COGNITIVE PSYCHOLOGY 7, 167–180 (1975)

# On Putting Apples into Bottles— A Problem of Polysemy

RICHARD C. ANDERSON AND ANDREW ORTONY

*University of Illinois at Urbana-Champaign*

An experiment was designed in such a manner that a sentence could be recalled given a certain cue only if the subject's encoding of the sentence included details and involved distinctions in the senses of words which could not have been part of the correct dictionary readings for these words. The most plausible interpretation of the results is that comprehension of a sentence entails constructing a particularized and elaborated mental representation, and that this process depends more heavily on knowledge of the world and analysis of context than is generally appreciated. It is claimed that existing associative or semantic network theories would be strained to accommodate the data.

A gedanken experiment: a group of subjects is presented with the sentence *The container held the apples;* another group sees *The container held the cola.* Later each group is given two retrieval cues, *bottle,* and *basket.* There are two questions: which cue will best facilitate recall for each group, and why? This paper addresses these questions.

Our thesis is that sentence comprehension and memory involve constructing particularized and elaborated mental representations, and that network models currently have no satisfactory way of accounting for this. That language comprehension is a constructive, elaborative process has been demonstrated already (cf. Bransford, Barclay & Franks, 1972; Bransford & McCarrell, in press). Previous investigators have stressed the importance of a person's using context to get a holistic interpretation of the to-be-comprehended sentences. Our emphasis is on the resulting mental representation. Specifically we claim that the representation is generally more detailed than the words in the utterance might appear to entail; that the words only loosely constrain the representation; and that one's store of knowledge about the world and analysis of context are

The research reported herein was supported in part by the Navy Personnel Research and Development Center and Advanced Research Projects Agency under Contract N61339-73-C-0078. We gratefully acknowledge Peter Zych, Claire Lieberman, and Michael Riviere for their assistance in collecting and analyzing the data. Thanks are also due to John Bransford, William Brewer and Henry Halff for their helpful comments.

Requests for reprints should be sent to Richard C. Anderson, Department of Educational Psychology, University of Illinois, Urbana, IL 61801.

heavily implicated in sentence comprehension and memory. Two corollary notions are that only in a vague, abstract sense could words be said to have fixed meanings and that it is impossible that the sense of an utterance could consist solely of a concatenation of the dictionary readings of its individual words.

Another consequence of our view is that the structure required to relate the concepts utilized in understanding language is context-dependent. Different contexts may necessitate different and in some cases even incompatible structures which cannot coexist. From this it would follow that our knowledge is not structured in a static manner but is reorganized during cognitive processing.

The intellectual ancestry of the view that words have flexible meanings can be traced to the late work of Wittgenstein, characterized by the dictum "don't ask for the meaning, look for the use." He argued persuasively that it is often impossible to state the necessary and sufficient conditions for the use of a word in ordinary discourse. Rather, he maintained that the various uses of a term are related by "family resemblance." Wittgenstein (1963, pp. 31–32) wrote:

> Consider for example the proceedings that we call 'games.' I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all?—Don't say: "There *must* be something common, or they would not be called 'games'"—but *look and see* whether there is anything common to all.—For if you look at them you will not see something that is common to *all,* but similarities, relationships, and a whole series of them at that . . . a complicated network of similarities overlapping and criss-crossing; sometimes overall similarities, sometimes similarities of detail.
>
> I can think of no better expression to characterize these similarities than "family resemblances"; for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc. etc. overlap and criss-cross in the same way.—And I shall say: 'games' form a family (italics in the original).

Linguists attacking semantic features theories (which we regard as vulnerable to our arguments in much the same way as network theories) have made a similar point. Most recently Labov (1974) has demonstrated the "fuzziness" of the word *cup.* Weinreich (1966) claimed that the attempt to deal with the problem of polysemy by providing for distinct "readings" was doomed to failure. Even when a word would be said to have the same reading, Weinreich argued its sense would change from sentence to sentence.

