Word Length Effects Are Not Due to Proactive Interference

Gerald Tehan
University of Southern Queensland
Toowoomba, Queensland, Australia

Josée Turcotte
Laval University
Québec, Québec, Canada

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

In immediate serial recall short words are better recalled than long words. The word length effect has become pivotal in the development of short-term memory models. The current research tests one explanation of the word length effect; that it is related to proactive interference (PI). We report two experiments in which the relationship is directly tested. In the first experiment we show that word length effects can be observed over the first few trials in an experiment and that the effect shows itself primarily in the number of omissions made. In the second experiment we simultaneously test for PI and word length effects. Strong word length effects were present but there was little evidence for PI influencing either overall levels of recall or the word length effect. In short, no empirical support was found for PI as an explanation of the word length effect.
Word Length Effects Are Not Due to Proactive Interference

Across the years there has been continual debate concerning whether forgetting is caused by decay or by interference. One interesting facet of this debate is that interference effects are usually seen as a prime source of forgetting in long-term memory settings, but decay is seen as the principle means in immediate memory settings (Baddeley, 1986; Burgess & Hitch, 1996; Henson, 1998; Page & Norris, 1998). Consequently, over the last thirty years or so, the vast majority of models of immediate memory have assumed that the short-term memory trace that supports immediate recall decays very rapidly, unless it is renewed by verbal rehearsal. Not surprisingly these models are now often referred to by the generic name of "trace decay plus rehearsal" models (Brown & Hulme, 1995).

For those who argue in support of decay, the word length effect, the fact that span for short words is larger than span for long words, is one of the key short-term phenomena. In the original research that established this effect, Baddeley, Thomson and Buchanan (1975) first established a span advantage for one-syllable words over five syllable words. However, in subsequent experiments they established that the prime determinant of the word length effect was the spoken duration of the words. Span for the short duration words was significantly larger than span for the long duration words. On the basis of such findings the decay plus rehearsal models assume that short-term memory traces rapidly decay in real time. Given that per given period of time more short words can be rehearsed than long words, more short words can be kept in an active state than long words.

In recent years, however, the trace decay plus rehearsal assumptions have come under increasing amounts of pressure. The initial research that implicated spoken duration has proved difficult to replicate (Lovatt, Avons & Masterson, 2000).
Word length effects have been found when pronunciation rates have been controlled for (Caplan, Rochon & Waters, 1992), or rehearsal has been prevented (LaPointe & Engle, 1990) and there are instances of where there are no-span differences where there are clear differences in pronunciation rates (Service, 1998). Cowan et al (1992, see also Dosher & Ma, 1998) have suggested that word length effects occur during the recall process itself rather than during rehearsal prior to recall. In short, simple decay notions appear to be inadequate as either a necessary or sufficient explanation for the word length effect.

Currently, there are three alternative explanations for the word length effect, two of which are based upon interference assumptions. Neath and Nairne (1995) suggested that word length effects might result from intra unit interference. Melton (1963) had demonstrated that items within a study list produced mutual interference on each other. Neath and Nairne extended Melton's idea to features within words. They suggested that words had to be compiled from sets of smaller features and that the more features that had to be compiled, the greater the likelihood that an error would be made. Since long words were assumed to require the compilation of more features than short words, these words would be more error prone. Neath and Nairne incorporated these ideas into the feature model (Nairne 1990) and were able to provide good fits of existing data. The feature model with its assumptions about intra unit interference has subsequently made novel predictions concerning the conditions under which the word length effect would and would not be found and the data have so far been consistent with the predictions of the model (Neath, Surprenant & LeCompte, 1998)

