

## Proactive interference plays a role in the word-length effect

JAMES S. NAIRNE and IAN NEATH  
*Purdue University, West Lafayette, Indiana*

and

MATT SERRA  
*Duke University, Durham, North Carolina*

The word-length effect is the advantage that lists of short words show over lists of long words in immediate recall. Typically, the word-length effect is used as strong support for decay-based theories of immediate retention. However, the present two experiments show that the traditional word-length effect emerges only after several trials in an experimental session, a result reminiscent of earlier findings in distractor-based short-term forgetting experiments. These results suggest that any complete account of the word-length effect will need to consider a role for proactive interference.

Research conducted over the past two decades has firmly established that the capacity of short-term memory is *not* seven-plus or minus-two items, or chunks (e.g., Miller, 1956). Instead, span, which is often used as a measure of memory capacity, is better described as a function of pronunciation or speaking time. The probability of correctly recalling, in order, a set of items immediately after presentation declines in a nearly linear fashion with the time that it takes to repeat the to-be-recalled list aloud. Lists containing long words are harder to recall than lists of short words (the *word-length effect*), and memory span is roughly equal to what we can repeat aloud in 2 sec (Baddeley, Thomson, & Buchanan, 1975; Schweickert & Boruff, 1986).

This relationship between speaking time and span strongly implies that it is time per se that is responsible for performance limitations. Indeed, many current theoretical conceptions of short-term memory propose that autonomous decay—the loss of information as a property of time—is a major source of forgetting in immediate memory tasks (e.g., Baddeley, 1992; G. D. A. Brown & Hulme, 1995; Burgess & Hitch, 1992; Cowan, 1995; alternatively, see Nairne, 1990; Neath & Nairne, 1995). In Baddeley's working-memory model, for example, memory traces are assumed to decay in approximately 2 sec unless actively refreshed by internal processes of rehearsal. Long words take up more rehearsal time than short words, so fewer long words can be refreshed within the 2-sec decay window. The size of the word-length effect, as well as a particular individual's memory span, is thereby determined by how successfully one can use the processes of rehearsal to offset the fixed forces of decay.

However, not all memory researchers are convinced that autonomous decay is a useful tool for explaining forgetting; in fact, the concept of decay has been repeatedly criticized historically in both the human and animal literatures (see Capaldi & Neath, 1995, for a review). Critics like McGeoch (1932) and Osgood (1953) rejected decay as the operative source of forgetting on the basis of, in part, two empirical facts: (1) Memory sometimes improves rather than declines over time (e.g., spontaneous recovery or reminiscence; see Wheeler, 1995), and (2) with time held constant, memory performance often depends on the nature of the events that intervene between presentation and recall (e.g., Proctor & Fagnani, 1978). In a classic analogy, forgetting was compared to the rusting of an iron bar over time—the accumulation of rust is correlated with the passage of time but is not caused by time (see McGeoch, 1932; Osgood, 1953). Time alone does not cause forgetting; it is the accumulation of other events *in* time (i.e., interference) that produces the performance loss.

Why, then, is decay so popular among researchers of short-term memory? The reason traces partly back to the classic short-term forgetting experiments of J. Brown (1958) and Peterson and Peterson (1959). In those experiments, subspan lists of letters were forgotten rapidly when subjects were distracted from rehearsal by a counting task that intervened between presentation and recall (e.g., counting backward by threes). Because letters and digits are so dissimilar, there seemed to be no obvious way that retroactive interference, which depends on similarity, could explain the significant forgetting. More recently, Cowan, Lichty, and Grove (1990) found that memory for unattended speech sounds declined rapidly during a period in which subjects were simply reading silently; because of the different modalities of the to-be-remembered and distracting tasks, as well as the differences in item types, there was no reason to expect that one could interfere substantially with the other.

---

Correspondence about this article may be addressed to J. S. Nairne, Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907-1364 (e-mail: nairne@psych.purdue.edu).

