

# Word-length effects in immediate memory: Overwriting trace decay theory

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

Memory is worse for items that take longer to pronounce, even when the items are equated for frequency, number of syllables, and number of phonemes. Current explanations of the word-length effect rely on a time-based decay process within the articulatory loop structure in working memory. Using an extension of Nairne's (1990) feature model, we demonstrate that the approximately linear relationship between span and pronunciation rate can be observed in a model that does not use the concept of decay. Moreover, the feature model also correctly predicts the effects of modality, phonological similarity, articulatory suppression, and serial position on memory for items of different lengths. We argue that word-length effects do not offer sufficient justification for including time-based decay components in theories of memory.

People can remember items that take less time to pronounce better than items that take longer to pronounce. Mackworth (1963) first reported the high correlation between reading rate and memory span, but because she was primarily interested in identifying limits in iconic memory (or the visual image, as she termed it), she did not measure word length precisely. Nonetheless, over five experiments with a variety of stimuli, including pictures, letters, digits, and colors, she found that "the amount reported was proportional to the speed of reporting the individual items" (Mackworth, 1963, p. 81). M. J. Watkins (1972) and M. J. Watkins and O. C. Watkins (1973) reported effects of word length as a function of modality and serial position in both free and serial recall. In all cases, recall was worse for words of four syllables than for words of one syllable.

Baddeley, Thomson, and Buchanan (1975) systematically explored the effect of word length in terms of pronunciation time. They demonstrated that even when items are equated for number of syllables and word frequency, if one set of words takes less time to pronounce than another set, memory will be better for the shorter items. Ellis and Hennesly (1980) showed that apparent differences in memory span between Welsh- and English-speaking subjects might be ascribed to the relatively longer time needed to say the Welsh digits than the English digits. Cowan et al. (1992) demonstrated a word-length effect even when the phonemic components of the long and short items were closely matched.

The predominant explanation of these findings involves Baddeley's (1986, 1992) concept of working memory. According to this account, short-term retention of verbal information depends on the articulatory loop. As part of this

structure, there is a passive, phonological store that is susceptible to time-based decay. A covert rehearsal process can refresh or reactivate the traces in the store to counter, temporarily, the effects of decay. If rehearsal of a particular item does not occur within a certain length of time, the memory trace for that item will have decayed too far to be usable. The amount of verbal information that can be retained is therefore a tradeoff between the decay rate (which is assumed to be fixed) and the covert rehearsal rate, which can vary. Specifically, the relationship between memory span,  $s$ , for verbal items of type  $i$ , is a linear function of pronunciation rate,  $r$ , and the duration of the verbal trace,  $\tau$ :

$$s_i = r_i \tau. \quad (1)$$

Perhaps the most compelling evidence supporting this view is the relative consistency of the presumed rate of decay,  $\tau$ . As Schweickert and Boruff (1986, p. 420) put it, "The mean of these trace duration estimates is 1.6 sec. Considering the variety of methods for presenting the stimuli and measuring the spans and pronunciation rates, one is struck more by the agreement than by the differences." The goal of the present paper is to demonstrate that the relationship expressed in Equation 1 need not arise from decay. Rather, we show that an interference-based model, Nairne's (1988, 1990) feature model, easily handles the relevant patterns of data.

## THE FEATURE MODEL

The feature model (Nairne, 1988, 1990) was designed to account for the major effects observed in immediate memory settings, including the recency effect, the modality effect, the terminal and preterminal suffix effects, and the effects of articulatory suppression, temporal grouping, and phonological similarity. Nairne distinguished between two types of features that comprise a memory trace: Modality-dependent features represent the conditions of presentation, including presentation modality, whereas modality-

---

We thank Nelson Cowan, Denny C. LeCompte, Randi C. Martin, Richard Schweickert, and Aimée M. Surprenant for comments on an earlier draft. Correspondence may be addressed to I. Neath, 1364 Psychological Sciences Building, Purdue University, West Lafayette, IN 47907-1364 (e-mail: neath@psych.purdue.edu).

independent features represent the nature of the item itself and are generated through internal processes such as categorization and identification. Modality-dependent information can interfere only with modality-dependent features, and modality-independent information can interfere only with modality-independent features. Rather than placing the locus of echoic memory in a separate structure, such as precategorical acoustic store, or PAS (Crowder & Morton, 1969), Nairne follows O. C. Watkins and M. J. Watkins (1980) in that echoic and nonechoic representations of an item are viewed as different aspects (or features) of a common memory trace.

Nairne (1990) assumed that recall would depend on the match between a degraded trace in primary memory and a particular undegraded trace in secondary memory. Thus, a major function of primary memory was to serve as a means of constructing and maintaining cues that might indicate which secondary-memory trace was present on a particular list (see also Raaijmakers & Shiffrin, 1981). Memory traces are represented by vectors of features, and each individual feature typically takes as a value +1, -1, or 0; the actual values used for each feature are generated randomly for each trial.

Primary memory traces do not exist in a vacuum; rather, they exist as part of a stream of ongoing mental activity (see Johnson & Raye, 1981; Nairne & McNabb, 1985). When interpreting a primary memory trace, the subject needs to discriminate the trace not only from other list traces but also from traces that are generated internally. By definition, traces that are generated by purely internal activity contain only modality-independent features. Externally produced list traces contain modality-dependent features because there is always a presentation modality. But it seems likely that there will be more modality-dependent features for auditory presentation than for visual presentation. The justification for this assumption rests on two lines of evidence. First, there is an extensive literature on the speech-like encoding of information presented visually and also for the speech-like nature of subvocal rehearsal (for reviews, see Baddeley, 1986; Crowder, 1976). Thus, it is reasonable to assume that the traces for internally generated items, which Nairne (1990) assumed contained only modality-independent features, are very similar to those for visually presented items. Second, there is an extensive literature that indicates that there is almost no visually based interference with sequential, single-mode presentation (e.g., Frick, 1985; Penney, 1989). Thus, for most presentation conditions, Nairne (1988, 1990) assumed that modality-dependent features play only a minor role in the identification of a visually presented item in memory. Typically, our simulations assume that there are 20 modality-independent features and either 2 (visual) or 20 (auditory) modality-dependent features (see Nairne, 1988, for further discussion on this point).

In the model, forgetting is due primarily to the overwriting of particular features. If feature  $i$  of item  $n$  is the same as feature  $i$  of item  $n+1$ , then feature  $i$  of item  $n$  is set to 0 with some probability,  $F$ . Thus, forgetting at this stage is implemented as a particular form of retroactive interfer-

ence and is not dependent on time per se. It is important to note, however, that it is not the loss of trace information in and of itself that lowers recall; rather, it is the reduction in similarity of a given degraded trace to its undegraded trace relative to the similarity of the other items to the same trace. In overwriting, no distinction is drawn between the interference produced by an externally based source (e.g., another list item) or an internally based source.

