Word Length Effects in Long-term Memory

Gerald Tehan

University of Southern Queensland

Toowoomba, Queensland, Australia

Georgina Ann Tolan

Australian Catholic University – McCauley Campus

Brisbane, Queensland, Australia

Mailing Address :
Gerry Tehan
Department of Psychology
University of Southern Queensland
Toowoomba, 4350
AUSTRALIA
Phone number : (07) 4631 2376
e-mail : tehan@usq.edu.au

Abstract

The word length effect has been a central feature of theorising about immediate memory. The notion that short-term memory traces rapidly decay unless refreshed by rehearsal is based primarily upon the finding that serial recall for short words is better than that for long words. The decay account of the word length effect has come under pressure in recent times. The current research tests alternative explanations of the word length effect, ones that suggest firstly, that word length effects should be found in long-term memory tasks. Secondly, one alternative predicts that the standard short-word advantage should only be observed in tasks in which participants use serial order memory. In tasks where only item memory is used, the approach predicts a long word advantage. Consistent with this notion, standard word length effects were found in long-term serial recall and free recall tasks, but a long word advantage was found in item recognition and cued recall tasks.

Lists containing short words are better recalled in immediate serial recall than lists containing long words (Baddeley, Thompson & Buchanan, 1975). This word length effect has been instrumental in the assumption that short-term forgetting is due to decay and consequently has become a cornerstone of trace decay plus rehearsal models of short-term memory (Brown & Hulme, 1995). The word length effect represents a key feature that most current formal models of the immediate serial recall task attempt to explain. Consequently, the word length effect is seen very much as a crucial aspect of theorising about short-term memory recall.

The current research is motivated by a number of concerns. Firstly, recent research suggests that the decay explanation of the word length effect has become less viable (Nairne, 2002; Hulme, Surprenant, Bireta, Stuart & Neath., 2004). Secondly, while these other accounts of the word length effect exist, none are widely accepted in the same way that the decay account has been. Thirdly, the locus of these alternative explanations is by and large still couched in the short-term memory domain. Lastly, and most specifically, the current experiments test one alternative explanation of the word length effect that predicts that word length effects should be found in long-term memory tasks as well as short-term tasks. That is, word length effects are not a defining characteristic of short-term recall, but reflect differential processing that is independent of retention interval.

Three accounts of the word length effect.

Current accounts of the word length effect fall into two broad camps. The first is a time based account and the second is a complexity/discrimination account. The intention here is to introduce a third account; an item order trade-off account.

The basis of the time based account is that short-term memory traces decay rapidly at a fixed rate, such that after about 2 seconds the trace is no longer reliable enough to support recall. However, the trace can be refreshed by rehearsal. Since forgetting is occurring in real time, the faster items can be rehearsed the greater the chance of the decaying trace being refreshed and updated. Given that long words take longer to rehearse than short words, long words are going to suffer more decay in a given time frame than short words. In short, the word length effect emerges as a result of a trade off between rehearsal speed and trace decay. Given the supposedly transient nature of the memory trace in the absence of rehearsal, a clear prediction from this perspective is that word length effects should not be present in long-term memory tasks.

The second general account of the word length effect is a complexity account (Caplan & Waters, 1994, Neath & Nairne, 1995). The notion here is that long words are more complex than short words and this has implications for memory. Neath and Nairne’s (1995) adaptation of the feature model (Nairne, 1990), explains the word length effect in terms of problems in feature assembly. That is, complex (long) words have more features to assemble at output than short words and as such there is a greater probability of an assembly error. Thus, from this perspective the word length effect is not attributable to decay, instead it is attributable to a form of interference.

Recently, Hulme et al. (2004) have combined the complexity argument with a discrimination account of memory. The discrimination dimension is derived from the SIMPLE model of memory, which asserts that items are located in a multidimensional psychological space. An item that is located in sparse areas of that space is considered to be more distinctive compared an item that is surrounded by other items in more dense regions of that psychological space. Increased distinctiveness is said to lead to better memory for that item and increased ease of item identification. In their treatment of the word length effect they utilise two dimensions in psychological space. First there is the dimension of temporal distinctiveness which is a function of item presentation rate and retention interval. For present purposes, the second dimension is more important and that dimension involves item discriminability. The assumption made here is that short, simple items are more distinctive because they have less complex phonological representations. Long words, because they are more complex, are less distinctive. In modelling word length effects, temporal distinctiveness parameters were equivalent for short and long words with the word length effect emerging as a function of changes in item distinctiveness.

The distinctiveness account is relevant to current concerns in that the assumptions underlying SIMPLE apply to the long-term memory domain as much as they do to the short-term domain. That is, the temporal dimension applies just as much over long retention intervals as over short retention intervals. Consequently, the model could well apply to long-term tests of memory. Furthermore, if distinctiveness is an attribute of the item, which it appears to be in this model, then short words should still be more discriminable in this multidimensional space than long words. If so, it seems reasonable to predict that standard word length effects should be apparent in long-term memory tasks.

The third account of the word length effect assumes that word length effects are an emergent property of the type of processing that occurs at study (Hendry & Tehan, 2005). At the heart of this explanation is the assumption that item and order information can be processed separately. Under some circumstances item processing and order processing can trade off and this produces dissociations on tasks that are assumed to be differentially sensitive to item or order information. Such assumptions have been used to explain the generation effect (Nairne, Reigler & Serra, 1991), the perceptual interference effect (Mulligan, 1999), the word frequency effect (DeLosh &

McDaniel, 1996; Hockley & Cristi, 1996) and the enactment effect (Engelkamp & Dehn, 2000) among others. Hendry and Tehan (2005) argued that the word length effect could be explained using these same principles.

The seminal demonstration of the item order processing trade off can be seen in a study by Nairne, Reigler and Serra (1991) that involved the generation effect. They presented participants with 24 trials each consisting of eight unique items. On half of the trials eight word fragments were presented on each trial; a fragment being a complete word with only one letter missing, e.g. umbr_lla. Participants were required to generate a word for each fragment. The remaining trials were presented with intact words that were read aloud. Each list was followed by a 30 second distractor activity before participants were asked to respond. A recall cue specified one of two recall options. If a line of asterisks was presented, participants were instructed to not respond and to prepare for the next trial. However, on the serial recall trials, the items from the list were presented in a random order and participants were requested to put these words into the original presentation order. This order reconstruction task was thus seen as the index of order information and results indicated that the read words were better recalled in order than the generated words. Following this phase of the experiment, a surprise single-item recognition test was given to participants to test for memory of the items that had been presented earlier on the asterisk trials. The recognition test comprised both target and distractor items and this task was assumed to be a measure of item information. On this test, the generated items were better recognised than the read items. These results clearly demonstrated that the generate / read manipulation dissociated on item and order tasks. The explanation that has subsequently been proposed for this dissociation is that the generated items receive more item-specific processing at the expense of order processing and that the read items receive more order processing at the expense of item processing.

Hendry and Tehan (2005) applied the same logic and a similar experimental procedure to the word length effect. They argued that long words take longer or are harder to process at the item level than short words. This results in good item memory for the long words but poor order memory because cognitive resources are focused on processing the item itself and not on encoding the order of the items. The relatively weak order processing of the long words should show up on tasks that utilise order information such as immediate serial recall, but the approach predicts that on tasks that require item information, long words should be better remembered than short words. Thus the approach predicts a double dissociation of word length across different types of memory tasks.