The second interference-based explanation is that proactive interference (PI) plays a role in producing the word length effect (Nairne, Neath & Serra, 1997).
Nairne et al. argued that if decay was the causal factor underpinning the word length effect the effect should be as strong on the first trial of an experiment as on the last. Consequently, they examined performance on a trial-by-trial basis. In their first experiment they presented subjects with four 5-word lists that consisted of two-syllable words that differed in spoken duration (Cowan et al., 1992). 220 students were given four trials of short words and another group of 220 students was given four trials of long words. No word length effects were present on any of the trials. This result could be simply interpreted as another failure to replicate previous findings, or simply be a cohort difference. However, these explanations were ruled out in the second experiment where subjects studied 24 trials of either short or long words. Again word length effects were absent on the first four trials, but they did emerge on the next block of four trials and were also present on the remaining blocks in the experiment. This finding that word length only emerged after four trials was clearly problematic for trace decay plus rehearsal explanations. Moreover, the result was reminiscent of Keppel and Underwood's (1962) finding that forgetting in the Brown-Peterson task gradually emerged over three or four trials. Given this correspondence and the fact that the Keppel and Underwood data are generally attributed to PI, they suggested that word length effects could be related to PI. Thus, both PI and word length effects build up over trials.

From a theoretical perspective, while the similarities between the Nairne et al and the Keppel and Underwood are enticing, it is hard to see how word length effects could be incorporated within standard trace discrimination explanations of PI. The standard explanation of PI in the Brown-Peterson task is that on the first trial there are only a small number of items available for recall. There are no problems in discriminating these items from other items in the experiment because as yet there are
no other items. However, on the second trial subjects have to discriminate the items from the current trial from the items on the first trial. Likewise on the third trial the items have to be discriminated from the items on trial one and two, and so on for subsequent trials. From the discrimination account it is hard to see why long words should start off being equally discriminable as short words on early trials but become less discriminable over latter trials.

There are also a number of methodological issues that require further examination before the role of PI in the word length effect can be unequivocally accepted. For a start, the use of short and long disyllabic words represents a weak manipulation of word length. Using a stronger manipulation of monosyllabic versus multi-syllabic words would provide a stronger test of word length effects. Secondly, Nairne et al. used a very slow presentation rate, effectively one word every four seconds rather than the standard rate of one word per second. Such a slow presentation rate allows participants time to use more sophisticated and elaborative mnemonic strategies than simple verbal rehearsal. If elaborative strategies were used there would be no reason to expect a word length effect. Fallon (1999) has shown that at a four-second presentation rate another marker of short-memory performance, the phonological similarity effect, is not observed, although in that case performance was close to ceiling. Thirdly, word length effects are sensitive to the word pool that is used to construct the word lists. Word length effects are much more robust when large pools of small and large words are used rather than small pools of words that are repeated from trial to trial (LaPointe & Engle, 1990). Nairne et al. used a small pool; different effects might be observed if a large word pool was employed. Lastly, Nairne et al. did not explicitly test for PI in their experiment. Without doing an error analysis it is difficult to establish that the lack of word length effects was due to PI rather than
some other factor. For instance it is possible that word length effects might be due to a
differential pattern of omission or order errors, as is the case with phonological
similarity effects (Fallon, Groves & Tchan, 1999).

In the following experiments we attempt to replicate and extend the Nairne et
al. research. In the first experiment we attempt a close replication of Nairne et al. save
that we use a standard presentation rate of one word per second and word length is
manipulated within subject rather than between subject. We also extend their results
by adding a stronger manipulation of word length and we also vary the type of word
pool that is used to construct the lists. In the second experiment we use a paradigm
that is better suited to exploring PI effects in short-term memory. Here we explicitly
examine PI effects and their relationship to word length.

Experiment 1

Method

Participants

Thirty psychology students from the University of Southern Queensland
volunteered to take part in the experiment.

Materials

Each trial in the experiment consisted of five words. Each subject studied
three blocks of short words and three blocks of long-words, with each block
containing four trials.

To construct the trials six different word pools were derived. The first two
were the Cowan et al (1992) short and long words that were used by Nairne et
al. (1997. The short words were decor, ember, hackle, pewter, wiggle and the long
words were coerce, humane, morphine, voodoo, zygote. These words were all
disyllabic words that had been matched for word frequency. The four trials of short
words were created by randomly assigning the five words to the five serial positions on each trial. The long-word trials were created in the same manner. Thus, each word appeared four times during the experiment.

