1-1-1973

The effect of syntactic form on retrieval from semantic memory.

Marilyn Bresler
University of Massachusetts Amherst

THE EFFECT OF SYNTACTIC FORM ON RETRIEVAL FROM SEMANTIC MEMORY

A Dissertation Presented

by

MARILYN BRESLER

Submitted to the Graduate School of the

University of Massachusetts in

partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 1973

Major Subject Psychology

THE EFFECT OF SYNTACTIC FORM ON

RETRIEVAL FROM SEMANTIC MEMORY

A Dissertation

By

MARILYN BRESLER

Approved as to style and content by:

Charles Clifton (Chairman of Committee)

Richard Conville (Member)

James Chumbley (Member)

Marvin Daehler (Member)

Ernest Dzendolet (Head of Department)

May 1973

Acknowledgement

The writer would like to acknowledge the fact that if not for Dr. Charles Clifton these experiments would never have been done and this paper never completed.

The Effect of Syntactic Form on Retrieval from Semantic Memory

(May 1973)

Marilyn Bresler, B. A., Brooklyn College

M. S., University of Massachusetts

Directed by: Dr. Charles Clifton

Two experiments are reported which deal with the effect of the syntactic form of a sentence on the reaction time (RT) of a true-false response to the sentence. Experiment one was a replication of a study by Collins and Quillian (1969) which supported a quasi-hierarchical conception of semantic memory. Sentences used in experiment one varied in their truth value, hierarchical level, and type of information presented. Each sentence was presented for true-false classification and RT was the response measure. The results supported the Collins and Quillian model. Experiment two was done to try to separate the effect of the syntactic form of the sentence from its semantic content. RT to sentences which express similar ideas, but in different syntactic forms, was measured. The results indicated that different kinds of semantic information, superset and property, are stored in different syntactic forms in memory.

TABLE OF CONTENTS

List of Tables

List of Figures

The topic of this paper is semantic memory. Semantic memory refers to the nature and the structure of the information which people know about words. In linguistics this problem is generally referred to as the nature of the lexicon or one's mental dictionary.

Tulving (1972) argued that the problems of semantic memory differ from those which have traditionally defined the area of verbal learning and memory. A brief summary of Tulving's position will be presented to relate the present experiments to previous research.

Traditional studies of memory have been concerned with episodic memory. Tulving (1972) proposed that episodic and semantic memory "differ from one another in terms of (a) the nature of the stored information, (b) autobiographical versus cognitive reference, (c) conditions and consequences of retrieval, and probably also in terms of (d) their vulnerability to interference resulting in transformation and erasure of stored information, and (e) their dependence upon each other [p. 385]."

Episodic memory consists of information about events or episodes in space and time which are part of an individual's experience. Episodic memory consists of autobiographical events. A list of nonsense syllables or words is a temporal sequence of events which the individual remembers as having been shown to him two minutes ago in an experimental room. It is a sequence of events which he experienced. Furthermore, episodic memory is not inferential or productive. The individual remembers only those events which happened to him. In reference to forgetting, Tulving (1972) suggested that the contents of episodic memory are probably more susceptible to forgetting due to loss of

spatial and temporal cues about the event than are the contents of semantic memory. Interference and decay theories of forgetting have been relevant to episodic memory.

Tulving (1972) defined semantic memory as "the memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and the other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of these symbols, concepts, and relations [p. 386]." In contrast to episodic memory, semantic memory is not autobiographical, is inferential and productive, and is probably less subject to forgetting. The information about a nation's capital city, for example, might constitute part of semantic memory.

Tulving (1972) also argued that much recent research has been concerned with the interaction of the two systems. For example, recent studies have been concerned with the free recall of word lists where specific semantic relationships, like hierarchical groupings, exist among the words (for example, Bower, Clark, Lesgold, and Winzenz, 1969). According to Tulving's analysis, these studies have been concerned with how Ss use information in semantic memory, in order to have accurate episodic memory. The emphasis of the research, however, has been on the processes which underlie accurate recall, rather than on the structure of semantic memory.

Tulving's distinction was accepted and used as a justification for omitting discussion of the free recall literature where semantic variables have been investigated.

While the writer accepted Tulving's (1972) distinction for its usefulness in narrowing the scope of the problem, there is one point which should be mentioned before considering some of the semantic memory models which have been proposed.

The problem is that there is a lack of agreement among theorists as to what knowledge is being modeled. In Tulving's (1972) definition, semantic memory is about our knowledge of verbal or linguistic symbols. Kintsch (1972) in introductory remarks about his own semantic memory model, stated that he was "concerned merely with the structure of a person's vocabulary, not his 'model of the world'. In other words, the empirical truth, falsity, likelihood, etc., of statements are irrelevant for present purposes. In terms of the distinction between pragmatic, semantic, and syntactic factors, we are only dealing with the latter two [p. 248]."

Others have felt that their models have broader generality. They were concerned with modeling one's "knowledge of the world" rather than just modeling one's knowledge of words. Quillian (1969) writing of a computer program which could read limited verbal material, stated: "'Comprehending' text is here defined as the relating of assertions made or implied in that text to information previously stored as part of the comprehender's general 'knowledge of the world'. Correspondingly, the central aim of . . . the model is the ability to appropriately relate text to the correct pieces of stored general knowledge of the world [p. 460]."

In a paper on the acquisition of mathematical knowledge, Greeno (1972) noted "the papers . . . in this volume provide important general

analyses of the way in which knowledge is stored in the mind. These papers dealing with semantic memory have dealt primarily with linguistic knowledge and knowledge about the world that is described in sentences. However, there is a sense of the term 'semantics' that refers to conceptual structures in a general way [p. 375]."

Collins and Quillian's Model of Semantic Memory

The impetus for the present paper was a semantic memory model proposed by Quillian (1967, 1968, 1969). Quillian's model concerns the organization of semantic information in memory and its subsequent retrieval. Quillian (1967) stated that the model is limited in that it "is not at present intended to handle all kinds of information that people presumably store in their heads. It is designed to hold only denotative, factual information . . . [p. 410]."

Semantic memory is conceptualized as a network of unidirectionally associated nodes. Each node may be considered as either a word or a set of property elements (features). The network is structured by the type of associations existing between nodes.

The definition of any word in memory begins with a node referred to as a type node. The type node is the word to be defined. Coming out of the type node or name word is a set of pointers to other nodes. These other nodes are called tokens. The associations from the type node to its tokens, and from token to token, define the name word. Among the tokens defining a word, there must first be one token which refers to a superset of the type node. Other token nodes refer to specific properties of the type node.

Figure 1 shows three type nodes, "plant," "tree," and "oak" and their token nodes. For the type node "tree" there is one token node for superset or superordinate information, i.e., "plant." Other token nodes branching out of "plant," "branches" and "has leaves," give the characteristics specific to "tree."

Insert Figure 1 about here

While the token nodes defining "tree" may occur in the definition of other words, the definition of any of these tokens occurs once. For example, "branch" may occur as a token in the definition of either "tree" or "river." As a type node, however, "branch" occurs once. For each occurrence of "branch" as a token node, there is a pointer or association which feeds into the type node for "branch."

Associations, then, are type-to-token and token-to-token, with an additional association from each token referring back to its own type node. This type-token-type structure reduces the amount of stored information, because tokens are defined only once, and not every time they occur as tokens.

Spatially, Quillian (1967) conceptualized the model as a set of planes. A type node with pointers to its token nodes exists in one of these planes. In going from a token to its type node, one moves into a new plane. This new plane consists of a type node and pointers to all its token nodes.

Two interesting characteristics of the model are its quasi-hierarchical structure and its nonredundant coding of property

information. One can trace the supersets which follow from each type node to define a hierarchy of superordinate terms. In Figure 1, there is a hierarchy of nouns: "oak"--"tree"--"plant"--"structure." The system is only quasi-hierarchical because there is always the potential for a higher order token to loop back down to a lower level in the hierarchy. In a more recent discussion of the model, Collins and Quillian (1972b) have stated that the hierarchical chains probably do not extend for more than three or four steps.

Related to the hierarchical structure is the idea that property information is stored in a nonredundant manner. Conrad (1972) has referred to this characteristic as cognitive economy. In Figure 1, note that even though the properties of "having roots," "having branches," etc., are true of oaks, these properties are not directly stored with the definition of "oak." Properties are generally stored with the highest level type node of which they are true. In short, one knows that an oak has roots indirectly, or through knowing that an oak is a tree, a tree is a plant, and that a plant has roots.

In a recent discussion of the model, Collins and Quillian (1972b) qualified this assumption of nonredundancy. They stated that properties of a superset word may be stored with a subset word, and that the degree to which this occurs depends on the conditions of learning the meaning of both words.

There are other characteristics of the model which are important for the present experiment.

Quillian (1967, 1969) stated that concepts, rather than word definitions are being represented. Concepts are distinguished from word definitions in that word definitions include only the information presented in one plane, i.e., a type node and its tokens. Concepts, on the other hand, may require a configuration of several type nodes and their tokens to be fully represented. Furthermore, some concepts might not have names. One may be able to define a concept in one plane, but there may be no one word in English which stands for that definition. Conversely, there may be several words which have the same definition, as with synonyms. Although Collins and Quillian (1972b) did not state this specifically, they seemed to imply that synonymous terms occupy the same type node.

Besides the structure of the information, another important aspect of the model is the form of the information. As has already been stated, the type node is directly associated with a token which defines its superset. This superset is then followed by a set of tokens in a particular configuration which shows how the superset must be modified in order to define the type node. The format of the word tokens is also of concern. In trying for greater economy of storage, all words are represented in canonical form (Quillian, 1967). Presumably, the most lexically simple form of a word is used. This is done both to reduce redundancy, and because the model is concerned with representing conceptual information rather than with just defining words. Quillian (1967) wrote: "Here it will only be noted that in encoding dictionary definitions all grammatical inflections . . . vanish, that is do not

become nodes themselves but instead dictate that various range-restricting tags be appended to the token nodes of certain other words. Removing all inflections during encoding permits all nodes in the memory model to represent canonical forms of words . . . [p. 419]."

For example, Collins and Quillian (1972b) remarked that "noun and verb forms of the same word must refer to the same concept . . . and should be treated similarly [p. 317]." However, they did not specify what they mean by treating noun and verb forms of the same concepts similarly. In Quillian (1967), there was an example of a concept which can be realized as either a noun or a verb, i.e., "plant." However, the definitions of the noun and verb forms of this concept were not specially marked to show the similarity. In this case, there were three planes for each of the three meanings of "plant." (Plant as a building is the third meaning given.) The only indication of the conceptual similarity between "plant" as a verb and "plant" as a noun was the notation that there are three meanings of "plant." Since there were separate definitions for the various meanings of the word, there is no economy of storage in this case.

The failure to note the conceptual similarity of the two nodes, and to achieve economy of storage, may, however, merely have been a function of this particular example. It may also have been due to the fact that this particular example comes from an earlier statement of the model. The model may have been modified subsequent to the Quillian (1967) paper in a way that would emphasize conceptual similarity.

Experimental Tests of the Model

The hypothesis of a nonredundant, hierarchical structure of memory has been tested by studying reaction time (RT) to true and false sentences (Collins and Quillian, 1969). Since the model specifies how information is organized, given some additional assumptions about how memory will be searched, the model predicts that sentence RT is a function of the number of nodes which separate the stored representations of the subject and predicate terms in memory.

Referring to Figure 1, consider the time required to respond "true" to (1) "An oak has acorns", and (2) "An oak has branches." Information relevant to the first sentence is stored as a token node under "oak" along with other properties specific to oaks. Properties of oaks as trees, however, are not stored under the type node "oak," but under the type node "tree." Therefore, making a "true" response to sentence (2) requires moving from the type node "oak" to the type node "tree." Since this decision requires traveling through an additional level of the hierarchy, RT should increase because of the additional nodes searched.

Before summarizing the method and results of the Collins and Quillian (1969) experiment, a discrepancy in the models presented in the various papers should be noted.

Figures 1 and 2 indicate the two kinds of memory networks which have been proposed. In Figure 1, based on Quillian (1967), each underlined word is a type node. The pointers from the type node are to superset tokens and property tokens. Coming out of each token node is the pointer to its type node indicated by the dashed arrow. In

moving from the token to its type, one moves to another plane or level. The important thing to note about the first model is that the information needed to respond "true" to "An oak is a tree," does not require moving to another plane or level. Superset information is present as a token within the plane of the type node.

Insert Figure 2 about here

Figure 2 shows the model presented in Collins and Quillian (1969). The memory net has been simplified in this paper. The type-token distinction is only implicit, and the hierarchical organization of words is emphasized rather than the movement to another plane. Nonredundant coding of information is clear, however. Note that deciding "true" to "An oak is a tree" requires moving up one level to another plane.

While the two models do not necessarily differ in predictions on RT, they do differ in terms of sentence classification and in where superset information is located. The implications of using the Quillian (1967) model to interpret the data of Collins and Quillian (1969) are presented later.

_Ss_' task in Collins and Quillian (1969) was to make true-false judgments about sentences that predicate property or superset relationships. Table 1 shows a sample set of the sentences used, based on the semantic hierarchy of words shown in Figure 2. P and S refer to property and superset relations, respectively. The 0, 1, and 2 following the S and P designate the number of levels which must be traveled

before the relevant information is found. Judgment on S0 sentences involves an identity operation and should require no search other than

----

Insert Table 1 about here

----

finding the initial node, i.e., recognizing the word. Judgment about S1 sentences requires moving up one level to retrieve relevant information, and judgment about S2 sentences moving up two levels to retrieve relevant information. In P0 sentences, "acorns," for example, is present at the same level as "oak". Judgment about P1 sentences requires moving up one level, and on P2 sentences moving up two levels.

A consistent increase in RT to true sentences from level 0 to level 1, and from level 1 to level 2 was predicted. The same increase should hold for property and superset relation sentences. Furthermore, if property search is assumed to be dependent upon first reaching the superset name, then property relation sentences should have longer RTs than superset relation sentences. More specific assumptions regarding search strategies are discussed later.

Three experiments were run to test the model. In each experiment <u>S</u>s were presented with sets of true and false sentences. The sentences were from either two- or three-level hierarchies. (In a two-level hierarchy, P2 and S2 sentences were omitted.) A typical hierarchy of sentences is shown in Table 1, except that instead of using "oak" in each sentence, names of different trees were used. Eight to twelve hierarchies of different semantic content were used.

Across the three experiments, false sentences were constructed in one of two ways. The first method, illustrated by the sentences in Table 2, involved predicating false properties and categories to the tree names. The false properties and categories were selected without regard to level of contradiction. For example, "A dogwood is lazy"

Insert Table 2 about here

---

is false because laziness contradicts a property of plants. The other two false property sentences in Table 2 are false because they contradict properties of particular trees, "hemlock" and "poplar." Therefore, under this method of constructing false sentences, false sentences are grouped together regardless of the level of the hierarchy at which the contradiction occurs.

The second method for constructing false sentences is illustrated by the sentences in Table 3 and the memory net in Figure 3. This

---

Insert Table 3 about here

Insert Figure 3 about here

---

method controlled for the level of contradiction. The P0 sentence shown in Table 3 is false because it contradicts a property of oaks. The P1 sentence shown contradicts a property of trees, and the P2

sentence contradicts a property of plants. Thus, there is a consistent increase in the level of contradiction.

Superset sentences were always of the syntactic form: "A (noun) is a (noun)." Property relation sentences were of one of the following forms: "A (noun) can (verb)," "A (noun) has (noun)," and "A (noun) is (adjective)."

Sentences were presented in blocks of 32 to 48. Each sentence appeared for two seconds, followed by a blank screen for two seconds. Ss pressed one of two response buttons to indicate "true" or "false."

The results for true sentences were that property relation sentences had consistently higher RTs than superset relation sentences for all levels. There was a fairly constant increase of 75 msec. This 75 msec. difference was interpreted as the time required to move from a node to its superset. The prediction of two parallel straight lines was confirmed except for one divergent point for SO sentences. RT to SO sentences was below the point predicted by parallel lines. This effect was attributed to Ss responding to SO sentences as a pattern matching task rather than as a sentence comprehension task. Ss respond to the subject and predicate nouns of SO sentences as visual patterns rather than understand them as words.

The average increase in RT from S1 to P1 and from S2 to P2 was 225 msec. This constant was considered to represent the time needed to retrieve a property from the level where it is stored.

The retrieval process indicated by the results for true sentences is a combination of parallel and serial processing. S starts at the initial category node and begins searching properties at level 0. If

the information is present, the average time to complete the property search is 225 msec. Whether the property search at any level is exhaustive or self-terminating cannot be determined from this experiment. Simultaneously with the property search of level 0, S is moving to level 1, and in 75 msec. reaches the level 1 superset node. Property search of level 1 then begins. Thus, there is simultaneous search of level 0 and 1 properties, except that the search of level 0 began 75 msec. earlier and, therefore, should end 75 msec. earlier. Seventy-five msec. after level 1 property search begins, S reaches the level 2 superset node and begins searching properties at level 2. Search of level 2 properties occurs simultaneously with search of level 0 and level 1, except that search of level 2 is 75 msec. behind that of level 1, and 150 msec. behind that of level 0.

Search is serial because S must pass through each superset node before the property search can begin at that level. Furthermore, Ss cannot skip nodes. Processing is parallel because property search occurs at several levels simultaneously, with 75 msec. lags between levels. Movement through the superset nodes occurs at the same time as lower level property search continues.

The results of the study can also be interpreted in terms of the Quillian (1967) model. This involves reinterpreting the level at which superset information is retrieved. In Figure 1, it can be seen that "An oak is a tree" would be classified as an S0 sentence. All the information for the decision is in one plane. In Figure 2, it can be seen that "An oak is a tree" would be classified as an S1 sentence because the relevant information is in two planes.

Similarly sentences classified as S2 by Collins and Quillian (1969) are reclassified as S1 sentences using the model in Quillian (1967). To respond "true" to "An oak is a plant" requires moving up one level to the type node "tree" in Figure 1. In this plane, one can retrieve the information that a tree is a plant and that, therefore, an oak is a plant.

Classification of property relation sentences with regard to level of retrieval does not change using the Quillian (1967) model.

One result of using the Quillian (1967) model to interpret the Collins and Quillian (1969) data is to reduce the difference in RT between property and superset sentences. Sentences previously classified as S1 are now S0, and S2 sentences are now S1. Therefore, the original difference of 225 msec. between property and superset relation sentences is now reduced by 75 msec., the change in RT between levels.

The resulting difference of 150 msec. between S0 and P0, and S1 and P1, is now interpreted as the time required to retrieve a property from a node once the superset information has been retrieved. The 75 msec. increase between levels is the time required to travel between the superset tokens of each type node.

Referring to Figure 1, if _S_ responds "true" to "An oak is a tree," this indicates that _S_ has retrieved the "tree" token under "oak." If that decision took, say 1000 msec., then in 1075 msec., _S_ can respond "true" to "An oak is a plant," indicating he has retrieved the "plant" token under "tree." Thus, 75 msec. is the time required to travel between superset tokens of the type nodes.

Again, if S responds "true" to "An oak is a tree," this indicates S has retrieved the "tree" token under "oak." If this decision took 1000 msec., then in 1150 msec., S can respond "true" to "An oak has acorns." The difference in RT of 150 msec., between the SO and PO sentences, is the time to retrieve other property tokens, once the superset token has been retrieved.

Probably, one can apply the Quillian (1967) model to the Collins and Quillian (1969) data in other ways. For example, one can assume that S cannot respond "true" to "An oak is a tree" until he reaches the "tree" type node. This kind of assumption would be inconsistent with the assumptions made about retrieval of other properties, however. One can see that by simplifying the model to that shown in Figure 2, the interpretation of the data becomes simpler.

A study by Collins and Quillian (1970b) provided some additional support for the model shown in Figure 2. This is the model tested by Collins and Quillian (1969) where the superset information is retrieved by moving up the hierarchy, and the properties specific to a type node are directly associated with it. In the earlier model, (Quillian, 1967, 1968, 1969) the properties specific to a type node are retrieved only after retrieving the superset of the type node (see Figure 1).

In this particular experiment, the locus of the superset information is important in making predictions about retrieval. Collins and Quillian (1970b) assumed that, if in retrieving information from the network, a particular pathway is used, then that pathway is momentarily more accessible. Assume that the sentence "An oak has branches" is presented for true-false classification. In searching

the hierarchy, depicted in Figure 2, S uses the path from "oak" to "tree" to retrieve the information that trees have branches. This inference should make the "oak" to "tree" pathway temporarily more accessible. If S is now shown the sentence "An oak is a tree," RT should be faster relative to some control condition where the path from "oak to "tree" had not been previously activated.

The sentences used in this study were similar to those used in the previous study (Collins and Quillian, 1969), except that pairs of sentences with the same subject noun were used. Eight combinations of P0, P1 and S1 sentences were used. A P0 sentence was preceded by a P0, P1, or S1. A P1 was preceded by a P0, P1, or S1. S1 sentences were preceded by either a P0 or P1 sentence.

Trial 1 data for P0, P1, and S1 sentences replicated the results of Collins and Quillian (1969) with one divergence. There was very little difference between P0 and S1 sentences in RT. This result is interesting because for both of these sentences, the information to be retrieved is one step away from the initial node. This result indicated that it is the number of steps traveled in the hierarchy that alone determines RT, and not also the difference between superset and property relations. In the Collins and Quillian (1969) study, the RT to S1 sentences was about 125 msec. faster than the RT to P0 sentences.

Collins and Quillian (1970b) made eight predictions about the conditional RTs. All the RT differences obtained were in the predicted direction; however, only five were significant. For example, it was found that RT to S1 sentences was significantly faster when

the S1 sentence had been preceded by a P1 sentence, in contrast to when the S1 sentence had been preceded by a P0 sentence. That is, it was easier to respond "true" to "An oak is a tree" when the sentence had been preceded by "An oak has branches," than when it had been preceded by "An oak has acorns."

Of the five significant effects, it should be noted that these effects were reduced when the data were adjusted to take into account the finding that mere repetition of the subject noun in a sentence reduced RT. (Collins and Quillian, 1970b, do not report what kind of adjustment was made in the data.) The results generally provided weak support for the model of semantic memory as being hierarchically organized with nonredundant coding of information.

Collins and Quillian (1969, 1972a) reported a number of experiments aimed at determining the strategies Ss use to decide that a sentence is false. They argued that Ss can employ one of two general strategies. They can search semantic memory for confirming evidence for a given amount of time, or through a given number of levels, and respond "false" if they find no confirming evidence. The other type of strategy is to search semantic memory for contradictory evidence, and then respond "false" when they find it.

Several lines of evidence led to the rejection of the first kind of strategy. Ss do not seem to set a time limit on how long they search (Collins and Quillian, 1972a). If there were a time limit, then there would have been less variability in RT to false sentences than to true sentences. The reverse was found.

Neither do Ss set a limit on the number of levels they search in the hierarchy. The evidence against a depth limit hypothesis was introspective. Ss could give an interpretation to the false sentences and could also explain why they are false (Collins and Quillian, 1972a). If they can give an interpretation, this means that they have found some path connecting the subject and predicate terms in the sentence. If they can explain why they give a "false" response, this means they have found contradictory evidence, rather than failed to find any confirming evidence.

