Word Fragments as Aids to Recall: the Organization of a Word

Leonard M. Horowitz, Margaret A. White, and Douglas W. Atwood

This paper discusses the organization of a single word. It shows that the beginning of a word is the best cue for eliciting that word; the middle is the poorest cue. Subject was shown a list of words one by one on a memory drum. (Some lists had six-letter words and some had nine-letter words.) Then subject saw a fragment of the word, and he had to recall the entire word. A beginning fragment elicited the correct response most readily and with the shortest latency. The middle elicited the correct response least readily and with the longest latency. These results are also related to the issue of associative symmetry.

The term "organization" has frequently appeared in the literature of verbal behavior, particularly in studies of grammar, associative meaning, and free recall. In general, this literature has examined the organization of words into larger segments of verbal behavior. The present study, though, considers the organization of a single word alone.[1]

Several studies have already hinted that a single word has a characteristic organization. That is, the parts of a word seem to vary in their importance to the word as a whole. Various techniques have shown, for example, that the middle of a word is less important than other parts for identifying that word. Miller and Friedman (1957) asked subjects to reconstruct a deleted letter

1. This research was supported in part by Grant GB-4561 of the National Science Foundation. The authors are indebted to the National Science Foundation for this support. The authors would also like to thank Michelle DeKlyen, William Newman, and Orrin Robinson for their help in collecting and analyzing the data.

from the body of a word. The subject performed best when the middle letter was the letter missing. As the missing letter moved to either end of the word, subject's performance progressively declined. Garner (1957, 1958, 1962), Carson (1961), and Garner and Carson (1960) have explained this result in informational terms: Letters which are adjacent in a word correlate more highly than letters which are one or more characters apart. Therefore, the best predictor of a missing letter in a word is the letter which lies next to it. A letter which is missing from the middle of a word has *two* "best predictors"—one on either side. But a letter which is missing from the end of a word only has one.

Method

The subject of the present study was shown four lists of words; two lists contained six-letter words and two contained nine-letter words. After subject studied a list, he saw a fragment of one of its words and he had to recall the word itself. The fragment was either the beginning, the middle, or the end of the word.

The four lists each contained nine words. Two lists had entirely six-letter words and two had entirely nine-letter words. List 6A contained the following words: SCHOOL, ACCEPT, FUTURE, CIRCLE, LENGTH, MINUTE, PUBLIC, TWELVE, FAMILY. List 6B contained: NATIVE, OBJECT, STRONG, SIMPLE, SECOND, FIGURE, GARDEN, DIVIDE, PERIOD. List 9A contained: RECOGNIZE, ESTABLISH, PASSENGER, PRESIDENT, VEGETABLE, UNIVERSAL, BEAUTIFUL, DISCOVERY, CHARACTER. List 9B contained: COMMUNITY, STRUCTURE, PRINCIPLE, RELIGIOUS, PHYSICIAN, OTHERWISE, NECESSARY, TERRITORY, KNOWLEDGE.

The words were all high-frequency words in the Thorndike-Lorge (1944) word count; their frequencies were 35 or more. There were no repeated consonants occurring consecutively in any of the words. As far as possible, letter combinations were avoided that violated the usual rules of English pronunciation. Words were also avoided that contained contractions or diphthongs. Only the present tense of verbs and the singular form of nouns were used. A six-letter word was not used if its beginning, middle, or final pair of letters made a word, or if the remaining four letters made a word. A nine-letter word was not used if its beginning, middle, or final trigram of letters made a word, or if the remaining six letters made a word.

First, subject was given a practice task. Three words (CAT, TOY, PEN) appeared one at a time on a Lafayette memory drum. Each word lasted two seconds, and a two-second blank space separated one word from the next. After the subject studied this short list, he was tested on an MTA Scholar teaching machine for "aided recall." A fragment of one word of the list appeared, and subject was asked to think of the word as fast as he could. A timer began operating as soon as the word appeared. When subject thought of the word, he pressed a button which stopped the timer and allowed examiner to record his latency. First the subject was tested with the fragment c a -, then with t - y, and then with - e n.

