Role of Semantic Memory In Storage and Retrieval of Episodic Information

Endel Tulving
University of Toronto

It is sometimes useful to assume that there are two separate memory systems. One, episodic memory, is involved in storage and retrieval of information about particular and concrete experiences localized in personal space and time; the other, semantic memory, deals with retention and utilization of general and abstract knowledge, which is independent of personal space and time, and of languages through which this knowledge can be transmitted. The major characteristics of the two systems are given in some detail in an earlier article (Tulving, 1972).

The distinction between episodic and semantic memory is useful to the extent that the postulated separation of the two systems helps us understand facts about cognitive activity of man that would not be readily understood otherwise. In this article, I will try to illustrate the usefulness of the distinction by considering certain recent findings from simple list learning experiments done in our laboratory.

We shall discuss two general classes of findings. One has to do with the determinants of memory traces of simple events, such as occurrences of words in to-be-remembered lists; the other is concerned with the retrieval of this stored information.

Encoding Operations and Storage

Let us begin by imagining a simple experiment in which two identical groups of people are asked to remember a list of unrelated words. Both groups are presented with identical sets of to-be-remembered materials. Each to-be-remembered word is, say, presented once, for the same length of time. Both groups are explicitly instructed that their memory for the material would be tested and that therefore they should try to do their best in learning the material. Both groups are tested after an identical retention interval, filled with identical activity, in an identical test. Thus, all variables with which classical experiments and theories of verbal learning have been concerned are held constant in our imaginary experiment: familiarity, frequency, meaningfulness, and other characteristics of the materials; ability of subjects; their intentions to learn; study time and frequency of presentation; length and nature of retention interval; the type of memory test.

Now, let us ask a question: is it conceivable that, with all these traditionally important conditions of learning equated, the two groups could differ in memory performance? If so, what other significant variables could the experiment manipulate?

Only ten or fifteen years ago, most students of verbal learning and memory would have been baffled by the question. Today, however, we know of many experimental situations in which it is possible to produce differences in memory performance under conditions where the learning situation for the two groups is equated with respect to the traditional variables. Experiments of this general kind—most, if not all, traditional variables held constant, something else manipulated, and differences in memory performance observed—have been described, among others, by Jenkins and his associates (e.g., Hyde & Jenkins, 1969, 1973; Jenkins, 1974; Tili & Jenkins, 1973; Walsh & Jenkins, 1973), by Shulman (1975), by Wood in this issue of the Bulletin de Psychologie, and by Craik in this issue of the Bulletin as well as elsewhere (Craik, 1985; Craik & Tulving, 1975).

The "something else" that is responsible for differences in memory performances
has been variously referred to as "orienting task," "incidental learning task," "processing task," or "encoding operation." Of these, the term "encoding operation" seems preferable because of its general applicability to a variety of relevant situations and because it lacks distracting connotations. The terms "orienting task" and "incidental learning task" imply that these activities entail processes that are somehow extrinsic to or different from the processes underlying encoding and trace formation. The term "encoding operation," on the other hand, implies a continuities among underlying processes. It accords with the hypothesis that intentional learning can be identified with or analyzed into more basic or elementary operations of the cognitive systems (Craik & Tulving, 1975; Ehrlich, 1970; Nadel, 1967).

There are many ways of encoding a given stimulus event, and many ways of experimentally manipulating encoding operations. Subjects can be asked to make conceptual judgments about the to-be-remembered words (e.g., Craik & Tulving, 1975; Hyde & Jenkins, 1969), they can be asked questions about these words (e.g., Schuman, 1974), the to-be-remembered words can be presented as related to context words (e.g., Tulving & Watkins, 1975), subjects can be led to expect a particular memory test (e.g., Tversky, 1973), they can be given instructions as to how to handle the material (e.g., Bower, 1973; Paivio & Caspary, 1973), they can be provided with information about certain aspects of the test (e.g., Watkins & Watkins, 1974), and so forth.

