

Interactive presentation in multitrial free recall

MICHAEL FRIENDLY and PAT FRANKLIN
York University, Downsview, Ontario M3J 1P3, Canada

This paper explores whether free recall performance and organization can be facilitated using computer-controlled interactive presentation, where stimulus presentation on Trial N+1 depends on the individual subject's prior recall history. A series of experiments investigated the effects of two types of response-contingent presentation order manipulations and of selective tagging manipulation designed to highlight the presentation of previously nonrecalled items. An initial experiment using 20-word lists showed no differences among groups, due to a ceiling effect. Two subsequent experiments, using longer lists (30 and 40 words), found that recall and subjective organization were increased by selective tagging and by presentation orders that preserved the subject's prior order of recall. The results demonstrate that subjects' previous recall histories contain information that can be used to facilitate recall. The potential of other interactive manipulations is discussed.

This paper describes a series of initial experiments using an interactive free recall procedure to study memory organization and the retrieval of information from memory. In standard multitrial free recall (SFR), the learner is presented a list of items to memorize, usually in random order, one item at a time. The learner then attempts to recall as many items as possible, in any convenient order. Interactive free recall (IFR) differs from this description principally in that what is presented to the subject on Trial N + 1 depends in some way on the subject's performance on one or more of the previous trials.

Interactive presentation is of interest for two reasons. First, free recall is a well-defined learning task in which the subject's active processing and use of linguistic knowledge are known to affect recall. In an interactive version, we can monitor some aspect of an individual's recall from one trial to the next and attempt to optimize that person's performance by presenting the items based on their earlier recall.

Second, a popular theory is that free recall learning reflects the grouping of distinct list items into functional memory units that grow in size, but not in number, over repeated trials (Postman, 1972; Tulving, 1962, 1966). According to this view, success in recall depends directly on the strength, coherence, and stability of this organizational scheme constructed by the learner. Interactive presentation provides a way to conduct strong tests of theories of organization and recall, since a manipulation derived from organization theory

This work was supported by National Research Council of Canada Grant A8615 to the first author. Address requests for reprints to Michael Friendly, Psychology Department, Room 210 BSB, York University, Downsview, Ontario M3J 1P3, Canada. We are grateful to Stephen Madigan, who provided us with some initial ideas for computerized memory experiments, and to Lisa Pollack and David Hoffman, who ran subjects and helped to prepare this report.

and applied interactively based on the subject's own recall history should have direct consequences on learning.

Some simple interactive manipulations have been previously examined, but, unfortunately, they have produced mixed results. As a consequence, the study of interactive presentation in free recall seems to have been abandoned without, in our view, having been adequately tested.

The most straightforward interactive manipulation is to make the order of presentation contingent on the subject's recall on the previous trial. Murdock, Penney, and Aamiry (1970) placed the items recalled on Trial N at the beginning or at the end of presentation on Trial N + 1. Compared with a control group that received randomized presentation (SFR), the experimental groups recalled slightly more. An analysis was performed using Tulving's (1964) CC and NC measures [the conditional probability that a word is recalled on Trial N + 1, given that it had been (CC) or had not been (NC) recalled on Trial N]. This showed that the group that had the previously recalled words presented at the beginning of the list had a higher NC, whereas the group that had the previously recalled words presented at the end of the list had a higher CC. Thus, the effect for both these manipulations was to elevate recall of the items that were selected to appear at the end of the list.

However, since recall in this experiment immediately followed presentation, with no distractor task, these results might merely reflect a strong recency effect, as Murdock, Anderson, and Ho (1974) subsequently suggested. Further experiments by these investigators used a sorting task to establish stable, subjective categories and a distractor task to wipe out any possible recency effects, but they showed no differential effects for beginning and end presentation

manipulations. In one experiment, the items were blocked according to the subject's own sorting categories, which should have facilitated recall in both interactive conditions. However, the lack of a control group receiving random presentation orders seriously weakens the conclusion of Murdock et al. (1974) that manipulating presentation order interactively is ineffective.