Since redundancy is greatest in the middle of a word, a typographical error in the middle of a word is not very disruptive. Bruner and O'Dowd (1958) tachistoscopically presented words which had an error in the beginning, middle, or end. The subject had to identify the words as fast as he could, and his latencies were measured. The data showed that typographical errors at the beginning of a word are the most disruptive, and those in the middle are the least disruptive.

Because a word has its redundancy in this form, subjects primarily examine the ends of a word during a brief exposure. Haslerud and Clark (1957) presented nine-letter words for brief durations, and the subject reported whatever he saw. In one analysis the authors examined subject's mistakes and determined how often the letter of each position was correctly identified. The subject identified the end letters best, and his performance declined towards the middle. Apparently a subject spends less time looking at the middle since the middle is less informative.

According to Marchbanks and Levin (1965), a similar conclusion holds for kindergarten children who have not yet begun to read. In their study a nonsense syllable was shown to a child and the child was asked to find a syllable like it from a set of alternatives. The child had to select a syllable which resembled the one he had seen. Marchbanks and Levin showed that the child's choice was mainly guided by the first letter, and that the middle letter had the smallest effect. Thus, again, the parts of a word contribute different amounts of importance to the word as a whole.

To speak of a word's "organization" also implies a second property: apparently, a fragment of a word can elicit the whole word faster than it can elicit some other ingredient fragment. Horowitz, Day, Light, and White (1967) studied common words as well as newly learned nonsense words; from their data, a subject who responds to a word-fragment can utter the entire word faster than he can utter a single missing letter.

But the different parts of a word vary in their importance; therefore, some fragments are probably better able than others to elicit the word as a whole. The initial part is the most informative, so it ought to be the best cue. The middle is the least informative, so it should be the poorest. Thus, the beginning of a word should be an excellent hint for helping a subject recall that word, while the middle should be a poor hint. This hypothesis is tested below.

After subject completed the practice task, he watched one experimental list appear on the memory drum. The nine words appeared one at a time. Each word lasted two seconds and a two-second blank space separated one word from the next. After subject studied the entire list, he was tested for aided recall on the teaching machine. He was shown a fragment of one word he had seen and he was asked to supply the entire word. He was allowed as much as 15 seconds for responding. When he responded (or when 15 seconds had elapsed), the next fragment appeared. The fragment was always one-third of the word—two consecutive letters of a six-letter word, or three consecutive letters of a nine-letter word. The fragment was always the initial part of the word, the middle part, or the final part. Blanks appeared in place of the missing letters. For example, the stimulus fragment for testing the recall of "twelve" was "t w - - - -," or "- - e l - -," or "- - - - v e." The fragment for testing the recall of "recognize" was "r e c - - - - - -," or "- - - o g n - - -," "- - - - - - i z e."

Each word was tested by all three fragments, so the test list contained 27 items. The first nine fragments each tested a different word—three with the initial fragment, three with the middle fragment, and three with the final fragment. The next nine items tested the words again but by a different fragment. The last nine items tested them again by the remaining fragment. The examiner recorded subject's latency plus any errors that occurred.

After subject completed one list, he immediately transferred to the next list. He studied each list on the memory drum and then, in the test of aided recall, he reconstructed the fragments into words. The subject worked on all four lists in this way. The order of lists varied from subject to subject; 16 different orders were used. Two subjects were tested with each order, making a total of 32 subjects. They were all students in the introductory psychology class at Stanford University

Results

Table I shows the mean number of items that subject recalled correctly to each type of stimulus fragment. Recall was best when the initial fragment served as the cue and poorest when the middle fragment served as the cue. Corresponding scores of Lists 6A and 6B were summed, and likewise, of Lists 9A and 9B. An analysis of variance was performed on these scores. This analysis showed that the fragment's position in the word was a significant source of variation: $F\ (2,62) = 122.32\ p < .001$. The difference in recall for the initial vs. the final fragment position was more than three times as great as the difference for the final vs. the

TABLE I: *Mean Number of Items Correctly Recalled on Each List (Maximum Score = 9)*

