JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 14, 392–397 (1975)

Within Word Structure in the Tip-of-the-Tongue Phenomenon

DAVID C. RUBIN

Lawrence University

Definitions of four rare words were read to 259 undergraduates. Those subjects who were in the tip-of-the-tongue (TOT) state recorded all the letters they knew. The within-word structure of the resulting 101 partial recalls was indistinguishable from that of similar sounding words from earlier studies. In both sets of data, morpheme-like clusters of letters were evident. The recall of high frequency clusters at the end of words could not be explained in terms of sophisticated guessing. The results support a distinct memory system for word names which is organized for use in the production and perception of speech and writing.

A memory for words, and not only for the concepts named by those words, is essential in the production and perception of speech and writing. The existence of distinct word-name and word-concept memories is supported by clinical studies (e.g., Geshwind, 1972; Howes, 1967), cases in which normal subjects cannot remember a word-name given the word-concept (e.g., Brown & McNeill, 1966), reaction time studies (Loftus & Cole, 1974), as well as by any evidence that requires the positing of linguistic levels more abstract than the word.

Loftus has conducted extensive research (e.g., Loftus & Cole, 1974) on the relationship between word-name memory and memory for the concepts named by those words. Her model is called the dictionary-network model, since its two major components are a dictionary (which contains word names and other lexical items of the language) and a network (which contains the concepts named by those lexical items). The model is silent on two issues: (1) What constitutes a lexical item? or, What is stored in the dictionary? and (2) How are the items in the dictionary organized? These issues are similar to the larger issues concerning any word-name memory. For example, are word-names stored in a unitary fashion or as collections of information which can be individually accessed where relevant? Are word-names stored as inseparable collections of letters, or as strings of subunits such as syllables or morphemes which could be recombined? Fortunately, the tip-of-the-tongue (TOT) phenomenon provides a rich source of data which can be used to study these aspects of word-name memory.

The TOT phenomenon has been brought under laboratory control by asking subjects for the name of a word given its definition (Brown & McNeill, 1966). Subjects are classified as being in the TOT state if they cannot quite retrieve a target-word, but feel its recall is imminent. Such subjects often know the target word's initial letter and length in syllables (Barton, 1971; Brown & McNeill, 1966; Yarmey, 1973). They are also more likely to know letters near the beginning or end of the target-word, tend to retrieve suffixes as units, and tend to know the syllabic stress pattern (Brown & McNeill, 1966).

Findings from TOT studies are limited in two major respects. First, they are all based on analyses that group data from different target-words and in this way lose any possible structure present in the individual target-words. Thus, for example, there is no way to tell whether syllable boundaries are respected, or if the memory consists of whole unanalyzed

I wish to thank Thomas Baker, Roger Brown, Reid Hastie, Elizabeth Loftus, Donald Olivier, Wayne Redenbarger, and Eric Wanner for their advice and encouragement. The data analyzed in Experiment II were generously provided by Roger Brown.

words as opposed to word stems plus inflections. The grouping of data could be avoided, however, if enough subjects participated to provide sufficient data on each target-word.

The second limitation is that the findings are all based either on answers to specific questions or on the analysis of similar sounding words produced by the subjects. For example, we know that subjects in the TOT state can guess the first letter of the target-word just because they were asked to guess at the first letter. Thus, we have no good way of assessing from the data whether subjects actually have knowledge of just the first letter, or knowledge of the first phoneme or the first syllable. We could try to answer these questions from an analysis of the similar sounding words, but these words may be subject to many confounding influences. A study intended to overcome these limitations was run in a large introductory course as a demonstration of the TOT phenomenon.

Experiment I

Method

Subjects. Two hundred fifty-nine Harvard and Radcliffe undergraduates performed the task as part of a classroom demonstration.

Material. Four definitions which produced the most tip-of-the-tongue states in an earlier unpublished study by Brown were chosen. They were presented in the following order:

1. A small boat, not the junk, used in the river and harbor traffic of China and Japan. (*sampan*)
2. One who collects postage stamps. (*philatelist*)
3. The first name of the character Scrooge in Dickens' *Christmas Carol*. (*Ebenezer*)
4. A secretion from the sperm whale used in the manufacture of perfume. (*ambergris*)

Procedure. The procedure used was a modification of Brown and McNeill's (1966) procedure. Subjects were read the definition of a low-frequency word. They were then asked to "Write down any letters that you know. Please feel free to guess. Please put in one dash for each letter you think is in the word, but that you do not know." After this, subjects indicated the number of syllables they thought were in the word and any similar sounding words that came to mind. The target-word was then read. Only subjects who indicated that their intended word was the target-word were scored. The procedure was repeated for all four words.