Yet, retroactive interference is not the only kind of interference that is likely to be present in a short-term memory environment. Information presented on early trials in a session can also interfere with one's ability to recall information from later trials. Bennett (1975) showed that when people are asked to recognize items from a just-presented memory list, items from previous trials lead to false alarms; moreover, for subjects, such false alarms are more likely to come from distractors from immediately preceding trials than from distractors from trials farther removed. Similarly, Estes (1991) found that when subjects intrude items into short-term recall, such intrusions tend to be items from immediately preceding memory lists (i.e., trial  $n-1$ ). These data suggest that short-term forgetting may arise, at least in part, from an inability to discriminate current trial information from information presented in the nearby past. It is not retroactive interference that produces forgetting in most immediate memory tasks; it is proactive interference.

Keppel and Underwood's (1962) classic demonstration that forgetting in the Brown-Peterson distractor task depends on where in the experimental session performance is assessed confirms that proactive interference plays an important role in short-term retention. If performance is examined only for the first trial or two in a session, little if any forgetting is found, even after 10–20 sec of counting backward. The characteristic short-term forgetting curve emerges only after subjects have experienced several memory trials. Note that an appeal to autonomous decay, at least in any simple form, cannot explain such proactive effects. Instead, additional assumptions about memory systems, or memory processes, are needed to explain the data (Atkinson & Shiffrin, 1968; Cowan, 1995; Waugh & Norman, 1965). In the general discussion, we will discuss some of these assumptions, and how they relate to the data reported in the two experiments of the present study.

The purpose of the present experiments was to investigate whether proactive effects of the type demonstrated for short-term forgetting by Keppel and Underwood (1962) also play a role in the word-length effect (the finding that short words yield higher memory spans than long words). Once again, the word-length effect is often used as prototypical support for decay-based forgetting in short-term memory: Long words produce smaller spans because long words take longer to articulate; more units, as a consequence, are likely to be lost through decay before they can be retrieved for covert articulation. Notice that if the word-length effect arises as a consequence of such decay processes, then we would expect it to occur throughout an experimental session. There is certainly no a priori reason to assume that the properties of decay will depend on the existence of prior trials.

## EXPERIMENT 1

In Experiment 1, we conducted a standard word-length experiment, using the same stimuli and procedures em-

ployed by others, but with subjects receiving only *four* immediate memory trials. Everyone was required to recall a five-item word list, immediately after presentation, and each list contained a collection of either all short or all long words. Word length was manipulated between subjects, and no practice trials of any kind were presented. We were interested primarily in learning whether a word-length effect would be observed under these conditions, when potential interference from prior trials was minimal. No word-length effect on the first few memory trials in a session would suggest that demonstrations of the standard word-length effect, like forgetting in the Brown-Peterson distractor task, depend on the presence of some kind of proactive interference.

## Method

**Subjects.** Four hundred and forty Purdue University undergraduates volunteered to participate as subjects in exchange for credit in introductory psychology courses. They were arbitrarily assigned to one of two groups; one group ( $n = 220$ ) received short words, and the other group ( $n = 220$ ) received long words.

**Materials.** The stimuli were selected from Cowan et al. (1992). The long words were *coerce*, *humane*, *morphine*, *voodoo*, and *zygote*. The short words were *decor*, *ember*, *hackle*, *pewter*, and *wiggle*. Cowan et al. (1992) had shown previously that the long words in this set took more time to pronounce than the short words and that these stimuli produced the standard word-length effect under conditions very similar to those employed in the present experiment. Importantly, the two sets were approximately matched for frequency, number of syllables, and number of phonemes. For the 26 phonemes in each word set, 19 occurred only once in the short set (i.e., in only one word) and 21 occurred only once in the long set.