To create the second set trials another two five-word pools were created. The short words in this pool were barn, boat, chin, pope, rice and the long words were blanket, cigarette, factory, magazine, telephone. These words were selected from the MRC Psycholinguistic Database. Short and long words were matched for word frequency and imagery but differed in the number of phonemes. The short words were all monosyllables of three phonemes in length and the long words were two or three syllables containing seven phonemes. Four trials of short and long words were created in the same manner as the first set. Again each word appeared four times during the experiment.

The third set of trials used larger word pools. In each case, twenty short and twenty long words were selected from the MRC database to have the same characteristics as those used in the previous set. That is, they were medium-frequency, high-imagery words that differed in the number of phonemes. Twenty words were one-syllable words that consisted of three phonemes and the remaining twenty were two or three syllable words that contained seven phonemes. The twenty words in each pool were randomly assigned to the five serial positions across the four trials such that each word only appeared once during the experiment.

As was the case in the Nairne et al. experiment, all subjects studied the same set of trials, in the same order.

Procedure

The participants studied the six sets of trials across two days. The subjects were all tested in a single group. The first session was held at 9.00 on the first
morning, the second at 11.00 am and the third at 1.00pm. The third, fourth and fifth were held at the same times on the second day. Loess and Waugh (1967) presented data that showed that PI would not be observed if the interval between trials was extended. That is, it was possible to show temporal release from PI. Consequently, a between subjects design could be used if the different materials could be studied with a long period between the different sessions. Given that it took less than two minutes to present the four trials and that there was at least a two-hour break between sets of trials, the expectation was that there should not be any PI between the different sets of materials.

Each trial in each session started with a READY sign presented for three seconds. The five words were then projected onto a screen at the rate of one word per second. The trial finished with a row of question marks at which point participants wrote what they could remember on a prepared answer sheet. The sheet contained five places for subjects to write down the items. Subjects were instructed to write from left to right and to leave blank spaces for any words they could not remember. They had 12 seconds to recall the five items.

Results

Recall was scored as correct only if the word was recalled in its correct serial position. The results of the experiment are summarized in Figure 1. The top panel presents the data for the Cowan et al. (1992) words, the middle panel represents the closed pool of three and seven phoneme words and the bottom panel represents the open pool of three and seven phoneme words. The top panel shows little evidence for word length effects and as such replicate the Nairne et al. findings. Word length effects do seem to be present in the lower two panels.
The initial analyses involved a 2*3*4 repeated measures ANOVA with two levels of word length, three levels of word pool and four lists. Overall, short words were better recalled than long words, $F(1,29) = 18.61$, $MSE = 1.57$, $p < .000$. Recall differed as a function of the word pool, $F(2,58) = 57.74$, $MSE = 1.87$, $p < .000$, and there were differences in recall across the four lists in each session, $F(3,87) = 9.28$, $MSE = 1.77$, $p < .000$. Only the list by word pool interaction reached significance, $F(6,174) = 2.29$, $MSE = 1.57$, $p < .05$.

To check the word length effects more closely, separate analyses were conducted for each of the three word pools. We replicate the Nairne et al. absence of word length effects with the Cowan et al. words, $F(1,29) = 2.28$, $MSE = 1.63$, $p = .14$, but reliable word length effects were obtained for the other two word pools, $F(1,29) = 5.92$, $MSE = 1.49$, $p < .05$ and $F(1,29) = 9.01$, $MSE = 2.18$, $p < .01$, for open and closed word pools respectively.

We also compared performance on the very first trial in each group. With the Cowan et al. pool there was no word length effect present, $t(29) = .33$, $p = .74$. However, there were strong word length effects on the first trial with three and seven phoneme word pools, $t(29) = 3.18$, $p = .003$.

In order to get a better understanding of what is happening in the task we analysed errors. The vast majority of errors were omissions or order errors. There was the odd extra list intrusion that usually had similar phonological characteristics to the forgotten target, the odd repetition of a word in the list, and the odd intrusion from a prior list or a perseveration of an earlier response, but these errors were very infrequent.