Hendry and Tehan adapted the Nairne et al. (1991) procedure by presenting participants with lists of short or long words with each list ending in a row of question marks or a row of asterisks. In this initial phase of the experiment, when the trial ended in question marks, participants were asked to recall the items in serial order. When the trials ended in an asterisk, participants were told not to make a response but to wait for the next trial to begin. On the serial recall trials the standard short word advantage emerged. In a second phase of the experiment, the items that had appeared on the asterisk trials were again presented intermixed with distractors that had not appeared anywhere else in the experiment, and participants were asked to recognise the items from the earlier list. On this long-term test reverse word length effect appeared, with long words being better recognised than short words. This dissociation of word length across order and item tasks is precisely what the item-order explanation predicts and no other theory does.

The following experiments explore word length effects in tasks that have traditionally been associated with long-term recall. We believe that the three accounts make different predictions about the form of word length effects in these tasks. From the decay perspective, one would predict that word length effects should not be present given the absence of maintenance rehearsal over substantial retention intervals. If our ideas of how the complexity/discrimination approach might be extended to the long-term domain are correct, then one might expect a short word advantage across all tasks. Likewise, a prediction of the item order approach is that word length effects should be found in standard long-term episodic memory tasks. Furthermore, word length should dissociate on the extent to which the long-term tasks require order memory or item memory. That is, the assumption being made is that if the task requires, or participants use, order information, short words should be better remembered than long words. However, if the task is one that relies only upon item information, a long word advantage should emerge.

In order to test the dissociation assumptions long-term tests of order memory and long-term tests of item memory are required. The traditional choice for a LTM order memory task is a serial list learning task, where a supraspan list is repeated across trials and the participant’s task it to recall that list in correct serial order to a criterion of where all the items in the list are recalled in their correct serial position. Our second choice is a long-term free recall task. While this task does not require serial order memory, it is still the case that participants often use order memory to recall the items. Our expectation is that a short word advantage will be found in these tasks.

The traditional choice for measuring item memory is a single item recognition task. The second item task is a cued recall task. In both cases it is assumed that memory for order is not required or likely to be used. As such, the expectation is that a long word advantage will be observed.

Experiment 1

Experiment 1 explores word length effects in a long-term serial list learning task. Participants were presented with three supra span lists of items; a list of 8 filler items in order for people to get to learn how to do the task, followed by a list of 8 short words and a list of 8 long words, the latter two lists being counterbalanced across participants. The task was to recall the 8 items in their correct serial position, and participants were given a series of trials in which to achieve this goal. Thus, the items on each trial were presented on the first occasion and participants attempted to recall the items in order. If unsuccessful, the list was again presented in the same order and participants then attempted to recall all items in correct order. These trials continued until success completion of recalling the eight items in order was achieved.

From the item / order perspective, one would predict that people would reach criterion on the short words sooner than the long words, since order encoding for the short words is assumed to be superior to that of the long words. Moreover, the short word advantage should be apparent across the learning trials.

Participants

Twenty students from the Australian Catholic University participated in this experiment for course credit.

Materials

Experiments 1, 3 and 4 in the current paper are based upon two word pools. These pools were constructed by selecting 120 short words and 120 long words from the Oxford Psycholinguistic database (Quinlan, 1992). The short words were all monosyllabic words consisting of three phonemes (e.g., mane, dad, vase, meal, patch) whereas the long words were two or three syllables in length containing six or seven phonemes (e.g., wholesaler, magazine, algebra, envelope, triangle). All items were low frequency and high imagery words. According to the Kucera and Francis (1967) norms the mean word frequency rating for short words was 7.79 (SD = 10.34) and for long words 8.22 (SD = 8.14). The mean imagery rating for short words was 550 (SD = 52.71) and long words 564 (SD = 64.64).

In the current experiment, for each participant 8 short words and 8 long words were randomly selected from the pools and then randomly assigned to the 8 serial positions to form a short and a long list. In addition, a list of 8 filler words were also selected and randomly allocated to the 8 serial positions. All participants saw the same filler list, but learned unique short and long lists.

Pilot testing indicated that during learning of the first list people often learned how to do the task and that if just short and long word trials were studied the word length effect was contaminated by a practice effect. Consequently, the filler trial was employed so that people could learn the task and establish whatever strategy they found to be effective.

Procedure

Each participant was presented with three 8-item trials to be learned in serial order. The procedure was identical for each lists. All trials began with a READY signal, which remained on the screen for a three second period. Then each word appeared in the middle of the computer screen at a rate of one word per second. At the end of the trial a row of question marks appeared on the screen and participants attempted to recall the 8 items in correct serial order. The experimenter recorded each participant's responses on a hard copy of their experimental trials. If the participant did not recall all the items correctly, then the study trial was again presented in the same way. If the participant did recall the sequence correctly, that list was discontinued and the second list was presented in the same manner. For each list, participants were given ten trials to reach criterion. If they did not, then training on that list was discontinued. All participants received the filler list first, and then half the participants received the list of short words first and the remainder received the list of long words first.

Results and Discussion

For the short lists, all participants reached the criterion of one correct recall within the 10 trials. For the long list there were 2 people who did not reach criterion. However, for scoring purposes we acted as if these people had reached criterion on the final trial. On average it took 4.70 (SD = 1.56) trials to reach criterion for the short words, and 5.80 (SD = 1.93) trials to reach criterion for the long words. This difference was reliable, t (19) = 2.46 at an alpha level of .05 which is used in all subsequent analyses in all experiments.

To examine performance over the first few trials we limited our analyses to the first three trials, because we had full data for all 20 participants. That is, it was at a trial three that we had the first participant reach criterion on one of the lists. The data are summarised in Figure 1. Short words were better recalled than long words across all three trials, F (1,19) = 12.81, MSE = 1.89, performance increased across trials, F (1,19) = 69.01, MSE = .83, but the interaction was not reliable, F (1,19) = .82, MSE = .85.

The standard word length effect where short words are better remembered than long words has primarily been explored in short-term serial recall tasks. We show that the standard word length effect can be observed also in a serial list learning task where the number of words to be recalled is supraspan. Word length effects are not limited to short-term tasks.

The outcomes are consistent with the complexity and item-order accounts. In the introduction we indicated that a pure trace decay model would predict that word length effects should not be apparent. However, it is possible that hybrid models of the phonological loop like that developed by Burgess and Hitch (1996) could account for the results. In the Burgess and Hitch model there are two types of associations formed between their context window and the list items. The first is a rapidly decaying one that characterises the forgetting characteristics of the phonological loop. The second is weaker but more durable and is formed when a previously recalled item is presented for a second or third time in an experimental session. It is this component of the model that is responsible for producing the Hebb repetition effect. Given that this component only comes into play when an item has successfully been recalled on an earlier list, it stands to reason that if more short words are recalled on the first list, this short word advantage should be maintained upon any repeated list. While the Burgess and Hitch model may account for the data, we would argue that it is not the rapid trace decay component of the model that is producing short word advantage on the latter trials.