For an illustration Weinreich used the verb *to eat.* Consider the phrases *eat steak, eat soup,* and *eat an apple.* Eating a steak requires a knife and fork. Soup is sipped with a spoon. Commonly an apple is eaten without a utensil. In each case the actions of the lips, teeth, and tongue are different. Further variations in sense are introduced when the agent is considered. Compare

*Lord Raleigh ate the soup.*
and
*The tramp ate the soup.*

or contrast the sense of *eat* in,

*The executive ate the steak.*
*The baby ate the steak.*
*The dog ate the steak.*

Each of these sentences gives rise to different suppositions about location, circumstance, manner, instrumentality, and antecedent and consequent conditions. The general point is that a word could have different meanings in a very large number of the sentences in which it might appear, even when there is some "core meaning" as in *eat*.

The *a priori* arguments for a dynamic knowledge structure are also, we believe, compelling. Consider, for example, the case of *piano* in the two sentences:

*Pianos can be pleasing to listen to.*
*Pianos can be difficult to move.*

Both sentences we presume to be true yet in the first case the fact that pianos are normally heavy pieces of furniture is irrelevant while in the second case it is all-important. And, of course, the reverse is true, namely that in the first case the "musical instrumentness" of the piano is the key feature while in the second it might as well be a lump of concrete. So, our point is that in one context piano is a member of the same category as, say, harmonica while in another it is certainly not. In the latter case perhaps sofa would be a cohyponym.

Now one might wish to argue that this is only to point out the need for cross-classification, that piano must appear as a member of both categories. But this argument turns out to be the thin end of a rather fat wedge. There are so many ways in which every object can be classified that most of the supposed advantages of a hierarchically structured classification scheme would be lost. If we want to classify fruits alphabetically, we can; if we want to classify them by color, we can; if we want to classify them by taste, or size, or shape, we can. It would seem unreasonable to suggest that they are permanently classified in all these ways simultaneously. But, if differing classifications like these were required by some context at some time they would have to be permanently classified in these ways, and in all other conceivable ways as well.

There may in some sense be preferred schemes based perhaps simply on frequency of use but this is not the same as saying that they are fixed and static. Indeed, there are cases in which only the context will help us to determine how to classify an object. Is, for example, a grandfather

clock a timepiece or an antique? Is black a color or is it not? The answer, we claim, is that such questions are pointless because the context will enable the construction of an appropriate representation, thereby effectively removing the choice.

It could be argued that we have stretched the concept of meaning too far, that the different interpretations that can be given to words in various contexts come to no more than optional embroidery. We believe the contrary: that elaborating and particularizing are routine and predictable in language comprehension. The purpose of the experiment reported in this paper is to provide an empirical test of these claims.

The strategy was to provide a cue which it was judged would be effective for retrieval of a previously-exposed sentence provided that the mental representation which evolved from the sentence included certain details, but ineffective if the representation embodied only the sentence's "literal meaning." For each target sentence there were two cues, neither of which was a constituent of the sentence. One cue was intended to be semantically relevant, considering the sense of the sentence as a whole, the other semantically less relevant. To illustrate, the idea was that *hammer* would be a more effective cue than *fist* for recall of *The accountant pounded the stake,* but less effective for recall of *The accountant pounded the desk.* A preliminary experiment showed that sentences were recalled more than one-and-one-half times as often given semantically "close" than semantically "remote" cues, min $F'$ $(1,47) = 22.98$, $p < .01$.

This research assumes that the semantically most relevant cue would not be differentially effective unless it related to the representation for the sentence constructed at the time of initial encoding (cf. Tulving & Thompson, 1973). For instance, it is assumed that *fist* would not be an effective cue for the second sentence above unless a fist were encoded as the instrument at the time the sentence was originally processed. That is not to say that we have to guess precisely the instrument the subject will have encoded, only that we provide a cue referring to an instrument (or location, manner, etc.) in the same semantic neighborhood. For instance, with respect to the first sentence above *hammer* should be an effective cue even if the subject had actually encoded the instrument as a hatchet or an axe.

What would a network theory predict about performance on the recall task that has just been sketched? We shall attempt a detailed answer to this question for the theory proposed by Anderson and Bower (1971, 1972, 1973), which they have named Human Associative Memory, or HAM. Anderson and Bower maintain that a proposition in memory can be represented as a tree-like structure consisting of nodes connected by labeled associations. While most of their attention is devoted to single

sentences, they believe that all of the knowledge one has acquired from experiencing events and hearing or reading can be represented as a network of interlocking propositions. An important assumption, one which Anderson and Bower lean on heavily when making experimental predictions, is that the associative connections in the network are independently formed or not formed; independently forgotten or not forgotten (cf. 1973, p. 284).

