A Note on the Two-Second Decay Conjecture in Verbal Working Memory

Shane T. Mueller

Indiana University Dept of Psychology, 1101 East 10th Street, Bloomington, IN 47409, USA. Phone: (812) 855-0626. Fax:(812) 856-5217

Abstract

Based on the experimental observation that span-length lists take about two seconds to articulate, experimenters have conjectured that the verbal short-term memory trace lasts about two seconds. However, this two-second decay conjecture is inconsistent with other experimental data, and a mathematical model that offered support for this conjecture actually estimates the lifetime of an intact list of words, rather than that of individual words from a list. To illustrate this point, I present an item-based model of trace decay which demonstrates that the duration of the short-term memory trace must be longer than the time taken to articulate a span-length list. For a set of words whose span-length lists can be rapidly articulated in about two seconds, the model offers a conservative estimate for their mean decay times of about four seconds.

Key words: short-term memory decay

In their influential paper on verbal short-term memory, Baddeley, Thomson, and Buchanan (1975) reported the articulatory duration effect in verbal short-term memory, in which lists comprised of shorter words are recalled more accurately than equal-length lists of longer words. They noted that the number of words that could be remembered (out of five) was equal to the number that could be articulated in about 1.5 seconds (across experiments and participants, their estimates ranged between 0.93 and 1.95 s). This result was later bolstered by the results of Standing, Bond, Smith, and Isely (1980), who showed that across a wide variety of stimuli and experimental conditions, the time it took to subvocalize memory-span length lists was between about 1.7 and 2.1 seconds. These behavioral findings suggest that information decays from short-term memory with time, and provided strong evidence for the articulatory loop

⋆ Sponsored in part by United States Office of Naval Research grant N00014-92-J-1173 to the University of Michigan

Email address: stmuelle@indiana.edu (Shane T. Mueller).

model of verbal working memory. However, the findings do not provide a direct estimate for the duration the verbal memory trace. Indeed, it was not until later that Baddeley and Lewis (1984) conjectured that the duration of the verbal short-term memory trace was roughly equal to the duration of a span-length list: a matter of about two seconds. A mathematical model of trace decay was soon proposed by Schweickert and Boruff (1986) which formulated this conjecture explicitly. When their model was fit to the data from a variety of stimuli, the duration of the verbal trace was consistently found to be around two seconds, validating the conjecture that information in verbal working memory disappears within about two seconds.

Since then, the two-second decay conjecture has entered the lore of experimental psychology, appearing in introductory textbooks and research articles alike. However, the conjecture is questionable for number of reasons. In this note, I will point out some of the problems with the conjecture, and show how it does not actually describe the lifetime of individual words in short-term memory, but rather the lifetime of an intact list of words. I will present a simple model that illustrates this, and demonstrate its ability to fit empirical data, producing a new estimate for the duration of the memory trace in verbal working memory.

1 Trace Decay

Trace decay is a process by which information in a memory trace is lost through the passage of time alone, independent of limitations in encoding, interference with other activity, capacity limits, or retrieval difficulty. The assumption that information decays from memory at all is itself controversial, and experiments designed to determine whether information does in fact decay have produced differing results (cf. Reitman, 1971; 1974, Shiffrin, 1973). Critics of trace decay theory have often supposed that other processes can account for some or all of the effects that appear to support time-based decay (e.g., Neath & Nairne, 1995) and that the concept of decay is unnecessary. I will not try to settle this debate here, but instead attempt to answer the conditional question, “If verbal information is lost from short-term memory because of decay alone, what are its properties?”. Consequently, I will leave open for now the questions of whether information does in fact decay, and the extent to which other factors limit memory performance as well. But even if we ignore the effects on memory of these other factors, there are a number of reasons to believe that the verbal memory trace lasts longer than just two seconds. These include the face validity of the two-second decay conjecture, experimental evidence to the contrary, and the misinterpretation of the model that has provided the strongest evidence for the two-second decay conjecture.

1.1 Face Validity

Although the two-second decay conjecture has become widely accepted, the claim may appear improbable on its face. Introspection and personal experience seem to contradict the conclusion that the contents of short-term memory last only about two seconds. If this were true, it should be very difficult to parse the meaning of sentences of any considerable length; we should have great difficulty maintaining focus on a task for any extended period of time; and we might find ourselves constantly rehearsing everything that needs to be remembered more than a few seconds. It might be argued that these tasks are all supported by long-term memory as well, but that would imply that short-term memory is good for little else but remembering phone numbers and artificially-constructed stimuli in memory experiments.

Despite these objections to the conjecture, it may be unwise to trust intuition in this matter. Our intuitions about many psychological phenomena have been found to be false, and even though we might hope to have better insight into short-term memory phenomena (cf. Ericsson & Simon, 1980), stronger support should be sought.