Rejection of the time limit and depth limit hypotheses led to a series of experiments testing the contradictory evidence hypothesis. The contradiction hypothesis states that Ss respond "false" because they have found information in memory which contradicts the information in the sentence presented.

The Collins and Quillian (1969) paper previously discussed included a test of the contradictory evidence hypothesis. False sentences were constructed so that the contradictory information was located at various levels of the hierarchy. RT to the false sentences should, therefore, have increased as the semantic distance of the contradictory information increased. This was not found. RT to false property sentences showed no change with hierarchical level. RT to false superset sentences showed a slight but not significant decrease with hierarchical level.

These results were inconsistent with a hierarchical theory of memory in which lower levels are searched for contradictory information before moving up to the higher levels to search for contradictory

information. The results of two other experiments indicated, however, that Ss do search for contradictions, and indicated why Collins and Quillian (1969) found that semantic distance decreased RT for false superset sentences.

In one experiment, (Collins and Quillian, 1972a) false property sentences were constructed with respect to two variables: (1) sentence anomalousness, and (2) presence of extraneous paths. A sentence is anomalous if contradictory evidence cannot be found within three levels of the hierarchy. Non-anomalous sentences have a connecting path within three levels of the hierarchy. Anomalousness, then, refers to semantic distance. The greater the semantic distance, the more anomalous the sentence. If a sentence has an extraneous path, it means that a path can be found between the subject and predicate terms, but the path is misleading, as in the sentences: "An almond has a fortune," and "A newspaper is red."

For these false sentences, anomalousness had no effect on RT, but the presence of an extraneous path significantly increased RT. That is, semantic distance did not increase RT, but the presence of misleading or confusing information did.

Collins and Quillian (1972a) stated that Ss always find at least one path, and possibly several paths, connecting the elements of the sentence. The Ss then compare the relationships found in memory, to the relationship posited in the sentence. If only one path is found, and that path contradicts the relationship stated in the sentence, then Ss respond "false." The presence of extraneous

paths slows down the decision making process because _Ss_ have to check out several paths.

This checking of extraneous paths, Collins and Quillian (1972a) stated, is the reason for the peculiar semantic distance effects obtained in Collins and Quillian (1969) for false sentences. In that study, it took somewhat longer to respond "false" to "A canary is an ostrich" than to "A canary is a fish." They attributed this effect to the presence of more extraneous paths between "canary" and "ostrich," than between "canary" and "fish." The time to reject false sentences depends not on semantic distance, but "on whether or not there is a neighboring node, which is confusable with the node given in the sentence. Where there is such a node,[they] predicted that RT would be longer, because the path through that node would need to be checked and rejected [Collins and Quillian, 1972a, p. 25]."

To test the effect of confusable nodes, confusable and nonconfusable false property and superset sentences were constructed. A confusable property sentence is one "where there exists within the superset a neighboring node which has the property specified in the sentence, (e.g., 'A tiger has a mane') [Collins and Quillian, 1972a, p. 25]." A confusable superset sentence is one "where the superset in the sentence is highly confusable with the correct superset, (e.g., 'A St. Bernard is a cat') [Collins and Quillian, 1972a, p. 26]." Confusable nodes, as Collins and Quillian define them, seem to be special cases of the occurrence of extraneous paths. That is, for confusable sentences, the extraneous paths occur between neighboring nodes.

The results of this study (Collins and Quillian, 1972a) were that the time to respond "false" to confusable property and superset sentences was longer than the time to respond "false" to nonconfusable sentences. Collins and Quillian (1972a) concluded that the "RT to decide whether a sentence is true or false depends on the number of paths considered, and the length of these paths. The consideration of more than one path explains why false sentences which contain misleading associations . . . or confusions . . . take longer to reject than sentences without such associations or confusions [pp. 27-28]."

Kintsch, Crothers and Berman (1970) reported three sentence comprehension experiments similar to those of Collins and Quillian (1969). Kintsch, et al. presented simple sentences which differed in (1) acceptability, (2) semantic distance, and (3) syntax.

Acceptability refers to whether the sentences are potentially true or false. Semantic distance refers to the degree of relationship between the subject and predicate terms. The two variables taken together give four kinds of sentences. (1) An acceptable sentence in which the subject and predicate terms are closely related is a definitional sentence, e.g., "A shark swims." (2) An acceptable sentence in which the subject and predicate terms are not closely related is a contingently true sentence, e.g., "A shark escapes." (3) Contradictory sentences are unacceptable, and the subject and predicate terms are closely related, e.g., "A shark growls." (4) Nonsensical sentences are unacceptable, and the subject and predicate terms are not closely related, e.g., "A shark ticks."

In addition to these four kinds of sentences, Kintsch, et al. (1970) used three syntactic forms similar to those used by Collins and Quillian (1969). They used both superset and property relation sentences. Superset sentences were of the form "A (noun) is a (noun)." Property relation sentences were of one of two forms: "A (noun) is (adjective)," and "A (noun) (verb)."

Kintsch, et al. (1970) reported that acceptable sentences had significantly faster RTs than unacceptable sentences. Acceptability and semantic distance interacted. Acceptable sentences had faster RTs when subject and predicate were closely related. Unacceptable sentences had faster RTs when subject and predicate were not closely related. This interaction is in agreement with Collins and Quillian's (1972a) findings, and with the results reported by other investigators (Meyer, 1970; Schaeffer and Wallace, 1970).

Kintsch, et al. (1970) reported that syntactic form did not significantly affect RT, nor did it interact with any other variables. Thus, they failed to replicate Collins and Quillian's (1969) finding that superset sentences were easier than property sentences. Kintsch, et al. concluded that this kind of syntactic variable is not very relevant in sentence RT studies. They attributed Collins and Quillian's syntactic form effect to poor control over stimuli. Kintsch, et al. controlled the Thorndike-Lorge frequency of the subject nouns, and of the predicate adjectives, verbs, and nouns. Furthermore, the same subject noun was used with twelve different predicates. Collins and Quillian (1969) reported that the predicate nouns in their superset sentences tended to

have higher Thorndike-Lorge frequencies than the adjectives and verbs in the property sentences. Therefore, the RT difference between superset and property sentences obtained by Collins and Quillian may have been due to an individual word frequency effect.

Although Kintsch, et al. (1970) failed to find a RT difference between property and superset sentences, this may have been due to another factor besides word frequency. Kintsch, et al. did not classify their sentences with respect to hierarchical level. Although their semantic distance variable is somewhat similar to Collins and Quillian's (1969) hierarchical level variable, as can be seen from the examples presented above, the two variables are not quite the same. The specific superset and property terms which Kintsch, et al. used may be from different levels of the hierarchy. Combining RTs from sentences without regard to hierarchical level may also have contributed to the absence of a RT difference between property and superset sentences.

Criticisms of the Collins and Quillian Model and Alternative Models

The Collins and Quillian model and the supporting research have been criticized by several investigators. Some of these investigators have proposed alternative models in explanation of the data. Landauer and Freedman (1968) have proposed a category size explanation of the data. Schaeffer and Wallace (1969, 1970) have proposed a word comparison model. Meyer (1970) has proposed a two-stage model of sentence comprehension. Finally, a host of criticisms all concerned with frequency effects have been made about their research.

Category size arguments. A study by Landauer and Freedman (1968, Exp. 1) provided an alternative explanation of some of the Collins and Quillian (1969) results. According to Landauer and Freedman, the RT differences which Collins and Quillian attributed to the hierarchical structure of memory, can be interpreted as effects of category size.

Landauer and Freedman (1968, Exp. 1) were concerned with the problem of long-term-memory search strategies. If memory search is a serial process, then the longer the list of items, the longer the search time to find a particular item in memory. List length was varied by using the following hierarchical set of categories: "word," "noun," "living thing," "animal," and "dog." They argued that when an item is presented, Ss must search through a list of items stored in memory under a category name in order to determine if that item is an instance of the category. Furthermore, they argued that in searching memory, the higher order categories must include all the items of the lower order categories. Therefore, search time for higher order categories should be longer than search time for lower order categories.

(The use of "word" and "noun" as categories on a continuum with "living thing," "animal," and "dog," is illustrative of the problem mentioned previously of what is being modeled. Landauer and Freedman, 1968, seem to imply that the model is of the Ss' knowledge of words.)

The five categories were grouped into sets of two by pairing adjacent categories. Thus for each set of categories there was a larger and a smaller category. Ss were told a category name, and were then given a positive or negative instance and required to make a classification

response. Each instance was classified twice by each S. The instance was presented once after the larger, and once after the smaller category. All positive instances were members of both categories, and the negative instances were members of neither category.

Landauer and Freedman (1968, Exp. 1) found that it took longer to classify the instance with respect to the larger category. This effect was significant for negative instances only. For positive instances, the difference was not significant, but in the predicted direction. This effect held for all of the nested category comparisons, except for the "word-noun" comparison where the RT was greater to the smaller category "noun" than to the larger category "word." Landauer and Freedman (1968) concluded that in making a classification response, Ss search lists of positive instances stored in memory to determine if the presented instance is a positive instance of a category.

The category size hypothesis provides an alternative explanation of the Collins and Quillian (1969) finding that RT increased with hierarchical level for superset sentences. That is, the reason that it took longer to decide that "A canary is a bird" than to decide that "A canary is an animal," is because the animal list in memory is longer than the bird list in memory.

Collins and Quillian (1970a, 1972a) have noted this alternative explanation of their results, and have argued that their own hierarchical model provides a better account of the data. Their interpretation of the Landauer and Freedman (1968, Exp. 1) data is that when a positive instance like "collie" is presented with the categories "animal" and "living

thing," Ss have to find a path between the instance and the category in the hierarchical memory network previously described. Classifying "collie" as an "animal" takes less time than classifying "collie" as a "living thing" because the first two items are closer in the semantic hierarchy. Collins and Quillian (1970a) attributed the failure of Landauer and Freedman to find a significant effect for positive instances, to the fact that Ss were continuously using the same categories. In Collins and Quillian (1970b) they showed that with repetition some paths become more accessible, thereby decreasing RT.

With negative instances, however, Ss cannot always use their previous responses to make inferences. The results of a previous decision can only be used if the S first saw the negative instance with the higher order category, e.g., if "tulip" was a negative instance first presented for classification with respect to the category "animal." When "tulip" is next presented with the lower order category "dog," Ss can use their last response as the basis for making their response. That is, they can decide that a tulip is not a dog, because they had previously decided that a tulip is not an animal, and, therefore, it cannot be a dog. Responding on the basis of this inference saves time, and thereby decreases RT for the smaller categories.

With the positive instances, then, Ss can always make use of their previous decisions to save time. With the negative instances, however, the inferential step can be used to save time only when the instance first occurs with the higher order category. This would explain the finding that there was a significant effect for negative instances only.

Collins and Quillian (1970a) stated that "for negative instances it might be that the Ss who had animal first and dog second would be faster on dog because they drew the correct logical conclusion, whereas the Ss who had dog first and animal second would not be faster on animal. This asymmetry in logic, then, might have produced the shorter times for the smaller nested category in deciding negative instances [p. 433]."

It was not clear from the Collins and Quillian (1970a) paper why they argued that RT to negative instances involves using a semantic hierarchy to make inferential decisions. In their previously discussed paper, (Collins and Quillian, 1972a) semantic relatedness, rather than hierarchical structure was the important determinant of RT to false sentences, or to negative instances if a word classification task is involved.

To test their explanation of Landauer and Freedman's (1968, Exp. 1) results, Collins and Quillian (1970a) ran two experiments. The purpose of experiment one was to determine whether category size, per se, affected classification time. To extricate the effects of category size from hierarchical ordering of categories, Collins and Quillian used three categories, two of which were independent of each other, but which were both nested within a third category. "Bird" and "dog" were selected as the two independent categories, where "dog" was considered to be the smaller category, because there were fewer dogs than birds listed in a thesaurus. "Animal" was used as the largest category. Lists of both positive and negative instances from these categories were presented once to Ss for "yes-no" classification. That is, S was given a category

name, then was shown a list of positive and negative instances, one at a time, for classification.

Landauer and Freedman's (1968) category size hypothesis predicts that classification time for positive and negative instances increases from "dog" to "bird" to "animal" since category size is increasing. Collins and Quillian (1970a, Exp. 1) predicted no difference in classification time across the three categories, since each positive instance is one step removed from its category name or superset. Their predictions for negative instances were not specified.

While Collins and Quillian (1970a, Exp. 1) concluded that "category size, in itself, is not a very critical variable in categorization time [p. 434]," the results were ambiguous. For negative instances, there were no significant differences in RT across the three categories, which argues against Landauer and Freedman's (1968) hypothesis. However, for positive instances, the RT for the dog category was significantly less than the RT for the bird and animal categories, a finding which supports the category size hypothesis. The absence of a significant difference in RT between "bird" and "animal," on the other hand, supports the Collins and Quillian model. Collins and Quillian (1970a, Exp. 1) argued that the higher RT for the bird and animal categories was due to Ss' forming inappropriate subclasses for these items, and subsequently being surprised at the range of the positive instances presented.

A second experiment by Collins and Quillian (1970a, Exp. 2), a partial replication of Landauer and Freedman's (1968, Exp.1) procedure,

was done to test Collins and Quillian's hypothesis about the effects of repeated classification of the same instance.

Ss were given positive and negative instances of dogs, birds, and animals which they were to classify with respect to the following pairs of categories: dogs vs. animals, birds vs. animals, animals vs. living things. Each instance was classified twice, once with respect to the lower order category, and once with respect to the higher order category. The Landauer and Freedman (1968) category size hypothesis predicts an increase in classification time across the dog to bird to animal categories for positive and negative instances. Collins and Quillian (1970a, Exp. 2) predicted an increase in classification time only between category pair members.

For positive instances, there was generally a significant increase in classification time between lower and higher order categories. There were no significant decreases in classification time as category size decreased, except for the dog category, which was faster than the others.

The results obtained for negative instances tended to support neither the Landauer and Freedman (1968) category size hypothesis nor the Collins and Quillian (1970a) hypothesis that the Ss remember their first classification response and use it in deciding on their second classification of the same instance. Collins and Quillian (1970a, Exp. 2) found that classification time for negative instances depends on the degree of semantic relatedness between the instances and the category. The greater the degree of semantic relatedness between the items, the longer the classification time is. Specifically, they argued that a negative instance takes longer to reject when its superset is somehow related to the category name which

was presented. For example, it is difficult to reject the negative instance "tree" with respect to the category "animal," because the superset of "tree," "plant," is semantically related to "animal."

This is similar to the finding previously reported in discussing how false sentences are rejected, (Collins and Quillian, 1972a). False sentences, or negative instances, are difficult to reject when there are confusable nodes involved. In this particular experiment (Collins and Quillian, 1970a, Exp. 2), "plants" and "animals" are confusable nodes. It is difficult to decide that "tree" is not an "animal," because a tree is a plant, and "plant" and "animal" are confusable, or semantically related.

Surprisingly, this interpretation was supported by the results of a similar study by Wilkins (1971, Exp. 2). Wilkins used the same classification procedure whereby a category name was presented, followed by a positive or negative instance, and Ss were timed on their classification responses. Wilkins used two types of negative instances. In one case, the negative instance was related to a superset of the category which was presented to Ss. In the other case, there was no such relationship. Classification of the negative instances took longer when there was a semantic relationship between the negative instance and a superset of the presented category.

With respect to the category size variable, however, Collins and Quillian (1970a) concluded that it is the hierarchical organization of the categories which determines classification time, rather than the number of instances in a category.

Probably one of the important considerations on the question of whether category size or hierarchical structure determines classification time, is the method used to determine category size. Landauer and Freedman (1968) used nested categories to increase category size. This procedure, as Collins and Quillian (1970a) have pointed out, confounds category size with hierarchical organization. Collins and Quillian (1970a) in trying to extricate these variables, used a dictionary and a thesaurus as external indices of category size. That is, since a dictionary lists more types of birds than dogs, "bird" was considered the larger category.

Wilkins (1971) used a "production method" to estimate category size in a word classification task. The production method for estimating category size is similar to a word association procedure. Ss are given a category name and asked to produce instances of that category. Wilkins used the Connecticut norms (described in Bousfield, Cohen, and Whitmarsh, 1958) which were generated in the following manner. Ss were given a category name and asked to produce four instances of that category. A large category from the norms is presumably one where Ss gave many different instances. A small category is one in which few instances were given.

The word classification procedure involved presentation of the category name, then the presentation of a positive or negative instance from a large or small category. Wilkins (1971, Exp. 1) reported that instances of large categories had significantly greater classification times than did instances from small categories, and that negative instances took significantly longer than did positive instances.

Since Wilkins (1971, Exp.1) did not confound category size and hierarchical structure, these results seemed to provide evidence for the hypothesis that classification time depends on the number of instances in a category. However, Wilkins' method of estimating category size makes his conclusion questionable.

Category size, as stated, was defined by the number of different instances listed in the Connecticut norms. However, the Connecticut norms were generated by a technique which would seem to place a limit on the total number of positive instances Ss produce. Limiting the number of positive instances produced, would invalidate the use of the norms as a way of estimating category size.

For example, Wilkins (1971, Exp. 1) used "vegetable" as a small category, as defined by the Connecticut norms. In collecting these norms, Ss were told to list only four instances of "vegetable." Given a limit on the number of instances to write, Ss probably list only the most common vegetables, rather than try to think of unusual instances. If all Ss list only the four most common vegetables, then one would conclude that "vegetable" is a small category.

In summary, it seems that the Landauer and Freedman (1968) argument that category size affects the time to search semantic memory has not been demonstrated because of methodological problems in isolating category size from hierarchical organization, and in adequately estimating category size.

Recently, Landauer and Meyer (1972) have reviewed the category-size literature, presented some additional relevant data, and proposed a category-search model to account for the results of word classification studies.

In their review of the literature, Landauer and Meyer (1972) criticized the Collins and Quillian (1970a) study for the use of a small number of categories, and they criticized the Wilkins (1971) study for the specific instances Wilkins used. Landauer and Meyer argued that the Collins and Quillian (1970a) semantic relatedness hypothesis has not been adequately tested. Landauer and Meyer (1972) also pointed out that the Collins and Quillian (1970a, Exp. 1) failure to find a category size effect for positive instances may have been due to _Ss_' redefining the categories _E_ presented. Collins and Quillian (1970a, Exp. 1) considered "animals" a large category. However, their _Ss_ might have considered "animals" to mean "mammals," thereby reducing category size and narrowing the category size difference between "animals," "dogs," and "birds."

Landauer and Meyer (1972) then presented data from two experiments which demonstrated category size effects. Meyer and Ellis (1970) had _Ss_ classify positive and negative instances with respect to nested categories. Meyer and Ellis included a condition where nonwords, e.g., "mafer," were presented as negative instances of the nested categories. Meyer and Ellis found a category size effect for the positive and negative instances, and also for the nonwords.

Landauer and Meyer (1972) stated that the category size effect for nonwords is evidence against Collins and Quillian's (1970a) semantic relatedness hypothesis for negative instances. However, as Landauer and Meyer (1972) pointed out, it does not seem reasonable to generalize about the treatment of negative instances on the basis of nonword data.

As a more appropriate test of the category size and semantic relatedness hypothesis about negative instances, Landauer and Meyer (1972) presented a reanalysis of some data which were collected in an unpublished replication of the Landauer and Freedman (1968, Exp. 1) study. In this study, Ss classified positive and negative instances (words) with respect to five pairs of nested categories. Both positive and negative instances showed a category size effect.

As a test of the semantic relatedness hypothesis, Landauer and Meyer (1972) had Ss rank the negative instances for "closeness in meaning" to the nested category members. Collins and Quillian (1970a) predict an interaction between semantic relatedness and category size for negative instances. That is, when the negative instances are semantically related to the nested categories, RT should be faster to the smaller of the nested categories. When the negative instances are not semantically related to the nested categories, there should be no difference, or at least a smaller difference, in RT to the nested categories.

Landauer and Meyer (1972) reported that RT was faster for the smaller categories than for the larger categories for both the closely related and the more remotely related negative instances. Therefore, they concluded that the Collins and Quillian (1970a) semantic relatedness hypothesis is insufficient to explain the observed category size effects for negative instances. Landauer and Meyer's argument is open to the criticism that even though the Ss were able to differentiate among the negative instances on the basis of "closeness in meaning" to the nested categories, the entire set of negative instances might have been semantically related to the

nested categories, so that the differentiation on the basis of ranking was of minimal significance. This possibility, however, seems rather remote.

Landauer and Meyer (1972) then argued that a simple model of the word classification effects assumes that S has a file of category names with lists of instances for each category stored in memory. When an instance is presented, S checks the relevant category list, and responds "yes" if the instance appears on the memory list, and "no" otherwise. The advantages of this model are that "it requires less information to be stored in memory, and embodies fewer assumptions about memory organization and search processes [Landauer and Meyer, 1972, p. 544]."

This model does seem to offer as adequate an explanation of the word classification effects as does the Collins and Quillian (1969, 1970a) model. Landauer and Meyer's (1972) discussion of the model, however, was ambiguous with respect to what items are included on the category lists in memory. In a discussion of an example case of the items on a category list in memory, they implied that only those items which are in a subset relationship to the category appear on the list, e.g., only "collie," "terrier," and "dachshund" appear on the "dog" list (Landauer and Meyer, 1972, p. 544). That is, in their example case, each category list is a two-level hierarchy.

In the appendix to the paper (Landauer and Meyer, 1972), however, they noted that they are discussing "categorization tasks in which the 'category members' denote either subsets of the specified category or sets that partially overlap it, whereas 'nonmembers' denote sets that are disjoint from the category. . . . In these tasks, it is sufficient to retrieve

information about whether the test word refers to a set that 'intersects' the specified category, that is, shares at least one exemplar [p. 548]." This general description of how list membership is determined implies that an item like "pets" occurs on the lists for "animals" and "dogs." This kind of category list is not hierarchical.

Previous word classification studies have used positive instances which are subsets of the nested categories. There have been no word classification studies where the instances overlap the nested categories, at least to the writer's knowledge, and a category size effect for this type of situation has not been demonstrated. Therefore, there is no data which would warrant positing the latter, less structured category lists for memory.

The former, two-level hierarchical category lists are similar to the hierarchies posited by Collins and Quillian (1969). The major difference between the models then is whether the RT increase in the classification of a positive instance with respect to nested categories derives from checking a list of category instances, or from moving from the instance to its higher order category.

Further evidence that category size does not affect retrieval time from semantic memory comes from a study by Freedman and Loftus (1971). Ss were presented with a category and asked to produce an instance of that category. For each category used, Ss were shown the category name plus one of two constraints on their response. The category was shown with either a constraining adjective, e.g., "flower-yellow," or with the first letter of a positive instance, e.g., "flower-p."

Category size was estimated by the use of the production method. Ss were shown the categories, adjectives, and letters used in the study, and were required to write down as many instances as possible in a one-minute period. This method of estimating category size decreases the likelihood that Ss give only the most common instances.