Experiment 2

Recently, word length effects (short word advantage) have been found in long-term free recall tasks (Russo & Grammatopoulou, 2003). At first glance this finding is problematic for the item-order perspective in that free recall is not typically viewed as a task that requires serial recall. If anything it is seen as an item memory task which should produce a null or reverse word length effect. From the item and order perspective the standard word length effect can only occur in free recall if participants are using a serial recall strategy even if the task instructions do not specifically request it. This presents the problem of how one measures the extent to which serial recall is being utilised in the free recall task.

Postman (1972) and Murdock (1968) argued that as list length decreases, participants are more likely to adopt a forward serial recall strategy and that it is only with longer lists that this strategy is abandoned in favour of a recency first protocol. However, even with long lists it is clear that serial information can be used in the free recall task (Kahana, 1996, Bhatarah, Ward, & Tan, 2006).

Measures of order memory in free recall are problematic. Here we use two measures; one based upon Kahana’s (1996) work, and the frequently used Asch-Ebenholz (1962) index.

Kahana (1996) examined free recall output protocols of long lists by dividing the output into pairs of items. Thus if 12 items were recalled, these 12 items can be considered as 11 pairs. He examined the relationship between the two items in each pair and found the second item of each pair was more likely to have occupied a nearby serial position than a remote position and that, forward recalls of near neighbours were more likely than backward recall of near neighbours. A constrained variant of the Kahana (1996) approach appears to provide a good index of order memory. One could examine the output and determine the frequency with which pairs that appeared in adjacent positions on the study list were recalled together.

The most widely used, and least restrictive, measure is the Asch-Ebenholz (1962) index. It is the one typically used when exploring order effects in generation and perceptual interference studies using free recall (Mulligan, 1999; Nairne et al., 1991). The measure addresses the extent to which relative rather than absolute serial order at study is maintained at recall. Like the Kahana measure, the output is considered as a series of pairs. Each pair is then scored as to whether the first item in the pair appeared before the second in the study list (it could be adjacent or it could be remote). The number of positive instances is then divided by the total number of pairs. Using this index a score of .50 would indicate chance level. A score of above .5 would reflect the use of forward serial recall, and a score of below .5 would reflect a strategy of backward recall. The difference between a correct response in this measure compared to the Kahana measure is that given recall of item from position n, an item is scored as correct as long at it is further along in the list, n+. In the Kahana method the recalled item has to be n+1.

Recently, Bhatarah, Ward and Tan (2006) have explored sequential recall in the free recall paradigm employing an immediate recall condition as well as a continuous distractor condition in which each item in the list was followed by six seconds of distractor activity. In their experiments they replicated Kahana's (1996) findings that forward recall was more likely than backward recall and that sequential recall of near neighbours was more likely than recall of remote neighbours. Most importantly for current purposes they found that recall of n+1 items was more pronounced in the immediate condition than in the distractor condition. That is, sequential recall was attenuated, but not eliminated in the distractor condition..

In this experiment we replicate Russo and Grammatopoulou (2003) by testing free recall of 8-item lists immediately or after 15 seconds of distraction. The assumption here is that on an immediate test with relatively short lists, participants are likely to be using a serial recall strategy. The short word advantage should be obtained. Given the Bhatarah et al. (2006) results, it is likely that after 15 seconds of distractor activity, participants would be less likely to use serial memory. If so, word length effects should be attenuated.

Method

Participants

Twenty-seven participants from the Australian Catholic University completed this experiment for partial course credit

Materials

The short-word and long-word experimental pools were similar in characteristics to those used Experiment 1. That is, the words were all concrete nouns that were either mono-syllabic or two or three syllables in length. Participants studied 40 eight-word trials, half of which were tested immediately and half after a filled retention interval. To create the 20 short-word trials for each participant, 160 short words were randomly selected and then randomly allocated to a particular trial, such that there were 10 immediate test trials and 10 delayed trials, and to a serial position within that trial. The 20 long-word trials were constructed in the same way by sampling from the 160 item long-word pool. The order of the 40 trials was then randomised. Unique sets were constructed for each participant.

Procedure

The participants were told that they would be presented with a number of trials, each of which consisted of 8 words. They were asked to try and remember them as best they could and that when it came time to recall, they could recall the items in any order they liked.

Each list started with a READY sign displayed for 3 seconds. The 8 words were then presented at a rate of one word per second. The end of the trial was signified by a row of question marks at which point participants attempted recall. On the delayed trials, 15 two-digit numbers were presented at a rate of one two digit number per second between the presentation of the eighth word and the row of question marks. Participants were asked to read each number aloud as it appeared on the screen.

Results and Discussion

The top two serial position curves in Figure 2 show free recall performance on the immediate test, and the bottom two show performance on the delayed test. On the immediate test, strong recency effects are present and modest primacy effects are also observable. On the delayed test, both these aspects are attenuated. More importantly, word length effects appear to be present on the immediate test, but not on a delayed test.

Free recall performance was analysed using a 2*2*8 repeated measures AVOVA. Performance was better on an immediate test than a delayed test, *F* (1,26) = 223.83, *MSE* = 2.94, there was no difference in recall of short and long words, *F* (1,26) = 1.35, *MSE* = 1.65, and there were serial position effects, *F* (7,182) = 28.92, *MSE* = 4.91. While a number of interactions (most involved serial position effects) were significant, the only one that impacted upon the hypotheses of the study was the word length by retention interval interaction, *F* (1,26) = 10.72, *MSE* = 1.50. Simple effects analyses of this interaction revealed that short words were better recalled than long words on an immediate test, *F* (1,26) = 7.23, *MSE* = 2.10, but there was no word length effect on the delayed test *F* (1,26) = 2.00, *MSE* = 1.06.

The next step in the analyses was to determine the extent to which participants were using a serial recall strategy. With the sequential recall measure, given that an item was recalled, the proportion of times recall of item n was followed by item n+1 was .30 for the short words and .21 for the long words on an immediate test. On a delayed test, the respective means were .15 and .13. A 2*2 repeated measures ANOVA, produced main effects for retention interval, *F* (1,26) = 61.50, *MSE* = .006, and word length, F (1,26) = 10.62, MSE = .007. The interaction was also reliable, F (1,26) = 4.87, MSE = .005. Simple effects analysis indicated that sequential recall was more likely for short words than long words on an immediate test, F (1,26) = 12.42, but there was no difference on a delayed test, F (1,26) = 1.03.

The current results for the immediate test replicate those of Russo and Grammatopoulou (2003) in showing word length effects in the free recall of 8 item lists. The word length effects on a delayed test are not significant in this study and do not replicate Russo and Grammatopoulou. The results are also consistent with the Bhatarah et al. (2006) data in that sequential recall is present on both immediate and delayed tests, but is much more pronounced on an immediate test. The crucial finding, however, is that word length effects are matched by the measures of sequential recall. Sequential recall was more likely for short words than long words on an immediate test. Delayed test performance was characterised by a reduction in the presence of sequential recall and that sequential recall was equivalent for short and long words.

In sum, the results indicate that on an immediate test there is a tendency for a serial recall strategy to be used as Postman (1972) and Murdock (1968) have shown. After a filled retention interval, participants are more likely to abandon that strategy. Crucially, where participants are using the strategy, word length effects are present. Where they don’t, word length effects also disappear.

