JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 5, 325-337 (1966)

The "Tip of the Tongue" Phenomenon

ROGER BROWN AND DAVID MCNEILL[1]

Harvard University, Cambridge, Massachusetts

The "tip of the tongue" (TOT) phenomenon is a state in which one cannot quite recall a familiar word but can recall words of similar form and meaning. Several hundred such states were precipitated by reading to *S*s the definitions of English words of low frequency and asking them to try to recall the words. It was demonstrated that while in the TOT state, and before recall occurred, *S*s had knowledge of some of the letters in the missing word, the number of syllables in it, and the location of the primary stress. The nearer *S* was to successful recall the more accurate the information he possessed. The recall of parts of words and attributes of words is termed "generic recall." The interpretation offered for generic recall involves the assumption that users of a language possess the mental equivalent of a dictionary. The features that figure in generic recall may be entered in the dictionary sooner than other features and so, perhaps, are wired into a more elaborate associative network. These more easily retrieved features of low-frequency words may be the features to which we chiefly attend in word-perception. The features favored by attention, especially the beginnings and endings of words, appear to carry more information than the features that are not favored, in particular the middles of words.

William James wrote, in 1893: "Suppose we try to recall a forgotten name. The state of our consciousness is peculiar. There is a gap therein; but no mere gap. It is a gap that is intensely active. A sort of wraith of the name is in it, beckoning us in a given direction, making us at moments tingle with the sense of our closeness and then letting us sink back without the longed-for term. If wrong names are proposed to us, this singularly definite gap acts immediately so as to negate them. They do not fit into its mould. And the gap of one word does not feel like the gap of another, all empty of content as both might seem necessarily to be when described as gaps" (p. 251).

The "tip of the tongue" (TOT) state involves a failure to recall a word of which one has knowledge. The evidence of knowledge is either an eventually successful recall or else an act of recognition that occurs, without additional training, when recall has failed. The class of cases defined by the conjunction of knowledge and a failure of recall is a large one. The TOT state, which James described, seems to be a small subclass in which recall is felt to be imminent.

For several months we watched for TOT states in ourselves. Unable to recall the name of the street on which a relative lives, one of us thought of *Congress* and *Corinth* and *Concord* and then looked up the address and learned that it was *Cornish.* The words that had come to mind have certain properties in common with the word that had been sought (the "target word"): all four begin with *Co*; all are two-syllable words; all put the primary stress on the first syllable. After this experience we began putting direct questions to ourselves when we fell into the TOT state, questions as to the number of syllables in the target word, its initial letter, etc.

Woodworth (1934), before us, made a record of data for naturally occurring TOT

[1] Now at the University of Michigan.

states and Wenzl (1932, 1936) did the same for German words. Their results are similar to those we obtained and consistent with the following preliminary characterization. When complete recall of a word is not presently possible but is felt to be imminent, one can often correctly recall the general type of the word; *generic* recall may succeed when particular recall fails. There seem to be two common varieties of generic recall. (a) Sometimes a part of the target word is recalled, a letter or two, a syllable, or affix. Partial recall is necessarily also *generic* since the class of words defined by the possession of any *part* of the target word will include words other than the target. (b) Sometimes the abstract form of the target is recalled, perhaps the fact that it was a two-syllable sequence with the primary stress on the first syllable. The whole word is represented in *abstract form recall* but not on the letter-by-letter level that constitutes its identity. The recall of an abstract form is also necessarily *generic,* since any such form defines a class of words extending beyond the target.

Wenzl and Woodworth had worked with small collections of data for naturally occurring TOT states. These data were, for the most part, provided by the investigators; were collected in an unsystematic fashion; and were analyzed in an impressionistic nonquantitative way. It seemed to us that such data left the facts of generic recall in doubt. An occasional correspondence between a retrieved word and a target word with respect to number of syllables, stress pattern or initial letter is, after all, to be expected by chance. Several months of "self-observation and asking-our-friends" yielded fewer than a dozen good cases and we realized that an improved method of data collection was essential.

We thought it might pay to "prospect" for TOT states by reading to *S* definitions of uncommon English words and asking him to supply the words. The procedure was given a preliminary test with nine *S*s who were individually interviewed for 2 hrs each.[2] In 57 instances an *S* was, in fact, "seized" by a TOT state. The signs of it were unmistakable; he would appear to be in mild torment, something like the brink of a sneeze, and if he found the word his relief was considerable. While searching for the target *S* told us all the words that came to his mind. He volunteered the information that some of them resembled the target in sound but not in meaning; others he was sure were similar in meaning but not in sound. The *E* intruded on *S*'s agony with two questions: (a) How many syllables has the target word? (b) What is its first letter? Answers to the first question were correct in 47% of all cases and answers to the second question were correct in 51% of the cases. These outcomes encouraged us to believe that generic recall was real and to devise a group procedure that would further speed up the rate of data collection.

Method

Subjects

Fifty-six Harvard and Radcliffe undergraduates participated in one of three evening sessions; each session was 2 hrs long. The *S*s were volunteers from a large General Education Course and were paid for their time.

Word List. The list consisted of 49 words which, according to the Thorndike-Lorge *Word Book* (1952) occur at least once per four million words but not so often as once per one million words. The level is suggested by these examples: *apse, nepotism, cloaca, ambergris,* and *sampan.* We thought the words used were likely to be in the passive or recognition vocabularies of our *S*s but not in their active recall vocabularies. There were 6 words of 1 syllable; 19 of 2 syllables; 20 of 3 syllables; 4 of 4 syllables. For each word we used a definition from *The American College Dictionary* (Barnhart, 1948) edited so as to contain no words that closely resembled the one being defined.