The present studies were undertaken to reevaluate the effects of simple interactive manipulations on recall performance. These experiments investigated two response-contingent presentation orders and a control condition. In the control condition (SFR), all words were randomized on Trial N + 1, regardless of the subject's recall on Trial N. One contingent presentation order (B) was based directly on Murdock et al. (1970). Words recalled on Trial N were presented at the beginning of the list, in the order of recall, on Trial N + 1, followed by the nonrecalled words in random order. The second contingent presentation order was based on the notion that physical contiguity between the recalled and the nonrecalled items might be more effective in integrating new items than the beginning manipulation, in which they are presented separately. Accordingly, in the second presentation order condition the items that the subject recalled on Trial N were presented in order of recall, but were meshed (M) with the nonrecalled items, interspersed in random order.

These presentation orders are based on the assumption that items recalled together are organized together. It should therefore be helpful to present list items in the order of recall. The question arises whether some other response-contingent manipulation that does not involve presentation order would also facilitate recall. We based a second interactive manipulation on Buschke's (1973) selective reminding procedure. On each trial, Buschke presented only those items that had not been recalled on the previous trial. This made the previously nonrecalled items more salient and facilitated recall. We did not wish to present only nonrecalled items, since this would confound presentation time with the interactive manipulations. So in our experiments we used a selective tagging condition, in which nonrecalled items from Trial N were made more distinctive by being printed with an asterisk (*) to the left of the word on Trial N + 1. If items are stored independently and retrieved on a strength basis alone (rather than through links with other items), then this "tag" manipulation should increase recall, without necessarily increasing consistency in subjects' output orders. Thus, the tag manipulation should facilitate recall, but it may or may not increase organization in recall. The tag manipulation was combined with the presentation order conditions, making six groups in all, three with and three without the tag.

An initial experiment involved lists of 20 words and used 10 subjects for each condition. Although

recall was least for the SFR baseline group, there were no significant differences among any of the groups, and the only significant effect was that of trials [$F(5,54) = 69.4, p < .01$]. It was clear that the lists had been too short, since there was a large ceiling effect. Accordingly, we treated this experiment as a pilot study and repeated it with longer lists.

EXPERIMENT 1

Method

Design. The pilot study and two further experiments had the same basic design and procedure. The design is formally a 3 by 2 factorial with six independent groups representing the combination of the three presentation orders (SFR, B, M), either with or without the tagging (*) of previously nonrecalled words. Thus Groups B* and M* received a response-contingent presentation order, together with selective tags, and SFR* received the tag manipulation alone. For purposes of analysis and interpretation, we are interested primarily in the separate effects of the tag and presentation order manipulations. Hence, it is more useful to regard the SFR condition (no tag) as the relevant control group for all comparisons (i.e., as a one-way design with six groups).

Apparatus. The entire experimental session was run under computer control using a CRT display terminal (Datamedia 1520A) linked to an IBM/370 computer by a 30-cps communication line. The computer system used in this research was derived from a general software system developed for running memory experiments (Friendly & Franklin, 1979). The system consists of a control program for running the experiment and computer files containing word pools and experimental parameters required by the program (e.g., selection parameters for the word pool, presentation rate, numbers of words, and trials). For each trial, the control program prepared the appropriate presentation order, displayed the items on the video-screen, and accepted and scored the subject's typed responses. In scoring the responses, the program used a spelling-correction algorithm based on work by Kellerman (Note 1) to minimize the influence of spelling errors. At the end of the session, the control program initiated the printing (off-line) of a detailed summary of the experimental session and stored a data record on file for subsequent analysis.

