

Effects of distorted auditory and of rhyming cues on retrieval of tip-of-the-tongue words by poets and nonpoets

LYNN T. KOZLOWSKI

Wesleyan University, Middletown, Connecticut 06457

Auditory cues (the target word distorted by a low-pass filter) and rhyming cues act as retrieval aids for people who have a word on the tip of their tongue. Parallels between the tip-of-the-tongue phenomenon (TOT) and the perception of well-put passages of poetry are also discussed. It is argued that the effects of these poetic passages derive in part from the engagement of TOT-like processes. In support of this hypothesis, poets (poetry appreciators) are shown to be more aware of being helped by TOT retrieval cues than are nonpoets (poetry nonappreciators); however, the retrieval cues do not differentially influence successful recall of TOT words by poets and nonpoets.

A person in the tip-of-the-tongue (TOT) state is on the track of a particular word (Brown & McNeill, 1966): Accurate indications are found along the way—similar sounds, the first or last letter, the number of syllables, or the primary stress. Freedman and Landauer (1966) have shown that supplying the first letter of a TOT word helps people remember the word. The present research explored the helpfulness of two other types of possible TOT retrieval cues, both derived from the partial knowledge that is often available during the TOT state.

In addition, the present research looked for differences between poets and nonpoets in their susceptibility to help from these TOT retrieval cues. Hymes (1960) and Lynch (1953) have presented evidence for the existence and importance of key words in lyric poetry. According to Hymes (1960), key words in sonnets (1) contain sounds which are dominant or very common in the poem, (2) express the theme of the poetic unit, and (3) are located in a position so as to have a culminating effect (cf. Smith, 1968).¹ The devices and constraints of poetry may provide the skilled reader with TOT-like knowledge of critical words in poems. The poetic devices in question include rhyme (giving knowledge of similar sounds), alliteration (giving knowledge of initial letters), and meter (giving knowledge of the number of syllables and the stress pattern). I propose that the full appreciation of key poetic passages depends upon a TOT-like process, such that, while reading (or rereading) poems, one can have the feeling of being part way through the constructive process of reading a specific word—before

the boundaries of that word are encountered. Then, when that word is read, it appears extremely and excitingly fitting, and some of the emotional accompaniments to solving a TOT may be engaged. Reading clearly involves a constructive process, highly influenced by systematic context effects (Hochberg, 1970; Rayner, 1975; Tulving & Gold, 1963). Brown (1970), in a largely ignored discussion of the causes of the TOT phenomenon, notes that reading makes use of a *noticing order*, that is, a regular order of processing the features of a word. Poems, I am arguing, through their constraints on meaning, rhyme, rhythm, and first letters, are capable of putting a reader part way through the noticing order of a word that is not yet read. When the word is read, it has already been constructed to such an extent that a special recognition experience takes place.

One other critical parallel between key words in poems and the TOT phenomenon should be mentioned. They each concern problems of word production. If either type of word could be produced easily, the situation, by definition, would not be a TOT, and, in the case of poetry, would not be interesting or original (Kammann, 1966).

If TOT and these poetic processes are equivalent, then people who differ in poetic processes should differ in TOT processes. Poets (poetry lovers) and nonpoets (poetry haters) should differ in the number of TOT experiences they have or in some other aspect of the TOT phenomenon.² Insofar as poetic constraints involve TOT-like processes which in turn affect the appreciation of poetic passages, it seemed reasonable that, when in the TOT state, poets may be more susceptible than nonpoets to being helped by cues, especially those cues normally supplied in poetry.