Response Sheet. The response sheet was laid off in vertical columns headed as follows:

Intended word (+ *One I was thinking of*).
(− *Not*).

[2] We wish to thank Mr. Charles Hollen for doing the pretest interviews.

Number of syllables (1-5).
Initial letter.
Words of similar sound. (1. *Closest in sound*)
(2. *Middle*)
(3. *Farthest in Sound*)
Words of similar meaning.
Word you had in mind if not intended word.

Procedure

We instructed *Ss* to the following effect.

In this experiment we are concerned with that state of mind in which a person is unable to think of a word that he is certain he knows, the state of mind in which a word seems to be on the tip of one's tongue. Our technique for precipitating such states is, in general, to read definitions of uncommon words and ask the subject to recall the word.

(1) We will first read the definition of a low-frequency word.

(2) If you should happen to know the word at once, or think you do, or, if you should simply not know it, then there is nothing further for you to do at the moment. Just wait.

(3) If you are unable to think of the word but feel sure that you know it and that it is on the verge of coming back to you then you are in a TOT state and should begin at once to fill in the columns of the response sheet.

(4) After reading each definition we will ask whether anyone is in the TOT state. Anyone who is in that state should raise his hand. The rest of us will then wait until those in the TOT state have written on the answer sheet all the information they are able to provide.

(5) When everyone who has been in the TOT state has signalled us to proceed, we will read the target word. At this time, everyone is to write the word in the leftmost column of the response sheet. Those of you who have known the word since first its definition was read are asked not to write it until this point. Those of you who simply did not know the word or who had thought of a different word will write now the word we read. For those of you who have been in the TOT state two eventualities are possible. The word read may strike you as definitely the word you have been seeking. In that case please write '+' after the word, as the instructions at the head of the column direct. The other possibility is that you will not be sure whether the word read is the one you have been seeking or, indeed, you may be sure that it is not. In this case you are asked to write the sign '—' after the word. Sometimes when the word read out is not the one you have been seeking your actual target may come to mind. In this case, in addition to the minus sign in the leftmost column, please write the actual target word in the rightmost column.

(6) Now we come to the column entries themselves. The first two entries, the guess as to the number of syllables and the initial letter, are required. The remaining entries should be filled out if possible. When you are in a TOT state, words that are related to the target word do almost always come to mind. List them as they come, but separate words which you think resemble the target in sound from words which you think resemble the target in meaning.

(7) When you have finished all your entries, but before you signal us to read the intended target word, look again at the words you have listed as 'Words of similar sound.' If possible, rank these, as the instructions at the head of the column direct, in terms of the degree of their seeming resemblance to the target. This must be done without knowledge of what the target actually is.

(8) The search procedure of a person in the TOT state will sometimes serve to retrieve the missing word before he has finished filling in the columns and before we read out the word. When this happens please mark the place where it happens with the words "Got it" and *do not provide any more data.*

RESULTS

Classes of Data

There were 360 instances, across all words and all *Ss*, in which a TOT state was signalled. Of this total, 233 were positive TOTs. A positive TOT is one for which the target word is known and, consequently, one for which the data obtained can be scored as accurate or inaccurate. In those cases where the target was not the word intended but some other word which *S* finally recalled and wrote in the rightmost column his data were checked against that word, his effective target. A negative TOT is one for which the *S* judged the word read out not to have been his target and, in addition, one in which *S* proved unable to recall his own functional target.

The data provided by *S* while he searched for the target word are of two kinds: explicit guesses as to the number of syllables in the target and the initial letter of the target; words that came to mind while he searched for the target. The words that came to mind were classified by *S* into 224 words similar in sound to the target (hereafter called "SS" words) and 95 words similar in meaning to the target (hereafter called "SM" words). The *S*'s information about the number of syllables in, and the initial letter of the target may be inferred from correspondences between the target and his SS words as well as directly discovered from his explicit guesses. For his knowledge of the stress pattern of the target and of letters in the target, other than the initial letter, we must rely on the SS words alone since explicit guesses were not required.

To convey a sense of the SS and SM words we offer the following examples. When the target was *sampan* the SS words (not all of them real words) included: *Saipan, Siam, Cheyenne, sarong, sanching,* and *sympoon.* The SM words were: *barge, houseboat,* and *junk.* When the target was *caduceus* the SS words included: *Casadesus, Aeschelus, cephalus,* and *leucosis.* The SM words were: *fasces, Hippocrates, lictor,* and *snake.* The spelling in all cases is *S*'s own.

We will, in this report, use the SM words to provide baseline data against which to evaluate the accuracy of the explicit guesses and of the SS words. The SM words are words produced under the spell of the positive TOT state but judged by *S* to resemble the target in meaning rather than sound. We are quite sure that the SM words are somewhat more like the target than would be a collection of words produced by *S*s with no knowledge of the target. However, the SM words make a better comparative baseline than any other data we collected.