Word Lists. The word pool used in these experiments was based on the Paivio, Yuille, and Madigan (1968) norms. This contains 925 unrelated words, rated for imagery, concreteness, meaningfulness, and frequency of occurrence; we added numbers of letters and syllables as additional selection variables. The word list for each subject was a random sample of all words in the pool that fell within the range specified for each of the six stimulus variables. In these experiments, the selection ranges were the same in all conditions and served merely to eliminate 72 words that had extreme ratings on one or more of the variables. The excluded items were largely long (more than 12 letters) low-frequency words. By this procedure, each subject learns a unique random sample of items from a known population. This avoids problems of generalizability of the results over samples of materials, as well as statistical problems that result from the use of a small, fixed set of lists (Clark, 1973).

Procedure. Each subject was given two practice trials on a 12-word list to familiarize him or her with the terminal and the procedure. They were then given the experimental condition, in which they learned a list of 30 items for four presentation-recall trials.

Subjects in the selective tagging condition were instructed that, "on the second and subsequent trials, you will notice a star (*) by some of the words; these are the words you

did not recall on the previous trial." No subjects were advised of the presentation order manipulations.

The first trial was identical for all subjects. The screen displayed "TRIAL 1," and then the words were presented one at a time, at a rate of 2.0 sec/item. After the last word was displayed, the subject was occupied with a series of two-digit additions for 15 sec to prevent recall from immediate memory. Then the word "RECALL" was displayed, and the subject typed as many of the words as he could remember, in any order. The words were displayed on the screen as he typed them in, so he was able to review his recall. The recall period was calculated at 8 sec/item, so the maximum time was 4.0 min.

Subjects. Ten subjects served in each of the six conditions. The subjects were paid volunteers obtained by advertisement from within the university. They were mostly undergraduates. Some subjects had taken part in the pilot experiment. These subjects did not receive the practice trials and were assigned to conditions at random, with the constraint that the group must be at a level of each factor that was different from the group in which they had previously participated.

Results and Discussion

Subjective organization scores for bidirectional pairs were calculated using the method described by Pellegrino (1971). Mean organization scores (observed-expected ITR) collapsed over trials are shown in Table 1. The mean recall for the tagged and nontagged conditions is shown in Figure 1. Because of the correlation between mean recall and mean subjective organization scores ($r = .74$, within cells), a multivariate analysis of variance (MANOVA) was performed, using mean recall and mean subjective organization as the primary dependent variables for between-groups effects (Bock, 1975). Since the data were collected over four trials, the slope over trials (linear trends) for recall and for organization were included in the analysis as additional variables.

The overall multivariate test of the baseline SFR group against the experimental conditions variable (tagged and nontagged) was highly significant, using Rao's F-transformation for Wilks' lambda [$F(4,51) = 137.52, p < .001$]. Univariate tests on the four variates showed not only that the experimental groups recalled more words and organized their recall more than the SFR control group, but also that the rate of increase of these variables over trials (slopes) was greater for the interactive conditions (all at the

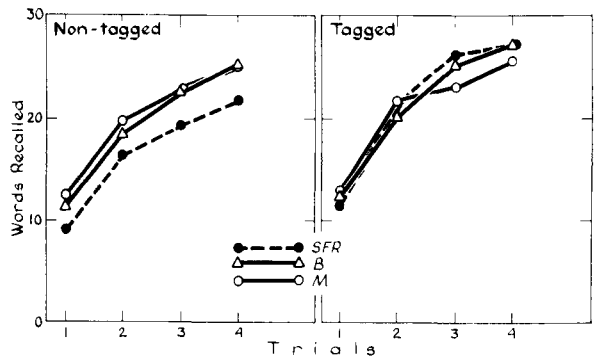


Figure 1. Mean number of words recalled in Experiment 1.

.001 level). In addition, the contrast of SFR (non-tagged) with SFR (tagged) was highly significant [$F(4,51) = 35.00, p < .001$]. In the univariate tests, this difference was significant for mean recall [$F(1,54) = 92.24, p < .001$], the mean subjective organization [$F(1,54) = 7.68, p < .01$], and the recall slope [$F(1,54) = 42.63, p < .001$], but not for the organization slope [$F(1,54) = 1.45, n.s.$].

