# Retrieval of long-term memory: "Tip-of-the-tongue" phenomenon[1]

J. L. FREEDMAN AND T. K. LANDAUER
*STANFORD UNIVERSITY*

*Ss were able to predict better than chance which common-knowledge question that they had missed on a recall test they would subsequently answer correctly on a multiple-choice test. Providing the initial letter of the correct answer significantly facilitated recall as compared to provision of an incorrect first letter or no letter clue.*

We are often unable to produce the correct response to a question even though we "know" the answer. It is, we say, "on the tip of our tongue." Although this phenomenon has been described extensively (Freud, 1955; James, 1950), few controlled experiments have been done to investigate its properties. In addition to the phenomenon being interesting in itself, it would seem that studying instances in which retrieval of long-term memory breaks down might be a fruitful method of attacking the general problem of retrieval.

This experiment was conducted to study two specific questions concerning the "tip-of-the-tongue" phenomenon: (1) To what extent can the S accurately report that he does know the answer even though he cannot produce it? Hart (1965) has shown that Ss can predict significantly better than chance which items that they cannot recall they will subsequently be able to identify in a recognition test. The current study replicated this finding. (2) To what extent does a specific clue, i.e., the initial letter of the correct answer, facilitate recall of the unavailable response.

**Method**

Twenty students served as Ss. Each S listened to a tape recording containing 150 simple questions, each of which could be answered with a one-word response. These questions were constructed on the basis of extensive pretesting which indicated that they were likely to produce a high percentage of blocked responses (i.e., the S would be unable to supply the answer but would report that he thought he knew the answer). Examples: Who was Secretary of Agriculture under Eisenhower? What was Buffalo Bill's real name? The questions were read at the rate of three per minute. After each question, S wrote his answer. He was urged to guess if he had any notion of what the correct answer was. He also noted how certain he was that he knew the answer on a four point scale ranging from certain that he knew it to certain that he did not know it.

The list of 150 was read, there was a brief rest period, and then the same list of questions was read again in a different order. For this second reading, each S was provided with an answer sheet on which one-third of the questions had printed the initial letter of the correct response, one third had an incorrect letter, and one-third had no letter. Three different forms were used so that each question was represented in each group. Ss were informed beforehand that some of the letters were correct and some were not. For this administration Ss merely wrote their answers; there was no confidence rating. As in the first series, Ss were urged to record an answer even if they were not certain of it.

Finally, all Ss were given a written multiple-choice form containing the 150 questions and six possible answers for each. The Ss circled what they thought was the correct response. They were urged to guess if they had a hunch, but not to make completely wild guesses.

**Results**

To begin with it should be noted that there were a large number of "tip-of-the-tongue" instances produced. These may be defined as items in excess of chance on which S was incorrect on test I, but correct on test III. Of an average of 95.2 incorrect on test I, an average of 46.0 were correct on test III. Since only 16.7 would be expected to be correct due to guessing, there were clearly many true blocked items.

The first question of interest then is the extent to which Ss were able to report accurately whether or not they knew the answer when they were unable to produce it. The relevant data are the number of items in each of the four confidence categories that were missed on test I and identified correctly on test III. The mean proportions correct were .73, .61, .51, and .35 for confidence ratings of "Definitely know it," "Probably know it," "Probably don't know it," and "Definitely don't know it," respectively. An analysis of variance was performed. The linear trend is highly significant ($F = 41.59$, $df = 1/57$, $p < .001$) as is the overall effect of confidence ($F = 16.93$, $df = 3/57$, $p < .001$). In other words, even though he could not supply the correct answer without additional help, S was able to judge quite well whether or not he "knew" the answer. It should be pointed out, however, that this judgment, while quite good, was far from perfect. Even when Ss were certain that they knew the answer, they were correct on only 73% of the items; and when they were certain they did not know the answer, they were still correct on 35%, over twice what would be expected by chance.

The second question was whether or not supplying the initial letter of the correct answer would facilitate production of the answer. The data relevant to this

question consist of the number of blocked items (as previously defined) that were right on test II, when the correct letter, incorrect letter or no letter was given. The mean proportions were, respectively: Right clue, .34; Wrong clue, .12; and No clue, .15. The difference between the correct letter items and each of the other conditions is significant ($t = 5.45$ and 4.30 respectively); but the difference between the wrong letter and no letter conditions does not approach significance ($t < 1.0$).

## Discussion

How does provision of a hint or clue aid in retrieval of a "blocked" response? One plausible view of the retrieval process is that it consists of some kind of an ordered search through a memory store. A block might then consist of a fixated search in a wrong part of the store. Provision of a new cue might restart the search in another locus, thus providing a second chance for a successful scan. However, in this case one would expect any cue, even a wrong one, to be better than none. This is not the case in our data. Another possibility is that the store is organized according to an ordered array of stimulus properties such that provision of a cue allows entry to the array at a point nearer the desired item. For example, if word responses were stored alphabetically, then provision of the correct initial letter would lead the search to an alphabetic listing containing the answer, while an incorrect initial letter would lead to a blind alley. Another, more traditional, hypothesis is that the additional cue, when correct, simply adds another modicum of association linkage to that provided by the original question, thus raising the response strength of the correct answer, and perhaps driving it over some emission threshold. The last two hypotheses (among others) do not seem separable on the basis of available data.

## References


**Freud**, S. *Psychopathology of everyday life*. New York: Mentor, 1955.

**Hart, J. T., Jr.** Recall, recognition, and the memory-monitoring process. Unpublished doctoral dissertation, Stanford University, 1965.

**James, W.** *The principles of psychology*. Vol. I. New York: Dover, 1950.


## Note


1. We thank Robert Gilbert for assistance in running the experiment, and Anthony Doob for help in the data analysis. The research was supported in part by National Science Foundation grant No. G. S. 196.