	Fragment Given As Cue					
	Initial		*Middle*		*Final*	
	M	*SD*	*M*	*SD*	*M*	*SD*
List 6A	7.78	0.93	5.78	1.49	6.25	1.30
List 6B	7.44	1.27	5.53	1.64	5.91	1.40
All 6-Letter Words	7.61		5.66		6.08	
List 9A	8.34	1.02	6.69	1.79	7.34	1.34
List 9B	8.56	0.66	6.28	1.28	6.53	1.41
All 9-Letter Words	8.46		6.49		6.94	

middle fragment position. Still, the final fragment-position differed significantly from both of the other two fragment positions. For six-letter words, the *t's* were 2.26 and 8.23. For nine-letter words, the *t's* were 2.45 and 8.15. All *t's* had $df = 62$ and $p < .05$.

Table I further shows that nine-letter words were better recalled than six-letter words. The effect of length was statistically significant; $F(1, 31) = 31.16$, $p < .001$. The interaction between length and position was not significant, $F < 1$, so fragment position seems to have had a consistent effect for words of both lengths.

When subject recalled a word correctly, his latency was recorded. The means of these latencies are reported in Table II. (Items that were not recalled are not included in this analysis.) The pattern of Table II strongly resembles that of Table I. On the average, items were recalled fastest when the initial fragment served as a cue and slowest when the middle fragment served as the cue. An analysis of variance showed that the effect of position was significant, $F(2, 62) = 77.19$, $p < .001$.

As Table II shows, the latency difference between the initial vs. the final fragment position was greater than the difference between the final vs. the middle fragment position. Again, on the average, the final fragment position differed significantly from both of the other two fragment positions. The *t*'s for the six-letter

TABLE II: *Mean Latency (in Seconds) of Items Correctly Recalled*

	Fragment Given As Cue					
	Initial		*Middle*		*Final*	
	M	*SD*	*M*	*SD*	*M*	*SD*
List 6A	1.72	0.65	3.54	1.15	2.75	1.32
List 6B	2.09	0.71	3.36	1.48	2.89	1.08
All 6-Letter Words	1.90		3.45		2.82	
List 9A	1.54	0.38	3.58	1.61	2.21	1.12
List 9B	1.52	0.66	2.69	0.87	3.02	1.54
All 9-Letter Words	1.53		3.14		2.62	

comparisons were 3.70 and 4.97. Those for the nine-letter comparisons were 2.89 and 5.87. All *t's* had $df = 62$ and $p < .01$. There was only one minor deviation from the overall pattern, and that occurred in the latency data of List 9B.

Table II also shows that nine-letter words were supplied faster than six-letter words. From an analysis of variance, $F\ (1, 31) = 12{:}30$, $p < .01$. This very consistent result may be due to the syllabic pattern of nine-letter words: in Lists 9A and 9B, a fragment of a word tended more often to be a syllable of the word. Thus, the syllable may comprise an important substructure within the word's organization. This view has been discussed by Hansen and Rodgers (1965). Two other factors may also explain the superior recall of nine-letter words. First, fewer nine-letter words exist in English, so there is a smaller set of nine-letter words from which subject draws. Second, a fragment of a nine-letter word contained more letters than one of a six-letter word; perhaps the extra length helped recall. The present study does not allow us to choose among these alternatives.

The latency of every subject on each word was also examined to determine which fragment position yielded the fastest recall. Then these data were pooled across subjects for each word. For example, in recalling "family," 20 subjects responded fastest to "f a - - - -"; two responded fastest to "- - m i - -"; and eight responded fastest to "- - - - l y." (Two subjects did not recall the word at all.) These values were then averaged across all 18 words of each length. For the six-letter words, on the average, 18.60 subjects responded fastest to the initial fragment, 3.16 responded fastest to the middle fragment, and 6.96 responded fastest to the final fragment. The corresponding means for the nine-letter words were 20.38, 3.84, and 6.84. The two sets of values are highly comparable and again reflect the pattern of Tables I and II. Their similarity is even clearer when each value is expressed as a proportion of the three-value total. They then tell the proportion of subjects who responded fastest to each type of fragment: For six-letter words, the values were .65, .11, and .24; for nine-letter words, the values were .66, .12, and .22.