1.2 Experimental Evidence

In addition to contradicting introspection, the two-second decay conjecture is also inconsistent with published experimental findings. For example, in an extensive review of literature concerning auditory memory, Cowan (1984) found evidence that auditory memory may persist five or more seconds. One example of the literature Cowan reviewed was Watkins and Todres (1980), who suggested that verbal information may even last up to 20 seconds. Subsequent research that measured neuromagnetic responses to tones (Sams, Hari, Rif, and Knuutila, 1993) showed that acoustic information lasts at least 10 seconds. Although some of this research has focused on auditory information rather than verbal information per se, it suggests that verbal information may be more persistent than the two-second conjecture supposes.

Other evidence suggesting that verbal information lasts longer than two seconds comes from experiments reporting the recall times for immediate serial recall. For example, Corballis (1969) showed that participants normally take between three and six seconds to recall lists of eight items. Similarly, Dosher and Ma (1998) and Dosher (1999) have noted that output delays during recall often extend to 4-6 seconds for a span-length list. For the two-second decay conjecture to explain these results, it must be supposed that participants were engaging in iterative rehearsal, even during recall, and that they were able to perform this rehearsal at a much faster rate than recall itself. Previous research has shown that subvocal speech does not happen considerably faster than overt speech (Landauer, 1962), and so this explanation is hard to accept. Furthermore, recall is only modestly impacted by articulatory suppression (cf. Baddeley, Buchanan, & Thomson, 1975), when it should be made nearly impossible when rehearsal is precluded, if the two-second conjecture were true.

Despite the trace decay conjecture's apparent lack of face validity and the its contradiction by experimental evidence, a quantitative model of Schweickert and Boruff (1986) embodying the trace decay conjecture has been quite successful at accounting for empirical data. Yet as I will show in the next section, some of the model's assumptions and conclusions have been misunderstood, adding further doubt to the veracity of the two-second decay conjecture.

1.3 Schweickert and Boruff's List-based Model

Initial empirical observations of the relationship between speech rate and memory accuracy (Baddeley, et al., 1975; Standing et al., 1980) noted that there appeared to be a linear relationship between the number of correctly-recalled words and speech rate, embodied by the equation $s = kr + c$. Here, memory span (s) is a function of speech rate (r), k is the fitted slope of the function, and (c) is the fitted intercept. The slope k has been interpreted as the duration of the verbal memory trace, and c as the contribution of other memory systems to performance. When fit to empirical data, the slope k is normally around 2.0, and the intercept c is normally (but not always) close to 0, suggesting that the verbal trace decays within about two seconds, and that decay is the primary manner in which information is lost from short-term memory.

Based on this observed relationship, Schweickert and Boruff (1986) proposed simple model of trace decay that accounted for 95% of the variance in the psychophysical functions relating list length to recall accuracy in the immediate serial recall task. This model made a number of simplifying assumptions: a list of items takes T_r seconds to rehearse or recall; the verbal trace of a list lasts for T_v seconds; and if T_r is less than T_v, the list should be recalled correctly. This should hold even if the two times T_r and T_v are random variables, based on which Schweickert and Boruff derived psychophysical functions predicting the probability of correct recall for lists of different lengths. These functions were able to fit empirical data quite well for a number of different stimuli, producing estimates of decay times that were slightly less than two seconds.

Schweickert and Boruff's (1986) model is both elegant in its simplicity and parsimonious in its ability to account for relationships between speaking times and list recall accuracies. Additionally, it has been used as a tool to investigate contributions to memory span by other factors, such as word frequency (e.g., Hulme, Roodenrys, Schweickert, Brown, Martin, & Stuart, 1997). However, it is also often misunderstood, because it does not describe the decay distribution of individual words in verbal short-term memory, but rather the functional decay distribution of an entire list of words, taken as a single entity. In the model, accuracy is a function of list duration and trace duration alone. Consequently, the list-length effect arises because lists with more words take longer to recall, but not simply because they contain more words. Schweickert and Boruff's model is able to estimate a decay time for a list of words, but the estimate it produces may be less interesting than another quantity: the decay time for individual words. Empirical evidence has indicated that these two quantities are not the same: individual words may be recalled correctly, even if the entire list is not (e.g., Drewnowski & Murdock, 1980). But in order for a list to be recalled correctly, each of the individual words must be recalled. Thus, two seconds may be a reasonable estimate for the average lifetime of the first word lost from memory, while the distribution of these lifetimes may be on much longer than two seconds. In the next section, I will describe a mathematical model that embodies this idea, in order to estimate the lifetime of individual items in verbal short-term memory.