One interesting feature of the immediate test performance is that with the final two serial positions, long words appear to be better recalled than short words. This difference can be attributed to the fact that on 20% of occasions participants recalled the eighth long word first in their recall protocol and the seventh word second. For short words, this occurred on only 9% of trials. In short, with long words there was a tendency for backward recall whereas for short words there was more of a tendency for forward sequential recall.

Experiment 3

The results of Experiment 2 show a strong correspondence between word length effects in free recall and the use of a serial recall strategy. One possible explanation of the word length effects in the previous experiments, is that we are using lists that are not all that much above span and we may well be detecting the decay function of the phonological loop. With longer lists, it is less likely that this explanation would hold. Certainly, Baddeley (1986) argued that the phonological loop was not involved in free recall. So in this experiment 30 items were studied for free recall.

Increasing the list length has the problem that participants may be less likely to use order memory to support recall. However, even with long lists it might be possible that participants start out with a serial rehearsal strategy but abandon it when the task becomes too difficult. For example, Rundus (1971) explored the effects of rehearsal in relation to the serial position curve and found that there was a very strong relationship between the number of rehearsals and the primacy effect. Bhatarah et al. (2006) have also demonstrated that the primacy items in long lists are rehearsed more often than the middle and recency items. They have also shown that across such lists, participants do in fact employ sequential recall to some extent. Their studies did not however, examine the possibility that sequential recall might be most pronounced where rehearsal is most pronounced.

To the extent that serial information is utilised in free recall, it is plausible that word length effects should be observed in the task. The Rundus and Bhatarah et al. data and the results of Experiment 2 suggest that the most likely candidate for such differences is in the primacy region of the free recall curve. If participants reduce encoding serial information for the latter items in the list then word length effects should not be apparent in the middle and recency sections of the free recall curve.

Method

Participants

Forty seven participants took part in the experiment for a ticket in a small cash lottery. They were tested in two large groups. One group of 20 was given the list of short words before the list of long words, and the other 27 were given the long words before the short words.

Materials

Two subsets of 30 words were randomly selected from the short and long word pools used in Experiment 1. A 30-item list of short words was created by randomly ordering the 30 short words. A similar list of long words was created in the same way. All participants saw the same lists.

Procedure

The participants were told that they would be presented with two lists, each of 30 words and that their task was to do their best to remember them. Each list was presented at a rate of one word every two seconds. At the end of each list participants were given 5 minutes to recall as many words as possible and the instructions stressed that the words could be recalled in any order.

Results

Free recall serial position curves generally show enhanced primacy and recency effects. On average the percentage of items recalled across the first six serial positions were 75, 82, 51, 47, 46 and 30; and the percentage of items recalled in the last six positions were 42, 21, 33, 47, 50, and 64. In short, standard primacy and recency effects were evident in the data.

The serial position curves for short and long lists are presented in Figure 3. The thirty items have been divided into fifths for sake of clarity and on the basis that word length effects, if present, should be limited to the primacy region. It is apparent in Figure 2 that word length effects are present in the first fifth of the list (serial positions 1 to 6) but not in the remainder of the list.

The data in Figure 3 were first divided on the basis of primacy region (1st fifth) and the rest (remaining 4 fifths) and were subjected to a 2*2 repeated measures ANOVA. Overall short words were better remembered than long words, $F(1,46) = 4.16$, $MSE = .02$, and recall was better in the primacy region than in the rest of the list, $F(1,46) = 61.59$, $MSE = .03$. The interaction however, was reliable, $F(1,46) = 6.56$, $MSE = .02$. Simple effects analyses indicated that the word length effect was significant in the primacy region, $F(1,46) = 6.39$, $MSE = .03$, but not in the remainder of the list, $F(1,46) < 1$, $MSE = .01$.

The word length effects are present in the predicted serial positions. However, if the theory is correct, one needs to demonstrate that order information is being used in the primacy region but not in the rest of the list. The Asch-Ebenholtz (1962) index for the primacy region was .68 and for the remainder of the list was .48. This difference was reliable, $t(86) = 5.58$.

Figure 4 presents the Kahana sequential recall measure for the primacy and recency regions. A 2*2 repeated measures ANOVA, produced main effects for serial position, $F(1,46) = 37.82$, $MSE = .018$, and word length, $F(1,46) = 24.46$, $MSE = .016$. The interaction was also reliable, $F(1,46) = 15.48$, $MSE = .015$. Simple effects analysis indicated that sequential recall was more likely for short words than long words in the primacy region, *F* (1,46) = 21.00; but there was still a significant, but attenuated short word advantage in the latter serial positions, *F* (1,46) = 7.22. Within the primacy portion there were strong serial position effects. The proportion of participants who recalled the n and n+1 pair in short word list was .51 for the positions 1-2 pair, .23 for the 2-3 pair, .10 for the 3-4 pair, .06 for the 4-5 pair and .01 for the 5-6 pair. For the long word list the respective proportions were .23, .11, 0, 0, 0. These results indicate that participants are sequentially recalling the short words in the primacy portion of the short lists.

Discussion

The results are straightforward. The two measures of the degree to which forward serial recall is being used in the free recall task converge on the same conclusion: Participants used a forward serial recall strategy when they attempted to recall the items that were first presented in the list. The sequential recall measure indicates that this was particularly true of the short words. However, for the remaining items in the list the use of a sequential strategy is reduced. In the places where sequential recall was being used, word length effects were apparent; in the places where it was not, the effect did not emerge. The results are thus consistent with the item-order distinction and show fairly clearly that standard word length effects only emerge where it can be demonstrated or it is reasonable to assume that participants are utilising order information. The fact that sequential memory is still used to some extent in the non-recency serial positions suggests that an item advantage for long words is being offset by an order advantage for short word. That is, in the latter serial positions the opposing effects of item and order information balance each other out.

Experiment 4

The results of Experiments 1 to 3 show that the standard word length effects emerge in situations where order information is being utilised and this is assumed to be a result of the short words receiving better order encoding. However, a key prediction of the item-order approach is that the reverse effects should be found on a task requiring only item information. The preferred "pure" test of item information has typically involved single item recognition procedures. Hendry and Tehan (2005), using the Nairne et al. (1991) methodology, found such reverse word length effects in a long-term item recognition. However, as mentioned earlier, this effect was confounded with retention interval. For the item order hypothesis to be strongly supported, it needs to be demonstrated that reverse item effects can be observed across a range of retention intervals and that they generalise to other memory tasks. The remaining experiments explore these issues.

In Experiment 4 standard long-term item recognition procedures were utilised. That is, participants first studied a long list of words containing short and long words. At the end of the study phase participants were given a standard recognition test containing the short and long items on the study list plus a number of short and long lures. Participants were asked to indicate which of the words had appeared on the study list. Based on the item-order approach, item recognition for long words should be better than for short words.

Method

Participants

Thirty-one participants from the University of Southern Queensland completed this experiment in exchange for a ticket in a small cash lottery. All subjects were tested in a single group session.

Materials

The experimental word pools from Experiment 1 were again used. Thirty of the short words and thirty of the long words were randomly selected from their respective pools to serve as the study items. These subsets of short and long words did not differ on word frequency or imagability measures. The 60 items were then randomised to form the study list that all participants viewed.

