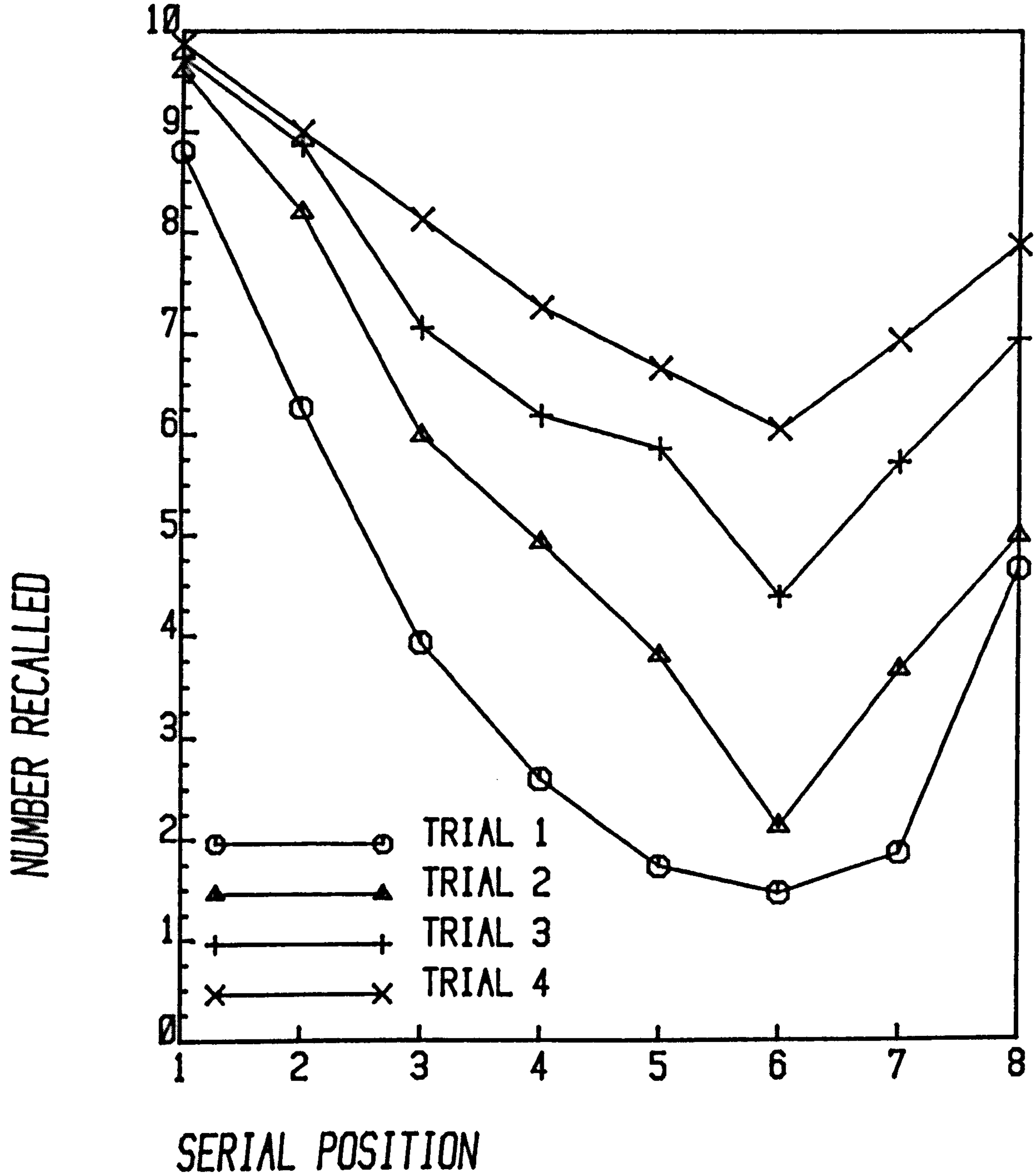
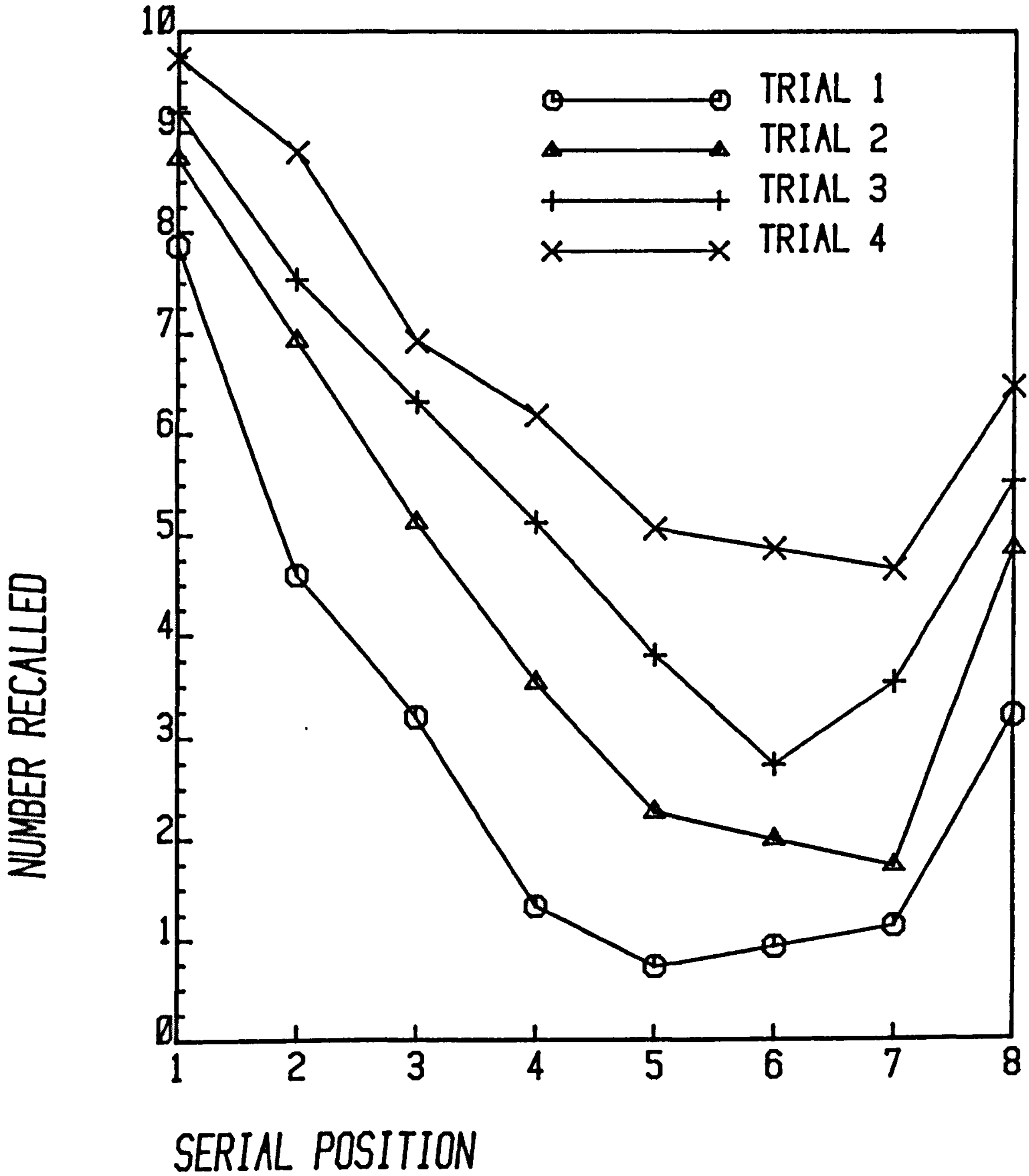


FIGURE 16
SERIAL POSITION CURVES FOR
LONG WORDS OVER THE 4 TRIALS



The interaction may be 'put into words' as: the change in the serial position curves over trials differs for the different word lengths. The simplest explanation of this effect would seem to be that there is a ceiling effect on the first list item and that this is greatest for the shorter words. This can be seen most clearly in figures 14, 15 and 16 where for the short words subjects score the same for serial position one on trials 2, 3 and 4. However, for the long words, scores on the first serial position are spread from approximately 10.75 down to 8.75 for the first three trials. From graphs 10 to 13 there appears to be no tendency for the word length effect to be reduced over the last few serial positions as occurred with experiment 1. Thus, the 'ceiling effect' explanation will be accepted and this interaction will not be further discussed.

DISCUSSION

The most important result of this experiment is that the slope of the NR vs RR function remained constant over trials while the intercept increased. This accords with the prediction made earlier and supports strongly the ideal that the slope is a measure of decay time in STM and the intercept the LTM component. The fact that the results are well described by the function $NR = RR.k_1 + c$ for a design in which word length was a between subjects variable has two implications. First it rules out the idea that the word length effect might be due to 'range effects' (Poulton, 1973) and secondly it suggests that the decay time of an item is constant across subjects. The implications of this result were discussed to some extent after the previous experiment and it does now seem that decay is the product of some neural process. However, further discussion of this result will be postponed until after the other experiments have been reported.

The mean slope for the four lines in figure 9 is 1.0815 compared to 1.87 for the previous experiment and it is not obvious why there should be this difference. This result will be discussed after the next experiment has been reported as it turns out to be a complicated issue and would cause the main argument to become difficult to follow.

To return to the main line of argument, experiment 5 provides evidence that the intercept reflects the LTM component of serial recall but only indirect evidence that the slope is a measure of decay time in STM component. The next experiment is designed to provide direct evidence that the slope is the STM component by showing that the word length effect is influenced by the same variables as the AS effect. In the introduction evidence was presented showing that a distractor task interpolated between presentation and recall of a list reduces performance and it was argued that this was primarily due to the loss of the STM component. The main evidence for this assertion was that the acoustic similarity effect disappeared with a distractor task (Conrad, 1967) and it follows that if word length is also mediated by STM, it should be abolished by a distractor task. This prediction would also follow from decay theory. The distractor task will stop rehearsal of the list items and, therefore, they will decay and, providing the distraction is maintained for longer than the decay time of an item, will have disappeared from STM at the start of recall. Thus, the next experiment investigates the change in the word length effect with a distractor task.

EXPERIMENT VI

The purpose of this experiment is to investigate the change in the NR vs RR function when a distractor task is interpolated between presentation and recall of serial recall lists. The prediction is that the slope of the function (the STM component) will be reduced to zero while there will be little change in the intercept (the LTM component).

Three levels of word length were combined with two recall conditions (immediate and delayed) in a 3 x 2 within subjects design. Subjects did eight six-word lists in each condition, half of them doing the immediate recall condition first and half the delayed recall condition first. Within a recall condition the lists were blocked within a word length and the ordering of the different word lengths was determined by a regular latin square.

The words used were taken from the 1, 3 and 5 syllable lists of experiment 4, except that it was found necessary to replace some of the three syllable words with slightly longer ones to obtain roughly equal spacing between the word pools in terms of reading rate. This may have been due to the subjects being younger than in previous experiments or, more likely, to their dialect being Yorkshire instead of Cambridgeshire. The words used are shown in table 13.

METHOD

Eight lists, six words in length, were constructed for

TABLE 13

Short	Medium	Long
Mumps	Leprosy	Tuberculosis
Maine	Wyoming	Louisiana
Stoat	Kangaroo	Hippopotamus
Switch	Television	Refrigerator
Scroll	Bulletin	Periodical
Zinc	Calcium	Aluminium
Maths	Biology	Physiology
School	Academy	University
Greece	Mexico	Yugoslavia
Crewe	Exeter	Weston-super-mare

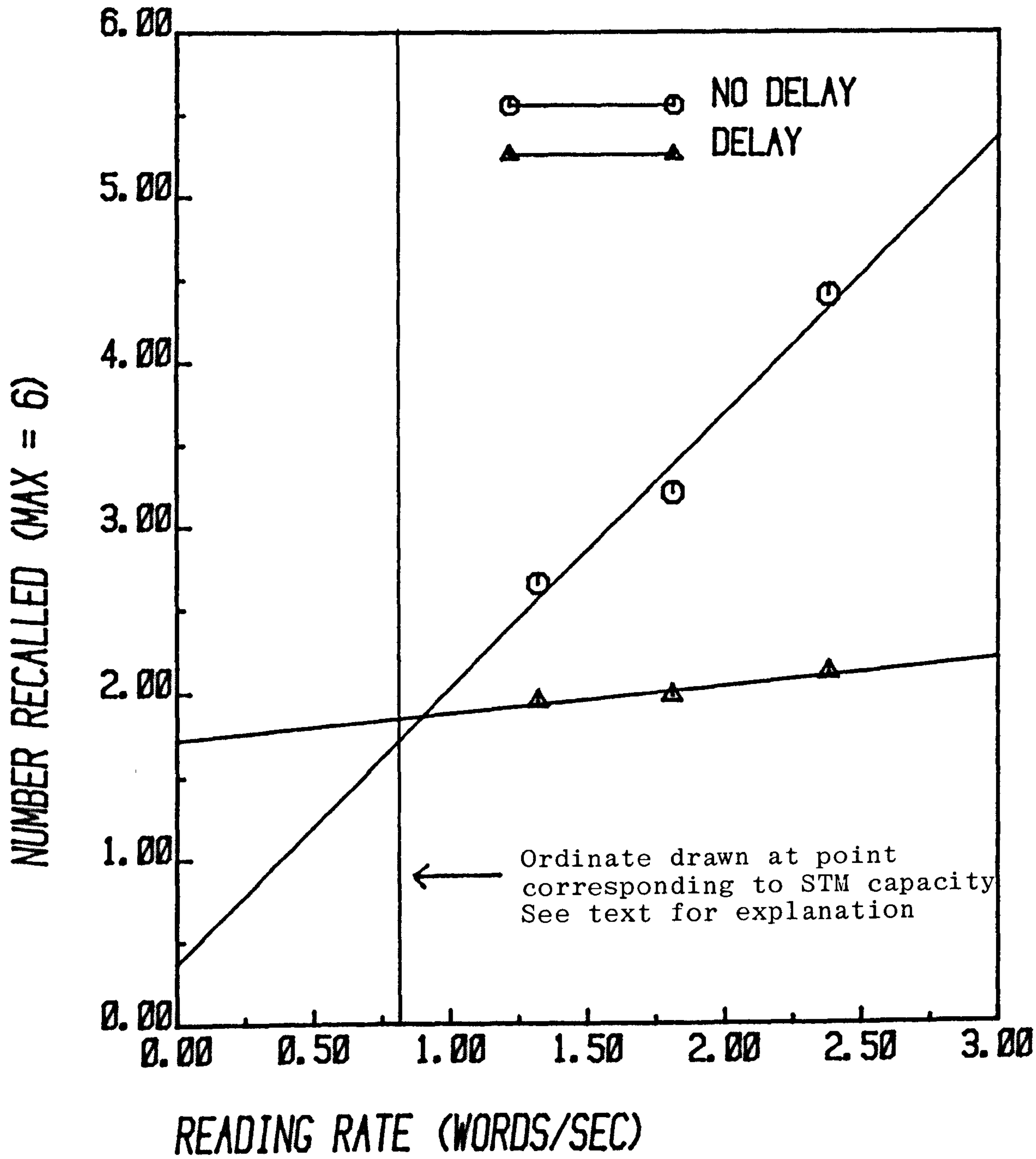
each word length by sampling at random without replacement from a pool six times. These lists were then recorded on to tape at a 1.2 second rate. The subjects were first allowed to read through the word pools to familiarise themselves with the material. In the immediate recall condition the lists were played to the subject who then tried to recall them in order as soon as the last list item had been presented. In the delayed recall condition subjects had to write down ten digits read by the experimenter at a one second rate immediately after the list. The experimenter used a visual metronome to keep to a one second reading rate. In both conditions recall was spoken and subjects were again told to say 'blank' if they could not remember a word. Because the procedure was more complicated than for previous experiments, subjects were given six practice trials, one for each of the conditions, before doing the experimental lists.