General Problems of Analysis

The data present problems of analysis that are not common in psychology. To begin with, the words of the list did not reliably precipitate TOT states. Of the original 49 words, all but *zither* succeeded at least once; the range was from one success to nine. The *S*s made actual targets of 51 words not on the original list and all but five of these were pursued by one *S* only. Clearly none of the 100 words came even close to precipitating a TOT state in all 56 *S*s. Furthermore, the *S*s varied in their susceptibility to TOT states. There were nine who experienced none at all in a 2-hr period; the largest number experienced in such a period by one *S* was eight. In out data, then, the entries for one word will not usually involve the same *S*s or even the same number of *S*s as the entries for another word. The entries for one *S* need not involve the same words or even the same number of words as the entries for another *S*. Consequently for the tests we shall want to make there are no significance tests that we can be sure are appropriate.

In statistical theory our problem is called the "fragmentary data problem."[3] The best thing to do with fragmentary data is to report them very fully and analyze them in several different ways. Our detailed knowledge of these data suggests that the problems are not serious for, while there is some variation in the pull of words and the susceptibility of *S*s there is not much variation in the quality of the data. The character of the material recalled is much the same from word to word and *S* to *S*.

Number of Syllables

As the main item of evidence that *S* in a TOT state can recall with significant success the number of syllables in a target word he has not yet found we offer Table 1. The entries on the diagonal are instances in which guesses were correct. The order of the means of the explicit guesses is the same as the order of the actual numbers of syllables in the target words. The rank order correlation between the two is 1.0 and such a correlation is significant with a $p < .001$ (one-tailed) even when only five items are correlated. The modes of the guesses correspond exactly with the actual numbers of syllables, for the values one through three; for words of four and five syllables the modes continue to be three.

When all TOTs are combined, the contributions to the total effects of individual *S*s and of individual words are unequal. We have made an analysis in which each word counts but once. This was accomplished by calculating the mean of the guesses made by all *S*s for whom a particular word precipitated a TOT state and taking that mean as the score for that word. The new means calculated with all words equally weighted were, in order: 1.62; 2.30; 2.80; 3.33; and 3.50. These values are close to those of Table 1 and *rho* with the actual numbers of syllables continues to be 1.0.

We also made an analysis in which each *S* counts but once. This was done by calculating the mean of an *S*'s guesses for all words

[3] We wish to thank Professor Frederick Mosteller for discussing the fragmentary data problem with us.

TABLE 1
ACTUAL NUMBERS OF SYLLABLES AND GUESSED NUMBERS FOR ALL TOTs IN THE MAIN EXPERIMENT

	Guessed numbers					No		
Actual numbers	1	2	3	4	5	guess	Mode	Mean
1	9	7	1	0	0	0	1	1.53
2	2	55	22	2	1	5	2	2.33
3	3	19	61	10	1	5	3	2.86
4	0	2	12	6	2	3	3	3.36
5	0	0	3	0	1	1	3	3.50

of one syllable, the mean for all words of two syllables, etc. In comparing the means of guesses for words of different length one can only use those *S*s who made at least one guess for each actual length to be compared. In the present data only words of two syllables and three syllables precipitated enough TOTs to yield a substantial number of such matched scores. There were 21 *S*s who made guesses for both two-syllable and three-syllable words. The simplest way to evaluate the significance of the differences in these guesses is with the Sign Test. In only 6 of 21 matched scores was the mean guess for words of two syllables larger than the mean for words of three syllables. The difference is significant with a $p = .039$ (one-tailed). For actual words that were only one syllable apart in length, *S*s were able to make a significant distinction in the correct direction when the words themselves could not be called to mind.

The 224 SS words and the 95 SM words provide supporting evidence. Words of similar sound (SS) had the same number of syllables as the target in 48% of all cases. This value is close to the 57% that were correct for explicit guesses in the main experiment and still closer to the 47% correct already reported for the pretest. The SM words provide a clear contrast; only 20% matched the number of syllables in the target. We conclude that *S* in a positive TOT state has a significant ability to recall correctly the number of syllables in the word he is trying to retrieve.

In Table 1 it can be seen that the modes of guesses exactly correspond with the actual numbers of syllables in target words for the values one through three. For still longer target words (four and five syllables) the means of guesses continue to rise but the modes stay at the value three. Words of more than three syllables are rare in English and the generic entry for such words may be the same as for words of three syllables; something like "three or more" may be used for all long words.

Initial Letter

Over all positive TOTs, the initial letter of the word *S* was seeking was correctly guessed 57% of the time. The pretest result was 51% correct. The results from the main experiment were analyzed with each word counting just once by entering a word's score as "correct" whenever the most common guess or the only guess was in fact correct; 62% of words were, by this reckoning, correctly guessed. The SS words had initial letters matching the initial letters of the target words in 49% of all cases. We do not know the chance level of success for this performance but with 26 letters and many words that began with uncommon letters the level must be low. Probably the results for the SM words are better than chance and yet the outcome for these words was only 8% matches.

We did an analysis of the SS and SM words, with each *S* counting just once. There were 26 *S*s who had at least one such word. For each *S* we calculated the proportion of SS words matching the target in initial letter and the same proportion for SM words. For 21 *S*s the proportions were not tied and in all but 3 cases the larger value was that of the SS words. The difference is significant by Sign Test with $p = .001$ (one-tailed).

The evidence for significantly accurate generic recall of inital letters is even stronger than for syllables. The absolute levels of success are similar but the chance baseline must be much lower for letters than for syllables because the possibilities are more numerous.