These analyses support the conclusion that the effects on recall are mediated by the five experimental conditions facilitating organization of the lists.

The nature of the difference between the experimental groups and the SFR nontagged control group can be explored using a conditional probability analysis of recall. If organization develops by the accretion or new items into an ongoing organizational scheme, then the differences between groups should appear in the conditional probability of recalling items that were not recalled on the previous trial (NC), rather than the probability of recalling items that were previously recalled (CC).

The conditional probability data shown in Figure 2 confirm that the locus of the effect was not in the maintenance of previously recalled words, but in the recall of words that were not recalled on the previous trial. The words that were previously recalled are presumed to be part of the developing list structure; they had a high probability of recall on the next trial, with no differences between the groups. The NC scores were greater when nonrecalled items were tagged [$F(1,54) = 4.7, p < .05$]; among the nontagged experimental conditions, the effect of presentation order was negligible until the end of the experiment. The difference among these three NC curves failed to reach significance overall, but the curves did differ in slope [$F(2,54) = 10.8, p < .05$], probably due to data from the last pair of trials.

EXPERIMENT 2

Experiment 2 was intended to replicate the results of Experiment 1 with a new group of subjects. Since the NC curves in Experiment 1 diverged only toward the end of learning, we increased the number of trials. We also

Table 1
Mean Subjective Organization Scores, Experiment 1

| Group | Trial Pair | | | Mean | Slope |
|--------|------------|------|-------|------|-------|
| | 1-2 | 2-3 | 3-4 | | |
| SFR | 1.42 | 3.01 | 4.04 | 2.82 | 1.31 |
| Begin | 1.38 | 4.99 | 6.39 | 4.25 | 2.50 |
| Mesh | 1.87 | 5.02 | 6.90 | 4.60 | 2.52 |
| SFR* | 1.73 | 5.02 | 5.71 | 4.15 | 1.99 |
| Begin* | 2.41 | 8.17 | 10.71 | 7.10 | 4.15 |
| Mesh* | 4.04 | 6.51 | 9.42 | 6.66 | 2.69 |

Note—The measure used is the bidirectional ITR for Unit Size 2, expressed as (observed - expected). *See text for explanation.

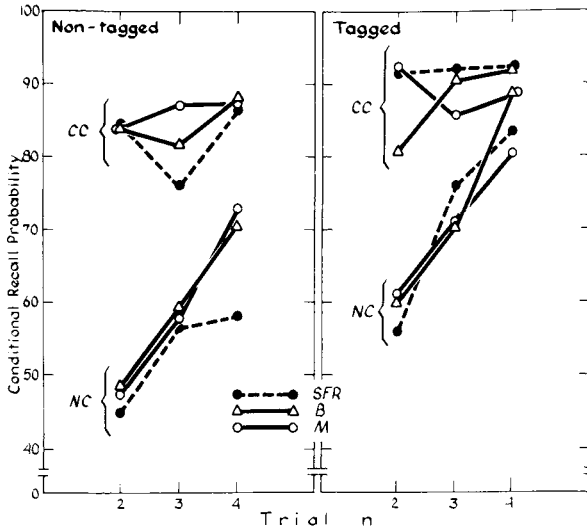


Figure 2. Conditional recall probability, Experiment 1. CC denotes $P(C_n | C_{n-1})$, the probability of recalling an item on trial, given its recall on Trial $N - 1$; NC denotes $P(C_n | N_{n-1})$, the probability of recall on Trial N , conditional on nonrecall on the previous trial.

increased the length of the list to see what would happen to those effects with longer lists.

Method

Apparatus, Design, and Procedure. The apparatus, design, and procedure were essentially the same as for Experiment 1, with the exception that each subject learned a unique list of 40 words over five trials. The total time allowed for recall was therefore 5.3 min.

Subjects. There were 60 subjects obtained by advertisement from within the university. They were paid \$2 for their participation in the experiment. Each experimental session began with a practice list, as in Experiment 1.

