Is Spoken Duration a Sufficient Explanation of the Word Length Effect?

Georgina Anne Tolan

Australian Catholic University

Gerald Tehan

University of Southern Queensland

Address correspondence to

Anne Tolan
Faculty of Arts and Sciences
McAuley Campus
Australian Catholic University
PO Box 247
Everton Park. QLD, Australia 4053
e-mail: a.tolan@mcauley.acu.edu.au

Abstract

The word length effect is one of the cornerstones of trace decay plus rehearsal models of memory. Words of long spoken duration take longer to rehearse than words of short spoken duration and as such suffer more decay and are thus less well recalled. The current experiment manipulates both syllable length and spoken duration within words of fixed syllable length in an aim to test the assumptions of the TDR model. Our procedures produced robust effects of both syllable length and spoken duration in four measures of the time it takes to pronounce the different types of words. Serial recall for the same materials produced robust syllable effects, but there was no evidence for duration effects. The results do not support TDR assumptions but are consistent with alternative explanations of the word length effect.

Is Spoken Duration a Sufficient Explanation of the Word Length Effect?

While interference effects are readily observable in tasks that employ long retention intervals, trace decay is still the preferred explanation for forgetting in immediate memory tasks. For example, many current models of immediate serial recall argue that short-term traces rapidly decay unless they are actively rehearsed. Brown and Hulme (1995) termed these models trace decay plus rehearsal (TDR) models. The widespread appeal of these models has recently led Nairne (2002) to term this approach the "standard model".

One of the cornerstones of the standard model is the word length effect; the finding that span for short words is better than span for long words. In the initial account of the word length effect, Baddeley, Thomson and Buchanan (1975) argued that it arose due to the fact that short words, because of their relatively short spoken duration, took less time to rehearse than long words. In this account rehearsal is required to preserve item availability in the face of rapid trace decay and since both decay and rehearsal happen in real time, the word length effect arises because more short words can be rehearsed in a given period of time than long words. While robust word length effects were found when syllable length was manipulated, the strong test of the decay assumption involved comparing memory span for words that had been matched on a number of dimensions, including number of syllables, but differing on spoken duration. Baddeley, et al., (1975) confirmed that span for the short-duration words was higher than that for the long-duration words. Trace decay plus rehearsal models that have developed since these early findings have been based on the assumption that the locus of the effect stems from the fact that longer duration items take longer to rehearse and suffer more decay as a result.

Cowan, Day, Saults, Keller, Johnson, and Flores (1992) have presented a variant of the duration account. While the Baddeley et al., account appears to stress decay during initial learning, Cowan et al., stresses forgetting during output. However, similar to Baddeley et al., the Cowan et al., explanation also emphasises the role of spoken duration in producing the word length effect; it is just that the locus of the effect differs.

In relatively recent times, growing dissatisfaction with the decay plus rehearsal account has emerged. Word length effects that are hard to explain from a rehearsal perspective have appeared in the literature (Coltheart & Langdon, 1998). There have been studies where there were no differences in spoken duration between different types of material, yet span differences have emerged (Schweickert, Grunnet & Hersberger, 1990; Tehan & Humphreys, 1988). Likewise, there have been instances where there were clear differences in spoken duration among items but no corresponding differences in span (Service, 1998).

Lovatt, Avons and Masterson (2000) have provided some of the most compelling evidence against the duration explanation. These researchers selected their stimuli along similar lines to Baddeley et al., (1975) and Cowan et al., (1992) in that they used a closed pool of bisyllabic short and long words that differed reliably in their spoken duration but were highly matched on other dimensions. With memory span as the dependent variable, they found no duration effects across four experiments. However, they did find duration effects when they used Baddeley et al's., original set of words as their stimuli. Lovatt et al., concluded that duration is not an important variable in determining the word length effect. Furthermore, they maintained that the emergence of duration effects when using the original set of Baddeley et al's., words was simply due to the idiosyncratic nature of these particular words.

Recently, Cheung and Wooltorton (2002) found the same pattern with one-syllable words. They varied the length of the vowel sound and were able to show reliable effects for spoken length of the short and long vowel item. In spite of the differences in spoken duration there were no differences in span for the short and long duration words.

The dissatisfaction with the standard model account of the word length effect has meant that alternative accounts have emerged (Caplan, Rochon & Waters, 1992; Neath & Nairne, 1995; Service, 1998). The particulars of these models are not overly important for current purposes, save that all these alternate explanations reject the rehearsal explanation; as such word length effects are predicted at the syllable level (or higher), but not necessarily within syllable. That is, these accounts predict span differences between one and two syllable words, for example, but not between short and long duration monosyllabic or short and long bisyllabic words.