As an illustration of the effect of encoding operations, let us consider the results of an experiment summarized by Craik in Table I of his contribution to this issue of the Bulletin. With the experimenter holding constant variables such as the to-be-remembered material, ability of subjects, intentionality of learning, frequency of presentation, study time, retention interval, and the nature of the memory test, subjects' recognition hit rate for studied words was as low as 23% or as high as 61%, depending upon the encoding operation. Recognition performance was very low when the subjects, at the time of studying the to-be-remembered word, made a judgment about the type of letters (lower or upper case) in which the word appeared on the display device; it was very high when the subjects had to answer a question about the to-be-remembered word's belongingness to a particular conceptual category. Craik's experiment shows that the effect of encoding operations on memory for simple word-events can be very large. The magnitude of the effects makes it possible to think of encoding operations as one of the most important determinants of the memory traces of such events. It is not even inconceivable that in some sense encoding operations constitute the sole determinant of the composition of memory traces, and that the effects of other variables known through earlier research—for instance, meaningfulness of the material, study time, frequency of presentation, intention to learn, and the like—reflect the consequences of whatever encoding operations are brought into play in any given situation.

Classical theory of verbal learning has little to say about the effects of encoding operations, primarily because these effects were not known to earlier workers. We can classify the research on encoding operations as research concerned with transfer and transfer effects, but the classical theories of transfer, such as formal discipline and identical elements, would not fit the new findings. Nor would theories of specific transfer, couched in terms of particular stimulus-response relations, have much to say about contemporary experiments showing that the application of general existing skills and knowledge can produce sizable effects on the remembering of particular episodes.

Contemporary students of learning and memory have interpreted their results in a variety of ways. The greater effect of some encoding operations than of others has been explained in terms of depth of processing (Craik & Lockhart, 1972),
code elaboration ( Craik, in this issue of the Bulletin de Psychologie), general or personal meaningfulness of the task (Jenkins, 1974), imaginal versus verbal memory codes (Paivio & Caspary, 1973), qualitatively different types of rehearsal of the material (Watkins & Watkins, 1974), and other similar notions. None of them is clearly superior to others, and all are, at this stage, somewhat vague and general. This is not surprising, since even today the experimental facts are still novel, and there has not been sufficient time for the integration of data and theory. The important book by Anderson and Bower (1973), for instance, which otherwise presents a comprehensive theory of memory, does not even refer to any experiments on encoding operations.

We will return to the problem of interpretation of these findings after we have considered the second class of experiments relevant to the present discussion. These have to do with effectiveness of retrieval cues, and particularly with the interaction between encoding conditions and cue effectiveness.

Encoding Specificity and Cue Effectiveness

Retrieval of stored information never takes place spontaneously. It is always initiated by a stimulus, a query, or a cue. Hence we think of retrieval as being critically dependent upon an interaction of information from two separate sources, (a) the memory trace, and (b) the rememberer's cognitive environment at the time of retrieval (Tuving, 1974, 1975). In the same way in which the memory trace of a to-be-remembered word can be influenced by experimentally manipulating encoding conditions, the cognitive retrieval environment can be influenced through experimental manipulation of retrieval instructions and specific retrieval cues.

In the course of experimental work that has been done to evaluate theories of retrieval (e.g., Thomson & Tulving, 1970; Tulving & Thomson, 1973) it has become clear that the effectiveness with which an extralist cue (a verbal item that has not been an explicit part of the to-be-remembered list) can evoke the memory of a studied list word is not always predictable from pre-experimental characteristics of the two words. Thus, for instance, whether the extralist cue word “white” can help the learner to remember the list word BLACK depends on particular experimental conditions. Similarly the cue word “black” can be differentially effective in reminding the learner of the nominally identical word BLACK from the list studied earlier, depending on specific experimental conditions.

These observations are difficult to understand in terms of classical theory, and are therefore of some interest. We will here briefly elaborate on these two kinds of experimental findings. One has to do with the effectiveness and lack of effectiveness of extralist cues that are strong associates of to-be-remembered words; the other concerns recognition failure of recallable words.