Syllabic Stress

We did not ask *S* to guess the stress pattern of the target word but the SS words provide relevant data. The test was limited to the syllabic location of the primary or heaviest stress for which *The American College Dictionary* was our authority. The number of SS words that could be used was limited by three considerations. (a) Words of one syllable had to be excluded because there was no possibility of variation. (b) Stress locations could only be matched if the SS word had the same number of syllables as the target, and so only such matching words could be used. (c) Invented words and foreign words could not be used because they do not appear in the dictionary. Only 49 SS words remained.

As it happened all of the target words involved (whatever their length) placed the primary stress on either the first or the second syllable. It was possible, therefore, to make a 2 × 2 table for the 49 pairs of target and SS words which would reveal the correspondences and noncorrespondences. As can be seen in Table 2 the SS words tended to stress the same syllable as the target words. The χ^2 for this table is 10.96 and that value is significant with $p < .001$. However, the data do not meet the independence requirement, so we cannot be sure that the matching tendency is significant. There were not enough data to permit any other analyses, and so we are left suspecting that *S* in a TOT state has knowledge of the stress pattern of the target, but we are not sure of it.

TABLE 2
SYLLABLES RECEIVING PRIMARY STRESS IN TARGET WORDS AND SS WORDS

		Target words	
		1st syllable	2nd syllable
SS Words	1st syllable	25	6
	2nd syllable	6	12

Letters in Various Positions

We did not require explicit guesses for letters in positions other than the first, but the SS words provide relevant data. The test was limited to the following positions: first, second, third, third-last, second-last, and last. A target word must have at least six letters in order to provide data on the six positions; it might have any number of letters larger than six and still provide data for the six (relatively defined) positions. Accordingly we included the data for all target words having six or more letters.

Figure 1 displays the percentages of letters in each of six positions of SS words which matched the letters in the same positions of the corresponding targets. For comparison purposes these data are also provided for SM words. The SS curve is at all points above the SM curve; the two are closest together at the third-last position. The values for the last three positions of the SS curve quite closely match the values for the first three positions. The values for the last three positions of the SM curve, on the other hand, are well above the values for the first three

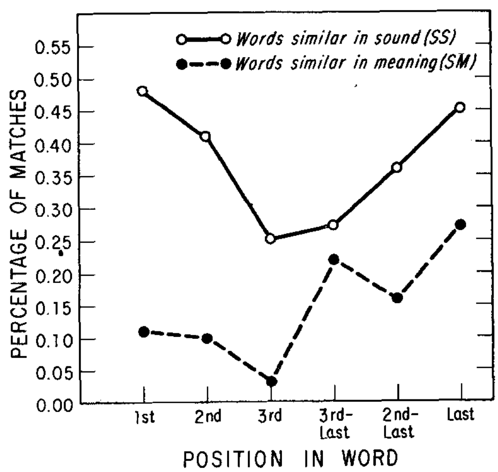

FIG. 1. Percentages of letter matches between target words and SS words for six serial positions.

positions. Consequently the *relative* superiority of the SS curve is greater in the first three positions.

The letter-position data were also analyzed in such a way as to count each target word just once, assigning each position in the target a single score representing the proportion of matches across all *S*s for that position in that word. The order of the SS and SM points is preserved in this finer analysis. We did Sign Tests comparing the SS and SM values for each of the six positions. As Fig. 1 would suggest the SS values for the first three positions all exceeded the SM values with p's less than .01 (one-tailed). The SS values for the final two positions exceeded the SM values with p's less than .05 (one-tailed). The SS values for the third-last position were greater than the SM values but not significantly so.

The cause of the upswing in the final three positions of the SM curve may be some difference in the distribution of information in early and late positions of English words. Probably there is less variety in the later positions. In any case the fact that the SS curve lies above the SM curve for the last three positions indicates that *S* in a TOT state has knowledge of the target in addition to his knowledge of English word structure.

Chunking of Suffixes

The request to *S* that he guess the initial letter of the target occasionally elicited a response of more than one letter; e.g., *ex* in the case of *extort* and *con* in the case of *convene*. This result suggested that some letter (or phoneme) sequences are stored as single entries having been "chunked" by long experience. We made only one test for chunking and that involved three-letter suffixes.

It did not often happen that an *S* produced an SS word that matched the target with respect to all of its three last letters. The question asked of the data was whether such three-letter matches occurred more often when the letters constituted an English suffix than when they did not. In order to determine which of the target words terminated in such a suffix, we entered *The American College Dictionary* with final trigrams. If there was an entry describing a suffix appropriate to the grammatical and semantic properties of the target we considered the trigram to be a suffix. There were 20 words that terminated in a suffix, including *fawning, unctuous,* and *philatelist.*

Of 93 SS words produced in response to a target terminating in a suffix, 30 matched the target in their final three letters. Of 130 SS words supplied in response to a target that did not terminate in a suffix only 5 matched the target in their final three letters. The data were also analyzed in a way that counts each *S* just once and uses only *S*s who produced SS words in response to both kinds of target. A Sign Test was made of the difference between matches of suffixes and matches of endings that were not suffixes; the former were more common with $p = .059$ (one-tailed). A comparable Sign Test for SM words was very far from significance. We conclude that suffix-chunking probably plays a role in generic recall.

