*Reprinted from the* Proceedings, 76th Annual Convention, APA, 1968

# PHONEMIC MODEL OF SHORT-TERM AUDITORY MEMORY

GEORGE SPERLING
***Bell Telephone Laboratories, Inc., Murray Hill, New Jersey***

*S*s generally recall strings of acoustically similar (AS) letters less accurately than strings of acoustically different (AD) letters (Conrad, 1963).[1] A deficit in recall of AS letters (*AS-deficit*) occurs not only with sequential auditory letter presentations, but also with sequential visual presentations and even with simultaneous visual presentation of the letters. It occurs in partial report, full report, and running memory span tests of recall.[1] The phonemic model of short-term auditory memory (STAM) proposed here predicts AS-deficits from the phonemic structure of an AS alphabet and from parameters established from AD recall scores, i.e., without requiring any parameters to be estimated from experiments with AS-stimuli.

*Estimating capacity.* The model distinguishes *capacity* and *performance.* Performance refers to the score observed in a recall task; capacity is a theoretical limit on all performances. For example, scores in total recall tasks are less than the capacity because, in attempting to report some letters, *S* forgets others.

The interference of recall with retention is minimized in partial report procedures which require *S* to report only a small number of letters—designated at random—from the stimulus. In the theoretical limiting case, when the number of letters to be reported is one, and when the cue designating the letter is easily interpreted (i.e., without its interfering with the contents of memory), the partial report procedure estimates capacity directly. In practice, partial reports of 4 letters from among 12 suffice to give estimates of capacity (after correction for the occasional inability to report just 4 letters).

Using this method, Sperling and Speelman (1967) estimated capacity of STAM as 7.5 AD-letters (plus an auditory poststimulus cue) and 5.3 AS-letters (plus the cue). The estimates varied less than ±¼ letter with presentation rate.

*Phonemic model of capacity.* The basic assumptions of the phonemic model are: (*a*) when letters are presented (or when they are rehearsed), all phonemes of the letters are stored in memory; (*b*) once in memory, constituent phonemes of letters are retained or lost independently; (*c*) at recall, if only 1 of the 2 phonemes of a letter is still available, a guess is made from among those letters of the alphabet which contain the retained phoneme in the same position (initial or terminal). For example, in the AS alphabet (b, c, d, g), retention of the phoneme $\bar{e}$ is of no value because it does not discriminate between the letters of the alphabet.

If the probability $p(i)$ of recalling a constituent phoneme $i$ of the first $n$ letters in memory is $p(i) = 1, 1 \leqslant i \leqslant 2n$, and of the remaining letters is $p(i) = 0, i > 2n$, then obviously no difference in memory will be observed for stimuli of different alphabets. In the phonemic model, differences between alphabets result only from partial recall of letters. To make an exact prediction of the difference between alphabets requires knowledge of how $p(i)$ varies from 1.0 to 0.0. While there is considerable basis for assuming exponential decay processes in memory (Atkinson & Shiffrin, in press; Norman, 1966), it is parsimonious to avoid the usual dual-parameter description of exponential decay, in which strength decays with one parameter and a second parameter relates strength to $p(i)$. Therefore, it is assumed simply that $p(i)$ decays exponentially with $i$ (see

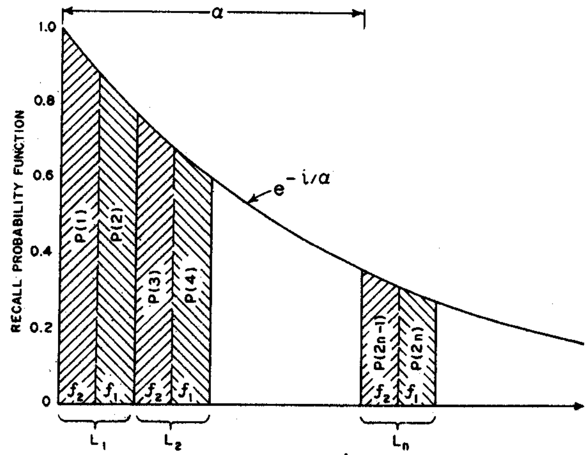


Fig. 1. Theoretical phonemic memory. (Abscissa: phonemic recency, most recent phoneme is numbered 1. Ordinate: recall probability defining function. Shaded areas indicate $p(i)$, the probability of recalling the $i$th phoneme. The f's indicate conditional probability of correct recall of a letter given recall of only its first ($f_1$) or its second ($f_2$) phoneme.)

Fig. 1). This choice of $p(i)$ yields a one-parameter description of the contents of memory in terms of the total number of phonemes (area under $p(i)$).

An AD letter capacity of 7.3 letters implies a phonemic capacity of $a$ = 16 or 25 phonemes, depending on whether or not the poststimulus cue is considered to be in memory. The AS capacity-deficit then is predicted to be 1.7 or 1.8 letters, again depending on whether or not the poststimulus cue is considered to be in memory. Fortunately, assumptions about the poststimulus cue or about the exact shape of distribution are not critical for predicting AS-deficits.

*Predicting performance.* To predict performance directly from the phonemic model of memory, it would be necessary to make detailed assumptions about *S*s' patterns of rehearsal, and about how rehearsed letters are stored in STAM. (It was not necessary to make these assumptions to estimate capacity because capacity of STAM appears to be independent of rehearsal—Sperling & Speelman, 1967). The procedure for predicting performance consists of two parts: (*a*) the phonemic model for capacity, (*b*) a set of rules—more closely related to the data than the model—to calculate performance from capacity.