In the current experiment memory for short and long duration monosyllables and short and long duration bisyllabic words is tested. By so doing, the aim is to extend the current empirical database and to discriminate the standard model from the emergent alternatives. The expectation from the standard model would be that a main effect for both syllable length and duration would be observed. From the alternative perspective, only a main effect for syllable length should be observed. These predictions are based on the assumption that reliable differences can be observed in duration measures for the different sets of items. It is to this issue that we now turn.

The key component of the standard model is the speed of the participant's rehearsal. Duration effects are not isomorphic with rehearsal. Spoken duration represents a property of a word whereas rehearsal speed is the property of a participant. Rehearsal speed is likely to depend in part upon the physical properties of the rehearsed items, but will also depend upon the interactions of those items with other items in lexical or semantic memory and is likely to also be constrained by motor factors involved in the planning and execution of rapid articulation. For instance, Tehan and Humphreys (1988) matched sets of high and low frequency words on the basis of their spoken duration as measured by digitised waveforms. If spoken duration was the sole determinant of rehearsal speed, rehearsal estimates should also have been equivalent. However, the high frequency words were read faster than the low frequency words. These differences in reading rate measures were accompanied by similar differences in span.

The above example should make it clear that one of the key issues in testing the standard model is selecting an appropriate measure of both duration and rehearsal. Moreover, it should be apparent that measures of spoken duration do not necessarily correspond to rehearsal measures, and that the standard model predicts that the latter should provide a better match to span differences than pure duration measures.

We have employed two measures of duration: human speech where no speed component was required and computer generated speech. Digitising unspeeded reading has been widely used to estimate spoken duration of words. Computer speech has not been used previously to our knowledge. The rationale was that in the human cognitive processing system, items are likely to interact with each other (e.g., spreading activation, target interactions in composite memories, etc.). The large literature on word identification and pronunciation confirms this likelihood. As such the human cognitive system may contaminate estimates of the purely physical characteristic of spoken duration. Removing the human component would well eliminate such contamination. We then tested rehearsal by using the two procedures most commonly used to evaluate span – rehearsal correspondences: the rapid and repeated articulation of three words and the rapid reading of fifty item lists.

Experiment 1

In this experiment we explored word length effects when participants were free to rehearse or under articulatory suppression. In contrast to most other experiments that have manipulated word length we used an open word pool in which each word only ever appeared once in the experiment. LaPointe and Engle (1990) showed that under suppression conditions, with an open pool, robust word length effects were still evident in the data. This does not happen when a closed pool is utilised (Baddeley et al., 1975). Moreover, we simultaneously explored duration and syllable effects by using short- and long-duration one-syllable words and short- and long-duration two-syllable words. At least when participants are free to rehearse, the standard model predicts that duration effects should be present in the data as well as syllable length effects. The non-duration models predict an absence of duration effects but the presence of a syllable length effect. Under suppression conditions, the standard model makes the prediction that all differences should be eliminated since rehearsal is prevented. Our reading of the alternative models is that syllable effects could well be expected to emerge in some of the models (e.g., Caplan, Rochon & Waters, (1992) because linguistic complexity does not differ under suppression conditions) but in others the differences might be expected to disappear as well (e.g., Neath & Nairne (1995) because suppression increases trace similarity).

As mentioned earlier in the introduction, there is prior research to expect that duration effects will not be observed for either the two syllable words (Lovatt et al., 2002) or monosyllabic words (Cheung & Wooltorton, 2002). Our expectation is that we will replicate these effects. In using an open word pool we also hope to replicate LaPointe and Engle's (1990) findings that under articulatory suppression, the syllable length effect is attenuated, but not eliminated.

Method

Participants

One hundred and ninety students from the Australian Catholic University and the University of Southern Queensland participated for partial course credit. The allocation of these people to the various conditions in the experiment is described below.

Materials

The manipulation of spoken duration was that used by Cheung and Wooltorton (2002). They generated pairs of monosyllabic words that had the same initial and terminal consonant phonemes but varied the length of the vowel sound. We used the same procedure to generate short and long one syllable and two syllable words.

The current set of one-syllable words was selected by going through the South Florida Rhyme Category Norms (Walling McEvoy, Oth & Neson, 1984) and selecting pairs of words that differed only on the length of the vowel sound. That is, the consonants were the same but the vowel was either short or long (e.g., **cut** and **cute)**. One hundred and fifty two pairs were generated in this way.