Francès (1958) showed that musically trained subjects were better than untrained subjects at making discriminations of rhythmic alterations in simple melodies. In addition, James' discussion of the TOT (which

Thanks go to K. Bryant, M. Savage, and J. Sommers for their assistance in Experiment 1; to B. J. Buckley, K. Wold, and M. Schulman for their assistance in Experiment 2; to J. Cutting, in part for suggesting the use of the auditory filter; and to Kathleen MacHan Kozlowski. This research was supported by a grant from Wesleyan University. Reprints are available from the author at the Department of Psychology, Wesleyan University, Middletown, Connecticut 06457.

includes a foreshadowing of the present argument about poetry) points to the significance of rhythm: "Every one must know the tantalizing effect of a blank rhythm of some forgotten verse, restlessly dancing in one's mind, striving to be filled out with words" (1890, p. 252). With this as a background, a cue with rhythmic information was sought. A technique such as having the rhythm of a word beat out on a bongo drum would have given a highly artificial cue. A distorted tape-recorded version of the potential TOT word was recommended as a cue which would be richer and more naturalistically related to language rhythms. The auditory distortion was accomplished by a low-pass filter, producing a stimulus sounding as words might to a person going deaf. The cue is complex, giving information about the stress pattern, rhythm, and number of syllables of the word along with some tonal information. This cue seemed to contain information that was more subtle and less consciously usable for retrieval than the first letter cue found in Freedman and Landauer (1966). The distorted auditory cue had another provocative property: When reading a list of the correct words while listening to the cues, it was difficult not to hear the words perfectly—they sounded unfiltered. However, with no list to follow, it was usually impossible to guess the correct word, and it sounded as if a man was speaking through a pillow.

EXPERIMENT 1

Method

Subjects. Subjects were college undergraduates and received \$5. Poets (five males, four females) were by self-report at least "moderately" involved in actively writing poetry and liked to read poetry. Nonpoets (eight males, two females) said that they disliked poetry and that they were no more than "slightly" involved in writing poetry (only 2 of the 10 nonpoets were "slightly" involved in writing).

Questions to produce TOTs. Questions covered a broad range of knowledge in both the sciences and the arts, so that neither poets nor nonpoets would have an advantage. Several questions were taken from Brown and McNeill (1966). Some sample questions were: Who invented the steamboat? What was the name of the woods in *Macbeth*?

Formulation of cues. All cues were forced through a low-pass crystal filter made by Allison Laboratory. The frequency cutoff was 222.0 Hz, falling off at 6 dB per octave. They were recorded on a Sony 353 tape recorder. Two Sony 353 speakers were used in the presentation of the cues.

Correct cues were the correct answers to the questions. Incorrect cues were English words that were selected by counting the number of syllables in the correct cue. If the number of syllables was even, the false cue had to be odd and start with a different first letter; if the number was odd, the cue had to be even and start with a different first letter. Noncues were recorded as white noise for 30 sec.

Three magnetic tapes were used to record the set of cues. On each tape there was one cue word for each of the 75 questions: 25 correct cues, 25 incorrect cues, and 25 noncues, arranged in random order. Across the three tapes, each question was paired once with each of the three cue types. A male read each of the cues three times to achieve as much rhythmic and syllabic clarity as possible. The cue was repeated 5 sec after its first read-

ing and 10 sec after that. A 15-sec pause was allowed between separate cue words.

The distorted cues were tested for intelligibility by letting naive subjects try to guess the original word. Among the three subjects in this control group, only 2 out of 225 words were guessed correctly.

Procedure. Subjects were run in groups of two to six, balanced as much as possible per session for poets and nonpoets. The use of the three different tapes was also balanced as much as possible. (Since the three tapes did not differentially help subjects solve TOTs, all analyses in the Results section ignore tapes as a variable.)

The experimenter read these instructions: "The experiment concerns a phenomenon of memory retrieval known as the tip-of-the-tongue phenomenon. The tip-of-the-tongue state is that state of mind in which a person is unable to think of a word that he is certain he knows. If you are unable to think of the word but feel sure you know it and that it is on the verge of coming back to you, then you are in a tip-of-the-tongue state.

"In front of you, you will notice a stack of 75 forms, each corresponding to an individual question. For each form, the top sheet presents a description or statement, a definition, if you will. What you are asked to do is to supply the appropriate *single word* corresponding to that description. All correct answers are single words—phrases, etc., are not appropriate.

"After you have had a few seconds to read the question and respond, we will play a short segment on this tape recorder. What is on the tape is a *distorted* recording of the correct answer. The words have been distorted beyond recognition by means of a special instrument. Hence, *though you will not be able to make out the proper pronunciation from the tape, you will still hear the correct rhythm and accent of the word—the proper cadence and intonation.* We ask you to listen to the recording and use it as a cue in helping you to recall the right answer. For certain questions there will be no cue given; for others a wrong cue, that is, the word, which has been distorted, is a word other than the correct one."