Proximity to the Target and Quality of Information

There were three varieties of positive TOT states: (1) Cases in which *S recognized* the word read by *E* as the word he had been seeking; (2) Cases in which *S recalled* the intended word before it was read out; (3) Cases in which *S recalled* the word he had been seeking before *E* read the intended word and the recalled word was not the same as the word read. Since *S* in a TOT state of either type 2 or type 3 reached the target before the intended word was read and *S* in a TOT state of type 1 did not, the TOTs of the second and third types may be considered "nearer" the target than TOTs of the first type. We have no basis for ordering types 2 and 3 relative to one another. We predicted that *S*s in the two kinds of TOT state that ended in recall (types 2 and 3) would produce more accurate information about the target than *S*s in the TOT state that ended in recognition (type 1).

The prediction was tested on the explicit guesses of initial letters since these were the most complete and sensitive data. There were 138 guesses from *S*s in a type 1 state and 58 of these, or 42%, were correct. There were 36 guesses from *S*s in a type 2 state and, of these, 20, or 56%, were correct. There were 59 guesses from *S*s in a type 3 state and of these 39, or 66%, were correct. We also analyzed the results in such a way as to count each word only once. The percentages correct were: for type 1, 50%; type 2, 62%; type 3, 63%. Finally, we performed an analysis counting each *S* just once but averaging together type 2 and type 3 results in order to bring a maximum number of *S*s into the comparison. The combining action is justified since both type 2 and type 3 were states ending in recall. A Sign Test of the differences showed that guesses were more accurate in the states that ended in recall than in the states that ended in recognition; one-tailed $p < .01$. Supplementary analyses with SS and SM words confirmed these results. We conclude that when *S* is nearer his target his generic recall is more accurate than when he is farther from the target.

Special interest attaches to the results from type 2 TOTs. In the method of our experiment there is nothing to guarantee that when *S* said he recognized a word he had really done so. Perhaps when *E* read out a word, *S* could not help thinking that that was the word he had in mind. We ourselves do not believe anything of the sort happened. The single fact that most *S*s claimed fewer than five positive TOTs in a 2-hr period argues against any such effect. Still it is reassuring to have the 36 type 2 cases in which *S* recalled the intended word *before* it was read. The fact that 56% of the guesses of initial letters made in type 2 states were correct is hard-core evidence of generic recall. It may be worth adding that 65% of the guesses of the number of syllables for type 2 cases were correct.

Judgments of the Proximity of SS Words

The several comparisons we have made of SS and SM words demonstrate that when recall is imminent *S* can distinguish among the words that come to mind those that resemble the target in form from those that do not resemble the target in form. There is a second kind of evidence which shows that *S* can tell when he is getting close (or "warm").

In 15 instances *S*s rated two or more SS words for comparative similarity to the target. Our analysis contrasts those rated "most similar" (1) with those rated next most similar (2). Since there were very few words rated (3) we attempted no analysis of them. Similarity points were given for all the features of a word that have now been demonstrated to play a part in generic recall—with the single exception of stress. Stress had to be disregarded because some of the words were invented and their stress patterns were unknown.

The problem was to compare pairs of SS words, rated 1 and 2, for overall similarity to the target. We determined whether each member matched the target in number of syllables. If one did and the other did not, then a single similarity point was assigned the word that matched. For each word, we counted, beginning with the initial letter, the number of consecutive letters in common with the target. The word having the longer sequence that matched the target earned one similarity point. An exactly comparable procedure was followed for sequences starting from the final letter. In sum, each word in a pair could receive from zero to three similarity points.

We made Sign Tests comparing the total scores for words rated most like the target

(1) and words rated next most like the target (2). This test was only slightly inappropriate since only two target words occurred twice in the set of 15 and only one S repeated in the set. Ten of 12 differences were in the predicted direction and the one-tailed $p = .019$. It is of some interest that similarity points awarded on the basis of letters in the middle of the words did not even go in the right direction. Figure 1 has already indicated that they also do not figure in *S*s' judgments of the comparative similarity to the target of pairs of SS words. Our conclusion is that *S* at a given distance from the target can accurately judge which of two words that come to mind is more like the target and that he does so in terms of the features of words that appear in generic recall.

Conclusions

When complete recall of a word has not occurred but is felt to be imminent there is likely to be accurate generic recall. Generic recall of the *abstract form* variety is evidenced by *S*'s knowledge of the number of syllables in the target and of the location of the primary stress. Generic recall of the *partial* variety is evidenced by *S*'s knowledge of letters in the target word. This knowledge shows a bowed serial-position effect since it is better for the ends of a word than for the middle and somewhat better for beginning positions than for final positions. The accuracy of generic recall is greater when *S* is near the target (complete recall is imminent) than when *S* is far from the target. A person experiencing generic recall is able to judge the relative similarity to the target of words that occur to him and these judgments are based on the features of words that figure in partial and abstract form recall.

Discussion

The facts of generic recall are relevant to theories of speech perception, reading, the understanding of sentences, and the organization of memory. We have not worked out all the implications. In this section we first attempt a model of the TOT process and then try to account for the existence of generic memory.

A Model of the Process

Let us suppose (with Katz and Fodor, 1963, and many others) that our long-term memory for words and definitions is organized into the functional equivalent of a dictionary. In real dictionaries, those that are books, entries are ordered alphabetically and bound in place. Such an arrangement is too simple and too inflexible to serve as a model for a mental dictionary. We will suppose that words are entered on keysort cards instead of pages and that the cards are punched for various features of the words entered. With real cards, paper ones, it is possible to retrieve from the total deck any subset punched for a common feature by putting a metal rod through the proper hole. We will suppose that there is in the mind some speedier equivalent of this retrieval technique.