In much the same manner, we searched through the Toronto Word Pool (Friendly, Franklin, Hoffman & Rubin,. 1980) to find pairs of two-syllable words that could differ only on one phoneme, usually one of the vowels. An example of one such pair would be **riffle** – **rifle.** Eighty-four such pairs were generated. Note that in neither word pools were other stimulus properties like word frequency or imagability controlled.

With these word pools three types of lists were generated: Fifty item lists to be used to measure spoken duration and reading rate, three item lists to be used to measure speeded articulation and five item lists to be used in the serial recall trials. In all lists the short and long duration items were yoked. Thus, if for one participant the word “cut” appeared in the fourth serial position on one of the serial recall trials, the word “cute” would also appear in the fourth serial position on another serial recall trial. Likewise if “cut” were the middle word of a set of three for speeded articulation measures, “cute” would also appear as the middle word in obtaining an estimate for one-syllable long duration words.

Twenty sets of 50 item lists were generated for the duration and reading rate measures. In each case the 50 pairs were randomly selected from the one-syllable word pool and the two-syllable word pool. Thus, each set contained a 50- item list that contained one-syllable short-duration words; a different list that contained the corresponding 50 one-syllable long-duration words; a list that contained 50 two-syllable short-duration words; and the list that contained the corresponding two-syllable long-duration words. In this way participants had their own unique set of trials.

Twenty sets of three item lists were again generated by randomly selecting six pairs from each pool and them randomly assigning the six words to two trials. Thus, each person studied eight three-word trials; two of each combination of syllable and duration length. Again participants had their own unique set of trials.

Twenty sets of serial recall trials were also generated. For each participant there were 40 trials, 10 of each syllable-duration combination. To create the trials, 50 pairs were randomly selected from each word pool and then randomly assigned to trials and serial position within trials. The order of the 40 trials was then randomised. Participants had their own unique set of trials.

Procedure

Duration Measures. In the case of the computer speech each 50-item list was opened in a word processing application on the computer screen. The Narrator utility of Windows Xp was then activated to read the list. The time it took the computer to read the 50 items on each list was measured via a stopwatch. Estimates were taken from 10 different sets of trials (the equivalent of testing 10 different participants).

In the case of the human measures, the 50 items in each list were presented on a computer screen at the rate of one word every two seconds. Participants were instructed to simply pronounce each word in a normal voice as it appeared on the screen. Participant's responses were recorded onto the computer and the waveform for each list was subsequently edited to remove the silence between words. The resultant file contained the time it took each participant to pronounce the fifty words in an unspeeded fashion. Ten participants took part in this portion of the experiment.

Rehearsal Measures. In the case of the rapid articulation measures, each participant was instructed that they would be given three words and that once they were given the instruction to start, they were to repeat the three words as quickly as possible 10 times. The time taken to produce these thirty utterances on each trial was recorded by stopwatch. The order of the eight trials in each session was counterbalanced with the proviso that the short and long duration trials for each syllable length never followed each other immediately. Twenty participants completed this task.

In the case of the reading rate measures, twenty participants were asked to read a list of 50 items as quickly as possible and the time to read each list was measured by stopwatch. Each participant read each list twice. Once at the beginning of an experimental session and then some twenty minutes later after data had been collected for a separate experiment. Again the order of presentation of the lists was counterbalanced such that the short and long duration pairs for each syllable length never immediately followed each other.

Immediate serial recall. Each trial was presented visually via a computer screen. The five words on each trial were presented at a one second rate and participants were required to verbally recall the items in order immediately after the last word had been presented. One group of twenty participants from the Australian Catholic University (ACU) were free to rehearse during presentation while rehearsal was prevented for another group of twenty participants from ACU, by having them constantly repeat the word "the" during presentation of the items.

Results

Estimates of the spoken duration of the four different types of material are presented in Table 1. It is clear that the estimates are quite similar for machine speech, normal spoken speech and for rapid reading. When participants are required to repeat three items 10 times in succession, the estimates of spoken duration are much lower.

Duration Measures. The results above provide an estimate of the duration of each type of material. However, the results were analysed in terms of the time it took to read, say, or articulate the entire list. The results for the two duration measures are presented in Figure 1. where it is evident that for both measures two syllable words take longer to pronounce that one-syllable words and the words with long vowels take longer to pronounce than words with short vowels within each syllable length. This pattern was obtained on all ten occurrences with the computer speech and all ten human participants produced a short duration word advantage over the long duration words at each syllable length.