Results and Discussion

As in Experiment 1, there was a high correlation between mean recall and mean subjective organization scores ($r = .83$). Therefore, a MANOVA was again performed, using mean recall, mean subjective organization, recall slope, and organization slope as dependent variables.

The mean recall scores for tagged and nontagged conditions are shown in Figure 3, and the mean organization scores are shown in Table 2. Relative to SFR (nontagged), all experimental groups were significantly different [$F(4,51) = 11.24, p < .001$]. A further analysis showed that SFR (tagged) recalled and organized significantly more than SFR (nontagged) [$F(4,54) = 140.45, p < .001$]. Among the nontagged groups, B and M had higher recall and organization scores than SFR [$F(4,51) = 147.20, p < .001$]. All individual variables were significant at the .01 level or better.

As in the previous experiment, the locus of the effect was in the NC data (see Figure 4), with both the tag and the presentation orders dramatically increasing the probability of recalling a word that was

not recalled on the previous trial [$F(1,54) = 11.42, p < .001$].

GENERAL DISCUSSION

Although interactive presentation has previously been used to impede a subject's learning (Mandler, Worden, & Graesser, 1974), the facilitative effects of response-contingent presentation have been harder to obtain. Madigan (Note 2) studied a condition in which, on each trial, two copies of each nonrecalled item were presented, randomly mixed with one copy of each previously recalled item. It was expected that the extra study time for nonrecalled items would facilitate learning, but learning was no faster than in SFR. However, the experimental subjects did not have their recall order preserved in the next presentation. This suggests that it is not increased study of some particular items themselves that is important, but rather study designed to attach these items to a mnemonic structure.

The main result of the present experiments was the demonstration of two distinct ways in which free recall learning can be facilitated by interactive presentation that is contingent upon the subject's prior recall. Relative to SFR, our experimental subjects recalled more words either when the order of presentation reflected their previous order of recall or when they were given explicit presentation cues indicating the words they

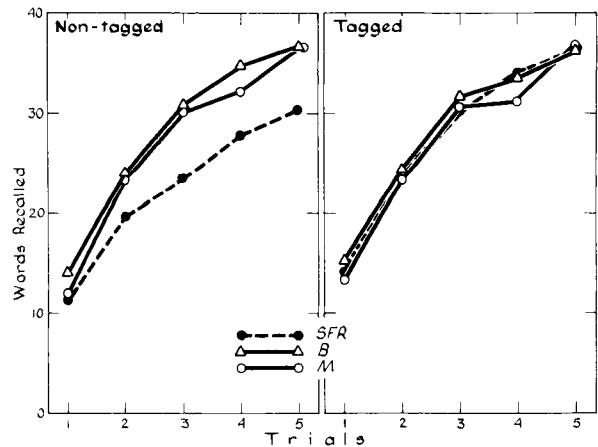


Figure 3. Mean number of words recalled in Experiment 2.

Table 2
Mean Subjective Organization Scores, Experiment 2

| Group | Trial Pair | | | | Mean | Slope |
|--------|------------|------|-------|-------|-------|-------|
| | 1-2 | 2-3 | 3-4 | 4-5 | | |
| SFR | 2.60 | 5.33 | 6.85 | 9.12 | 5.97 | 2.11 |
| Begin | 3.37 | 8.01 | 12.50 | 16.79 | 10.17 | 4.48 |
| Mesh | 2.86 | 5.95 | 9.79 | 13.85 | 8.11 | 3.68 |
| SFR* | 2.82 | 5.48 | 9.83 | 12.55 | 7.67 | 3.35 |
| Begin* | 5.10 | 8.96 | 12.06 | 15.17 | 10.32 | 3.33 |
| Mesh* | 2.35 | 5.56 | 9.27 | 13.04 | 7.55 | 3.58 |

Note—The measure used is the bidirectional ITR for Unit Size 2, expressed as (observed - expected). *See text for explanation.

