N88-23386

SHORT-TERM MEMORY LOAD AND PRONUNCIATION RATE

Richard Schweickert
and
Cathrin Hayt
Purdue University
West Lafayette, IN 47907

One of the major components determining mental workload is the amount of material that must be maintained in short term memory. Some tasks, such as air traffic control, involve coordination between people, and the main communication is verbal. Critical parts of the communication require memory not only for the gist or meaning of the material, but for verbatum recall (ref. 1). Even tasks which do not involve communication between people often have a verbal component. Communication between humans and computers often requires the human to remember certain information verbatim which has disappeared from the screen (ref 2). Everyone has had the experience of looking up the call number of a book in a library, and rehearsing it while trying to find the shelf.

The capacity of short term memory was described in a classic paper by Miller (ref. 3) as 7 plus or minus 2 chunks, where a chunk is a meaningful unit of material. This gives a good rule of thumb, but it has at least two problems. First, the number seven is an estimate of the memory span, that is, the number of items that can be immediately recalled correctly half the time. But there is nothing special about probability one-half. In most practical situations, we would like to be able to predict probability of correct recall over a range of probabilities, or at least be able to estimate the length of a list that can be recalled with a high probability, say, .99. The second problem is that the probability of correct recall depends on the type of material. The memory span is greater for color names, such as red and orange, than it is for shape names, such as circle and square. Although one can define the capacity of the short term memory to be 7 chunks, this leads to the curious notion that there are more chunks in the name of a shape than in the name of a color.

Another approach is to assume the short term memory is limited in the time for which it can hold items. The support for this has waxed and waned over the years, but the decay hypothesis has enjoyed renewed interest recently. This is because Mackworth, Baddeley, and others have found that the memory span for a type of material can be predicted quite well from the amount of material that can be pronounced in about 1.5 seconds (refs. 4, 5, 6). For example, the memory span for digits is 7.98 and that for four-letter concrete nouns is 5.76 (ref. 7). It turns out that these are the number of digits and nouns, respectively, that a typical subject can pronounce in 1.5 seconds.

This result can be summarized by saying

$$S_i = 1.5 \text{ sec} \times r_i, \qquad (1)$$

where S_i is the memory span for items of type i and r_i is the rate of pronunciation of items of type i, in items/sec.

PRECEDING PAGE BLANK NOT FILMED

PAGE 230 INTENTIONALLY BLANK

The explanation is straightforward. Suppose when a subject is presented with material for immediate recall, he forms a verbal trace, and the trace begins to decay. If the subject can emit the items before the trace has deteriorated, recall will be correct, otherwise it will be incorrect. Evidently, on the average, the trace decays after 1.5 seconds, which determines the span.

Equation 1 resolves the second problem, accounting for differences in memory span for different types of material in terms of differences in their pronunciation rates. Schweickert and Boruff (ref. 6) proposed a resolution to the first problem by saying the probability of correct recall is simply the probability that the duration of recall is less than the duration of the verbal memory trace,

$$P = \text{Prob}\,[T_r < T_v], \tag{2}$$

where P is the probability of correct recall, T_r is the time the subject requires to recall the list, and T_v is the duration of the memory trace. In an experiment, subjects were presented with 6 list lengths of 6 types of material. A good account of the data was given by Equation 2. Normal distributions were assumed for T_r and T_v. The mean and variance of the trace duration were estimated to be 1.88 sec and .187 sec^2, respectively.

An equally good, but more easily calculated, estimate of the probability of correct recall was found, based on linear regression,

$$z = -2.02\ T_r + 3.87. \tag{3}$$

Here z is the standard normal deviate of the probability of correct recall of a list, and T_r is the average amount of time required to read the list aloud.

The correlation between the z-score for correct recall and pronunciation time was .977, so 95% of the variance is accounted for by pronunciation time. In contrast, the analogous linear regression equation using the number of items in the list as the predictor yielded a correlation of .849, so only 72% of the variance is accounted for by list length.

It is of interest to note that Equations 2 and 3 underestimated the probability of correct recall for digits, the material subjects had most experience with in daily life, and overestimated the probability of correct recall for nonsense syllables, the material least familiar to the subjects. The subjects in the experiment were not particularly practiced. They came for three one hour sessions, and learned only 60 lists of each material type. The nonsense syllables are hardly chunks, in the usual sense. The following experiment was done to investigate memory in highly practiced subjects.

Method

<u>Subjects</u>. Two subjects completed 4 practice sessions followed by 30 test sessions. They were paid by the hour. Each session lasted about an hour and a half.

<u>Materials</u>. Five types of material were used: consonants, color names, prepositions, shape names, and three letter concrete nouns. To make the

probability of correct guessing low, each set contained 20 items. This precluded the use of digits, a commonly used material in immediate memory studies. Lists of a given material were all presented together in a block. The order of presentation of materials within sessions was governed by six 5 x 5 Latin squares. The lengths of the lists were from 3 to 9 items, inclusive. List lengths were randomized within the blocks.