Other factors besides position and word-length might also help a fragment elicit the whole word. One correlate, for example, might be the fragment's frequency of occurring as a letter-combination in English. An *unusual* letter combination might elicit the whole word faster. To examine this hypothesis, the frequency in English was examined of each digram of the six-letter words and of each trigram of the nine-letter words. Underwood and Schulz (1960) have reported the frequencies of various digrams and trigrams in English, and each fragment's frequency was converted to a logarithm. Means latencies were also recorded, showing how fast each digram or trigram elicited the entire word. These mean latencies were converted to logarithms, and they were correlated with the log frequencies. Separate correlations were computed for each fragment position and word length. These six correlations were not significant. They ranged from $-.03$ to $+.44$, averaging $+.27$. For all r's, $p > .05$. Thus, a letter combination's frequency in English did not reliably predict how well it aided recall.

The digram frequency of word beginnings in Lists 6A and 6B averaged 942.2; of middle digrams, 969.1; of final digrams, 2,964.0. The corresponding means for the trigrams in Lists 9A and 9B were: 223.7 for initial trigrams; 122.4 for middle trigrams; and 291.8 for final trigrams. Thus, again, the frequency of letter combinations was not consistently related to subject's performance in recall: whereas a word typically ends with a more common letter combination, these end fragments had an intermediate ability to aid recall.

Data of free association. Some letter combinations might also have a greater *chance-probability* of eliciting the word in question. For example, "t w - - - -" might make a subject think of "twelve" without any recent exposure to the word. On the other hand, some fragments might elicit so many different alternative responses that they could interfere with subject's recall.

Therefore, free-association data were collected to determine what words each fragment would elicit. Test booklets were prepared with the various fragments of the experiment. The subject was asked to write all the words he could think of that fit each frame. He was allowed 20 seconds for each fragment. A five-second pause occurred between fragments, and a minute rest occurred after each set of nine fragments.

A different fragment appeared on each page of the booklet, and subject was only tested with one of the fragments of any particular word. Therefore, three different booklets were needed. Each booklet contained 36 items: there were 12 initial fragments, 12 middle fragments, and 12 final fragments. Half of the fragments were from six-letter words and half were from nine-letter words. The order of items in a booklet was systematically varied from one subject to the next. The subjects were 24 students from the class in introductory psychology.

The resulting data were scored in three different ways. One measure reported the mean number of words that a subject produced to any one stimulus. These values are reported in Table III. The pattern of means in Table III resembles the pattern found earlier in the recall data. Responses occurred most readily to an initial fragment and least readily to a middle fragment. Furthermore, the largest differences were between initial position and each of the other positions. This result again shows how very effectively the beginning of a word elicits the whole word: The stimulus "t w - - - -" helped subject think of words of the English language more effectively than the stimulus "- - - - v e." As shown above, though, word *endings* are more common as letter combinations, so "ve" probably occurs in more different words than "tw" does.

TABLE III: *Mean Number of Words Supplied in Free Association to Each Fragment*

	Position of Fragment					
	Initial		*Middle*		*Final*	
	M	*SD*	*M*	*SD*	*M*	*SD*
List 6A	1.47	0.44	0.63	0.48	0.75	0.48
List 6B	1.58	0.77	0.92	0.55	0.65	0.52
All 6-Letter Words	1.52		0.78		0.70	
List 9A	0.96	0.39	0.21	0.30	0.46	0.38
List 9B	1.00	0.42	0.14	0.23	0.39	0.36
All 9-Letter Words	0.98		0.18		0.42	

If a subject responds to an initial fragment more readily, then he is more apt to think of any one particular word. A second measure reported the probability that subject in free association emitted the particular word which had been used in the experiment proper. These probabilities are shown in Table IV. They reveal the same pattern as the means of Table III. The result might be summarized this way: An initial fragment elicits more responses, so subject's recall is more apt to include some particular word.