After this phase of the experiment, subjects read the lists aloud as fast as possible using the same procedure as in the previous experiment. They read through each set of eight lists twice providing four measurements of reading rate for each word length. Half the subjects read the lists in ascending order of word length and half in descending order of word length. Reading times were measured by stopwatch. Twelve subjects, all undergraduates at the University of York, were tested and were paid for their services.

RESULTS

The NR vs RR functions are plotted in figure 17 for

FIGURE 17
NUMBER RECALLED VS READING RATE
FOR DELAY AND NO DELAY CONDITIONS



the two recall conditions. An analysis of variance was performed on the data with number recalled as the dependent variable and word length and recall condition as the independent variables. The results are shown in table 14.

The main effects of word length and type of recall were significant and, in line with the predictions made, there was a significant interaction between these two variables. A regression analysis was performed on each recall condition and for immediate recall fave a slope of 1.663, and intercept value of 0.369 and the standard error of the slope was 0.278. For delayed recall condition the slope was 0.160, the intercept was 1.721 and the standard error of the slope was .050. The results may be summarized by saying that the addition of a distractor task reduces the slope of the NR vs RR function and increases the intercept.

DISCUSSION

The reduction of the word length effect with a distractor task accords with the predictions of decay theory and supports the idea that the slope is a measure of decay time and STM component of recall. Thus, there is now evidence that the slope represents the STM component and the intercept the LTM component. These are important results as they imply that word length provides a means of separating these two components in serial recall without changing their magnitude. Consequently, it will be possible to examine the effect of a variable on either or both components. However, before too much weight is placed on this idea, the unpredicted change in intercept with the distractor task that occurred in this experiment must be considered.

TABLE 14

Source	df	SS	MS	F	Probability
Subjects (S)	11	4881.500			
Recall (R)	1	2289.389	2289.389	69.450	<.001
R x S	11	362.611	32.965		
Word length (WL)	2	741.333	370.667	22.800	<.001
WL x S	22	357.666	16.258		
R x WL	2	500.111	250.056	31.456	<.001
R x WL x S	22	174.889	7.949		

It would not be unreasonable to expect the distractor task to reduce the LTM component for reasons discussed in the introduction, viz. that PI could increase the longer the items are in LTM. However, the results seem to show an increase in the intercept with a distractor task and this is not easily accommodated by any of the ideas previously discussed. However, as discussed previously (experiment 4), the ordinate should be placed at a reading rate corresponding to the decay time of an item as the relationship between NR and RR can only be expected to hold for word durations that are equal to or less than this decay time. If a single word exceeds the decay time then, without making further assumptions, it is not possible to predict from the present theory whether that word will be recalled. Thus, the equation relating RR and NR should be properly written as:

$$NR = RR.k_1 + c \ : \text{ where } 1/RR < k_1$$

This is quite reasonable as it states that the decay time of words in STM can only be predicted from the durations of those words when they themselves do not exceed the decay time. If the decay time is calculated from the 'no delay' condition and converted into a reading rate score (by taking the reciprocal) then an "ordinate" can be drawn at this point. This has been done in figure 18 and it can be seen that there is virtually no difference in the intercept value of the two recall conditions when measured at this point.

At this point it is appropriate to summarize the results obtained so far. The first set of experiments showed that chunking theory was not an adequate explanation of STM storage capacity and that decay theory consistently produced the correct predictions. The last two experiments provide evidence that it is STM that is sensitive to word length while LTM is insensitive to this variable. This allows the slope of the NR vs RR function to be used as a measure of decay time in STM and the intercept as an estimate of LTM capacity. These results are also completely consistent with decay theory. The last experiment in the thesis was designed to strengthen the case for word length being an STM effect but also provides some problems for a decay theory explanation. However, before reporting this experiment there remains the problem raised by the results of experiments 4 and 5, concerning the large change in the STM component which occurred between these experiments. This is an important issue as one of the major assumptions of this research has been that the capacity of STM is constant and so any sizeable changes in capacity require explanation. The following discussion and the next experiment are an attempt to resolve this problem.

The mean slope for the four regression lines in experiment 5 is 1.082 compared with 1.87 for experiment 4 and it is not immediately clear why there should be this difference. There were three major changes in procedure between the experiments which might have produced this result: presentation rate, modality of presentation and list length, and each of these will now be considered.

The effects of presentation rate on immediate serial recall are complex and not well understood (e.g. Aaronson, 1967). This may well be due to rate having different effects on STM and LTM components, but in general small changes in rate such as those that occurred between experiments 4 and 5 (2 seconds per word to 1.5 seconds per word) do not have large affects. Further, in experiment 6 a faster rate was used (1.2 seconds per word) than for experiment 5 and yet the slope was much closer to that obtained for experiment 4, which used the slowest rate. (The slope for experiment 6 was 1.663). Thus, it seems unlikely that the change in presentation rate alone caused the change in slope although it remains possible that it could have been an interaction between the change in modality and the change in rate as it is known that these two variables interact (Posner, 1964).

The change in presentation modality seems a more likely explanation as it has been claimed that a modality specific store (pre-categorical acoustic store; PAS) is involved when material is presented auditorially but not when presented visually. (Crowder & Morton, 1969). However, this could only account for the change in slope if PAS was insensitive to word length and if some of the list items were being recalled from this store. If this were the case then STM would be making a relatively smaller contribution to recall producing a reduction in the slope of the function while the contribution of PAS would cause an increase in the intercept. In fact the intercept value for the first presentation trial of experiment 5 was higher than that obtained for experiment 4 and so at first glance this would seem a plausible explanation. Further, there is evidence from Watkins and Watkins (1972) that PAS is insensitive to word length. They found that there was an interaction between serial position and word length such that there was no word length effect over the recency region of the serial position curve, a result similar to that obtained for experiment 1 in this thesis.

Unfortunately, there is a problem with this explanation. It assumes that the last few list items are recalled from PAS alone and, therefore, that the contribution of STM is relatively smaller. Crowder and Morton (1969) argue convincingly that this cannot be the case; that PAS adds to rather than replaces the contribution of STM for the last list items. The evidence for this assertion comes from the fact that a stimulus suffix which erases the PAS contribution does not reduce recall to zero, but to the level obtained for visually presented material, i.e. to a level where STM alone determines recall. Thus, the addition of a PAS component may well increase the intercept of the NR vs RR function if it is insensitive to word length, but should not reduce the STM contribution and hence the slope. Watkins and Watkins' results and those of experiment 1 must then be explained by assuming that the last list items were either not registered in STM or had been lost by the time they were to be recalled. This is very much in line with the explanation of the interaction in experiment 1 that was provided at the time. It was assumed that the slow pacing of recall caused the last items to have decayed to a point where they could not be retrieved when their time for recall came. A similar explanation can be offered for Watkins et al's results in that they used fairly long lists and subjects used written recall which takes appreciably longer than spoken recall. It is worth noting that had they not obtained an interaction between serial position and word length they would not have been able to conclude that PAS is insensitive to word length; the result could have been due to the STM component for the last items. Similarly, the fact that the interaction disappeared in experiments 2 and 3 does not discredit the idea that PAS is insensitive to word length. Indeed, only by assuming this can the results of Watkins et al and experiment 1 be explained.

Thus in conclusion of this rather complicated issue, it does not seem likely that the change in slope between experiments 4 and 5 can be attributed to the influence of PAS. It does seem likely, however, that PAS caused the change in intercept particularly as experiment 6 which, also used auditory presentation, had a higher intercept than experiment 4. The final reason for dismissing PAS as the cause of the change in slope is that experiment 6 in which auditory presentation was used produced a slope more comparable with experiment 4 than with experiment 5.

The final change in procedure between experiments 4 and 5 which will be considered is the change in list length. Crannell and Parrish (1957) showed that serial recall performance decreased as list length increased. However, it is not clear from their experiment whether this decrease occurred in the STM or LTM components and, if it was the former, then this could be the cause of the reduction in slope as experiment 5 used a greater list length than experiment 4. It is not easy to see why an increase in list length should cause a reduction in the STM component. Subjects would presumably fill STM and then rehearse these items regardless of how many others were presented. However, this may not occur and subjects may use more complicated strategies which involve a compromise between the efficient use of STM and LTM. The nature of these strategies is unknown and at present it is not possible for decay theory to predict precisely what will happen to the STM component with changes in list length. The LTM component of recall should be reduced by increases in list length as the greater the number of items the more chance there is for interference. Thus, for a number of reasons it would seem pertinent to run an experiment looking at the effect on the STM and LTM components of recall produced by changes in list length and the next experiment is an attempt to do this. The main purpose of the experiment is to determine whether the change in the STM component between experiments 4 and 5 can be attributed to the change in list length between the experiments.

EXPERIMENT VII

Three levels of word length (1, 3 and 5 syllables) were combined with three levels of list length (5, 6 and 7 words) in a 3 x 3 within subjects design. Each subject did six replications of each condition making a total of fifty-four (3 x 3 x 6) lists per subject. Lists were blocked within a condition so that subjects did all six replications for a particular condition one after the other. Twelve subjects were used, half receiving the different word lengths in ascending order (1, 3 then 5 syllable lists) and half in descending order. The ordering of the different list lengths was determined by a latin square shown in table 15. Subjects did all the word length

TABLE 15

Order of List Lengths

Subjects

1, 4, 7, 10	2, 5, 8, 11	3, 6, 9, 12
5 words	6 words	7 words
6 words	7 words	5 words
7 words	5 words	6 words

conditions of one list length before doing the next list length. This was done because the experiment was designed to determine the effects of list length and subjects might not find the optimum strategy for a particular list length if this variable was randomised.

METHOD

The lists were constructed from the same pool of words used in experiment 5. From each pool six lists of each list length were constructed by sampling at random without replacement.

The lists were read to the subjects at a 1.5 second rate paced by a visual metronome. Subjects were required to recall the lists in the order presented immediately after the last item had been presented. They were instructed to say 'blank' if they could not remember a word so that subsequent words were not assigned the wrong ordinal position. Before the memory task subjects were allowed to read through the word pools to familiarise themselves with the experimental material.

In order to check that speech rate and reading rate are equivalent for the purpose of predicting recall (as found in experiment 3) both were measured in this experiment. Reading rate was measured by timing the subject reading a list of forty-eight words of one word length (3 occurrences of each word in the pool) typed in random order in upper-case on a sheet of paper. This measure was taken three times after the memory task and was done for each word length. Speech rate was measured in the same way as in experiment 3. Subjects were given three different words of a particular word length and asked to repeat them as fast as possible consistent with a correct pronunciation of each word. The subject was timed for ten repetitions of the sequence and this was done with three different sets of words for each word length. All timing was done with a stop-watch.

Twelve subjects were tested, all members of the A.P.U. subject panel and they were paid for their services.

RESULTS

The mean number recalled is plotted as a function of reading rate in figure 18 and as a function of speech rate in figure 18a. An analysis of variance performed on the memory scores with list length and word length as the independent variables produced the results shown in table 16.

This analysis shows quite clearly that while both word length and list length produced significant main effects, they did not interact with each other. This implies that list length affects the LTM and not the STM component, i.e. it causes changes in the intercept of the NR vs RR (or SR) function but no change in the slope. This is also clear from figures 18 and 18awhere the best fitting straight lines have been drawn from the results of a regression analysis shown in table 17.

These results confirm the fact that both speech rate and reading rate are good predictors of serial recall performance and, when taken with the previous experiments, show the word length effect to be a reliable and fairly robust phenomenon. Unfortunately, the results do not shed

TABLE 16

Source	df	SS	MS	F	Probability
Subjects (S)	11	46.6111			
List Length (LL)	2	7.4578	3.7289	6.4974	<.01
LL x S	22	12.6260	0.5739		
Word Length (WL)	2	38.9656	19.4828	51.7367	<.001
WL x S	22	8.2847	0.3766		
LL x WL	4	1.4124	0.3531	1.4270	ns
LL x WL x S	44	10.8875	0.2474		

TABLE 17

Regression of Reading Rate upon Memory Scores

List Length	Slope	Intercept	Standard Error of Slope
5	0.881	1.233	.1122
6	1.115	0.552	.0595
7	1.067	0.208	.2867

Regression of Speech Rate upon Memory Scores

List Length	Slope	Intercept	Standard Error of Slope
5	0.662	1.676	.1983
6	0.864	1.056	.0963
7	0.856	0.625	.0844