Freedman and Loftus (1971) reported finding a small, negative correlation between category size and RT. Moreover, this correlation dropped to near zero, when other variables were taken into account. They also reported a faster RT for the noun-adjective pairs than for the noun-letter pairs.

Freedman and Loftus (1971) argued that the absence of a category size effect indicates that Ss do not search lists of possible instances successively in retrieving words from memory. They felt that the data were consistent with Quillian's (1968) model which provides for immediate access to the category name, after which the search can spread out to specific instances of the category.

The model which seems to account for the data is the spreading activation model. Collins and Quillian (1969, 1970b) argued that when the search process starts at some particular node, it continues along several paths simultaneously. Collins and Quillian (1969) have shown that the spreading activation model is a viable model when memory search is moving up the hierarchy, and Freedman and Loftus (1971) have shown that spreading activation or parallel search is a viable description of memory search when moving down the hierarchy.

Schaeffer and Wallace's word comparison model. Schaeffer and Wallace (1969, 1970) have reported a series of word classification experiments which they felt disconfirmed the Collins and Quillian (1969) model, and supported their own model of a word comparison process.

In the first experiment reported in Schaeffer and Wallace (1969, Exp. 1), Ss were shown two instances from the four categories: mammal, flower, fabric, and metal. Their task was to respond "same" if both instances were living things, or if both instances were nonliving things, and to respond "different" if one instance was living and the other instance was a nonliving thing.

Schaeffer and Wallace (1969, Exp. 1) found that classification time was faster when both instances were from the same category. For example, classification time was faster when both instances were fabrics, than when one instance was a fabric and one a metal. Although they did not report the results of a significance test, "same" judgments were somewhat faster than "different" judgments.

There are probably several models which can explain these results, including the Collins and Quillian (1969) model. Interpretation in terms of the hierarchical model stresses how much of the hierarchy has to be searched before enough information is retrieved to make a decision. With two instances of the same category, search need proceed only one step to their common superordinate in order for Ss to decide "same." With instances from two categories, however, Ss have to search the hierarchy up to the "living thing" node before a "same" decision can be made. Similarly, "different" judgments require more time than "same" judgments, because

more of the hierarchy has to be searched before a connection can be found between the two instances.

Schaeffer and Wallace (1970) explained the results of this study (1969, Exp. 1), using a word comparison model. They found that Ss could more rapidly decide that "elephant" and "lion" were living things, than that "elephant" and "daisy" were living things. When the items are presented, they argued Ss retrieve a set of semantic elements which constitute the meaning of these words, including the element "living." All three words have this element in common. However, "elephant" and "lion" also have other elements in common, e.g., "animal," and "mammal." In Schaeffer and Wallace's (1970) terms, there is a greater degree of semantic overlap between "lion" and "elephant." Because these concepts overlap, the threshold for making a "same" response is reached faster, than in the comparison of "lion" and "daisy."

One problem with this explanation is that it was not clear why the determination that "lion" and "elephant" are both animals and mammals is relevant to deciding that they are both living things. The experimental task required only the decision that they both be living things. While it is true that there is more semantic overlap between "lion" and "elephant," than between "lion" and "daisy," Schaeffer and Wallace (1970), by omitting assumptions about structural organization of memory, have not provided Ss with a way of determining the relevance of the semantic overlap. They seemed to be arguing that any kind of semantic overlap can influence the decision, regardless of the relevance of the information.

Schaeffer and Wallace (1970, Exp. 1) reported another study in which they varied the degree of semantic similarity among items when Ss were required to make "different" judgments. Ss were presented with two instances selected from the four categories: tree, flower, bird, and mammal. They were to respond "same" if the instances were from the same category, and "different" if the instances were from different categories. In this task, there were two kinds of "different" judgments, semantically similar "differents," as in the comparison of "hemlock" and "daisy," and semantically dissimilar "differents," as in the comparison of "hemlock" and "parrot." (In the terminology of a hierarchical model, Schaeffer and Wallace, 1970, Exp. 1, varied the semantic distance between nodes when making "false" judgments.)

Schaeffer and Wallace's (1970) word comparison model predicts that the greater the semantic similarity, the more difficult it is to decide that the instances are different. Therefore, RT is longer for semantically similar "different" comparisons than for semantically dissimilar "differents." Their predictions were confirmed. That is, it took longer to decide that "hemlock" and "daisy" were not members of the same category, than to decide that "hemlock" and "parrot" were not members of the same category.

Schaeffer and Wallace's (1970) analysis of the results was as follows. When a word is presented, all the semantic elements defining the word are retrieved. The semantic elements in "hemlock" are compared to those in "daisy," and S finds the element "tree" for the word "hemlock," and the element "flower" for the word "daisy." Therefore, he

has sufficient information to respond "different." However, he also finds that the element "plant" is listed with both words. This semantic overlap confuses the S and makes the "different" response more difficult. Again, it is not clear why these extraneous semantic elements enter into the decision making process to increase response time.

In a second experiment, reported in Schaeffer and Wallace (1970, Exp. 2), it was found that the semantic similarity did not significantly affect RT for "different" judgments when one item was a category instance, and the other was a category name. Judgment time for the pair "hemlock-flower" was very slightly, but not significantly, longer than the judgment time for "hemlock-bird." For an instance-category comparison, Schaeffer and Wallace (1970) argued that Ss do not have to retrieve the semantic elements of the category word. Therefore, there is no overlap between the semantic elements of the instance and the semantic elements of the category to confuse Ss in making a decision.

One problem with Schaeffer and Wallace's (1970) interpretation of experiment two is that they gave no reason why in this particular task the Ss did not retrieve the semantic elements which define the category word. They did not explain why when "hemlock-bird," for example, was presented, Ss did not retrieve the elements which define "bird," but only retrieved the elements which define "hemlock."

Schaeffer and Wallace (1970) noted this problem with their theory. They remarked that in another of their studies (Schaeffer and Wallace, 1969, Exp. 2), when Ss were presented with category names, they apparently retrieved the semantic elements which define the categories, even though

it was not necessary. In Schaeffer and Wallace (1969, Exp. 2), Ss were shown two of the following category names together: bird, mammal, fruit, and grain. If "bird" and "mammal," or "fruit" and "grain," were the categories presented together, then the categories were semantically similar. For all other combinations of these four categories, there was semantic dissimilarity between the two categories presented. Following the presentation of semantically similar or semantically dissimilar categories, Ss were shown a positive instance of one of the two categories, for example, "parrot," and the Ss were to indicate the correct category. Classification times were longer when the categories had been similar then when the categories had been dissimilar. It was easier to classify "parrot" as a bird, after having seen the categories "bird-grain," than after having seen the categories "bird-mammal."

Now Schaeffer and Wallace (1970) could explain this result in terms of overlap of semantic elements, but they could not explain why Ss were retrieving semantic elements for the categories in this experiment, (Schaeffer and Wallace, 1969, Exp. 2), when it was unnecessary to retrieve the semantic elements in the previously reported experiment using category names (Schaeffer and Wallace, 1970, Exp. 2).

There are two problems with their word comparison model then. First, the model does not explain how, when there is semantic overlap, Ss are able to eventually decide which elements are significant for a decision, because their model does not make any assumptions about the structure of semantic information. Second, the model cannot predict when semantic elements are retrieved given that a word has been presented.

Nevertheless, Schaeffer and Wallace (1970) felt that the effects of semantic similarity on "different" judgments provided evidence against the Collins and Quillian (1969) model. The reason for this, in the language of the hierarchical model, is that if semantic distance increases the time to make a "true" response, semantic distance should also increase the time to make a "false" response. This would seem to be the most straightforward prediction of the Collins and Quillian model. This prediction has not generally been confirmed, and, as has been previously discussed, Collins and Quillian (1972a) have had to modify their model to explain how people respond to false sentences.

Meyer's two-stage model of sentence verification. Meyer (1970), on the basis of a study on RT to true and false sentences, presented another model of the sentence verification process. Meyer's is a two-stage model.

Meyer (1970, Exp. 1) presented universal affirmative sentences, i.e., "All (noun) are (noun)." Meyer used four types of logical relations between the subject and predicate noun classes: subset, superset, overlapping, and disjoint. In a subset relation, the subject noun class is a subset of the predicate noun class, as in the sentence "All thrones are furniture." These universal affirmative sentences, where subject and predicate noun classes are in a subset relation, are analagous to Collins and Quillian's (1969) true superset relation sentences. In Meyer's superset relation sentences, the subject noun class is a superset of the predicate noun class, as in "All stones are rubies." In the overlapping relation, the subject and predicate nouns overlap in the set of elements they refer to, as in "All mothers are writers." In the disjoint relation,

the subject and predicate noun classes have no elements in common, as in "All houses are vacuums."

As can be seen from the example, the universal affirmative sentences in which the subject and predicate noun classes are in a superset, overlapping, or disjoint relationship, are all false. It is difficult to determine which of the Collins and Quillian (1969) false sentences these correspond to, since Collins and Quillian defined false sentences in terms of a hierarchical structure. However, the universal affirmatives where subject and predicate are in a disjoint or overlapping relationship seem to be closest to those which Collins and Quillian classified as false superset sentences.

Besides the logical relationship between subject and predicate noun classes, Meyer (1970) also manipulated the size of the classes denoted by the subject and predicate nouns. This manipulation involved replacing the noun with a higher order category. In terms of the four examples presented, increasing subject noun class size gives the following sentences: "All chairs are furniture;" "All females are writers;" "All solids are rubies;" and "All structures are vacuums." Predicate noun class size was also manipulated by replacing the predicate noun with its higher order category.

Increasing predicate noun class size is comparable to what Collins and Quillian (1969) referred to as increasing hierarchical level, or semantic distance. That is, Collins and Quillian used the sentences: "A canary is a bird," and "A canary is an animal," where "canary" is further from "animal" than from "bird" in the semantic hierarchy. Collins and

Quillian manipulated only the predicate noun class size, rather than the size of both subject and predicate noun classes.

Meyer (1970) found that RT varied with the type of logical relationship. The order of RT from fastest to slowest was: disjoint, subset, overlap, and superset.

Increasing subject noun class size (decreasing semantic distance): (1) significantly decreased RT for subset and overlapping relation sentences; (2) significantly increased RT for superset relation sentences; and (3) had no effect on disjoint relation sentences. The decrease in RT for subset sentences is in agreement with Collins and Quillian's (1969) model. Increasing subject noun class size means that the distance between the two nodes in the hierarchy has been decreased. Collins and Quillian predict that for true sentences decreasing semantic distance decreases RT.

Increasing predicate noun class size (increasing semantic distance): (1) significantly increased RT for subset, superset and disjoint relation sentences; and (2) increased, but not significantly, RT for overlapping relation sentences. Again, the effect of increasing predicate noun class size for subset relation sentences was as predicted by the Collins and Quillian (1969) model.

The Collins and Quillian (1969, 1972a) model makes no clear-cut predictions with respect to the false sentences used by Meyer (1970). Falsity, according to Collins and Quillian, is determined by finding contradictions between the sentence and information in the semantic hierarchy. Searching for and finding contradictions depends on the particular semantic information presented, and is not simply related to the

type of logical relationship between subject and predicate nouns, or to the semantic distance between subject and predicate nouns.

On the basis of these results, and Ss' introspections, Meyer (1970) presented a two-stage model of the processes involved in sentence verification. Ss reported that they first tried to decide if the two words were at all related. Meyer stated that this is the first stage of sentence verification, attempting to find a relation between the subject and predicate nouns. If they are not related, as in the disjoint relation sentences, Ss respond "false." If they are related in some way, Ss must then decide during stage two if the subject noun is a subset of the predicate noun class. If so, he responds "true." If not, he responds "false" to the overlapping and superset relation sentences. During stage one, then, Ss discriminate between disjoint relations and the other three, and during stage two, Ss discriminate between subset relations and the other two.

Meyer (1970) tested this two-stage model by using a task that would theoretically eliminate stage two, thereby decreasing RT. The verification of particular affirmative sentences, i.e., "Some (noun) are (noun)," is such a task. When the sentences are particular affirmatives, Ss have only to decide if the subject and predicate nouns are related. If they have any elements in common, that is, if they intersect, Ss can respond "true." If they do not have any elements in common, as in the case of disjoint relations, Ss can respond "false." For the particular affirmative sentences, only differentiation between disjoint relation sentences and the other three is necessary for correct responding. Stage two should be eliminated.

If the two-stage model is correct, Meyer (1970) predicted the following effects. (1) For disjoint relation sentences, RT to particular and universal affirmatives should be the same. This follows from the idea that stage one only is required for differentiating between the disjoint sentences and the other three. (2) For the subset, superset, and overlapping sentences, RT to particular affirmatives should be less than the RT to universal affirmatives. Again, the reason for this prediction is that stage two is eliminated for particular affirmatives, but not for universal affirmatives. Eliminating stage two, should decrease RT. (3) For disjoint relation sentences, changes in subject and predicate noun class size will affect RT to particular affirmatives in the same way as they were found to affect RT to universal affirmatives.

Consequently, Meyer (1970) presented Ss with particular affirmative sentences, for "true-false" judgments. Type of logical relation, as well as subject and predicate noun class size were varied.

Type of logical relation affected RT for the particular affirmatives. The ordering of RTs from fastest to slowest was: subset, superset, overlapping, and disjoint.

Increasing subject noun class size: (1) significantly decreased RT for subset relation sentences; (2) significantly increased RT for superset and disjoint relation sentences; and (3) did not significantly change RT for overlapping relation sentences. Increasing predicate noun class size: (1) significantly decreased RT for superset and disjoint relation sentences; (2) significantly increased RT for subset relation sentences; and (3) did not significantly change RT for overlapping relation sentences.

Predictions (1) and (2) of the two-stage model were confirmed. RT to particular affirmatives was not significantly different from RT to universal affirmatives for disjoint relation sentences. For the subset, superset, and overlapping relation sentences, however, RT to particular affirmatives was significantly less than RT to universal affirmatives.

Prediction (3) concerning the effect of changes in subject and predicate noun class size on disjoint relations was not completely confirmed. Increasing predicate noun class size increased RT for both universal and particular affirmatives. However, increasing subject noun class size increased RT for particular affirmatives, and had no significant effect on universal affirmatives. The two-stage model predicts that increasing subject noun class size should not have affected RT to particular affirmatives, since it did not have an effect on universal affirmatives. Meyer (1970) attributed this non-confirmation of the model to reading errors which Ss could have made on the particular affirmatives. He said that in reading particular affirmatives, Ss may have interchanged or confused the subject and predicate nouns. For example, they might have read "Some vacuums are houses," instead of "Some houses are vacuums." This type of error would not change the truth value of particular affirmatives, as it would have for universal affirmatives, and in reading the universal affirmatives, perhaps Ss were more careful not to make this kind of interchange. If, however, they did make the interchange, then when the E increased subject noun class size, S changed it into an increase in predicate noun class size. Since increasing predicate noun class size increased RT for universal affirmatives, then it is not surprising that the

S-defined increase in predicate noun class size increased RT for particular affirmatives.

Meyer (1970) concluded that the two-stage model was supported by the results of experiment two. Furthermore, he used the category size effects to determine what kind of retrieval processes are involved in each stage, and what kind of information is stored in memory.

Stage one, it will be recalled, involves determining whether the elements in the subject and predicate noun categories intersect. This stage, according to Meyer (1970), requires that S retrieve the names of categories which intersect with the predicate noun category. S then compares these categories to the subject noun. If one of the predicate intersections matches the subject noun, S responds "true," or if universal affirmatives are presented, S goes on to stage two. Stage one is referred to as the predicate-intersections stage.

Meyer (1970) further speculated about the organization of the predicate-intersections file which S consults. Figure 4 (from Meyer, 1970) shows the quasi-hierarchical structure he postulated. If the

---

Insert Figure 4 about here

---

predicate noun is "stone," S retrieves the following kinds of items. S retrieves items which are supersets of the predicate noun, like "solid." S retrieves subsets of "stone," like "gem." S retrieves items like "gift" which overlap with "stone." Items closest to "stone," like "solid" and "gem," are retrieved first for comparison to the subject noun.

Meyer (1970) stated that this structure is very similar to the one postulated by Collins and Quillian (1969). However, the structure differs from the Collins and Quillian model in that "if they exist, these pointers [to categories which intersect the predicate noun] probably differ in an important way from those Collins and Quillian . . . propose: they apparently cannot be used for searching either the names of P- [predicate] supersets alone or those of P-subsets . . . . In fact, these pointers may convey no precise information about set relations [Meyer, 1970, p. 274]." Thus, Ss do consult a quasi-hierarchical structure, but they use the structure only to determine whether there is a relationship between the two categories, but not the type of relationship between the two categories. Determination of the type of relationship involves other processes which occur during stage two.

This description of how the predicate-intersections file is used raises a problem for the processing of disjoint relation sentences. Presumably, Ss consult the intersections file to determine that certain categories, e.g., "typhoons" and "grains" are disjoint. In Meyer's (1970) terms, this would mean there is no path relating these two words in the intersections file. Using Webster's New World Dictionary (The World Publishing Co., 1960), the writer found the following relationship between "typhoon" and "grain": typhoon--wind--air--gas--substance--matter--particle--seed--grain. This relationship, while it is long, and while it probably uses superset, subset, and overlapping relations, is nevertheless a relationship, or path between the two words. If S retrieved this relationship during stage one, he would respond erroneously. In short, the

predicate-intersections file concept does not provide S with sufficient information to decide that "typhoon" is not "grain."

Meyer (1970) also tested a number of models describing stage two, wherein subset relation sentences are differentiated from the superset and overlapping relation sentences. Three potential models were tested: the exemplar model, the attribute model, and the exemplar-attribute model.

In the exemplar model, S retrieves examples of the subject and predicate noun categories, and responds "true" if each subject exemplar is also a predicate exemplar.

In the attribute model, S retrieves defining attributes of the subject and predicate noun categories, and responds "true" if each predicate noun attribute is also an attribute of the subject noun.

In the exemplar-attribute model, both exemplars and attributes are available for retrieval. S responds "true" if the attributes of all subject exemplars match the attributes of the predicate noun.

The models were tested by deriving their predictions for the effects of category size on stage two duration. Stage two duration was estimated by the subtraction method. That is, the RT to universal affirmatives where the subject and predicate are in a subset, superset, or overlapping relationship can be broken down into three components: (1) the time to complete stage one; (2) the time to complete stage two; and (3) the time duration of any additional processes required by the task, like sentence encoding.

The RT to particular affirmatives where the subject and predicate are in a subset, superset, or overlapping relationship can be broken down

into only two components: (1) the time to complete stage one; and (2) the time duration of any additional processes required by the task. That is, for particular affirmatives, stage two is omitted. Therefore, stage two duration can be estimated by subtracting the mean RT to particular affirmatives from the mean RT to universal affirmatives when the subject and predicate noun classes are in a subset, superset, or overlapping relationship.

It was found that the attribute model made the most accurate predictions. The model predicts that increasing subject noun class size does not change the duration of stage two for subset relation sentences. The attribute model makes no predictions for superset and overlapping relations. The attribute model also predicts that increasing predicate noun class size decreases stage two duration for subset relation sentences. Both of these predictions were confirmed.

Meyer's (1970) results provided confirming evidence for the Collins and Quillian (1969) model for those cases where the Collins and Quillian model makes clear-cut predictions. That is, for true universal affirmative sentences the effects of increasing subject and predicate noun class size were as predicted by the Collins and Quillian model. Increasing subject noun class size decreased semantic distance and, therefore, should have decreased RT. Increasing predicate noun class size increased semantic distance and, therefore, should have increased RT. Both effects were obtained.

It is not possible to make predictions about the false sentences Meyer (1970) studied, using the Collins and Quillian (1969) model, for reasons previously stated.

With particular affirmatives there is a change in the syntax of the sentence, and Collins and Quillian have not discussed the effects of this type of change. However, it should be noted that for particular affirmative sentences, where subject and predicate are in a subset relation, the results Meyer (1970) obtained for changes in subject and predicate noun class size were the same as those he obtained for universal affirmatives. Increasing subject noun class size decreased RT, and increasing predicate noun class size increased RT. These results are consistent with the Collins and Quillian (1969) model.

Collins and Quillian (1972) have commented that Meyer's (1970) model is similar to theirs. They stated that "comparing concepts involves a semantic search outward in parallel from both the concepts to all associated properties, including superset properties [Collins and Quillian, 1972, p. 330] ." This search seems roughly analogous to stage one of Meyer's model. Then "any connection found must be checked to see if the relation between the concepts meets the constraints of syntax and context (including the instructions) [Collins and Quillian, 1972, p. 330]." This might correspond to stage two in Meyer's model.

In fact, Collins and Quillian (1972), stated "we doubt that our differences with his model are very substantive except in one respect. He considers several decision strategies . . . . But he treats decision strategies as if people use one of the strategies consistently, at least in any one task, whereas we are arguing that the decision strategy will depend on the connections found. Our position weakens the kind of experimental predictions that can be made, but we think it is unavoidable [p. 334]."

Frequency freaks. A possible source of confounding in this type of research has to do with word frequency effects. Three potential word frequency effects have been considered as contributing to the obtained RT differences.

Howes (see Miller, 1951) reported that high frequency words can be recognized more quickly than low frequency words. Individual word frequencies may, therefore, have contributed to RT differences obtained in the sentence studies, or to the RT differences obtained in the word classification studies. The argument would be that a sentence like "A canary is a bird" is easier than a sentence like "An ostrich is an animal," because "canary" has a higher frequency than "ostrich." This particular type of frequency effect has been controlled for, or found not to have contributed much to RT differences.

In their first study on the hierarchical organization of memory, Collins and Quillian (1969) did not control individual word frequencies. However, they subsequently computed a weighted average of the individual word frequencies for each sentence, using Thorndike and Lorge (1944) word tables. They reported that the predicate nouns in the superset sentences tended to have high Thorndike-Lorge frequencies. This difference in word frequency could have affected the finding that superset sentences had lower RTs than property sentences. This argument is supported by the results of Kintsch, et al. (1970). Kintsch, et al. controlled word frequency and found no RT difference between superset and property sentences.

Meyer (1970) reported that "an attempt was made . . . to equate the lengths and frequencies (Thorndike and Lorge, 1944) of: (1) S-category names within pairs where S-size was varied, (2) P-category names within pairs where P-size was varied, and (3) both S- and P-category names between pairs involving different set relations [p. 261]."

Landauer and Freedman (1968) and Schaeffer and Wallace (1969, 1970) did not report controlling word frequencies in their word classification studies. However, Collins and Quillian (1970a) partially replicated the Landauer and Freedman study and did control word frequency. In constructing lists of positive and negative instances of the categories "dog," "bird," and "animal," Collins and Quillian (1970a) matched word frequencies among the three lists of positive instances, and between lists of positive and negative instances.

Since semantic distance has been shown to affect RT in studies where word frequency has been controlled, (e.g., Meyer, 1970), it is concluded that word frequency, per se, is not the determining variable. However, word frequency may have been the basis of the RT difference between property and superset sentences.

Wilkins (1971) argued that in sentence classification and word classification studies, controlling individual word frequencies is an inadequate procedure. A potentially more important frequency effect, Wilkins stated, is conjoint frequency. The "frequency of co-occurrence of category and instance in English usage is a critical variable. This conjoint frequency can, of course, be independent of the individual frequencies of the category and instance [Wilkins, 1971, p. 382]."