To determine which items will be sampled and then potentially recovered for recall, Nairne (1990) proposed a similarity-based choice rule (cf. Luce, 1963), the same sort of rule that has proven useful in exemplar models of categorization (e.g., Nosofsky, 1986). Formally, the probability that a particular secondary memory trace,  $SM_j$ , will be sampled as a potential recall response for a primary memory trace,  $PM_i$ , is given by

$$P_s(SM_j|PM_i) = \frac{w_{ij}s(i,j)}{\sum_k w_{ik}s(i,k)}, \quad (2)$$

where  $w_{ij}$  and  $w_{ik}$  are possible response-bias weights and  $s(i,j)$  represents the computed similarity between primary memory trace  $PM_i$  and secondary-memory trace  $SM_j$ . The parameters  $w_{ij}$  and  $w_{ik}$  are included in Equation 2 for generality; in all simulations reported here and in previous work, they were set to 1.0 and so had no effect on the model's performance. Typically,  $k$  represents and is determined by the set of items presented just on the current trial, although it can also represent larger sets, as when modeling set-size effects (Neath, 1994). Similarity is related to distance ( $d_{ij}$ ) by the following equation:

$$s(i,j) = e^{-d_{ij}}. \quad (3)$$

The distance between two trace vectors, in some psychological space, is calculated by adding the number of mismatched features,  $M$ , and dividing by the number of compared features,  $N$  (see Equation 4).  $M_k$  is simply the number of times feature position  $x_{ik}$  does not equal feature position  $x_{jk}$ ,  $a$  is a scaling parameter, and  $b_k$  is an attentional parameter that could be used to weight particular feature comparisons. For example, more attention might be given to certain modality-dependent features in a data-driven task than in a conceptually driven task (e.g., Roediger, Weldon, & Challis, 1989). In all simulations reported,  $b_k$  was set to 1.0.

$$d_{ij} = \frac{a \sum_k b_k M_k}{N}. \quad (4)$$

Output interference occurs by assuming that an item needs to be recovered prior to actual recall. The probability of recovery is related to the number of times the item has already been recalled. Let  $P_r$  be the probability of recovering a sampled item,  $c$  be a scale constant, and  $r$  be the number of times a sampled item has already been recalled on the current trial:

$$P_r = e^{-cr}. \quad (5)$$

Thus, even though a subject may have correctly sampled  $SM_j$ ,  $SM_j$  may not be recovered if it has already been re-

called, either correctly or incorrectly, earlier in the trial. This process captures the finding that subjects tend not to recall an item more than once, even if that item actually appeared more than once on the original list (e.g., Hinrichs, Mewaldt, & Redding, 1973).

A description of a typical auditory trial follows to make clear the operation of the model. First, 20 modality-independent and 20 modality-dependent features are each randomly set to values of +1 or -1; together, these 40 features represent the first item. The modality-independent features represent the results of internal processes, such as identification and categorization, that are independent of presentation modality; auditory and visual presentation of the same item are assumed to produce identical modality-independent features. The aspects of the stimulus peculiar to its mode of presentation (e.g., male or female speaker, Texan or English accent, etc.) are represented by the modality-dependent features; different presentation modalities afford different physical features that the memory system encodes. Each additional item in the list is then represented by its own randomly generated set of 40 features, and overwriting of identical features in adjacent traces is assumed to occur. For example, if modality-independent feature number 5 of item 2 has the same value as modality-independent feature number 5 of item 1, then item 1's feature 5 is overwritten by replacing the original value with a value of 0, with a certain probability,  $F$ .

At the end of list presentation, primary memory contains a trace of each of the items that was presented. In the typical case, these traces will be degraded because certain features will have been overwritten. For each trace in primary memory, the subject tries to select an appropriate recall candidate by comparing the degraded trace with intact traces in the secondary memory search set. Typically, the secondary memory search set will consist of items presented on the most recent list, but this is not necessarily always the case. For example, when the same to-be-remembered items are repeated on each list or if the items comprise a well-established group (e.g., the digits 1 through 9), the subject will use the appropriate secondary memory-search group. When, however, the to-be-remembered items come from a larger group (e.g., unique items on each trial), the search group is likely to be larger (Neath, 1994).

The probability of sampling a particular secondary memory trace as the recall response for a particular degraded primary memory trace has been described in Equation 2. Because the two weighting parameters are included only for generality, the sampling probability depends primarily on the similarity between traces. In addition, because the attentional parameter  $b_k$  is set to 1.0, similarity essentially depends upon the number of matching features between the two traces compared (see Equations 3 and 4). Thus, the secondary memory trace chosen as the response for a particular degraded primary memory trace will typically be that trace with the largest proportion of matching features, relative to the other choices currently available in the secondary memory search set. The absolute number of features is not as important as the proportion of matching features. If a subject selects as a response to a degraded primary

memory trace an incorrect secondary memory trace, the subject will typically not choose that particular secondary-memory trace again. This output interference is implemented through Equation 5.

Using the equations and procedures described above, Nairne (1990) has shown that the model accounts for many of the major effects of immediate memory. The pronounced recency effect seen in serial recall of auditory items arises because although the modality-independent features of the last list item are subject to overwriting by subsequent internal activities, the modality-dependent features are not; these undegraded features then result in a better match between the trace and the original item, relative to the other traces available. Items presented visually do not enjoy this advantage because they have fewer modality-dependent features. According to the theory, if there are no modality-dependent features, there is no recency. In addition, for a recency effect to be observed, a presentation modality must afford the encoding of *useful* physical features. If a list of auditory homophones is presented (e.g., *pear, pare, pair*), there should be no recency (Crowder, 1978), because the undegraded modality-dependent features of the final item will contribute almost no useful information for matching the degraded trace to the original item. Similarly, in the typical case, there is little or no recency effect observed with visual presentation (LeCompte, 1992) because the few modality-dependent features do not represent useful physical information. However, when visual presentation affords useful features, recency effects can be obtained, as with mouthed or lipread items.

If an additional item, a stimulus suffix, is presented after the final item, recency will be reduced if the suffix is grouped with the list items and if the modality-dependent features overwrite the features of the final item. Thus, visual suffixes should have little or no effect on auditory list items, and physically similar suffixes should have far larger effects than semantically similar suffixes (because the locus of recency lies in residual modality-dependent features). If the suffix is grouped in with the list items, then there will also be a preterminal suffix effect due to the increased search set (see Nairne, 1990). This approach has been used to explain the context-dependent suffix effect (see Neath, Surprenant, & Crowder, 1993) as well as modality and suffix effects that are observed in modalities with no apparent acoustic component, such as tactile (Nairne & McNabb, 1985; M. J. Watkins & O. C. Watkins, 1974) or mouthed or lip-read (Nairne & Crowder, 1982; Spoehr & Corin, 1978) stimuli.

The feature model also naturally accounts for the observation that phonological similarity impairs serial recall performance (Crowder, 1978). Because sampling probability is conceived as the ratio of similarities, any increase in similarity increases the value of the denominator in Equation 2 relative to the numerator, resulting in overall worse performance. Phonological similarity is simulated in the feature model by manipulating the number of overlapping features.

When subjects engage in articulatory suppression, they repeatedly say a constant item out loud, and this constant

piece of information was assumed by Nairne (1990) to be incorporated into the memory trace of each individual item. Because articulatory suppression reduces performance for both auditory and visual items, articulatory suppression is implemented by setting half of the modality-independent features of each item to a constant value. This has the net result of increasing the similarity of the items, producing decrements in the ability to match a degraded trace correctly with the appropriate undegraded trace. As Nairne (1990) demonstrated, the feature model produces the appropriate interaction between phonological similarity effects in the two modalities when tested with and without articulatory suppression.

Finally, the feature model can predict the appropriate modality-based grouping effects. When a temporal gap is inserted in a list, performance is enhanced for auditory items but there is little or no effect on visual items (Frankish, 1985; Ryan, 1969). A temporal gap preserves the modality-dependent features of the item immediately prior to the gap in the auditory case, and because of this the auditory condition can be conceived of as two (or more) smaller lists. The reduction in the search set size provides an overall increase in performance relative to the ungrouped condition, and the removal of overwriting of the modality-dependent features of each end-boundary item produces miniserial position functions (see Figures 11 and 12 of Nairne, 1990).

Thus, with a small, core set of assumptions about how information is represented (as modality-independent and modality-dependent features), forgotten (by featural overwriting due to similarity), and retrieved (by searching secondary memory for the trace most similar to a particular degraded memory trace, as compared with other traces), the feature model can simulate the main data of immediate memory. Its most notable omission to date has been its inability to account for what appears to be time-based word-length effects.

#### IMPLEMENTING THE WORD-LENGTH EFFECT IN THE FEATURE MODEL

As originally implemented, the feature model had no plausible mechanism to account for word-length effects. It might be supposed that simply manipulating the number of features, with longer words having more features, would do the trick. However, successful recall depends on the relative similarity of a given primary memory trace to a particular secondary memory trace. Increasing the number of features may affect the absolute similarity between a primary memory trace and a specific secondary memory trace, but adding more randomly generated features will not, on average, affect the similarity relative to the other items very much.