The mean number of errors collapsed across serial position is depicted in Figure 1, and were analysed in the same way as the targets. Overall, there were more
omissions with long words than with short words, $F(1,29) = 40.11$, $\text{MSE} = 3.07$, $p < .000$. Omissions differed as a function of the word pool, $F(2,58) = 108.32$, $\text{MSE} = 4.47$, $p < .000$, and the interaction was significant, $F(2,58) = 5.46$, $\text{MSE} = 2.32$, $p = .007$. There were more omissions on the long words in each word pool, $t(19) = 2.83$, $p = .008$, $t(19) = 4.89$, $p < .000$ and $t(19) = 4.22$, $p < .000$, for the Cowan pool, the open pool and the closed pool respectively. Thus the interaction reflects the weaker effect with the Cowan words and the stronger effects with the three and seven phoneme pools.

For transposition errors, the only significant effect was for word pool, $F(2,58) = 10.59$, $\text{MSE} = 4.22$, $p < .000$. Transposition errors were equally frequent for short and long lists in all three conditions.

Discussion

The results are quite straightforward. We replicate the Nairne et al. (1997) findings that word length effects do not emerge over four trials when the Cowan et al. words are used as experimental stimuli. The lack of a word length effect with these lists does not appear to be attributable to the relatively slow presentation times used by Nairne et al.; faster presentation rates typical of STM tasks produce the same null effect.

However, contrary to the Nairne et al position, word length effects are observed when a stronger manipulation of word length is utilised and it does not matter if a small word pool is used or a large pool of items. Furthermore, word length effects can be observed on the first trial in a block and across the subsequent three trials. The simplest explanation of the current results is that the absence of a word length effect in the Nairne et al. results is that they used a weak manipulation of word
length, albeit a manipulation that is at the heart of the decay plus rehearsal perspective.

The error analysis raised two additional points of interest. Firstly, PI effects in the form of intrusions from prior lists or perseverations of earlier responses were very rare. That is, there is very little direct evidence for PI effects in the task, yet there are strong word length effects with the three and seven phoneme words. Secondly, the major source of the word length effect is in terms of omission errors. Irrespective of the word pool, subjects were more likely fail to recall a long word than a short word. To maintain a PI explanation for word length effects it would be necessary to argue that omissions are attributable to PI. This possibility will be addressed later.

The results of the first experiment show that word length effects are observable across the first four trials in an experiment provided that a strong manipulation of word length is used. Whatever is producing the word length effect is operating on the very first trial. Secondly, the difference seems to be attributable primarily to the differential numbers of omission errors for short and long lists. Thirdly, there does not seem to be much evidence for PI in the task.

**Experiment 2**

In Experiment 2 we explore the same issues but with a task that directly measures PI effects. The task we use is an adaptation of the Tehan and Humphreys (1995, 1996, 1998) cued recall task. Subjects study trials that consist of either one block or two blocks of four words but they never know in advance whether a trial will be a one-block trial or a two-block trial. They are instructed to study each block of words as it is presented on the computer screen but they are asked to remember only the most recent block of four words; either the first block on a one-block trial or the
second block on a two-block trial. If the trial turns out to be a two-block trial, they are to forget the first block and concentrate on remembering the second block.

The two-block trials are the critical trials for manipulating PI. On these trials the second block always contains four words that have to be recalled in serial order. PI is manipulated by varying the material in the first block. In the no-interference condition the first block contains four consonants. For example a no-interference trial might look like "f d s w ! chair book cloud clock". In the interference conditions the first block contains another four words: "can window tile switch ! chair book cloud clock". The choice of consonants and words as appropriate materials was based on the fact that these materials produce reliable release from PI effects in the Brown-Peterson task (Wickens, 1972). PI can be observed or not observed in two ways. Firstly, recall of the words in the second block may be influenced by the characteristics of the first block, that is there might be differential recall of "chair book cloud clock" in the two conditions. Secondly, and more importantly, recall of words from the first block (can, window, tile or switch being recalled) or words from previous trials provides direct evidence for the influence of prior memories. If word length effects are due to differential amounts of PI, then the effect should be readily evident in different numbers of block-1 or prior list intrusions for short and long words.

Method

Subjects.

Twenty introductory level psychology students participated for course credit. None had participated in Experiment 1.