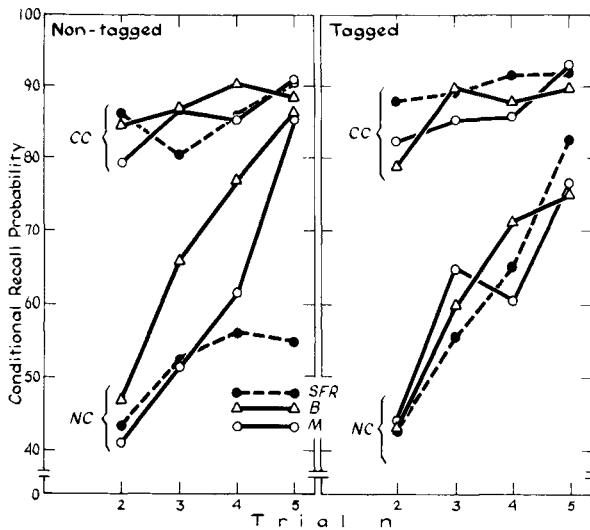


Figure 4. Conditional recall probabilities in Experiment 2. CC and NC are defined as in Figure 2.

had previously failed to remember. Comparison of the three experiments also suggests that the effects of these manipulations increase with length. That is, the effects were negligible for 20-item lists (pilot experiment), moderate for 30-item lists (Experiment 1), and numerically largest for 40-item lists (Experiment 2).

These results are consistent with our original reasoning. Since both manipulations are presumed to aid long-term memory, we would expect them to provide more benefit with longer lists. Of course, these interactive learning procedures may also facilitate recall to a degree due to other characteristics or to the novelty for the subject. Nevertheless, our belief that the dominant effects here reflect configural properties of information in memory is supported by the differences among groups in organization scores, since these are largely parallel to the recall results. Moreover, the major effects in the present study occur as increments to the probability of recalling new items. These findings clearly do not support the position of Murdock et al. (1974), who state that "the characteristics of list presentation seem surprisingly unimportant in affecting learning itself" (p. 528).

Of some interest is the SFR (tagged) condition, in which words were presented in random order from trial to trial, but words not recalled on the previous trial were presented with an asterisk at the left-hand corner on the current trial. This was a condition in which, if one assumes independent storage and retrieval of items, it was possible for subjects to increase the "strength" of an item without increasing consistency in their output order. However, the organization scores for this condition are comparable with the organization scores for the other four experimental conditions. This is consistent with the assumption that increasing

the strength of the item increases the probability of its being integrated in a developing organization structure, thereby facilitating recall.

Still, the interactive conditions used here are not overwhelmingly powerful. While the manipulations were quite successful, and statistically reliable, in increasing both recall and organization scores, these effects are of the same magnitude as many noninteractive manipulations of list presentation. Is this all one can expect from response-contingent training procedures? We think not. An intriguing question, therefore, is what things one might try in order to optimize the subject's learning.

For one thing, these types of interactive conditions use only the information available from the previous recall trial. Yet, if all the information required to predict (and possibly modify) performance on Trial $N + 1$ were available from Trial N data alone, it would be possible to model this performance exactly by a Markov process. Markov models, however, have been shown to be generally inadequate for both basic free recall data (Miller & McGill, 1952) and linguistic structure (Chomsky & Miller, 1958), on which recall organization might be based. We believe, therefore, that a more powerful intervention will require making use of more of the subject's own data.

Second, our interactive order manipulations are somewhat indirect. We have varied aspects of list presentation believed to affect the individual's organization, rather than trying to manipulate that organization directly. However, informed intervention into a subject's developing learning structure requires a precise knowledge of that structure. While subjective organization measures can indicate when an effect has occurred, they are of little use in determining how to produce a stronger one. One possibility would be to construct a spatial mapping (Friendly, 1977) of the subject's organizational structure from the data available from all previous trials. The materials could then be presented in a way that is maximally consistent with this structure. Since presenting categorized materials spatially in the form of a tree-structure diagram has been found to increase recall dramatically (Bower, Clark, Lesgold, & Winzenz, 1969), it can be hoped that spatial presentation of the subject's own learning structure will increase the magnitude of facilitation found in the present experiments.