The feature model can be applied to word-length effects if one simply assumes that word-length effects operate in a manner analogous to list-length effects. When the length of a list of to-be-remembered items is increased, the number of times all items within a particular list are correctly recalled in order will systematically decrease (Ebbinghaus, 1885/1964; Schweickert & Boruff, 1986). The rea-

son, according to the feature model, is straightforward: the longer the list, the larger the denominator in Equation 2 and the lower the sampling probability for each primary memory trace. Moreover, it is also true that there are nine opportunities to make an error in a nine-item list, but only two opportunities in a two-item list. Even if the probability of making an error is identical in these two cases, as long as it is less than 1.0, correct recall of the long list is less likely than correct recall of the short list. Note that neither time nor decay is required in this explanation.

In our new simulations, we applied a similar line of reasoning to *items* rather than to lists (see also Melton, 1963). Just as lists are made up of multiple items, each of which has to be successfully recalled for correct reproduction of the list, one can conceive of words as being made up of multiple segments (Brown & Hulme, in press), each of which has to be assembled in the correct order for identification of the item. As used in this paper, the term "segment" is intended to be theoretically neutral, although the number of segments in a particular item is presumed to be linearly related to its pronunciation rate.<sup>1</sup> To initiate recall, the subject has to assemble the segments of the degraded traces in primary memory into serviceable retrieval cues. Many visual perception theorists make an analogous assumption concerning object parts and visual perception (e.g., Hildreth & Ullman, 1989; Hoffman & Richards, 1985). Although the details of this assembly process may be complex, we assume for simplicity that there is a small, but nonzero, probability that an error will occur, and that when an error does occur, the usefulness of the trace as a retrieval cue is reduced.

The number of segments in an item is assumed to be uncorrelated with the number of features. For example, the word "cat" has only one syllable, whereas the word "elephant" has three; the former should have fewer segments than the latter. However, both words have approximately the same ratings for concreteness, imagery, and meaningfulness, and both have similar word frequencies (Paivio, Yuille, & Madigan, 1968). The modality-independent features, which arise through processes such as identification and categorization, would be different in terms of the particular values of each trace element, but need not differ in the number of features required to code the information. The modality-dependent features, which encode primarily the conditions of presentation, would also be similar (assuming the same speaker or font, etc.). Thus, two items may have the same number of features but may require different numbers of segments. Consider two more items. The word "Kierkegaard" has many segments but, for the person with relatively little knowledge of philosophy, may require very few features to code the modality-independent information (cf. Muter, 1984). On the other hand, for a lexicographer, the word "set" has over two dozen different meanings and may require many more features to code all this information even though it has only one segment. When words do differ in terms of concreteness or some other dimension that affects an item's memorability, the feature model captures this by extending the range of values that each feature element can take rather than by al-

tering the number of features (Neath, 1994). Thus, for the simulations that follow, we assumed a fixed number of features regardless of the number of segments.

In all of our simulations, the probability of a segment assembly error was held constant at 0.10. The error probability was applied to each segment in a trace; as a result, long words had a greater probability of being assembled incorrectly because they possessed more segments. If a segment error occurred at any point in the assembly process, half of the modality-independent features of the trace were set to 0 and the remaining half were left unchanged. Although some segments may undoubtedly be more critical to the identification of a particular item than others (e.g., Brown & Hulme, in press), we chose the simplifying assumption that any segmental error would result in an equal loss of information. Regardless of the number of segmental assembly errors, it seems likely that some information will remain, even if that information is only disassembled segments; thus, we again chose a simplifying assumption that no more than half of the modality-independent features could be affected. Because word-length effects obtain in both the visual and auditory modalities, the assembly process was assumed to apply only to modality-independent features. As a result, when modeling word-length effects, the probability of an individual segment error was the same for auditory and visual items, as well as for long and short words. Long words simply had more segments than short words.<sup>2</sup>

### Demonstration 1

Baddeley et al. (1975, Experiment 5) showed subjects 5-item lists of words, presented visually, for immediate serial recall. The words were matched for number of syllables, word frequency, and number of phonemes (given Scottish pronunciation). The top portion of Table 1 lists the mean number of words recalled and the pronunciation rate (in items per second) for both short and long items observed in that experiment. Subjects recalled, on average, approximately 3.59 short items and 2.67 long items, and the measured pronunciation rate was 2.15 and 1.65 items per second, respectively. Note that this produces an almost identical estimate of  $\tau$ , the presumed duration of the verbal memory trace.

**Table 1**  
Average Number of Items Correctly Recalled ( $s$ ) as a Function of Pronunciation Rate ( $r$ , in Items per Second) and Verbal Trace Duration ( $\tau$ , in Seconds) in Demonstration 1

Items	$s$	$r$	$\tau$
Data			
Short	3.59	2.15	1.67
Long	2.67	1.65	1.62
Feature Model			
Short	3.56	2.15	1.65
Long	2.67	1.65	1.62

Note—Data are from Baddeley, Thomson, and Buchanan (1975). For the feature model, the pronunciation rate,  $r$ , was calculated using the equation  $r = 2.19 - 0.04n_s$ , where  $n_s$  is the number of segments. The duration of the verbal trace was estimated using the calculated pronunciation rate and the number of items recalled in the simulation.

Demonstration 1 has two purposes: First, we demonstrate that as the number of segments increases, performance decreases linearly—the basic word-length effect. Second, we suggest how the number of segments can be related to articulation rate and thus to Equation 1. The parameter values that we used to simulate these data with the feature model are listed in the appendix. They were selected on the basis of the parameter values used in Nairne (1990) to simulate the major characteristics of immediate memory, including primacy, recency, and modality effects. Note that the first three parameters listed were set to 1.0, and that they therefore play no role in producing the phenomena of interest, either in the original simulations or in the ones reported here. Visually presented items were again represented by trace vectors that contained 20 modality-independent features and 2 modality-dependent features. Each element in each vector was randomly set to  $\pm 1$  at the beginning of each trial. The only substantive change from previous work with the feature model was setting the parameter  $a$ , a scaling parameter, to 11 to increase the overall level of performance. Thirteen simulations, each of 1,000 trials, were run, with the number of segments contained within each list item systematically increasing from 1 (short words) to 13 (long words). There was a probability of 0.10 of a segmental error, regardless of the number of segments or the presentation modality. When such an error occurred, half of the modality-independent features were set to 0.

Using these parameters, when the number of segments in an item is presumed to be 1, average recall of the items within a list was approximately the same as recall of the short words in the Baddeley et al. (1975) study (3.56 out of 5 items for the model versus 3.59 in the study); when the number of segments was increased to 13, average recall was the same as recall of the long words (2.67 for both; see Table 1). Moreover, as Figure 1 shows, as the number of presumed segments in a word increases, there is an approximately linear relationship between word length and number of items recalled. The best-fitting straight line for the simulated data plotted in Figure 1 ( $y = 3.52 - 0.072x$ ) accounted for approximately 96% of the variance.<sup>3</sup> Just as successful recall of an entire list depends on recalling each of the items, successful identification of a word can be assumed to depend on correctly assembling each of the segments. Note that the model is producing the basic word-length effect without recourse to any kind of decay.