The model will be described in terms of a single example. When the target word was *sextant*, *S*s heard the definition: "A navigational instrument used in measuring angular distances, especially the altitude of sun, moon, and stars at sea." This definition precipitated a TOT state in 9 *S*s of the total 56. The SM words included: *astrolabe, compass, dividers,* and *protractor*. The SS words included: *secant, sextet,* and *sexton*.

The problem begins with a definition rather than a word and so *S* must enter his dictionary backwards, or in a way that would be backwards and quite impossible for the dictionary that is a book. It is not impossible with keysort cards, providing we suppose that the cards are punched for some set of semantic features. Perhaps these are the semantic "markers" that Katz and Fodor (1963) postulate in their account of the comprehension of sentences. We will imagine that it is somehow possible to extract from the definition a set of markers and that these are, in the present case: "navigation, instrument, having to do with geometry." Metal rods thrust into the holes for each of these features might fish up such a collection of entries as: *astrolabe, compass, dividers,* and *protractor*. This first retrieval, which is in response to the definition, must be semantically based and it will not, therefore, account for the appearance of such SS words as *sextet* and *sexton*.

There are four major kinds of outcome of the first retrieval and these outcomes correspond with the four main things that happen to *S*s in the TOT experiment. We will assume that a definition of each word retrieved is entered on its card and that it is possible to check the input definition against those on the cards. The first possible outcome is that *sextant* is retrieved along with *compass* and *astrolabe* and the others and that the definitions are specific enough so that the one entered for *sextant* registers as matching the input and all the others as not-matching. This is the case of correct recall; *S* has found a word that matches the definition and it is the intended word. The second possibility is that *sextant* is not among the words retrieved and, in addition, the definitions entered for those retrieved are so imprecise that one of them (the definition for *compass*, for example) registers as matching the input. In this case *S* thinks he has found the target though he really has not. The third possibility is that *sextant* is not among the words retrieved, but the definitions entered for those retrieved are specific enough so that none of them will register a match with the input. In this case, *S* does not know the word and realizes the fact. The above three outcomes are the common ones and none of them represents a TOT state.

In the TOT case the first retrieval must include a card with the definition of *sextant* entered on it but with the word itself incompletely entered. The card might, for instance, have the following information about the word: two-syllables, initial s, final t. The entry would be a punchcard equivalent of S__ __T. Perhaps an incomplete entry of this sort is James's "singularly definite gap" and the basis for generic recall.

The *S* with a correct definition, matching the input, and an incomplete word entry will know that he knows the word, will feel that he almost has it, that it is on the tip of his tongue. If he is asked to guess the number of syllables and the initial letter he should, in the case we have imagined, be able to do so. He should also be able to produce SS words. The features that appear in the incomplete entry (two-syllables, initial s, and final t) can be used as the basis for a second retrieval. The subset of cards defined by the intersection of all three features would include cards for *secant* and *sextet*. If one feature were not used then *sexton* would be added to the set.

Which of the facts about the TOT state can now be accounted for? We know that *S*s were able, when they had not recalled a target, to distinguish between words resembling the target in sound (SS words) and words resembling the target in meaning only (SM words). The basis for this distinction in the model would seem to be the distinction between the first and second retrievals. Membership in the first subset retrieved defines SM words and membership in the second subset defines SS words.

We know that when *S* had produced several SS words but had not recalled the target he could sometimes accurately rank-order the SS words for similarity to the target. The model offers an account of this ranking performance. If the incomplete entry for *sextant* includes three features of the word then SS words having only one or two of these features (e.g., *sexton*) should be judged less similar to the target than SS words having all three of them (e.g., *secant*).

When an SS word has all of the features of the incomplete entry (as do *secant* and *sextet* in our example) what prevents its being mistaken for the target? Why did not the *S* who produced *sextet* think that the word was "right?" Because of the definitions. The forms meet all the requirements of the incomplete entry but the definitions do not match.

The TOT state often ended in recognition; i.e., *S* failed to recall the word but when *E* read out *sextant* *S* recognized it as the word he had been seeking. The model accounts for this outcome as follows. Suppose that there is only the incomplete entry S__ __T in memory, plus the definition. The *E* now says (in effect) that there exists a word *sextant* which has the definition in question. The word *sextant* then satisfies all the data points available to *S*; it has the right number of syllables, the right initial letter, the right final letter, and it is said to have the right definition. The result is recognition.

The proposed account has some testable implications. Suppose that *E* were to read out, when recall failed, not the correct word *sextant* but an invented word like *sekrant* or *saktint* which satisfies the incomplete entry as well as does *sextant* itself. If *S* had nothing but the incomplete entry and *E*'s testimony to guide him then he should "recognize" the invented words just as he recognizes *sextant*.

The account we have given does not accord with intuition. Our intuitive notion of recognition is that the features which could not be called were actually in storage but less accessible than the features that were recalled. To stay with our example, intuition suggests that the features of *sextant* that could not be recalled, the letters between the first and the last, were entered on the card but were less "legible" than the recalled features. We might imagine them printed in small letters and faintly. When, however, the *E* reads out the word *sextant*, then *S* can make out the less legible parts of his entry and, since the total entry matches *E*'s word, *S* recognizes it. This sort of recognition should be "tighter" than the one described previously. *Sekrant* and *saktint* would be rejected.