In one experiment of the first kind (Thomson & Tulving, 1970, Experiment 2) subjects studied to-be-remembered words, such as CHAIR, under one of two conditions. In one, the to-be-remembered words were presented singly, one at a time, and the subjects expected a free-recall test. In the other, the to-be-remembered words appeared as members of pairs, the other member of each pair being a word weakly associated with the to-be-remembered word. (I will refer to these words as "weak" cues.) There were two test conditions, crossed with the two study conditions. In one test, subjects were given weak cues as aids to retrieval, in the other they were provided with strong extralist associates of the to-be-remembered words.

The design of the experiment (Groups 2, 3, 5 and 6, List 3, in Experiment 2 by Thomson & Tulving, 1970) is schematically presented in Table 1, together with the results. In all four conditions, the target word was nominally the same (CHAIR).

Table 1 about here
The fact that the weak cue "glue" was very effective (62%) in retrieving the target word CHAIR in Condition III, but less effective (43%) in Condition I, is easily explained by any theory. It represents the well known case of association by contiguity: in Condition III, "glue" was the 1st cue for CHAIR and paired with it in the study list, whereas in Condition I it was not present in the list. On the other hand, the fact that the strong cue "table" was quite effective in Condition II (68%), and quite ineffective in Condition IV (23%) does not find a ready explanation in most extant theories. Explanation of these results in terms of association by contiguity is precluded in both conditions, because the cue word "table" was not present in the list in either case. The other kind of association known in classical theory, association by similarity, also fails as an explanatory device here, since the similarity between "table" and CHAIR was nominally the same in both cases.

The results of the experiment summarized in Table I illustrate an "encoding specificity phenomenon": The effectiveness of a given retrieval cue ("table") in evoking the memory of a given list word (CHAIR) depends on the specific conditions under which the target word was encoded into memory (Tulving & Thomson, 1973). Although the stimulus word CHAIR was exactly the same in both Conditions II and IV, the fact that one and the same cue ("table") varied greatly in its effectiveness can only mean that the memory trace of CHAIR was different in the two cases. Thus, it is not the pre-experimental similarity of the cue and the target that determines the effect of the cue, but rather the similarity between the cue (or, more precisely, its encoded version) and the trace of the target word resulting from the specific encoding of the target word in a particular situation.

The major difficulty the classical theory has in handling this experimental result has its source in the assumption that each familiar word is represented
in "permanent memory" in a single, relatively fixed form or location which is defined by its relations to other components of the overall structure. The appearance of the word, such as CHAIR, in a to-be-remembered list, according to the theory, results in the activation, marking, or tagging of the word's representation in the permanent memory. At retrieval, the pre-experimentally established connections between "table" and "chair" are used to gain access to CHAIR through "table" as cue, and the state of activation of, or the list-tag attached to, the representation of CHAIR is examined with a view to its recallability. But such a theory cannot readily account for the fact that the effectiveness of an extralist cue depends upon the conditions under which the target word was encoded. Why was the cue "table" quite effective in Condition 11 and quite ineffective in Condition IV in the experiment summarized in Table 17?

In an attempt to rescue the classical theory, some theorists have argued that most words have multiple representations in memory, each representation corresponding to a different semantic sense (Anderson & Bower, 1974; Fender, Anderson, & Bjork, 1974). One implication of this position is that the effectiveness of extralist retrieval cues should be more pronounced and less dependent on encoding conditions for target words that have fewer representations in memory. In support of this hypothesis, Fender, Anderson, and Bjork (1974) have shown that extralist retrieval cues, in the Thomson and Tulving (1970) paradigm, are much more effective for low frequency than for high frequency words. But Fender, Anderson, and Bjork's explanation runs into difficulties with the data showing specific encoding effects in experiments with to-be-remembered words that have only a single semantic meaning. Thus, for instance, Bartley, Bransford, Franks, McCarrell, and Nitsch (1974) presented the target word PIANO to subjects as a part of a sentence. When the sentence was, The man lifted the piano, then the retrieval cue "something heavy" was quite effective, whereas the cue "something with a nice sound" was quite ineffective. On the other hand, when the subjects had studied the target word as a part of the sentence, The man tuned the piano, then the latter cue was effective and the former one was ineffective. These and other similar results (e.g., Anderson & Ortony, 1975) suggest that specific encodings can be experimentally created for all kinds of target words, and not only for those that have a multitude of semantic senses.