HAM does not currently include mechanisms for retrieval unless the cue is a constituent of the to-be-recalled proposition. This remark is not in itself intended as a criticism. Nevertheless, a complete theory of human memory would have to deal with more than verbatim recall given verbatim stimuli. Thus it is not unreasonable to ask whether the theory is readily extendable. Here we give what we consider to be the most simple and natural extension of HAM. Alternatives will be discussed later.

According to our reading of the Anderson and Bower theory, the probability that a certain cue will allow retrieval of a target sentence, $t$, ought to equal the probability that there exists an intact path from the cue to any constituent of the target sentence. Following Anderson and Bower, sentences were analyzed into just two main constituents, subject and predicate. Letting $s$ be the probability of an intact path from the cue to the subject of the target sentence and $p$ the probability of an intact path to the predicate, and allowing for the possibility that paths to both elements exist, the prediction from HAM can be expressed as follows:

$$t = s + (1 - s)p. \tag{1}$$

However, as will be explained in a moment, because of the way in which the experiment was designed, a path to either the subject or the predicate, but not both, could manifest itself. Hence, in this special case the prediction from HAM is,

$$t = s + p. \tag{2}$$

Our view is that the senses of the individual words interact in the process of constructing an elaborated meaning for an utterance. Hence, we predict,

$$t > s + p. \tag{3}$$

To test this hypothesis, a number of sets of four sentences were created. Below is an example.

A. *Nurses are often beautiful.*
B. *Nurses have to be licensed.*
C. *Landscapes are often beautiful.*
D. *Taverns have to be licensed.*

When the cue was *actress,* A was the target sentence, B was the subject control sentence, C was the predicate control sentence, while D was called the "double control." On the other hand when the cue was *doctor,* then B was the target, A the subject control, D the predicate control, and C the double control. From each set of sentences any one person saw what for him was a target sentence and a double control sentence; his performance to the alternate cue provided the subject control and predicate control for another person.

It was assumed that there was a zero probability of an intact path from the cue to either constituent of the double control sentence. A violation of this assumption would bias the experiment in favor of the associative theory.

It may be objected that the experiment did not respect the integrity of the sentence. Let us be clear that in our view that the sense of a sentence is a particularized and elaborated mental construction to which words flexibly contribute meanings, there is an indissoluble integrity to a sentence. However, this is definitely not a feature of network theories, at least not the version espoused by Anderson and Bower (see 1972; 1973, pp. 337–341). In their model a sentence is represented as a colligation of independently associated elements; it has no special semantic or memorial cohesion. If there is an explicitly formulated associative theory that includes a principle of "belongingness" it is unknown to us. Anyway, such a theory would be less associative and more constructive in character. There would be nothing left to quarrel about but taste in jargon.

## METHOD

### *Subjects*

Participating were 84 undergraduates enrolled in four sections of an introductory educational psychology course. The experiment was run on a group basis during regularly scheduled class periods.

### *Materials*

Fourteen sets of four sentences were created. The subject noun of every sentence was a general term used in such a way that it evidently referred to indefinitely many members of some class. Two of the sentences within each set contained the same subject noun. While these two sentences were designed to cause different interpretations of the subject noun, in most cases it could be interpreted as referring to the same persons or objects regardless of which of the two sentences it appeared in. To put this another way, the two sentence contexts tended to effect the intension rather than the extension of the common term. As already

explained there were two distinct ways in which the sentences in each set could be classified as target, subject control, predicate control, or double control, depending upon which of two cues were considered. Here is one additional example of a sentence set:

*Coins can be flipped.*
*Coins are difficult to forge.*
*Pancakes can be flipped.*
*Signatures are difficult to forge.*

The cues were *odds* and *dollar.*

*Design and Procedure*

A target sentence and a double control sentence from each set of four sentences were assigned to List A; the other pair of sentences from the set (also a target and double control when seen from the perspective of the other cue) was assigned to List B. The assignment of sentence pairs to lists was made in a way judged likely to minimize intralist intrusions. Subjects were exposed to either List A or List B. In order to further reduce the likelihood of intrusions and also to increase level of learning, each list was split into 14 sentence (seven target-double control pairs) sublists. The sublists were parallel in the sense, for instance, that Sublist A1 contained sentence pairs from the same sentence sets as Sublist B1.