A third measure told how many *different* words each fragment elicited throughout all the subjects. On the average, the initial fragment of a six-letter word elicited 8.67 different words; the middle fragment, 4.72 words; and the final fragment, 4.89 words. The corresponding values for the nine-letter words were: 4.78, 1.06, and 2.94. These values again reflect the pattern of Tables III and IV.

If a fragment elicits many different responses, more words could have interfered during the recall task of the experiment proper. Therefore, each fragment was given two scores which were then correlated. One score was described above as the number of different words the fragment elicited in free association. The other score told how often the fragment elicited a correct recall in the experiment proper. Six separate r's were computed for each fragment position at each word length. Each r was based

TABLE IV: *Probability of Eliciting List Word in Free Association*

	Position of Fragment					
	Initial		*Middle*		*Final*	
	M	*SD*	*M*	*SD*	*M*	*SD*
List 6A	0.19	0.16	0.01	0.07	0.00	0.00
List 6B	0.17	0.19	0.04	0.11	0.07	0.14
All 6-Letter Words	0.18		0.02		0.04	
List 9A	0.26	0.20	0.06	0.12	0.10	0.15
List 9B	0.33	0.30	0.06	0.16	0.06	0.12
All 9-Letter Words	0.30		0.06		0.08	

on 18 pairs of scores. The values of r for the six-letter words were: $-.41$ for initial fragments, $+.12$ for middle fragments, and $+.17$ for final fragments. The corresponding values for the nine-letter words were: $-.52$, $+.03$, and $-.01$. Only two values, those for the initial fragments, even approached significance: the former just missed significance at the 10% level and the latter reached significance at the 5% level. Thus, a relationship does hold for initial fragments: when there are more responses to compete, recall is poorer. However, the relationship is not very strong.

The number of correct responses on the recall task was also correlated with the response's probability in free association. For the six-letter words, the values of r were: $+.64$ for the initial fragments, $-.04$ for the middle fragments, and $+.03$ for the final fragments. The corresponding r's for the nine-letter words were: $+.52$, $+.11$, and $+.39$. Only the r's for initial fragments were significant; the six-letter value had $p < .05$ and the nine-letter value had $p < .01$. Thus, again, a relationship holds for initial fragments: A word which is likely to occur in free association is also likely to be recalled. Again, though, the relationship is not very strong.

Discussion

These data show that the different parts of a word have different abilities to elicit the whole word. Furthermore, the beginning's advantage over the end is greater than the end's advantage over the middle.

How does this organization come about? In part, it may be a matter of perceptual learning. In learning to read, a child may come to know that the middle of a word is more redundant. Therefore, as one efficient strategy, perhaps he examines the beginning of a word first; if the word is not recognized, he might then look to the end; and if the end is not recognized, he might then examine the middle. Thus, word beginnings would be called upon more often to elicit the whole word, the middles would be called on least often. Brown and McNeill (1966) have offered a similar explanation for recall patterns that operate in the tip-of-the-tongue phenomenon.

This explanation does have one shortcoming, though. It does not explain why kindergarten children, before they learn to read, show the adult pattern in judging the similarity of nonsense syllables: according to Marchbanks and Levin (1965), kindergarten children mainly judge similarity by the beginnings of syllables and least by the middles. Why should this pattern appear *before* the children learn to read? Possibly such habits develop through prereading tasks before formal reading begins.

The present result has two implications for research in verbal learning. The first concerns the formal similarity between words. If two nonsense words only differ in their middle letters, they should seem to be more similar than two nonsense words which only differ in their initial letters. The nonsense words *neglan* and *nebran* should seem more alike than, say, *neglan* and *moglan.*

Second, the present result is related to the issue of associative symmetry. Suppose a subject learns an association between two nonsense syllables, A and B, which are equally available. According to the principle of associative symmetry, A and B elicit each other with equal strength. Data on this issue have been reviewed by Ekstrand (1965). Throughout this literature, forward and backward recall do not always differ *significantly,* but most studies have shown that forward recall is usually better than backward recall, even when item availabilities are equal. Can the principle of associative symmetry be reconciled with this slight advantage to forward recall?