Results

A total of 101 tip-of-the-tongue states were obtained: 21 from *sampan*, 41 from *philatelist*, 23 from *Ebenezer*, and 16 from *ambergris*. The scoring of the letters guessed correctly resulted in the same general findings as Brown and McNeill's scoring of similar sounding words. In particular, there was a high percentage of first letters guessed correctly and there was a greater probability of letters at the beginning and end of words being recalled. The guesses at the number of syllables in the target words also replicated Brown and McNeill's (1966) findings: Two-thirds of the 58 guesses were correct.

Table 1 presents the transition frequencies of a letter being guessed correctly, calculated from both the left and the right of each target-word. That is, a letter was scored as correct only if the letters before it (in the case of left) of the letters after it (in the case of right) were scored as correct. For example, in the case of *ambergris* six subjects guessed an initial *a*, and of those six subjects only one guessed an *m* following that *a*. Similarly, three subjects guessed the final *s*, and none of them guessed an *i* before that *s*. This method of scoring, while missing a few of what could be considered to be correct responses, simplifies the scoring considerably. It also provides an analysis which is especially sensitive to initial and final letter clusters. Instances of clusters occurring in the middle of words would be likely to be lost, but none were noted in the raw data.

TABLE 1

NUMBER OF CORRECT LETTERS

Number of subjects guessing any letters	Direction of scoring	Target-word
9		*a m b e r g r i s*
	left	6 1 0 0 0 0 0 0 0
	right	0 0 0 0 0 0 0 0 3
14		*E b e n e z e r*
	left	10 4 1 1 0 0 0 0
	right	0 0 0 0 0 0 4 4
37		*p h i l a t e l i s t*
	left	27 24 16 16 2 0 0 0 0 0 0
	right	0 0 0 0 0 0 0 0 18 18 18
13		*s a m p a n*
	left	7 4 2 1 0 0
	right	0 1 1 5 6 6

Table 1 shows evidence of several clusters including possibly the *er* of *Ebenezer*, the *phil* and *ist* of philatelist and the *pan* of *sampan*. Thus examining the data from individual target words does reveal internal structure in generic recalls. Discussion of the nature of this internal structure, however, requires more data.

Because of the large number of subjects needed to obtain generic recalls, similar sounding words from two previous studies of the TOT phenomenon were reanalyzed. It has been noted that possible artifacts might have arisen from allowing subjects to demonstrate their knowledge of the target word only through the use of similar sounding words. The magnitude of any such artifacts will be evaluated by comparing the analysis of such words with the write-any-letter-you-know method used in Experiment I.

EXPERIMENT II

Method

Subjects. Approximately 300 Harvard and Radcliffe undergraduates participated in two separate studies. Fifty-six were in Brown and McNeill's (1966) study. The remaining subjects were in a later unpublished study by Brown.

Materials and procedure. Subjects were read the definition of a low-frequency word. Those who were unable to name the word, but felt that they were on the verge of doing so filled in a response sheet including columns for information on the number of syllables, initial letter, words of similar sound, or words of similar meaning. After this was completed, subjects were informed of the correct word and asked to indicate if it was their intended word. Only subjects who answered in the affirmative were considered in this analysis. The procedure was repeated for 49 definitions in the published study and 12 in the unpublished study. A more detailed description of the procedure appears in Brown and McNeill (1966).

Results

Table 2 presents the transition frequencies of a letter in a similar sounding word matching the corresponding letter in its target-word. To avoid having frequently occurring similar sounding words bias the results, each distinct similar sounding word is counted only once, regardless of how many instances of that word were observed. The nine target-words which produced seven or more matches to one letter are listed. The target-word *Ebenezer* is also indicated to allow comparisons with Table 1.

Examination of the four words common to Tables 1 and 2 reveals substantial similarity. A quantitative measure of this similarity was obtained by correlating the number of matches for each letter (i.e., both left and right combined). The average for the four words was .85 (*ambergris*, .85, $p < .01$; *Ebenezer*, .93, $p < .01$; *philatelist*, .81, $p < .01$; *sampan*, .76, $p < .05$). It therefore appears that similar sounding words do provide a valid measure which can be used to examine the structure of generic recall.