We did not try the effect of invented words and we do not know how they would have been received but among the outcomes of the actual experiment there is one that strongly favors the faint-entry theory. Subjects in a TOT state, after all, sometimes recalled the target word without any prompting. The incomplete entry theory does not admit of such a possibility. If we suppose that the entry is not S____T but something more like S*ex tan*T (with the italicized lower-case letters representing the faint-entry section) we must still explain how it happens that the faintly entered, and at first inaccessible, middle letters are made accessible in the case of recall.

Perhaps it works something like this. The features that are first recalled operate as we have suggested, to retrieve a set of SS words. Whenever an SS word (such as *secant*) includes middle letters that are matched in the faintly entered section of the target then those faintly entered letters become accessible. The match brings out the missing parts the way heat brings out anything written in lemon juice. In other words, when *secant* is retrieved the target entry grows from S*ex tan*T to SE*x t*ANT. The retrieval of *sextet* brings out the remaining letters and *S* recalls the complete word—*sextant*.

It is now possible to explain the one as yet unexplained outcome of the TOT experiment. Subjects whose state ended in recall had, before they found the target, more correct information about it than did *S*s whose state ended in recognition. More correct information means fewer features to be brought out by duplication in SS words and so should mean a greater likelihood that all essential features will be brought out in a short period of time.

All of the above assumes that each word is entered in memory just once, on a single card. There is another possibility. Suppose that there are entries for *sextant* on several different cards. They might all be incomplete, but at different points, or, some might be incomplete and one or more of them complete. The several cards would be punched for different semantic markers and perhaps for different associations so that the entry recovered would vary with the rule of retrieval. With this conception we do not require the notion of faint entry. The difference between features commonly recalled, such as the first and last letters, and features that are recalled with difficulty or perhaps only recognized, can be rendered in another way. The more accessible features are entered on more cards or else the cards on which they appear are punched for more markers; in effect, they are wired into a more extended associative net.

The Reason for Generic Recall

In adult minds words are stored in both visual and auditory terms and between the two there are complicated rules of translation. Generic recall involves letters (or phonemes), affixes, syllables, and stress location. In this section we will discuss only letters (legible forms) and will attempt to explain a single effect—the serial position effect in the recall of letters. It is not clear how far the explanation can be extended.

In brief overview this is the argument. The design of the English language is such that one word is usually distinguished from all others in a more-than-minimal way, i.e., by more than a single letter in a single position. It is consequently *possible* to recognize words when one has not stored the complete letter sequence. The evidence is that we do not store the complete sequence if we do not have to. We begin by attending chiefly to initial and final letters and storing these. The order of attention and of storage favors the ends of words because the ends carry more information than the middles. An incomplete entry will serve for recognition, but if words are to be produced (or recalled) they must be stored in full. For most words, then, it is eventually necessary to attend to the middle letters. Since end letters have been attended to from the first they should always be more clearly entered or more elaborately connected than middle letters. When recall is required, of words that are not very familiar to *S*, as it was in our experiment, the end letters should often be accessible when the middle are not.

In building pronounceable sequences the English language, like all other languages, utilizes only a small fraction of its combinatorial possibilities (Hockett, 1958). If a language used all possible sequences of phonemes (or letters) its words could be shorter, but they would be much more vulnerable to misconstruction. A change of any single letter would result in reception of a different word. As matters are actually arranged, most changes result in no word at all; for example: *textant, sixtant, sektant.* Our words are highly redundant and fairly indestructible.

Underwood (1963) has made a distinction for the learning of nonsense syllables between the "nominal" stimulus which is the syllable presented and the "functional" stimulus which is the set of characteristics of the syllable actually used to cue the response. Underwood reviews evidence showing that college students learning paired-associates do not learn any more of a stimulus trigram than they have to. If, for instance, each of a set of stimulus trigrams has a different initial letter, then *S*s are not likely to learn letters other than the first, since they do not need them.

Feigenbaum (1963) has written a computer program (EPAM) which simulates the selective-attention aspect of verbal learning as well as many other aspects. ". . . EPAM has a *noticing order for letters of syllables,* which prescribes at any moment a letter-scanning sequence for the matching process. Because it is observed that subjects generally consider end letters before middle letters, the noticing order is initialized as follows: first letter, third letter, second letter" (p. 304). We believe that the differential recall of letters in various positions, revealed in Fig. 1 of this paper, is to be explained by the operation in the perception of real words of a rule very much like Feigenbaum's.

Feigenbaum's EPAM is so written as to make it possible for the noticing rule to be changed by experience. If the middle position were consistently the position that differentiated syllables, the computer would learn to look there first. We suggest that the human tendency to look first at the beginning of a word, then at the end and finally the middle has "grown" in response to the distribution of information in words. Miller and Friedman (1957) asked English speakers to guess letters for various open positions in segments of English text that were 5, 7, or 11 characters long. The percentages of correct first guesses show a very clear serial position effect for segments of all three lengths. Success was lowest in the early positions, next lowest in the final positions, and at a maximum in the middle positions. Therefore, information was greatest at the start of a word, next greatest at the end, and least in the middle. Attention needs to be turned where information is, to the parts of the word that cannot be guessed. The Miller and Friedman segments did not necessarily break at word boundaries but their discovery that the middle positions of continuous text are more easily guessed than the ends applies to words.

Is there any evidence that speakers of English do attend first to the ends of English words? There is no evidence that the eye fixations of adult readers consistently favor particular parts of words (Woodworth and Schlosberg, 1954). However, it is not eye fixation that we have in mind. A considerable stretch of text can be taken in from a single fixation point. We are suggesting that there is selection within this stretch, selection accomplished centrally; perhaps by a mechanism like Broadbent's (1958) "biased filter."