FIGURE 18
MEAN NUMBER OF WORDS RECALLED VS
READING RATE FOR EACH LIST LENGTH

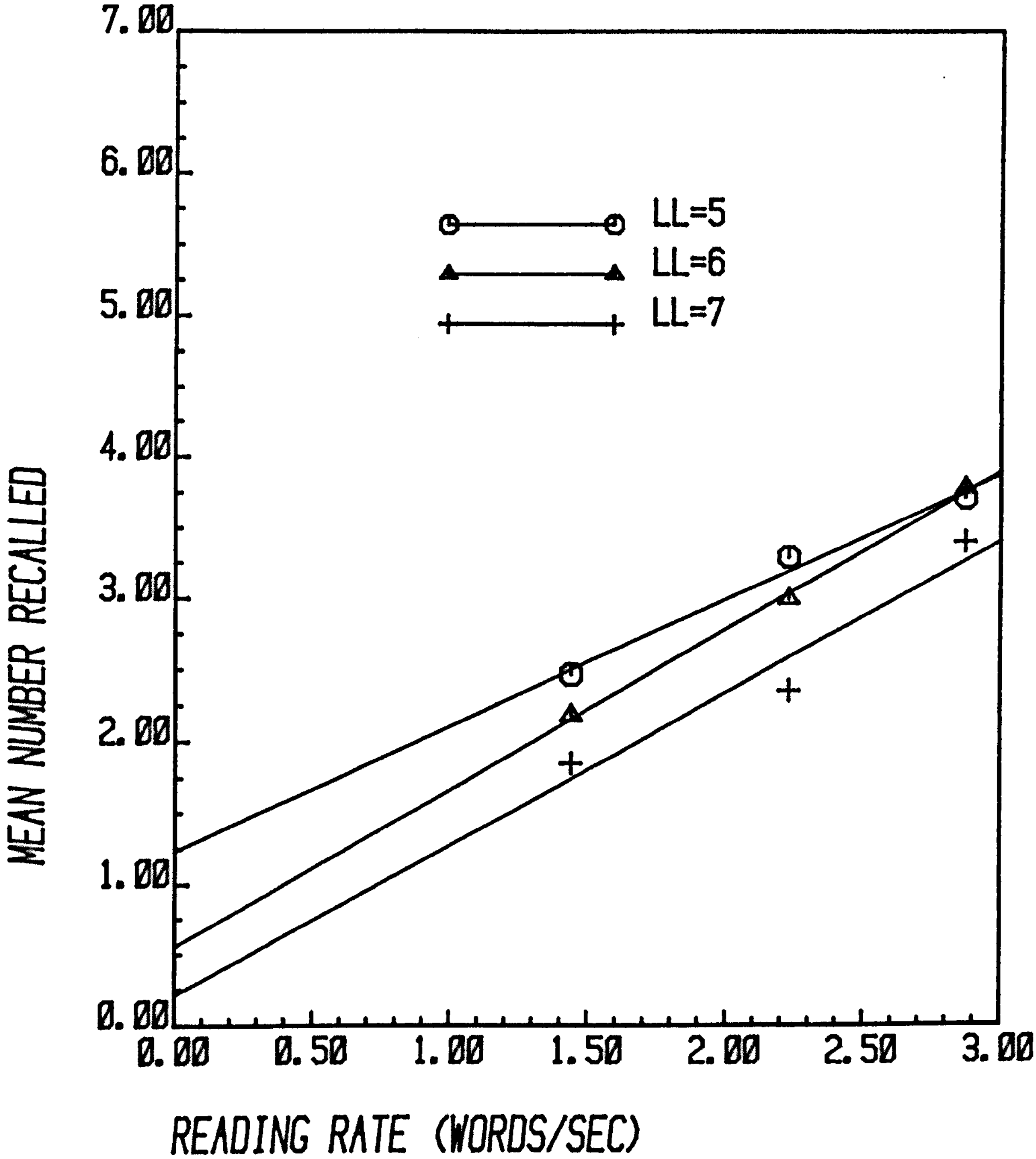
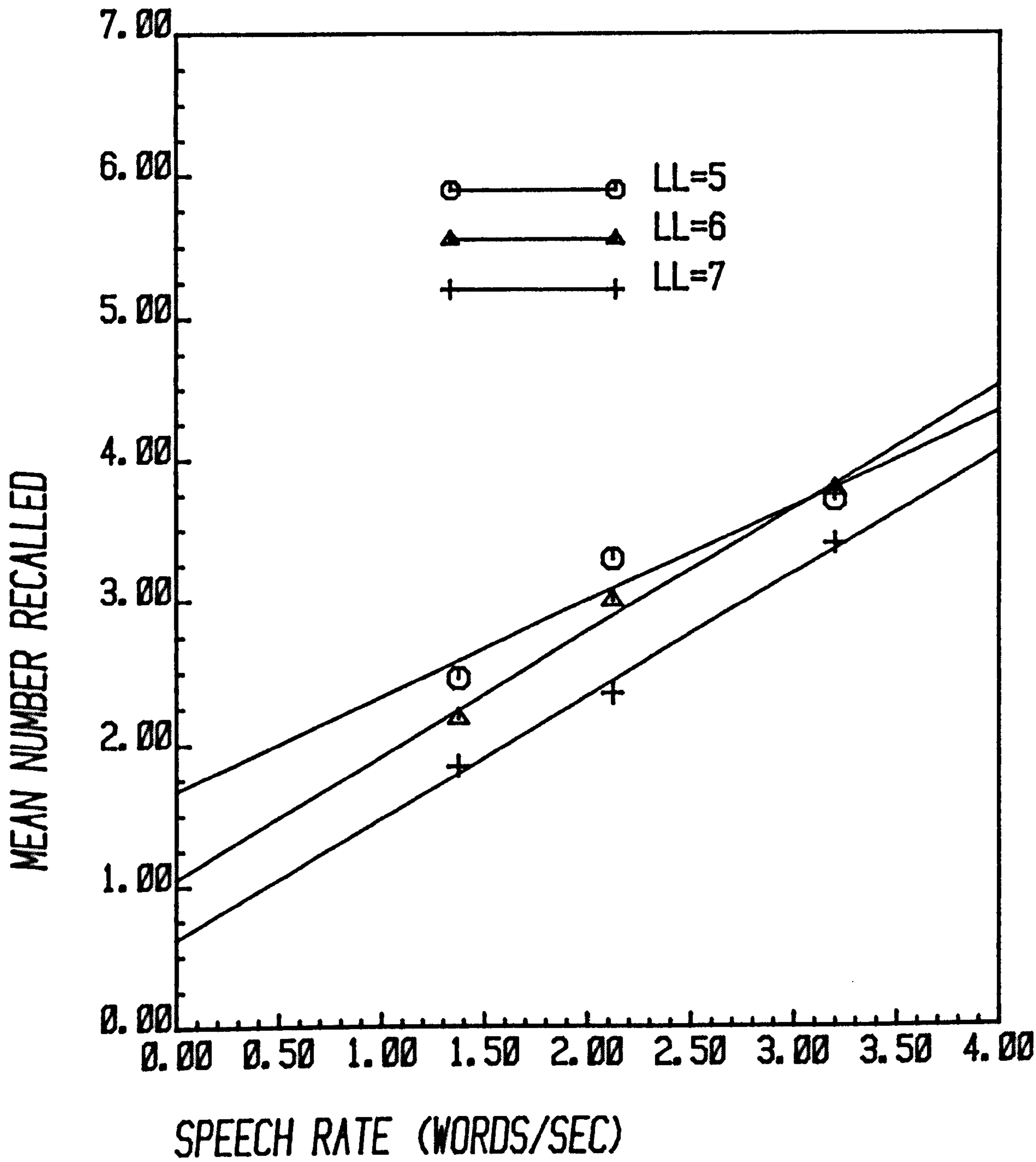


FIGURE 18A
MEAN NUMBER OF WORDS RECALLED VS
SPEECH RATE FOR EACH LIST LENGTH



much light on the difference in slope between experiments 4 and 5. The results show quite clearly that the reduction in performance with increases in list length is due to a reduction of the LTM component and not the STM component. This would not be readily predicted by theories which assume that forgetting in STM is caused by interference as longer lists should produce greater inter-item interference and hence a smaller STM component. However, the result does accord with an interference theory of forgetting from LTM although it is almost certainly the case that most theories would predict that the amount of forgetting from LTM will increase with the number of items to be remembered.

The results provide no problem for decay theory for, as argued earlier, the theory makes no prediction about the effects of changes in list length unless some additional assumptions are made concerning the nature of the strategies used by subjects. Before leaving this experiment and returning to the main line of inquiry, it is worth examining the serial position data in an 'ad-hoc' manner to see whether any further information is present.

The serial position curves are shown in figures 19 to 21 for each list length and it would appear that they are very similar to those obtained in the previous experiments (excepting experiment 2) in that the word length effect is present at all serial positions.

It was not possible to perform a single analysis of variance with serial positions as a dependent variable due to the different list lengths used. The results of separate analyses are shown in table 18 where the dependent variable