**Procedure.** The subjects were tested individually. The words were displayed in the center of an IBM-compatible or Apple Macintosh LC screen, in lower case, one at a time, for 2 sec, with a 2-sec interpresentation interval. Following convention, a given subject received the same word set (containing either all short or all long words) throughout the experimental session, but the words were displayed in a new random order on each trial. Regardless of the group, subjects were read the same instructions:

In this experiment, the computer will show you five words, one at a time, in the center of the screen. Please do not say anything out loud while the words are being displayed. After the final word has been shown, the computer will pause for 20 seconds. During this interval, we would like you to write down the words that you saw in the exact order in which they were presented.

An experimenter verified that the strict serial recall instructions were followed. Twenty seconds were allowed for written recall; following this, the computer beeped to signal that a new trial was about to begin, and 2 sec later, the next list was presented. Each subject received four lists and no practice lists.

## Results and Discussion

Figure 1 displays the proportion of words correctly recalled, separated for short and long lists, as a function of trial number. An item was scored as "correct" only if it was reported in its proper serial position. Overall, the mean proportion of correct recall was .73 in the long-word condition and .72 in the short-word condition. Clearly, no indication of the standard word-length effect (short recalled better than long) can be seen on any of the four trials.

This conclusion is supported by a 2 (word length)  $\times$  4 (trial)  $\times$  5 (serial position) analysis of variance (ANOVA).

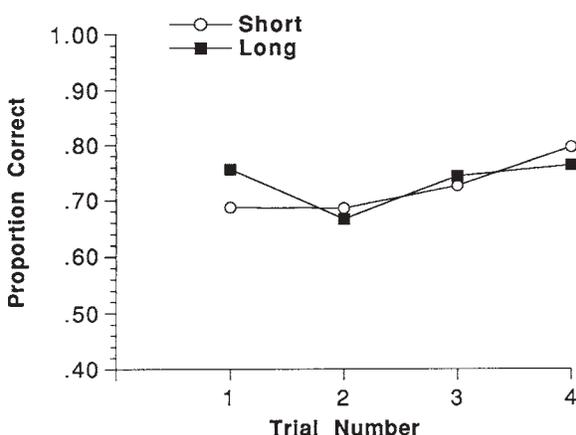


Figure 1. Mean proportion correct for groups receiving lists of short words or lists of long words in Experiment 1, plotted as a function of trial number in the session.

Reliable effects were found for trial [ $F(3,1314) = 13.11$ ,  $MS_e = 0.31$ ,  $p < .01$ ] and for serial position [ $F(4,1752) = 208.72$ ,  $MS_e = 0.14$ ,  $p < .01$ ], but not for word length [ $F(1,438) < 1$ ]. Word length did interact significantly with trial [ $F(3,1314) = 3.71$ ,  $MS_e = 0.31$ ,  $p < .05$ ], but the interaction seemed attributable to a long-word advantage on the very first trial in the session—that is, a reverse word-length effect [ $F(1,438) = 8.41$ ,  $MS_e = 0.06$ ,  $p < .01$ ]. Both the trial  $\times$  serial position and the three-way interaction were also reliable [ $F(12,5256) = 7.07$ ,  $MS_e = 0.12$ ,  $p < .01$ , and  $F(12,5256) = 1.75$ ,  $MS_e = 0.12$ ,  $p = .05$ , respectively]. A close look at the data revealed that the former was due to an increased recall of end-of-list items with trial number, but no straightforward interpretation of the three-way interaction seemed possible. The group  $\times$  position interaction was not reliable [ $F(4,1752) = 1.30$ ,  $MS_e = 0.14$ ,  $p > .25$ ].

Regardless of how one chooses to interpret the interactions, the overall results are inconsistent with a straightforward decay account of the word-length effect. No short-word advantage was present on any trial, even though performance was excellent and, presumably, the subjects were engaging in strategic rehearsal to help maintain list items in memory. Because it was possible to collect only a small number of observations from each subject, it is of course possible to raise the specter of power to explain our failure to obtain the traditional word-length result. However, it is important to keep in mind that we used a relatively large number of subjects ( $N = 440$ ), and statistically significant effects were obtained for serial position and for the trials variable. There was even a significant effect of word length on the very first trial, but the advantage was for the long rather than the short words. Consequently, in our view, it is unlikely that adding more subjects would significantly change the pattern of results.