2 An Item-based Trace Decay Model

For this new model, I will make several assumptions that, although unrealistic, simplify calculations and enable the decay distributions of individual items to be estimated. I will ignore, for instance, the possible effects of factors such as capacity limitations, interference, encoding and retrieval failure, and guessing. Although many (and perhaps all) of these other factors may influence immediate serial recall accuracy, I will not incorporate them so as to offer a purer extension of Schweickert and Boruff's (1986) model. Additionally, although Schweickert and Boruff used both the mean and variance of list rehearsal times, and Schweickert et al. (1996) estimated quadratic functions relating pronunciation times to list length, I will use only the slope of the best-fit line between list-length and total pronunciation time as an index of articulatory duration. However, I will consider two distinct performance conditions: iterative rehearsal and articulatory suppression. Furthermore, I will allow more general assumptions to be made about the underlying decay distributions.

2.1 Item Decay with Rehearsal Suppressed

To model the decay of individual items, I will initially assume that rehearsal is eliminated through articulatory suppression. Additionally, I will assume that decay for each item is independent and identically distributed according to an arbitrary distribution with density $f(t)$, where t indicates the time between complete item presentation and some point in the future. The distribution function of $f(t)$ is denoted as $F(t) = \int_0^t f(t)dt$, which indicates the probability of an item having decayed after t seconds, and its survival function is denoted as $S(t) = 1 - F(t)$, which indicates the probability that an item remains intact after t seconds. Let I_p indicate the presentation time between items, I_d indicate the delay between presentation and recall (often equal to I_p), I_r indicate the recall time between items (assuming that all items take the same amount of time to utter), and N indicate the length, in words, of the list.

Based on these assumptions, the probability of recalling an item depends on the time elapsed since its presentation was completed. For an item in serial position n, this time can be computed to be:

$$I_{T_n} = I_p \cdot (N - n) + I_d + (n - 1) \cdot I_r \tag{1}$$

which can be refactored as:

$$I_{T_n} = (I_d - I_r) + N \cdot (I_p) + n \cdot (I_r - I_p), \tag{2}$$

where the first part is a constant for all positions of all lists using the same word set, the second part depends on the number of words in the list, and the third part depends on a word's serial position in the list.

Although the relationship between memory accuracy and pronunciation time was initially shown as a proportion of items correctly recalled out of five (cf. Baddeley et al., 1975), it is typically understood as a relation between the pronunciation time and memory span (the length at which probability of recalling the entire list correctly is .5). For an entire list to be recalled correctly, each of the individual items must be recalled correctly as well. Consequently, an N-item list should be recalled correctly with probability:

$$\prod_{n=1}^{N} S(I_{T_n}). \tag{3}$$

The values $S(I_{T_n})$ may not map directly onto the serial position function for a set of words, because an error earlier in the list may reduce the accuracy of later positions. In fact, Schweickert et al. (1999) found empirical support for the assumption that the product of conditional probabilities of consecutive items accounts well for the probability of correct list recall. To conform to this finding, $S(I_{T_n})$ might be considered the conditional probability of correct recall, so that item n will be recalled correctly with probability $S(I_{T_n})$ if the previous item was recalled correctly.

Equation 3 can be combined with Equation 2 to produce a direct formula for item-based recall accuracy when participants are engaging in articulatory suppression:

$$P(I_p, I_r, N) = \prod_{n=1}^{n} S((I_d - I_r) + N \cdot (I_p) + n \cdot (I_r - I_p)). \tag{4}$$

Although this equation is useful, many experiments do not require participants to engage in articulatory suppression, and so do not discourage iterative rehearsal Equation 4 may not be appropriate for these cases. Consequently, in the next section I will compute a similar formula for immediate serial recall accuracy when articulatory rehearsal is being used.

2.2 Articulatory Rehearsal

The processes involved in iterative articulatory rehearsal are quite complex (e.g., Kieras, Meyer, Mueller, & Seymour, 1999) and cannot be modeled easily with a simple mathematical formula. However, some progress may be made if it is assumed that when a rehearsal strategy is adopted, few errors are made during iterative rehearsal occurring while the list is being presented and that the time between consecutive rehearsals of an item is approximately equal to the duration of the list. This time can be estimated to be roughly equal to $I_r * N$ for each list item. Consequently, for this model, the inter-stimulus and post-stimulus intervals do not matter when rehearsal is allowed. This prediction is most likely wrong, although Baddeley and Lewis (1984) showed that presentation rate has little effect on accuracy when articulatory rehearsal is allowed. Given these assumptions, the probability of recalling item n correctly is $S(I_r * N)$, and that the probability of recalling an entire list correctly is

$$\prod_{i=1}^{N} S(I_r \cdot N) = (S(I_r \cdot N))^N. \tag{5}$$

When the presentation, post-stimulus delay, and recall times are equal ($I_p = I_d = I_r$), Equations 4 and 5 are identical, which makes intuitive sense: in that case, item presentation essentially replaces rehearsal.