TABLE 2

NUMBER OF CORRECT LETTERS

Number of distinct words	Direction of scoring	Target-word
11		*a l t i m e t e r*
	left	2 2 2 1 0 0 0 0 0
	right	0 0 0 0 9 9 9 9 9
19		*a m b e r g r i s*
	left	13 6 3 2 2 0 0 0 0
	right	0 0 0 0 0 0 0 0 2
21		*c a d u c e u s*
	left	3 2 1 0 0 0 0 0
	right	0 0 0 0 0 0 9 12
25		*E b e n e z e r*
	left	5 1 0 0 0 0 0 0
	right	0 0 0 0 0 0 3 3
22		*e m e t i c*
	left	7 2 1 1 0 0
	right	2 2 2 3 8 9
20		*m i s t r a l*
	left	10 5 2 1 0 0 0
	right	0 0 0 0 0 1 2
13		*p h i l a t e l i s t*
	left	5 5 5 5 2 1 1 1 0 0 0
	right	0 0 0 0 0 0 0 0 7 8 8
27		*s a m p a n*
	left	10 6 1 1 0 0
	right	0 1 2 13 14 16
7		*u s u r e r*
	left	5 4 1 1 0 0
	right	0 0 0 1 7 7
19		*v a q u e r o*
	left	5 5 0 0 0 0 0
	right	0 0 0 0 8 11 13

It is clear from Table 2 that several groups of letters appear to be retrieved as clusters. That is, these groups of letters are retrieved more often than their neighbors, and are all retrieved with about equal frequency. To quantify this notion of a cluster the following two criteria were adopted: (1) Letters in a cluster must be retrieved with a frequency greater than twice that of the surrounding letters, and (2) all letters in a cluster must be retrieved with at least two-thirds the frequency of the most often retrieved letter. Table 3 shows the clusters that are found by adopting these criteria.

TABLE 3

CLUSTERS FROM SIMILAR SOUNDING WORDS

Word	Initial cluster	Final cluster
altimeter	—	meter
ambergris	—	—
caduceus	—	us
Ebenezer	—	er
emetic	—	ic
mistral	—	—
philatelist	phil	ist
sampan	—	pan
usurer	us	er
vaquero	va	ero

Most of the clusters in Table 3 are common combinations of letters for the initial or final portions of words. There is no evidence that these clusters are syllables. According to the phonetic breaks in *Webster's New Collegiate Dictionary* (1963) five clusters were syllables and six were not. Using the printer's syllabic divisions of words in that dictionary, four clusters were syllables and seven were not. There does, however, tend to be a large number of clusters that are morphemes, specifically the *meter* of *altimeter*, the *us* of *caduceus*, the *ic* of *emetic*, the *phil* and *ist* of *philatelist*, the *er* of *usurer*, and the *ero* of *vaquero*. There are also several clusters which would function as morphemes in different contexts, specifically the *er* of *Ebenezer*, the *pan* of *sampan*, and the *us* of *usurer*. One rule can account for most of the observed clustering. If there is an initial or final morpheme, it is retrieved, otherwise only one or two letters are retrieved. Under this explanation, the scarcity of initial morpheme clusters is due to the fact that the initial morpheme clusters are usually word stems

and their retrieval would often lead to the retrieval of the whole word, and thus no TOT state (Peretti, 1974).

The high redundancy–low information of the final clusters might lead to the speculation that these clusters are not really known, but are only guessed. As subjects were given the definitions of the words this is not so unlikely. In order to demonstrate that this is not, in fact, the case, the endings of each distinct similar sounding word given to each of the three definitions that ask for "a person who . . ." were tabulated. Data from all three TOT studies were combined to form Table 4.

TABLE 4

FINAL CLUSTERS FROM TARGET-WORDS DEFINING "ONE WHO . . ."

	Number of different similar sounding words ending in		
Target-word	ist	er	ero
philatelist	8	2	0
usurer	0	7	0
vaquero	0	0	8

Twenty-three distinct similar sounding words produced to the three words which are definitions of "a person who . . ." match their respective target-words ending while two similar sounding words match the ending of another target-word listed in Table 4 ($\gamma^2_{(1)} = 38.71$, $p < .001$). As the definition for *vaquero* might have provided information about the Spanish etymology of the ending of that word, independent of subjects' knowledge of the word, a test using only philatelist and usurer was also performed. These results were also significant ($p < .005$, Fisher exact probability test).