As the number of segments increases, there will be more mismatching features between a degraded primary memory trace and the correct undegraded trace in secondary memory. For the short items, an average of 11.31 features were mismatched ( $M$  in Equation 4); for the long items, the value was 14.60. At the same time, there was an average of 17.94 features mismatched between the degraded primary memory trace of a short item and the incorrect undegraded secondary memory traces; there were 19.17 mismatches for the long items. So, short items match both the correct and incorrect secondary memory traces better than do long items. The locus of the word-length effect in the feature model is not the absolute number of mismatches, but rather the mismatch ratio, the ratio between

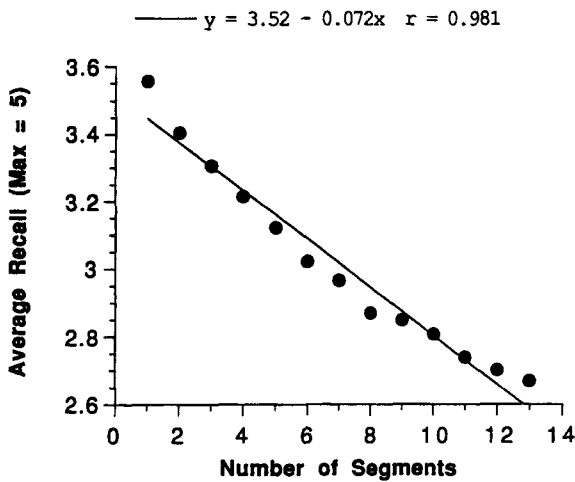


Figure 1. Simulation results from the feature model demonstrating that the linear relationship between word length and number of items recalled can be produced without recourse to decay: as the number of segments per item increases, the average recall decreases.

number of mismatches with the correct secondary memory trace and the number of mismatches with the incorrect secondary memory traces. For the short items, this ratio is 0.63; for the long items, it is 0.76. Generally, the smaller the mismatch ratio, the larger the probability of correctly matching a degraded primary memory trace with the correct undegraded secondary memory trace.

Having demonstrated that, with the additional assumption of segments, the feature model could account for the main aspects of the word-length effect, we next sought to determine a relationship between the number of segments and articulation rate. Once this relationship is expressed as an equation, we can solve for the variables in Equation 1. Let  $r$  be the articulation rate in items per second and  $n_s$  be the number of segments. Baddeley et al. reported pronunciation rates of 2.15 items per second for the short words and 1.65 items per second for the long words. Using these data, we established that the best-fitting linear relationship for all 13 segment sizes was  $r = 2.19 - 0.04n_s$ . Using this equation, when there is 1 segment, the articulation rate is calculated to be 2.15; when there are 13 segments, the articulation rate is 1.65. Using these calculated articulation rates, we can use the mean number of items recalled correctly, as determined by the simulation, to solve for the estimated duration of the trace. These values are 1.65 and 1.62 for short and long items, respectively, and they correspond with the estimates calculated by Baddeley et al. (see Table 1).

Demonstration 1 resulted in two notable findings. First, and most importantly, an interference model can produce the main aspects of the word-length effect: As Figure 1 shows, as the length of the to-be-remembered items increases, recall will linearly decrease. The only additional assumption required by the feature model is that longer items have more segments than shorter items—no appeal to any process of decay is required. The second result of interest is that we can mimic Equation 1 by first independently establishing

the relationship between the number of segments and the articulation rate, and then by using this calculated pronunciation rate to arrive at estimates of the duration of the verbal trace that are in accord with published estimates.

Having shown that the model can simulate the basic word-length effect, we next sought to test predictions of the feature model by determining how the word-length manipulation interacted with already implemented aspects of the model. A strong test of the model, then, would be to observe the predicted interactions between word length, as shown in Demonstration 1, with phonological similarity effects in Demonstration 2, articulatory suppression in Demonstration 3, articulatory suppression and phonological similarity effects in Demonstration 4, modality in Demonstration 5, and serial position in Demonstration 6.

**Demonstration 2**

Schweickert, Guentert, and Hersberger (1990) measured span for a set of phonologically similar and dissimilar consonants. The mean number of items correctly recalled was 5.62 for the similar set and 7.06 for the dissimilar set. An important finding was that there was essentially no difference in measured pronunciation rate between the similar (3.01 per second) and dissimilar (2.92 per second) items. The relevant data are displayed in the top part of Table 2. Schweickert et al. expressed their data in terms of ratios in order to highlight the relationships among span, articulation rate, and trace duration for short and long words and for phonologically similar and dissimilar items. The ratio of the number of items recalled in the similar and dissimilar conditions was 0.796, whereas the ratio of the measured pronunciation rates was approximately 1. In terms of Equation 1, this means that the ratio of the estimates of the duration of the verbal trace was 0.773. In the Baddeley et al. (1975) data, where items did not differ phonologically but did differ in length, the ratio of the number of items recalled in the long and short conditions was 0.744 and the ratio of the measured pronunciation rates was approximately 0.767. This means that the ratio of the estimates of the duration of the verbal trace was approximately 1.0. On the basis of these relationships, Schweickert et al. (1990) concluded

**Table 2**  
Measured Span ( $s$ ) as a Function of Pronunciation Rate ( $r$ , in Items per Second) and Verbal Trace Duration ( $\tau$ , in Seconds) in Demonstration 2

Condition	$s$	$r$	$\tau$
Data			
Similar	5.62	3.01	1.87
Dissimilar	7.06	2.92	2.42
Ratio	0.796	1.031	0.773
Feature Model			
Similar	2.76	2.15	1.28
Dissimilar	3.56	2.15	1.65
Ratio	0.775	1.000	0.775

Note—The data, from Schweickert, Guentert, and Hersberger (1991), show measured articulation rate. The simulation results, from the feature model, were calculated on the basis of the number of segments in exactly the same way as in Demonstration 1. The dissimilar condition for the feature model is the short condition from the previous simulation.

that although word length affects  $r$  in Equation 1, phonological similarity affects  $\tau$ . The same conclusion was reached independently by Hulme and Tordoff (1989).

The main purpose of Demonstration 2 was to show that the effects of the number of segments are independent from the effects of phonological similarity. Nairne (1990) modeled the effects of phonological similarity in the feature model by manipulating the number of overlapping features contained in the primary memory traces. For example, in a control or phonologically dissimilar condition, each of the 20 modality-independent features is randomly set to be  $\pm 1$ . Under these conditions, there will be some featural overlap, but it will be essentially random. In a phonologically similar condition, however, 13 of the 20 features are set to the same value, +1 for example, and the remaining 7 are randomly set to  $\pm 1$ . There will be at least 13 overlapping features, but some of the other 7 may also overlap.

The feature model predicts the complementary relationship reported by Schweickert et al. (1990) and Hulme and Tordoff (1989), because word length and phonological similarity are modeled in two different ways. Word length is modeled by increasing the number of segments, which is correlated with pronunciation rate, but the number of segments is unrelated to the number or type of features (see above). Phonological similarity, on the other hand, is modeled by increasing featural overlap, but featural overlap is unrelated to the number of segments. Because featural overlap is independent of the number of segments, when items are made more similar by increasing the featural overlap but keeping the number of segments constant, the calculated pronunciation rate will not be altered. Thus, the addition of the assumptions concerning segments is independent from the way in which the model accounts for phonological similarity.

Demonstration 1 showed that increasing the number of segments reduces memory performance in an approximately linear fashion. For the current demonstration, the short items from Demonstration 1 were considered phonologically dissimilar, because there was no preset minimum overlap of features, and these were used to represent the consonants that Schweickert et al. used as stimuli. Only one parameter was changed to represent the phonologically similar condition: the minimum number of identical modality-independent features was increased from 0 to 13. Thus, the model will still be generating predictions for a five-item visual list.<sup>4</sup>

When the minimum number of similar modality-independent features was set to 13, performance on the short items from Demonstration 1 decreased from 3.56 to 2.76. Increasing phonological similarity decreases performance, and here resulted in a ratio of 0.775, approximately the same as that observed by Schweickert et al. (1990). Using the relationship between the number of segments and articulation rate developed in Demonstration 1 (e.g.,  $r = 2.19 - 0.04n_s$ ), the calculated pronunciation rate will be identical, 2.15 items per second, because only one segment is used per item. Using this calculated rate, the estimate of the verbal trace duration is 1.28 for similar items and 1.65 for dissimilar items, a ratio of 0.775, again approximately the same as that observed by Schweickert et al.