Procedure. At the beginning of each trial, a list appeared on a TV monitor. In pronunciation trials, subjects read the list aloud with no requirement to remember it. In memory trials, subjects read the list aloud, and then attempted to recall it by speaking aloud. Voice keys indicated the onset and offset of their speaking, and the durations of the utterances were timed with a microcomputer. The pronunciation and recall times are beyond the scope of this paper.

During recall, the experimenter recorded whether the list was correctly recalled or not.

Results

The reading time for a list is the time from when the subject started to read the list until he finished. Reading was followed immediately by recall. Mean reading times and probability of correct recall are given in Tables 1 and 2.

Recall that for the unpracticed subjects in the experiment of Schweickert and Boruff (ref. 6), reading time was a much better predictor of recall than the number of items in the list. Here, the number of items is a better predictor, although only slightly.

For subject 1, the correlation between the z-score for correct recall and the number of items in the list is -.95, so 90% of the variance in recall is accounted for by list length. The correlation between the z-score for correct recall and reading time is -.90, so 80% of the variance in recall is accounted for by reading duration.

For subject 2, the results are similar. The correlation using the number of items in the list was -.95, so 90% of the variance is accounted for by list length. The correlation using reading time is -.92, so 85% of the variance is accounted for by reading time. In each case, list length does slightly better as a predictor than reading time.

The regression equation for predicting the z-score for correct recall is

$$z = b_0 + b_1 n,$$

where n is the number of items. For subject 1, the regression coefficients were $b_0 = 5.50$ and $b_1 = -.83$. For subject 2, they were $b_0 = 5.40$ and $b_1 = -.80$. The coefficients agree remarkably well for the two subjects.

In the calculations, conditions with recall probabilities of 0 or 1 were ignored, since the corresponding z-scores are infinite.

Is there an advantage of practice? One way to evaluate this is to note that the duration of a list recalled half the time was about 2.4 seconds, compared with 1.8 seconds for the unpracticed subjects in the previous experiment.

Increasing the length of the items leads to two competing tendencies. First, the longer the items, the greater the time required to output the list, so the greater the chances of trace decay before recall is completed. But, second, the longer the items, the more distinctive they tend to be, and hence the greater the chances of guessing an item correctly from a partial trace. Highly practiced subjects are probably better able to reconstruct the partially decayed trace of an item to make a correct guess. The more familiar the items are, the better subjects are able to discriminate the fragments remaining in the traces.

For unpracticed subjects, reading time is a notably better predictor of immediate recall than the number of items in the list. For practiced subjects, the two predictors do about as well, with a slight advantage for the number of items. In either case, about 90% of the variance is accounted for, so for most practical purposes, good estimates of recall probability are available. If the items that must be recalled are likely to be unfamiliar, and likely to remain unfamiliar, then it is advantageous to keep the items short. For example, codes for identifying airplanes or pilots encountered only once in a while should be short to pronounce. On the other hand, if the same items will be encountered over and over again, it is advantageous to concentrate efforts on making them distinctive, even at the cost of adding to the number of syllables.

REFERENCES

1. Lenorovitz, D. R., & Phillips, M. D.: Human factors requirements engineering for air traffic control systems. In G. Salvendy (Ed.), Handbook of human factors. New York: John Wiley, 1986.

2. Card, S. K., Moran, T. P., & Newell, A. The psychology of human-computer interaction. Hillsdale, N. J.: Lawrence Erlbaum, 1983.

3. Miller, G. A. The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97, 1956.

4. Mackworth, J. F. The duration of the visual image. Canadian Journal of Psychology, 17, 62-81, 1963.

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

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

TABLES

Mean Reading Times and Probability of Correct Recall

Table 1: Subject 1

List Length		3	4	5	6	7	8	9
Colors:	Read	.889	1.273	1.656	2.099	2.499	2.945	3.430
	Recall	1.000	1.000	.987	.832	.500	.191	.000
Letters:	Read	.667	.951	1.297	1.659	2.072	2.451	2.896
	Recall	1.000	1.000	.967	.846	.592	.242	.023
Preps:	Read	.867	1.212	1.617	2.018	2.428	2.840	3.275
	Recall	1.000	.993	.940	.805	.415	.113	.020
Shapes:	Read	1.254	1.831	2.399	2.972	3.537	4.037	4.637
	Recall	1.000	.987	.866	.513	.128	.014	.000
Words:	Read	.827	1.195	1.561	1.967	2.380	2.817	3.254
	Recall	1.000	1.000	.931	.685	.281	.055	.000

Table 2: Subject 2

List Length		3	4	5	6	7	8	9
Colors:	Read	.920	1.353	1.764	2.234	2.696	3.151	3.635
	Recall	1.000	1.000	.967	.839	.476	.148	.020
Letters:	Read	.677	1.104	1.468	1.959	2.293	2.743	3.130
	Recall	1.000	.987	.980	.890	.710	.345	.094
Preps:	Read	.883	1.208	1.647	2.067	2.488	2.897	3.287
	Recall	1.000	.993	.967	.879	.537	.208	.053
Shapes:	Read	1.621	2.200	2.790	3.356	3.993	4.574	5.039
	Recall	.993	.953	.800	.547	.157	.013	.000
Words:	Read	.873	1.239	1.664	2.145	2.581	3.010	3.433
	Recall	.993	.993	.927	.627	.366	.088	.007