An alpha level of .05 was used to evaluate statistical significance in all analyses. The results of separate 2 * 2 repeated measures ANOVA confirmed that there was a reliable effect of syllable length, F (1,9) = 2704.85, MSe = .28, η^2 = .99, and F (1,9) = 12.46, MSe = 9.50, η^2 = .58, for machine and human speech respectively. Likewise there was a significant effect for spoken duration for machine, F (1,9) = 303.26, MSe = .15, η^2 = .97, and for human speech F (1,9) = 221.96, MSe = .67, η^2 = .96. There was a significant duration by syllable length interaction for machine speech, F (1,9) = 21.69, MSe = .14, η^2 = .70 but not for human speech, F (1,9) = .04, MSe = 1.49, η^2 = .00.

Rehearsal Measures. Because the pattern differed for articulation and reading time measures, the results are reported separately. However, it is apparent in Figure 1 that speeded articulation produces the same pattern of results as the duration measures. With reading rate the pattern changes substantially.

With speeded articulation, one syllable words were read faster than two-syllable words, F (1,19) = 10.23, MSe = 3.83, η^2 = .35, and the short duration words were articulated faster than the long duration words, F (1,19) = 4.72, MSe = 2.67, η^2 = .19. The interaction was not significant, F <1.

With reading rate one syllable words were articulated faster than two syllable words, F (1,19) = 59.56, MSe = 7.85, η^2 = .57, there was no difference in reading rates between the short and long duration words, F (1,19) = 2.59, MSe = 2.93, η^2 = .12. However the interaction just reached statistical significance, F (1,19) = 4.31, MSe = 2.21, η^2 = .18. For one syllable words, the short duration words were read faster than the long duration words, F (1,19) = 7.30, MSe = 2.36, but for the two-syllable words there was no reliable difference, F < 1.

Serial Recall: The proportion of items recalled in their correct serial position are presented in Figure 2. A 2*2*2 mixed design ANOVA with rehearsal condition as the between subjects factor and syllable length and duration as the within subjects factors. Performance was better when participants were free to rehearse than when suppression was required, F (1,38) = 47.76, MSe = 1.87, one syllable words were better recalled than two syllable words, F (1,38) = 50.09, MSe = .003, η^2 = .57, but there were no differences in spoken duration, F (1,38) = .69, MSe = .008, η^2 = .02. The syllable length by suppression interaction was significant, F (1,38) = 47.76, MSe = 1.87, η^2 = .17, suggesting that syllable length effects were strong when people were free to rehearse, F (1,19) = 32.53, MSe = .002, and weaker, but still reliable when suppression was required, F (1,19) = 18.83, MSe = .001. Importantly the three way interaction was significant, F (1,38) = 5.01, MSe = .004, η^2 = .12. The reason for this interaction is evident in Figure 1. Duration effects were not present in any condition save with two-syllable words under suppression conditions F (1,19) = 6.69, MSe = .001. However, in this instance the long duration words were better recalled than the short-duration words, the opposite of what is predicted.

Discussion.

In spite of our concerns about the possible contamination of physical measures of spoken duration by other cognitive influences, the effects of spoken duration were relatively consistent across the four measures used. In all four measures there were strong effects of syllable length. More importantly, there were strong duration effects for all measures save for short and long two-syllable words that were presented in long lists that had to be read as quickly as possible. It is clear that the spoken duration manipulation we employed was effective and all four measures displayed sensitivity to the manipulation.

Our concerns that spoken duration and rehearsal measures need not be equivalent also appear to receive support in the data. When the task was unpaced as was the case with the machine speech and with the normal pronunciation tasks, the effects of syllable length and duration were extremely robust as is evident by the number of participants who showed the expected pattern and by the level of power in the experiment as measured by the partial eta squared. However, when the task was paced as was required in the articulation and reading rate measures, fewer participants demonstrated the expected pattern and power estimates were reduced. Thus, while the physical properties of the material still produced an influence upon the tasks, they accounted for less of the variance in the task. Presumably this reduction in explained variance is due to the impact of other variables on these tasks that are not in operation in the unspeeded task. The data suggest that rehearsal speed is determined in part by the physical properties of the stimulus like spoken duration, but is also affected by other cognitive or articulation processes.