Materials
Subjects again studied two sets of 50 trials in which 10 were one-block trials and 40 were two-block trials. To construct the lists three pools were generated. The letter pool involved all English consonants except for the letter “y”. 120 short and 120 long words were selected from the MRC psycholinguistic database. The short words were all three phonemes in length whereas the long words were 7 phonemes in length. The words in each pool were matched for frequency and concreteness; all being low frequency and high imagery words.

Individual study lists were created for each subject. Each subject studied a set of lists that had been constructed from the open word pools and one that had been constructed from the closed word pools. The order of each set was counterbalanced across subjects. Open pool and closed pool variants were created for each subject.

The trials in the open pool condition were created in the following way. To create the first block items in the no-interference trials, four letters were randomly sampled from the letter pool 20 times. Forty items were then randomly selected from the short word pool and assigned to the four serial positions in the second block for 10 trials. Combining the letters and the short words produced the 10 short, no-interference trials. Another set of letters were produced and paired with forty long words that were randomly sampled from the long word pool and assigned to the four serial positions of the second block of another 10 trials. This produced the 10 long, no-interference trials. The remaining 80 short words were then randomly allocated to the four positions in both the first and second blocks for another 10 trials; the short interference trials. Finally, the remaining 80 long words were allocated to the two blocks to form 10 long, interference trials.

Five of one-block trials consisted of short words and five consisted of long words. The order of the 50 trials was then randomised for each subject.
To create the items in the closed pool list, we first randomly selected 16 items from each of the 120 item word pools and these 16 items then became the new short and long word pools used for creating the interference and no-interference trials. Creating the lists involved similar procedures to those used in the open pool condition, save that on each trial the relevant 16-item word pool was accessed and items were selected without replacement on each trial. This ensured that the short and long words were repeated frequently during the course of the experiment. Again 10 one-block trials were added and the order of the trials was randomised.

Procedure.

Participants were told that they would study a series of one or two-block trials and they only had to remember the most recent block that they had seen. Thus, when they discovered that a trial was a two-block trial they were to forget the first block and concentrate on remembering the second block because it would be on this block that they would be tested. It was also stressed that they would not know in advance whether that trial would be a one-block trial or a two-block trial, thus it was in their best interests to attend to the first block on each trial just in case it was a one-block trial.

Each trial started with the word READY presented in the middle of the screen. The list items were then presented at a rate of 1 item per second. In the case of two block trials, an exclamation mark (!) appeared for one second after the final item in the first block to indicate that the trial was a two-block trial. At the end of each trial a row of question marks (???) appeared on the screen as a cue for participants to write down the items of the most recent block in correct serial order. Subjects had a prepared answer sheet on which to record their answers.

Results
The results of the present experiment, where performance has been collapsed over serial position, are shown in Figure 2. A 2*2*2 repeated measures ANOVA with interference, word pool and word length as within-subject was conducted on correct recall. There was no reliable difference in recall between no-interference trials and interference trials, $F(1,19) = 2.12$, $\text{MSE} = .011$; recall was better for closed pool lists than for open pool lists, $F(1,19) = 8.21$, $\text{MSE} = .05$, and recall was better for short words than for long words, $F(1,19) = 78.04$, $\text{MSE} = .15$. No interactions were significant.

In looking at omission errors, there was no reliable difference in omissions between no-interference trials and interference trials, $F(1,19) > 1$, $\text{MSE} = .001$; there were fewer omissions in closed pool lists than for open pool lists, $F(1,19) = 35.92$, $\text{MSE} = .002$, and there were fewer omissions for short words than for long words, $F(1,19) = 75.79$, $\text{MSE} = .011$. The interaction between word length and word pool was significant, $F(1,19) = 6.04$, $\text{MSE} = .002$. Main effects analyses indicated that word length effects were present for both open pool, $F(1,19) = 92.66$, $\text{MSE} = .002$, and for closed pool, $F(1,19) = 41.62$, $\text{MSE} = .008$, but were obviously stronger in the open pool condition.

Transposition errors were next analysed. There was no main effect for type of interference trial, $F(1,19) > 1$, $\text{MSE} = .001$; there were fewer transpositions in closed pool lists than for open pool lists, $F(1,19) = 4.59$, $\text{MSE} = .002$, and there were fewer transpositions for short words than for long words, $F(1,19) = 26.49$, $\text{MSE} = .003$. None of the interactions were reliable.