Again, the model is capturing the important aspects of the relationship described by Equation 1: increasing the length of an item affects  $r$  but not  $\tau$ , and increasing the similarity of the items affects  $\tau$  but not  $r$ , replicating the data observed by both Baddeley et al. (1975) and Schweickert et al. (1990). Because there are only two factors in the Schweickert et al. model, the default assumption was that phonological similarity affected the duration of the trace decay, although there could be some other causal factor. As implemented in the feature model, phonological similarity does not affect the number of segments—a similar conclusion to that made by Schweickert et al.—but neither does it affect trace decay. Rather, phonological similarity affects the degree of nonrandom overlap of features; in this way, the model predicts the complementary effects of increasing pronunciation time and increasing phonological similarity. In Demonstration 3, we show that the feature model also correctly predicts the effects of articulatory suppression on the word-length effect.

### Demonstration 3

Baddeley et al. (1975, Experiment 8) presented five-word lists for immediate serial recall. There were two word lengths (one and five syllables), two presentation modalities (auditory and visual), and two articulatory conditions (suppression and no suppression). According to the working-memory view, articulatory suppression prevents subjects from registering visually presented items in the phonological store because the articulatory control process is not available. Although auditory items have automatic access to the phonological store, articulatory suppression prevents rehearsal of these items, because the articulatory control process is not available. Baddeley et al. observed appropriate effects of word length for the no-suppression groups, but articulatory suppression eliminated the word-length effect *only* for the visual group: the effect of word length remained in the auditory modality with the articulatory suppression group. Because articulatory suppression should remove the word-length effect regardless of presentation modality, the finding that the word-length effect remained for auditory items is problematical for the working-memory view.

Baddeley (1986) argued that subjects might be rehearsing during recall, refreshing the decaying trace in the phonological store while simultaneously recalling items. This is more likely for auditory items, because they are automatically registered in the phonological store and do not require conversion using the articulatory control process during presentation. If this is the case, then requiring articulatory suppression not only during presentation but also throughout recall should eliminate the word-length effect for auditory items. Baddeley, Lewis, and Vallar (1984, Experiment 4) presented five-item lists of auditory items for immediate serial recall. The short words had one syllable and the long words had five; the pool of words was drawn from Baddeley et al. (1975, Experiment 6). Subjects engaged in articulatory suppression not only during presentation but also throughout the recall period. With this change, the word-length effect was removed for audi-

tory presentation (see also Baddeley & Lewis, 1984, Experiment 2; Baddeley et al., 1984, Experiment 5). To the extent that subvocal rehearsal is prevented during both presentation and recall, there will be no word-length effect, just as working memory predicts (Baddeley, 1992).

When subjects engage in articulatory suppression, they repeatedly say a constant item out loud, and this constant piece of information was assumed by Nairne (1990) to be incorporated into the memory trace of each individual item. Because articulatory suppression reduces performance for both auditory and visual items, articulatory suppression is implemented by setting half of the modality-independent features of each item to a constant value of 1. The model assumes that no additional rehearsal occurs after list presentation, a situation comparable to when articulatory suppression is required during both input and recall. We kept the parameter values identical to the short and long settings used in Demonstration 1, but ran these simulations with articulatory suppression selected for both auditory and visual five-item lists, or with no suppression. The results are displayed in Figure 2: articulatory suppression eliminated the word-length effect for both auditory and visually presented words.

For both modalities, when a segment assembly error occurs, half of the modality-independent features are set to 0 and there is more opportunity for a segment assembly error for long items. A segment assembly error will increase the number of mismatching features between the degraded primary memory trace and the correct undegraded secondary memory trace because there are no vector elements in secondary memory that have a value of 0. Overall performance is better for auditory items than for visual items primarily because there are more useful modality-dependent features with auditory presentation. The main locus of this advantage is in the recency portion of list recall (see Nairne, 1990). When articulatory suppression occurs, half of the modality-independent features of every item are set to a constant value of 1. Overall performance will decline, relative to that of the no-suppression group, because the degraded primary memory traces are rendered more similar to the incorrect secondary memory traces. For example, with no articulatory suppression, there was an average of 17.94 features mismatched between the degraded primary memory trace of a short item and the incorrect undegraded secondary memory traces; there were 19.17 mismatches for the long items. With articulatory suppression, the number of mismatched features decreases to 14.60 for both short and long items. Because so many features have already been modified by articulatory suppression, the effects of word length will be masked.

#### Demonstration 4

Articulatory suppression during both input and recall removes the word-length effect for both visual and auditory items, but articulatory suppression does not remove the phonological similarity effect for auditory items, even when articulatory suppression is continued throughout recall. Longoni, Richardson, and Aiello (1993, Experiment 4) presented subjects with six-item lists of either phonolog-

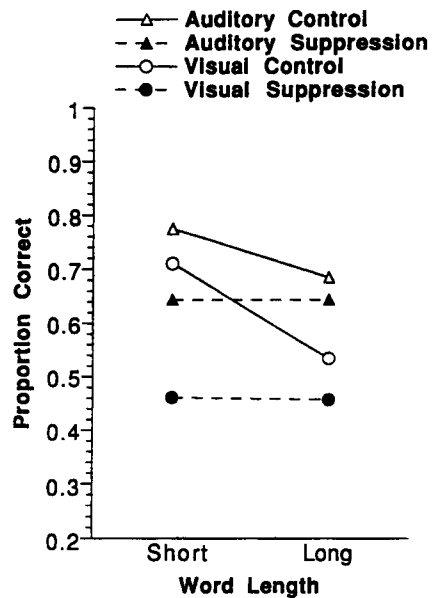


Figure 2. Simulation results demonstrating that articulatory suppression removes the word-length effect for both auditory and visual items.

ically similar (e.g., *can, cap, cat*) or dissimilar (e.g., *bun, day, few*) items for immediate serial recall. In addition, half of the time subjects listened to the words silently and half of the time they engaged in articulatory suppression during both presentation and recall, whispering the word “the” repeatedly as fast as possible. The results are shown in the top segment of Table 3. Although articulatory suppression affects the overall level of performance, the phonological similarity effect remains.

For this simulation, the minimum guaranteed number of overlapping modality-independent and -dependent features was set to 10 for the phonologically similar items and to 0 for the phonologically dissimilar items. The words were assumed to have 1 segment, there were six words per list, and auditory presentation was simulated. The simulation produced the appropriate pattern of results, as can be seen in the bottom segment of Table 3. We ran another series of simulations, this time changing the number of segments from 1 to 13 to reflect long words; all other parameters remained the same. As can be seen, the phonological similarity effect remains for both short and long words with and without articulatory suppression, but the word-length effect is eliminated by articulatory suppression. This demonstration produces the appropriate interactions between phonological similarity, word length, and articulatory suppression (Baddeley et al., 1984, Experiments 2, 3, 4, and 5; Longoni et al., 1993, Experiments 1 and 4).

Table 4 shows the average number of mismatching features between a degraded primary memory trace and the appropriate undegraded secondary memory trace (labeled “Correct”) and between a degraded primary memory trace and the other, incorrect undegraded secondary memory traces (labeled “Incorrect”). Generally, the smaller the mismatch ratio, the more likely a particular primary mem-



**Table 3**  
**Mean Recall of Phonologically Dissimilar and Similar Items Presented With and Without Articulatory Suppression**

Items	Silent		Articulatory Suppression	
	Dissimilar	Similar	Dissimilar	Similar
		Data		
Short	70.0	51.1	56.9	37.8
		Model		
Short	72.0	48.2	58.1	37.5
Long	62.5	41.7	57.4	37.0

Note—The top row shows mean percentage of items correctly recalled in an immediate serial recall test of six-item auditory lists of either phonologically similar or dissimilar items presented with or without articulatory suppression (from Longoni, Richardson, & Aiello, 1993, Experiment 4, fast presentation condition). These items were all one-syllable words such as *can*, *cap*, *cat* or *bun*, *day*, *few*. The bottom rows show the mean percentage of items correctly recalled in a simulation using the feature model which also simulated results for short (1 segment) or long (13 segments) items, showing that although articulatory suppression removes the word-length effect, it does not remove the phonological similarity effect.

ory trace will be matched with the correct secondary memory trace. Looking first at the silent-dissimilar condition, the word-length manipulation increases the number of mismatches between corresponding primary and secondary memory traces, reducing the likelihood of a correct match. In a comparison of the silent-dissimilar condition with the silent-similar condition, the phonological similarity manipulation produces a larger mismatch ratio by increasing the number of mismatching features between a primary memory trace and the correct secondary memory trace. This increase occurs because more overwriting will occur if all the items in the list are phonologically similar: overwriting changes an element's value to 0, and the secondary memory traces do not contain any zeros.