FIGURE 19
PERCENTAGE RECALLED VS SERIAL POSITION
FOR EACH WORD LENGTH - LIST LENGTH 5

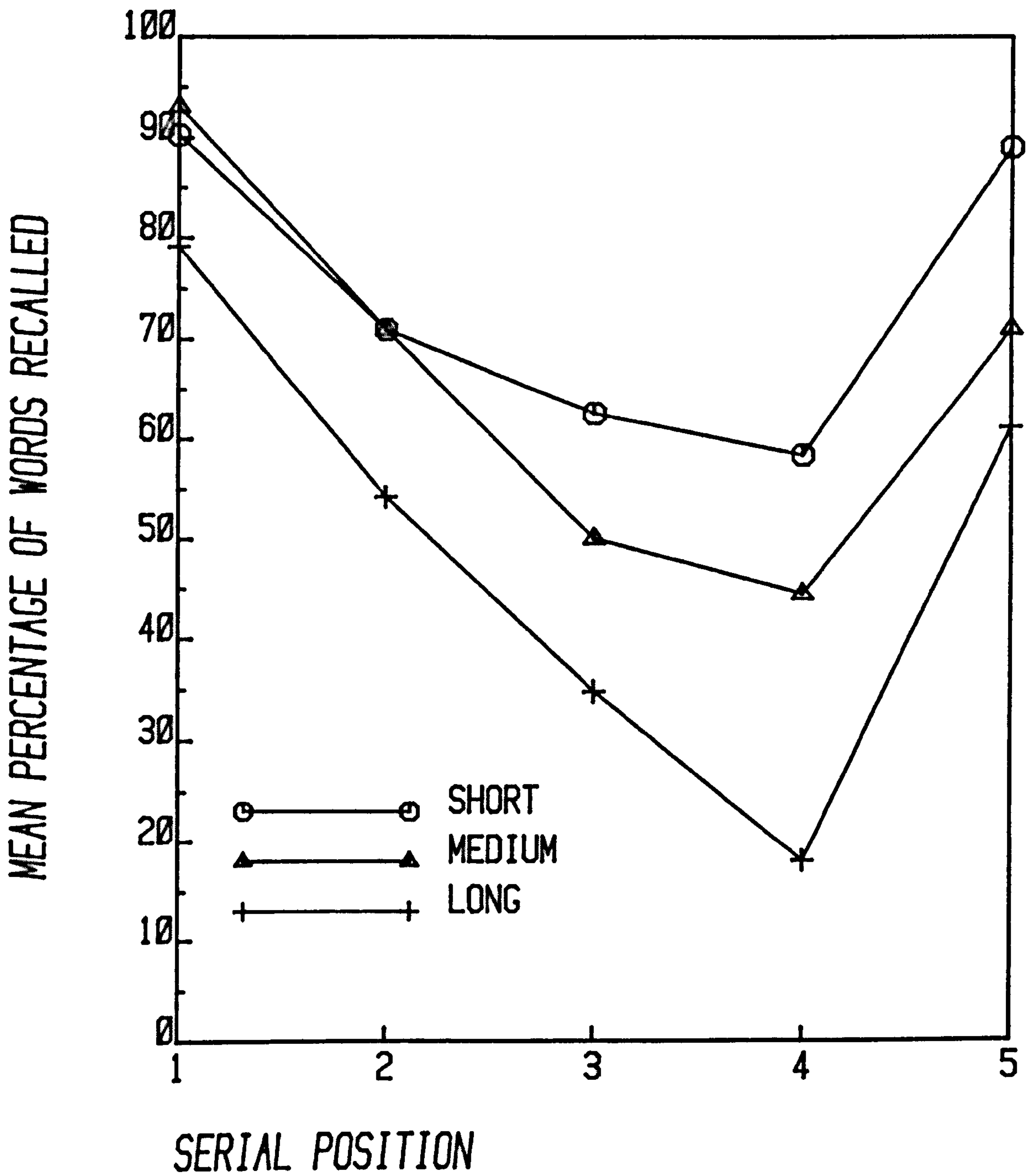


FIGURE 20
PERCENTAGE RECALLED VS SERIAL POSITION
FOR EACH WORD LENGTH - LIST LENGTH 6

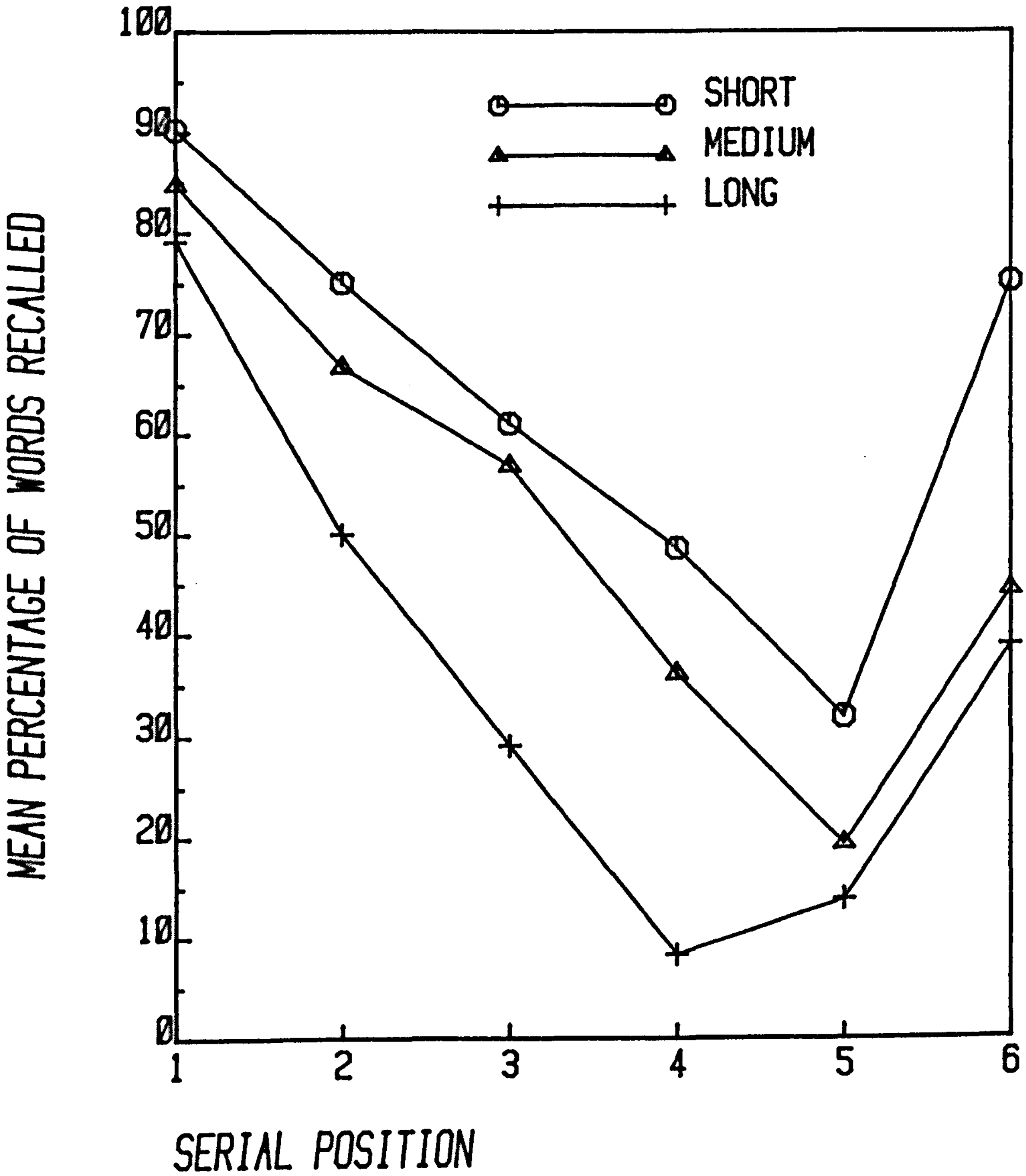


FIGURE 21
PERCENTAGE RECALLED VS SERIAL POSITION
FOR EACH WORD LENGTH - LIST LENGTH 7

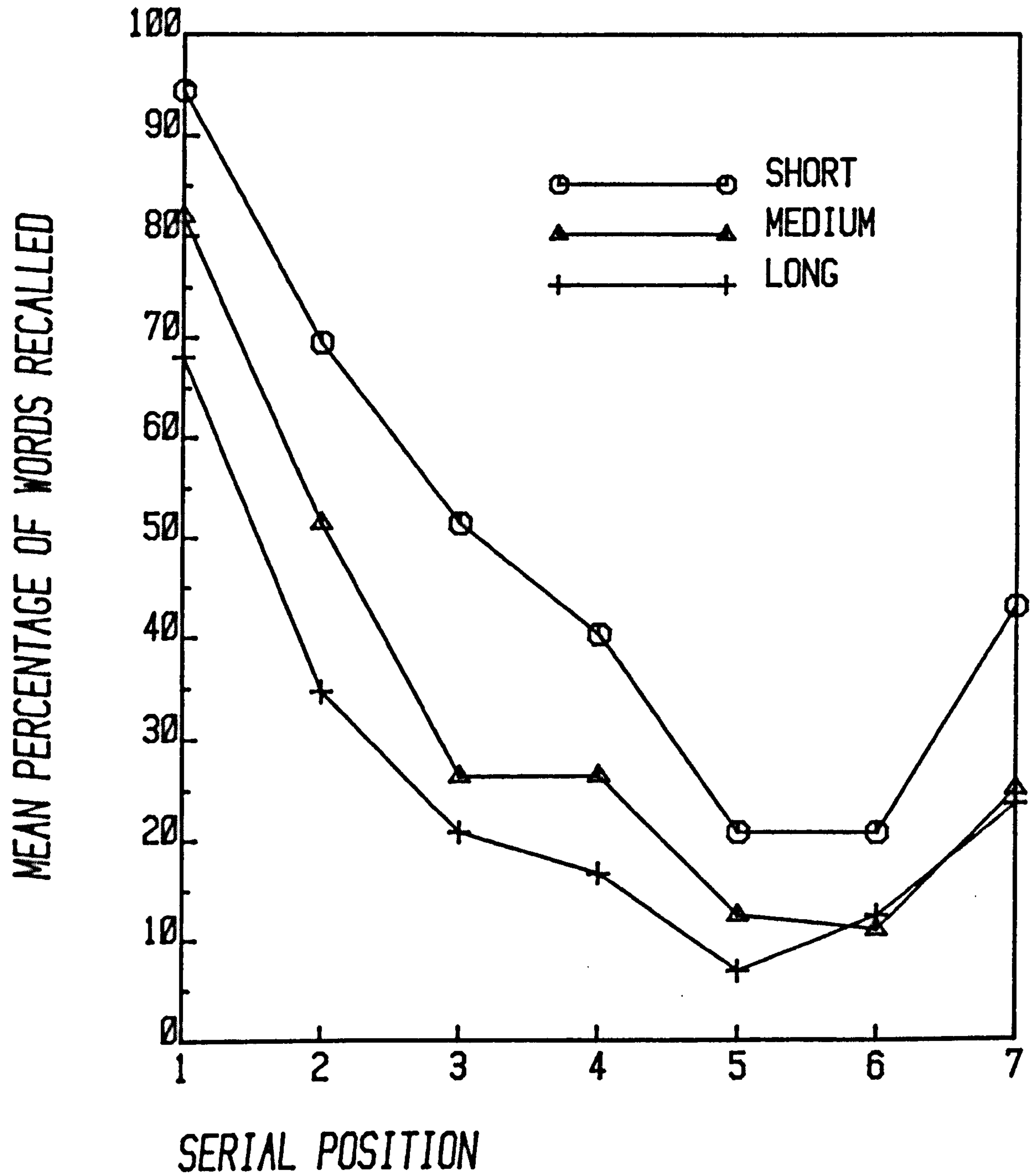


TABLE 18

Source	df	SS	MS	F	Probability
Analysis for List Length 5					
Subjects (S)	11	167.844			
Word Length (WL)	2	68.344	34.172	31.603	<.001
WL x S	22	23.789	1.081		
Serial Position (SP)	4	185.089	46.262	34.482	<.001
SP x S	44	59.044	1.342		
WL x SP	8	14.378	1.797	1.941	ns
WL x SP x S	88	81.489	0.926		
Analysis for List Length 6					
Subjects (S)	11	125.833			
Word Length (WL)	2	97.028	48.514	20.803	<.001
WL x S	22	51.306	2.332		
Serial Position (SP)	5	332.389	66.478	51.537	<.001
SP x S	55	70.944	1.290		
WL x SP	10	24.917	2.492	2.552	<.05
WL x SP x S	110	107.417	0.977		
Analysis for List Length 7					
Subjects (S)	11	89.467			
Word Length (WL)	2	79.024	39.512	26.513	<.001
WL x S	22	32.786	1.490		
Serial Position (SP)	6	438.722	73.120	43.818	<.001
SP x S	66	110.135	1.669		
WL x SP	12	13.921	1.160	1.406	ns
WL x SP x S	132	108.936	0.825		

is number recalled and the independent variables are serial position and word length.

The analyses show that serial position and word length produce significant effects for every list length and, apart from list length six, the two variables do not interact. The interaction for list length six does not seem readily explicable and does not look like that found for experiment 2 where the word length effect disappeared over the recency part of the curve, or like those in experiments 4 and 5 where there was a ceiling effect on the first serial position. However, given that this analysis is purely ad-hoc and the significance levels have not been adjusted accordingly, it would be inappropriate to attach much importance to this result and it will not be considered further.

To conclude the discussion of this experiment it must be admitted that it did not provide the answer that was expected, viz. that changes in list length would cause changes in slope of the NR vs RR function. The next experiment, although designed for a different purpose, allows another explanation of the slope change to be investigated. It is possible that the changes in slope are caused by changes in subjects' rehearsal strategy. In the next experiment it is possible to compare the size of the word length effect when subjects can and cannot rehearse. If this produces changes in the slope, then it is possible that the slope changes between experiments 4, 5, 6 and 7 were due to changes in rehearsal strategy. This in turn would require explanation but it might be a first step towards understanding these results.

Further discussion of this point will,

therefore, be postponed until the results of the next experiment have been reported.

To return to the main line of investigation, it was decided to run a further experiment to test the idea that the word length effect is mediated by STM. This is one of the most important issues in the thesis and it was felt insufficient to have only one experiment specifically designed to test this assertion. The rationale for experiment 6 was to show that word length was affected by the same variables as AS which is based on the well established assumption that the AS effect is mediated by STM. The same rationale is used for the next experiment where the effects of articulatory suppression upon word length are investigated. In addition to providing further evidence that the word length effect is mediated by STM, the suppression paradigm also allows the decay theory explanation of the word length effect to be tested.

The articulatory suppression paradigm was discussed in the introduction where results were reported which showed that suppression abolished the AS effect when material was presented visually but not when it was presented auditorially. One interpretation of this result was that with visual presentation and suppression the material could not be encoded into STM. It would follow that the word length effect, if mediated by STM, would also be abolished when the material was presented visually with suppression. This prediction is tested in the next experiment where the size of the word length effect is measured for visually presented material with and without suppression.

The fact that the AS effect did not disappear when the material was presented auditorially provides a means of testing decay theory. If word length behaves in the same way as AS then it should also be unaffected by suppression with auditory presentation. However, if it is assumed that suppression stops subjects rehearsing because it is impossible 'to say two things at once' then decay theory should predict that the word length effect will be abolished under these conditions. This is because the explanation of the word length effect in terms of decay theory was that items decay over time and subjects use rehearsal to 'refresh' the trace. Because more short words than long words can be rehearsed within the decay time of an item, subjects recall more short words than long ones. Thus, it follows that if rehearsal is prevented both long and short words should decay at the same rate and the word length effect will disappear. Adding an auditory condition, with and without suppression, to the next experiment allows this prediction to be tested.

EXPERIMENT VIII

In this experiment there is no need to use more than two levels of word length as the point of interest is whether the word length effect is abolished or not by suppression. Thus, two levels of word length (short and long) were combined with two presentation modalities (visual and auditory) and two levels of suppression (suppression and no suppression) in a 2 x 2 x 2 within subjects design. Subjects did five lists in each condition making thirty lists per subject in total. The lists were blocked within a presentation modality with half the subjects doing the visually presented lists first and half doing the auditorially presented lists first. Within a presentation modality the ordering of the conditions was determined by the same Williams latin square used in experiment 1 (see figure 1).

METHOD

The experimental material was the one and five syllable word pools of experiment 4. The lists, all five words in length, were constructed by sampling at random from a pool without replacement. Fifteen lists of each word length were constructed.

In all conditions the words were presented at the rate of one word every two seconds. In the auditory conditions the lists were read to the subjects, a visual metronome being used to keep to a constant presentation rate. For the visual condition, the words were presented on a memory drum. The warning signal for the auditory condition was the word 'ready' and for the visual condition three asterisks appeared in the memory drum window. For both conditions subjects were instructed to recall the words in order as soon as the last list item had been presented. They were instructed to say 'blank' if they could not remember a word. In the no suppression conditions subjects either listened or looked at the words and when the list had finished attempted to recall them. In the suppression conditions subjects started saying 'the the the . . .' at a regular rate before the list was presented. When they were suppressing at a regular rhythm the list was started. Subjects continued to suppress while the list was being presented but stopped immediately they had received the last item and then attempted to recall the list in the usual manner. The rate of suppression was not closely controlled but subjects were encouraged to try to achieve a rate of approximately 4 words per second. There is no information in the literature on the effects of different rates of suppression, but from my own observations with subjects it would seem far more important to ensure that subjects do not pause or hesitate when suppressing than to control the rate of their speech.

Only two words lengths were used so no measure of reading rate was taken as it is not possible to plot the NR vs RR function with only two points on the graph. Sixteen subjects were tested, all members of the A.P.U. subject panel and were paid for their services.

RESULTS

The mean percentage correctly recalled as a function of word length for each condition is shown in figure 22. An analysis of variance was performed on the data with number recalled as the dependent variable and word length, modality and suppression as the independent variables. The results are shown in table 19.

From figure 22 it can be seen very clearly that the word length effect is present in all conditions except that of visual presentation with suppression. This is confirmed by the analysis of variance in which the significant three-way interaction can be expressed as: the change in the word length effect produced by suppression is different for the different presentation modalities. The significance of the three-way interaction makes it unnecessary to consider the main effects and two-way interactions.

FIGURE 22
PERCENTAGE CORRECTLY RECALLED AS A FUNCTION OF WORD LENGTH FOR EACH CONDITION

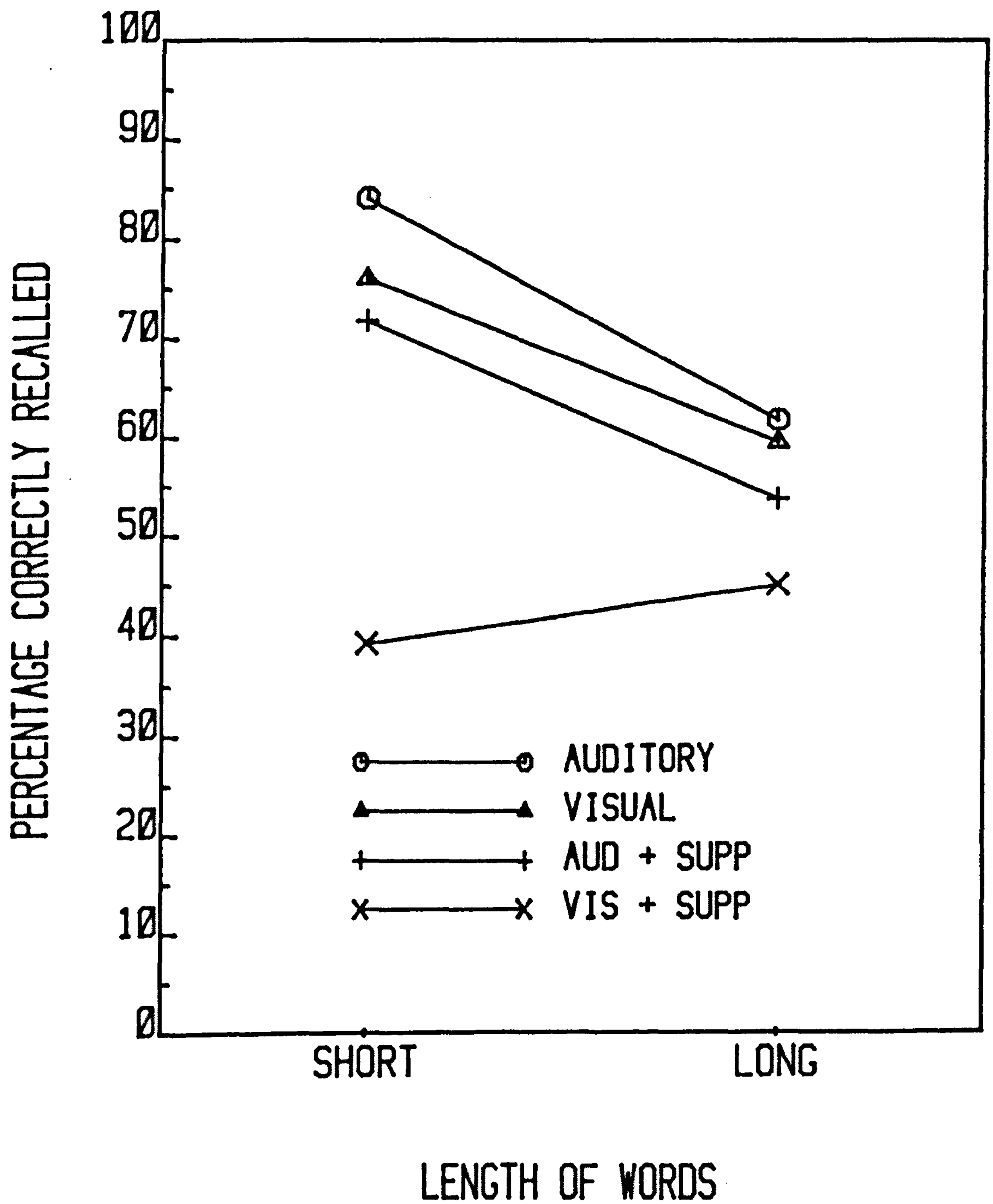


TABLE 19

Source	df	SS	MS	F	Probability
Subjects (S)	15	497.618	33.175		
Word Length (WL)	1	341.259	341.259	14.0198	<.005
WL x S	15	365.118	24.341		
Modality (M)	1	347.821	347.821	35.6588	<.001
M x S	15	131.555	8.770		
Suppression (Sp)	1	616.884	616.884	85.6844	<.001
S x Sp	15	107.992	7.200		
WL x M	1	118.195	118.195	8.813	<.01
WL x M x S	15	201.180	13.412		
WL x Sp	1	79.695	79.695	4.491	ns
WL x Sp x S	15	131.555	8.770		
M x Sp	1	130.008	130.008	33.1274	<.001
M x Sp x S	15	58.867	3.925		
WL x Sp x M	1	46.320	46.320	6.2284	<.05
WL x Sp x M x S	15	111.555	7.437		

The serial position curves for each condition are shown in figures 23 to 26. It is very clear from these curves and from the data (see appendix 2) that there were substantial ceiling effects on the first serial position, especially for the short words. This would cause serial position to interact with the other variables but the analysis would be meaningless as far as any theoretical points were concerned. For this reason no analysis of this data was performed. However, one point is worth noting: if the serial position curves for the no suppression conditions are considered it is very clear that the word length effect is the same for the visual and the auditory conditions. This would seem to rule out the change in presentation modality as the cause of the slope change between experiments 4 and 5. Further, given that suppression in the auditory condition produces no change in the size of the word length effect, it seems unlikely that slope changes in previous experiments were due to changes in rehearsal strategy. However, this point is complicated by the results with visual presentation and further discussion of this point is better left until these have been dealt with.