At face value, these results seem directly compatible with those obtained by Keppel and Underwood (1962), who showed that forgetting in the Brown–Peterson task depended on prior trials, emerging only after several trials

in the session. If the word-length effect is due to increased forgetting of long words, compared with short words, then a failure to obtain the effect under conditions in which short-term forgetting is minimized is not particularly surprising. To the extent that forgetting processes are minimized or inoperative, perhaps due to the absence of significant proactive interference, we might expect comparable performance between short and long words. Still, given the importance of these results for theory, we decided to conduct another experiment in an effort to (1) replicate the results of Experiment 1, and (2) establish that the standard word length would emerge in our experimental context provided that a sufficient number of trials were presented.

## EXPERIMENT 2

Experiment 2 was an exact replication of Experiment 1 except that the subjects received 24 trials instead of only 4. Once again, we were interested in comparing the immediate retention of lists containing either short or long words as a function of the number of trials into the session. The number of trials was extended to 24 in Experiment 2 in order to confirm that the word-length effect would indeed emerge later into the session.

### Method

**Subjects.** The subjects were 118 Purdue undergraduates who participated for course credit. They were arbitrarily assigned to either the short ( $n = 59$ ) or long ( $n = 59$ ) word group.

**Materials and Design.** Exactly the same materials and experimental procedures were used in Experiment 2. The only difference was that subjects received 24 rather than 4 trials in the session. As in Experiment 1, no practice trials were included.

### Results and Discussion

Figure 2 shows immediate memory performance, separated by list type, as a function of trials. Performance is collapsed into blocks of four trials, so the first two points on the left in the figure can be compared with average performance in Experiment 1. An ANOVA of the data re-

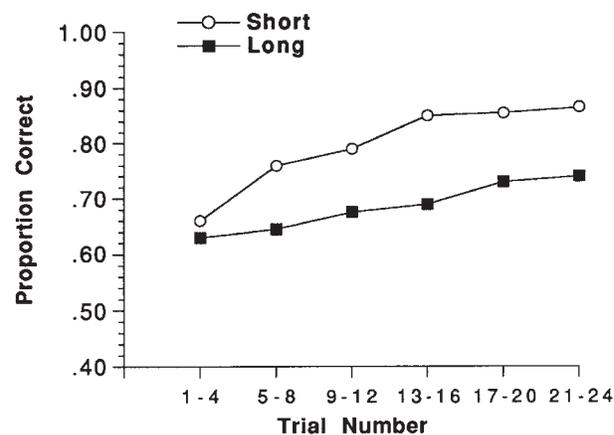


Figure 2. Mean proportion correct for groups receiving lists of short words or lists of long words in Experiment 2, plotted as a function of trial. Each data point represents average performance collapsed over a block of four trials.

vealed significant main effects of trial block [ $F(5,580) = 22.89$ ,  $MS_e = 1.52$ ,  $p < .001$ ], serial position [ $F(4,464) = 104.06$ ,  $MS_e = 1.01$ ,  $p < .001$ ], and, importantly, word length [ $F(1,116) = 14.45$ ,  $MS_e = 12.74$ ,  $p < .002$ ]. The critical word length  $\times$  trial block interaction was also reliable [ $F(5,580) = 2.35$ ,  $MS_e = 1.52$ ,  $p < .05$ ], as was the interaction of serial position  $\times$  trial block [ $F(20,2320) = 5.13$ ,  $MS_e = .406$ ,  $p < .001$ ]. None of the other interactions reached statistical significance.

Two important findings emerged from this experiment. First, we were able to replicate the main results of Experiment 1: No word-length effect was found on early trials in the session. For the very first trial in the session, average retention performance for lists containing all short words was .70; for long words it was .68. This difference did not approach statistical significance [ $t(58) < 1$ ]. For the first block of four trials, as in Experiment 1, again there was no significant difference between short and long words [ $F(1,276) = 1.20$ ,  $MS_e = 3.39$ ,  $p > .25$ ]. Clearly, when one looks only at performance on the early trials, there is little, if any, evidence for the standard short-word advantage.