The experimenter then went over the answer sheets. A sample question and the correct cue were given. The answer sheets included a question on how confident they were in their response (not at all = 0 to completely confident = 4) and the crucial question: Would you say that the correct answer is on the "tip of your tongue"? (not at all = 0 to completely confident = 4).

After the cue was played, subjects were asked for the correct answer again and asked: Did hearing the recorded cue help or interfere with your finding the answer? (interfered = -3 to helped = +3, in 7 points).

At the completion of the set of 75 questions, a six-option multiple-choice questionnaire was administered to see if subjects could recognize the correct answers (as in Freedman & Landauer, 1966). If a subject could not recognize the answer and had claimed a TOT for that answer, data for that item were discarded on the grounds that cues were only assuredly correct and incorrect in relation to the answer intended by the experimenter. Each session lasted about 3 h.

Results and Discussion

Number of TOTs. A TOT was measured as anything from a slight to an extreme TOT experience. When not at all in a TOT state, subjects recognized an average of 46% of the correct answers on the multiple-choice test; when they had moderate TOTs (1 or 2, by the scale), they were 61% correct; when they had extreme TOTs (3 or 4, by the scale), they were 84% correct [linear trend, $F(1,34) = 12.91, p < .01$]. As can be seen, subjects were not perfectly accurate in their estimates of words being on the tips of their tongues. (Remember,

Table 1
Mean Percentage of TOTs Correctly Solved as
a Function of Distorted Auditory Cue

	Distorted Word	
	Correct	Incorrect
Poets*		
Mean	36.44	18.11
SD	15.36	14.50
Nonpoets**		
Mean	34.60	12.20
SD	19.89	10.17

* $n = 9$ ** $n = 10$

mistaken TOT questions were dropped from a particular subject's data.) Inaccurate TOT appraisals seemed to be due to two factors. First, as in a signal detection situation, people were highly motivated to have TOTs and there were no obvious penalties for mistakes. This acted to artificially lower the threshold for TOTs and increase the frequency of false-positive judgments. Second, unfortunately, some questions had more than one plausible answer; for example, "An instrument used in navigating for measuring altitudes?" could be an astrolabe, an altimeter, or a sextant (the intended answer). On the average, subjects had 5.42 accurate TOTs per cue condition ($SD = 2.44$). There was no hint of poet-nonpoet differences here ($F < 1$).

Effect of cues on recall during TOT state. The incorrect- and no-cue conditions produced nearly identical results ($F_s < 1$). To get a more stable estimate of control effects, these conditions were combined, and an average of the two was taken as the datum for the control condition. (Freedman & Landauer, 1966, also found no differences between their incorrect- and no-cue conditions.) Table 1 shows the means for the percentage of correct responses by people in the TOT state, as a function of the correct-cue and control conditions. Correct cues did help greatly [$F(1,17) = 24.40, p < .01$]; contrary to expectation, the cues did not help poets more than nonpoets ($F < 1$).

Effect of cues on recall when not in TOT state. When not in the TOT state and when unable to recall the answer (but able to recognize it on the multiple-choice test), subjects solved an average of 3.4% of the questions as a function of correct cues and 0% of the questions as a function of control conditions. Correct cues did not produce significantly more recall than did control cues [$F(1,17) = 2.21, p > .05$]. Again, there were no poet-nonpoet effects ($F_s < 1$).

When subjects were in a TOT state, correct cues did help them more than when they were not in a TOT state [interaction, $F(1,17) = 14.03, p < .01$].