DISCUSSION

The main results of this experiment are very clear - word length behaves in exactly the same way as acoustic similarity under articulatory suppression. The word length effect is unchanged in all but the condition where the material is

FIGURE 23
SERIAL POSITION CURVES FOR
AUDITORY/NO SUPPRESSION CONDITION

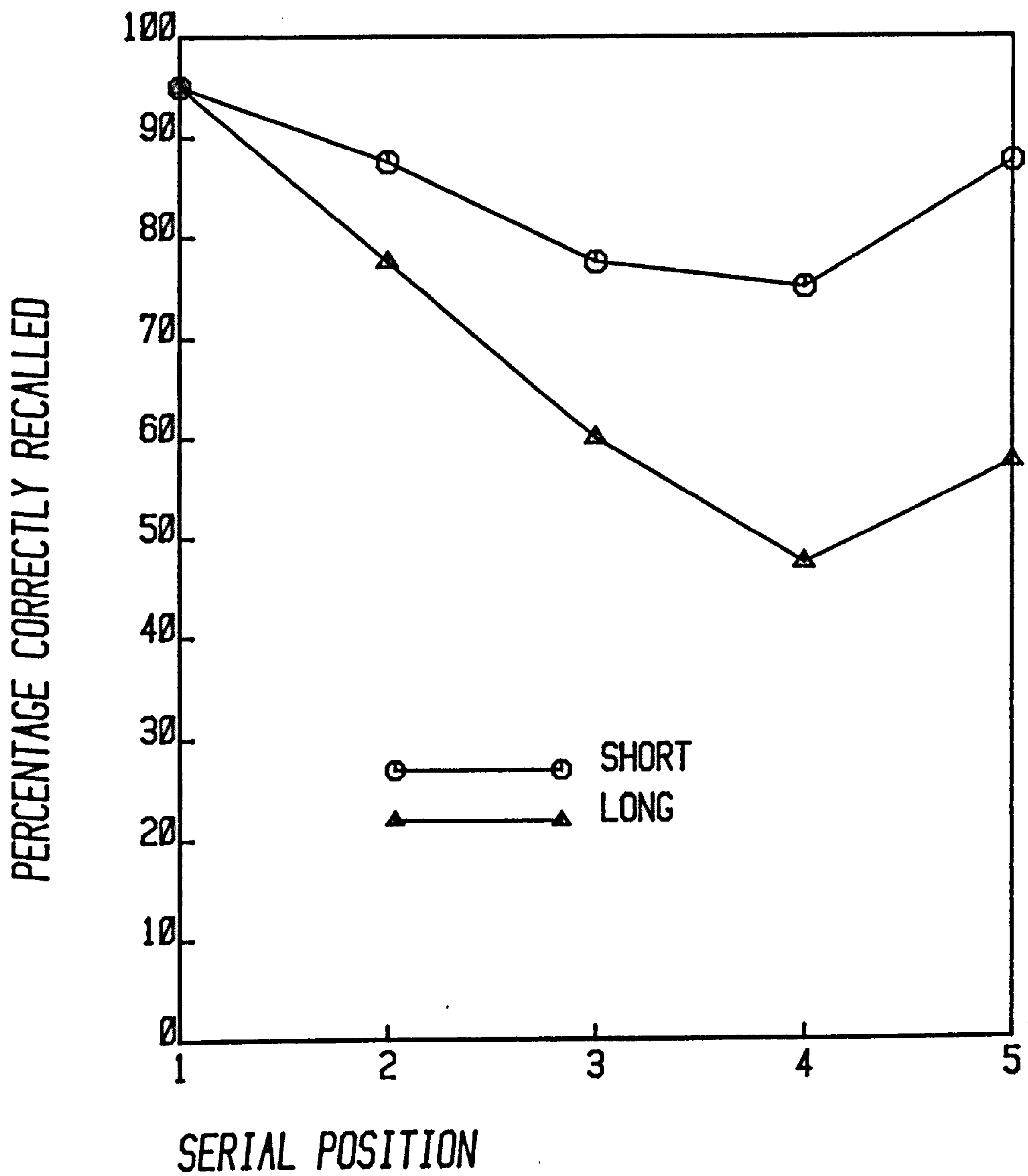


FIGURE 24
SERIAL POSITION CURVES FOR
VISUAL/NO SUPPRESSION CONDITION

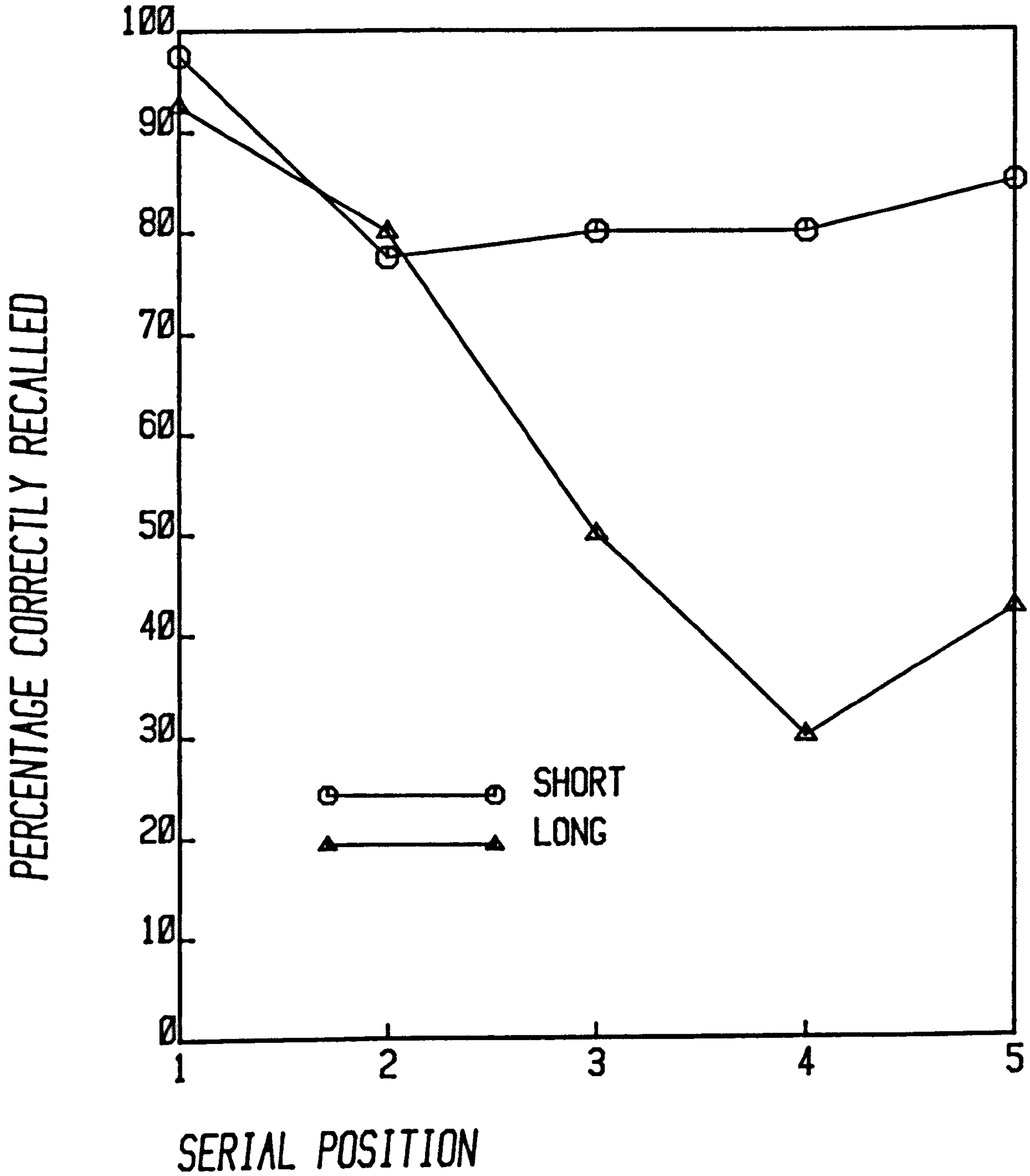


FIGURE 25
SERIAL POSITION CURVES FOR
AUDITORY/SUPPRESSION CONDITION

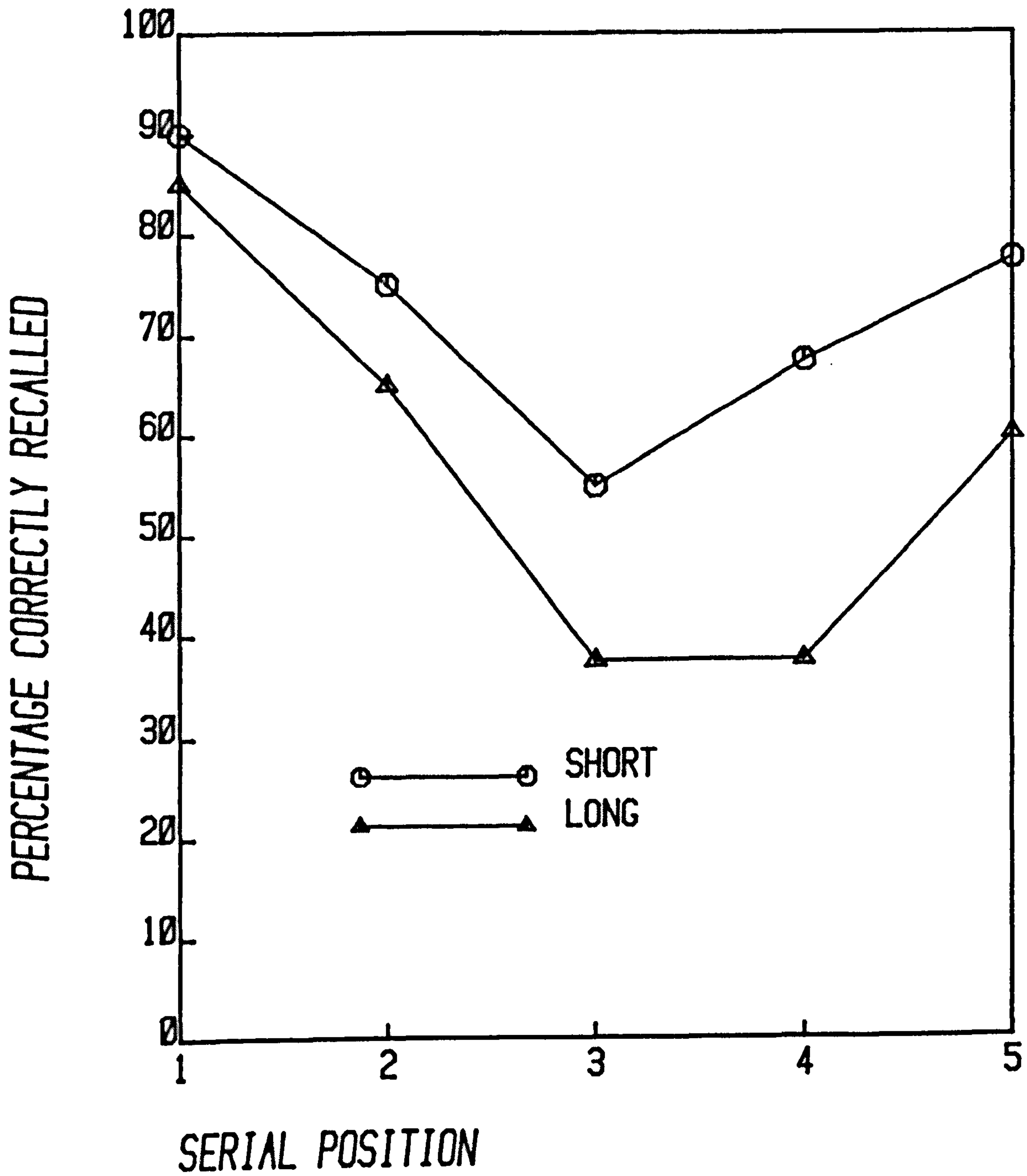
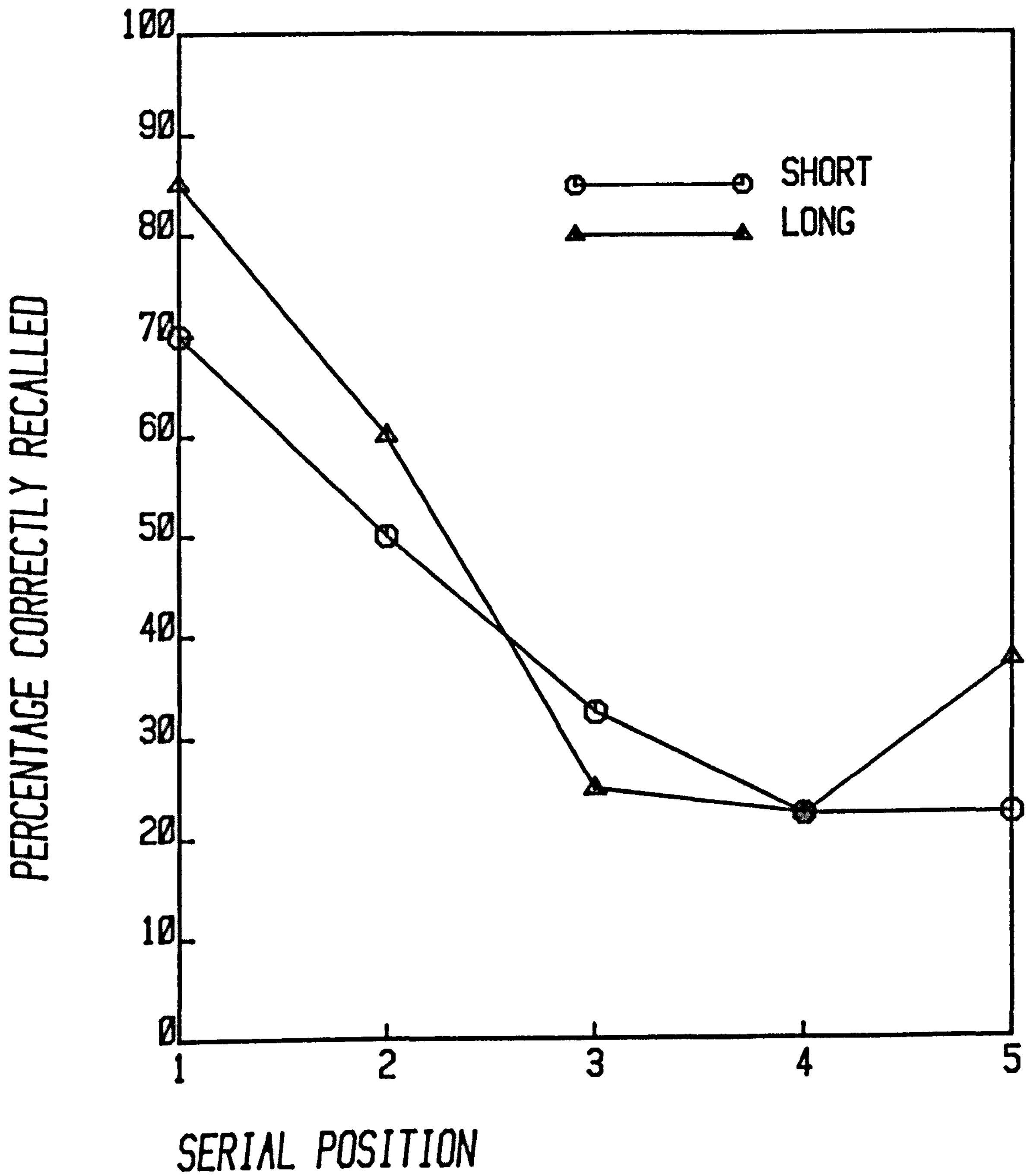


