Recognition failure of recallable words in semantic memory

PAUL MUTER

University of Toronto, Toronto, Ontario, Canada M5S 1A1

In an experiment in which there was no study phase, 54 subjects were tested for recognition of famous surnames and then were tested for cued recall of the same surnames. Subjects failed to recognize 53.4% of names that they subsequently recalled. Recall was significantly higher than recognition. The relationship between overall recognition rate and recognition rate of recallable words closely resembled that reported by Tulving and Wiseman (1975) for episodic memory experiments. The present data therefore extend the generality of this relationship, and of the principle that the probability of retrieval from memory depends critically on the cues provided. It is argued that the similarity between results for episodic memory experiments and the present semantic memory experiment can be more parsimoniously accommodated by tagging theory than by episodic theory.

It has been demonstrated repeatedly that under certain circumstances subjects fail to recognize a substantial proportion of recently memorized words that they later recall (e.g., Postman, 1975, Experiment 2; Tulving & Thomson, 1973; Watkins & Tulving, 1975). The theoretical interpretation of this phenomenon of recognition failure of recallable words, however, remains in question. According to episodic theory (Tulving, 1972, 1976; Tulving & Thomson, 1973; Watkins & Tulving, 1975), at the time of input a unique episodic memory trace is formed, and whether a retrieval cue will be effective depends on the relation between the cue and this episodic trace. If the information in the retrieval cue "matches" the information in the episodic trace, the item will be remembered (recognized, if the cue is a copy of the target, recalled if the cue is not). Under some circumstances, it is argued, it happens that a target item is encoded in such a way that a copy cue is not effective, but a noncopy cue is.

According to tagging theory, recently modified to accommodate the finding of recognition failure of recallable words (Anderson & Bower, 1974; Martin, 1975; Reder, Anderson & Bjork, 1974), there is no unique episodic trace. When an item is presented for study in an episodic memory experiment, an occurrence tag is attached to a node in the mental lexicon. A given word (e.g., water) which has several shades of meaning, may be represented in the lexicon by several nodes, but normally they will not all be tagged during input. At the time of the recognition

test, one node for each of the test words is automatically accessed, and if it has a tag a positive response is emitted. At the time of the recall test, a subset of all nodes is accessed and examined, and those which have a tag are output. If one of the "water" nodes is tagged during input, not accessed in the recognition test (because a different, untagged "water" node is accessed), but successfully accessed in the recall test, recognition failure of a recallable word will occur. (For more complete descriptions of episodic theory and tagging theory, see Watkins & Tulving, 1975.)

Tulving and Wiseman (1975) summarized results from 40 conditions from 12 different episodic memory experiments. Tulving and Wiseman discovered an orderly and remarkable relationship between overall recognition rate and amount of recognition failure of recallable words. This relationship is demonstrated by the curve in Figure 1, which fits the data quite well. The obtained relationship is remarkable because according to some prominent views of memory (Bahrick, 1970; Kintsch, 1970), the proportion of recallable words recognized should be 1.0 regardless of the overall level of recognition.

If the same pattern of recognition failure of recallable words could be demonstrated in a test of semantic memory (e.g., a test of ability to recognize and recall famous surnames), tagging theory could interpret the result in the same terms that it uses to interpret the analogous result in episodic memory experiments: several nodes for a single nominal surname; the examination of nodes for information indicating fame (rather than recent occurrence); and, for a given surname, the possibility of accessing a different node in the recognition test than in the recall test. Furthermore, such a finding would add credibility to tagging theory in that it would prove that the observed pattern of recognition failure of recallable

This research was supported by Research Grants APA146 from the National Research Council of Canada and OMHF164 from the Ontario Mental Health Foundation to Bennet B. Murdock, Jr., and by a National Research Council Postgraduate Scholarship. The author is grateful to F. I. M. Craik, A. J. Flexser, B. B. Murdock, Jr., and E. Tulving for helpful comments.

Figure 1. Recognition rate of recallable words as a function of overall recognition rate, with 95% confidence intervals indicated for each point (standard error = $\sqrt{p(1-p)/n}$).

words can occur when there is no unique episodic trace, and therefore that the concept of a unique episodic trace is not necessary to explain this pattern. Episodic theory, on the other hand, could not interpret such a result in the same terms that it interprets the analogous effect in episodic memory experiments, and such a result would suggest a degree of similarity between retrieval from episodic memory and retrieval from semantic memory that presumably would be unpalatable to episodic theorists, who posit two different memory systems, episodic memory and semantic memory.

The present experiment was similar in many respects to the episodic memory experiments which demonstrated recognition failure of recallable words (e.g., Tulving & Thomson, 1973); the essential difference was that the tests were of semantic memory.

METHOD

Fifty-four students in an undergraduate cognitive psychology summer course volunteered to be subjects.