Bruner and O'Dowd (1958) studied word perception with tachistoscopic exposures too brief to permit more than one fixation. In each word presented there was a single reversal of two letters and the *S* knew this. His task was to identify the *actual* English word responding as quickly as possible. When the *actual* word was AVIATION, *S*s were presented with one of the following: VAIATION, AVITAION, AVIATINO. Identification of the actual word as AVIATION was best when *S* saw AVITAION, next best when he saw AVIATINO, and most difficult when he saw VAIATION. In general, a reversal of the two initial letters made identification most difficult, reversal of the last two letters made it somewhat less difficult, reversal in the middle made least difficulty. This is what should happen if words are first scanned initially, then finally, then medially. But the scanning cannot be a matter of eye movements; it must be more central.

Selective attention to the ends of words should lead to the entry of these parts into the mental dictionary, in advance of the middle parts. However, we ordinarily need to know more than the ends of words. Underwood has pointed out (1963), in connection with paired-associate learning, that while partial knowledge may be enough for a stimulus syllable which need only be recognized it will not suffice for a response item which must be produced. The case is similar for natural language. In order to speak one must know all of a word. However, the words of the present study were low-frequency words, words likely to be in the passive or recognition vocabularies of the college-student *S*s but not in their active vocabularies; stimulus items, in effect, rather than response items. If knowledge of the parts of new words begins at the ends and moves toward the middle we might expect a word like *numismatics,* which was on our list, to be still registered as NUM___ ___ICS. Reduced entries of this sort would in many contexts serve to retrieve the definition.

The argument is reinforced by a well-known effect in spelling. Jensen (1962) has analyzed thousands of spelling errors for words of 7, 9, or 11 letters made by children in the eighth and tenth grades and by junior college freshmen. A striking serial position effect appears in all his sets of data such that errors are most common in the middle of the word, next most common at the end, and least common at the start. These results are as they should be if the order of attention and entry of information is first, last, and then, middle. Jensen's results show us what happens when children are forced to produce words that are still on the recognition level. His results remind us of those bluebooks in which students who are uncertain of the spelling of a word write the first and last letters with great clarity and fill in the middle with indecipherable squiggles. That is what should happen when a word that can be only partially recalled must be produced in its entirety. End letters and a stretch of squiggles may, however, be quite adequate for recognition purposes. In the TOT experiment we have perhaps placed adult *S*s in a situation comparable to that created for children by Jensen's spelling tests.

There are two points to clarify and the argument is finished. The *S*s in our experiment were college students, and so in order to obtain words on the margin of knowledge we had to use words that are very infrequent in English as a whole. It is not our thought, however, that the TOT phenomenon occurs only with rare words. The absolute location of the margin of word knowledge is a function of *S*'s age and education, and so with other *S*s we would expect to obtain TOT states for words more frequent in English. Finally the need to produce (or recall) a word is not the only factor that is likely to encourage registration of its middle letters. The amount of detail needed to specify a word uniquely must increase with the total number of words known, the number from which any one is to be distinguished. Consequently the growth of vocabulary, as well as the need to recall, should have some power to force attention into the middle of a word.

References

BARNHART, C. L. (Ed.) *The American college dictionary*. New York: Harper, 1948.

BROADBENT, D. E. *Perception and communication.* New York: Macmillan, 1958.

BRUNER, J. S., AND O'DOWD, D. A note on the informativeness of words. *Language and Speech,* 1958, **1**, 98-101.

FEIGENBAUM, E. A. The simulation of verbal learning behavior. In E. A. Feigenbaum and J. Feldman (Eds.) *Computers and thought.* New York: McGraw-Hill, 1963. Pp. 297-309.

HOCKETT, C. F. *A course in modern linguistics.* New York: Macmillan, 1958.

JAMES, W. *The principles of psychology,* Vol. I. New York: Holt, 1893.

JENSEN, A. R. Spelling errors and the serial-position effect. *J. educ. Psychol.,* 1962, **53**, 105-109.

KATZ, J. J., AND FODOR, J. A. The structure of a semantic theory. *Language,* 1963, **39**, 170-210.

MILLER, G. A., AND FRIEDMAN, ELIZABETH A. The reconstruction of mutilated English texts. *Inform. Control,* 1957, **1**, 38-55.

THORNDIKE, E. L., AND LORGE, I. *The teacher's word book of 30,000 words.* New York: Columbia Univer., 1952.

UNDERWOOD, B. J. Stimulus selection in verbal learning. In C. N. Cofer and B. S. Musgrave (Eds.) *Verbal behavior and learning: problems and processes.* New York: McGraw-Hill, 1963. Pp. 33-48.

WENZL, A. Empirische und theoretische Beiträge zur Erinnerungsarbeit bei erschwerter Wortfindung. *Arch. ges. Psychol.,* 1932, **85**, 181-218.

WENZL, A. Empirische und theoretische Beiträge zur Erinnerungsarbeit bei erschwerter Wortfindung. *Arch. ges. Psychol.,* 1936, **97**, 294-318.

WOODWORTH, R. S. *Psychology.* (3rd ed.). New York: Holt, 1934.

WOODWORTH, R. S., AND SCHLOSBERG, H. *Experimental psychology.* (Rev. ed.). New York: Holt, 1954.

(Received January 15, 1965)