Finally, when it came time to score intrusions in the closed pool condition it became clear that we could not say with certainty whether a particular intrusion came from the first block or from the previous list because it was often the case that
appeared in both places. To solve this problem we collapsed block-1 and prior list intrusions into one group.

For combined block-1 and prior list intrusions, there were more intrusions in the interference condition than the no-interference condition, $F(1,19) = 20.29$, $MSE = .014$, and there were more intrusions on closed pool lists than open pool lists, $F(1,19) = 10.76$, $MSE = .007$. This latter main effect was moderated by an interaction with interference condition, $F(1,19) = 4.54$, $MSE = .003$. Simple effects indicated that intrusion effects were very strong in the closed pool condition, $F(1,19) = 15.33$, $MSE = .014$, although they were still present in the open pool condition, $F(1,19) = 5.63$, $MSE = .002$.

**Discussion**

This experiment used a serial recall task in which PI was manipulated. PI can show up in this task either by differences in target recall or differences in intrusions from the first block or from previous trials. In the current experiment there is very little indication that PI is influencing recall at all. There is very little difference in correct recall between control and interference trials. Secondly, although there are differences in the number of intrusions, intrusions represent a minor source of error in the task. Immediate test performance does not seem to overly influenced by the material in the first block. The current results are thus consistent with a number of other findings that show very low levels of PI on an immediate test (Dempster & Cooney, 1986; Sanders & Willemsen, 1975; Tehan & Humphreys, 1995; Wickens, Moody & Dow, 1981)

While PI effects are not readily observable, word length effects are. Using three and seven phoneme words we replicate Experiment 1 by showing that short words are better recalled in position than long words. Furthermore, there is no
difference between short and long words when it comes to intrusions. This is not the pattern that would be expected if word length effects were caused by PI.

As was the case in Experiment 1, more omission errors were made with long words that with short words. However, in the current experiment more order errors were made with long words than the short words. In Experiment 1, this was not the case. This pattern suggests that the emergence of a word length effect after four trials in the Nairne et al. study might well be attributable to an increase in order errors for the long words as the number of lists increases. That is, the Nairne et al. data might not reflect PI per se, but reflect the fact that as the number of lists increase order memory for long words decreases.

The Nairne et al. (1997) assertion that word length effects are due to the build up of PI is hard to maintain. In their defence, there are signs in the bottom panel of Figure 2 that with a closed pool of long words, subjects are more likely to make a previous list intrusion, which is what one would expect from the Nairne et al. results. However, at this point there are so few of these intrusions that it is hard to argue for a significant role of PI in the word length effect.

General Discussion

The current research was designed to test one explanation for the word length effect. Nairne et al.'s (1997) findings that word length effects were not present across the first four trials in an experiment was taken as evidence for the proposition that word length was related to the build-up of PI. Evidence for word length effects across the first four trials and particularly on the first trial would weaken this argument. The results of Experiment 1 show that word length effects can be observed across the first four trials if a stronger manipulation of word length is employed. Finding word length effects on the very first trial is problematic for a PI account. Varying the number of
phonemes to determine word length produces its own set of problems in terms of
differential complexity (Caplan et al., 1992; Service, 1998), but empirically the data
show that word length effects can and do emerge on the very first trial.

The second issue is that in both experiments there is little direct evidence for
PI effects in the current experiments. In neither experiment were intrusions from prior
lists or prior blocks a strong source of error, and there were no differential effects of
word length associated with this measure. In Experiment 2 there was no difference in
target recall between the interference and no interference trials. The results of this
experiment are consistent with other research that indicates that immediate recall of
sub-span lists is immune to the effects of PI (Dempster & Cooney, 1986; Sanders &
Willemsen, 1975; Tehan & Humphreys, 1995; Wickens, Moody & Dow, 1981).