Articulatory suppression is modeled by adding a constant value of 1 to a random half of the modality-independent features, making the degraded primary memory traces more similar to each other. This will make the correct discrimination more difficult because there is less of a difference in the number of mismatching features between a primary memory trace and correct and incorrect secondary memory traces.

### Demonstration 5

The results of the simulation shown in Figure 2 suggest that the word-length effect may be larger for visually presented items than for auditory items. M. J. Watkins and O. C. Watkins (1973) presented eight-item lists for immediate serial recall. Lists were either auditory or visual, and the items within the lists had either one or four syllables. There was a larger effect of word length for visual items than for auditory items. Baddeley et al. (1975, Experiment 8) also found a reliable modality  $\times$  word length interaction.

The data from M. J. Watkins and O. C. Watkins (1973) are presented in Table 5, as are simulation results. The only parameter changed was  $\alpha$ , the scaling parameter, to lower overall level of performance. The simulations used eight-

item lists, as did Watkins and Watkins, and modeled serial recall of long and short words using both auditory and visual presentation conditions. The feature model correctly predicts smaller effects for auditory presentation than for visual presentation as a function of serial position. The reason is straightforward: the word-length effect is implemented in the feature model as affecting only modality-independent features. Because auditory presentation results in more modality-dependent features than does visual presentation, the effects of a word-length manipulation will affect a smaller proportion of elements in a vector that represents an auditory item, other things being equal.

### Demonstration 6

The above demonstrations illustrate the ability of the feature model to account for the basic effects of word length. Because the model was designed originally to account for serial position effects, it is also possible to test the feature model against findings that show that word length interacts with serial position. For example, Cowan et al. (1992, Experiment 1) found that, for visual presentation of five-item lists, the magnitude of the word-length effect increased across serial positions. The feature model predicts this result nicely. We ran simulations with a list length of five for short and long items presented visually, using the same parameter values as in Demonstration 1. The results are shown in Table 6.

The effect of word length is larger at the final position than at the first position, and this is due neither to ceiling effects nor to floor effects. The only difference in the pattern of results between the simulation data and those observed by Cowan et al. (1992, Experiment 1) is that the feature model predicts a difference at the first position whereas there was no difference at the first position in the Cowan et al. data. As noted in the original paper, this was probably due to ceiling effects. Baddeley et al. (1975, Experiment 3) found the opposite pattern, that the effects of

**Table 4**  
**Average Number of Mismatching Features in Demonstration 4**

Type of PM-SM Mismatch	Silent		Articulatory Suppression	
	Dissimilar	Similar	Dissimilar	Similar
	Short Words			
Correct	18.85	25.38	18.33	21.28
Incorrect	31.21	32.98	28.00	26.86
Mismatch ratio	0.60	0.77	0.65	0.79
	Long Words			
Correct	22.02	27.76	18.35	21.21
Incorrect	32.48	33.93	28.01	26.82
Mismatch ratio	0.68	0.82	0.66	0.79

Note—The average number of mismatching features between a degraded primary-memory trace and the correct undegraded secondary-memory trace ("Correct") and between a degraded primary-memory trace and the incorrect undegraded secondary-memory traces ("Incorrect") as a function of word length, phonological similarity, and presence or absence of articulatory suppression. These values come from the simulations that produced the results shown in Table 3: Articulatory suppression during input and recall eliminates the word-length effect but does not eliminate the phonological similarity effect. Generally, the larger the mismatch ratio, the worse performance will be.

**Table 5**  
**Proportion Correct as a Function of Presentation Modality and Word Length**

Word Length	Data		Feature Model	
	Auditory	Visual	Auditory	Visual
Short	0.368	0.300	0.363	0.336
Long	0.328	0.237	0.301	0.250
Difference	0.040	0.063	0.062	0.085

Note—Average proportion of items correctly recalled under immediate serial recall as a function of modality of presentation and word length for eight-item lists of short (one syllable) or long (four syllables) words. Data are from Watkins and Watkins (1973); the values were extracted from the original figures using the computer program by Huyser and van der Laan (1994). The feature model simulated the same procedure, immediate serial recall of eight-item lists of short or long words presented either aloud or visually. The row labeled "Difference" shows the magnitude of the word-length effect, which is larger for visual items than for auditory items.

word length were apparent only in the early part of the list. There are many differences between the two studies, however, with Cowan et al. using visual presentation and stimuli that were matched for both the number of syllables and number of phonemes and Baddeley et al. using auditory presentation and stimuli that were matched only for the number of syllables.

The feature model predicts this interaction between serial position and word length because output interference builds up over serial positions. If a secondary memory trace is mistakenly sampled as the match for a degraded primary memory trace early in the recall process, that particular secondary memory trace is less likely to be successfully recovered even if, later on in the recall process, it is correctly sampled as the match for another degraded primary memory (see Equation 5). Because there are fewer intact features, on average, for long items than for short items, the long items are more likely to be mismatched early in the recall process, and thus less likely to be successfully recovered when they are the correct match. In the current simulations and in all previous work, the number of recovery attempts was set to two. The larger this value, the more constant the effect of word length over serial positions.

## EVALUATION

### Empirical Discrepancies

There are certain aspects of the word-length effect that the current version of the feature model cannot account for, although some of these findings rely on testing methods that have not yet been implemented in the model. For example, Cowan et al. (1992, Experiments 2 and 3) presented subjects with lists that contained both long and short words. They found that lists that contained the short items in the first half and long items in the second half were recalled slightly better than lists in which the long items came first. Performance was better when the short items were recalled first, regardless of recall direction; that is, a list being recalled backward would be recalled better if the short words occurred in the second half of the list.

The feature model can simulate these situations only for forward serial recall, but it predicts that performance for the short items should be better than for the long items regardless of where in the list they occur, at least for the parameters used in the demonstrations reported here. For example, when simulating a six-item mixed list (three short items and three long items) with forward serial recall, the feature model predicts no reliable differences for the first four positions, but does predict an advantage for the last item in the long-short condition relative to the last item in the short-long condition. Cowan et al. (1992) observed differences in both list halves and that at the final position the ordering of conditions was the reverse. It is possible that the addition of overwriting during recall could remedy this incorrect prediction, although we have not explored this possibility in detail.

A second finding that the feature model cannot currently explain is the difference, when testing for backward serial recall, between results when the items are presented for immediate recall and when they are presented in a modified form of the continual distractor paradigm. Cowan, Wood, and Borne (1994) used both presentation procedures. In the immediate presentation condition, each word was presented for 1 sec, with a 1-sec interitem presentation interval. In the continual distractor condition, subjects engaged in 15 sec of digit shadowing after every item, including the final item, but also for 15 sec prior to presentation of the first item. Subjects' recall responses were spoken, and presentation of the to-be-remembered items was also spoken. The subjects were again required to recall the items in strict backward order. With the immediate presentation procedure, Cowan et al. replicated their earlier findings of an advantage for lists with a second half comprising short, rather than long, items but found no such effect with the modified continual distractor procedure.