*Rules.* The pivotal concept of the rules is the memory span, which is derived from the letter capacity $a$ of STAM but which is reduced by an amount $r$ which depends on presentation rate. The mean memory span $m$ is given by $m = a - r$, where $r$ equals 0.7, 1.0, 2.3 letters for presentation rates of 1/sec., 2/sec., 4/sec. Memory span is assumed to be

[1]For additional references and for details see Sperling and Speelman (1967).

normally distributed with $\sigma = 1.2$ letters. (*a*) When a subject attempts to report all the letters he can from a presented list of length $n$, the predicted score, span-$n$, is given by the smaller of $(m,n)$. (*b*) When he makes a partial report of $n$ letters from among $k$ presented letters, the predicted mean score is (span-$n$) $(a/k)$, for $k > a > n$. (Usually the partial report score is converted into "number of letters available" by multiplying it by $k/n$.) (*c*) In a running memory experiment ($S$ reports only the last letters of a long list) the predicted score is simply $.6m$.

Because alphabet does not enter into the rules, all differences in performance between AD and AS stimuli derive from the initial differences in letter capacity for these alphabets. The rules translate the initial capacity difference into predicted AS-deficits. A large capacity difference does not necessarily imply a large predicted AS-deficit in performance. For example, predicted and observed performance on lists of length-3 is virtually perfect for AD and AS alphabets.

*Goodness of fit.* Predictions were based on 38 conditions studied by Sperling and Speelman (1967) which included 3 procedures (whole report with list lengths of 4 to 12 letters, running memory span, and partial report), 3 presentation rates (1/sec., 2/sec., 4/sec.) and 2 alphabets (AD = f, h, k, l, m, q, x, y; AS = b, c, d, g, p, t, v, z). An iterative optimization procedure was used to obtain the best overall values of the parameters; these are the values given in the text. The predictions account for .96 of the variance of the 38 observed scores (Fig. 2).

These same parameters are not optimal for predicting the 19 observed AS-deficits, nevertheless they account for .78 of the variance of these data. Considering the small range of observed AS-deficits and their greater error (each is the difference of 2 scores), this also is a satisfactory prediction.

How well can the rules account for the AS scores when their parameters are estimated only from AD scores, and the capacity deficit is estimated simply from the phonemic structure of the alphabets? This prediction uses no data whatsoever from the 19 AS-scores; it accounts for .90 of their observed variance.

The value of the AS capacity-deficit which gives a best fit to all the data is 1.79 letters, which happens to be identical to the value predicted from the difference in phonemic efficiencies of the AD and AS alphabets. However, when the AS capacity-deficit is estimated from only the partial report scores—which theoretically give the best

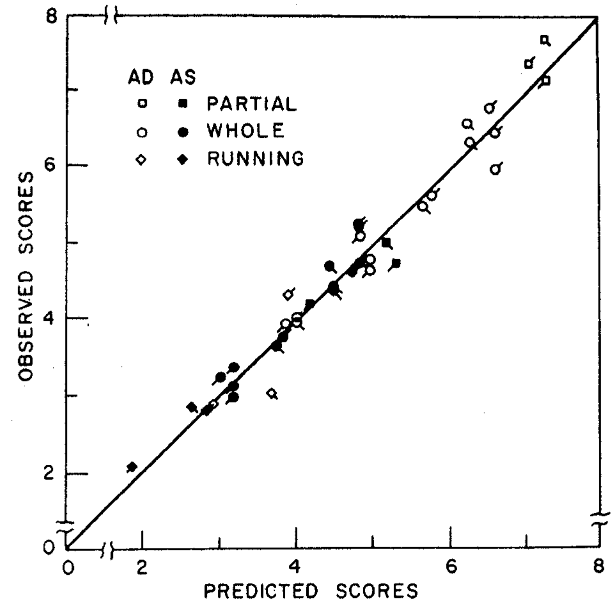


Fig. 2. Scatter plot of observed vs. predicted scores in 38 auditory conditions. (Diagonal strokes indicate presentation rate: up-right = 1/sec., down-right = 2/sec., down-left = 4/sec. Procedures indicated are: partial report [scored for number of letters available], whole report, and running memory span.)

estimates of capacity—it is 2.25 letters. This discrepancy, and some other smaller ones, suggest that even the good observed fit of theory to data is not without strain, that a better theory is needed.

*Conclusion.* The twin concepts of memory for phonemes and of phonemic efficiency of an alphabet yield good predictions of observed AS-deficits. Simple phonemic-inefficiency of AS alphabets accounts for most of the adverse effect of acoustic similarity on recall.

## REFERENCES

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* Vol. II. New York: Academic Press, in press.

Conrad, R. Acoustic confusions and memory span for words. *Nature,* 1963, 197, 1029-1030.

Norman, D. A. Acquisition and retention in short-term memory. *Journal of Experimental Psychology,* 1966, 72, 369-381.

Sperling, G., & Speelman, R. G. The effect of sound of stimuli on short-term memory. *Bell Telephone Laboratories Technical Memorandum,* 1967.

For a fuller acount of the experiments and theory described herein see:

Sperling, G. and Speelman, R. G. Acoustic similarity and auditory short-term memory: experiments and a model. In D. A. Norman (Ed.), *Models of Human Memory.* New York: Academic Press, 1970. Pp. 149-202.