Effect of cues on subjective sense of cue helpfulness. The following analysis was confined to the correct-cue condition, because there were too few correct answers in the incorrect-cue condition for an adequate comparison: Fifty-three percent of the subjects had no correct

answer. For all the TOTs a given subject solved, the perceived degree of helpfulness of the auditory cue (helped = +3, interfered = -3, all 7 points used) was averaged. This gave an estimate of how much the subjects felt the cue contributed to their being able to answer the question. Due to the non-normal distribution of scores, a median split was performed (one nonpoet had no correct answers and was excluded). Seventy-eight percent of the poets and 22% of the nonpoets scored above the median, that is, felt that the correct cues were very helpful (difference by Fisher exact test, $p = .028$, one-tailed).³ Although recall performance does not differ between poets and nonpoets, the phenomenological experience of the benefit of the cues may differ. The important dissimilarities between poets and nonpoets, then, may concern their awareness of the information processes involved in recall rather than concern their overt success at recall.

How sensitive are the poets? When they do not recall the answer, can they detect help or interference from the cues? There is no evidence that either poets or nonpoets can do this (Fisher exact test, $ps > .20$).

EXPERIMENT 2

This study was identical to the first except that a rhyming cue was used. Such a cue has little of the informational subtlety of the distorted auditory cue, but was of obvious relevance to poetry.

Method

Subjects. These college undergraduates received homemade cookies in payment. Poets (four males, seven females) had published at least one poem in a campus literary magazine or had given an invited poetry reading on campus. Nonpoets (eight males, seven females) said they disliked poetry and were not at all involved in writing poetry.

TOT questions. Some of the best target words from Experiment 1 had to be discarded because they were too difficult to rhyme or because they had too many variant pronunciations (e.g., Nebuchadnezzar). From the remaining questions, the 50 which had produced the most TOTs were selected.

Cues. A panel of five assistants helped select the rhymes according to the following criteria: (1) They could not have the same first letter; (2) their meaning should have no strong association with the correct answers, and (3) the number of syllables and rhymes should be as similar as possible to those of the correct answers. The rhymes sometimes consisted of more than one word (e.g., *Excalibur* was rhymed by *tall conifer*). As much as possible, the incorrect rhymes were from a similar semantic domain as the correct rhyme. Incorrect rhymes did not have the same first letter as the correct answer, and had the same number of syllables, but often a different stress pattern, than the correct answer (e.g., *sampan* was rhymed by *can-can* and not rhymed by *fox-trot*).

A no-cue condition was omitted.

Three different random orders of cue arrangement were used. These orders made no difference to the pattern of results and are not discussed further.

Procedure. Subjects were again run in groups of two to six, balanced as much as possible for sex and poet-nonpoet status.

The answer sheets were almost identical to those in the first study. Nine-point rather than 7-point bipolar scales were used

Table 2
Mean Percentage of TOTs Correctly Solved
as a Function of Rhyming Cue

	Rhyme	
	Correct	Incorrect
Poets*		
Mean	51.64	5.91
SD	39.98	10.20
Nonpoets**		
Mean	36.57	8.93
SD	26.58	16.14

n* = 11*n* = 15

for the question on the "helpfulness" of the rhyming cue. The cues were stapled on the back of the answer sheet. When subjects worked through the answer sheet to the point where the cue was supposed to be given, they turned the page over and administered the cue to themselves. Instructions were tape recorded and identical to those in Experiment 1, except that the cue was said to rhyme.

Results and Discussion

Number of TOTs. Subjects averaged 6.32 accurate TOTs in total (*SD* = 2.75). There were no poet-nonpoet differences ($F < 1$).

Effects of cues on recall when in the TOT state. Table 2 shows the means for the percentage of correct responses by poets and nonpoets in the TOT state, as a function of rhyming cue. Correct rhymes did help everyone solve the TOT problems [$F(1,24) = 18.75$, $p < .01$]. Poets were not helped more than nonpoets ($F < 1$). The poet-nonpoet difference in the correct-cue condition did not approach statistical significance ($F < 1$).

Effects of cues on recall when not in the TOT state. The rhyming cue seems to present more information than does the auditory cue. When given a correct rhyme, subjects solved an average of 39.9% (*SD* = 28.43) of the questions; when given an incorrect rhyme, they solved 0%. The correct rhyme helped recall more than did the incorrect rhyme [$F(1,24) = 50.41$, $p < .01$]. There were no poet-nonpoet differences ($F < 1$).