Wilkins (1971) suggested that conjoint frequency may be estimated by using the Connecticut word association norms. These norms were developed by presenting Ss with a category, and having Ss produce instances of that category.

Conjoint frequency provides an alternative explanation to semantic distance in accounting for RT effects. According to a conjoint frequency explanation, the sentence "A canary is a bird" has a faster RT than "A canary is an animal," because "canary" has a higher frequency of co-occurrence with "bird" than with "animal."

Wilkins (1971, Exp. 1) tested this hypothesis in a word classification task. Ss were shown a category name, then a positive or negative instance, and were required to make a "yes-no" classification response. The Connecticut word norms were used as an index of conjoint frequency. Positive instances of a category were selected, and these instances had either a high or low frequency of co-occurrence with the category word. The Thorndike-Lorge frequency of the individual words was held constant. Wilkins found that instances with a high conjoint frequency of occurrence had faster classification response times than instances with a low conjoint frequency of occurrence.

In a second experiment, Wilkins (1971, Exp. 2) varied the frequency of occurrence, using Thorndike-Lorge tables, while keeping conjoint frequency constant. There was no significant difference in classification times for the more frequent versus the less frequent words.

Wilkins' (1971) results indicated that conjoint frequency of category and instance, and hierarchical organization of words in memory, are

potential explanations of the observed increase in RT between sentences like "A canary is a bird" and "A canary is an animal." His results also supported the previously stated conclusion that individual word frequency is not an important determinant of RT for superset sentences.

The conjoint frequency of categories argument is relevant to sentences Collins and Quillian (1969) classified as superset sentences. An analogous argument has been made with respect to Collins and Quillian's property sentences, e.g., "A canary is yellow," and "A canary has wings." One could argue that for the category "canary," the properties "is yellow" and "has wings" are both directly associated with "canary," but the two properties differ in their frequency of co-occurrence. Property frequency, rather than semantic distance might be the basis of RT difference between the sentences. Conrad (1972) has presented this argument as an alternative hypothesis to Collins and Quillian's (1969) hypothesis of nonredundant storage, or cognitive economy.

Conrad's (1972) criticism of the cognitive economy concept was a frequency argument. She had Ss write descriptions of categories like "canary," "bird," and "animal." The frequency of occurrence of each property was then rated as high, moderate or low, depending on how many Ss stated a particular property. Her Ss were describing many of the same words used by Collins and Quillian (1969). Conrad then classified Collins and Quillian's level 0, level 1, and level 2 property sentences with respect to how frequently these properties were produced in her Ss' descriptions of a category. Conrad reported that the properties Collins and Quillian classified as level 0 were most frequent in Ss' descriptions.

Those properties classified as level 1 properties were less frequently produced, and those classified as level 2 properties were least frequently produced. According to Conrad's analysis, the difference between level 0, 1, and 2 properties, is not where they are stored in the hierarchy, but how difficult these properties are to retrieve with respect to a given category. Presumably, retrieval difficulty depends on how frequently these properties have occurred with a particular category in Ss' experience.

To test this idea, Conrad (1972, Exp. 1) used Ss' descriptions of categories from three levels of a hierarchy, e.g., "canary," "bird," and "animal." For each level, properties were grouped with respect to frequency of occurrence. True property sentences were then constructed using only level 0 categories, e.g., "canary," paired with high, moderate, and low frequency properties from each level of the hierarchy.

Conrad (1972) found that property frequency had a significant effect on RT. RT increased as property frequency decreased. There was no main effect for hierarchical level, however. This absence of a main effect for level is not unexpected. The significance test reported included true and false sentences. As previously mentioned, RT does not increase with hierarchical level for false sentences, but only for true sentences. Conrad confirmed this finding. She reported a significant level by truth value interaction. For true sentences, RT increased slightly with increasing hierarchical level, but for false sentences RT decreased with increasing hierarchical level. She also reported that the increase in RT with hierarchical level for true sentences, was strongest for moderate and low frequency properties. For high frequency property sentences, RT

increased from level 0 to level 1, but decreased from level 1 to level 2.

Conrad (1972) concluded that Collins and Quillian (1969) confounded hierarchical level and property frequency. She stated that the level 0 properties which Collins and Quillian used were high frequency properties, that their level 1 properties were moderate frequency properties, and their level 2 properties were low frequency properties.

Thus far, Conrad's data do not indicate that property frequency rather than hierarchical level determines RT. The data showed that both property frequency and hierarchical level contribute to RT for property sentences.

The method Conrad (1972) used to determine property frequency is subject to criticism. It is possible that differences in property frequency are dependent upon hierarchical organization of property information. In describing the category "canary," for example, Ss might start with properties peculiar to canaries, then some Ss might go on to list properties of birds, and fewer Ss might go on to list properties of animals as well. From a structural point of view, it is the frequency effects which are explained by assuming cognitive economy, rather than cognitive economy being explained in terms of frequency. (A similar argument may be made for the category frequency hypothesis. Wilkins', 1971, data on frequency of co-occurrence of categories may be explained in terms of hierarchical organization of categories.)

Conrad (1972) seemed to be aware of this criticism in that she stated that the results "are subject to the criticism that the selection

of properties to be assigned at any given level might be biased [p. 153]." She reported a second experiment which is a more adequate test of the property frequency versus cognitive economy hypotheses. In this study, Conrad manipulated the level of the superordinate noun, while keeping property level constant. For example, RTs to the following sentences were compared: "A canary can move," "A bird can move," and "An animal can move." This manipulation of subject noun class size is similar to Meyer's (1970) manipulation of subject noun class size for superset sentences like "A canary is an animal," and "A bird is an animal." According to the Collins and Quillian (1969) model, RT should decrease as the hierarchical level of the subject noun increases.

This prediction was not confirmed. Conrad (1972, Exp. 2) found no consistent decrease in RT as the number of steps separating the subject noun and property decreased. She concluded that while there is evidence for the hierarchical organization of semantic categories, properties are stored directly with every relevant category. In those studies where cognitive economy of storage was found, Conrad said that the effect was due to property frequency. This conclusion seems to contradict her own finding that both hierarchical level and frequency affected RT in the first experiment. Furthermore, the sentences used in this second experiment were produced by using Ss' descriptions, and the previous criticism still seems relevant.

An experiment reported by Collins and Quillian (1972a) is relevant to the property frequency hypothesis. Collins and Quillian tested their model in a manner similar to Conrad. Using the property sentences of

their original study (1969), they simultaneously manipulated hierarchical level of the subject noun and the property. That is if the sentences in the original study were: "A canary can sing," "A robin can fly," and "An eagle has skin," then some of the sentences were transformed so that there was only one step separating the subject noun and the property. The transformed sentences were: "A canary can sing," "A bird can fly," and "An animal has skin." The cognitive economy hypothesis predicts that RT for these three types of sentences should be equal, since the semantic distance between the subject noun and the property is equal for all three sentences.

This prediction was not confirmed. RT to the transformed level 1 property sentences like "A bird can fly," was significantly less than RT to the other two types of property sentences. Collins and Quillian (1972a) attributed this finding to a word frequency effect. They stated that the subject nouns in the transformed level 1 sentences were high frequency words. The use of a high frequency word reduces RT because it facilitates word recognition.

Although Collins and Quillian (1972a) did not find equality in RT among the three transformed sentences, reducing semantic distance did have an effect. The RTs for the transformed level 1 and level 2 sentences were lower than the RTs to the original sentences. Thus, "A bird can fly," was responded to faster than "A robin can fly," and "An animal has skin" was responded to faster than "An eagle has skin." Collins and Quillian attributed this effect to a reduction in semantic distance, but it may also be due to a word frequency effect.

While this result does tend to confirm the cognitive economy hypothesis, it is not inconsistent with Conrad's (1972) property frequency hypothesis. Conrad stated that the retrieval difficulty of properties is not just a function of the specific property. Retrieval difficulty for any given property can vary with the category. For example, "has skin" may be a difficult property to retrieve for the category "canary," but it may be an easy property to retrieve for the category "animal." Thus, she would not argue that for the transformed sentences "A bird can fly," and "An animal has skin," RT will increase as it did in the original study by Collins and Quillian (1969).

In summary, it seems that the results of these three experiments indicate that there is not complete economy of property storage. Properties like "can fly" may be stored with "canary" as well as with "bird." In discussing the development of the model, it has already been noted that Collins and Quillian (1972b) have modified the model. "There is nothing in the theory, however, that prevents storing superset properties with particular instances, and we certainly think it is a common practice [Collins and Quillian, 1972b, p. 345]." Rather than stating that property sentences must be responded to on the basis of inference, i.e., by traveling through the hierarchy, they state that they were merely trying to demonstrate that such may be the case. In fact, "in constructing sentences for our original study (Collins and Quillian, 1969) we made an effort to choose instances (e.g., wren) where the superset property (e.g., has wings) was not particularly associated with the instance.

Hence, by design the sentences used were ones likely to be decided by inference [Collins and Quillian, 1972b, p. 345]."

The Collins and Quillian position, then, with respect to criticisms of their model seems to be that they are developing a model of competence rather than of performance in one experimental situation.

Present Experiments

The present experiments concerned the nature of the tokens which define a type node in the Collins and Quillian model. In their experiments, Collins and Quillian (1969, 1970b) distinguished between property and superset sentences. This distinction parallels the two types of tokens which define a type node or word. It will be recalled that the first token always provides superset or superordinate information. This is followed by a set of "properties stating how the superset must be modified to constitute the concept [word] intended [Quillian, 1969, p. 462]."

In testing the model, these two types of information were always presented in syntactically different forms. Sentences which test superset information were of the form "A (noun) is a (noun)." Property sentences had three possible forms: (1) "A (noun) is (adjective)"; (2) "A (noun) has (noun)"; and (3) "A (noun) can (verb)." Superset sentences were found to have faster RTs than property sentences. Kintsch, et al. (1970), however, were not able to replicate this effect. They found no difference between property and superset sentences in RT.

Both of these studies, however, were subject to a confounding. This confounding was between the "idea" being presented, and the syntactic realization of this idea. When Collins and Quillian (1969) or

Kintsch, et al. (1970) compared RTs to "A shark is a fish" and "A shark is vicious," they were not only changing syntax, but they were also changing semantic content. The question raised was whether one can separate syntactic form from the semantic content of a sentence.

One way to separate syntactic form from semantic content is to keep semantic content constant, and vary syntactic form. That is, one can present the same idea in two syntactic forms. For example, one can present the sentences "A teacher is an educator" and "A teacher can educate." Here we have the same idea, or at least very similar ideas, presented in two syntactic forms.

RTs to sentence pairs like this would enable one to distinguish between two possible explanations of the Collins and Quillian (1969) finding that superset sentences were faster than property sentences. The two explanations are referred to as the syntactic hypothesis and the semantic hypothesis.

The syntactic hypothesis states that the previously observed difference in RT between property and superset sentences was due to their syntax. This hypothesis predicts that if the same idea is presented in two syntactic forms, then the superset realization of the idea is easier. That is, "A teacher is an educator" has a faster RT than "A teacher can educate."

The semantic hypothesis states that the previously observed difference in RT between property and superset sentences was due to their semantic content. This hypothesis predicts no difference in RT between

the superset realization, "A teacher is an educator," and the property realization, "A teacher can educate."

This paper presents two experiments. The first study was a replication of Collins and Quillian's (1969) study. The results were basically similar to those of Collins and Quillian. Since their general findings proved reliable, a second experiment was done to test the semantic and syntactic hypotheses.

In the second experiment, two-level hierarchies were constructed. For each hierarchy, there were property relation and superset relation sentences, where the terms property and superset relation refer to the syntactic form of the sentence. There were two forms of each hierarchy. In form A, a particular idea was expressed as a property sentence, and in form B, that same idea was expressed as a superset sentence. All Ss were presented with forms A and B, so that they responded to the same idea in both syntactic forms.

The important effect for distinguishing between the two hypotheses was the presence of a main effect for type of relation. The syntactic hypothesis predicts a main effect for type of relation, specifically a faster RT for superset relation sentences. The semantic hypothesis predicts no difference in RT between property and superset relation sentences.

Experiment 1: A Replication of Collins and Quillian (1969)

A pilot study replicating the Collins and Quillian (1969) experiment was carried out, with two changes in design. Syntactic form was balanced within the property sentences, so that there were equal numbers of "can

(verb)," "has (noun)," and "is (adjective)" sentences. The second change concerned the variability in the predicates of the false sentences. All the false sentences were constructed in a manner similar to that used by Collins and Quillian (1969) where the level of contradiction was controlled. A sample of these sentences was shown in Table 3.

It was thought that one explanation for the RT effects to false sentences found by Collins and Quillian (1969) was that the false sentences covered a wider range of the entire lexicon than did the true sentences. In verifying true sentences, Ss may learn to narrow down the area of the lexicon to be searched. Although Collins and Quillian (1969) randomly presented sentences from three or four hierarchies of the lexicon within a trial block, once the S encounters a true sentence, he can focus on a particular area of the lexicon to search.

For example, if the first true sentence S sees is "An oak is a tree," he can predict that within the same trial block he will again have to search the "tree" area of the lexicon to find verifying information about particular kinds of trees. S can also predict that he will have to search for properties of trees in general, and if a three-level hierarchy were used in the experiment, he will have to search for properties of plants.

If, in the false sentences, Ss are trying to find the node which contains the information in the predicate of the sentence, then the search cannot be narrowed down to one area of lexicon. Assume that S is given the false P1 sentence, "A salmon can fly." This sentence predicates a property of birds to a particular fish. S can first check the lexicon

for properties of salmon, then for properties of fish, and not find the relevant information about what animal does fly. S can stop searching now and respond "false," or he can continue searching the lexicon to locate the node which does contain the property "can fly." If he continues to search, he has no idea of the level at which to continue the search. S can check properties of particular fish like "shark" or "tuna," or he can check properties at the level of "bird" or "mammal," or he can continue moving up the lexicon from "fish" to "animal" to "organism."

Given a false sentence, if S tries to locate the contradictory information, there are many possible locations to search. Variability in the location of contradictory information can lead to variability in search strategies and consequently to variability in RTs.

To test this possibility, two forms of the sentences used in this study were constructed. The main form is described first, and then the alternate form.

Figure 5 shows a hierarchy from which true and false sentences about kinds of fish were constructed. It was assumed that "fish," "mammal," "insect," and "reptile" are equivalent categories in terms of their level in the hierarchy. That is, they all have "animal" as their next level category node. Table 4 shows sentences that were constructed from the hierarchy.

Insert Figure 5 about here

Insert Table 4 about here

A false P0 sentence about a particular fish was constructed by using a property of another fish, mammal, insect or reptile. "A herring is dangerous" predicates a property of sharks. "A trout has a shell" predicates a property of turtles. This second sentence was also considered a false P0 sentence in that "trout" and "turtle" are at the same level of the hierarchy. False P1 sentences about kinds of fish were constructed by predicating a property of mammals, insects, or reptiles. The sentence in Table 4, "A goldfish can crawl," predicates a property of reptiles.

False S0 sentences were constructed by using the name of another fish, mammal, insect, or reptile. "A swordfish is a turtle" and "A swordfish is a flounder" were both considered false S0 sentences, since "swordfish," "flounder," and "turtle" are at the same level of the hierarchy. A false S1 sentence stated that a particular fish was a mammal, insect, or reptile.

Table 5 shows the eight sentences constructed as the alternate forms of those in Table 4. These sentences were constructed by changing

---

Insert Table 5 about here

---

the subject nouns of the sentences in Table 4, so that the true sentences are now false, and false sentences are now true. The effect of this change, it was hoped, would be to increase the area of the lexicon which the true sentences cover. The true sentences in Table 5 have fish names, reptile names, and insect names. With this form, if the S first sees "An ant is an insect," he is not able to predict that the following true sentences are also about specific kinds of insects. The primary purpose of using this alternate form was to get more variability in the true sentences so that Ss would not be able to predict what part of the lexicon the true sentences were from.

Something of the opposite effect holds for the false sentences in Table 5. If S first sees "A horse is a sardine," then he can predict that the following predicates all come from that part of the lexicon relating to fish, and so restrict his search of the lexicon somewhat.

Method

Subjects. Fifteen male and female undergraduates at the University of Massachusetts participated as Ss to fulfill an introductory psychology course requirement. Three Ss were eliminated because of equipment failures during the experimental session.

Apparatus. Sentences were typed in capital letters on index cards and slides were made of each card. Sentences were presented one at a time by Kodak Carousel slide projector on an 18" x 24" screen. Each slide showed white letters on a gray-black background. There were 10 to 29 letters per sentence. Ss sat in a sound-proofed room facing the screen. On the table in front of S was a six-button response panel. Two of the response buttons

were labeled "TRUE" and "FALSE." Labeling of the right- and left-hand response buttons was balanced across Ss. Below the screen a set of seven alphanumeric Nixie tubes was used to inform Ss of the beginning and end of a one-minute rest period between trial blocks. Presentation of stimuli and recording of responses was controlled by programming of a DEC PDP-8/I computer.

Sentences. Twelve two-level hierarchies were constructed. Many of the true sentences were those originally used by Collins and Quillian (1969) or were very similar. None of the false sentences used by them were available, so these were constructed by predicating false properties or category names at appropriate levels of the hierarchy. For each hierarchy, there were eight sentences representing the following conditions: (1) true property relation level 0, (2) true property relation level 1, (3) true superset relation level 0, (4) true superset relation level 1, (5) false property relation level 0, (6) false property relation level 1, (7) false superset relation level 0, and (8) false superset relation level 1.

For the property sentences, one-third were "can (verb)" sentences, one-third were "has (noun)" sentences, and one-third were "is (adjective)" sentences. These three verb phrases appeared equally often for true and false property sentences at levels 0 and 1.

From the initial set of twelve hierarchies, an alternate form was constructed by the method previously described, so that an originally false sentence was made true. For the original set of hierarchies, the true sentences covered a narrow range of the lexicon, and in the alternate form, the true sentences covered a broader range of the lexicon. The

original set of hierarchies is referred to as the narrow range hierarchies, and the alternate form as the broad range hierarchies.

Four additional sentence hierarchies, two narrow, and two broad, (32 sentences) were constructed for practice trials.

To check on E's classification as true or false, the entire set of 224 sentences, including practice sentences, was presented to eleven psychology graduate students. Sentences were typed on sheets of paper, sixteen sentences to a sheet. Each sentence was followed by a "T" or "F." Ss were told to read each sentence and circle "T" or "F" if the sentence was generally true or false. Ss were told to go through the list of sentences fairly rapidly and not to spend too much time on any one. If they could not come to a decision easily, they were to leave the item blank.

If any sentence received two out of eleven answers which differed from E's classification, or were left blank, these sentences were changed or new ones were substituted for them. Thirty-two sentences were changed for either of these reasons. There was no further check on the thirty-two sentences substituted for those deleted. The final set of 224 sentences used is shown in Table A of the Appendix.

Procedure. Practice and experimental sentences together constituted seven blocks of thirty-two sentences per block. Following each block of sentences, "BREAK" appeared for seven seconds on the Nixie tubes. After approximately 45 seconds, "READY" appeared on the Nixies for seven seconds, and the next trial block of sentences was shown. Each sentence appeared for two seconds, followed by a blank screen for two seconds. S could respond any time within the four second interval. An additional 300 msec.

delay followed the four second interval. This delay was interposed between trials because of problems in programming the slide projector for faster presentation rates. Responses occurring within this delay were not recorded.

Ss were tested individually. Each S was told that sentences would be presented for about two seconds followed by a blank screen for about two seconds, and that they should respond as fast as they could within this interval. Ss were told to decide if the sentence was generally true or generally false as rapidly and as accurately as possible, and then to press the appropriate response button. Ss were told that there was a rest period indicated by "BREAK" and "READY" on the Nixie tubes. At the end of the practice block, E asked for questions. A few Ss were still not sure how far to push the meaning of "true" and "false" and were told not to search for remote or metaphorical meanings.

All Ss were shown 224 sentences. The same practice set was used for all Ss. Practice sentences were presented in a different random order for each S.

Following the practice block, six Ss were shown the 96 sentences from the twelve narrow hierarchies followed by the 96 sentences from the broad hierarchies. The other six Ss were shown the sentences from the broad hierarchies first, followed by the sentences from the narrow hierarchies. The twelve hierarchies in each form were randomly divided into blocks of four hierarchies. The 32 sentences thus constituting a block were randomly presented. Order of the three blocks of 32 sentences was counterbalanced across Ss.

Results

Data corrections. In analyzing the RT data, a constant of 28 msec. was subtracted from all scores. This was done to compensate for a lag in shutter opening time.

The following RT data are for correct responses only. For those trials on which S made an error or failed to respond, his mean RT for that condition was substituted. Overall error rate was 7%, ranging from 0.5% to 15.4%. This error rate includes failure to respond as well as wrong responses. Mean error rate for wrong responses only was 6.2%. Error rates for experimental conditions are presented later.

Reaction time for correct responses. Figure 6 shows the RTs for true and false, property and superset relation sentences, at levels 0 and 1, averaged over broad and narrow hierarchies. The interaction among type

---

Insert Figure 6 about here

---

of relation (property and superset), truth value (true and false), and level (0 and 1) was significant, $F(1, 10) = 8.81$, $p < .025$, when analyzed in an overall analysis of variance. (A fuller presentation of the results of the analysis of variance appears later.)

For the true sentences, property relation sentences had a higher RT than superset relation sentences. Both property and superset relation RTs increased from level 0 to level 1, the increase being greater for superset relation sentences. The effects shown in Figure 6 for true sentences were fairly consistent for individual Ss and hierarchies. The relative

position of the two curves was as predicted for eight out of twelve Ss, and for nine out of twelve hierarchies.

The difference between P0 and P1, 124 msec., is an estimate of the time to travel from a node to its superset node. In Collins and Quillian (1969), this time was 75 msec. The increase in RT of 157 msec. from S1 to P1, is an estimate of the time to retrieve a property from a node. The time estimated for this process by Collins and Quillian was 225 msec.

In the false sentences, there was a slight increase of 11 msec. from level 0 to level 1 in RT for both property and superset relation sentences. RT to property relation sentences was about 72 mesc. longer than RT to superset relation sentences at both levels. While the changes in RT for false sentences were consistent with those found for true sentences, the effects were not stable across Ss or across hierarchies.

False property relation sentences were consistently higher than superset relation sentences for six out of twelve Ss. For the other six Ss, there were cross-over effects, with property relation sentences having higher RTs at one level and superset relations having higher RTs at the other level. RTs to false property relation sentences did not always increase from level 0 to level 1 for all Ss. For eight Ss, there was an increase in RT from level 0 to level 1, and RT decreased from level 0 to level 1 for four Ss. RTs for false superset relation sentences increased from level 0 to level 1 for six Ss, and decreased from level 0 to level 1 for six Ss.