In turning to performance on the immediate serial recall task, we replicate a number of previous findings. Firstly, when people are free to rehearse we find a difference between one-syllable and two-syllable words. When word length is operationalised in this way we replicate the effects that have been found in many other studies. With articulatory suppression we again replicate the universal finding that suppression has a deleterious affect upon performance. Likewise, we replicate the LaPointe and Engle (1990) finding that with an open word pool, word length effects are attenuated but not eliminated. In the current experiment one syllable word are still better recalled than two syllable words.

The primary finding for current purposes is that there is no evidence for the predicted duration effect in any part of the experiment, even though there are real and robust differences in spoken duration for these items. Thus, while our duration and rehearsal measures were sensitive to duration effects, the serial recall task was not. In fact, the only place where duration influenced serial recall performance was with the two-syllable words under suppression conditions. However, here the result was in the opposite direction to that predicted by the TDR models.

The absence of duration effects replicates the results of Cheung and Wooltorton (2002) with monosyllabic words and those of Lovatt et al., (2000) with bisyllablic words. The current results show that the absence of duration effects is not just limited to instances where a closed pool of items has been used. The results generalise to open pool conditions. Moreover, finding that syllable length differences remain in serial recall when duration effects are absent provides direct support for those who argue that word length effects are due to other factors and not to differences in rehearsal speed (Caplan, Rochon & Waters, 1992; Service, 1998; Neath & Nairne, 1995).

Experiment 2.

While the TDR models see the locus of the word length effect occurring during the study phase of each trial, Cowan and his colleagues (1992) argue that the word length effect results from differential amounts of decay that occur during the recall process itself. Thus while recalling long items, there is more time available for decay to occur than while recalling short items. Cowan et al., presented mixed list data that supported this conclusion. However, Lovatt, et al. (2002) carried out replications of the Cowan et al., procedures using different sets of materials and was unable to replicate the Cowan et al., results.

In the current experiment we explore word length and duration effects with an order reconstruction task. Here the items are presented as for immediate serial recall but at the point of test, the items are presented again in a different order to that studied. The task is for the participant to reconstruct the original order of the items.

Cowan and his colleagues (1992) argue that the locus of the word length effect is due to problems in items availability associated with differential trace decay and not a function of differential problems in order discriminability. Thus, with an order reconstruction task both long and short items are represented intact and as such presenting the items again should offset the effects of trace decay. Under such conditions, there is probably no compelling reason to assume differential trace discriminability from the Cowan et al assumptions.

The decay during output position of Cowan and his colleagues (1992) was based upon verbal output of the list items. With an order reconstruction task it is possible to use alternative response procedures like computer mouse clicks that do not necessarily involve differential response times. With mouse clicks for example, there is no inherent difference in response times for short and long words as there is with vocal or written responses.

Of course, it is possible that participants are subvocally recalling the list items in serial order, which might produce differential times between mouse clicks for short and long words. In other words, if people are subvocally doing normal serial recall in a reconstruction task, then word length effects still might be present due to differential decay during output, but this should show up as differences in response times.

We believe that this explanation requires further empirical testing and as such we use a reconstruction test in the second experiment and record response times. A response time and recall difference between short and long words would provide support for the Cowan et al.,

(1992) argument that the word length effect occurs at the output stage. The presence of duration or syllable effects in the absence of corresponding differences in response time would suggest that decay was not a sufficient explanation for the effect.

Method

Participants.

Forty-two students from the Australian Catholic University participated in the experiment for partial course credit.

Materials.

The twenty sets of trials that had been generated for Experiment 1 were again used in this experiment.

Procedure

The items in each list were presented at a rate of one item per second on the computer screen. Following the last item, the items were again presented in alphabetical order simultaneously in a single column on the screen. Participants were asked to place the items in their original order of presentation by using the mouse to click on the item that was first presented and then click on the item that was presented second and so on. As each item was selected it was removed from the screen. Both accuracy and response times were measured. The items that were correctly recalled in position were determined for each type of list and the response times for these correct responses were averaged across serial position. These times included the preparation times, that is, the time from the presentation of the list items in alphabetical order to the first item being selected.

Results.

The results of the experiment are summarised in Figure 3. Overall, one-syllable words were better recalled than two-syllable words, F (1,41) = 4.69, MSe = 0.01; there was no effect of spoken duration, $F(1,41) = .25$, $MSe = 0.01$, however, the syllable length by duration interaction was significant, $F(1,41) = 5.11$, $MSe = 0.01$. Simple effects analysis revealed that there was no significant duration effect for the one syllable words, $F(1,41) = 2.19$, $MSe = 0.01$, but for the two-syllable words, the long duration words were better recalled than the short duration words, $F(1,41) = 4.96$, $MSe = 0.01$.