FIGURE 26
SERIAL POSITION CURVES FOR
VISUAL/SUPPRESSION CONDITION



presented visually with suppression and this poses serious problems for decay theory if the assumption that suppression stops rehearsal is accepted. The assumption is based on the idea that it is impossible to say two things at once and thus implies that rehearsal uses some of the same mechanisms as overt speech. Given the evidence for the use of a speech code in STM this assumption appears reasonable, although there is at present no independent evidence that it is correct.

Decay theory might be saved if it is assumed that recall of the list constitutes a single rehearsal of the material. Taken with the assumption that suppression stops a visual code being translated into a speech code the results can now be explained by decay theory. Recall of the items has to be performed serially and because the longer words take longer to recall, more of them will decay during recall than for the shorter words. The visually presented material would never enter STM and, therefore, could not be expected to show a word length effect.

Thus it is possible to retain the idea that suppression stops rehearsal and yet explain the results in terms of decay theory. This explanation could be tested if there was a means of equating recall time for both short and long words. One method would be to use paced recall but from the results of experiments 2 and 3 it is not very clear what subjects are doing in this paradigm. Another possibility is to have the subject use abbreviations of the long words in recall, devised so that they have the same spoken length as the short words. Experiments of this kind were conducted by

M. Buchanan at Stirling (unpublished undergraduate project) but the results were difficult to interpret as it was not clear at what stage subjects translated the words into their abbreviations. The evidence suggested that they did this at input rather than recall. Thus, until a satisfactory test of this idea can be devised the results cannot be taken as crucial for decay theory. There will be some further discussion of this issue in the next chapter.

Perhaps the most important aspect of these results is the fact that word length has again behaved in the same way as AS. This provides further support for the idea that the word length effect is mediated by STM and suggests that the capacity of STM must be measured in units of time rather than units of structure. The experimental work done for this thesis has now been reported and in the last chapter some of the issues raised by these results and their implications for models of serial recall will be considered.

GENERAL DISCUSSION

The purpose of this chapter is to bring together the results of the previous chapters and to discuss their theoretical implications. The first two sections are a precis of the results and the inferences drawn from them.

Summary of Experimental Results

(i) The spoken duration of material in a serial recall task predicts subjects' performance independently of the number of syllables or phonemes contained in the list.

(ii) Performance in a serial recall task is well described by the function NR = RR.k + c, where 'NR' is the number recalled, 'RR' is reading rate in words per second and 'k' and 'c' are constants.

(iii) The improvement in performance over trials that occurs with a multi-trial serial recall procedure is due solely to an increase in the intercept of the NR vs RR function, i.e. the component that is insensitive to word length.

(iv) If a rehearsal preventing task is interpolated between the presentation and recall of a list in a serial recall task, the slope of the NR vs RR function is greatly reduced, i.e. the size of the word length effect is reduced.

(v) The reduction in recall that results from increasing list length in a serial recall task is due to a reduction of the intercept of the NR vs RR function.

(vi) Articulatory suppression abolishes the word length effect when the material is presented visually but has very little effect when the material is presented auditorally.

These were the main results of the experiments reported in this thesis and before examining them further it will be useful to summarize the most important inferences that were made in the discussions of the experiments. These are numbered so as to correspond with the experimental results.

Inferences Drawn from the Results

(i) The chunking hypothesis fails to explain the data and in general produces incorrect predictions concerning the capacity of STM.

(ii) Decay theory accurately predicts the results and correctly suggests that the capacity of STM is constant in units of time rather than units of structure.

(iii) The intercept of the NR vs RR function is an estimate of the LTM component of serial recall.

(iv) The slope of the NR vs RR function is a measure of the STM component of serial recall.

(v) STM capacity is not affected by changes in list length while the LTM component of recall is reduced as list length increases.

(vi) Assuming:-

(a) That articulatory suppression prevents rehearsal, and

(b) That articulatory suppression prevents the translation of a visual code into a speech code,

then it was inferred:-

(a) That with visual presentation material does not enter STM when the subject has to suppress and hence there is no word length effect.

(b) That with auditory presentation the word length effect is due to the short words taking less time to recall than the long words and so having less opportunity to decay.

The results and the inferences drawn from them have been purposely presented in a very abbreviated form so that the main points are not lost in a mass of detail. However, the rather simple picture of STM that has been expounded for the purposes of providing a coherent structure to the thesis now has to be examined in more detail. It will be found that decay theory cannot cope with certain aspects of the data and an alternative explanation in which STM is conceived as a system for the storage of "motor programmes" will be considered.

Some Theoretical Implications

The explanation of the word length effect was that material in STM decays over time but can be refreshed by rehearsal. Consequently, the more items that can be rehearsed within the decay time, the more will be recalled. Assuming that rehearsal can be equated with speech, short words will take less time to rehearse than long words and so more of them can be maintained in STM.

On the positive side, this account is able to explain the linear relationship between word duration and serial recall and, unlike interference theories, predicts that the structural complexity of the material is irrelevant in serial recall. It also explains the loss of the word length effect with a distractor task as, by preventing rehearsal for ten seconds at the end of a list, the material will have decayed from STM.

On the negative side, decay theory provides no explanation of the serial position effects, cannot readily account for the slope changes between experiments and has some trouble with the data from experiment VIII where articulatory suppression was used. These issues will now be discussed.

The serial position effect occurs in both free and serial recall and in both paradigms it is widely assumed that primacy and recency are mediated by different stores or processes. However, as discussed in the introduction, there is now serious doubt as to whether recency in free recall may be ascribed to STM. For serial recall, it has been established that recency only occurs for auditorially presented material (the modality effect) although it is common to get some recency with visual presentation if the lists are short. Further, the recency effect in serial recall is limited to the last item compared to the last three items for free recall. The most widely accepted explanation of the recency effect in serial recall is that due to Crowder and Morton (1969) who proposed that material enters an auditory sensory store (PAS) prior to entry into STM. Material in this store is displaced by subsequent items and thus only the last list item benefits at recall as no other items follow it. Being a modality specific store, visual material does not enter PAS, rather it enters visual sensory memory which has a very fast decay and thus cannot contribute to performance in this paradigm. The data reported in this thesis are not directly relevant to this issue as no attempt was made to control the contribution of PAS to recall. The only point that can be made was discussed after Experiment VI in which it was tentatively suggested that the present results and those of Watkins et al. (1972) showed that PAS was insensitive to word length. Thus, the recency effect in serial recall may not be a problem for a decay theory of STM as it seems likely that it is mediated by a separate store.

There is one aspect of the present results that makes such a conclusion suspect. This is the small but reliable recency effect obtained in experiments using visual presentation (e.g. Experiment IV, fig. 7). This effect does not occur when longer lists are used (Crowder and Morton typically used nine item lists) and this could be because subjects do not try very hard to store the last few list items in STM with long lists as its capacity would be exceeded and they would lose the earlier items. Where list length is closer to their span, subjects do attempt to encode all items into STM and this produces the recency effect at short list lengths. If this is the case, it must be argued that STM does exhibit recency effects and it is not clear how decay theory could explain them. The basic problem being that decay theory predicts that recall is only a function of time since presentation or rehearsal, and, therefore, must explain the effect of any variable in these terms. However, it does not seem likely that the last list item spends less time in STM than other items: subjects have to recall the items in order and do not appear to "rush" through the list to get to the last item before it is lost. Even if they recalled at a rate similar to their rehearsal rate, the last item would be stored for the same length of time as all other items. Thus, it must be concluded that if STM does produce recency effects, they cannot be explained by decay theory. There remains the possibility that these recency effects are mediated by LTM but, if this were true, it is difficult to understand why they should only occur at short list lengths where it would be expected that subjects would rely more on STM than LTM. However, it is the case that when a distractor paradigm is used, with short list lengths, some recency occurs (e.g. Estes (1972) Experiment 1) and so it remains possible that LTM is mediating the serial position effects observed. Some support for this view might be derived from the serial position curves for the visual presentation condition in Experiment VIII. It was argued that, when subjects had to suppress, the items did not enter STM and this caused the word length effect to disappear. Thus, the suppression condition should reflect recall from LTM alone while the no suppression condition should produce recall from both STM and LTM. Because of ceiling effects, no analysis was performed on this data but it can be seen that in both conditions there is a strong primacy effect and some recency. To resolve this issue it would be necessary to repeat these two conditions with longer lists so that ceiling effects do not occur and examine the interaction of suppression with serial position.

As regards primacy, it is commonly accepted that in free recall this is mediated by LTM. Given the above arguments concerning the serial position effects in the visual suppression condition of Experiment VIII, it could be claimed that these data also support the view that primacy is mediated by LTM. However, it has not been shown that STM produces no primacy, but unlike recency it is possible that decay theory could account for it. If subjects cumulatively rehearse from the beginning of the list and continue to add items to their rehearsal string until they can add no more without losing early items, then recall would be highest for the first few serial positions. Because data are averaged over lists and over subjects, the characteristic decline in recall from the first item would be predicted. In common with explanations of a number of phenomena in terms of decay theory, this account relies on assumptions about the manner in which subjects rehearse. Given a lack of knowledge on this question, the explanation must remain speculative.

No firm conclusions can be drawn from this discussion of serial position effects. The presence of a recency effect is a problem for decay theory, although more research is needed before it could be said to be crucial. This state of affairs will recur throughout this discussion; no data will entirely discredit decay theory, but often it will be necessary to make untested, although not necessarily untestable, assumptions usually about the effect of a variable on rehearsal in order to save the theory.

The second problem the present data poses for decay theory is the changes in slope that occurred between experiments. Although the results of Experiment IV showed sizeable differences in the slope of the NR vs RR function for different subjects, this would not seem sufficient to account for the changes between experiments. It will be helpful at this point to tabulate the slopes obtained in the different experiments along with some aspects of the designs that would seem relevant:

Experiment	List	Modality	Rate (secs.)	Within/ between subs.	Intercept	Slope
IV	5	Visual	2	Within	0.17	1.87
V	8	Auditory	1.5	Between	1.26	1.08
VI	6	Auditory	1.2	Within	0.37	1.66
VII	5, 6, 7	Auditory	1.5	Within	0.66	1.02

In addition to these data it should be noted that there was no difference between the size of the word length effect for the visual and auditory no suppression conditions in Experiment VIII. This would seem to rule out presentation modality as an explanation. The results presented above seem to divide into two groups: Experiments IV and VI where the slope is around 1.7 seconds and Experiments V and VII where the slopes are near to

1 second. Neither presentation rate nor presentation modality seems able to account for this fact, nor is it evident that an interaction between these variables is important in determining slope. However, one possible explanation may be found by appealing to the idea that people vary in the amount they use STM as opposed to LTM in these tasks.

Where list length is short as in Experiments IV and VI, subjects attempt to store all the material in STM by relying strongly on rehearsal and hence a sizeable STM component arises. Where list length is considerably greater than span, subjects rely more on LTM and hence show a smaller slope.

This would explain why the slope obtained in Experiment V was lower as lists of eight words were used. However, Experiment VII was designed to test this hypothesis by using three different list lengths ranging from 5 to 7. The results showed that the slope was unaffected by list length and in magnitude resembled that obtained in Experiment V. However, it is possible that this result is an artifact of the design of Experiment VII. A within subjects design was employed and it is possible that range effects (Poulton, op. cit.) were present. In other words subjects may have chosen the strategy that produced optimal performance in the experiment as a whole rather than changing their strategy with changes in experimental conditions. If the changes in slope between the experiments are due to changes in the relative contributions of STM and LTM, then it should follow that the experiments producing the lower slopes should produce the higher intercept values. This is based on the assumption that the intercept is a measure of the LTM component. As can be seen from the table this turns out to be the case. The intercept value on trial 1 is reported for Experiment V as this is more comparable with the other results. For Experiment VII the value is the mean for the three list lengths: for list length 5 the value was 1.233 compared to 0.17 for experiment IV which also used list length 5. For list length 6 the value was 0.55 which is not much greater than 0.37 obtained in Experiment VI although the difference is in the right direction. For list length 7 the value was 0.208. Note that the drop in intercept with increases in list length does not invalidate the argument as it is assumed that a range effect occurred in this experiment causing subjects to rely on LTM to the same extent for all list lengths. Thus the change in the intercept does not imply a change in the utilization of LTM, but rather that list length causes a change in the level of recall from LTM possibly because of more inter-item interference. The change in the relative utilization of STM and LTM which it is claimed is occurring between the different experiments, is shown by a rise in one component (slope or intercept) and a corresponding fall in the other component.

A test of this explanation of the slope changes would be to repeat experiment VII using a between subjects design. It would be expected that as list length increased so the slope decreases, although the behaviour of the intercept would be less predictable, the increased use of LTM causing it to increase while the increased storage demands causing it to decrease. Should list length turn out to be the determinant of the slope changes, it would not necessarily be a problem for decay theory as it could be argued that when subjects choose to rely on STM for recall, they rehearse more quickly and/or reliably as they are not trying to perform more complex semantic encoding of the material.