RTs to false sentences for the individual hierarchies showed consistently higher RTs to property relations at both levels for five out of twelve hierarchies. For five hierarchies, property relations had higher

RTs at one level, and for two hierarchies, superset relation RTs were higher at both levels. RTs to property relation sentences increased from level 0 to level 1 in seven out of twelve hierarchies, and decreased from level 0 to level 1 in five out of twelve hierarchies. RTs to superset relation sentences increased from level 0 to level 1 in six hierarchies, and decreased in six hierarchies.

Overall analysis of variance results on RT data. Table 6 shows the results of an analysis of variance on RT for the following variables: (1) range of the hierarchy, narrow and broad; (2) order of presentation of the forms, narrow then broad and broad then narrow; (3) type of relation, property and superset relation; (4) truth value, true and false; (5) hierarchical level, level 0 and level 1; and (6) hierarchy, twelve hierarchies covering different semantic content areas.

Insert Table 6 about here

The following significant main effects were found. (1) RT to superset sentences, 1365 msec., was significantly lower than RT to property sentences, 1525 msec., $F(1,10) = 42.32$, $p < .001$. (2) RT to true sentences, 1390 msec., was significantly less than RT to false sentences, 1500 msec., $F(1,10) = 46.72$, $p < .001$. (3) RT significantly increased with hierarchical level, from 1389 msec. for level 0, to 1501 msec. for level 1 sentences, $F(1,10) = 111.04$, $p < .001$. (4) RT varied significantly among the twelve hierarchies, $F(11, 110) = 6.47$, $p < .025$. RT varied from 1314 msec. for hierarchy four, to 1583 msec. for hierarchy five. Hierarchy four (see

Table A of Appendix) was based on birds, and hierarchy five was based on medical specialities. Hierarchy five contained the longest sentences used.

Neither the range of the hierarchies, narrow or broad, nor the order in which they were presented had a significant effect. However, these two variables did interact significantly, $F(1,10) = 14.69$, $p < .005$. The narrow form had the longest RT when presented first, 1529 msec., and the shortest RT, 1368 msec., when presented second. In contrast to this decrease of 261 msec., the RT for the broad hierarchies decreased only 72 msec., from 1477 msec. when they came first, to 1405 msec. when they came second.

The range variable was included in the present study to determine if the changes in RT for false sentences found by Collins and Quillian (1969) were related to differences in search strategies for true and false sentences. It was thought that with false sentences, Ss search a wider range of the lexicon to find contradictions. This hypothesis would be supported if range of hierarchy interacted with truth value, or if range interacted with truth value, level, and type of relation. Neither of these interactions was significant.

For both narrow and broad hierarchies, RT to true sentences was 1390 msec. For false sentences, RT was 1508 msec. for narrow, and 1492 msec. for broad hierarchies. This difference, while in the expected direction, is small.

As stated, the type of relation by level by truth value interaction depicted in Figure 6 did not change significantly when the range variable was included. For true sentences, the means were essentially the same for

both narrow and broad hierarchies. For false sentences, there were slight changes in RT for narrow and broad hierarchies.

Table 7 shows the cell means for all combinations of range, order of presentation of forms, type of relation, truth value, and hierarchical level.

Insert Table 7 about here

RT data for property relation sentences. An analysis of variance was done on the property relation sentence RTs only, to determine if the three verb phrases used had different effects on RT. The variables in the analysis were: (1) range, narrow and broad; (2) order of presentation of ranges, narrow then broad and broad then narrow; (3) truth value, true and false; (4) hierarchical level, level 0 and level 1, and (5) verb phrase, "has (noun)," "can (verb)," and "is (adjective)." Results of the analysis are shown in Table 8.

Insert Table 8 about here

Verb phrase was significant, $F(2,20) = 3.60$, $p < .05$. The shortest RT, 1493 msec., was for "can (verb)" property relations. RT for "is (adjective)" was slightly higher, 1503 msec., and RT for "has (noun)" was 1573 msec.

There was a significant interaction among verb phrase, hierarchical level, and truth value, $F(2,20) = 4.27$, $p < .05$. Figure 7 shows that for

true sentences, RT increased for each of the three verb phrase property relation sentences. The largest increase in RT was for "can (verb)" sentences which increased 184 msec. from level 0 to level 1. "Is (adjective)" sentences increased 107 msec., and "has (noun)" sentences increased 80 msec.

Insert Figure 7 about here

For false property relation sentences, RT for "has" and "is" sentences increased 153 and 109 msec., respectively. However, for "can" sentences, RT decreased 125 msec. from level 0 to level 1.

Error rates for sentence conditions. Figure 8 shows the mean proportion of errors for true and false, property and superset relation sentences for levels 0 and 1. For true sentences, increases in errors paralleled increases in RT for correct responses. That is, errors were higher for

---

Insert Figure 8 about here

---

property relation sentences than for superset relation sentences. Errors increased from level 0 to level 1, and the increase in errors was greater for superset relation sentences than for property relation sentences.

For false sentences, errors also increased from level 0 to level 1 for both property and superset relation sentences. However, there was a slight cross-over in the data. At level 0, property relation sentences had a higher error rate, but at level 1, superset relations had the higher error rate.

Figure 9 shows the proportion of errors for the true and false property relation sentences using the three different kinds of predicates. For the true sentences, proportion of errors increased for "can (verb)" and "is (adjective)" sentences, but decreased for "has (noun)" sentences from level 0 to level 1. For the false sentences, proportion of errors increased for the "can (verb)" and "has (noun)" sentences, but was constant for the "is (adjective)" sentences from level 0 to level 1.

Insert Figure 9 about here

Discussion

The results showed a fairly clear replication of the Collins and Quillian (1969) data. The relative position of the property and superset relation curves and the increase in RT with hierarchical level confirmed the effects found previously for true sentences. The effects held over the wider range of sentences used in the present study. Furthermore, the data for individual Ss indicated that the effects were not due to averaging.

The major divergence between the results of the present study and those of the Collins and Quillian (1969) study was in the absolute values of RT. RTs in the present study were generally longer than those reported by Collins and Quillian (1969). The estimates for time to retrieve properties, and time to move between levels, differed. In the present study, it took more time to move between levels, 124 msec., and less time to retrieve properties, 157 msec. The comparable data for Collins and Quillian were 75 msec. and 225 msec.

Results for false sentences were not clear. Apparently, property relation sentences had longer RTs than superset relation sentences. This general result confirmed that reported by Collins and Quillian (1969). However, changes in RT were not consistent across Ss or across hierarchies. Apparently, RTs to false sentences did not vary with differences in the range of the lexicon S searched. RTs to false sentences did not differ when S was searching narrow or broad hierarchies.

Results on variation in verb phrase or property relation sentences indicated that syntactic differences of the kind used here had some effect on RT. "Has (noun)" sentences had the longest RTs. These sentences are superficially similar to superset relation sentences in that both contain a noun in the verb phrase. Since "has (noun)" sentences had the longest RTs, it is not the mere presence of a noun in the verb phrase which makes superset relation sentences easiest to process.

All true property relation sentences showed increases in RT regardless of verb phrase. An unexplained effect was that RT decreased from level 0 to level 1 for false "can (verb)" sentences. These level 1 sentences might possibly have been simpler than the level 0 sentences, for some reason. However, if the proportion of errors is an indication of sentence difficulty, then there was no support for this explanation. There was no decrease in errors corresponding to the decrease in RT for false "can (verb)" sentences.

Experiment Two

The purpose of experiment two was to determine whether the previously obtained difference in RT between property and superset sentences was due to the syntactic differences between the sentences, or the different ideas the two sentence types express. These two hypotheses, the syntactic and the semantic, were tested by presenting similar ideas in two syntactic forms, e.g., "A teacher can educate" and "A teacher is an educator." The syntactic hypothesis predicts a difference in the RT to these two sentences, and the semantic hypothesis predicts no difference in RT to these two sentences.

Method

Experimental sentences. Twelve two-level hierarchies were constructed. The semantic content of all hierarchies concerned roles or occupations, e.g., writers, businessmen, athletes. Figure 10 shows a two-level hierarchy on writers and Table 9 shows the sixteen sentences which were constructed on the basis of this hierarchy.

Insert Figure 10 about here

Insert Table 9 about here

Sentences one through eight were true sentences and sentences nine through sixteen were false. The P or S before each sentence indicates the type of relation, property or superset. Superset relation sentences were

of the form, "A (noun) is a (noun)." Property sentences were of the form, "A (noun) is (adjective)," or "A (noun) can (verb)." Property sentences of the form "A (noun) has (noun)," which were used in the pilot study, were eliminated because they had the highest RTs. The true sentences were constructed with regard to hierarchical level and are marked level 0, level 0', and level 1. (The meaning of the levels is explained later.) The false sentences were constructed without regard to level of contradiction.

Besides the hierarchy shown in Figure 10, eleven other hierarchies were constructed with sixteen sentences per hierarchy, for a total of 192 sentences. Half of the sentences made up form A (left column of Table 9) and the other 96 sentences constituted form B (right column of Table 9).

Each row of Table 9 shows the sentences as semantically corresponding sentence sets. Set 1 consisted of the property sentences of form A, and their semantically corresponding sentences which were superset sentences in form B. Set 2 consisted of those sentences which were superset sentences in form A, and their semantic correspondents, the property sentences of form B.

In the experiment, Ss saw all of form A and then all of form B, or vice versa. The sentences in Table 9 are shown grouped by forms and by sets. However, the comparison of property and superset sentences within a set was more critical in this experiment than was the comparison of property and superset sentences within a form.

While the hierarchy in Figure 10 shows just two levels, the true sentences in Table 9 are marked for three possible levels, 0, 0', and 1.

This classification of levels was sometimes the same and sometimes different from the classification used by Collins and Quillian (1969).

1. The superset level 0 sentences, S0, were similar to the superset level 0 sentences used by Collins and Quillian (1969).

2. The superset and property level 1 sentences, S1 and P1, were similar to those used by Collins and Quillian (1969).

3. The property level 0 sentences, P0, were different. In the Collins and Quillian (1969) study, a P0 sentence contained information about properties specific to that word. In the present experiment, a P0 classification resulted from trying to express the same information presented in an S0 sentence, in a property relation sentence form. Therefore, sentence (1) in Table 9 was constructed because there were sentences like sentence (2).

4. The property level 0' sentences, P0', as in sentence (6), corresponded to the Collins and Quillian (1969) property level 0 sentence. That is, sentence (6) in this experiment contained property information that was specific to playwrights, i.e., that they "can dramatize."

5. The superset level 0' sentences, S0', as in sentence (5), were used because they were the semantic correspondents to the P0' sentences.

The main difference between the presented sentences and those used by Collins and Quillian, was in the classification of level 0 sentences. In the present experiment, there were two types of level 0 sentences.

In the sentence hierarchy shown in Table 9, there are P1 and S1 sentences in both forms A and B. However, form A has a P0 sentence but no P0' sentence, and an S0' sentence but no S0. There were six hierarchies

in form A like this. The other six hierarchies in form A had a PO' sentence but no PO, and had an SO sentence but no SO'. Each hierarchy had either a level 0 or a level 0' sentence. Table 10 summarizes the total number of each type of sentence in each form.

Insert Table 10 about here

Practice sentences. An additional group of 32 sentences, representing two hierarchies, were constructed for use as practice sentences to familiarize Ss with the task.

Standardization of sentences. To determine the validity of E's judgment on the truth and falsity of the sentences, lists of sentences were presented to groups of Ss to be scored true or false. In the pilot study, the error rate per S was 7%, indicating some disagreement with E's judgment. Collins and Quillian (1969) reported an error rate of 8%. Therefore, sentences were selected until an acceptable error rate of about 7% was reached.

In the first standardization group, Ss were presented with a booklet containing 276 sentences to be marked true or false. The sentence types were randomly arranged in the booklet. The first group of Ss were ten high school students who were paid $1.25 for participating.

On the top sheet of the booklet was an instruction sheet which read as follows:

> Read each sentence and decide if you think it is a generally true or generally false statement. If true, circle the T after the sentence. If false, circle the F.
>
> For example, if the sentence reads:
> 'A warrior is aggressive' T F,
> you would circle the T. If the statement reads:
> A warrior is gentle T F,
> you would circle the F.
>
> If you read a sentence and can't make a decision fairly quickly, leave it blank and go on to the next sentence.
>
> If you can't make a decision because you don't understand one of the words in the sentence, circle the word you don't understand, and leave it blank.

Ss worked at their own pace and a session lasted from 20 to 60 minutes.

The mean proportion of errors in the first group was .21 per S, with scores ranging from a low of .04 to a high of .57.

Since this error rate was considered too high, the sentences were modified, deleted or new ones were added and a new booklet of 264 sentences randomly arranged was made. These sentences were presented to a group of three high school students who were paid $1.25 for participating. Instructions and other conditions were the same as for the previous group.

The mean proportion of errors for the second group was .14 per S, with scores ranging form a low of .12 to a high of .17. This error rate was still considered too high, so further changes were made in the sentences.

A modified set of 246 sentences was presented to a group of seven Ss. One S was a high school student, paid $1.25 for participating, and six Ss were psychology graduate students who volunteered to participate.

Other conditions and instructions were the same as for the two previous groups.

Mean proportion errors on the 246 sentences was .03 per S, ranging from .01 to .05. From these 246 sentences, a set of 224 sentences were selected to be used in the experiment. The mean proportion of errors was .03 per S for the final set of sentences. Proportion of errors on the 32 sentences selected for the practice trials was .035 per S, and the proportion errors on the 192 experimental sentences was .025. Most of the errors were due to one S making an error on a single sentence, but in three cases two Ss made an error on the same sentence.

Table B (see Appendix) shows the final set of practice and experimental sentences.

Subjects. Forty-one undergraduate students at the University of Massachusetts were Ss. The design of the study required 24 Ss. However, the additional 17 Ss were run because equipment problems resulted in the loss of data from some Ss. For other Ss, data was lost because of equipment problems coupled with high error rates. Ss were paid $1.25 to participate in a session which lasted approximately 45 minutes.

Apparatus. Preparation of the sentences on slides, the Ss' response panel, and programming of the equipment were the same as described in the pilot study.

Procedure. In most respects, the procedure was the same as in the pilot study. Sentences were presented in blocks of 32, beginning with a practice block of 32 sentences. There was a one-minute break between trial blocks, which was signaled by the words "BREAK" and "READY" on the Nixie

tubes below the screen. Each sentence appeared for two seconds, followed by a blank screen for two seconds. Ss were required to respond within this four second period. There was a one second delay before the presentation of the next sentence.

Twelve of the Ss were shown the sentences constituting form A first, followed by form B, and the other twelve Ss were shown form B first, followed by form A.

A block of 32 trials consisted of the random presentation of sentences from four of the twelve hierarchies within a particular form. Various randomization and balancing procedures were used to insure that the twelve hierarchies were not always presented in the same sequence.

Instructions. Ss were read the following instructions.

> You'll see a series of sentences one at a time. You should decide if each sentence is generally true or generally false. Once you've decided 'true' or 'false,' press the corresponding button.
>
> Each sentence will be on for two seconds, then a blank screen for two seconds. You can answer as soon as you decide 'true' or 'false,' but it has to be within the four second period. At the end of the four seconds, another sentence will come on and you'll have to respond to that one. So, if you haven't responded by the time the next sentence comes on, forget about it, and only respond to the sentence that is being presented.
>
> At the end of 32 sentences, there'll be a one-minute break, and you can relax for a while.
>
> The first set of sentences will be for practice so that you can get used to the timing of the slides and learn which button to push for 'true' and which button to push for 'false.'
>
> Answer as quickly and accurately as you can in deciding if the statement is generally true or false.
>
> Any questions?

E returned after the practice block to see if S had any questions.

Results

RT data. Table 11 shows the results of an analysis of variance done on the mean RT to true and false property and superset sentences. There were eight mean RTs from each of the 24 Ss. These eight mean RTs consisted of the RTs to the four sentence types presented in form A, (true property, true superset, false property, and false superset), and the RTs to the four sentence types again presented in form B. The four variables included in the analysis were: (1) form, A and B; (2) order of forms, AB order and BA order; (3) truth value, true and false sentences; and (4) type of relation, property and superset sentences. (The hierarchical level variable was meaningful only for true sentences and is discussed later.)

Insert Table 11 about here

The syntactic hypothesis predicts that property sentences have higher RTs than superset sentences. RT to property sentences was 1818 msec., and RT for superset sentences was 1807 msec., a difference of only 11 msec., which was not significant, $F(1,22) = 0.53$.

A significant main effect was found for truth value, $F(1,22) = 19.82$, $p < .001$. RT for true sentences was 1737 msec., and RT for false sentences was 1888 msec. As shown in Figure 11, truth value interacted with type of relation, $F(1,22) = 5.04$, $p < .05$. For true sentences, property relations took 38 msec. longer to respond to than superset relations.

This reversed for false sentences, where superset relations took 17 msec. longer than property relations.

Insert Figure 11 about here

A significant interaction between form and type of relation, $\underline{F}(1,22) = 12.31$, $\underline{p} < .005$, is shown in Figure 12. The interaction shown was that in form A, RT for superset sentences was 38 msec. longer than RT for property sentences. When the syntax of the form A sentences was changed, so that the property sentences of form A became the superset sentences of form B, and vice versa, then the relative position of the superset and property RTs reversed. In form B, RT for superset sentences was 59 msec. faster than RT for property sentences.

Insert Figure 12 about here

This same effect is also shown in Figure 13, perhaps more clearly. Figure 13 shows the RTs grouped with respect to set and type of relation rather than with respect to form and type of relation. The sentences

----

Insert Figure 13 about here

----

of set 1, it will be recalled, consisted of property sentences of form A, and their semantic counterparts, which were the superset sentences of form B. Set 2 consisted of the superset sentences of form A, and their semantic counterparts, which were the property sentences in form B.

The mean RT for set 1 was 1788 msec., and 1837 msec. for set 2. This difference was significant, $F(1,22) = 12.31$, $p < .005$.

As can be seen from Figure 13, there was no difference in RT between the property and superset sentences of set 2. For set 1, the RT for property sentences averaged 22 msec. longer than for their superset counterparts. This difference of 22 msec. was not significant using a Tukey test for multiple comparisons.

Figures 12 and 13, then, indicate that syntactic differences between sentence pairs did not have a major effect on RT.

Table 11 also shows that there was a significant form by order of forms interaction, $F(1,22) = 61.89$, $p < .001$. This interaction, shown in Figure 14, was interpreted as a practice effect. In the AB order group, form A was shown first, and apparently this resulted in the longer RT for form A. In the BA order group, form B was shown first, and form B had the longer RT for these Ss.

Insert Figure 14 about here

This practice effect was somewhat different for true and false sentences as indicated by the significant interaction among forms, order of forms, and truth value, $F(1,22) = 4.40$, $p < .05$. Figure 15 shows the interaction among forms, order of forms, and truth value.

Insert Figure 15 about here

To study the effect of hierarchical level, an analysis of variance was done on data for true sentences, since only true sentences were constructed with regard to hierarchical level. The data consisted of twelve RTs from each of the 24 Ss, six RTs from form A and six RTs from form B. The sentence types represented in each form were, T-PO, T-PO', T-Pl, T-SO, T-SO', and T-Sl. The RTs of each S for the T-PO, T-PO', T-SO, and T-SO' sentences were based on six RTs, since there were six hierarchies that contained level 0 sentences, and six hierarchies that contained level 0' sentences. The RTs for each S for the T-Pl and T-Sl sentences were based upon twelve RTs, since each of the twelve hierarchies contained level 1 sentences. The variables in the analysis of variance shown in Table 12 were: (1) form, A and B; (2) order of forms, AB order and BA order; (3) type of relation, property and superset; and (4) level of hierarchy, level 0, and level 0', and level 1.

Insert Table 12 about here

RT increased with hierarchical level. RT at level 0 was 1522 msec. RT increased 264 msec. to 1786 msec. at level 0'. There was a further increase of 38 msec. to 1824 msec. at level 1. The effect for level was significant, $F(2,44) = 61.28$, $p < .001$. Only the RT for level 0 sentences was significantly different from the other two RTs by a Tukey test for multiple comparisons.

The RT for property sentences was 1725 msec., and the RT for superset sentences was 1696 msec. As in the previous analysis of variance, there was not a significant effect for type of relation, $F(1,22) = 3.89$.

Table 12 also shows that there was a significant form by order of forms interaction, $\underline{F}(1,22) = 40.89$, $\underline{p} < .001$. This interaction was similar to that previously shown in Figure 13, and was interpreted as a practice effect.

Figure 16 shows the RT for property and superset relation sentences in forms A and B. In form A, superset sentences had an 11 msec. higher RT than property sentences. In form B, property sentences had a 69 msec. higher RT than superset sentences. This interaction was significant, $\underline{F}(1,22) = 7.03$, $\underline{p} < .025$. This type of interaction, along with the absence of a main effect for type of relation, was consistent with the semantic hypothesis, when set 1 sentences are easier than set 2 sentences. Since set 1 had a significantly faster RT than set 2 sentences, $\underline{F}(1,22) = 7.03$, $\underline{p} < .025$, the data again seemed to support the semantic hypothesis.

Insert Figure 16 about here

That is, the form by type of relation interaction shown in Figure 16 seemed to be due to the fact that the superset sentences in form B were responded to faster than the superset sentences in form A. Figure 17 shows the RTs to true property and superset sentences grouped according to set rather than form. There was no significant interaction between sets and type of relation, $\underline{F}(1,22) = 0.28$.

Insert Figure 17 about here

So far the data presented have tended to support the semantic hypothesis. There were two other significant interactions, indicated in Table 12, which complicated the interpretation of the results somewhat.

Figure 18 shows the RTs for property and superset sentences as a function of hierarchical level. At levels 0 and 1 property sentences had higher RTs than superset sentences. At level 0' superset sentences had higher RTs than property sentences. This interaction was significant, $\underline{F}(2,44) = 8.94$, $\underline{p} < .001$. Using a Tukey multiple comparison test, it was found that there was a significant increase in RT from level 0 to level 0', and from level 0' to level 1 for property sentences. For superset sentences, RT significantly increased from level 0 to level 0', but not from level 0' to level 1. In fact, there was a slight, 17 msec., decrease in RT from level 0' to level 1 for superset sentences.

Insert Figure 18 about here

Figure 19 shows the RTs for property and superset sentences at each hierarchical level in forms A and B. From inspection of Figure 19, it

---

Insert Figure 19 about here

---

seemed that the type of relation by level interaction changed from form A to form B. This interaction was significant, $\underline{F}(2,44) = 4.15$, $\underline{p} < .025$. The change in interaction with forms seemed to be due to what occurred at

level 1. In form A at level 1, the superset sentences had the higher RT, whereas in form B at level 1, the property sentences had the higher RT. At levels 0 and 0', the relative positions of the superset and property RTs were the same in both forms A and B. It was apparently this reversal at level 1 of the superset and property sentences which was the basis of the observed interaction between forms and type of relation previously discussed (see Figure 16).

While this interaction seemed to be very complex, Figure 20 shows how it was made to disappear. Figure 20 shows the RTs to property and superset relations at each hierarchical level, grouping the data on the basis of set rather than on the basis of form. The set by type of relation by hierarchical level interaction was not significant, $\underline{F}(2,44) = 0.97$. That is, the type of relation by hierarchical level interaction shown in Figure 18 was the same for sentence sets 1 and 2. Since grouping the sentences by sets was a more pertinent comparison than grouping by forms, the interaction depicted in Figure 19 was ignored.