The 16 target surnames used in the experiment are given in Column 1 of Table 1. These targets were selected to meet the following criteria: They are names of people who are listed in the *Random House Dictionary* (1969) and who, in the estimation of the experimenter, achieved their fame before 1950; there are no other people listed in the dictionary with the same surname that the experimenter recognized; they are names sufficiently common to take up at least half a column in the 1976 Toronto telephone directory; and, names with obvious variations in spelling, such as "Johnson," were excluded. For the recall tests, given names and a brief description were

provided as cues. These cues are listed in Column 2 of Table 1. The lures for the recognition test were Ash, Bedford, Bradshaw, Clarkson, Conway, Ferguson, Foley, Good, Lane, Lynn, Olson, Richmond, Sanderson, Trent, Turnbull, and Walker. These lures were generated quite arbitrarily; the only explicit constraint was that they not be listed in the *Random House Dictionary*.

The subjects were run in a group in a large classroom. Each subject received a self-explanatory booklet and proceeded through it at his own pace. Page 1 of the booklet had the following instructions: "Please go through the booklet and complete each page one at a time. Never look back to a previous page. At all stages of the experiment, you have as much time as you need."

At the top of Page 2 were these instructions: "Some of the following surnames belong to people who are or were famous (i.e., they are in the *Encyclopedia Britannica*) and some do not. Please circle those names which you recognize as belonging to a person (real, not fictional) who was famous before 1950. (Please do not try to guess what proportion are intended to be famous; simply call them as you see them.)" It was felt that subjects would be more familiar with the *Encyclopedia Britannica* than with the *Random House Dictionary*. Below these instructions was the list of 32 surnames, 16 targets and 16 lures, in a different randomized order for each subject.

Pages 3 and 4 of the booklet were blank. On Page 5, the subject was instructed to "Please try to fill in the blanks with appropriate surnames. Do not turn back to previous pages." Below were the 16 recall cues, in a different randomized order for each subject.

Finally, on Page 6 were the following instructions: "Were there any names which you recalled on the second test, and recognized on the first test but did not circle because you thought they were not sufficiently famous or were after 1950? (Do not turn back to previous pages.) If so, please write the name(s) and explain what happened."

RESULTS

The results collapsed across subjects and words are presented in Table 2. Substantial recognition failure of

Table 1
Materials

Targets	Cues
Bryan	U. S. politician and religious leader: William Jennings
Burke	British 18th century political thinker: Edmund
Cooper	Author of <i>Last of the Mohicans</i> : James Fenimore
Davis	President of the Confederacy during the U. S. Civil War: Jefferson
Doyle	Author of the Sherlock Holmes stories: Sir Arthur Conan
Ellis	Sexologist and psychologist: Havelock
Foster	U. S. song writer (<i>Swanee River</i> , <i>Old Black Joe</i>): Stephen
Franklin	U. S. statesman and inventor: Benjamin
Harvey	Discoverer of circulation of the blood: William
London	U. S. author of <i>Call of the Wild</i> and <i>The Sea Wolf</i> : Jack
Morton	U. S. jazz pianist and composer: Jelly Roll
Ross	Maker of the first U. S. flag: Betsy
Spencer	English 19th century philosopher: Herbert
Thomas	Welsh poet: Dylan
Whitney	Inventor of the cotton gin: Eli
Young	U. S. Mormon leader: Brigham

Table 2
Overall Frequencies of Response Classes

	Recalled	Not Recalled	Total
Recognized	169	85	254
Not Recognized	194	416	610
Total	363	501	864

recalled words occurred—53.4% of recalled items were not recognized. Of the 16 words, every one was recalled but not recognized by at least two subjects. Except for two subjects who correctly recalled only two words each, all of the 54 subjects failed to recognize at least one recalled word. Furthermore, recall was significantly higher than recognition [$t(53) = 4.54$, $p < .0001$]. (Recognition failure of recallable words does not necessarily imply superiority of recall over recognition; see Wiseman & Tulving, 1976.) The data for individual words are presented in Table 3.

On the recall test, five subjects wrote "Bryant" instead of "Bryan," and one subject wrote "Davies" instead of "Davis." These recall attempts were scored as incorrect. If they had been scored as correct, the rate of recognition failure of recallable words would have increased, since the names were not recognized in any of the six cases. Overall, there were 23 extralist intrusions and 14 intralist intrusions, including six inappropriate targets, five lures from the recognition test, and three given names from elsewhere in the recall test. On the recognition test, overall there were 78 false positives (9.0% of all lures).

In response to the posttest query, no subjects reported that there were names which they recalled on the second test and recognized on the first test but did not circle because they thought the names were not famous enough or were famous after 1950.

Figure 1 portrays the relationship between overall recognition rate and recognition rate of recallable words. Each point represents the data for four names collapsed across subjects. Names were ranked according to overall recognition rate and then were blocked in groups of four. The curve in Figure 1 is the function used by Tulving and Wiseman (1975) to fit the data from 40 conditions from episodic memory experiments: $P(R_n \text{ given } R_c) = P(R_n) + .5[P(R_n) - P(R_n)^2]$, where R_n is successful recognition and R_c is successful recall.