Subjects received two presentations of Sublist 1, then the tests over this sublist, then two presentations of Sublist 2, and finally the tests over this sublist. The sentences were presented by a tape recorded male voice at an eight second rate. Subjects were simply told to learn the sentences. There were two tests per sublist, each of which consisted of a mimeographed booklet with one cue word on each page. The first test assessed learning. The cues were the subject nouns of the presented sentences; the task was to complete each sentence. The second test consisted of the remote and close cues.

Recall instructions for the second test indicated that each cue would be related to one of the sentences and that sometimes the relationship would be obvious, sometimes not so obvious. The instructions also indicated that each sentence would be the correct answer twice. Instructions for both tests stated that while subjects should try to recall sentences exactly paraphrases would be counted correct, and that parts of sentences should be given even when the whole could not be remembered. The tests were subject paced.

A blocking procedure was used to minimize recall from short-term, nonsemantic memory. The sublists were randomly divided into two, seven-sentence blocks. The same order of blocks was retained in each list presentation and test; however, there was a different random order

of items within blocks on each occasion, subject to the constraint that for each randomization of an A sublist or test there was an exactly parallel randomization of a B sublist or test, such that the comparable items from the two sublists or tests appeared in identical positions. Overall there were two completely different list presentation orders and four completely different orders of each test.

Whole, intact classes were randomly assigned to lists and presentation orders. Subjects within classes were assigned to test orders by distributing test booklets from randomly ordered stacks.

*Scoring*

Recall was scored by the sentence in order to have roughly comparable measures for the two tests. There were four different levels of scoring. Level I was an exact verbatim reproduction. The only permissible deviation was the obvious abbreviation of an individual word, for instance, TV for *television*. Level II was a verbatim reproduction except for the following: deletion, addition or changes of articles; changes in number; changes of tense; changes of auxiliary verbs; meaning-preserving changes of prepositions; meaning-preserving changes of word order; and contractions or expansions. Level III permitted substitutions of synonyms, close superordinates (e.g. *officer* for *lieutenant*), hyponyms (e.g. *pistol* for *gun*) and close cohyponyms (e.g. *shirt* for *sweater*) for any of the substantive words of the original (cf. Anderson, 1972, 1974). At Level IV the reproduction was unmistakably derived from a presented sentence but the meaning was distorted by, for instance, reversal of agent and object, omission of a negative, omission of an element, substitution of an incorrect element, or the introduction of new material. The score at any level included all sentences which were counted as correct at preceding levels. Several judges scored the recall protocols after studying a two page manual and passing a 25 item test that involved classifying instances and noninstances of each of the scoring categories. Another judge independently rescored 15 randomly selected protocols. The interrater reliability coefficients for the various levels ranged from .97 to .99.

## RESULTS

Table 1 shows mean numbers of sentences recalled divided by total number of sentences. As can be seen, proportion recalled increased with the leniency of the scoring procedure. In order to adjust for possible differences in the learnability/memorability of sentences, each subject's score for each type of sentence was divided by the number of sentences of that type that he had recalled on the first test. The means of these

TABLE 1
MEAN PROPORTIONS OF SENTENCES RECALLED

| Sentence type | Scoring procedure[a] | | | |
|---|---|---|---|---|
| | I | II | III | IV |
| Target | .36 | .43 | .66 | .69 |
| Subject control | .27 | .31 | .44 | .46 |
| Predicate control | .05 | .07 | .11 | .12 |
| Double control | .00 | .01 | .01 | .01 |

[a] See text for description of procedures.

proportions appear in Table 2. Each proportion involves numerators and denominators to which the same scoring procedure has been applied. Since it was a rare event for a subject to respond correctly on the second test without having responded correctly on the first test, the figures in Table 2 can be read as conditional probabilities.