The final and most serious problem for decay theory presented by these results comes from the last experiment which looked at the effects of suppression. In the discussion of the results it was argued that the presence of the word length effect with auditory presentation and suppression might have been due to the short words being recalled faster than the long words. However, this was rather a weak argument especially as the words were presented at a 2 second rate and, if no rehearsal could occur, and decay time is between 1 and 2 seconds, then only the last list item should have been recalled. The main problem in interpreting these results is the uncertainty about the effects of articulatory suppression, in particular whether or not it prevents subjects rehearsing. An experiment that needs to be done is to use an alternative method of preventing rehearsal, for example, using a fast presentation rate (cf. Estes, 1972). In the absence of such data, the most plausible suggestion must be that subjects cannot say two things at once and, therefore, articulatory suppression prevents rehearsal.

If this is true, then the results of Experiment VIII discredit a decay theory explanation of the word length effect and alternative explanations should be examined. Before leaving these results, it is worth noting that suppression, although not interacting with word length in the auditory condition, did still reduce performance. Tukey's HSD test for p = 0.01 produced a value of 5.91 while the difference between the means for the auditory suppression and no suppression conditions was 12.25 for for the short words and 8.0 for the long words. Thus it can be claimed that with auditory presentation suppression significantly reduces performance. The reasons for this are not clear. It could be that preventing rehearsal reduces the probability of items being encoded into LTM or, it could be that it is a very general effect of trying to do two things at once. The present data do not allow any decision on this point.

To summarize the arguments so far, it would seem that the decay theory explanation of the results although not totally discredited has some difficulty with certain aspects of the data. Before considering alternative explanations it will be useful to extend the discussion beyond the present data and examine the success or otherwise of decay theory in accounting for other data on serial recall. The most explicit and well developed decay theory is that proposed by Conrad (1965) and the data will be examined in the context of this model.

Conrad proposed that material was stored in an ordered array of "bins" and was subject to decay.

Because the material was in the form of a speech code, it was possible for some parts of words to decay faster than others. Thus the word "keep" after some period of decay might only remain as "ee". When this occurred, Conrad assumed that subjects searched through memory for a plausible candidate, obviously this would be a word containing the "ee" sound. If no candidate was found (as might occur in a list containing acoustically dissimilar words) then subjects reported that they could not remember the word. However, if the list contained another word with the "ee" sound then subjects were likely to report this in place of the correct one. This would give rise to an order error and this should be more frequent when different words in the list have similar speech codes, i.e. when the list is of high acoustic similarity. This prediction was confirmed by Conrad and Hull (1964) and by Conrad (1967) where it was shown that increasing the acoustic similarity of a list produced more order errors.

Conrad assumed that the role of rehearsal was to refresh the material and so stop decay and for this reason the theory would seem capable of explaining the word length effect. Unfortunately, there are some results which have been claimed to discredit Conrad's theory, most notable of which is an experiment by Posner and Konick (1966) using the Brown-Peterson distractor paradigm. They looked at performance with difficult and easy distractor tasks as well as with lists of high and low acoustic similarity. They assumed that a difficult distractor task would cause the trace to decay further than an easy one, as they had previously found that the level of performance in this paradigm was determined by the difficulty of the distractor task. They then compared the size of the AS effect over different retention intervals for the different difficulties of distractor task. They found that the AS effect increased with retention interval but was not at all affected by the difficulty of the distractor task.

This result is frequently taken as evidence against Conrad's theory (e.g. Crowder, 1976), but, given some later results, this position does not seem so viable. The main problem with the Posner and Konick experiment is that their retention intervals were 0, 5, 10 and 20 seconds and, from the results of Baddeley and Scott (1971) previously discussed, it seems likely that material in STM would have decayed completely after 5 seconds. In fact, the present results suggest that the decay time of an item is between one and two seconds and, consequently, it seems unlikely that Posner and Konick were looking at STM at all except for the zero second retention interval. Further, their experiment which established that performance in the Brown-Peterson paradigm was a function of distractor task difficulty also used retention intervals of over 5 seconds and, consequently, their results might be interpreted as showing that forgetting in LTM is a function of distractor task difficulty. However, there remains the problem of why they obtained the acoustic similarity effect at the long delays. It has been assumed that AS occurs within STM so it should not be possible for the STM component to be removed without also removing the AS effect. However, Conrad (1967) and Estes (1973) both found that a distractor task did remove the AS effect which directly contradicts Posner and Konick's results. In conclusion, it seems reasonable to claim that Posner and Konick's experiment cannot be accepted as crucial evidence against Conrad's theory in particular or decay theory in general.

Another set of results which poses a problem for any decay theory is that demonstrating the effects of temporal grouping (e.g. Ryan, 1969a). In these experiments serial recall lists were presented with sets of words temporally grouped. Wickelgren, who has done considerable work on this topic (Wickelgren, 1964; 1967), found that subjects perform best when the list items are grouped in threes. In an unpublished Ph.D. thesis Frankish found that subjects were extremely sensitive to the temporal structure of lists and when, in a regularly spaced list, the interval between two of the words was increased by as little as 80 msecs. subjects perceived the list as two groups of words. Further, he found that this perceptual phenomenon was reflected in performance as the serial position curve for the list took on the appearance of two separate serial position curves, one for each of the groups. Thus, there is ample evidence that subjects are very sensitive to the temporal structure of lists in serial recall tasks, but it is not easy to see how these results might be explained by Conrad's theory.

The reason is that the theory assumes that the probability of recalling an item is a simple function of the time between presentation or rehearsal and recall. Changing the temporal structure of the list should only improve performance if this time interval is decreased, but most experiments investigating temporal grouping have ensured that the overall time to present the list was the same over different temporal structures. The theory might be saved if it could be shown that temporal grouping alters subjects' rehearsal strategies in such a manner as to produce less decay of the list items and there is some evidence that suggests this might be the case.

Wilkes (1972) used a multi-trial procedure in which the list was grouped differently on each trial. He found, as had Bower and Winzenz (1969) that there was very little learning compared to the condition where the same grouping was preserved over trials. However, he also taped the subjects' reading and recall of the list and found that they grouped the words in accordance with the presented groupings. For the condition in which the presented groupings changed over trials, the subjects' spoken groupings changed accordingly and the deficit in learning was attributed to a lack of a consistent rehearsal strategy. This is based on the plausible assumption that the grouping used by the subjects at recall was the same as they used in rehearsal. The important point for the present purposes is that grouping was shown to alter subjects' rehearsal strategy. More evidence is needed before this may be taken as an explanation compatible with decay theory and, in particular, it needs to be shown that the most advantageous groupings of words (i.e. in threes) changes the rehearsal strategy so as to reduce the amount of decay undergone by each item. Some tentative support for this idea was obtained by myself in an informal study where subjects were asked to read lists of words which were either regularly spaced on the paper or grouped in threes. It was found that all eight subjects read the list faster when the words were grouped. This could, of course, have been due to the grouped words being more "readable" than the ungrouped list, e.g. it might be easier to keep one's place in the list when the material is grouped. However, if it did reflect the fact that subjects can rehearse lists faster when they are grouped in threes, then decay theory would predict that recall should be improved.

A further complication with the work on grouping is that it would seem that temporal grouping affects the LTM component of recall. In a recent paper Glanzer (1976) showed that the secondary memory component of free recall was sensitive to the temporal structure of the list and concluded that the manner in which a list was grouped determined the way it was encoded in LTM. The experiment by Wilkes (1972) discussed above also provides evidence that grouping affects LTM as he showed that irregular grouping depresses performance in a multi-trial paradigm. In conclusion, more work is necessary before the effects of temporal grouping can be understood and, in particular, it will be necessary to separate the effects of grouping on STM and LTM and to investigate the changes in rehearsal strategy and rehearsal rate produced by grouping. In the case of Conrad's model, it cannot be said to explain grouping phenomena at present but until more is known about grouping, it is not quite clear exactly what has to be explained.

The most commonly questioned prediction of Conrad's model is that item and order errors should be correlated. This arises from the assumption that order errors occur because of a loss of item information through decay. Conrad (1965) stated that:- "What would be crucial is a variable that, in the defined case, could be shown to affect order of items differentially from the items themselves." Murdock and Von Saal (1967) claimed that category membership was such a variable. They found that when words in a serial recall task were from the same conceptual category, the number of item errors was reduced, but the number of transposition errors increased compared to lists containing words from different categories. They argued that this result was inconsistent with Conrad's model as it must predict that the increase in order errors be accompanied by an increase in item errors.

This result certainly contradicts Conrad's model in its present form, but it is possible to develop the model slightly so that it may account for these data. The main weakness in Conrad's model is that no contribution to recall from LTM is assumed and, as discussed above, there is good reason for believing this to be the case. Indeed, if this assumption is accepted then an explanation of Murdock et al.'s data becomes possible. Category membership is evidently a semantic attribute of words and, to be consistent with the previous arguments, could not be encoded in STM. Thus, the improvement in recall with categorized lists must be due to an increase in the LTM component and the increase in order errors could be ascribed to LTM being poorer at encoding order information for linguistically unstructured material. It is also necessary to assume that subjects relied more heavily on LTM to the decrement of the STM component with the categorized lists, an assumption that might be readily tested using the word length effect to separate the STM and LTM components. The prediction could be that categorized lists should produce a shallower slope but a greater intercept compared to uncategorized lists when NR is plotted as a function of RR.

To summarize this discussion of other data on serial recall, it would seem that again, decay theory has not been completely discredited. However, it has been frequently necessary to appeal to changes in rehearsal strategies and/or changes in the relative utilization of STM and LTM. Decay theory although requiring such assumptions to remain viable cannot itself provide any explanation for them. In many respects the theory is shallow and provides only a low level of explanation of the data. For this reason and, for the doubt raised by the suppression data in particular, alternative explanations of the word length effect will be considered. It is not claimed that the following account is a developed theory, but only that it might suggest more interesting and fruitful lines of research than decay theory.

The alternative explanation to be proposed starts with a consideration of the function of STM rather than a statement about the nature of forgetting in the store. It was argued in the introduction that STM stored material in an articulatory code and it was suggested that it might be involved in the production of speech. The experiments reported in this thesis cannot be said to support this view over the alternative idea that STM stores acoustically encoded material; the effect of word duration could be due to the different spoken duration of the words or to their different acoustic lengths. However, Dr. P. Barnard at the Applied Psychology Unit, Cambridge, designed an experiment to investigate this question. He did this by making the words have the same acoustic length at presentation but still have different spoken lengths when uttered by the subjects. This was achieved by using a computer to artificially shorten the waveform of the long words. It is possible to remove repetitive pulses within a speech waveform and still retain the intelligibility of the signal. With this technique he was able to equate the acoustic length of words like "stoat" and "kangaroo", although, when spoken by the subject, they still took different amounts of time to say. His main result was that this procedure did not alter the magnitude of the word length effect and he concluded that the important variable was the time it took subjects to say the words and not their acoustic length.

This supports the idea that STM is part of the speech production system and accordingly employs an articulatory code. It is not consistent with the idea that STM employs an acoustic code because, if this were true, it should be sensitive to the acoustic length of the words rather than their spoken length. Thus, there is good reason to believe that STM stores articulatory commands and, given the arguments of Ellis (1976) discussed in the introduction, it is feasible that STM is a buffer store concerned with the storage of speech prior to its production.

If STM is ascribed this function, then it is clear that it must be capable not only of storing articulatory commands, but also their temporal structure. One major characteristic of speech is prosody, i.e. the temporal patterning of the articulatory units. This is not merely an artefact of speech, but rather is used as a means of conveying meaning and, consequently, is interrelated to the linguistic content of the message. Consequently, the timing and stress patterning of an utterance must be planned in conjunction with the planning of the words and the syntax and a store responsible for holding the utterance during production must be able to preserve this information.

From this point of view, it is not surprising that STM is sensitive to temporal grouping and word duration as one of the functions of the store is the preservation of prosodic information. Until more is known about the nature of the store, this view cannot provide an explanation of the serial recall data. In fact, it cannot be claimed to be a theory about the nature of STM, it only makes assumptions about its function. However, this might turn out to be a useful approach as the function a system performs will determine its nature and it would follow that the recent interest in speech prosody might well provide information relevant to STM.

Prior to this view, little concern was paid to the function that STM was supposed to perform (except the commonly stated view that it was used to remember telephone numbers) and, consequently, any theory that accounted for the serial recall data was acceptable. This approach produces very abstract theories which have little relevance outside the data that generated them. It is hoped that the present approach might produce theories that are of more general interest.

Such a theory was recently proposed by Baddeley and Hitch (1974) in which STM was ascribed an important role in many cognitive tasks such as reading, mental arithmentic and verbal reasoning. The details of the model are not relevant to the present discussion, but the proposed nature of STM should be mentioned as it is rather similar to the view outlined here.

Short-term memory was conceived as two systems, a working memory which was responsible for storage, retrieval and generally performed the role of a central processor, and an articulatory loop which was a small capacity store that was used by working memory for temporary storage. As the name implies, material on the articulatory loop was encoded in an articulatory form and it was maintained on the loop by means of rehearsal. The loop was likened to a tape loop revolving at a constant speed and this aspect allows an explanation of the word length effect. The amount that can be stored on a tape loop will be a simple function of the duration of the material. Consequently, more short than long words could be placed on the loop before previous items were overwritten. However, the model requires further development before it can explain other serial recall data such as serial position effects and temporal grouping phenomena.

To summarize this section, it seems that while better than most alternatives, decay theory does have problems in coping with the data on serial recall. It was suggested that a better approach might be to consider the function of STM, and the possibility that it is involved in the production of speech was considered.

Some General Implications

One of the most important results of this research is the ability to describe serial recall performance as a very simple function of subjects' reading rate or speech rate. Apart from the consequences for theories of STM, the fact that the equation contains two parameters, the slope and intercept, implies that there are two distinct components underlying serial recall. It is possible that many of the difficulties that have occurred in trying to understand the effect of a variable on serial recall (e.g. rate of presentation) are a product of the variable affecting the two components in different ways. Word length provides a means of separating these two components and might allow some of the anomalies to be resolved. A related point is that the data are in good agreement with the view that both STM and LTM contribute to serial recall and might provide a means of investigating the way these stores interact which, at present, is not well understood.

A second consequence of the results is that they suggest a novel explanation of why memory span increases with age throughout childhood (e.g. Belmont and Butterfield, 1969). Rather than proposing that there is some structural change in the capacity of STM, it is possible that the increases in memory span are due to an increase in speech rate. Some recent data by Nicholson (1979) supports this suggestion. He showed, using exactly the same design and material as Experiment 4 of this thesis, that children's scores were lower than adults', but that when plotted against reading rate fell on exactly the same line as the adult scores. This implies that the capacity of STM is exactly the same for children and adults but that the latter perform better because they can rehearse faster. If this result is confirmed, it would appear to have implications beyond the immediate problems of devising a model of serial recall. As mentioned after Experiment 4, it might be the case that improving a child's speech fluency could cause his performance to improve on a number of tasks including some intelligence tests.