Insert Figure 20 about here

To summarize the important effects thus far, it seemed that when RTs to property and superset sentences were contrasted with the RTs to their semantic counterparts, the same type of interaction emerged. At levels 0 and 1, property sentences had higher RTs than their superset counterparts. At level 0', however, supersets had higher RTs than their property counterparts. This can be seen in Figures 18 and 20.

The level 0 and level 1 data indicated support for the syntactic hypothesis, since property sentences had the higher RTs at these levels. However, the high RT for level 0' superset sentences, indicated that some types of information were more difficult to handle when presented in the syntactic form of a superset sentence, than when presented in the syntactic form of a property sentence.

The overall results of the first analysis of variance discussed, (see Table 11), indicated general support of the semantic hypothesis, i.e., that the content of the sentence had a greater effect on RT than the syntax. However, the results of the second analysis of variance, which included the hierarchical level variable, indicated support for a compromise position. That is, the syntax of a sentence influences RT, but not in the simple straightforward way predicted by the syntactic hypothesis.

Several additional analyses were done to try and get more information about the basis of the type of relation by hierarchical level interaction.

Since RTs were generally longer in this experiment than in the pilot study, an analysis of variance was done to determine if the Ss' speed of responding interacted with any of the other variables. Using overall median RT as a dividing point, Ss were split into two groups, fast and slow responders. Table 13 shows the results of an analysis of variance on RT data which included the following variables: (1) speed of responding, fast and slow; (2) form, A and B; (3) type of relation, property and superset; and (4) truth value, true and false.

Insert Table 13 about here

Figure 21 shows the mean RTs for fast and slow responders to true and false, property and superset relation sentences in forms A and B. The interaction was significant, $\underline{F}(1,22) = 5.39$, $\underline{p} < .05$.

Insert Figure 21 about here

Table 14 shows the results of an analysis of variance on RT for true sentences only. The variables in the analysis were: (1) speed of responding, fast and slow; (2) type of relation, property and superset; (3) form, A and B; and (4) hierarchical level, level 0, level 0', and level 1.

Insert Table 14 about here

The data represented in Table 14 show that speed of responding did not interact significantly with the other variables in a way that seemed related to the type of relation by hierarchical level interaction shown in Figure 18.

Generally, the results of the last two analyses reported indicated that the trends reported thus far held for fast and slow responders.

In the next two analyses of variance, the between-hierarchy variability in the RT data was investigated. Each datum was the standard deviation of a set of RT scores. These were the same scores that were used for calculating the mean RTs of each S in the previous analyses. It will be recalled that the RTs on which the mean RT was calculated, were the RTs

to sentences which were all of the same sentence type, but from different hierarchies. RT_{SD} was the variability in this set of RTs.

Table 15 shows the results of an analysis of variance done on RT_{SD} for true and false sentences. The variables were: (1) form, A and B; (2) order of forms, AB order and BA order; (3) type of relation, property and superset; and (4) truth value, true and false.

Insert Table 15 about here

The mean RT_{SD} was 403 msec. Comparison of the data in Table 15 with that in Table 11, shows that increases in mean RT were not always accompanied by corresponding changes in RT_{SD}.

Of the two significant interactions, the one of interest was the form by type of relation interaction, depicted in Figure 22. Comparison

---

Insert Figure 22 about here

---

of Figure 22 to Figure 12 shows that as the mean RTs increased, so did the variability of scores on which the means were calculated, $\underline{F}(1,22) = 15.35$, $\underline{p} < .001$. This interaction again was interpreted as a main effect for sentence set. Set 1 (property sentences in A and superset sentences in B) showed less variability in RTs than did set 2 (superset sentences in A and property sentences in B.)

Table 16 shows the results of an analysis of variance on RT_{SD} for true sentences only, where the variables included were: (1) form, A and B;

(2) order of forms, AB order and BA order; (3) type of relation, property and superset; and (4) hierarchical level, level 0, level 0', and level 1.

Insert Table 16 about here

RT_{SD} increased from 302 msec. at level 0, to 343 msec. at level 0', to 389 mesc. at level 1. This hierarchical level effect was significant, $F(2,44) = 13.83$, $p < .001$. That is, there was more variability in the RT between hierarchies as hierarchical level increased. There seemed to be three possible explanations for this increase in variability with hierarchical level. The increase in variability may have been due to the information stored in memory. That is, the further away one gets from level 0 information, the less likely it is that Ss' memory is organized like E's (the person who made up the hierarchies). Secondly, the increase in variability may have been due to increased variability in Ss' search strategies as distance from level 0 information increases. Thirdly, differences between E and Ss in actual information available might necessitate the use of different search strategies.

This analysis was done partly to get more information on the interaction between hierarchical level and type of relation (see Table 12 and Figure 18). While the trend in the RT_{SD} indicated that variability increased with increases in RT, as shown in Figure 23 the interaction was not significant, $F(2,44) = 2.51$.

Insert Figure 23 about here

A series of experiments by Kintsch (1972) suggested that lexical complexity provides a partial explanation for some of the RT differences observed between property and superset sentences. Lexical complexity refers to whether a word may be decomposed into simpler elements. For example, "soul" is lexically simpler than "ability." A summary of Kintsch's experiments is first presented, and then the present data are considered in terms of lexical complexity.

Kintsch (1972) found that lexical complexity affected paired-associate learning. In learning lists of nouns and digits, Ss made fewer errors when the nouns were lexically simple. Kintsch further investigated two possible origins of the lexical complexity effect. One explanation of the results was that lexically derived nouns, e.g., "ability," were difficult because Ss had to reduce them to simpler elements. This reduction process might have interfered with paired-associate learning. A second explanation, which Kintsch tested, was that Ss reduced the lexically derived nouns to simpler elements which may be verbs or adjectives. Ss then learned the paired-associate lists, but treated the lexically derived nouns as if they were simpler verbs or adjectives. The reason for the greater difficulty of the lexically complex paired-associate lists in contrast to the lexically simple noun lists, was that paired-associate learning was more difficult with verbs and adjectives than with nouns.

To test these explanations, Kintsch (1972) presented paired-associate lists with either simple nouns or with verbs and adjectives from which the lexically complex nouns derive. Ss made more errors learning the verb and

adjective lists, than in learning the simple noun lists, a result supporting the second hypothesis.

Although lexical complexity was not systematically controlled in the present experiment, the sentence pairs were examined to determine whether there were any systematic differences in the lexical complexity of the predicate nouns in the superset sentences versus the predicate verbs and adjectives in the property sentences.

The 48 sentence pairs used in the experiment were classified as to whether the predicate noun, verb or adjective was lexically simple or derived. With respect to superset and property sentences, four categories were differentiated.

1. It was found that in 27 of the 48 sentence pairs, the property sentences had the lexically simpler form of the word in the predicate. For example, "A mechanic can repair" was considered lexically simpler than "A mechanic is a repairman." According to Kintsch's lexical complexity hypothesis, these 27 property sentences should have had faster RTs than their superset counterparts.

2. In 9 of the 48 sentence pairs, the superset sentences had the lexically simpler word in the predicate, as in the set "A poet is a poet," and "A poet is poetic." For these sentence pairs, the lexical complexity hypothesis predicts faster RTs for the superset sentences.

3. In 7 of the 48 sentence pairs, both the property and superset predicate words seemed to be derived from a simpler word. In the set "A playwright can dramatize" and "A playwright is a dramatist," both "dramatize" and "dramatist" would derive from the lexically simpler "drama."

Unless one starts to consider degree of derivation, it seemed that for these sentence pairs, the lexical complexity hypothesis predicts no difference in RT between superset and property sentences.

4. In 5 of the 48 sentence pairs, the same word, whether it was lexically simple or lexically complex, was used in both the property and superset sentences. For example, in the sentence pair "A murderer is criminal" and "A murderer is a criminal," the same lexically derived word was used in both the property and superset sentences.

Table 17 shows the RT differences between the property and superset

----

Insert Table 17 about here

----

sentence pairs which fell into these four categories for level 0, 0', and 1. For example, row one of Table 17 shows the RT to property and superset sentences in those pairs where the property sentences had the lexically simpler form of a word than the superset sentences. At level 0, there were eight sentence pairs which fell into this category, and the RT to the property and superset sentences is shown as well as the RT difference. The number in parentheses below the RT differences gives the number of sentences the difference is based on. The final column of Table 17 shows the overall RTs to property and superset sentences, and the RT differences for each sentence pair category.

The data in the first two rows of Table 17 are most relevant to Kintsch's (1972) lexical complexity hypothesis. While the number of sentence pairs in some cells is only one, and the differences in RT

were small, the direction of the differences in RT between property and superset sentences was as predicted by the lexical complexity hypothesis.

For example, there were 27 sentence pairs where the verb or adjective of the property sentences was lexically simpler than the noun of the superset sentence. For these sentences, the RT for the property sentences was 56 mesc. less than the RT for the superset sentences. On the other hand, when the noun of the superset sentences was lexically simpler than the verb or adjective of the property sentence, superset sentences had a 233 msec. faster RT.

This patterns of results, which is predicted by Kintsch's (1972) lexical complexity hypothesis, held at each hierarchical level.

The pattern of RT differences also suggested that nouns are easier than verbs and adjectives. This conclusion is supported by several comparisons. When the property sentences were lexically simpler, the RT difference was 56 msec. However, when the superset sentences were lexically simpler, the RT difference was even greater, i.e., 233 msec. The data in rows three and four of Table 17 supported the idea that nouns are easier than verbs or adjectives. It can be seen in row three, that when both the superset and property sentences contained derived words, the superset or noun sentences were easier, i.e., a RT difference of 87 msec. Similarly, in row four of Table 17, it can be seen that when both property and superset sentences contained the same word, the sentences were easier when the word was used as a noun than when it was used as a verb or adjective. The RT difference here was 22 msec.

While the pattern of RT differences suggested that sentences with lexically simple words, and sentences with nouns, whether simple or derived, have faster RTs generally, this did not seem to hold for level 0' sentences. Consider the RT differences from rows three and four at level 0', in contrast to the overall RT differences for these rows. At level 0', when both property and superset predicates contained derived words, the property sentences were faster by 83 msec. Similarly, when the property and superset sentences contained the same word, the property sentences were faster by 82 msec. These differences contradict the overall differences for these rows.

Therefore, at level 0', it seemed that property sentences were easier than superset sentences, whereas at levels 0 and 1, superset sentences were easier. The origins of this effect may lie in the way level 0' property sentences and superset sentences were constructed. The level 0' property sentences presented property information pertinent to a specific instance, e.g., "A canary can sing." The level 0' superset sentences were generally constructed by first identifying properties specific to an instance, and then determining which properties could be worded in superset form. The greater ease of level 0' property sentences reflects the fact that they were primary constructions.

Error data. The overall proportion of errors was .05. Figure 24 shows the proportion of errors for true and false sentences. The false

---

Insert Figure 24 about here

---

set 1 sentences showed a higher error rate for property sentences, and the false set 2 sentences showed a higher error rate for superset sentences. The false sentences, then, showed no overall syntactic effect with error rate.

Figure 25 shows the breakdown in the proportion of errors for true sentences at each hierarchical level. Error rates for true sentences indicated that superset sentences were more difficult than property sentences in both sets. The two extreme points in Figure 25 occurred for superset sentences. True superset level 0 sentences had a very low error rate. The interpretation offered earlier was that these sentences do not really represent a sentence comprehension task to the Ss, but rather a pattern matching task. The highest proportion of errors occurred for true SO' sentences. This high error rate for these sentences is consistent with the high RT to these sentences, indicating that these sentences were difficult to process.

Insert Figure 25 about here

Discussion

The results of experiment two did not support the semantic or syntactic hypotheses as originally stated. The semantic hypothesis predicted no difference in RT between semantically corresponding property and superset sentences, while the syntactic hypothesis predicted that superset sentences would have faster RTs than their semantically corresponding property sentences. Neither hypothesis was completely supported. While

there was no overall difference in RT between property and superset sentences, a result supporting the semantic hypothesis, there were RT differences between property and superset sentences at each level of the hierarchy. At levels 0 and 1, superset sentences had faster RTs, and at level 0', property sentences had faster RTs. This effect indicated that syntactic differences influenced RT, but not in the simple way of property sentences always having longer RTs than superset sentences, as predicted by the syntactic hypothesis.

The following factors seemed to have influenced RT to sentences in this study.

1. Hierarchical level. RT generally increased with hierarchical level. This effect was expected as it had been reported by Collins and Quillian (1969), and Meyer (1970), and was replicated in experiment one of this study.

2. Lexical complexity. RT to sentences with lexically complex nouns, verbs, or adjectives was generally higher than RT to sentences with simpler lexical items in the predicate. This result is consistent with Kintsch's (1972) finding that the presence of lexically complex words increased the difficulty of paired-associate list learning.

3. Syntactic form of the sentence and nature of the information presented. The data also indicated that certain kinds of information, (level 0' information), were simpler to process in the property sentence syntactic form, while other types of information (level 0 and level 1 information), were easier to process in superset sentence syntactic form.

4. Syntactic form. There was some indication that superset sentences were somewhat easier than property sentences overall. That is, the presence of a lexically simple noun in the predicate of superset sentences reduced RT more than the presence of a lexically simple verb or adjective. This effect is also consistent with Kintsch's (1972) results. This effect is very tentatively suggested.

Figure 26 shows a memory network which can account for the RT effects obtained in this study, with some notable exceptions. The Figure shows

----

Insert Figure 26 about here

----

a two-level hierarchy about athletes. It is similar to the Collins and Quillian (1969) network in that it starts at an S0 node. The S0 node has a property specific to that noun. In this network the property at level 0 is labeled P0', rather than P0 as in Collins and Quillian's (1969) network. The S0 node is also directly associated with its S1 node, and the S1 node has a higher order property directly associated with its noun.

The network in Figure 26 has four primary nodes: S0, S1, P0', and P1. Each of these nodes is subscripted with an "s," indicating that the primary node is a lexically simple item. The primary nodes are connected by solid lines.

Dangling from each primary node are the lexically complex forms of each of the lexically simple items. For example, $S0'_c$, "is a tumbler," was the lexically complex form of the property, $P0'_s$, "can tumble." The broken line connecting the complex form to the simple form is used to

indicate that the complex form is not a primary node in the network. It is a form derived from the primary node by grammatical operation(s).

In short, the nodes in Figure 26 are an attempt to incorporate the RTs in the first two rows of Table 17 into a single network which can account for the ordering of the twelve RTs. The number in parentheses at each primary or derived node in Figure 26 is a mean RT from Table 17. The relevant cells in Table 17 have been labeled in accord with the network in Figure 26. For example, Figure 26 shows that the RT for PO'_s sentences was 1723 msec. This RT comes from Table 17. It was the RT to level O' property sentences where the property form of the sentence had a lexically simpler predicate than the superset form of the sentence. The RT to its derived counterpart, SO'_c, was 1839 msec. This was the RT to level O' superset sentences, where the predicate of the property sentence was lexically simpler than the predicate of its corresponding superset. That is, both means are from row one of Table 17 from the column headed level O'.

The network in Figure 26 can account for some of the observed RTs in the following sense. RT to a particular sentence increases depending upon the number of steps one must travel from SO_s to the node indicated by the sentence type. For example, take the $P1_s$ and the $S1_c$ sentences. $P1_s$ is two steps removed from the initial SO_s node. The RT to this sentence type was 1751 msec. $S1_c$, however, is two steps plus one set of grammatical operations removed from the SO_s node. Therefore, its RT should be greater than that for the $P1_s$, which it was. $S1_c$ had a mean RT of 1793 msec. as indicated in Figure 26 or Table 17.

Similarly, the Sl_s sentences are one step removed from the SO_s node. The RT to these sentences was 1730 msec. The Pl_c sentences which are removed from the SO_s node by one step plus one set of grammatical operations, had a higher RT, 1870 msec.

RT in this network is dependent upon the number of steps between a sentence type and the SO_s node, and whether or not the sentence is derived by grammatical operations.

There are several inadequacies in this model, however.

The most important problem with the model concerns its inability to explain the RTs to Pl_c and Sl_c sentences. A Pl_c sentence is one step plus one set of grammatical operations away from the SO_s node. Its RT was 1870 msec. An Sl_c sentence is two steps plus one set of grammatical operations removed from the SO_s node. Therefore, an Sl_c should have a higher RT than a Pl_c sentence. It did not. RT to Sl_c was only 1793 msec.

A second problem with the model presented in Figure 26 is that it does not represent all the RTs in the first two rows of Table 17. There should be twelve nodes in the network representing the twelve RTs in the first two rows of Table 17. Figure 26 has incorporated only eight of the sentence types. SO_c, PO_s, SO'_s, and PO'_c sentence types have not yet been accounted for in terms of a network.

An example of a PO_s sentence is "A legislator can legislate." Its RT from Table 17 was 1550 msec. The SO_c would be "A legislator is a legislator," with an RT of 1557 msec. (See row one of Table 17.)

There was only one sentence used in this study which was classified as an SO'_s, "An infant is a baby." The RT to this sentence was 1567 msec.

Its corresponding PO'_c sentence, "An infant is babyish," had an RT of 1889 msec. (See row two of Table 17.)

It was not clear to the writer how the RTs to these sentence types can be simply incorporated into the network. However, some suggestions are offered.

First, the network may be modified to explain the RTs to PO_s and SO_c sentences. RT to PO_s was slightly lower than RT to SO_c, a difference of only 7 msec. However, the direction of the difference was consistent with the idea that SO_c is derived from PO_s by grammatical operations. For these two sentence types, the SO node is not primary. It is suggested that when the PO form of the sentence is lexically simpler than the SO form, the lexicon is entered not at a noun, but at a lexically simple verb or adjective.

Figure 27 illustrates this suggestion. The lexicon is entered at

----

Insert Figure 27 about here

----

the node "to legislate." This node is more closely associated with a sentence of the form "A legislator can legislate," than the sentence "A legislator is a legislator." This network is inadequate in that it does not specify who is doing the legislating. A network of this form, where PO_s is closer to the initial node seems warranted by the finding that PO_s is faster than SO_c.

Furthermore, this kind of structure also seemed indicated by the fact that SO_c sentences had a much higher RT than SO_s sentences. SO_s sentences

had an RT of 1257 msec., in contrast to an RT of 1557 msec. for SO_c. Even though both sentences can be responded to as pattern matching tasks, the SO_c had a much higher RT than SO_s. The network in Figure 27 suggests that SO_c sentences are not treated as pattern matching sentences, but require more semantic processing than SO_s sentences.

Another problem with this explanation is that it assumes Ss process SO_s and SO_c sentences to the extent that Ss can determine whether the words involved are lexically simple or lexically complex. This obviously contradicts the assumption (Collins and Quillian, 1969) that SO sentences are not semantically processed, but are responded to as patterns.

Finally, there are the sentence types SO'_s and PO'_c. These types are represented by only one sentence pair. "An infant is a baby," and "An infant is babyish." It has been suggested that this sentence pair may be disposed of by arguing that it involves neither a property nor a superset relationship, but a synonymous relationship between "infant" and "baby." That is, a network of the type shown in Figure 28 is suggested. "Infant" and "baby" are considered to be both SO_s nodes, and the terms "infantile" and "babyish" are derived from them by grammatical operation. This network can only be suggested since it is based on RTs to two sentences, and the corresponding RTs to the sentences "A baby is an infant" and "A baby is infantile" were not obtained in this experiment.

Insert Figure 28 about here

One basic "however" about the generality of the findings seems to be in order. The semantic content of the sentences used in experiment two

was limited to occupational or social roles. This content was selected because of the ease of expressing the same idea in two syntactic forms. The ease of transforming the human semantic content sentences may reflect the flexibility with which we categorize and recategorize people. The content of the semantic hierarchies used by Collins and Quillian (1969) is not so easily transformed into alternative sentence types. For example, the sentences "A canary is a bird" and "A canary is an animal" cannot both be changed into property sentences. "A canary is animate" seems fairly reasonable, but "A canary is birdlike"(?) does not. Therefore, the results and conclusions reported may not generalize to other semantic content areas.

References

Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychological Reports, 1958, 4, 39-44.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 323-325.

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. Does category size facilitate categorization time? Journal of Verbal Learning and Verbal Behavior, 1970, 9, 432-438. (a)

Collins, A. M., & Quillian, M. R. Facilitating retrieval from semantic memory: The effect of repeating part of an inference. Acta Psychologica, 1970, 33, 304-314. (b)

Collins, A. M., & Quillian, M. R. Experiments on semantic memory and language comprehension. In L. W. Gregg (Ed.), Cognition in learning and memory. New York: John Wiley & Sons, 1972. Pp. 117-137. (a)

Collins, A. M., & Quillian, M. R. How to make a language user. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972. Pp. 309-351. (b)

Conrad, C. Cognitive economy in semantic memory. _Journal of Experimental Psychology_, 1972, _92_, 149-154.

Freedman, J. L., & Loftus, E. F. Retrieval of words from long-term memory. _Journal of Verbal Learning and Verbal Behavior_, 1971, _10_, 107-115.

Greeno, J. On the acquisition of a simple cognitive structure. In E. Tulving & W. Donaldson (Eds.), _Organization of memory_. New York: Academic Press, 1972. Pp. 353-377.

Kintsch, W. Abstract nouns: Imagery versus lexical complexity. _Journal of Verbal Learning and Verbal Behavior_, 1972, _11_, 59-65.

Kintsch, W. Notes on the structure of semantic memory. In E. Tulving & W. Donaldson (Eds.), _Organization of memory_. New York: Academic Press, 1972. Pp. 247-308.

Kintsch, W., Crothers, E. J., & Berman, L. N. The effects of some semantic and syntactic properties of simple sentences upon latency of judgments of semantic acceptability. Technical Report, University of Colorado, November, 1970.

Landauer, T. K., & Freedman, J. L. Information retrieval from long-term memory: Category size and recognition time. _Journal of Verbal Learning and Verbal Behavior_, 1968, _7_, 291-295.

Landauer, T. K., & Meyer, D. E. Category size and semantic-memory retrieval. _Journal of Verbal Learning and Verbal Behavior_, 1972, _11_, 539-549.

Meyer, D. E. On the representation and retrieval of stored semantic information. Cognitive Psychology, 1970, 1, 242-300.

Meyer, D. E., & Ellis, G. B. Parallel processes in word recognition. Paper presented at the meeting of the Psychonomic Society, San Antonio, November, 1970.

Miller, G. Language and communication. New York: McGraw-Hill, 1951.

Quillian, M. R. Word concepts: A theory and simulation of some basic semantic capabilities. Behavioral Science, 1967, 12, 410-430.

Quillian, M. R. Semantic memory. In M. Minsky (Ed.), Semantic information processing. Cambridge, Mass.: MIT Press, 1968. Pp. 227-270.

Quillian, M. R. The teachable language comprehender: A simulation program and theory of language. Communications of the ACM, 1969, 12, 459-476.