Analyses of variance were computed in which the random factors were subject and sentence set, and the fixed factor of interest was target ($t$) versus the sum of subject control and predicate control ($s + p$). The dependent variable was proportion of sentences recalled given learning. When there was strict verbatim (Level 1) scoring, a significant, $F(1,82) = 8.41$, $p < .01$, result was obtained in the analysis in which subject constituted the random factor but not in the one in which sentence set was the random factor, $F(1,13) = 1.78$, so of course the minimum quasi $F$ ratio (cf. Clark, 1973) was not significant. At every other level of scoring $t$ was significantly greater than $s + p$. For instance, considering the Level III scoring, min $F'$ $(1,22) = 12.89$, $p < .01$. There were no significant main or interaction effects involving presentation order.

TABLE 2
MEAN PROPORTIONS OF SENTENCES RECALLED GIVEN LEARNING

| Type of sentence | Scoring procedure[a] | | | |
|---|---|---|---|---|
| | I | II | III | IV |
| Target | .67 | .68 | .73 | .73 |
| Subject control | .47 | .48 | .49 | .50 |
| Predicate control | .11 | .13 | .12 | .13 |
| Double control | .01 | .01 | .01 | .01 |

[a] See text for description of procedures.

## GENERAL DISCUSSION

The present research suggests that sentence comprehension and memory involve constructing particularized representations whose sense cannot be reliably predicted from the dictionary readings of the constituent words (see also Anderson & McGaw, 1973). In our view, the most plausible explanation for the results is that words loosely and flexibly constrain the building of a representation; that intimately involved in language comprehension is knowledge of the world as well as knowledge of the language (assuming one wishes to make this distinction at all); and that an essential process is analysis of context. Effects were obtained in the experiments described herein by manipulating just intrasentence context. Obviously, though, in the ordinary course of affairs extrasentence linguistic and nonlinguistic context play a large role.

Existing theories of semantic memory and language comprehension cannot do a good job of accounting for our results. In this discussion most attention will again be paid to HAM, though the general line of argument applies to several other models as well. As already explained, according to our interpretation HAM predicts that the probability of recalling a target sentence given a certain cue is equal to the probability of an intact path from that cue to any constituent of the sentence. The evidence failed to confirm this prediction but not, we believe, because of defects which could be remedied with small repairs or simple extensions.

It might be argued that HAM could be made to cope with the present data by incorporating a different decision rule for recall. Instead of a path to either constituent, a subject might recall a sentence only if there were paths to both. So, for instance, *fist* would be an effective cue for *The accountant pounded the desk* only if it were associatively related to both *accountant* and *pounded the desk*. How in HAM could *fist* be related to both sentence constituents? By strict interpretation *fist* could have a direct connection only if reference had been made to a fist of *that* accountant—in which case *fist* would be part of either the subject substructure or the predicate substructure. Under this interpretation the joint path criterion is indistinguishable from the single path criterion and, incidentally, the main part of our argument is conceded as well.

The idea of a joint criterion seems to presuppose a traditional associationism in which words (or ideas) are more or less strongly associated. It has to be supposed, for example, that *fist* in general is associated with *accountant* in general, *pound* in general, and/or *desk* in general. HAM becomes unrecognizable when so loosely interpreted. The distinction between class and specific instance has to be obliterated and the labels on the links have to be ignored. In short, the joint criterion rule would not appear to be a promising augmentation of HAM to account for the superior recall of target sentences observed in the present experiment.

Semantic theories (cf. Katz, 1972) typically attempt to handle polysemy by providing a different reading for each distinct sense of a word. This is the tack which Anderson and Bower take; each polysemous word relates to a number of different ideas (1973, p. 207). For the purposes of representing the crude distinctions between different meanings—as for *ball,* a dance, and *ball,* a round thing—this method is adequate. But the problem is not with the macro-distinctions embodied in homonyms but with micro-distinctions, like *ball* in *The quarterback threw the ball, Arnold Palmer lost his ball, The guard dribbled the ball* and so on. Consider one of Anderson and Bower's illustrations (1973, p. 195): *John kicked Bill.* It is said that *kicked* is a token which points to, or accesses, a definition in HAM, the English paraphrase of which is an "animal moving its foot forcefully against an object." Now try, *The donkey kicked Bill.* Presumably the same generic definition of *kick* will be accessed, unless we are to have as many definitions of *kick* as there are animals that can kick and objects that can be kicked. From whence will come the distinctions between John kicking and the donkey kicking which any reader or listener would make? If this subtle form of polysemy is to be handled in the fashion Anderson and Bower suggest, then they would be forced to include an unmanageable number of word to idea links, one for each of the indefinitely many fine gradations in meaning that a word in context can have.