The test list involved the 60 items from the study list and 30 of the remaining unused items from each of the small and large word length pools. The 60 targets and the 60 lures were then randomly ordered and presented on a sheet of paper in four columns of 30 words.

Procedure

The participants were instructed that they would be given a long list of words to learn and that their memory for these words would be tested later. The study list was then presented via computer projection onto a large screen. Each item on the study list was presented for 2 seconds. Following list presentation, the test list was distributed and participants were instructed to circle the items on the test sheet that had appeared on the study list.

Results

The probability of correctly recognising a long word was .77 (SEM =.03) compared to .69 (SEM=.03) for short words. This difference was reliable, t (30) = 4.50. There was no difference in false alarms, t (30) = 1.80, with means of .03 (SEM=.01) and.05 (SEM=.01) for short and long words respectively. Recognition accuracy for long words (d’ = .92) was better than for short words (d’ = .55), t (30) = 4.36.

Discussion

The findings from Experiment 4 clearly show that long words produce better item recognition than short words and this result is consistent with the item-order trade off explanation. The current results replicate Hendry and Tehan (2005) and indicate that reverse word length effects can be observed under standard LTM tests.

In many ways these results mirror those of the word frequency effect. In immediate serial recall high frequency words are better recalled than low frequency words (Watkins, 1977). However, in long term recognition studies like that conducted here, the standard finding is that low frequency words are better recognised than high frequency words. Hockley and Christi (1996) have suggested that the pattern of word frequency effects in recognition and recall could well be explained in terms of differential item and relational processing.

Experiment 5

The item part of the item order approach has almost universally been tested using single item recognition procedures (Hendry & Tehan, 2005; Mulligan, 1999, 2000; Nairne et al., 1991). In an attempt to generalise these findings we turn to cued recall as a measure of item memory. To the best of our knowledge, cued recall has never been used in the item order context. Consequently, in this experiment we adopt a conservative approach by utilising the Nairne et al., (1991) paradigm but replace the final item recognition test with a cued recall test. Hendry and Tehan (2005) found the word length dissociation with recognition using this paradigm so the expectation that a similar dissociation should emerge with a cued recall task.

Method

Participants

Thirty seven undergraduate students volunteered to participate in the experiment, either for course credit or a ticket in a cash prize draw.

Materials

Each participant studied 16 trials with each trial comprising five items. The complete set of trials consisted of 4 trials of short words and 4 trials of long words for immediate serial recall and 4 trials of short words and 4 trials of long words for cued recall. The trials were constructed by first selecting 40 monosyllabic words and 40 polysyllabic words from the University of South Florida taxonomic category norms (McEvoy & Nelson, 1982). Two sets of trials were generated. To construct the first set of trials the 40 words from each word pool were first randomly allocated to the five serial positions across 8 trials for the short and long trials. Each set of 8 trials was randomly allocated to be tested via immediate serial recall (4 trials) and via cued recall (4 trials). The order of the 16 trials was then randomised.

To create the second set of trials, all that differed was that those trials that had been allocated to the serial order condition were now allocated to the cued recall condition and vice versa, the items in each trial and their serial position remained unchanged across the two sets. Half the participants studied the first set of trials and the remaining half studied the alternative set.

Procedure

Participants were instructed that they would be presented with a list of five words, and that after the fifth word in each trial a prompt would be presented in the form of exclamation marks or question marks. If the prompt was a row of exclamation marks (!!!!) participants were told to forget the words. If the prompt was a row of question marks (????) participants were to recall five items in the order that they were presented. The sixteen trials were then presented.

Each trial commenced with a beep and the words for the trial were then presented in lowercase in the middle of the computer screen, at a rate of one word per second. At the end of each trial a prompt was presented followed by a 10-second delay during which the participant recalled the words in serial order or waited for the commencement of the next trial.

After completion of the immediate serial recall task, participants were told to try to recall the items they had previously been told to forget. Participants were presented with a sheet of paper that contained 40 category cues and were told that each of the cues was related to one of the items on the exclamation trials. They were instructed to use these cues to help them recall the items, and then write any words that they remembered from the earlier list in the place provided beside the appropriate cue.

Results and Discussion

Preliminary inspection of the data indicated that with five item lists all participants adopted a forward serial recall strategy as requested. The results of the experiment are summarised in Figure 5.

Two planned comparisons indicated that memory performance for short words was superior to long words for immediate serial recall, t (36) = 4.85. Long words were better recalled than the short words on the cued recall task, t (36) = 5.11.

The results are clear. Reverting back to the standard paradigm in which item order trade offs have been most frequently explored, word length dissociates across order and item tasks. The reverse word length effect generalises from an item recognition task to a cued recall task. The pattern of results is totally consistent with the item order hypothesis.

General Discussion

The word length effect is a pivotal piece of evidence used to support the notion firstly, that short-term memory traces actually exist and secondly, those traces rapidly decay in the absence of rehearsal. Given this role it is not surprising that word length effects have largely been ignored in long-term recall. The current research has examined the word length effect across a range of long-term memory tasks and shown that the effect generalises to the long-term domain. Long-term word length effects question the assertion that the word length effect is a uniquely short-term phenomenon. Parsimony would suggest that if word length effects are observed across both short and long retention intervals, an explanation that crosses retention interval boundaries is desirable.

In the introduction we suggested that the phonological loop account would predict the absence of word length effects in any of the tasks. Several facets of our results require a re-examination of this prediction. In Experiment 1, word length effects were present on the first trial and were maintained across the first three trials where performance was below criterion. We have suggested that the Burgess and Hitch (1996) model of the phonological loop could well account for the results but that it was likely to be the long-term connections in their model that were producing the effect. An alternative might be that the word length effect and the increase in overall levels of recall across trials might be due to independent mechanisms or processes. For example, the word length effects may be due to rapid decay (the phonological loop), but increments in performance might be due to recall from the episodic buffer. This account is very speculative in that the interaction between phonological loop and episodic buffer in recall tasks has not yet been specified in any testable way. Moreover, the interaction between components is unlikely to be straightforward. Broadbent (1987) describes one attempt to integrate the output of different stores (phonological loop and central executive) in a sequential fashion. He concluded that such attempts, while logically and theoretically sound, did not fit the known data.

The free recall data in Experiment 3 show a strong correspondence between serial position, word length and rehearsal. If participants in our experiment are performing the task in the same way as those in the Rundus (1971) and Bhatarah et al. (2006) studies, then it seems to be the case that the primacy portion of the free recall curve has three characteristics. The items are rehearsed more frequently than other list items, order memory appears to be used most frequently here, and word length effects are stronger here as well. The link between rehearsal, word length, and serial recall has almost all the hallmarks of the phonological loop save rapid trace decay. Given that the trace decay assumption means that the phonological loop component cannot account for the primacy effects, an alternative account of the co-occurrence of the three effects is required. One possibility is that rehearsal's primary function is not to refresh a decaying trace, but to refresh or maintain order information. Baddeley has almost always utilised closed sets of items in his experiments precisely to measure order memory unconfounded with item memory. Likewise, it has long been argued that with closed sets order memory is lost more rapidly than item memory such that errors dominate item errors in frequency of occurrence (Bjork & Healy, 1974). It is thus plausible that the primary feature of rehearsal is to maintain order memory, not item memory. If, as we contend, short words receive more or better order encoding, word length effects should co-occur with rehearsal.