When a correct cue is given, being in a TOT state does not lead to better recall than does not being in a TOT state [interaction, $F(1,24) = 1.4$, *n.s.*]. It should be remembered that the auditory (Experiment 1) cue was not powerful enough to produce recall in people when they were not in the TOT state.

Effects of cues on subjective sense of cue helpfulness. There were several reasons why the main analysis from Experiment 1 could not be repeated here. Poets (mean = +3.61, *SD* = .67) and nonpoets (mean = +3.5, *SD* = .85) both agreed that the correct cues helped them get the correct answer. In the first study the distorted auditory cue was subtle and transitory. Even if the cue helped the subjects get the correct answer, it was no longer available for comparison with the answer. However, with the printed rhyming cue, the cue was available for easy comparison with the answer. Either it rhymed or it did

not. The obviousness of the rhyming cue's relationship to the correct answer may account for everyone's ability to say that it helps them (and for the ability of the cue to help people even when they are not in the TOT state).

If we look at the feeling of helpfulness when the TOT word was not actually recalled, we are able to evaluate more subtle cue effects. In this case, since the correct answer has not been found, it is not available for comparison to the rhyming cue. In the correct-cue condition, there were so many successful recalls that there are too few instances of no answers to do any analyses. In the incorrect-cue condition, however, there were many failures to get the correct answer. More poets (73%) than nonpoets (40%) were sensitive to the interference of the incorrect cues, by the criterion of having a negative mean score on the helped (+)/interfered (−) scale (Fisher exact test, $p = .036$, one-tailed).

Since rhymes influenced recall even when there was not a TOT state, it seemed important to see if the sense of help or interference of the cues extended to non-TOT conditions. There were no systematic poet-nonpoet effects in their responses to the helpfulness question for correct or incorrect rhymes (Fisher exact test, $ps > .20$).

An Additional Result from Experiments 1 and 2

TOT and the feeling of knowing. People who fail to recall a word are able to reliably say whether they could recognize it (Blake, 1973; Koriat, 1975; Tulving & Thomson, 1971). The TOT state is a special case of the feeling of knowing in which recall is felt to be imminent. In the present experiments, a 5-point question about whether the subjects believed that they could recognize the correct answers (to as yet unanswered questions) correlates significantly with the ratings of the TOT state (mean $r = +.87$): This correlation derives from subjects' confidence and TOT ratings for each question not solved before the cue was given, and was averaged across subjects in both experiments.

GENERAL DISCUSSION

It is claimed that some of poetry's characteristic effects derive from TOT-like processes. In the present project, support for this claim requires (1) that poets be shown to be more sensitive than nonpoets to assistance from TOT-retrieval cues and (2) that the TOT state be shown to have special status in the cuing process.

Evidence for the increased sensitivity of poets to retrieval cues was found only for their awareness of the cue's aid and not for actual word retrieval. The self-report findings are encouraging, but far from compelling: Perhaps poets and nonpoets differ in their willingness to report what is happening to them, rather than in their ability to monitor their cognitive processes. Nevertheless, the present results do coincide with work showing perceptual or cognitive style differences (e.g., Klein, 1970) between appreciators and nonappreciators of art

or between artists and nonartists (e.g., Child, 1965; Stumberg, 1928). It is interesting to speculate that nonpoets suffer a type of cognitive myopia to the workings of poetic language. They may be able to play the verbal game involved in poetry, as evidenced by their recall performance in the present study, but they may lack an appreciation of all that goes on during the game.