2.3 Examining the Item-based Decay Model

For both the suppression and rehearsal models, the probability of recalling an entire list correctly is the product of the probability of recalling each individual item correctly. As with Schweickert and Boruff (1986), the probability of correct recall depends on the articulatory duration of the words in the list. However, in Equations 5 and 4, the decay distribution describes the probability of a single item in memory existing over time, rather than the entire list. Thus, the estimate of the decay distribution for individual items may be different than the decay distribution estimated for the list as a whole.

Using the formulas developed so far it is simple to show that individual items in memory-span length lists (i.e., those at 50% accuracy) must be recalled with considerably greater accuracy than 50%. For example, under suppression, I_{T_n} will have N different values on which $S(t)$ is evaluated, and under rehearsal the values are identical. For both conditions, when a span-length list is being recalled, the product of $S(t)$ evaluated at these N points must equal 0.5, and so a rough estimate of the probability of recalling any single item in a span-length list is $.5^{1/N}$. For lists whose memory span is 2, 3, 4, 5, 6, and 7, these values are .71, .79, .84, .87, .89, and .91, respectively. Thus, depending on the length of the span list, the survival function must be considerably greater than .5 when recall accuracy of the complete list is at .5.

Interestingly, this demonstration is not sufficient to conclude that the mean of the decay distribution for individual items is greater than the list rehearsal time: any random variable whose survival function jumps discontinuously from above about 91% to below 50% will produce lists with memory spans between 2 and 7 whose recall times are equal to the mean decay times. The simplest such distribution has a "step" survival function at the maximum value of I_{T_n}. Any list that can be recalled in less time than I_{T_n} will be recalled correctly, and any list that takes more time will not be recalled correctly. This can be easily interpreted in terms of a "tape loop" that records the last few seconds reliably but overwrites everything before it, a metaphor that motivated Baddeley's original phonological loop model. Thus, in terms of the item-based decay model proposed here, the only way the statement "Memory span is equivalent to the number of items that can be repeated rapidly in about two seconds" implies that the decay time of items in working memory is about two seconds is if decay happens like a tape-loop step function, in which all items are always around for a short fixed period of time (e.g., two seconds), after which they always disappear completely.

2.4 Decay Distributions

Although the previous section described a "tape-loop" decay function, it is unrealistic to suppose that memory decay could work as mechanically as a tape loop suggests. Yet little consensus exists in the choice of decay distributions for modeling short-term memory. Many models do not even specify a decay distribution, but instead propose an underlying process through a decay distribution can be derived (e.g. Brown & Hulme, 1995; Anderson, Bothell, Lebiere, & Matessa, 1998). Of those models that assume specific statistical distributions for information decay, the most common is probably the exponential distribution. (cf. Cowan, 1984; Byrne, 1996). The exponential distribution has the property that it is history-less, so that the time an item has existed in the past has no effect on its expected future lifetime. For such a distribution, the conditional decay distribution given that an item has survived to time t is the same as if it were encoded at time t. This suggests that if words decayed from short-term memory according to an exponential distribution, rehearsal would not improve performance. Apparently, the exponential distribution is not a realistic candidate to consider as a decay distribution because if items truly did decay from short-term memory according to it, rehearsal would not improve recall.

Other distributions also have the property that rehearsal would not improve performance, and for some of these rehearsal would actually hurt performance. Consider the survival function $S(t)$ of the hypothetical distribution $f(t)$ shown as a thick solid curve Figure 1. To determine whether rehearsal improves the chance of recalling an item correctly, the survival function of a newly-rehearsed word (the original distribution) must be compared to the survival function if the word is not replaced by rehearsal (the distribution conditionalized on the item having survived to the point of rehearsal). For example, suppose that a list of items took approximately one second to rehearse, and the next item to rehearse has survived this long. If the item is rehearsed at this point, its new lifetime will conform to the original distribution $S(t)$, but if it is not rehearsed, it will conform to the conditional distribution whose survival function $S(t+1|t>1)$ (shown as a dashed line in Figure 1). The survival function for the rehearsed item lies above the survival function for the unrehearsed item for the next two seconds, and so it is clearly advantageous to rehearse the item. However, suppose that a list of items takes two seconds to rehearse. In this case, the survival function of the rehearsed item (again, the original distribution $S(t)$) lies below the survival function of the unrehearsed item (shown as a thin dotted line in Figure 1) at two seconds, when the next rehearsal is expected. In this case, rehearsing the word could actually harm performance, because the original word would be more durable than the newly rehearsed item. Because it is known that rehearsal does improve performance in many short-term memory tasks, it seems reasonable to eliminate such distributions from consideration prima facie.