Second, as the word length  $\times$  trial interaction indicates, the word-length effect did emerge over the course of the session. In fact, by the second block of four trials, lists of short words were recalled significantly better than lists of long words [ $F(1,276) = 10.22$ ,  $MS_e = 3.39$ ,  $p < .002$ ], and the short-word advantage continued for the remainder of the trial blocks. This finding is important because it verifies that our experiment was sensitive enough to detect the standard advantage of short over long words—it simply requires several trials before evidence for the word-length effect will emerge.

## GENERAL DISCUSSION

The results of these two experiments are clear. Using standard experimental procedures and stimulus materials identical to those used by other laboratories (e.g., Cowan et al., 1992), we were unable to find evidence for a word-length effect on the first few trials in the experimental session. Lists of long words were recalled as well as, and in some cases better than, lists of short words. Moreover, it appears unlikely that one can attribute this null result to any simple scale or power problem. Performance was well below ceiling on the first few trials in both experiments, and other effects, such as serial position and trials, reached acceptable levels of statistical significance. In addition, Experiment 2 revealed that the standard word-length effect was easily replicated in our experimental context, provided that subjects had experienced several prior trials in the session. Consequently, these results add an important boundary condition to the empirical database on the word-length effect.

From a theoretical standpoint, the absence of a word-length effect early in the session is significant because it is not easily explained from current working-memory perspectives (see, e.g., Baddeley, 1992, 1996). According to working-memory advocates, the word-length effect arises because it is easier to rehearse short words internally, and thus more memory representations for short (compared with long words) can be refreshed within the fixed decay window. Decay is assumed to be a fixed property of the memory system, acting in the same way that a force like gravity acts to limit the ability of a juggler to keep balls aloft; there is no reason to expect it to change as a function of where in the session the testing occurs. Decay should be in force on every trial in the session, from Trial 1 onward, working to dissolve the traces of to-be-remembered items.

It is conceivable that one's ability to rehearse strategically might change throughout an experimental session. The current data could be explained, then, if one assumes that either (1) subjects fail to rehearse on early trials in the session, or (2) subjects do not rehearse short and long words differentially until later in the session. If one "learns" to rehearse, then perhaps it takes a few trials for subjects to become relatively more efficient at rehearsing short than long words. Performance does improve over trials in the session, and this improvement is greater for short words, a finding that is consistent with a rehearsal-learning account. On the other hand, performance is quite good on the first trial in the session (averaging around 70%), which is difficult to explain without assuming some strategic rehearsal process.

Another way to explain these data is to assume that performance is mediated by different memory stores on early and later trials in the session. Researchers have frequently proposed that performance in immediate memory tasks reflects some combination of retrieval from short-term and long-term memory systems (e.g., Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). For example, as Cowan (1995) has recently argued, early in an experimental session, when there is little chance of interference from other trials, subjects may be able to recover trial information easily from long-term memory. As the number of trials increases, retrieval from long-term memory becomes more difficult—because current trial items cannot be discriminated from previously presented items—and subjects might then strategically shift entirely to recovery from short-term store. Once this shift occurs, however, word-length effects will emerge because of the role that decay presumably plays in the short-term-memory system.

This "shifting stores" account is attractive because it was proposed initially by Atkinson and Shiffrin (1968) as an explanation for the forgetting data of Keppel and Underwood (1962). That is, one finds little evidence for distractor-based forgetting early in the session because subjects are able to recover trial information from long-term memory. In fact, one could argue that the present results—no word-length effect on early trials—confirms a prediction of this account. At the same time, there is, at present, no direct empirical evidence to confirm that strategic "shifts" from short- to long-term retrieval occur within an experimental session, so the account remains primarily speculative. Moreover, the explanatory appeal of this account is gained at some cost—it essentially assumes that subjects rely on short-term store only as a "back-up" system to be used when recovery from long-term store is rendered ineffective by prior trials. A subject's first line of defense is to recover information from long-term memory (or a combination of recovery from short- and long-term stores); pure recovery from short-term store proceeds only when long-term representations become difficult to access.