For response times, there was no significant effect of syllable length, $F(1,41) = .64$, $MSe = 0.03$, or duration, $F(1,41) = 2.35$, $MSe = 0.04$, nor was the interaction significant, $F(1,41) = .68$, $MSe = 0.02$.

When questioning participants after the task, a number mentioned that they had adopted a first letter strategy. That is, instead of reading the words, they concentrated on remembering the initial consonant or consonant cluster and reordered the items on this basis. If all participants had adopted such an approach one would not expect to see strong word length effects of either type, nor would one expect to see corresponding differences in response times.

In response to this problem we divided the sample into three groups on the basis of the strength of the syllable length effect. One group of 14 participants generated reverse word length effects, another group of 10, showed weak effects (ranging from 2% to 8% advantage) and the final group of 18 participants showed a strong one-syllable advantage (ranging from 10% to 38% advantage). It is this latter group that would provide a reasonable test of the Cowan et al., (1992) assumptions. While this group was selected on the basis of their overall differences between one-syllable and two-syllable words, there was no corresponding difference in duration effects in their serial recall, $F(1,17) = .50$, $MSe = 0.01$. In looking at the response times for this group one difference did emerge. Responses were slower for the two-syllable words than the one-syllable words, $F(1,17) = 5.12$, $MSe = 0.03$, but there was no effect of duration, $F(1,17) = 2.32$, $MSe = 0.06$, nor was the interaction significant.

Discussion.

The first thing to note about this experiment is that again there is absolutely no evidence that supports spoken duration having any influence upon performance. Across the group as a whole and even in the group that produce strong recall differences in terms of syllable length, there is no hint that short duration words are better recalled than long duration words.

The second aspect of performance is that again syllable effects in order reconstruction were apparent across the group as a whole without a corresponding difference in response times. However, in the group that did show a strong syllable length effect in reconstruction there was a corresponding difference in response times. Thus, there is some support for the Cowan et al., (1992) notion that the word length effect occurs during output. However, while this might be the case, there is no compelling evidence that the effect is due to differential decay during output. The fact that the items are represented at test should offset the effects of decay and there are no duration effects, so the locus of the syllable effect is unlikely to be due to decay.

General Discussion.

Testing the duration explanation of the TDR models is reliant upon there being measurable differences in spoken duration. In this experiment we have used four different measures to assess duration differences across three different and independent samples (i.e., four samples if one considers the computer to be a sample). Across all four measures there are strong syllable effects and in seven of the eight opportunities, there are strong duration effects. The fact that duration effects replicate across different samples and across different measures convinces us that there are real physical differences in the spoken traces for the short and long words. The fact that the duration effects are weaker in the articulation and speeded reading also indicate that physical attributes of words, although still detectable, interact with other variables when the task more closely approximates speeded rehearsal. In short, our measures capture the patterns of performance that people would assume to operate when participants verbally rehearse during presentation of list items on a serial recall task.

When it comes to serial recall, we have three separate and independent samples. Again the pattern is very consistent across groups. Syllable length effects are present in all groups and the expected duration effects are present in none. It is hard to conclude that spoken duration has any influence upon recall. On the basis of the current findings, there is strong support for those recent explanations of the word length effect that posit the locus of the effect in terms of other processes besides trace decay.

One possible argument in favour of the decay position would be to maintain that the absolute differences in spoken duration are less pronounced when the vowel is manipulated as opposed to when syllable length is manipulated and as a consequence recall differences should be stronger for syllable effects than duration effects. Duration estimates are smaller than syllable estimates when reading rate and computer speech are used to measure duration, but when unspeeded human speech and rapid articulation measures are used there is very little difference in absolute differences between syllable and duration measures. Given that articulation speed probably has the best face validity of our four measures, as a measure of rehearsal speed, the data do not seem to support the argument. That is, in Figure 1 there appears to be a monotonic increase of the same magnitude across the four types of material. There is no corresponding monotonic increase of similar magnitude in serial recall.

The current research indicates that word length effects, based upon the number of syllables in the word, are readily apparent in immediate serial recall under a range of experimental conditions. Moreover, word length effects based upon differential rehearsal rates, that are a function of differences in the temporal duration of words, does not appear to be supportable. The absence of duration effects in serial recall has been reported in the literature before and has been taken as evidence against trace decay plus rehearsal models. The current research makes the same argument but it also provides a specific test of the non-decay models that syllable-based word length effects should be observed in the absence of duration-based word length effects. The results are consistent with these models but do not provide a basis for discriminating between them.