Schaeffer, B., & Wallace, R. Semantic similarity and the comparison of word meanings. Journal of Experimental Psychology, 1969, 82, 343-346.

Schaeffer, B., & Wallace, R. The comparison of word meanings. Journal of Experimental Psychology, 1970, 86, 144-152.

Thorndike, E. L., & Lorge, I. The teacher's word book of 30,000 words. New York: Columbia University Press, 1944.

Tulving, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972. Pp. 381-403.

Webster's new world dictionary of the American language. (College ed.) Cleveland: The World Publishing Co., 1960.

Wilkins, A. J. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 382-385.

TABLE 1

Sample Set of True Sentences used by Collins and Quillian (1969)

Hierarchical level	Type of relation	
	Property	Superset
Level 0	P0 An oak has acorns	S0 An oak is an oak
Level 1	P1 An oak has branches	S1 An oak is a tree
Level 2	P2 An oak has roots	S2 An oak is a plant

TABLE 2

False Sentences Constructed Without Reference to Level of Contradiction used by Collins and Quillian (1969)

Type of relation	
Property	Superset
A hemlock has buckeyes	A pine is barley
A poplar has thorns	A juniper is grain
A dogwood is lazy	A willow is grass

TABLE 3

False Sentences Constructed With Reference to Level of Contradiction used by Collins and Quillian (1969)

Hierarchical Level	Type of relation	
	Property	Superset
Level 0	An oak has pine cones	A cedar is an elm
Level 1	An elm has petals	A birch is a flower
Level 2	A maple can breathe	A spruce is an animal

TABLE 4

Sample Set of True and False Sentences from Narrow Hierarchy Form used in Experiment One

Sentence type	Truth value	
	True	False
P0	A salmon is pink	A herring is dangerous
P1	A tuna can swim	A goldfish can crawl
S0	A sardine is a sardine	A swordfish is a flounder
S1	A trout is a fish	A mackeral is an insect

TABLE 5

Sample Set of True and False Sentences from Broad Hierarchy Form used in Experiment One

Sentence type	Truth value	
	True	False
PO	A shark is dangerous	A beetle is pink
P1	A snake can crawl	A butterfly can swim
SO	A flounder is a flounder	A horse is a sardine
S1	An ant is an insect	A bear is a fish

TABLE 6

Analysis of Variance on Reaction Time for Correct Responses of Experiment One

Source	df	MS	F
Order of forms (O)	1	1,140,891.02	0.15
Range of hierarchy (R)	1	35,595.11	0.06
Hierarchy (H)	11	910,406.06	6.47****
Type of relation (Y)	1	14,761,772.64	42.32****
Hierarchical level (L)	1	7,285,595.56	111.04****
Truth value (T)	1	7,010,580.06	46.72****
Subjects (S/O)	10	7,374,958.65	
O x R	1	7,861,014.06	14.69***
O x H	11	270,988.27	1.93*
R x H	11	474,647.48	3.78****
O x Y	1	0.29	<0.01
R x Y	1	43,472.25	0.39
H x Y	11	530,252.30	8.07****
O x L	1	117,306.25	1.79
R x L	1	4,005.84	0.05
H x L	11	130,053.63	1.70
Y x L	1	1,172,347.56	6.61*

TABLE 6 cont'd.

Source	df	MS	F
L x T	1	5,846,925.50	62.84***
R x T	1	36,337.89	0.83
H x T	11	470,772.74	6.02****
S x R/O	10	535,054.03	
S x H/O	110	140,675.51	
S X Y/O	10	346,712.99	
S x L/O	10	65,617.77	
S x T/O	10	150,059.17	
O x R x H	11	104,902.73	0.83
O x R x Y	1	497,142.51	4.45
O x H x Y	11	83,891.59	1.28
R x H x Y	11	232,026.77	3.11***
O x R x L	1	51.96	<0.01
O x H x L	11	107,184.42	1.40
R x H x L	11	216,035.64	2.24**
O x Y x L	1	60,598.03	0.34
R x Y x L	1	11,156.64	0.16
H x Y x L	11	169,843.50	1.76
O x R x T	1	22,337.79	0.51
O x H x T	11	62,409.89	0.80

TABLE 6 cont'd.

Source	df	MS	F
R x H x T	11	231,442.71	2.11*
O x Y x T	1	531,198.03	1.86
R x Y x T	1	13,953.52	0.20
H x Y x T	11	215,147.41	2.71***
O x L x T	1	130,652.13	1.40
R x L x T	1	12,637.51	0.10
H x L x T	11	162,067.20	2.00*
Y x L x T	1	1,136,089.51	8.61
S x R x H/O	110	125,521.34	
S x R x Y/O	10	111,714.19	
S x H x Y/O	110	65,881.51	
S x R x L/O	10	79,286.72	
S x H x L/O	110	76,385.91	
S x Y x L/O	10	177,392.39	
S x R x T/O	10	43,861.62	
S x H x T/O	110	78,265.05	
S x Y x T/O	10	285,058.39	
S x L x T/O	10	93,046.76	
O x R x H x Y	11	67,887.89	0.91
O x R x H x L	11	86,505.38	0.90

TABLE 6 cont'd.

Source	df	MS	F
O x R x Y x L	1	331.54	0.01
O x H x L x Y	11	141,441.32	1.47
R x H x Y x L	11	157,093.25	1.99*
O x R x H x T	11	78,486.49	0.72
O x R x Y x T	1	34,518.54	0.50
O x H x Y x T	11	90,958.13	1.14
R x H x Y x T	11	99,905.64	1.43
O x R x L x T	1	24,232.11	0.20
O x H x L x T	11	79,561.11	0.98
R x H x L x T	11	236,750.82	2.96***
O x Y x L x T	1	21,230.97	0.16
R x Y x L x T	1	143,830.56	3.44
H x Y x L x T	11	287,259.29	2.73***
S x R x H x Y/O	110	74,633.78	
S x R x H x L/O	110	96,442.35	
S x H x Y x L/O	110	96,421.88	
S x R x H x T/O	110	109,523.58	
S x R x Y x T/O	10	68,808.37	
S x H x Y x T/O	110	79,452.26	
S x R x Y x L/O	10	69,505.19	

TABLE 6 cont'd.

Source	df	MS	F
S x R x L x T/O	10	121,637.38	
S x H x L x T/O	110	81,058.08	
S x Y x L x T/O	10	131,904.02	
O x R x H x Y x L	11	88,093.25	1.12
O x R x H x Y x T	11	42,906.97	0.62
O x R x H x L x T	11	77,882.24	0.97
O x R x Y x L x T	1	250.71	0.01
O x H x Y x L x T	11	36,027.78	0.34
R x H x Y x L x T	11	128,356.84	1.73
S x R x H x Y x L/O	110	78,909.98	
S x R x H x Y x T/O	110	69,702.17	
S x R x H x L x T/O	110	79,924.83	
S x R x Y x L x T/O	10	41,852.81	
S x H x Y x L x T/O	110	105,245.38	
O x R x H x Y x L x T	11	110,391.89	1.49
S x R x H x Y x L x T/O	110	74,152.96	

*p<.05.

**p<.025.

***p<.005.

****p<.001.

TABLE 7

Cell Means for Reaction Time to Correct Responses in Msec. from Experiment One

Truth value	Hierarchical level	Order of forms							
		Narrow-Broad				Broad-Narrow			
		Range of hierarchy				Range of hierarchy			
		Narrow		Broad		Narrow		Broad	
		Relation		Relation		Relation		Relation	
		Property	Superset	Property	Superset	Property	Superset	Property	Superset
True sentences	0	1526	1159	1358	1080	1391	1070	1530	1150
	1	1608	1475	1489	1400	1519	1367	1683	1427
False sentences	0	1640	1543	1521	1402	1401	1420	1561	1465
	1	1704	1579	1529	1458	1404	1372	1530	1471

TABLE 8

Analysis of Variance on Reaction Time for Correct Responses to Property Relation Sentences from Experiment One

Source	df	MS	F
Order of forms (O)	1	130,452.79	0.13
Range of hierarchy (R)	1	1,537.67	0.01
Truth value (T)	1	51,908.84	1.08
Level (L)	1	335,130.07	7.80**
Verb phrase (V)	2	183,297.70	3.60*
Subjects (S/O)	10	1,019,811.95	
O x R	1	1,430,139.32	13.58***
O x T	1	473,321.53	9.83**
R x T	1	1,458.59	0.12
O x L	1	113.34	<0.01
R x L	1	357.85	0.05
T x L	1	222,426.16	15.56***
O x V	2	20,242.87	0.40
R x V	2	75,859.91	3.05
T x V	2	42,893.11	1.27
L x V	2	36,398.46	1.30
S x R/O	10	105,301.25	

TABLE 8 cont'd.

Source	df	MS	F
S x T/O	10	48,129.19	
S x L/O	10	42,949.17	
S x V/O	20	50,895.70	
O x R x T	1	312.19	0.03
O x R x L	1	229.35	0.03
O x T x L	1	44,195.38	3.09
R x T x L	1	21,753.68	1.12
O x R x V	2	5,417.57	0.22
O x T x V	2	17,589.09	0.52
R x T x V	2	3,946.98	0.10
O x L x V	2	40,892.58	1.46
R x L x V	2	23,149.18	0.83
T x L x V	2	178,436.30	4.27*
S x R x T/O	10	11,963.54	
S x R x L/O	10	7,212.76	
S x T x L/O	10	14,294.26	
S x R x V/O	20	24,886.65	
S x T x V/O	20	33,729.05	
S x L x V/O	20	27,916.43	
O x R x T x L	1	3,168.81	0.16

TABLE 8 cont'd.

Source	*df*	*MS*	*F*
O x R x T x V	2	3,152.73	0.08
O x R x L x V	2	12,930.38	0.46
O x T x L x V	2	5,600.09	0.13
R x T x L x V	2	36,885.72	0.80
S x R x T x L/O	10	19,365.80	
S x R x T x V/O	20	38,773.21	
S x R x L x V/O	20	27,982.35	
S x T x L x V/O	20	41,830.90	
O x R x T x L x V	2	29,227.62	0.63
S x R x T x L x V/O	20	46,321.19	

*p < .05.

**p < .025.

***p < .005.

TABLE 9

Sample Set of True and False Sentences Corresponding to Hierarchy in Figure 10

Truth value	Form: A	Form: B
True sentences	(1) P0 A poet is poetic	(2) S0 A poet is a poet
	(3) P1 A journalist can write	(4) S1 A journalist is a writer
	(5) S0' A playwright is a dramatist	(6) P0' A playwright can dramatize
	(7) S1 An autobiographer is an author	(8) P1 An autobiographer can author
False sentences	(9) P A reporter is cowardly	(10) S A reporter is a coward
	(11) P A columnist is hallucinatory	(12) S A columnist is a hallucinator
	(13) S A correspondent is an illiterate	(14) P A correspondent is illiterate
	(15) S An essayist is a fool	(16) P An essayist is foolish

Note.--Sentences 1, 2, 3, 4, 9, 10, 11, and 12 constitute set 1. Sentences 5, 6, 7, 8, 13, 14, 15, and 16 constitute set 2.

TABLE 10

Number of Each Type of Sentence in Forms A and B

Form			
A		B	
Number of sentences	Type of sentence	Number of sentences	Type of sentence
6	T-PO	6	T-PO
6	T-PO'	6	T-PO'
12	T-P1	12	T-P1
6	T-SO	6	T-SO
6	T-SO'	6	T-SO'
12	T-S1	12	T-S1
24	F-P	24	F-P
24	F-S	24	F-S

TABLE 11

Analysis of Variance on Mean RTs to True and False, Property and Superset Sentences in Experiment Two

Source	df	MS	F
Order of forms (O)	1	2,588,138.35	1.83
Form (F)	1	6,709.74	0.24
Truth Value (T)	1	1,089,232.65	19.82***
Type of Relation (R)	1	5,518.87	0.53
Subjects (S/O)	22	1,414,190.34	
O x F	1	1,759,636.11	61.89***
O x T	1	2.67	<0.01
F x T	1	469.50	0.07
O x R	1	414.48	0.04
F x R	1	114,174.47	12.31**
T x R	1	35,903.99	5.04*
S x F/O	22	28,430.21	
S x T/O	22	54,947.60	
S x R/O	22	10,441.63	
O x F x T	1	30,730.39	4.40*
O x F x R	1	1,828.79	0.20
O x T x R	1	2,574.89	0.36

TABLE 11 cont'd.

Source	*df*	*MS*	*F*
F x T x R	1	1,384.28	0.35
S x F x T/O	22	6,990.47	
S x F x R/O	22	9,275.96	
S x T x R/O	22	7,129.74	
O x F x T x R	1	4,889.81	1.24
S x F x T x R/O	22	3,948.00	

*p < .05.

**p < .005.

***p < .001.

TABLE 12

Analysis of Variance on Mean RTs to True Sentences at

Levels 0, 0', and 1 in Experiment Two

Source	df	MS	F
Order of forms (O)	1	3,760,193.77	1.78
Form (F)	1	14,390.14	0.28
Type of relation (R)	1	62,990.31	3.89*
Hierarchical level (L)	2	2,605,774.00	61.28***
Subjects (S/O)	22	2,115,298.12	
O x F	1	2,090,489.36	40.89***
O x R	1	693.01	0.04
F x R	1	116,109.39	7.03**
O x L	2	78,013.98	1.83
F x L	2	22,410.98	0.97
R x L	2	159,410.86	8.94***
S x F/O	22	51,121.73	
S x R/O	22	16,213.73	
S x L/O	44	42,524.51	
O x F x R	1	34,519.00	2.09
O x F x L	2	61,866.15	2.69
O x R x L	2	18,023.97	1.01

TABLE 12 cont'd.

Source	df	MS	F
F x R x L	2	100,251.33	4.15**
S x F x R/O	22	16,510.14	
S x F x L/O	44	23,008.75	
S x R x L/O	44	17,829.70	
O x F x R x L	2	3,243.67	0.13
S x F x R x L/O	44	24,166.67	

*p < .10.

**p < .025.

***p < .001.

TABLE 13

Analysis of Variance on Mean RTs to True and False, Property and Superset Sentences for Fast and Slow Responders in Experiment Two

Source	df	MS	F
Median split groups (M)	1	21,930,048.67	40.99***
Form (F)	1	6,709.74	0.66
Truth value (T)	1	1,089,232.65	20.23***
Type of relation (R)	1	5,518.87	0.53
Subjects (S/M)	22	535,012.60	
M x F	1	38,425.74	0.36
M x T	1	24,584.32	0.46
F x T	1	469.50	0.06
M x R	1	996.09	0.10
F x R	1	114,174.47	12.35**
T x R	1	35,903.99	5.57*
S x F/M	22	106,667.05	
S x T/M	22	53,830.25	
S x R/M	22	10,415.19	
M x F x T	1	15,289.96	1.99
M x F x R	1	2,443.06	0.26
M x T x R	1	17,728.14	2.75

TABLE 13 cont'd.

Source	df	MS	F
F x T x R	1	1,384.28	0.41
S x F x T/M	22	7,692.31	
S x F x R/M	22	9,248.02	
S x T x R/M	22	6,440.95	
M x F x T x R	1	18,064.50	5.39*
S x F x T x R/M	22	3,349.15	

*p < .05.

**p < .005.

***p < .001.

TABLE 14

Analysis of Variance on Mean RTs to True Sentences at Levels 0, 0',

and 1 for Fast and Slow Responders in Experiment Two

Source	df	MS	F
Median split groups (M)	1	30,757,572.87	34.63***
Form (F)	1	14,390.14	0.10
Type of relation (R)	1	62,990.31	4.20*
Hierarchical level (L)	2	2,605,774.80	59.68***
Subjects (S/M)	22	888,144.53	
M x F	1	17,950.22	0.12
M x R	1	27,184.30	1.81
F x R	1	116,109.19	6.80**
M x L	2	53,015.19	1.21
F x L	2	22,410.87	0.87
R x L	2	159,410.87	8.63***
S x F/M	22	145,328.05	
S x R/M	22	15,009.09	
S x L/M	44	43,660.82	
M x F x R	1	22,132.87	1.30
M x F x L	2	3,143.01	0.12
M x R x L	2	4,018.94	0.22

TABLE 14 cont'd.

Source	df	MS	F
F x R x L	2	100,251.33	4.65**
S x F x R/M	22	17,073.14	
S x F x L/M	44	25,677.98	
S x R x L/M	44	18,466.29	
M x F x R x L	2	61,005.09	2.83
S x F x R x L/M	44	21,541.13	

*$p < .10$.

**$p < .025$.

***$p < .001$.

TABLE 15

Analysis of Variance on Standard Deviation of RTs for True and False, Property and Superset Sentences in Experiment Two

Source	df	MS	F
Order of forms (O)	1	54,529.63	0.53
Form (F)	1	21,692.64	2.40
Truth value (T)	1	12,626.13	1.21
Type of relation (R)	1	3,469.08	0.58
Subjects (S/O)	22	102,478.77	
O x F	1	3,203.21	0.35
O x T	1	4,284.56	0.41
F x T	1	1,145.09	0.11
O x R	1	3,172.73	0.53
F x R	1	70,489.26	15.35**
T x R	1	83.67	0.02
S x F/O	22	9,051.42	
S x T/O	22	10,454.95	
S x R/O	22	5,963.95	
O x F x T	1	73.64	0.01
O x F x R	1	8,419.71	1.83
O x T x R	1	8,352.57	2.11

TABLE 15 cont'd.

Source	*df*	*MS*	*F*
F x T x R	1	4,125.43	1.15
S x F x T/O	22	10,261.90	
S x F x R/O	22	4,593.46	
S x T x R/O	22	3,965.19	
O x F x T x R	1	24,388.96	6.83*
S x F x T x R/O	22	3,572.56	

*p < .025.

**p < .001.

TABLE 16

Analysis of Variance on Standard Deviation of RTs for True Sentences at Levels 0, 0', and 1 in Experiment Two

Source	df	MS	F
Order of forms (O)	1	183,163.28	1.11
Form (F)	1	14,890.92	0.48
Type of relation (R)	1	18,560.38	1.12
Hierarchical level (L)	2	182,061.91	13.83**
Subjects (S/O)	22	165,482.41	
O x F	1	5,616.32	0.18
O x R	1	4,319.15	0.26
F x R	1	203,310.98	21.12**
O x L	2	12,977.38	0.99
F x L	2	2,532.80	0.13
R x L	2	41,230.49	2.51*
S x F/O	22	31,342.41	
S x R/O	22	16,623.39	
S x L/O	44	13,163.71	
O x F x R	1	124,310.54	12.92**
O x F x L	2	3,040.07	0.15
O x R x L	2	39,375.31	2.39

TABLE 16 cont'd.

Source	df	MS	F
F x R x L	2	2,818.42	0.19
S x F x R/O	22	9,625.00	
S x F x L/O	44	19,725.67	
S x R x L/O	44	16,456.76	
O x F x R x L	2	2,048.62	0.14
S x F x R x L/O	44	15,154.48	

*p < .10.

**p < .001.

TABLE 17

Mean RTs to Property and Superset Sentence Pairs at Levels 0, 0', and 1. Sentence Pairs are Classified into Four Categories Depending on Lexical Complexity of Predicate Words

Lexical complexity classification relationship	Level 0			Level 0'			Level 1			Overall		
	P	S	RT diff. (N)	P	S	RT diff. (N)	P	S	RT diff. (N)	P	S	RT diff. (N)
Property lexically simpler than superset predicate	1550 PO_s	1557 SO_c	- 7 (8)	1723 PO'_s	1839 SO'_c	-116 (9)	1751 $P1_s$	1793 $S1_c$	- 42 (10)	1682	1738	- 56 (27)
Superset lexically simpler than property predicate	1585 PO_c	1257 SO_s	328 (3)	1889 PO'_c	1567 SO'_s	322 (1)	1870 $P1_c$	1730 $S1_s$	140 (5)	1777	1544	233 (9)
Both superset and property predicates derived	1588	1439	149 (1)	1784	1867	- 83 (1)	1932	1823	109 (5)	1861	1771	87 (7)
Superset and property predicates are the same	----	----		1815	1897	- 82 (1)	1896	1848 (4)	48	1880	1858	22 (5)

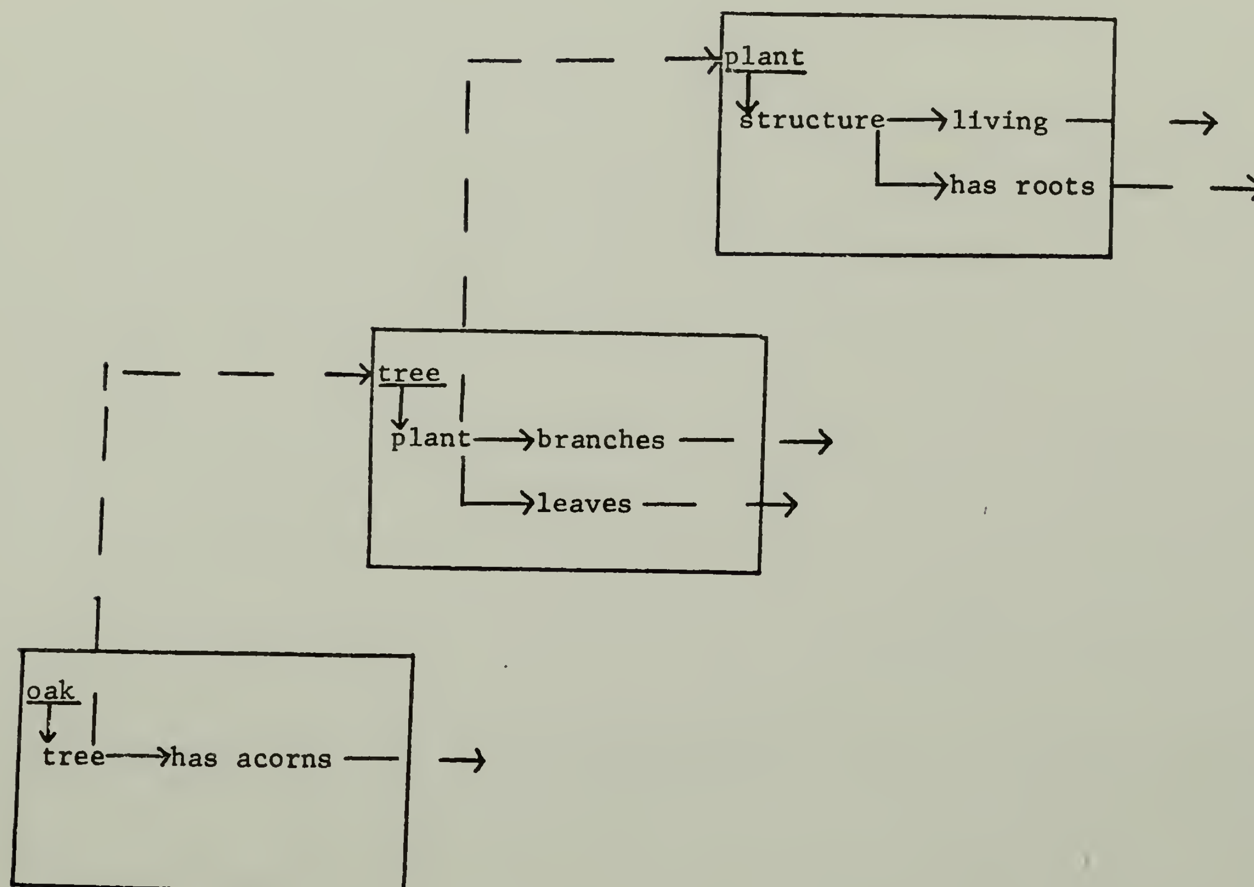

FIG. 1. Memory net model presented in Quillian (1967).