There is at least one other desirable property that a decay distribution should have: its density should be entirely above zero, because information cannot decay before it is presented. But even distributions without this property have been used with success: Schweickert and Boruff (1984) assumed that list decay times are normally distributed. Although in this model, some small proportion

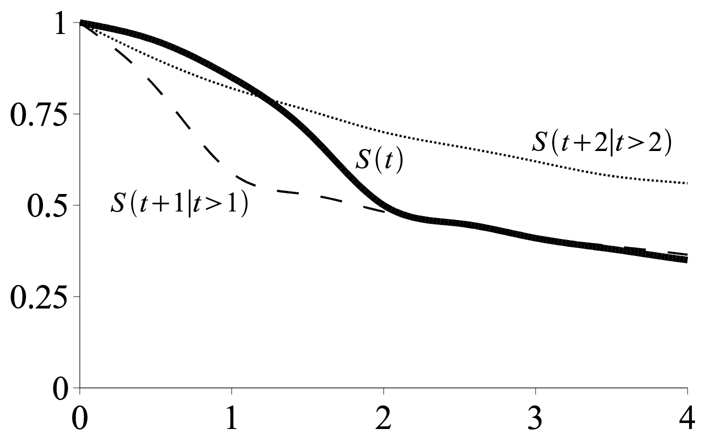

Fig. 1. Survival functions for a hypothetical decay distribution and its conditional distributions after 1 and 2 seconds.

of lists were decaying before they were even presented, the assumption proved to be quite successful in practice, because the mean and variance of the decay distribution were such that the proportion of lists decaying before they were presented were vanishingly small.

The choice of decay distribution may influence the estimates one obtains about the mean lifetime for items in short-term memory. For example, for a set of items whose memory span is seven, if items decay according to an exponential distribution, it must have a mean 5.7 times as large as the list rehearsal time: for a list that takes 2 seconds to articulate, the exponential decay distribution for which $S(2.0) = .5^{1/7} = .91$ has a mean of 11.4 seconds. Another simple decay distribution is the uniform distribution with density $f(t) = 1/T$ for $t < T$; $f(t) = 0$ elsewhere, which has mean $T/2$. For this distribution, the mean decay time needs to be 5.3 times as large as the duration of the list: 10.6 seconds. Many other interesting distributions have multiple parameters and so the proportion depends on the specific combination of parameter values, but as a general rule of thumb, the more a decay distribution deviates from a tape-loop step function, the longer the mean decay time needs to be relative to the list recall duration.

For the item-based decay model to make predictions about empirical data, a decay distribution must be specified. The two-parameter version of the log-normal distribution (Hastings & Peacock, 1975) offers an attractive option. Its density is:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \, \mathrm{e}^{\frac{-(\log(t/M))^2}{2\sigma^2}} \qquad \sigma > 0; M, t \geq 0; \tag{6}$$

and its survival function can be expressed as a function of the standard normal distribution function Φ, which can be approximated numerically by functions found in most scientific computing packages.

$$S(t) = 1 - \Phi(\log(t/M)/\sigma), \tag{7}$$

This parameterization of the log-normal distribution is useful for representing decay times because it has no density below zero, and it has two parameters that are easily interpretable (M, the median and σ the spread or shape of the distribution). Combinations of these parameters offer a considerable amount of flexibility in the shape and scale of the distribution: when σ is close to 0, the distribution approaches a step-like "tape loop" distribution, and when it grows large, the survival function drops rapidly to near .5, after which it decays slowly with a long tail. Furthermore, it is relatively simple to compute random numbers with log-normal distributions, which is useful for Monte Carlo simulations of decay. The mean of the log-normal distribution is $M \cdot \exp\{\sigma^2/2\}$. Other one-tailed distributions could have been chosen as well, although it may be nearly impossible to distinguish between various distributions with empirical data.

3 Modeling empirical data

The utility of the list-based trace decay model of Schweickert and Boruff (1986) has been shown by its ability to fit to numerous data sets (e.g., six experiments in Schweickert & Boruff, 1986; five experiments with two participants in Schweickert et al., 1996). To demonstrate that the item-based decay model is able to make similar predictions about the psychophysical functions produced by immediate serial recall, I fit the model to these same 16 data sets. In those experiments, visual stimuli were presented simultaneously. Presumably, participants read the list to themselves, at about the same rate that they were later rehearsed and recalled. To estimate the average time required to rehearse each word, I estimated a single slope parameter of the relationship between list length and total recall time for each of the 16 data sets. Then, based on the Equation 4, and using a log-normal decay distribution with two free parameters, I fit psychophysical functions relating list recall accuracy and list length for each experimental condition, using numerical techniques that minimized root-mean-squared-error for each condition. The item-based decay model fit the data extremely well, with an overall R^2 value of .998. This is not entirely surprising, because the original psychophysical functions were very smooth and new parameters were fit for each condition. The more interesting result is found by examining the decay distributions required to fit these data. For each of the 16 data sets, the survival functions of the decay distributions