Recognition Failure of Recallable Words

Another defense of the classical theory might take the form of the assumption that a particular encoding operation, or the processing of the to-be-remembered item in a particular context, changes the associative connections that the to-be-remembered item has with other components of permanent memory. The reduced effectiveness of a strong semantic associate of the target item might then result from the impaired access routes to the target item's representation in permanent memory. Existing data do not show any such impairment in a free-association test (Tulving & Thomson, 1973; Watkins & Tulving, 1975), but a more direct test of the hypothesis is possible. In such a test, we use as retrieval cues literal copies of target items which, according to classical theory, provide direct access to the item's representation in memory. The access problem should thereby be circumvented. The question then is: Does the specific encoding of the target item change its recognizability? Experiments designed to explore this question have led to demonstrations of recognition failure of recallable words, another "encoding specificity phenomenon" on which classical theory founders.

The experimental paradigm under which recognition failure of recallable words has been observed has the following components: (a) The to-be-remembered word is presented on a single trial together with another word, the list cue. (b) The subject studies the to-be-remembered word in the expectation that its recall is
going to be tested in the presence of the list cue. (c) Following the study trial, the subject is given a recognition test in which copies of to-be-remembered words are presented together with distractor words, and the subject is required to identify the to-be-remembered words. (d) After the recognition test, all list cues are presented to the subject, and he is asked to recall the corresponding to-be-remembered words.

Recognition failure of recallable words, under the conditions of the paradigm just described, was first described by Tulving and Thomson (1973). The Tulving and Thomson experiments involved the four steps as just enumerated, although the general procedure was somewhat more complex. In the first experiment they described, recognition hit rate was found to be .23, and cued recall was .61. Thus there were many words that the subject could recall but could not recognize. Subsequent experiments have shown the phenomenon of recognition failure of recallable words to be very robust (Tulving, 1974b; Watkins & Tulving, 1973; Wiseman & Tulving, 1975). Recently Tulving and Wiseman (1975) have summarized evidence from 12 different previously published experiments, representing 40 different experimental conditions. A certain amount of recognition failure of recallable words was present in every one of the 40 experimental conditions. The extent of the recognition failure of recallable words in a given condition was highly correlated with the overall level of recognition in that condition; this correlation was independent of the procedural details used in the experiments. Since the classical theory denies the possibility that a recallable item cannot be recognized, the finding of recognition failure of recallable words proves the classical theory wrong.

Episodic and Semantic Memory

It seems that the major source of the difficulty the classical theory has with the facts briefly described here lies in its basic assumption that there is a single "permanent" memory, and that hence information about personally experienced episodes is stored in the same memory structure that retains general knowledge of the world and the symbolic systems for representing this knowledge. A corollary of this assumption is that a word, or a certain semantic sense of a word, has but a single form of representation in this permanent memory. Experimental facts of greatly attenuated effectiveness of strong associates as retrieval cues, and recognition failure of recallable words following certain encoding operations performed on to-be-remembered words, strongly imply that the internal representations of the word the subject cannot recognize is different from the representation of the word that he can recall (cf., Martin, 1975). This conclusion suggests that we should abandon the assumption of a single memory system.

The experimental findings we have discussed become more readily understandable if we assume a distinction between episodic and semantic memory. Information stored in semantic memory is used in the course of encoding operations, to construct a unique trace of an event in the episodic system, and it is again used, at the time of retrieval, to interpret the retrieval query and cues. But the episodic trace as such exists independently of the semantic system, and may remain unaffected by the activity of the semantic system.