Investigators who hypothesize associative symmetry seem to assume that the associates get merged into a larger whole. Suppose that two associated syllables (A and B) united as a single six-letter unit; call the unit AB. Now "forward recall" may simply mean that A has to elicit AB, so subject can select B as his response. Likewise, "backward recall" may mean that B has to elicit AB so subject can select A as his response. Thus, the forward or backward recall tasks may be measuring whether A or B better elicits AB. If so, forward recall may seem superior only because A elicits the whole more effectively.

Evidence in the literature of gestalt psychology already supports this interpretation. Meyer (1939), replicating an earlier study by Müller and Pilzecker (1900), presented nonsense-syllable triplets to subject. Let us denote the syllables ABC. The subject studied all triplets for 10 trials. Then he was shown one syllable of a triplet and he had to report "whatever came to consciousness." Meyer's data yielded two interesting results: First, Item B was more apt to elicit Item A than it was to elicit Item C. In other words, the "backward" association B-A seemed to be stronger than B-C, the "forward" association. Second, Item C was more apt to elicit Item A than it was to elicit Item B. In other words, the "remote backward" association C-A seemed to be stronger than C-B, the "adjacent backward" association. These results can be used to further refute the chaining hypothesis of serial learning. More interesting, though, they show a tendency for the subject to respond to any ingredient of the triplet with Item A. This tendency, the *initial reproducing tendency* (Müller and Pilzecker, 1900), is perhaps the first step towards recalling the entire ABC complex.

REFERENCES

Brown, R., and McNeill, D. The "tip of the tongue" phenomenon. *J. Verbal Learn. Verbal Behav.*, 1966, *5* 325-337.

Bruner, J. S., and O'Dowd, D. A note on the informativeness of parts of words. *Lang. Speech*, 1958, *1* 98-101.

Carson, D. H. Letter constraints within words in printed English *Kybernetik*. 1961, *1* 46-54.

Ekstrand, B. R. Backward associations. *Psychol. Bull.*, 1966, *65*, 50-64.

Garner, W. R. Symmetric uncertainty analysis and redundancy of printed English. In *Proceedings of the 15th International Congress of Psychology*, Brussels, 1957, 104-110.

Garner, W. R. Symmetric uncertainty analysis and its implications for psychology. *Psychol. Rev.*, 1958, *65*, 183-196.

Garner, W. R. *Uncertainty and structure as psychological concepts*. Wiley: New York, 1962.

Garner, W. R., and Carson, D. H. A multi-variate solution of the redundancy of printed English. *Psychol. Rep.*, 1960, *6*, 123-141.

Hansen, D., and Rodgers, T. S. An exploration of psycholinguistic units in initial reading. Technical Report No. 74, 1965, Stanford University Institute for Mathematical Studies in the Social Sciences.

Haselrud, G. M., and Clark, R. E. On the redintegrative perception of words. *Amer. J. Psychol.*, 1957, *70*, 97-101.

Horowitz, L. M., Day, R. S., Light, L. L., and White, M. A. Availability growth and latent verbal learning. *J. Gen. Psychol.*, 1968, *78*, 65-83.

Marchbanks, G., and Levin, H. Cues by which children recognize words. *J. Educ. Psychol.*, 1965, *56*, 57-61.

Meyer, G. Temporal organization and the initial reproductive tendency. *J. Psychol.*, 1939, *7*, 269-282.

Miller, G.A., and Friedman, E. A. The reconstruction of mutilated English texts. *Inf. Control*, 1957, *1*, 38-55.

Müller, G. E., and Pilzecker, A. Experimentelle Beiträge zur Lehre vom Gedächtnis. *Z. Psychol.*, 1900, Ergbd. *1*.

Thorndike, E. L., and Lorge, I. *The teacher's word book of 30,000 words.* New York: Teachers College, Columbia University, 1944.

Underwood, B. J., and Schulz, R. W. *Meaningfulness and verbal learning.* Lippincott, 1960.