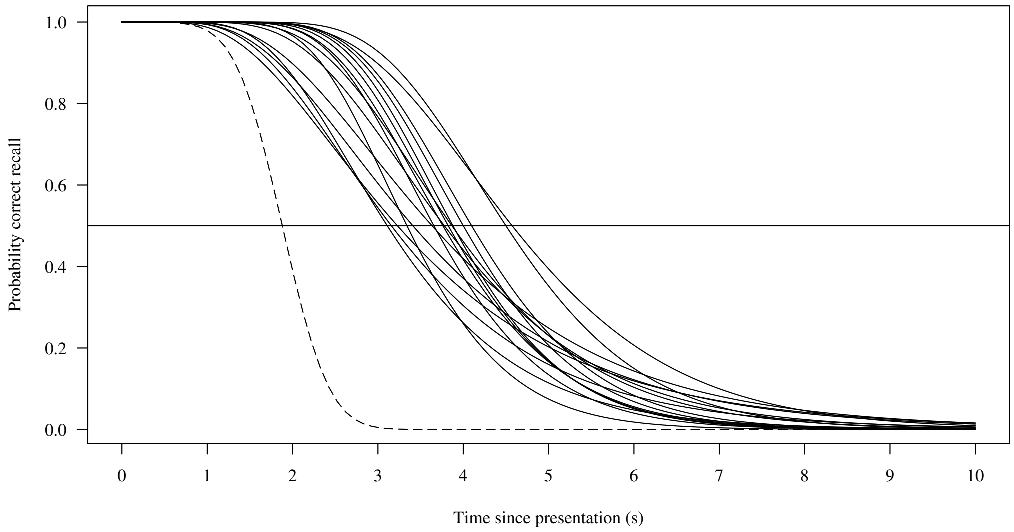

Fig. 2. Survival functions of the best-fit log-normal decay distributions for the six conditions reported by Schweickert and Boruff (1986), and five conditions with two participants reported in Schweickert et al. (1996), found using the item-based decay model. For comparison, the dashed curve shows the survival function of a normal distribution with a mean of 1.88 s and variance of 0.187 s^2, the best-fitting values obtained by Scheickert and Boruff (1986).

are shown in Figure 2.

Figure 2 shows the estimated survival functions that best fit the psychophysical function in each of the 16 data sets. Overall, the decay distributions had median parameters M which ranged between 3.1 and 4.5, with an average of 3.7 and spread parameters σ which ranged between .27 and .52 with an average of .35. Across the 15 data sets, the fitted values of the two parameters had a correlation of -.58, indicating that some trade-offs existed in the parameter values. Overall, the means of these distributions were between 3.4 and 4.8 seconds, with an average of 4.0 seconds, even though the span-length lists could be articulated in less than two seconds. In Schweickert and Boruff's list-based decay model, list decay times were 1.88 s on average, with a variance of 0.187 s^2, less than half of the mean decay times produced in the item-based decay model. For comparison, the survival function of Schweickert and Boruff's (1986) distribution is shown as a dashed curve in Figure 2.

Although the fitted values of the item-based model are longer than two-seconds, they may still be conservative estimates of the duration of items in verbal working memory, because the model assumes mechanically perfect encoding and recall processes, and it attributes all memory loss to decay. To the extent that other factors affect recall and rehearsal is less precise than prescribed by this model, the actual decay distributions would likely be longer. In support of this, when more realistic assumptions about rehearsal were made in a model by Kieras et al (1999), they determined that memory span was best accounted for by a decay distribution with a mean between 5 and 8 seconds.

4 The relationship between span and speech rate

As mentioned earlier, the two-second decay conjecture had its genesis in the observation of a linear relationship between memory span and speech rate. The item-based decay model makes predictions about this relationship as well, examples of which are shown in Figure 3. Here, the relationship predicted by Equation 4 between memory span and recall rate is plotted as a function of variations in the two different parameters of the log-normal decay distribution. For fixed values of the log-normal distribution's parameters σ and M, the relationship is close to linear, although there is some curvature at low numbers of words per second. As the spread (σ) parameter grows small, the distribution approaches a sharp cutoff (i.e., a tape-loop distribution), and the intercept approaches 0. In contrast, as the shape of the distribution spreads out more, the intercept obtained by the best-fit line increases. Although the predicted functions are not linear, data generated from such a function would probably be difficult to distinguish from a linear function. Furthermore, a linear function that was fit to such data could easily produce an intercept greater than 0.