The discrimination account of the word length effect allows for the possibility of word length effects in long-term retention tasks. Given the assumption that short words are more discriminable than long words, the presence of a short word advantage in Experiments 1, 2 and 3 would be expected. However, from this perspective there is no reason to suppose that long words should be better recognised or better remembered in a cued recall task than short words. If anything one would predict a similar short word advantage. The results of Experiments 4 and 5 are not readily explained by the discrimination account.

The item-order perspective appears to fare better than the other two explanations of the word length effect. The results across experiments were consistent with this expectation in that word length produced the dissociative effects on item and order tasks. The standard short word advantage was obtained in a long-term serial list learning task where by definition participants are required to use order memory. In the free recall tasks, word length effects were present but most importantly, were strongest in those serial positions where it could be demonstrated that participants were using some form of order memory to facilitate recall. In those instances where evidence for the use of order memory was absent or weak, there was a corresponding attenuation of the short word advantage in recall. The long word advantage predicted by the item-order account has been demonstrated in a standard long-term recognition task that is widely seen as a task that relies almost exclusively upon item memory and is devoid of order memory. We have extended this finding to long-term cued recall using the standard item-order paradigm. In short, standard word length effects are found in long-term tasks where order memory is being used. Where item memory is being accessed, a long word advantage emerges. The item-order account of the word length effect does generalise across retention intervals and across tasks.

While we think that the item/order account provides a better explanation of the data than the discrimination perspective, our current results could well be incorporated within the SIMPLE framework by adding an “order” dimension. That is, recall would now involve discrimination within a three dimensional (temporal, order, and item) space. What is currently labelled the item dimension in SIMPLE we would argue is our order dimension, that is, short words are more discriminable on an order dimension than long words because of better order encoding. Long words would be more discriminable than short words on our proposed item dimension. In all other respects the model would stay the same. Given that the weightings along each dimension must sum to one in SIMPLE, tasks that require order memory would have a higher weight than the item dimension and item memory tasks would have a higher weight on the item dimension than the order dimension. The dissociation of word length with type of task should emerge from the model.

Applying the Item-order account to other word length effects

The current results provide further evidence that the word length effect extends beyond immediate serial recall and as such a parsimonious explanation would posit that a short word advantage should be evident in any task where order information is being used. The item-order approach thus readily accounts for the findings that a short word advantage has been found in complex span tasks where items and distractor activity alternate with each other (Tehan et al, 2001) and in the Brown-Peterson task where a 12 second filled retention interval was utilised (Tehan et al., 2001). Furthermore, word length effects in backward recall (Cowan, Wood & Borne, 1994) are explainable given the assumption that backward recall is accomplished via a series of forward recalls (Thomas, Milner & Haberlandt, 2003). Likewise, word length effects in serial recognition (Baddeley, Chincotta, Stafford & Turk, 2002) are expected given that the task requires participants to utilise order information.

There are, however, a number of potentially problematic findings. Firstly, word length effects have been observed in probed recall tasks where ordered recall of multiple items is not required (Avons, Wright & Pammer, 1994; Henry, 1991; Henry, Turner, Smith & Leather, 2000). As we have demonstrated with free recall, participants can adopt a serial recall strategy even though the instructions do not stress the use of order information, and to the extent that this is so, one would expect the standard word length effect. A probe recall study by Henry et al. (2000) illustrates this point. They examined the emergence of word length effects as a function of age. In one experiment serial recall was requested and the subjects in the three age groups used (4-year-old, 7-year-old and 10-year-old) all produced reliable word length effects. In a second experiment, probe recall was required rather than serial recall. In the case where only one item had to be recalled, the 4-year-olds did not exhibit word length effects. In fact there was a tendency for reverse word length effects to emerge. With the two older groups, word length effects were present in this task but were stronger for those students who reported using a serial recall strategy. Those students who reported using a simple naming strategy produced very weak effects. Thus, even though the probe recall task does not require serial order to be utilised, participants still adopted a serial rehearsal strategy. Interestingly, as implied by the item-order trade off perspective, word length effects were not strong when serial rehearsal was not being used.

A second problematic study is that of Cowan, Nugent, Elliott and Geer (2000). They manipulated output speed by having participants articulate their responses either at a fast speed or an exaggeratedly slow speed. Given that the same items were involved in each case, the deficit for the slowly articulated responses suggests that forgetting was occurring during output. However, this research has not been free from criticism on the basis of differential attentional demands (Service, 2000).

Another potentially problematic finding is that reported by Cowan, Wood and Borne (1994) where six-item lists were studied for backward recall. They studied word length effects under immediate serial recall and under continuous distractor conditions where each list item was preceded and followed by 15 seconds of distractor activity. They found the standard short word advantage on the immediate test, but the advantage was reversed on the continuous distractor condition. On the basis of these findings they argued that separate short and long term memory systems were required. From the item-order perspective the continuous distractor task is one that required serial order, yet given all the previous evidence that order information is lost relatively quickly (transposition errors rapidly become omission errors), it seems reasonable to suppose that with a one and a half minute retention interval for the early list items that such order information is rapidly lost. As such even though the task requires serial information, participants may be relying predominantly upon item information with the result that a long word advantage emerges. This argument is the mirror of that used to explain standard word length effects in free recall. In that task we suggested that participants use order information although the task does not demand it. With the continuous distractor task, participants are required to use serial order information, but the task may be so difficult that item information is used to a greater extent than order information.

Finally, Hulme, et al. (2004) found that word length effects disappeared in serial recall of lists containing both short and long words mixed together. That is, there were strong word length effects in the pure lists, but no word length effect in the mixed lists. This finding is less problematic in that precisely the same pattern emerges with the generation and perceptual interference effects. That is, the effects are readily observable in pure lists but are attenuated in mixed lists. This is the case because differential item and order processing are equivalent in mixed lists

Assumptions underpinning the Item-order account

To this point we have not addressed the fundamental assumption we are making that short words are easier to process at the item level than long words, and that this facilitates order encoding. There is nothing in our results that directly bear on this issue all we can do is re-iterate the points that were made by Hendry and Tehan (2005). In the visual word recognition literature, short words are recognised more quickly than long words in lexical access tasks like word naming, perceptual identification and lexical decision (Balota & Chumbley, 1985; Forster & Chambers, 1973; Samuels, Laberge & Bremer, 1978). Thus, there is direct evidence that long words take longer to identify and respond to than short words but there is no direct evidence supporting the notion that increased identification time results in enhanced item processing.

On the positive side, the individual differences literature has indicated that one of the prime determinants of immediate memory span is the speed at which items can be identified. For instance, in a review of the literature to that time, Dempster (1981) examined ten possible sources of individual differences in memory span. His review indicated that the item identification speed was the most reliable source of individual differences in span among children. The relationship between item identification time and span has since been demonstrated on a number of occasions with both children (Case, Kurland & Goldberg, 1982; Hitch, Halliday & Littler, 1989; 1993) and adults

(Tehan & Lalor, 2000) as participants (but see Henry and Millar, 1993, for an alternative view).