References

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

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.

Caplan, D., Rochon, E., & Waters, G. (1992). Articulatory and phonological determinants of word length effects in span tasks. *Quarterly Journal of Experimental Psychology, 45A*, 177-192.

Cheung, H., & Wooltorton, L. (2002). Verbal short-term memory as an articulatory system: Evidence from an alternative paradigm. *Quarterly Journal of Experimental Psychology, 55 (1),* 195 - 223.

Coltheart, V., & Langdon, R. (1998). Recall of short word lists presented visually at fast rates: Effects of phonological similarity and word length. *Memory and Cognition, 26*, 330-342.

Cowan, N., Day, L., Saults, J.S., Keller, T.A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language, 31*, 1-17.

Friendly, M., Franklin, P. E., Ho man, P., & Rubin, D. C. (1982). The Toronto word pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods and Instrumentation, 14,* 375 - 399.

LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1118-1133.

Lovatt, P., Avons, S. E., & Masterson, J. (2000). The word-length effect and disyllabic words. *Quarterly Journal of Experimental Psychology, 53A*, 1-22.

Lovatt, P., Avons, S. E., & Masterson, J. (2002). Output decay in immediate serial recall: Speech time revisited. *Journal of Memory and Language*, *46*, 227-243.

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

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

Schweickert, R., Guenert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science, 1*, 74-77.

Service, E. (1998). The effect of word length on immediate serial recall depends upon complexity, not articulatory duration. *Quarterly Journal of Experimental Psychology, 51A*, 283-304.

Tehan, G., & Humphreys, M. S. (1998). Creating proactive interference in immediate recall: Building a dog from a dart, a mop and a fig. *Memory and Cognition, 26,* 477 - 489.

Walling, J. R., McEvoy, C. L., Oth, J. E., & Nelson, D. L. (1984). The University of South Florida rhyme category norms. Unpublished manuscript, University of South Florida.

Table 1

Estimates of spoken duration (msec) for one and two syllable words with long and short vowels.

	Machine	Speaking	Articulation	Reading
1-syllable short	0.445	0.472	0.389	0.487
1-syllable long	0.488	0.548	0.416	0.513
2-syllable short	0.617	0.542	0.437	0.594
2-syllable long	0.666	0.620	0.462	0.593

Figure Captions:

Figure 1.

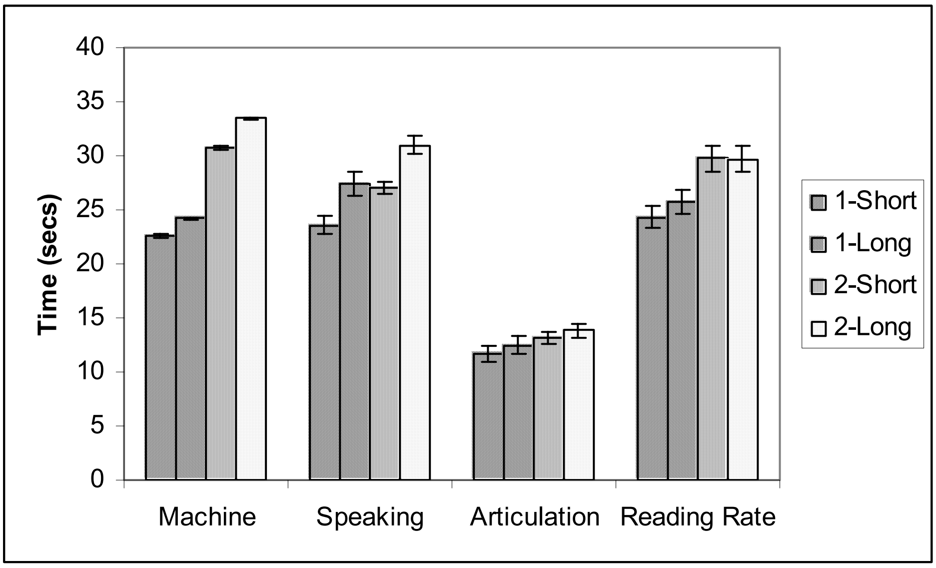
Pronunciation time estimates for one and two syllable words with short or long vowels. Error bars represent standard error of the mean.