A third finding that the feature model cannot predict is the results reported by LaPointe and Engle (1990, Experiment 5). When a fixed set of items is used, the word-length effect for visually presented items is eliminated when subjects engage in articulatory suppression and are tested in a span procedure using immediate serial recall. However, when unique items are used on every trial—an essentially unlimited set size—the word-length effect is *not* eliminated by articulatory suppression. The feature model predicts no word-length effect regardless of set size when subjects en-

**Table 6**  
**Word-Length Effect as a Function of Serial Position**

Items	Serial Position				
	1	2	3	4	5
Short	0.776	0.742	0.692	0.669	0.677
Long	0.612	0.584	0.511	0.489	0.473
Difference	0.164	0.158	0.181	0.18	0.204

Note—Simulation results of the feature model showing the proportion of short and long items correctly recalled as a function of serial position for a five-item list of visually presented items. The row labeled "Difference" shows that the magnitude of the word-length effect is larger at the end of a list than at the beginning.

gage in articulatory suppression. Although the feature model does predict fewer mismatches between the degraded primary memory trace and the correct secondary memory trace for short items than for long items with a large set size, this did not result in a consistent recall advantage for the short items. The reason for this is that the mismatch ratios were approximately the same for short and long items. However, it is worth noting that the working-memory view also cannot easily predict these data; attributing the word-length effect for a large set size under articulatory suppression to contributions from long-term memory may explain the results but does not predict the pattern.

### Hidden Decay Assumptions

A different criticism might be that we are sneaking in decay “through the back door” by being vague about what exactly constitutes a segment. Because our fundamental assumption is that there is a correlation between the number of segments and pronunciation rate, this may give segments the appearance of being time based units. Even if these segments are time based, this does not logically imply decay. For example, even though items that take longer to pronounce have more segments, the loss of segmental information does not automatically happen over time. Segmental loss is probabilistic, and therefore it is possible to have no loss regardless of the time interval. Time-based decay in the phonological store of working memory logically entails that in the absence of such other processes as elaborative rehearsal, forgetting always happens and performance will *always* be worse over time.

To reiterate our main view, we assume that recalling an item involves a process similar to that used to recall a list. When tested by serial recall, successful recall of a list depends on recalling each individual item in order. As the number of items in the list increases, the probability of correct recall will decrease. The reason for this is that items are represented in memory by traces that comprise features. Successive features can overwrite features of earlier items, an instantiation of retroactive interference. To explain word-length effects, we assume that a similar explanation holds for items (Melton, 1963).

We assume that long items are “longer” in that they have more segments than short items. For simplicity, we assumed a fixed probability of making an error per segment. When a segmental assembly error occurs, information about particular features is lost. Loss of featural information, whether from a segmental assembly error or from featural overwriting by successive items, generally reduces the similarity of the degraded trace to its undegraded trace in secondary memory. Even under these circumstances, it is still possible for perfect recall if the remaining features are still sufficiently informative. For example, Neath (1994) has shown that increasing the range of values that individual elements can take—from  $\pm 1$  used here to  $-2$  through  $+2$ —results in better performance. This manipulation was used to model differences between abstract (a range of two values) and concrete (a range of four values) items. If we were including a hidden time-based decay process, this

particular aspect of the feature model’s performance would not be possible.

There are few successful simulation models of trace-decay theory in general or of the word-length effect in particular (although see Brown & Hulme, in press, for an exception). Trace-decay theories, of which working memory is the most dominant, remain primarily verbal theories that do not make precise predictions and that apply only to situations lasting a few seconds, so-called short-term memory. Not only can the feature model produce appropriate results for much of the word-length literature, it can also simulate many other effects, including primacy and recency effects, suffix effects, modality effects, effects of articulatory suppression and phonological similarity, and temporal grouping effects. Moreover, many of these factors interact in important ways, and the feature model can account for these also. Thus, the feature model not only makes specific, unambiguous predictions but also applies to a far wider variety of situations.

Baddeley’s (1986) working-memory view was originally designed to account for four phenomena: (1) the word-length effect, (2) the phonological similarity effect, (3) effects due to articulatory suppression, and (4) the unattended speech effect, although its ability to handle the latter has been recently called into question (e.g., Jones, Macken, & Murray, 1993; Martin, 1993). Both the feature model and the working-memory view can predict the effects of phonological similarity and articulatory suppression (Baddeley, 1986; Nairne, 1990), and we have shown that the feature model can also predict the main aspects of the word-length effect. Although we cannot definitively rule out a time-based decay mechanism, we have shown that such a mechanism is not necessary. Even if the working memory view of word-length effects is accurate, it is still incomplete and must appeal either to a vague “central executive” or to “contributions from long-term memory” to account for many of the results that the feature model can simulate. Working memory may offer an explanation of why short words tend to be better remembered than long words, but it does so in a less precise way than the feature model, and it does not address how these effects are related to other immediate memory phenomena.

### SUMMARY AND CONCLUSIONS

McGeoch (1932) detailed many empirical and philosophical problems of attributing forgetting to time-based decay. His analysis was so influential that almost no theories of long-term memory currently include decay. Even Thorndike (1914), who was “credited” by McGeoch (1932, p. 354) with a time-based decay view of forgetting, did not actually maintain that forgetting over the long term was due to decay (despite many citations to the contrary). Thorndike’s (1914) implied “law of disuse” referred to changes in habit strength rather than to the time-based fading of memories.

When it comes to short-term memory, however, virtually all theories employ decay. For example, the central mechanism for forgetting in Baddeley’s (1986) working memory view, at least for verbal information, is a phonological-

based store whose contents are susceptible to time-based decay. One of the principal relationships that seemed to require such a decay-based explanation was the relationship between pronunciation time and memory span.

Using a specific form of retroactive interference, the feature model (Nairne, 1990) can account for serial position and modality effects, can explain both the terminal and preterminal suffix effects, and can also predict the appropriate effects of grouping and modality-like effects seen with mouthed and lip-read stimuli. In addition, it correctly simulates the effects of phonological similarity and articulatory suppression. We have shown here that the feature model easily handles the finding that word length influences pronunciation time, but that phonological similarity does not (Demonstration 2). Using the same parameter settings, it also predicts that the word-length effect will be abolished under articulatory suppression, as long as all rehearsal is prevented (Demonstration 3). It can also account for the finding that the phonological similarity effect remains for both long and short words under articulatory suppression, but that the word-length effect is abolished (Demonstration 4). Finally, we have demonstrated that it correctly predicts that word-length effects are larger for visual than for auditory items (Demonstration 5) and that word-length effects will be larger at the end of a list than at the beginning (Demonstration 6). The word-length effect, then, does not require a time-based decay model.