Rules for main stress in English are written in a way that count syllables from the end, as opposed to the beginning, of words (Chomsky & Halle, 1968). Thus a similar sounding word and a target-word are much more likely to have the same main stress if they have the same ending than if they have the same beginning. An empirical test of this hypothesis was made by examining all three and four syllable similar sounding words that (1) had the same number of syllables as their target-words, (2) matched either the first or last letter of their target word but not both, and (3) had a stress pattern listed in *Webster's New College Dictionary* (1963). Twenty words met these requirements: Of the seven beginning with the same letter as their target-word three matched the stress pattern and four did not, of the 13 ending with the same letter as their target-word 12 matched the stress pattern and one did not. The results were significant at the .05 level (Fisher exact probability test). Although this finding may be more a statement about English words *per se* than about word-name memory, it does demonstrate that for subjects in the TOT state knowing the ending of a word provides information about the word's stress pattern.

DISCUSSION

Previous work has shown that subjects in the TOT state often know the beginning and ending of their target-word, the number of syllables in the target-word, and the target-word's stress pattern. If the target-word has a suffix they tend to retrieve the suffix as a unit. This study has shown that for individual target-words clusters of letters tend to be retrieved together, that these clusters are often morphemes, that these clusters do not appear to be artifacts of using similar sounding words or the constraints imposed by definitions, and that the retrieval of one of these clusters at the end of a word increases the probability that the main stress will be placed correctly.

What are the implications of these findings for a model of word-name memory? First, it appears that word-name memory does not consist solely of inseparable strings of letters, but rather includes at least morphemes or morpheme-like clusters. Recent studies on the

perception of tachistoscopically presented words support this claim (Gibson & Guinet, 1971; Jarvella & Snodgrass, 1974; Murrell & Morton, 1974).

Second, the clusters are organized in a manner which allows for efficient utilization of information in the production and perception of language. Brown and McNeill (1966) considered their results solely in the light of a printed-word perception model. While no argument is being made against this here, word-name memory must also be organized in terms of a speech perception-production model in order to account for all the findings. Predictions can be made from a speech production model if we assume that the aspects of a word-name that are needed earliest in production will be the easiest to retrieve.

As breathing has to be controlled, the number of syllables and stress pattern would have to be known earliest in the planning of speech. Next the beginning of the word would be needed in order to prepare the transition from the previous word. A strong case cannot be made for a direct need for the endings of words in production. However, if word stems and morphemic endings are stored separately or if stress patterns are calculated by rule as opposed to being stored as separate facts about words, knowledge of words' endings would be needed before the stress could be assigned. Similar arguments could be made for speech perception. In this case, however, the end of a word would play a direct role in segmentation.

References

BARTON, M. I. Recall of generic properties of words in asphasic patients. *Cortex*, 1971, **7**, 73–82.

BROWN, R., & MCNEILL, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 325–337.

CHOMSKY, N., & HALLE, M. *The sound pattern of English*. New York: Harper and Row, 1968.

GESHWIND, N. Language and the brain. *Scientific American*, 1972, **226** (4), 76–83.

GIBSON, E. J., & GUINET, L. Perception of inflections in brief visual presentations of words. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 182–189.

HOWES, D. Hypotheses concerning the functions of the language mechanism. In K. Salzinger & S. Salzinger (Eds.), *Research in verbal behavior and some neurophysiological implications*. New York: Academic Press, 1967.

JARVELLA, R. J., & SNODGRASS, J. G. Seeing ring in rang and retain in retention: On recognizing stem morphemes in printed words. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 590–598.

LOFTUS, E. F., & COLE, W. Retrieving attribute and name information from semantic memory. *Journal of Experimental Psychology*, 1974, **102**, 1116–1122.

MURRELL, G. A., & MORTON, J. Word recognition and morphemic structure. *Journal of Experimental Psychology*, 1974, **102**, 963–968.

PERETTI, P. O. Initial and final word fragments as an aid to recall of whole words. *Psychological Studies*, 1974, **19**, 78–82.

Webster's Seventh New Collegiate Dictionary. Springfield, MA: G. & C. Merriam, 1963.

YARMEY, A. D. I recognize your face but I can't remember your name: Further evidence on the tip-of-the-tongue phenomenon. *Memory & Cognition*, 1973, **1**, 287–290.

(Received February 18, 1975)