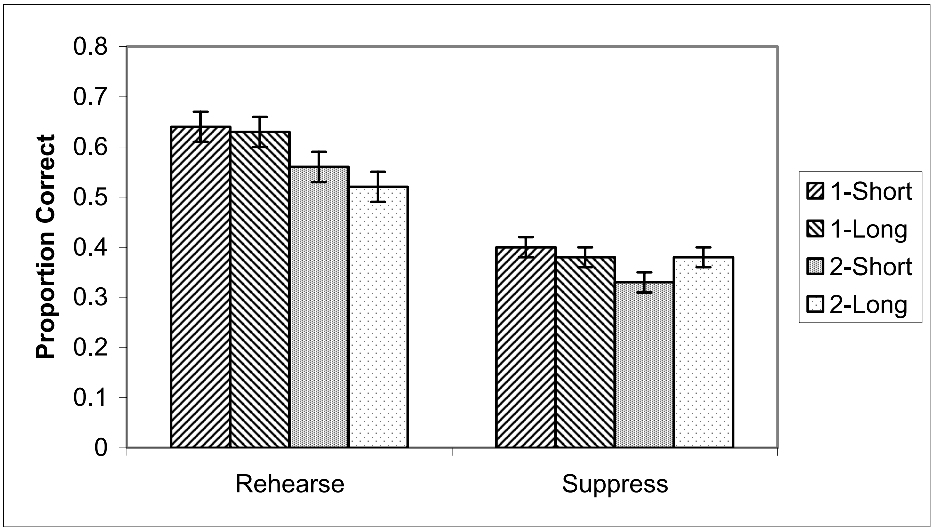
Figure 2.

Correct recall for one and two syllable words with short or long vowels as a function of rehearsal conditions. Error bars represent standard error of the mean.

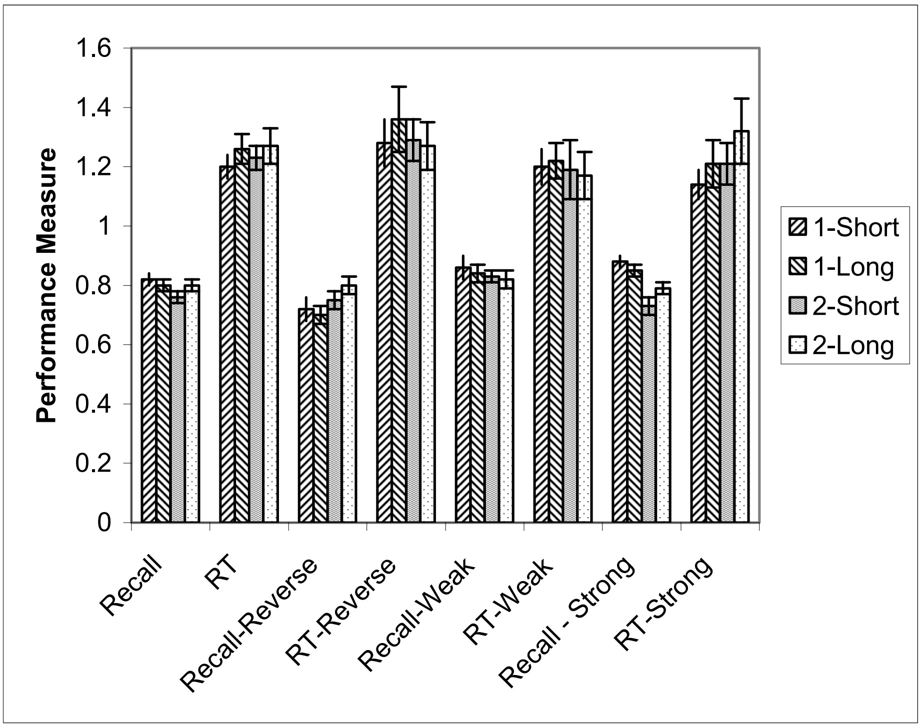
Figure 3.

Proportion of items placed in correct serial position and response time (secs) for correctly ordered items as a function of word length. Note y axis simultaneously represents proportion correct (0-1.0) for serial recall and response time (0-1.4secs)


40
35
30
25
20
15
10
5
0
Time (secs)
Machine
Speaking
Articulation
Reading Rate
1-Short
1-Long
2-Short
2-Long


0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Proportion Correct
Rehearse
Suppress
1-Short
1-Long
2-Short
2-Long

.


1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Performance Measure
Recall
RT
Recall-Reverse
RT-Reverse
Recall-Weak
RT-Weak
Recall - Strong
RT-Strong
1-Short
1-Long
2-Short
2-Long