When distorted auditory cues are used, the TOT state seems to be privileged: Correct cues increase recall only when the individual is in the TOT state. However, correct rhymes increase recall in both TOT and non-TOT states. Looking more closely at Experiment 2, support can be found for the special status of the TOT state. First, it should be stressed that the perceived interference of incorrect rhymes was felt relatively more by poets than nonpoets *only* when they were in the TOT state. Second, the fact that correct rhymes help even when there is no TOT state is probably due to (1) the power of the rhyming cue and (2) the delicacy of the boundaries between failed recall, the TOT state, and successful recall. Correct rhymes give more information than do corresponding auditory cues. The number of single words rhyming with a correct answer is smaller than the number of words sufficiently matching a distorted auditory cue. In addition, the rhymes place the subject immediately in the realm of a set of English words (as distinct from sounds and rhythms), while the auditory cues are capable of supplying the listener with words only when the person is well on the track of the correct answer (see comments at end of Introduction). Also, if asked the name of, say, King Arthur's sword and given the rhyme *tall conifer*, a person might answer *Excalibur* if he or she knew only that it was the name of a famous sword: The TOT state is certainly not required for this to happen. It would seem that more subtle cues (the distorted words) are needed to preserve the boundaries between failed recall and the TOT state. Too strong a hint will give someone the answer even if they are not in a TOT state, but subtle hints will be effective only when people are in the TOT state.

Poetic Devices and Key Passages in Poems

The present paper argues that the efficacy of poetic punch lines is the result of a TOT-like process which is engaged by the devices and constraints of poetry. Those who consider poetic devices as adornments and poetic constraints as rigors for the sake of rigor should attend to this possible function of poetic technique (cf. Smith, 1968, on clausal effects in poetry).

Levin (1962), in a linguistic analysis, discusses poetry in a way consistent with the TOT hypothesis: "A poem . . . has built into it such equivalences . . . that it assists the individual in re-encoding it uniquely; it prompts, as a result of its own systemic pressure, the same selections from the language code" (p. 61-62).

Retrieval Cues and TOTs

One of the most interesting results of this project has been the subtle effectiveness of the distorted auditory

retrieval cues. Such stimuli are naturalistically related to normal auditory information and may prove useful in other memory research.

REFERENCES

- BLAKE, M. Prediction of recognition when recall fails: Exploring the feeling of knowing phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 311-319.
- BROWN, R. Psychology and reading: Commentary on chapters 5 to 10. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.
- BROWN, R., & McNEILL, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 325-337.
- CHILD, I. L. Personality correlates of esthetic judgment in college students. *Journal of Personality*, 1965, 33, 476-511.
- FRANÇÈS, R. *La perception de la musique*. Paris: Vrin, 1958.
- FREEDMAN, J. L., & LANDAUER, T. K. Retrieval of long-term memory: "Tip-of-the-tongue" phenomenon. *Psychonomic Science*, 1966, 4, 309-310.
- HOCHBERG, J. Components of literacy: Speculations and exploratory research. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.
- HYMES, D. H. Phonological aspects of style: Some English sonnets. In T. A. Sebeok (Ed.), *Style in language*. Cambridge, Mass: M.I.T., 1960.
- JAMES, W. *The principles of psychology*. New York: Holt, 1890.
- KAMMANN, R. Verbal complexity and preferences in poetry. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 536-545.
- KLEIN, G. S. *Perception, motives, and personality*. New York: Knopf, 1970.
- KORIAT, A. Phonetic symbolism and feeling of knowing. *Memory & Cognition*, 1975, 3, 545-548.
- LEVIN, S. R. *Linguistic structures in poetry*. The Hague: Mouton, 1962.
- LYNCH, J. J. The tonality of lyric poetry: An experiment in method. *Word*, 1953, 9, 211-224.
- RAYNER, K. The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 1975, 7, 65-81.
- SMITH, B. H. *Poetic closure: A study of how poems end*. Chicago: University of Chicago Press, 1968.
- STUMBERG, D. A study of poetic talent. *Journal of Experimental Psychology*, 1928, 11, 219-234.
- TULVING, E., & GOLD, C. Stimulus information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology*, 1963, 66, 319-327.
- TULVING, E., & THOMSON, D. M. Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology*, 1971, 87, 116-124.

NOTES

1. I think that key words in poems are composed of the intersection of more elements than provided for in Hymes' treatment (cf. Levin, 1962), but Hymes' discussion helps describe the phenomenon at issue.

2. Poets are of interest here not so much as creators but as appreciators of poetry. Those nonpoets who dislike poetry are used primarily in their capacity as nonappreciators of poetry.

3. Fisher exact tests are used for all subsequent analyses of helpfulness effects, because of a general pattern of non-normal distributions of scores; t tests were also performed and disclosed a comparable pattern of results.