Anderson and Bower believe that particularization does occur. Indeed, few have given such eloquent affirmation (1973, p. 460):

> Our words spoken to a listener are like the cryptic directions a playwright provides for a play director, from which a competent director is expected to construct an entire setting, an expressive mood, or an action episode. To illustrate, in the course of reading a story, you might read the sentence "James Bond ran to his car and drove to the casino." As you read you can concretize that sentence by bringing to bear all sorts of facts and sensory images about running, about getting into cars, about driving, and so forth. . . . What the sentence does is merely mention a couple of signposts (source, instrument, goal) along the way in the description of an event sequence; the listener interpolates or fills in all the interstitial events between the mentioned signposts.

Nevertheless, HAM does not include a mechanism for filling in the "interstitial events" between the "signposts" mentioned in sentences.

If such a model is going to be able to accommodate the elaborated and instantiated representations which have been shown here and elsewhere (Anderson & McGaw, 1973; Rosch, 1973), it appears to us that it will have to make deeper and more intricate use of its knowledge base. Suppose that on encountering *Nurses are often beautiful* an intersection search (cf. Quillian, 1968, 1969) were conducted starting from *nurse* and *beautiful,* and with *Nurses have to be licensed,* from *nurse* and *license.* It would be reasonable to expect that the former would be more likely to

arrive at *woman* and the latter at *health-professional.* The mechanism would have to incorporate these new concepts into the representation. If such a procedure were to be applied recursively, then perhaps successively refined representations could be elaborated, each of which might be regarded as having involved another level of processing. There would thus be two potential advantages: first, it might account for elaboration; and, second, it might lead to some new ideas about levels of processing.

Attractive as this may be, HAM would have to pay a high price for it. To make inferences from existing knowledge at the time of encoding new experiences would violate one of HAM's most basic assumptions, namely that there is a strategy-free component of memory. HAM, having extracted an all-too-close-to-surface representation would be forced to interpose some interfacing, inferential mechanism in order to provide a detailed memorial representation. So perhaps, as they recognize, it is "impossible to extricate memory from such matters as problem-solving and inference," in which case it may be that their "whole theoretical enterprise will come crashing down on [their] heads" (1973, p. 141). The extent to which what one knows affects what one understands and stores should not be underestimated.

Theories which assume fixed hierarchical relationships among concepts will have trouble explaining our data. A theory which does this quite explicitly is that of Collins and Quillian (1969, 1972). Static (though not necessarily hierarchical) organization of the knowledge base is a feature of all current network theories.

The network model of Collins and Quillian has as an essential element a direct pointer from a concept to its superset. The problem of determining what this superset is already arises when they note (1969, p. 262) that there may be unutilized distinctions:

> . . . hierarchies are not always clearly ordered, as exemplified by dog, mammal, and animal. Subjects tend to categorize a dog as an animal, even though a stricter classification would interpose the category mammal between the two.

The expression "a stricter classification" seems to suggest that there is in some sense a "correct" structure. On the model they propose one would expect that if someone does in fact have the concept *mammal* then it must be represented as the superset of, for example, *dog.* The theory is not sufficiently flexible to be able to have it both ways; our contention is that any theory of semantic organization must in fact be sufficiently flexible to have it both ways—and more.