The identity and characteristics of a to-be-remembered item only partly determine the composition of its trace in episodic memory. Other, and probably more important, factors that shape the trace are subsumed under the concept of encoding operations. It is in this sense in which we can think of the trace of an item as "a record of various pattern-recognition and interpretative analyses carried out on the stimulus event" in its particular context (Craik &
Tulving, 1975, p. xxx). The characteristics of the trace do not seem to depend as much on the "intensity" of these analyses as they do on their qualitative nature. It is the qualitative nature of encoding operations that determines the "depth," "spread," and "elaborateness" of the final product of the encoding of an event. This final product can be described in terms of features, attributes, or "trace elements" (Tulving & Wicker, 1974; Tulving & Watkins, 1975).

Retrieval, or recency (Tulving, 1975), of an experience is conceptualized as a joint product of information from two sources, the memory trace and the retrieval cue. The mechanism of retrieval is not yet known, but we can think of it as a matching or pattern completion operation (Kintsch, 1974). The success of retrieval depends on how well the information in the retrieval cue matches the information contained in the episodic trace of the item, rather than on the characteristics of the cue and target items as defined in semantic memory. This is why memory performance, in any given retrieval situation, depends so greatly on the encoding operations. Retrieval cue has to contact the trace of the episode; access to the representation of any item in semantic memory is not sufficient.

Given these general ideas, it is not too difficult to make sense out of the observation that strong retrieval cue is quite effective in evoking the memory of a target word when the latter was presented as a member of a list to be recalled under free recall conditions, but not effective after the subject had studied the target word as a member of a cue/target pair. The traces of the two events in the episodic system are quite different because of different study contexts and test expectations on the part of the subjects. The finding of recognition failure of recallable words, too, makes sense. The informational contents of the trace of the cue/target pair of words overlap more with the information contained in a copy of the list cue than with the information contained in a copy of the target word. Obviously, this interpretation does not explain the facts completely, but it can be elaborated and it does make it possible to ask further meaningful questions about the facts.

Kintsch (1974) has recently proposed a theory of episodic memory that can account for the findings of many list-item experiments, including phenomena of encoding specificity. The distinction between episodic and semantic memory plays a useful role in the theory. In an entirely different context, recent clinical observations have suggested that the distinction is heuristically useful in understanding mental functioning of amnesic and agnostic patients (Kinsbourne & Woods, 1975; Warrington, 1975). Although current ideas about different memory systems and their interrelations may change as we learn more about cognitive activity of man, at the present time they appear to be not only necessary but promising.

Summary

In this paper, the distinction between episodic and semantic memory has been applied to the problem of interpretation of experimental results from two broad classes of experiment. In one, effects of different encoding operations performed on to-be-remembered stimulus events were considered in simple list learning experiments. In these experiments, to-be-remembered words, whose nominal identity is held constant, are presented for study under conditions in which all classical variables—such as ability of subjects, frequency of presentation, study time, intention to learn, nature and length of the retention interval, and the memory test—are held constant. The major finding is that differences, sometimes large differences, are observed in memory performance, depending upon encoding operations that the subjects have been induced to perform on the to-be-remembered material. The other class of experiments demonstrates the differential effectiveness of retrieval cues; in particular, the interaction between cues and conditions.
under which a particular word event is encoded into memory. Classical theory has difficulty explaining these data, primarily because it assumes a unitary memory structure in which words are represented as fixed elements and in which episodic occurrences of these words are registered in the form of activation or tagging of the words' representations. Given this assumption, certain encoding specificity phenomena, such as recognition failure of recallable words, are difficult to explain. A more promising approach to understanding these phenomena is provided by a view that assumes a distinction between semantic and episodic memory systems. According to this view, skills and knowledge stored in semantic memory are used in encoding operations for the construction of unique memory traces in the episodic system. The exact composition, and hence the exact conditions under which the trace is retrievable, are to a large extent determined by the qualitative nature of the encoding operations. The specific information stored about an event determines the effectiveness of retrieval cues. Retrieval queries and cues are cognitively analyzed in terms of the contents of the semantic system, and the resultant retrieval information is combined with the trace information from the episodic system to create the memory of an experienced event. Thus, memories of events are stored in the episodic system, but both encoding of memory traces and their retrieval may depend greatly on the semantic system.

References


Hyde, T. S., & Jenkins, J. J. Recall for words as a function of semantic, graphic,