The Tehan and Lalor (2000) data are relevant here in that they demonstrated that performance on lexical decision, word naming and other item identification tasks made a significant contribution to individual differences in span performance; a contribution that was more important than traditional rehearsal and output time measures. Importantly, the tasks they used to establish the relationship between span and lexical access were the same tasks that were showing word length differences in the lexical access literature and the same tasks that are used in the item identification literature. Thus, the available literature does support the notion that item processing speed is important in immediate serial recall and that short words are processed faster than long words. Given that short words are identified more quickly than long words, there is more time available for encoding order information. Thus, our argument that the word length effect in serial recall is due to differences in item processing does have support in the literature.

In conclusion, the results of the current experiments show that word length effects are not specific to immediate serial recall. They extend across a range of tasks that involve the use of order information. Reverse word length effects can also be observed across a range of conditions where item information is accessed. These latter effects are problematic to current theoretical accounts of the word length effect but are readily explained by the item-order trade-off hypothesis.

Asch, S. E. & Ebenholtz, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *Journal of Psychology, 54*, 3-31.

Avons, S. E., Wright, K. L. & Pammer, K. (1994). The word-length effect in probed and serial recall. *Quarterly Journal of Experimental Psychology, 47A*, 207-231.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

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

Baddeley, A. D., Chincotta, D., Stafford, L., & Turk, D. (2002). Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *Quarterly Journal of Experimental Psychology, 55A*, 353-369.

Balota, D. A. & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24*, 89-106.

Batharah, P., Ward, G., & Tan, L. (2006). Examining the relationship between free recall and immediate serial recall: The effect of concurrent task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 215-229.

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior, 13,* 80-97.

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

Broadbent. D. E. (1987). Simple models for experimentable situations. In P. Morris. (Ed.) *Modelling Cogntion* (pp.169-185). London: John Wiley & Sons.

Caplan, D. & Waters, G. (1994). Articulatory length and phonological similarity in span tasks: A reply to Baddeley and Andrade. *Quarterly Journal of Experimental Psychology, 47A*, 1055-1062.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386-404.

Cowan, N., Wood, N.L., & Borne, D. N. (1994). Reconfirmation of the short-term storage concept. *Psychological Science, 5*, 103-106.

Cowan, N., Nugent, L. D., Elliott E. M. & Geer T. (2000). Is there a temporal basis of the word length effect?: A response to Service (1998). *Quarterly Journal of Experimental Psychology, 53A*, 647-660.

DeLosh, E. L. & McDaniel, M. A. (1996). The role of order information in free recall: Application to the word frequency effect. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1136-1146.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*, 63-100.

Engelkamp, J. & Dehn, D. M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 671-682.

Forster, K. I., & Chalmbers, S. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behaviour, 12*, 627-635.

Hendry, L. & Tehan, G. (2005). An item and order analysis of word length and generation effects in the short-term domain. *Memory, 13*, 364-371.

Henry, L. A. (1991). Development of auditory memory span: The role of rehearsal. *British Journal of Developmental Psychology, 9,* 493-411.

Henry, L. A., & Millar, S. (1993). Why does memory span improve with age? A review of the evidence for two current hypotheses. *European Journal of Cognitive Psychology, 5,* 241-287.

Henry, L. A., Turner, J. E., Smith, P. T. & Leather, C. (2000). Modality effects and the development of the word length effect in children. *Memory, 8,* 1-17.

Hitch, G. J., Halliday, M. S., & Littler, J. E. (1989). Item identification time and rehearsal rate as predictors of memory span in children. *The Quarterly Journal of Experimental Psychology, 41,* 321-337.

Hitch, G. J., Halliday, M. S., & Littler, J. E. (1993). Development of memory span for spoken words: The role of rehearsal and item identification processes. *British Journal of Developmental Psychology, 11*, 159-169.

Hockley, W. E., & Cristi, C. (1996). Tests of encoding trade-offs between item and associative information. *Memory and Cognition, 24,* 202-216.

Hulme, C., Surprenant, A.M., Bireto, T.J., Stuart, G. & Neath, I. (2004). Abolishing the word length effect. *Journal of Experimental Psychology: Learning, Memory & Cognition, 30,* 98-106.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24,* 103-109.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence. RI: Brown University Press.

McEvoy, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology, 95*, 581-634.

Mulligan, N. W. (1999). The effects of perceptual interference at encoding on organization and order: Investigating the roles of item-specific and relational information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*, 54-69.

Mulligan, N. W. (2000). Perceptual interference and memory for order. *Journal of Memory and Language, 43,* 680-697.

Murdock, B. B. (1968). Recent developments in short-term memory. *British Journal of Psychology*, *58,* 421-433.

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

Nairne, J. S. (2002). Remembering over the short term: The case against the standard model. *Annual Review of Psychology*, *53,* 53-81.

Nairne, J. S., Reigler, G. L. & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 702-709.

Neath, I. & Nairne, J. S. (1995) Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin and Review, 2,* 429-441.

Postman, L. (1972). A pragmatic view of organisation. In E. Tulving and W. Donaldson. (Eds). *Organization of memory*. (pp. 3-48.) San Diego, CA: Academic Press.

Quinlan, P. T. (1992). *The Oxford psycholinguistic database*. Oxford: Oxford University Press.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*, 63-77.

Russo, R. & Grammatopoulou, N. (2003). Word-length and articulatory suppression affect short-term and long-term recall tasks. *Memory and Cognition, 31,* 728-737.

Samuels, S. J., Laberge, D. & Bremer, C. D. (1978). Unit of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behaviour, 17,* 715-720.

Service E. (2000). Phonological complexity and word duration in immediate recall: Different paradigms answer different questions-A comment on Cowan, Nugent, Elliott, and Geer. *Quarterly Journal of Experimental Psychology, 53A,* 661-665.

Tehan, G. & Lalor, D. M. (2000). Individual differences in memory span: The contribution of rehearsal, access to lexical memory and output speed. *Quarterly Journal of Experimental Psychology*, *53A*, 1012-1038.

Tehan, G, Hendry, L & Kocinski, D. (2001) Word length and phonological similarity effects in simple, complex and delayed serial recall task: Implications for working memory. *Memory, 9*, 333-348.

Thomas, J. G., Milner, H. R. & Haberlandt, K. (2003). Forward and backward recall: Different retrieval time patterns, same retrieval order. *Psychological Science, 14*, 169-174.

Watkins, M. J. (1977). The intricacy of memory span. *Memory & Cognition, 5*, 529-534.

Figure Captions

Figure 1. Recall of short and long words across the first three trials in serial learning task. Note: Error bars represent standard error of the mean.

Figure 2. Recall of short and long words as a function of serial position in a free recall task. Note: Error bars represent standard error of the mean.

Figure 3. Recall of short and long words as a function of serial position in a free recall task. Note: Error bars represent standard error of the mean.

Figure 4. Proportion of sequential recalls for short and long words. Note: Error bars represent standard error of the mean.

Figure 5. Serial and cued recall of short and long words. Note: Error bars represent standard error of the mean.

Figure 1.