DISCUSSION

The relationship between overall recognition rate and amount of recognition failure of recallable words in the present experiment is strikingly similar to that obtained in the episodic memory experiments summarized by Tulving and Wiseman (1975). Given the considerable amount of scatter of the data in Figure 1 of Tulving and Wiseman, there is no evidence

that the present semantic memory data are different from the episodic memory data with respect to the relationship between overall recognition rate and amount of recognition failure of recallable words. Tulving and Wiseman reported that the amount of recognition failure of recallable words is a certain function of the overall level of recognition in a particular episodic memory paradigm. The present data suggest that, in fact, their function applies more generally than they realized; it extends even to semantic memory. The present data also reinforce Tulving's more general position (e.g., 1974) that many memory theorists have not ascribed sufficient importance to the role of the retrieval environment in memory phenomena.

The objection could be raised that the present experiment may tap episodic memory, not semantic memory. Admittedly, the fact that there is no experimental input does not necessarily mean that the episodic system is not involved. However, the present task is clearly not an episodic memory task according to Tulving's own definitions: Episodic memory is memory for "autobiographical events, describable in terms of their perceptible dimensions or attributes and in terms of their temporal-spatial relations to other such events" (Tulving, 1972, p. 387). Even if some subjects could remember episodes in their lives relating to some of the names in Table 1, it would not follow that the test was of episodic memory: "The assignment of a task to one or the other category [semantic or episodic memory] depends on the kind of memory query addressed to the person" (Tulving, 1972, p. 386). The queries in the present experiment are about the subject's "knowledge," not his "personal experience," and therefore tap semantic memory (Tulving, 1972, p. 387).

The possibility of a more conservative response criterion on the recognition test than on the recall test is

Table 3
Frequencies of Response Classes for Each Word

	Recognized		Not Recognized	
	Recalled	Not Recalled	Recalled	Not Recalled
Franklin	44	0	10	0
Whitney	18	8	5	23
Burke	15	10	4	25
Doyle	20	3	16	15
Spencer	1	18	3	32
Harvey	12	6	2	34
Cooper	10	7	13	24
Thomas	16	0	36	2
London	9	5	13	27
Ross	11	2	21	20
Morton	0	9	6	39
Bryan	3	4	6	41
Davis	3	4	15	32
Ellis	2	4	7	41
Foster	4	2	12	36
Young	1	3	25	25

not a problem because of the following considerations: There were more than twice as many false positives in recognition as intrusions in recall; a conservative criterion would affect the recognition rates of both recalled and nonrecalled words; and similar potential criterion differences are normally present in the episodic memory experiments to which the present experiment is being compared.

There are a number of other comments that could be made with respect to the present experiment which apply equally to the episodic memory experiments. Some examples are: in the recognition test, subjects may try to generate the cues from the test words; the memory "trace" may consist of the cue-target combination, not just the target; the recognition test may enhance performance on the later recall test. The primary purpose of the present paper is to compare results from tests of episodic and semantic memory, and therefore these comments and others that apply equally to both situations are largely irrelevant.

If, as Tulving (1972) has suggested, there are two memory systems, episodic memory and semantic memory, the present data suggest that in some respects retrieval from semantic memory is remarkably similar to retrieval from episodic memory. An alternative conclusion would be that perhaps, on grounds of parsimony, the qualitative distinction between episodic memory traces and semantic memory traces should be abandoned; for example, perhaps tagging theory should be favored over episodic theory. The present experiment is a convincing empirical demonstration of a point that tagging theorists have made on theoretical grounds: It is not necessary to invoke the concept of a unique episodic trace to accommodate the pattern of recognition failure of recallable words found in episodic memory experiments.

REFERENCES

- ANDERSON, J. R., & BOWER, G. H. A propositional theory of recognition memory. *Memory & Cognition*, 1974, 2, 406-412.
- BAHRICK, H. P. A two-phase model for prompted recall. *Psychological Review*, 1970, 77, 215-222.
- KINTSCH, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- MARTIN, E. Generation-recognition theory and the encoding specificity principle. *Psychological Review*, 1975, 82, 150-153.
- POSTMAN, L. Tests of the generality of the principle of encoding specificity. *Memory & Cognition*, 1975, 3, 663-672.
- Random House dictionary* (unabridged ed.). New York: Random House, 1969.
- REDER, L. M., ANDERSON, J. R., & BJORK, R. A. A semantic interpretation of encoding specificity. *Journal of Experimental Psychology*, 1974, 102, 648-656.
- TULVING, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- TULVING, E. Cue-dependent forgetting. *American Scientist*, 1974, 62, 74-82.
- TULVING, E. Ecphoric processes in recall and recognition. In J. Brown (Ed.), *Recognition and recall*. London: Wiley, 1976. Pp. 37-73.
- TULVING, E., & THOMSON, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.
- TULVING, E., & WISEMAN, S. Relation between recognition and recognition failure of recallable words. *Bulletin of the Psychonomic Society*, 1975, 6, 79-82.
- WATKINS, M. J., & TULVING, E. Episodic memory: When recognition fails. *Journal of Experimental Psychology: General*, 1975, 104, 5-29.
- WISEMAN, S., & TULVING, E. Encoding specificity: Relation between recall superiority and recognition failure. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 349-361.

(Received for publication August 9, 1977;
accepted September 12, 1977.)