Finally, our experiments have tested the basic hypothesis of organization theory that an increase in organization will produce an increase in recall. However, the experiments have not provided a test of various models of organization against each other. As Murdock et al. (1974) have pointed out, the ultimate test of such models would be to develop some means of identifying the subjective clusters generated by an individual and then to apply some manipulation designed to

facilitate or impede their formation and growth. In principle, this experiment could be carried out by using a technique (Friendly, 1977, 1979) for discovering a subject's recall groupings through an analysis of the proximities among items in the subject's order of recall. This procedure determines the best fit of a subject's organization to an underlying model of memory structure. The model may be hierarchical, as is often assumed in free recall (Bower et al., 1969; Mandler, 1967, 1968), but organizational structure can also be represented in dimensional and network models using the same technique. This technique implemented interactively may provide the basis for optimizing recall and experimentally contrasting the theoretical usefulness of alternative structures of mnemonic organization.

REFERENCE NOTES

1. Kellerman, E. *An adaptive query system*. Paper presented at the Fourth International APL Users' Conference, Atlanta, Georgia, June 1972.
2. Madigan, S. A. *Limited-capacity retrieval in free recall*. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, November 1973.

REFERENCES

- BOCK, R. D. *Multivariate statistical methods in behavioral research*. New York: McGraw-Hill, 1975.
- BOWER, G. H., CLARK, M. C., LESGOLD, A. M., & WINZENZ, D. Hierarchical retrieval schemes in recall of categorized work lists. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 323-342.
- BUSCHKE, H. Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 543-550.
- CHOMSKY, N., & MILLER, G. A. Finite state languages. *Information and Control*, 1958, 1, 91-112.
- CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 335-359.
- FRIENDLY, M. L. In search of the M-Gram: The structure of organization in free recall. *Cognitive Psychology*, 1977, 9, 188-249.
- FRIENDLY, M. L. Methods for obtaining graphic representations of associative memory structures. In C. R. Puff (Ed.), *Memory organization and structure*. New York: Academic Press, 1979.
- FRIENDLY, M. L., & FRANKLIN, P. Computer control of memory experiments on a large timesharing system. *Behavior Research Methods & Instrumentation*, 1979, 11, 212-217.
- MANDLER, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1). New York: Academic Press, 1967.
- MANDLER, G. Association and organization: Facts, fancies, and theories. In T. R. Horton & D. L. Dixon (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
- MANDLER, G., WORDEN, P. E., & GRAESSER, A. C. Subjective disorganization: Search for the locus of list organization. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 220-235.
- MILLER, G. A., & MCGILL, W. J. A statistical description of verbal learning. *Psychometrika*, 1952, 17, 369-396.
- MURDOCK, B. B., JR., ANDERSON, R. E., & HO, E. Effects of presentation order on learning in multitrial free recall. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 522-529.
- MURDOCK, B. B., JR., PENNEY, C., & AAMIRY, A. Interactive presentation in multitrial free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 679-683.
- PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. Concreteness, imagery and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, 1968, 76(1, Part 2).
- PELLEGRINO, J. W. A general measure of organization in free recall for variable unit size and internal sequence consistency. *Behavior Research Methods & Instrumentation*, 1971, 3, 241-246.
- POSTMAN, L. A pragmatic view of organization theory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- TULVING, E. Subjective organization in free recall of "unrelated" words. *Psychological Review*, 1962, 69, 344-354.
- TULVING, E. Intratrial and intertrial retention: Notes toward a theory of free recall verbal learning. *Psychological Review*, 1964, 71, 219-237.
- TULVING, E. Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 193-197.

(Received for publication August 13, 1979;
revision accepted February 8, 1980.)