5 Comparison to Schweickert and Boruff's Model

The model presented here and the list-based model of Schweickert and Boruff (1986) are fairly similar, and make similar predictions for list recall accuracies. They are each also quite simplistic approximations to reality, so it may be difficult to argue that one model is better than the other based on goodness-of-fit metrics alone. Neither model attempts to accurately capture all aspects of the immediate serial recall task, and so deviations from empirical data are expected. Yet they do differ in several ways, and empirical testing may be able to show that some of these differences do in fact matter.

One way in which the models differ is in how they incorporate articulatory duration. Schweickert and Boruff's model attempts to use individual recall times for each list length, and accommodates the variability of these times as well. In contrast, the item-based model presented here uses a single value for the mean articulation time of a word from a list. Certainly, the weight of evidence favors Schweickert and Boruff's approach: list pronunciation times are often curvilinear functions of list length (cf. Sternberg, Monsell, Knoll, & Wright, 1978; Mueller, et al., 2003), and the variability of these times may

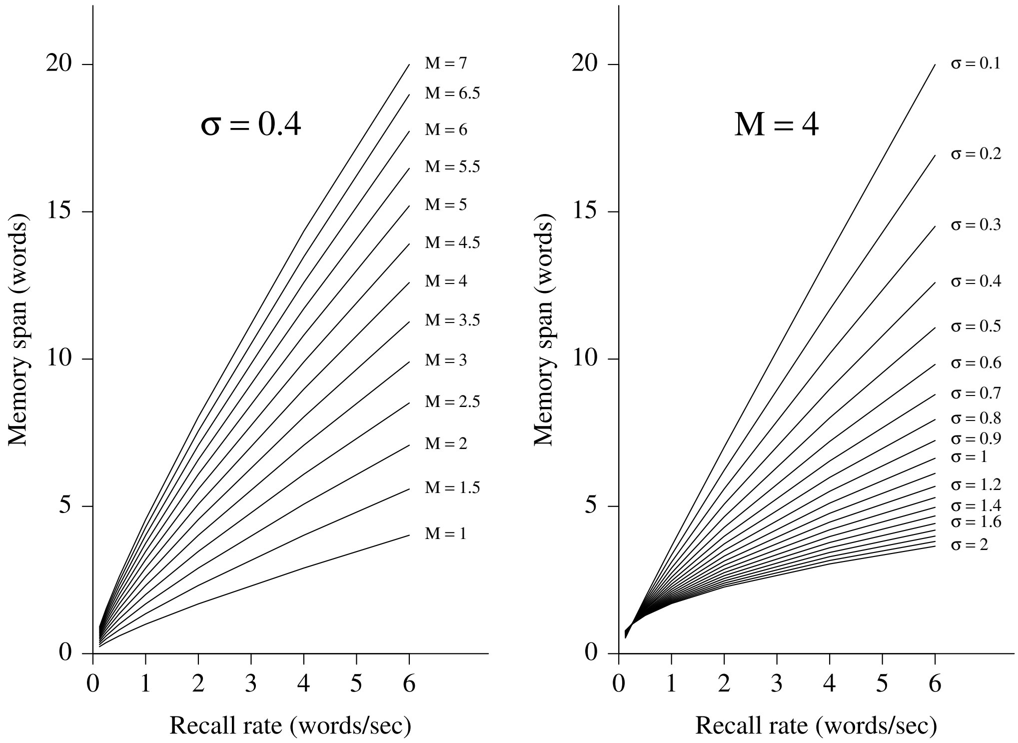

Fig. 3. The relationship between memory span and recall rate, as a function of changes in parameters M and σ of the log-normal decay distribution, which affect its median and shape. The left panel shows how changes in M affect the relationship when σ is fixed at 0.4; the right panel shows how changes in σ affect the relationship when M is fixed at 4 s.

influence memory accuracy. By not accounting for this variability, the item-based decay model may overestimate the variability of the decay times. Yet these differences appear only when the model is being fit to aggregated data; if fit on a trial-by-trial basis, the curvature and variability would be incorporated implicitly into the item-based decay model. Thus, Schweickert and Boruff's model enjoys a justifiable approach to fitting averaged data, which the item-based model lacks.

Another way in which the models differ is in their specification of a decay distribution. Schweickert and Boruff's model assumes that list decay times have a normal distribution, whereas the decay distribution in the item-based model is not specified explicitly. To fit empirical data, I chose a two-parameter log-normal decay distribution, which produced excellent fits to individual psychophysical functions, but it may be possible to improve the goodness-of-fit by using different decay distributions. Thus, Schweickert and Boruff's model is clearly more testable, at the cost of being less flexible. Yet Schweickert and Boruff's selection of a normal decay is probably not a central assumption of their model, and another decay distribution could be easily chosen.