At face value, the results of the present two experiments, like those of Keppel and Underwood (1962), suggest that proactive interference is an important source of forgetting in immediate memory tasks. The other obvious implication, of course, is that proactive interference must play a role in any complete explanation of the standard word-length effect. As it stands, these results are inconsistent with decay-based accounts like those currently proposed by most working-memory theories (although more sophisticated versions of decay theory, perhaps in conjunction with assumptions about interference, can certainly handle the results). The present results also add to a growing body of evidence suggesting that the word-length effect is a rather complex phenomenon, influenced by multiple factors (see G. D. A. Brown & Hulme, 1995; Cowan, 1995; Nairne, 1996). It is clear that the word-length effect should and will continue to play an important role in shaping theories of immediate retention, but the exact nature of that role remains uncertain at the moment.

## REFERENCES

- ATKINSON, R. C., & SHIFFRIN, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89-105). New York: Academic Press.
- BADDELEY, A. D. (1992). Working memory. *Science*, **255**, 556-559.
- BADDELEY, A. D. (1996). The concept of working memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 1-27). Hove, U.K.: Psychology Press.

- BADDELEY, A. D., THOMSON, N., & BUCHANAN, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **14**, 575-589.
- BENNETT, R. W. (1975). Proactive interference in short-term memory: Fundamental forgetting processes. *Journal of Verbal Learning & Verbal Behavior*, **14**, 123-144.
- BROWN, G. D. A., & HULME, C. (1995). Modeling item length effects in memory span: No rehearsal needed? *Journal of Memory & Language*, **34**, 594-621.
- BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, **10**, 12-21.
- BURGESS, N., & HITCH, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory & Language*, **31**, 429-460.
- CAPALDI, E. J., & NEATH, I. (1995). Remembering and forgetting as context discrimination. *Learning & Memory*, **2**, 107-132.
- COWAN, N. (1995). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.
- COWAN, N., DAY, L., SAULTS, J. S., KELLER, T. A., JOHNSON, Y., & FLORES, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory & Language*, **31**, 1-17.
- COWAN, N., LICHTY, W., & GROVE, T. R. (1990). Properties of memory for unattended spoken syllables. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 258-269.
- ESTES, W. K. (1991). On types of item coding and source of recall in short-term memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 155-174). Hillsdale, NJ: Erlbaum.
- KEPPEL, G., & UNDERWOOD, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior*, **1**, 153-161.
- MCGEOCH, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, **39**, 352-370.
- MILLER, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.
- NAIRNE, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, **18**, 251-269.
- NAIRNE, J. S. (1996). Short-term/working memory. In E. L. Bjork & R. A. Bjork (Eds.), *Handbook of perception and cognition: Vol. 10. Memory* (pp. 101-126). New York: Academic Press.
- NEATH, I., & NAIRNE, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review*, **2**, 429-441.
- OSGOOD, C. E. (1953). *Method and theory in experimental psychology*. New York: Oxford University Press.
- PETERSON, L. R., & PETERSON, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, **58**, 193-198.
- PROCTOR, R. W., & FAGNANI, C. A. (1978). Effects of distractor-stimulus modality in the Brown-Peterson distractor task. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 676-684.
- SCHWEICKERT, R., & BORUFF, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 419-425.
- WAUGH, N. C., & NORMAN, D. A. (1965). Primary memory. *Psychological Review*, **72**, 89-104.
- WHEELER, M. A. (1995). Improvement in recall over time without repeated testing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 173-184.

(Manuscript received February 28, 1997;  
revision accepted for publication April 14, 1997.)