There is some indirect evidence for PI effects in the first experiment. In Figure
1, it would appear that in all conditions, performance on the second trial was not as
good as on the first trial. This is the pattern that Keppel and Underwood (1962) found
and is typical of performance on the build-up trials in the release from PI paradigm. If
one accepts this trial-2 decrement as evidence for PI, then it first indicates that the
manipulation of a minimum two-hour break between the different sets of lists
effectively produced temporal release from PI. On the assumption that the trial-2
decrement might reflect PI, we did a post-hoc analysis of word length effects across
the first two trials to see if word length effects changed as PI increased. The analysis
confirmed that short words were better remembered than long words and that
performance on trial one was reliably better than performance on trial two. None of
the interactions were significant. Thus, in the case where there is some evidence for
PI, there are no differential word length effects.
One way that the PI explanation could be maintained is if it is assumed that in Experiment 1 the PI effects are experiment wide and word length effects emerge across the experiment. If the first experiment is conceptualised as a single experimental session consisting of 24 trials in which 4-trial blocks of short and long words alternated, then word length effects in the latter four blocks could be explained in terms of the build up of PI across those trials as was the case in the Nairne et al study. From this perspective, the word length effects that emerged with the open and closed pool, three and seven phoneme lists, would be attributable to PI effects rather than being attributable to a stronger manipulation of word length.

We conducted a post-hoc test of this position by presenting 36 students with a single trial of either five three-phoneme words or five seven-phoneme words (the first trials of the closed pool condition in Experiment 1). On average, participants recalled 4.38 of the five short words and 3.38 of the five long words. This difference in mean recall was highly reliable, \( t(34) = 3.07, p < .05 \). Thus, robust word length effects were obtained when word length was manipulated between subjects and only one trial was presented. These word length effects are not attributable to PI.

On balance, the research indicates that robust word length effects can be observed where there is little evidence of PI. The error analyses of Experiments 1 and 2 show that the primary source of the word length effect is in the number of omission errors made. Long words are more likely to be omitted from the recall protocol than short words. Current models of immediate recall are not overly specific about what produces omission errors; they tend to concentrate more upon order errors. Given that there are probably many different possible factors that could produce an omission error, the finding of differential omission errors is probably not all that helpful. Having said this, it is possible to produce a PI-type explanation of omission errors.
Humphreys, Tehan, Boland and O'Shea (2000) tested parallel distributed processing assumptions of some memory models by exploring PI effects where similar and dissimilar items were subsumed by the same cue. They found that under some circumstances the two items could effectively block each other such that neither item could be recalled. That is, response competition resulted in an omission error. For this explanation to be relevant to the current data, the assumption would have to be made that words from previous trials were simultaneously active in memory at the time of the current trial and that these items were competitors for the targets. As such they could block recall of the target, such that neither the target not the competitor was produced. Furthermore, one would have to assume that this process was more likely to occur with long words than with short words. Given that the conditions for blocking in the Humphreys et al research bear little resemblance to those used here, such an explanation, though possible, remains unlikely.

In the introduction we indicated that there were a number of different explanations for the word length effect. The results do not speak to the other explanations in any definitive way. The current results using the three and seven phoneme words are explainable in terms of differential decay rates (although many other PI effects are not) as proposed by the "trace decay plus rehearsal" models, they are consistent with notions of linguistic complexity (Caplan, Rochon & Waters, 1992; Service, 1998) and they are consistent with intra unit interference (Neath & Nairne, 1995). However, we are reasonably sure that word length effects are not related to proactive interference effects in any substantial way.
References


Correspondence regarding this paper can be addressed to the second author via e-mail (tehan@usq.edu.au). The research was supported a FCAR research internship scholarship to the first author. The second author would like to thank Jim Nairne and the staff at Purdue University for support during sabbatical leave when the first draft of this paper was written. We would also like to thank Jim Nairne and two reviewers for helpful comments on this research.
Figure Captions

**Figure 1.** Correct serial recall and errors as a function of word length and word pool. Correct recall is the average number of items recalled per trial (max = 5) and errors represent the total number of errors per condition (max = 20). Error bars are standard error of the mean.

**Figure 2.** Correct serial recall and errors as a function of interference condition, word length and word pool. Correct recall is proportion of items correctly recalled in position and errors represent the proportion of errors per condition. Error bars are standard error of the mean.