If every concept is to point to its superset in a direct and rapidly accessible manner then the hierarchical structures in the network must be fixed and every element in it must bear the same constant relationship to every other element in it, regardless of the context. Thus, if the superset

of *television* is *appliance* then, on the Collins and Quillian model there will always be a direct and maximally accessible path between the two. Now, in the experiment, half the subjects saw the sentence,

*Televisions need expert repairmen,*

and about 80% of those who recalled the sentence (25 of 31 subjects) did so when the cue was the superset, *appliance*. The superset name was, therefore, relatively much more effective as a retrieval cue. But, relative to what? In fact it was relative to another, different superset, namely *furniture*. But this cue itself became very effective in a different context. Of those who recalled the sentence,

*Televisions look nice in family rooms,*

over 70% (28 of 39 subjects) did so when the cue was *furniture* rather than *appliance*.

We have already argued that to attempt to account for facts such as these by appealing to the need for cross-classification renders the model unrecognizable. The difficulty is that fixed-structure models are inept at handling cross-classification. A more reasonable model starts with the assumption that the context imposes on the particular word in question, in this case *television*, an interpretation which requires that it have certain relationships to other terms. Other contexts may require quite different relationships. Thus we would say that context dynamically determines semantic relationships.

Our data suggest that people's internal representations for sentences often contain elements which could not be derived solely from the constituent words. There are only two possible sources of the information imported into representations: context and existing knowledge. Both would appear to be indispensable. If models of human memory are to cope with elaboration and particularization, they will have to allow for an interaction between incoming information and existing knowledge, for presumably it is the richness and variety of existing knowledge that contribute to the uniqueness and subtlety of the ultimate representation. We have tried to show that rather than merely recording inputs, models of memory should invoke understanding, and that understanding is not just parsing; it is processing to a level whose depth depends on the degree of interaction with the context and the existing knowledge base.

## REFERENCES

ANDERSON, J. R., & BOWER, G. H. On an associative trace for sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 673–680.

ANDERSON, J. R., & BOWER, G. H. Configural properties in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 594–605.

ANDERSON, J. R., & BOWER, G. H. *Human associative memory.* Washington, DC: Winston & Sons, 1973.

ANDERSON, R. C. Semantic organization and the retrieval of information from sentences. *Journal of Verbal Learning and Verbal Behavior,* 1972, **11,** 794–800.

ANDERSON, R. C. Substance recall of sentences. *Quarterly Journal of Experimental Psychology,* 1974, **26,** 530–541.

ANDERSON, R. C., & MCGAW, B. On the representation of the meanings of general terms. *Journal of Experimental Psychology,* 1973, **101,** 301–306.

BRANSFORD, J. D., BARCLAY, J. R., & FRANKS, J. J. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology,* 1972, **3,** 193–209.

BRANSFORD, J. D., & MCCARRELL, N. S. A sketch of a cognitive approach to comprehension: Some thoughts about what it means to comprehend. In W. B. Weimer and D. S. Palermo (Eds.), *Cognition and symbolic processes.* New York: Winston & Sons, in press.

CLARK, H. H. The language-as-fixed effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* 1973, **13,** 335–359.

COLLINS, A. M., & QUILLIAN, M. R., Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior,* 1969, **8,** 240–247.

COLLINS, A. M., & QUILLIAN, M. R. How to make a language user. In E. Tulving and W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press, 1972.

KATZ, J. J. *Semantic theory.* New York: Harper and Row, 1972.

LABOV, W. The boundaries of words and their meanings. In C. J. Bailey and R. Shuy (Eds.), *New ways of analyzing variation in English.* Washington DC: Georgetown University Press, 1973.

QUILLIAN, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing.* Cambridge, MA: M.I.T. Press, 1968.

QUILLIAN, M. R. The teachable language comprehender. *Communications of the ACM,* 1969, **12,** 459–476.

ROSCH, E. H. On the internal structure of perceptual and semantic categories. In T. M. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press, 1973.

RUMELHART, D. E., LINDSAY, P. H., & NORMAN, D. A. A process model for long-term memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press, 1972.

TULVING, E., & THOMSON, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 1973,* **80,** 352–373.

WEINREICH, U. Explorations in semantic theory. In T. A. Sebeok (Ed.), *Current trends in linguistics,* Vol. 3. The Hague: Mouton, 1966.

WITTGENSTEIN, L. *Philosophical investigations,* translation by G. E. Anscombe. Oxford: Blackwell, 1963.

(Accepted September 17, 1974)