The most important way in which the models differ is in how they view the memory trace: in Schweickert and Boruff's model, the decay time represents the functional decay time of the entire list, whereas in the item-based decay model the decay time distribution is of individual items. The critical test of this difference is to determine whether the number of items in a list matter, once total list production time is controlled for. This could hypothetically be tested by constructing lists of different numbers of words with the same total articulation time. However, this type of manipulation is notoriously difficult to carry out (cf. Mueller, et al., 2003, for a review), because it almost necessarily introduces confounds. Whether any experiment could offer a compelling test of this difference between models remains to be seen.

6 Summary and Conclusions

The two-second decay conjecture has become part of the lore of short-term memory, even though it defies introspective and experimental evidence. The strongest theoretical support for the conjecture was provided by the trace decay model of Schweickert and Boruff (1986), but this model accounted for decay of lists as holistic entities, rather than as a set of individual items. If, instead, an item-based decay model is assumed, the best estimate of the mean decay time is longer, and its length increases as the variability of the decay distribution increases and deviates from a "tape-loop" decay distribution. In this new model, list recall accuracy decreases as list length increases because (1) more words must be recalled correctly, and (2) the mean delay between presentation and recall increases. When items were assumed to decay according to a log-normal distribution, the mean decay time appears to be around four seconds, more than twice that of earlier estimates. These conclusions are all based on the assumption that decay is the only way information disappears from short-term memory; to the extent that information is lost through the action of other mechanisms, the true decay distributions of individual items in verbal working memory is likely to be even longer, in greater agreement with both introspective and empirical evidence.

7 References

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341–380.

Baddeley, A. & Lewis, V. (1984). When does rapid presentation enhance digit span? *Bulletin of the Psychonomic Society, 22*, 403–405.

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

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

Byrne, M. D. (1998). Taking a computational approach to aging: The SPAN theory of working memory. *Psychology & Aging, 13*, 309–322.

Corballis, M. C. (1969). Patterns of rehearsal in immediate memory. *British Journal of Psychology*, 60, 41–49.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin*, 96, 341–370.

Dosher, B. (1999). Item interference and time delays in working memory: Immediate serial recall *International Journal of Psychology. Special Issue: Short-term/working memory, 34*, 276–284.

Dosher, B., & Ma, J. (1998). Output loss or reharsal loop? Output-time versus pronunciation-time limits in immediate recall for forgetting-matched materials. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24*, 316–335.

Drewnowski, A., & Murdock, B. B. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning & Memory, 6*, 319–332.

Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.

Hastings, N. A. J., & Peacock, J. B. (1975). *Statistical distributions.* London: Butterworth.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1217–1232.

Hulme, C., Newton, P., Cowan, N., Stuart, G., & Brown, G. (1999) Think before you speak: Pauses, memory search, and trace redintegration processes in verbal memory span. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*, 447–463.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 183–223), New York: Cambridge University Press.

Landauer, T. K. (1962). The rate of implicit speech. *Perceptual and Motor Skills, 15*, 626.

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1353–1380.

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

Reitman, J. (1971). Mechanisms of forgetting in short-term memory. *Cognitive Psychology, 2*, 185–195.

Reitman, J. (1974). Without surreptitious rehearsal, information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior, 13*, 365–377.

Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: Neuromagnetic Evidence. *Journal of Cognitive Neuroscience, 5*, 363–370.

Schweickert R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology, Learning, Memory, and Cognition, 12*, 419–425.

Schweickert, R., Hayt, C., Hersberger, L., & Guentert, L. (1996). How many words can working memory hold? A model and a method. In S. E. Gathercole (Ed.), *Models of Short-term Memory.* East Sussex, UK: Psychology Press, 267–293.

Schweickert, R., Chen, S., & Poirier, M. (1999). Redintegration and the useful lifetime of the verbal memory representation. *International Journal of Psychology, 34*, 447–453.

Shiffrin, R. (1973). Information persistence in short-term memory. *Journal of Experimental Psychology, 100*, 39–49.

Standing, L., Bond, B., Smith, P., & Iseley, C. (1980). Is the immediate memory span determined by subvocalization rate? *British Journal of Psychology, 71*, 525–539.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelrnach (Ed.) *Information processing in motor control* (pp. 117–152). New York: Academic Press.

Watkins, M. J., & Todres, A. K. (1980). Suffix effects manifest and concealed: Further evidence for a 20-second echo. *Journal of Verbal Learning and Verbal Behavior, 19*, 46–53.