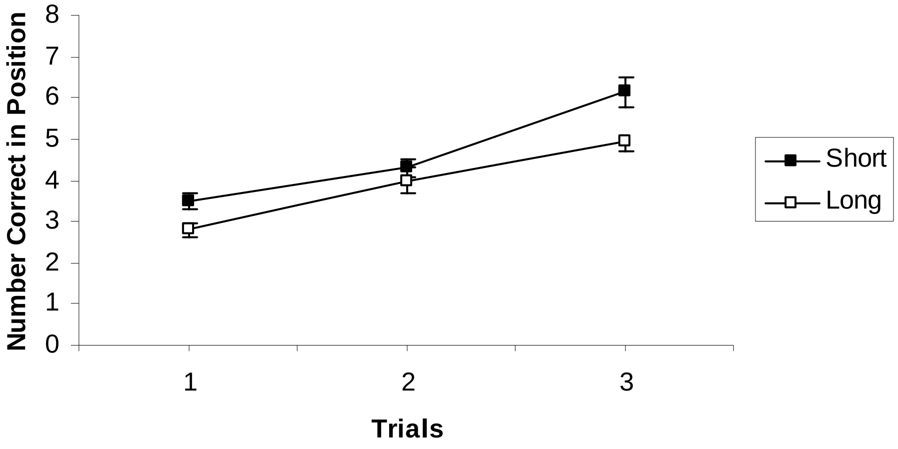

Number Correct in Position
8
7
6
5
4
3
2
1
0
1
2
3
Trials
Short
Long

Figure 2

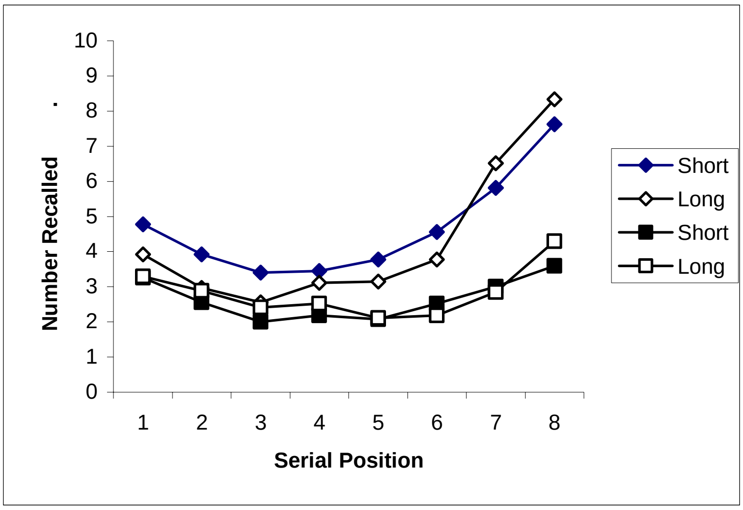

10
9
8
7
6
5
4
3
2
1
0
Number Recalled
1
2
3
4
5
6
7
8
Serial Position
Short
Long
Short
Long

Figure 3

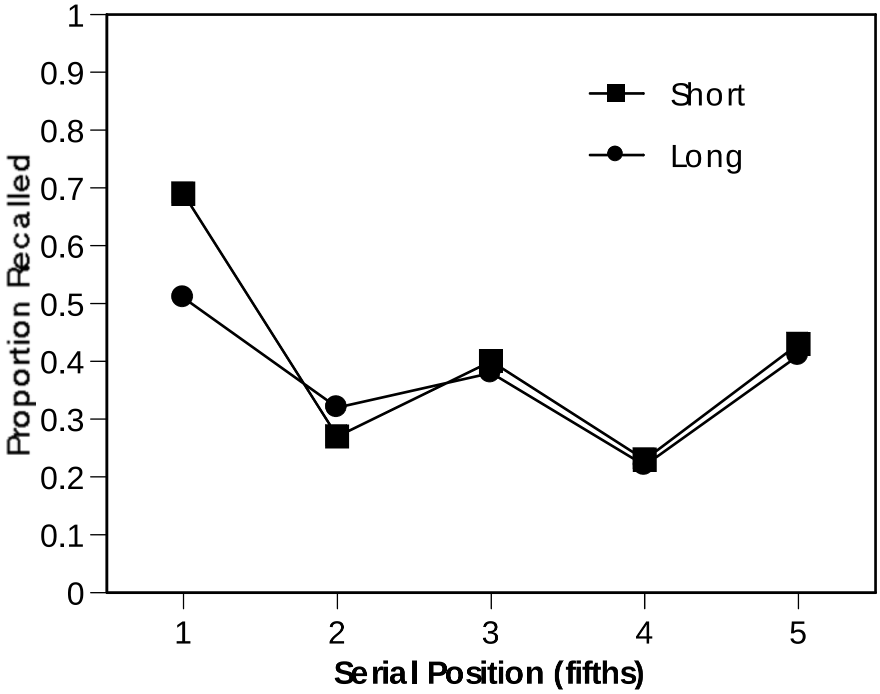

Proportion Recalled
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Short
Long
1
2
3
4
5
Serial Position (fifths)

Figure 4.

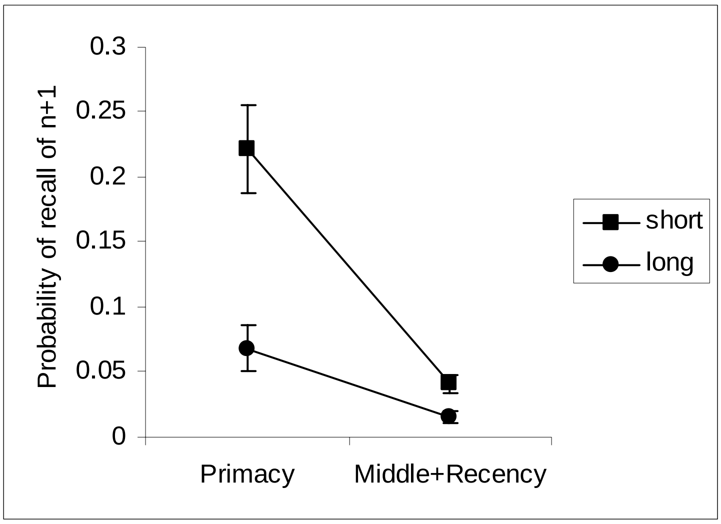
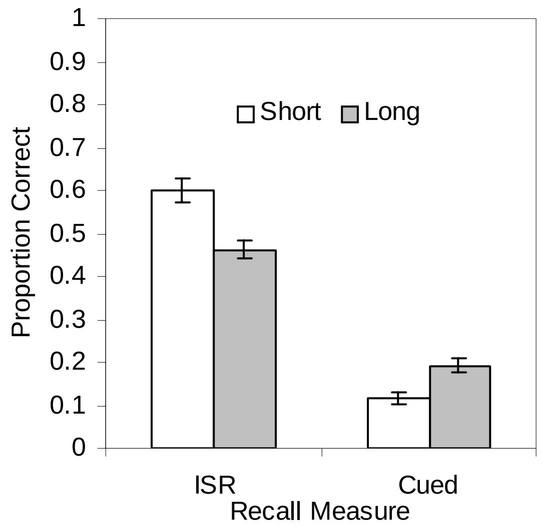

0.3
0.25
0.2
0.15
0.1
0.05
0
Probability of recall of n+1
Primacy
Middle+Recency
short
long

Figure 5


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Proportion Correct
Short
Long
ISR
Cued
Recall Measure