#### REFERENCES

- BADDELEY, A. D. (1986). *Working memory*. New York: Oxford University Press.
- BADDELEY, A. D. (1992, January 31). Working memory. *Science*, **255**, 556-559.
- BADDELEY, A. D., & ANDRADE, J. (1994). Reversing the word-length effect: A comment on Caplan, Rochon, and Waters. *Quarterly Journal of Experimental Psychology*, **47A**, 1047-1054.
- BADDELEY, A. D., & LEWIS, V. J. (1984). When does rapid presentation enhance digit span? *Bulletin of the Psychonomic Society*, **22**, 403-405.
- BADDELEY, A. D., LEWIS, V. J., & VALLAR, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, **36**, 233-252.
- 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.
- BROWN, G. D. A., & HULME, C. (in press). Modeling item length effects in memory span: No rehearsal needed? *Journal of Memory & Language*.
- CAPLAN, D., ROCHON, E., & WATERS, G. S. (1992). Articulatory and phonological determinants of word-length effects in span tasks. *Quarterly Journal of Experimental Psychology*, **45A**, 177-192.
- CAPLAN, D., & WATERS, G. S. (1994). Articulatory length and phonological similarity in span tasks: A reply to Baddeley and Andrade. *Quarterly Journal of Experimental Psychology*, **47A**, 1055-1062.
- 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 & Language*, **31**, 1-17.
- COWAN, N., WOOD, N. L., & BORNE, D. N. (1994). Reconfirmation of the short-term storage concept. *Psychological Science*, **5**, 103-106.
- CROWDER, R. G. (1970). The role of one's own voice in immediate memory. *Cognitive Psychology*, **1**, 157-178.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- CROWDER, R. G. (1978). Memory for phonologically uniform lists. *Journal of Verbal Learning & Verbal Behavior*, **17**, 73-89.
- CROWDER, R. G., & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, **5**, 365-373.
- EBBINGHAUS, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger, Trans.). New York: Dover. (Original work published 1885).
- ELLIS, N. C., & HENNELLY, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculations in Welsh and English. *British Journal of Psychology*, **71**, 43-52.
- FRANKISH, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory & Language*, **24**, 200-209.
- FRICK, R. W. (1985). Testing visual short-term memory: Simultaneous versus sequential presentations. *Memory & Cognition*, **13**, 346-356.
- HILDRETH, E. C., & ULLMAN, S. (1989). The computational study of vision. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 581-630). Cambridge, MA: MIT Press.
- HINRICHS, J. V., MEWALDT, S. P., & REDDING, J. (1973). The Ranschberg effect: Repetition and guessing factors in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **12**, 64-75.
- HOFFMAN, D. D., & RICHARDS, W. A. (1985). Parts of recognition. In S. Pinker (Ed.), *Visual cognition* (pp. 65-96). Cambridge, MA: MIT Press.
- HULME, C., & TORDOFF, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology*, **47**, 72-87.
- HUYSER, K., & VAN DER LAAN, J. (1994). *DataThief 2.0* [Computer program]. Amsterdam: Computer Systems Group of the Nuclear Physics Section of the National Institute for Nuclear Physics and High Energy Physics (NIKHEF-K).
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, **88**, 67-85.
- JONES, D. M., MACKEN, W. J., & MURRAY, A. C. (1993). Disruption of visual short-term memory by changing-state auditory stimuli: The role of segmentation. *Memory & Cognition*, **21**, 318-328.
- LAPOINTE, L. B., & ENGLE, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1118-1133.
- LECOMPTE, D. C. (1992). In search of a strong visual recency effect. *Memory & Cognition*, **20**, 563-572.
- LONGONI, A. M., RICHARDSON, J. T. E., & AIELLO, A. (1993). Articulatory rehearsal and phonological similarity in working memory. *Memory & Cognition*, **21**, 11-22.
- LUCE, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103-189). New York: Wiley.
- MACKWORTH, J. F. (1963). The duration of the visual image. *Canadian Journal of Psychology*, **17**, 62-81.
- MARTIN, R. C. (1993). Short-term memory and sentence processing: Evidence from neuropsychology. *Memory & Cognition*, **21**, 176-183.
- MCGEOCH, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, **39**, 352-370.
- MELTON, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning & Verbal Behavior*, **2**, 1-21.
- MUTER, P. (1984). Recognition and recall of words with a single meaning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 198-202.
- NAIRNE, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition*, **16**, 343-352.
- NAIRNE, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, **18**, 251-269.
- NAIRNE, J. S., & CROWDER, R. G. (1982). On the locus of the stimulus suffix effect. *Memory & Cognition*, **10**, 350-357.
- NAIRNE, J. S., & MCNABB, W. K. (1985). More modality effects in the absence of sound. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 596-604.
- NEATH, I. (1994, May). *Temporal encoding and the auditory superiority hypothesis*. Paper presented at the meeting of the Midwestern Psychological Association, Chicago.
- NEATH, I., SURPRENANT, A. M., & CROWDER, R. G. (1993). The context dependent stimulus suffix effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 698-703.
- NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.
- PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, im-

- agery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monographs*, **76** (1, Pt. 2), 1-25.
- PENNEY, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, **17**, 398-422.
- RAAIJMAKERS, J. G. W., & SHIFFRIN, R. M. (1981). Search of associative memory. *Psychological Review*, **95**, 93-134.
- ROEDIGER, H. L., III, WELDON, M. S., & CHALLIS, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.
- RYAN, J. (1969). Grouping and short-term memory: Different means and patterns of groups. *Quarterly Journal of Experimental Psychology*, **21**, 137-147.
- SCHWEICKERT, R., & BORUFF, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 419-425.
- SCHWEICKERT, R., GUENTERT, L., & HERSBERGER, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science*, **1**, 74-77.
- SPOEHR, K. T., & CORIN, W. J. (1978). The stimulus suffix effect as a memory coding phenomenon. *Memory & Cognition*, **6**, 583-589.
- THORNDIKE, E. L. (1914). *The psychology of learning*. New York: Columbia University, Teachers College Press.
- WATKINS, M. J. (1972). Locus of the modality effect in free recall. *Journal of Verbal Learning & Verbal Behavior*, **11**, 644-648.
- WATKINS, M. J., & WATKINS, O. C. (1973). The postcategorical status of the modality effect in serial recall. *Journal of Experimental Psychology*, **99**, 226-230.
- WATKINS, M. J., & WATKINS, O. C. (1974). A tactile suffix effect. *Memory & Cognition*, **2**, 176-180.
- WATKINS, O. C., & WATKINS, M. J. (1980). The modality effect and echoic persistence. *Journal of Experimental Psychology: General*, **109**, 251-278.

## NOTES

1. Caplan, Rochon, and Waters (1992; see also Caplan & Waters, 1994) have suggested that the word-length effect does not result from differences in articulatory duration but rather from differences in the complexity of planning the phonological form of a word for speech production. They point out that there would be a correlation between phonological planning and pronunciation rate because planning is a necessary prerequisite for articulation. Although a controversial conclusion (see, e.g., Baddeley & Andrade, 1994), this interpretation would be consistent with the feature model, as the concept of a segment could easily be mapped onto phonological complexity.

2. Because an entirely new implementation of the feature model was used throughout this paper, we checked its performance against the results described in Figures 3, 4, 5, 6, 7, 9, 10, and 11 of Nairne (1990). The only substantive difference concerned simulating the effects of acoustically identical items such as *pair*, *pare*, *pear*, etc. The original version produced slight recency effects (see Figure 7 of Nairne, 1990) even when all the modality-dependent features were set equal. The current version did not show any recency, resulting in figures more closely approximating the empirical data.

3. Although Figure 1 shows a highly linear relationship between the number of segments and correct recall, there does appear to be a non-linear trend. At some point, when the number of segments becomes too large, there will be no additional loss in performance, and so the model does predict a quadratic function for extreme values of length. When we use the term "predict" in conjunction with the feature model, we mean that the assumptions in the model logically entail a particular outcome.

4. Schweickert et al. (1990) presented items visually and had subjects read them aloud, a presentation procedure that results in essentially the same effects as pure auditory presentation (Crowder, 1970). We chose to continue using visual items so that only one parameter would be changed between Demonstrations 1 and 2. Demonstration 4 examines phonological similarity effects with auditory items.

## APPENDIX

## Parameter Settings for All Reported Simulations

With the exception of the scaling parameter  $a$ , the parameters listed below were the same as those used in the majority of simulations reported by Nairne (1990). Note that the first three parameters were set to 1.0 and therefore play no role in predicting the phenomena of interest either in the original simulations or in those reported here.

Parameter	Demo 1	Demo 2	Demo 3	Demo 4	Demo 5	Demo 6
Attention constant [ $b$ ]	1.0	1.0	1.0	1.0	1.0	1.0
Response bias weight $w_{ij}$	1.0	1.0	1.0	1.0	1.0	1.0
Response bias weight $w_{jk}$	1.0	1.0	1.0	1.0	1.0	1.0
Recovery scaling constant [ $c$ ]	2.0	2.0	2.0	2.0	2.0	2.0
Number of recovery attempts	2	2	2	2	2	2
Number of simulations	1,000	1,000	1,000	1,000	1,000	1,000
Probability of overwriting [ $F$ ]	1.0	1.0	1.0	1.0	1.0	1.0
Distance scaling constant [ $a$ ]	11	11	11	11	6	11
Number of modality-independent features	20	20	20	20	20	20
Number of modality-dependent features:						
Auditory	NA	NA	20	20	20	NA
Visual	2	2	2	NA	2	2
Minimum Similarity:						
Modality Independent	0	13	10	0	0	0
Modality Dependent	0	0	10	0	0	0
List Length	5	5	5	6	8	5
Number of Segments:						
Short	1	1	1	1	1	1
Long	13	13	13	13	13	13
Probability of Segment Error	0.10	0.10	0.10	0.10	0.10	0.10

Note—NA, not applicable.