WORD FRAGMENTS AS AIDS TO RECALL: THE
ORGANIZATION OF A WORD ¹

LEONARD M. HOROWITZ, MARGARET A. WHITE,²
AND DOUGLAS W. ATWOOD ³

Stanford University

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. S was shown a list of words 1 by 1
on a memory drum. (Some lists had 6-letter words and some had
9-letter words.) Then S saw a fragment of the word, and he had
to recall the entire word. The fragment was the beginning, or the
middle, or the end of the word. 1 list, for example, contained the
word “recognize,” and S was shown the fragment “rec-----,”
or “ogn--,” or “-----ize.” 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 fre-
quently 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 behav-
ior. The present study, though, consi-
ders the organization of a single word
alone.

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 tech-
niques 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 S's 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 col-
lecting and analyzing the data.
² Now at the University of Michigan.
³ Now at the University of Illinois.
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, Ss primarily examine the ends of a word during a brief exposure. Haslerud and Clark (1957) presented nine-letter words for brief durations, and the S reported whatever he saw. In one analysis the authors examined S's mistakes and determined how often the letter of each position was correctly identified. The S identified the end letters best, and his performance declined towards the middle. Apparently an S 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, an S 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 the 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.

**Method**

The S of the present study was shown four lists of words; two lists contained six-letter words and two contained nine-letter words. After S 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: 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 S was given a practice task. Three words (CAT, TOY, PEN) appeared one at a time on a Lafayette memory drum. Each
word lasted 2 sec., and a 2-sec. blank space separated one word from the next. After the S 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 S was asked to think of the word as fast as he could. A timer began operating as soon as the word appeared. When S thought of the word, he pressed a button which stopped the timer and allowed E to record his latency. First the S was tested with the fragment ca-, then with t-y, and then with -en.

After S 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 2 sec. and a 2-sec. blank space separated one word from the next. After S 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 sec. for responding. When he responded (or when 15 sec. 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 "tw---", or "-el--", or "ve..."

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 E recorded S's latency plus any errors that occurred.

After S 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 S worked on all four lists in this way. The order of lists varied from S to S; 16 different orders were used. Two Ss were tested with each order, making a total of 32 Ss. They were all students in the introductory psychology class at Stanford University.

**RESULTS**

Table 1 shows the mean number of items that S 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 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 1 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
Table 2

<table>
<thead>
<tr>
<th>Fragment Given As Cue</th>
<th>Initial</th>
<th>Middle</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>List 6A</td>
<td>1.72</td>
<td>0.65</td>
<td>3.54</td>
</tr>
<tr>
<td>List 6B</td>
<td>2.09</td>
<td>0.71</td>
<td>3.36</td>
</tr>
<tr>
<td>All 6-Letter Words</td>
<td>1.90</td>
<td>0.38</td>
<td>3.58</td>
</tr>
<tr>
<td>List 9A</td>
<td>1.54</td>
<td>0.38</td>
<td>3.58</td>
</tr>
<tr>
<td>List 9B</td>
<td>1.52</td>
<td>0.66</td>
<td>2.69</td>
</tr>
<tr>
<td>All 9-Letter Words</td>
<td>1.53</td>
<td>3.14</td>
<td></td>
</tr>
</tbody>
</table>

significant, $F < 1$, so fragment position seems to have had a consistent effect for words of both lengths.

When $S$ recalled a word correctly, his latency was recorded. The means of these latencies are reported in Table 2. (Items that were not recalled are not included in this analysis.) The pattern of Table 2 strongly resembles that of Table 1. On the average, items were recalled fastest when the initial fragment served as the 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 2 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 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 2 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 $S$ 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 $S$ on each word was also examined to determine which fragment position yielded the fastest recall. Then these data were pooled across $S$s for each word. For example, in recalling “family,” 20 $S$s responded fastest to “fa - - - -”; 2 responded fastest to “- - m i - -”; and 8 responded fastest to “- - - - - ly.” (Two $S$'s 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 $S$s 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 1 and 2. Their similarity is even clearer when each value is expressed as a proportion of the three-value total. They then tell the proportion of $S$s 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. Mean 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 S’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, “tw-----” 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 S’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 S was asked to write all the words he could think of that fit each frame. He was allowed 20 sec. for each fragment. A 5-sec. 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 S 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 S to the next. The Ss 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 an S produced to any one stimulus. These values are reported in Table 3. The pattern of means in Table 3 resembles the pattern found earlier in the recall data. Responses occurred most readily to an initial
TABLE 3
MEAN NUMBER OF WORDS SUPPLIED IN FREE ASSOCIATION TO EACH FRAGMENT

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Middle</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>List 6A</strong></td>
<td>1.47</td>
<td>0.77</td>
<td>2.53</td>
</tr>
<tr>
<td><strong>List 6B</strong></td>
<td>1.38</td>
<td>0.78</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>All 6-Letter Words</strong></td>
<td>1.52</td>
<td>0.78</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>List 9A</strong></td>
<td>0.56</td>
<td>0.39</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>List 9B</strong></td>
<td>1.00</td>
<td>0.42</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>All 9-Letter Words</strong></td>
<td>0.98</td>
<td>0.18</td>
<td>0.42</td>
</tr>
</tbody>
</table>

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 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.

TABLE 4
PROBABILITY OF ELICITING LIST WORD IN FREE ASSOCIATION

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Middle</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>List 6A</strong></td>
<td>0.19</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>List 6B</strong></td>
<td>0.17</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>All 6-Letter Words</strong></td>
<td>0.18</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>List 9A</strong></td>
<td>0.20</td>
<td>0.06</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>List 9B</strong></td>
<td>0.33</td>
<td>0.06</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>All 9-Letter Words</strong></td>
<td>0.30</td>
<td>0.06</td>
<td>0.08</td>
</tr>
</tbody>
</table>
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, and 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 an $S$ 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 $S$ can select B as his response. Likewise, “backward recall” may mean that B has to elicit AB so $S$ 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 S. Let us denote the syllables ABC. The S 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 & Pilzecker, 1900), is perhaps the first step towards recalling the entire ABC complex.

REFERENCES


(Received December 28, 1966)