Conclusions

The main conclusion from this research is that the capacity of STM is constant when measured in units of time. This result is predicted by decay theory of forgetting and by an articulatory loop model of STM. However, neither model can cope with other data on serial recall and it was suggested that it might be more useful to consider the function STM performs especially as there is considerable evidence suggesting that it is a speech output buffer.

REFERENCES

Aaronson, D. (1967) Temporal factors in perception and short-term memory. _Psychological Bulletin_, _73_, 130-144.

Atkinson, R.C. and Shiffrin, R.M. (1968) Human Memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.) _The Psychology of Learning and Motivation_, Vol. 2. New York: Academic Press, 89-105.

Baddeley, A.D. (1966a) The influence of acoustic and semantic similarity on long-term memory for word sequences. _Quarterly Journal of Experimental Psychology_, _18_, 302-309.

Baddeley, A.D. (1966b) Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. _Quarterly Journal of Experimental Psychology_, _18_, 362-365.

Baddeley, A.D. (1972) Retrieval rules and semantic coding in short-term memory. _Psychological Bulletin_, _78_, 379-385.

Baddeley, A.D. (1979) The trouble with levels: A re-examination of Craik and Lockhart's framework for memory research. _Psychological Review_, _85_, 139-152.

Baddeley, A.D. and Hitch, G. (1974) Working Memory. In G. H. Bower (Ed.) _The Psychology of Learning and Motivation_, Vol. 8, 47-90.

Baddeley, A.D. and Hitch, G. (1977) Recency re-examined. In S. Dornic (Ed.) _Attention and Performance VI_. Hillsdale N.J.: Erlbaum.

Baddeley, A.D. and Levy, B.A. (1971) Semantic Coding and Memory. _Journal of Experimental Psychology_, _89_, 132-136

Baddeley, A.D. and Scott, D. (1971) Short-term forgetting in the absence of proactive inhibition. _Quarterly Journal of Experimental Psychology_, _23_, 275-283.

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

Belmont, J.M. and Butterfield, E.C. (1969) The relations of short-term memory to development of intelligence. In L. Lipsitt and H. Reese (Eds.) Advances in Child Development and Behavior. New York: Academic Press, Vol. 4.

Bower, G.H. and Winzenz, D. (1969) Group structure, coding and memory for digit series. Journal of Experimental Psychology Monograph Supplement, 80,(Pt. 2) 1-17.

Broadbent, D.E. (1954) The role of auditory localization in attention and memory span. Journal of Experimental Psychology, 47, 285-293.

Brown, J. (1959) Information, redundancy and decay of the memory trace. In: The mechanization of thought processes. National Physical Laboratory Symposium, No. 10. London: H.M.S.O.

Brown, J. (1958) Some tests of the decay theory of immediate memory. Quarterly Journal of Experimental Psychology, 10 12-21.

Cermack, L.S. (1972) Human Memory. New York: Ronald Press.

Colle, H.A. and Welsh, A. (1976) Acoustic masking in primary memory. Journal of Verbal Learning and Verbal Behavior, 15, 17-31.

Conrad, R. (1964) Acoustic confusions in immediate memory. British Journal of Psychology, 55, 75-84.

Conrad, R. (1965) Order errors in immediate recall of sequences. Journal of Verbal Learning and Verbal Behavior, 4, 161-169.

Conrad, R. (1967) Interference or decay over short retention intervals. Journal of verbal Learning and Verbal Behavior, 6, 49-54.

Conrad, R. (1970) Short-term memory processes in the deaf. British Journal of Psychology, 61, 179-195.

Conrad, R. and Hull, A.J. (1964) Information acoustic confusion and memory span. British Journal of Psychology, 55, 429-432.

Conrad, R. and Hille, B.A. (1958) The decay theory of immediate memory and paced recall. Canadian Journal of Psychology, 12, 1-6.

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

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

Craik, F.I.M. and Lockhart, R.S. (1972) Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Crannell, C.W. and Parrish, J.M. (1957) A comparison of immediate memory span for digits, letters and words. Journal of Psychology, 44, 319-327.

Crowder, R.G. (1976) Principles of Learning and Memory, Hillsdale, N.J. LEA.

Crowder, R.G. and Morton, J. (1969) Precategorical acoustic storage (PAS). Perception and Psychophysics, 5, 365-373.

Ellis, A.W. (1976) The role of a phonemic response buffer in speech and short-term memory. Paper presented to the Autumn Conference of the Institute of Acoustics, Edinburgh.

Estes, W.K. (1972) An associative basis for coding and organization in memory. In A.W. Melton and E. Martin (Eds.) Coding Processes in Human Memory. Washington, D.C.: Winston. 161-190.

Estes, W.K. (1973) Phonemic coding and rehearsal in short-term memory. Journal of Verbal Learning and Verbal Behavior, 12, 360-372.

Glanzer, M. (1976) Intonation grouping and related words in free recall. Journal of Verbal Learning and Verbal Behavior, 15, 85-92.

Glanzer, M., Koppenaal, L. and Nelson, R. (1972) Effects of relations between words on short-term storage and long-term storage. Journal of Verbal Learning and Verbal Behavior, 11, 403-416.

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

Goldstein, E.F. (1975) Selective phonemic and semantic coding in short-term recall. Memory and Cognition, 3, 619-626.

Hebb, D.O. (1961) Distinctive features of learning in the higher animal. In J.F. Delafresnage (Ed.) Brain Mechanisms and Learning. New York: Oxford University Press.

Hintzman, D.L. (1967) Articulatory coding in short-term memory. Journal of Verbal Learning and Verbal Behavior, 6, 312-316.

Huggins, S.W.F. (1964) Distortion of the temporal pattern of speech: interruption and alternation. Journal of the Acoustical Society of America, 36, 1055-1064.

Huttenlocher, J. and Newcombe, N. (1976) Semantic effects on ordered recall. Journal of Verbal Learning and Verbal

Behavior, 15, 387-399.

Keppel, G. and Underwood, B.J. (1962) Proactive inhibition in short-term retention of single items. Journal of Verbal Learning and Verbal Behavior, 1, 153-161.

Klein, G.A. and Klein, H.A. (1974) The influence of serial retention and theme identification paradigms on encoding. Quarterly Journal of Experimental Psychology, 26, 556-560.

Lashley, K.S. (1951) The Problem of Serial Order in Behavior. In L.A. Jeffress(Ed.) Cerebral Mechanisms in Behavior. New York: Wiley, 112-136.

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

MacNulty, J.A. (1966) The measurement of adopted chunks in free recall learning. Psychonomic Science, 4, 71-72.

Marcel, A.J. (1979) Unconscious reading: Experiments on people who do not know that they are reading. Visible Language, in press.

Melton, A.W. (1963) Implications of short-term memory for a general theory of memory. Journal of Verbal Learning and Verbal Behavior, 2, 1-21.

Miller, G.A. (1956) The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Miller, G.A. and Nicely, P.E. (1955) An analysis of perceptual confusions among some English consonants. Journal of the Acoustical Society of America, 27, 338-352.

Miller, G.A. and Selfridge, J.A. (1950) Verbal context and the recall of meaningfull material. American Journal of Psychology, 63, 176-185.

Morton, J. (1970) A functional model for memory. In D.A. Norman, (Ed.) Models of Human Memory. New York: Academic Press. 203-254.

Murdock, B.B. Jr. and vom Saal, W. (1967) Transpositions in short-term memory. Journal of Experimental Psychology, 74, 137-143.

Murray, D.J. (1968) Articulation and acoustic confusability in short-term memory. Journal of Experimental Psychology, 78, 679-684.

Nelson, D.C., Wheeler, J.W., Bordan, R.C. and Brooks, D.M. (1974) Levels of processing and cueing: Sensory versus meaning features. Journal of Experimental Psychology, 103, 971-977.

Nelson, T.O. (1977) Repetition and depth of processing. Journal of Verbal Learning and Verbal Behavior, 16, 151-172.

Nicholson, R. (1979) The relationship between memory span and processing speed. To appear in M. Friedman, J.P. Das and N. O'Connor, Intelligence and Learning. Plenum.

Peterson, L.R. and Peterson, M.J. (1959) Short-term retention of individual verbal items. Journal of Experimental Psychology, 58, 193-198.

Posner, M.I. (1964) Rate of presentation and order of recall in immediate memory. British Journal of Psychology, 55, 303-306.

Posner, M.I. and Konick, A.W. (1966) On the role of interference in short-term retention. Journal of Experimental Psychology, 72, 221-231.

Posner, M.I. and Rossman, E. (1965) Effect of size and location of informational transforms upon short-term retention. _Journal of Experimental Psychology_, _70_, 496-505.

Poulton, E.C. (1973) Unwanted range effects from using within-subject experimental designs. _Psychological Bulletin_, _80_, 113-121.

Reitman, J.S. (1971) Mechanisms of forgetting in short-term memory. _Cognitive Psychology_, _2_, 185-195.

Reitman, J.S. (1974) Without surreptitious rehearsal, information in short-term memory decays. _Journal of Verbal Learning and Verbal Behavior_, _13_, 365-377.

Richardson, J.T.E. and Baddeley, A.D. (1975) The effect of articulatory suppression in free recall. _Journal of Verbal Learning and Verbal Behavior_, _14_, 623-629.

Ryan, J.A. (1969) Grouping and short-term memory: different means and patterns of grouping. _Quarterly Journal of Experimental Psychology_, _21_, 137-147.

Salame, P. and Wittersheim, G. (1978) Selective noise disturbance of the information input in short-term memory. _Quarterly Journal of Experimental Psychology_, _30_, 693-704.

Shulman, H.G. (1971) Similarity effects in short-term memory. _Psychological Bulletin_, _75_, 399-415.

Sperling, G. (1960) The information available in brief visual presentations. _Psychological Monographs_, whole No. 498.

Tulving, E. (1966) Subjective organization and effects of repetition in multi-trial free recall learning. _Journal of Verbal Learning and Verbal Behavior_, _5_, 193-197.

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

Watkins, M.J. and Watkins, O.C. (1973) The post-categorical status of the modality effect in serial recall. Journal of Experimental Psychology, 99, 226-230.

Watkins, M.J., Watkins, O.C. and Crowder, R.G. (1974) The modality effect in free and serial recall as a function of phonological similarity. Journal of Verbal Learning and Verbal Behavior, 13, 430-447.

Waugh, N.C. and Norman, D.A. (1965) Primary memory. Psychological Review, 72, 89-104.

Wetherick, N.E. (1975) The role of semantic information in short-term memory. Journal of Verbal Learning and Verbal Behavior, 14, 471-480.

Wickelgren, W.A. (1964) Size of rehearsal group and short-term memory. Journal of Experimental Psychology, 68, 413-419.

Wickelgren, W.A. (1965) Acoustic similarity and intrusion errors in short-term memory. Journal of Experimental Psychology, 70, 102-108.

Wickelgren, W.A. (1967) Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. Quarterly Journal of Experimental Psychology, 19, 97-102.

Wickelgren, W.A. (1969) Auditory or articulatory coding in verbal short-term memory. Psychological Review, 76 232-235.

Wickens, D.D. (1970) Encoding categories of words: an empirical approach to meaning. Psychological Review, 77 1-15.

Wickens, D.D. (1972) Characteristics of word encoding. In A.W. Melton and E. Martin (Eds.) Coding Processes in Human Memory. Washington D.C.: Winston, 191-215.

Wickens, D.D., Born, D.G. and Allen, C.K. (1963) Proactive inhibition and item similarity in short-term memory. Journal of Verbal Learning and Verbal Behavior, 2, 440-445.

Wilkes, A.L. (1972) Reading pauses during serial list learning with fixed or randomly changing groups. Journal of Experimental Psychology, 94, 206-209.

Wingfield, A. and Byrnes, D.L. (1972) Decay of information in short-term memory. Science, 176, 690-692.

Worthington, A.G. (1964) Differential rates of dark adaptation to "taboo" and "neutral" stimuli. Canadian Journal of Psychology, 18, 257-268.

APPENDICES

APPENDIX 1

RAW DATA FOR EXPERIMENT IV

Mean Memory Score and Reading Rate

For Each Subject In Each Condition

	1 Syllable		2 Syllable		3 Syllable		4 Syllable		5 Syllable	
Subject	MS	RR	MS	RR	MS	RR	MS	RR	MS	RR
1	4.5	2.38	4.2	2.27	4.0	2.00	4.3	1.79	3.3	1.29
2	4.5	2.50	4.7	2.38	3.5	2.17	3.1	1.92	2.7	1.67
3	4.6	1.79	4.0	1.67	2.5	1.52	2.5	1.35	2.4	1.14
4	3.4	2.17	2.9	1.67	1.6	1.67	1.4	1.52	1.4	1.22
5	4.9	2.17	4.0	2.08	3.0	1.79	1.4	1.61	1.2	1.22
6	4.8	2.38	5.0	2.00	4.0	1.79	4.0	1.47	2.2	1.11
7	4.8	2.38	4.5	2.27	4.3	2.00	3.1	1.79	3.3	1.25
8	4.8	2.38	4.9	2.17	4.5	1.85	4.7	1.61	3.7	1.29
9	4.6	2.00	3.9	1.92	4.3	1.67	3.6	1.35	2.5	1.09
10	3.4	1.85	3.1	1.79	2.1	1.61	2.8	1.35	2.4	1.00
11	4.9	3.13	5.0	2.78	4.8	2.50	4.3	1.92	3.5	1.56
12	4.5	2.17	4.0	1.92	3.1	1.85	3.0	1.56	2.3	1.22
13	4.1	1.79	3.2	1.72	2.9	1.56	2.9	1.29	2.2	1.14
14	4.8	2.94	4.9	2.78	4.6	2.50	4.0	2.17	3.0	1.56
Means	4.47	2.29	4.16	2.10	3.51	1.89	3.22	1.62	2.58	1.27

Maximum Memory Score = 5

Key

MS = Memory Score
RR = Reading Rate

APPENDIX 2

RAW DATA FOR EXPERIMENT VIII

No. Correct (Max. = 25)

SUBJECTS	SHORT WORDS				LONG WORDS			
	Auditory		Visual		Auditory		Visual	
	SUPP.	NO SUPP.	SUPP.	NO SUPP.	SUPP.	NO SUPP.	SUPP.	NO SUPP.
S1	19	17	6	18	15	21	9	15
S2	19	20	7	18	13	18	15	15
S3	22	25	8	22	15	10	12	13
S4	17	21	8	24	19	21	12	11
S5	14	24	5	18	13	13	15	16
S6	18	20	15	24	8	16	6	16
S7	15	19	11	18	7	12	11	12
S8	22	23	17	24	24	24	12	20
S9	20	21	9	18	17	8	10	11
S10	13	20	10	18	16	12	11	14
S11	18	22	4	20	9	11	6	8
S12	13	25	9	21	12	15	10	12
S13	21	16	10	14	8	11	12	12
S14	19	19	10	12	17	25	15	25
S15	16	20	18	17	14	16	13	21
S16	21	24	10	18	8	14	11	17
Mean % Correct	71.75	84.00	39.25	76.00	53.75	61.75	45.00	59.50