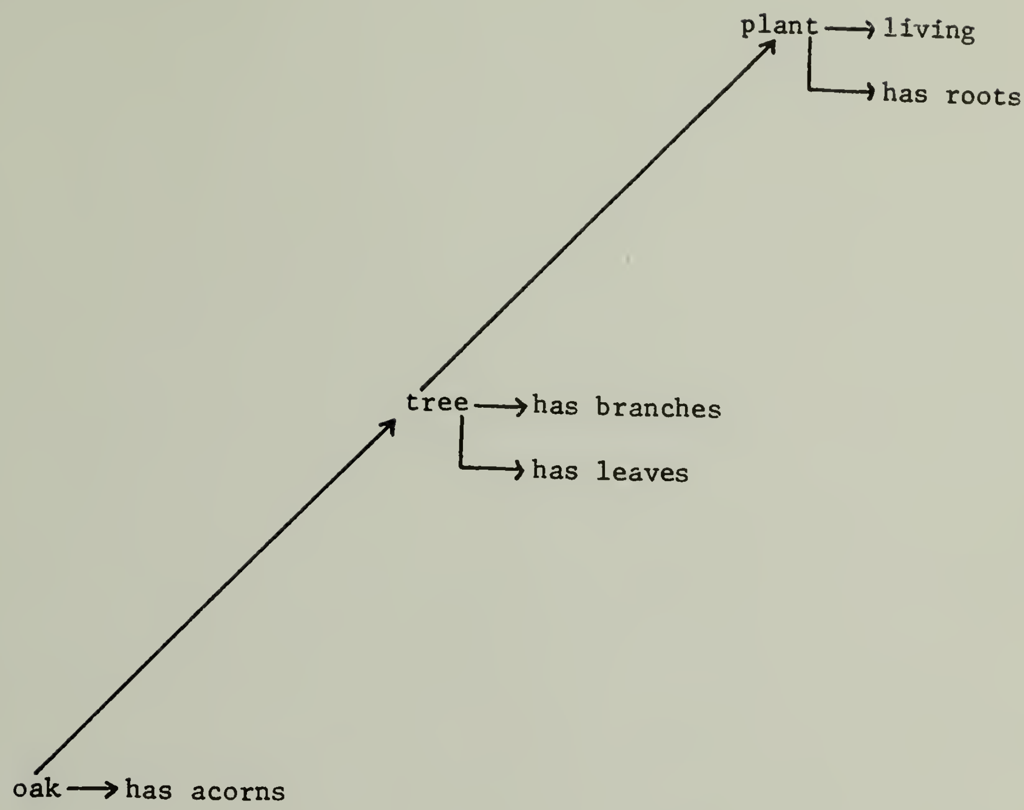

FIG. 2. Memory net model presented in Collins and Quillian (1969).

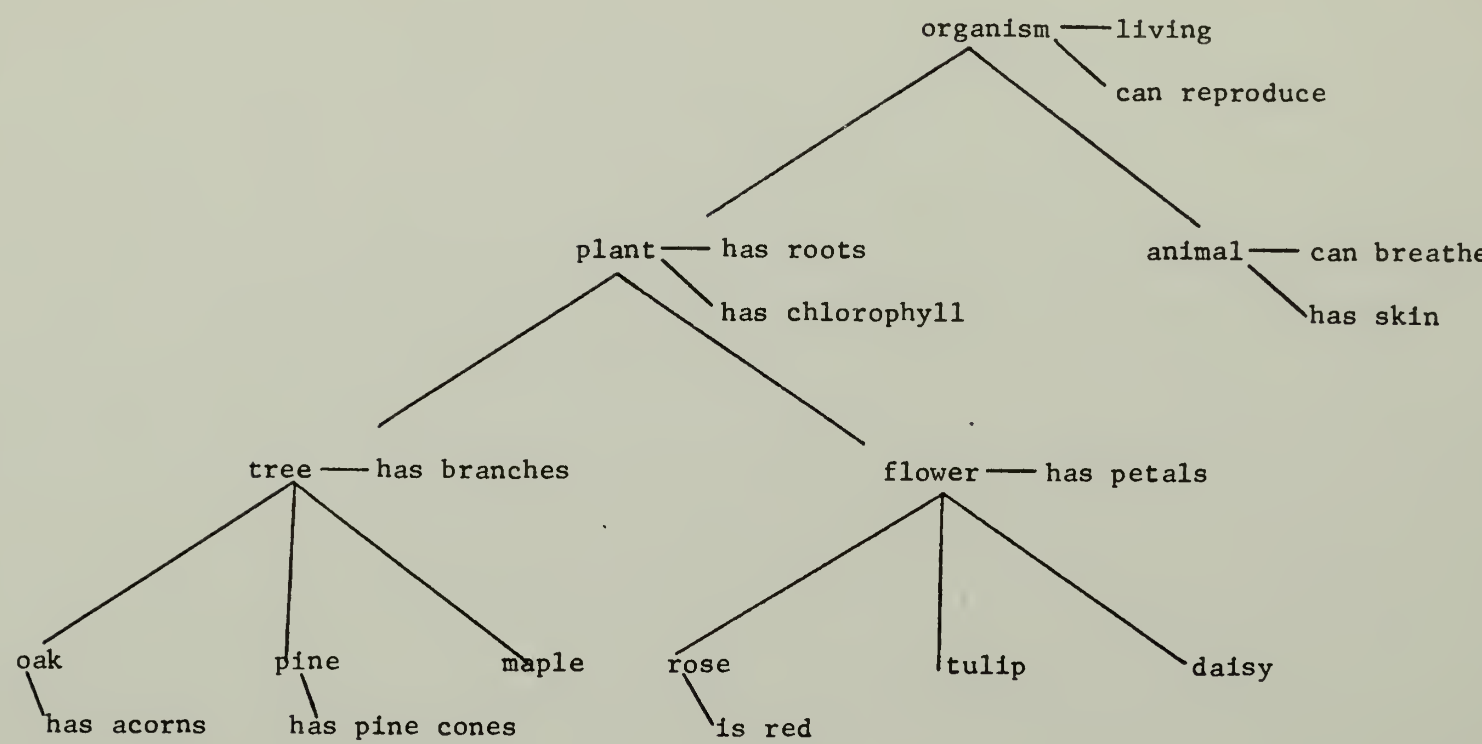

FIG. 3. Semantic hierarchy used to construct sentences from Collins and Quillian (1969).

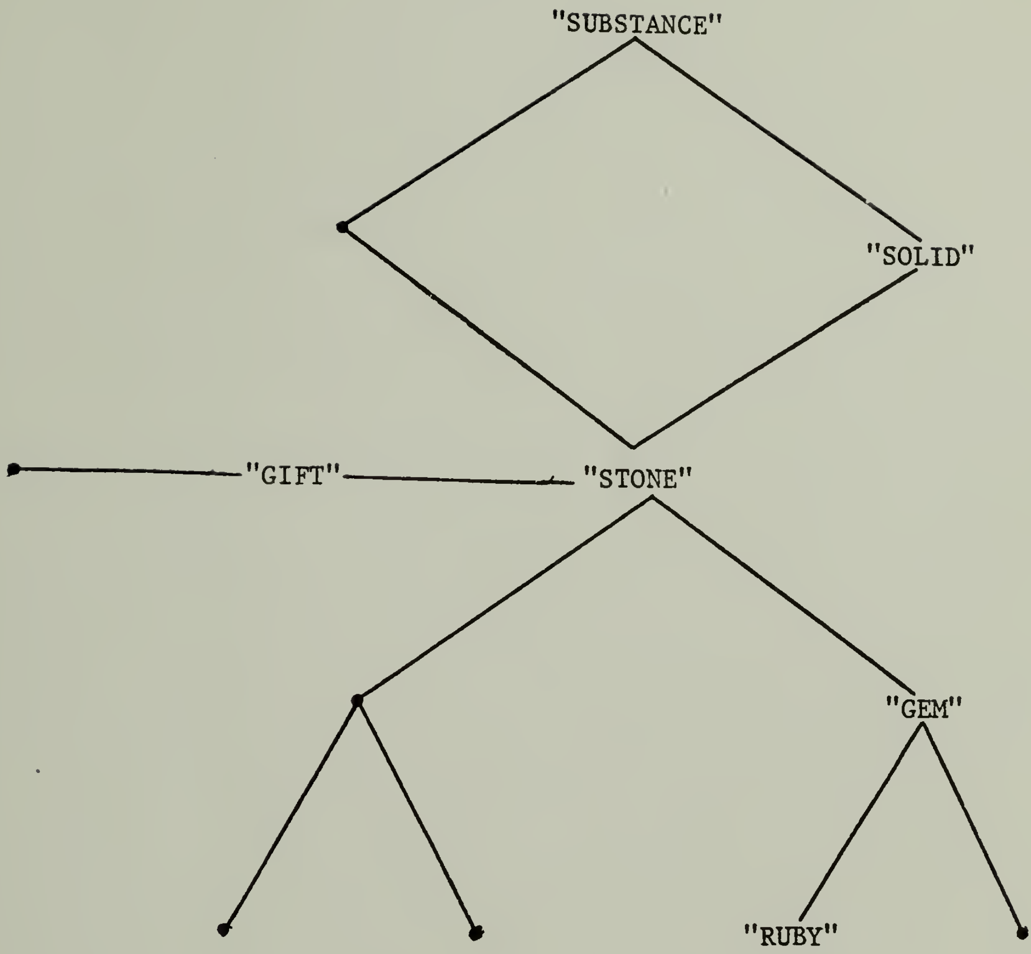

FIG. 4. Predicate intersections file from Meyer (1970).

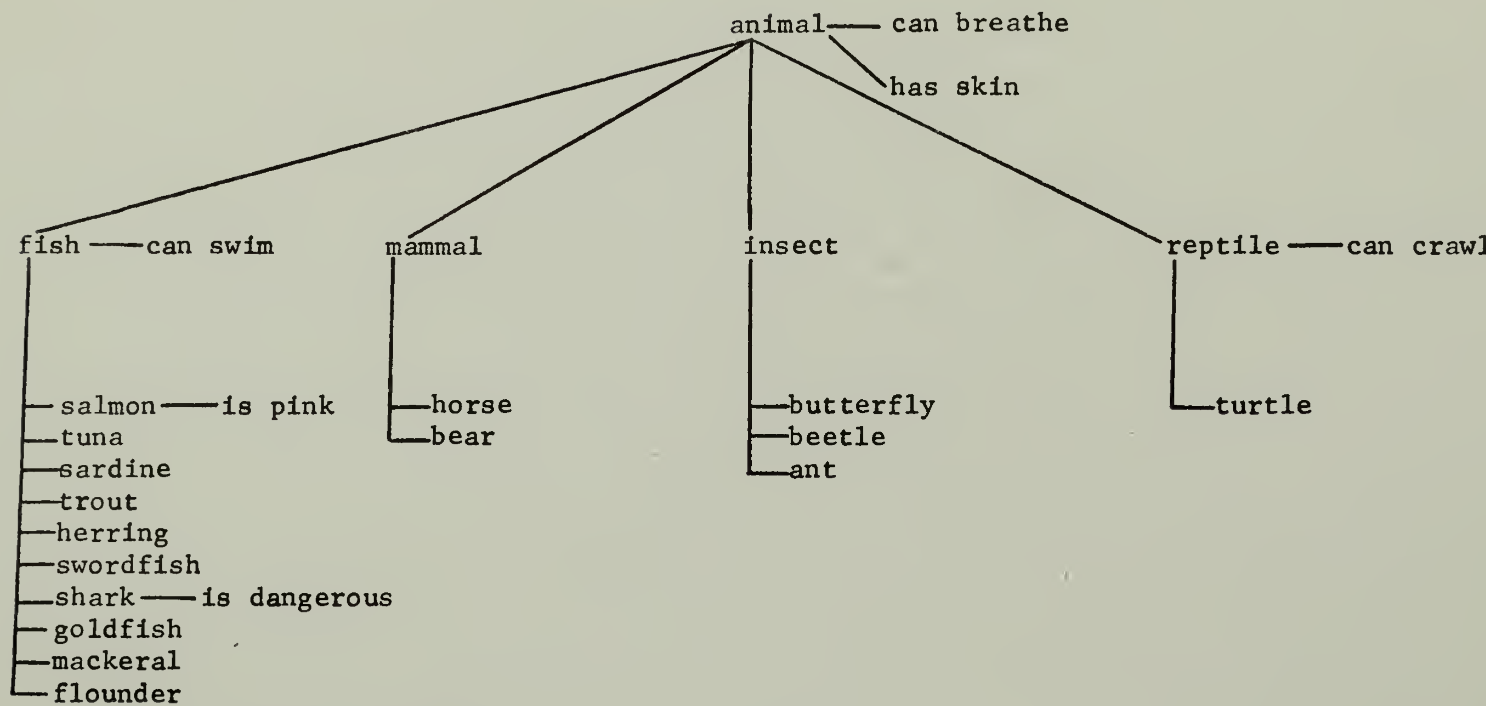

FIG. 5. Semantic hierarchy used to construct true and false sentences in experiment one.

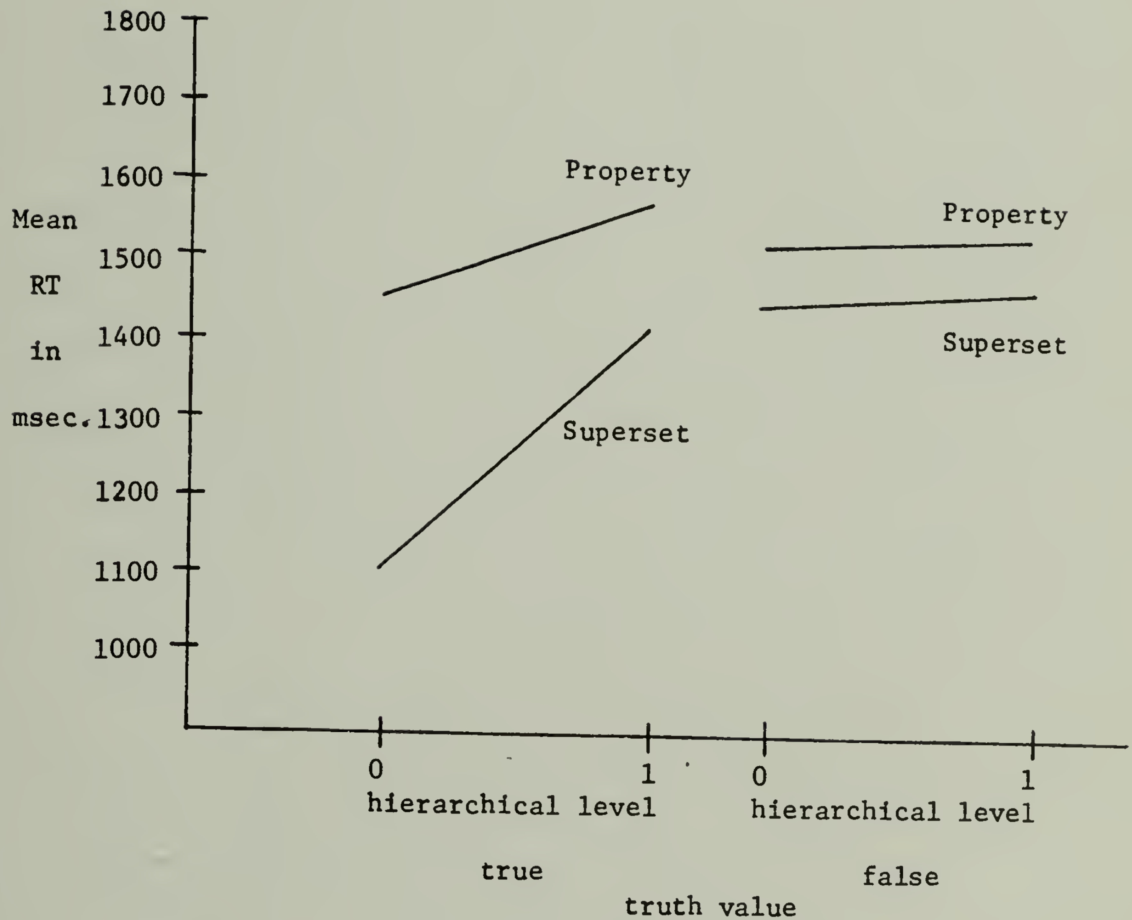

FIG. 6. Mean RT for true and false, property and superset sentences at levels 0 and 1, averaged across broad and narrow hierarchies from experiment one

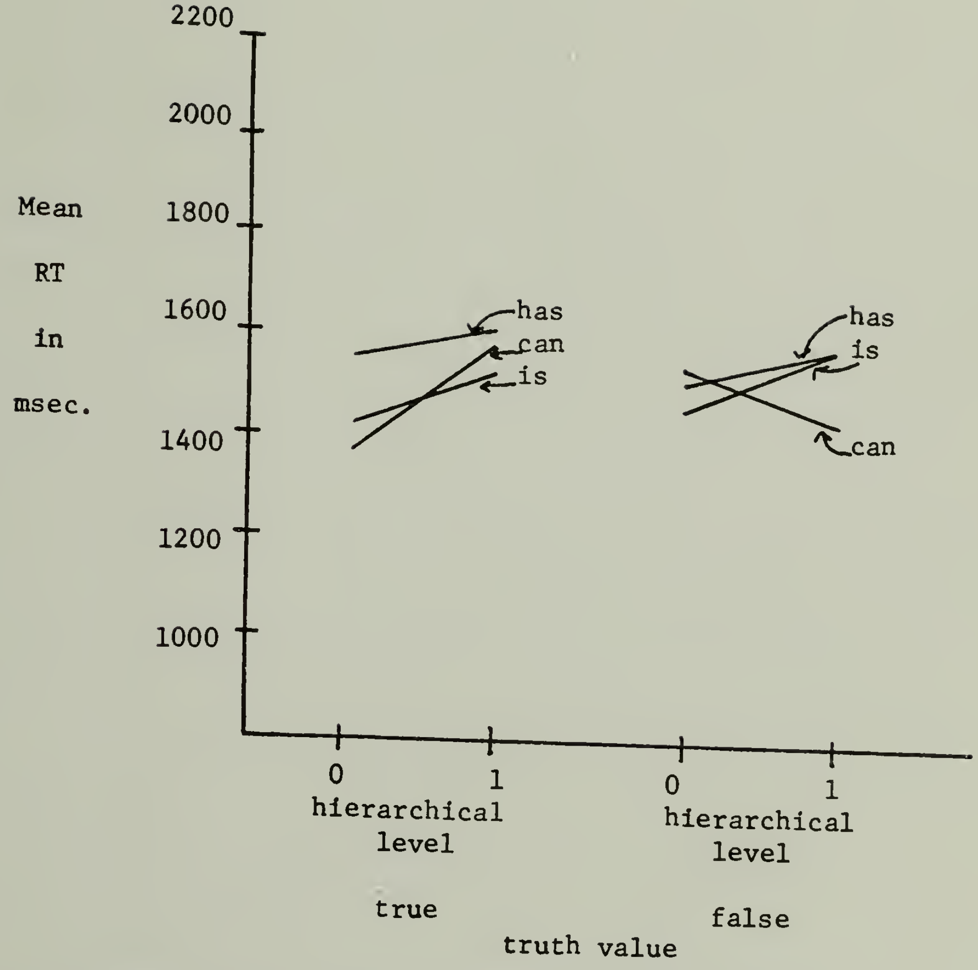

FIG. 7. Mean RT for property relation sentences with different verb phrases from experiment one.

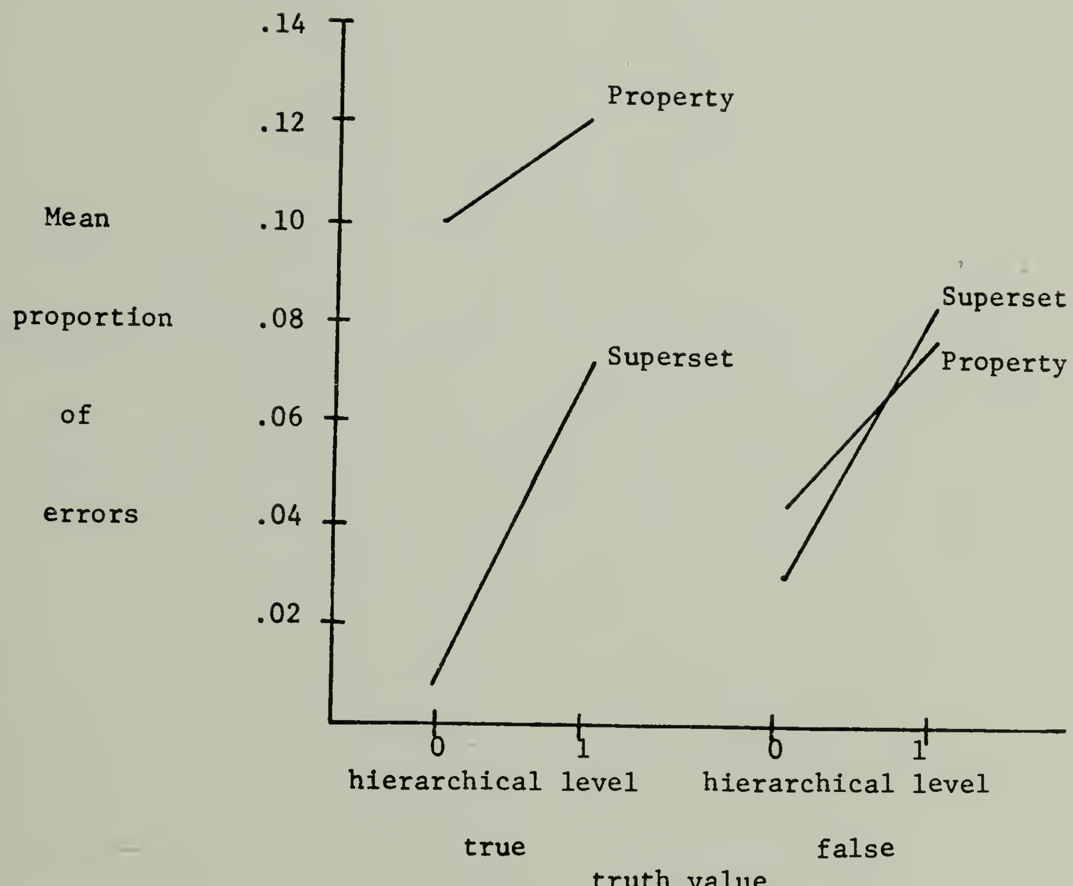

FIG. 8. Mean proportion of errors for all sentence conditions from experiment one.

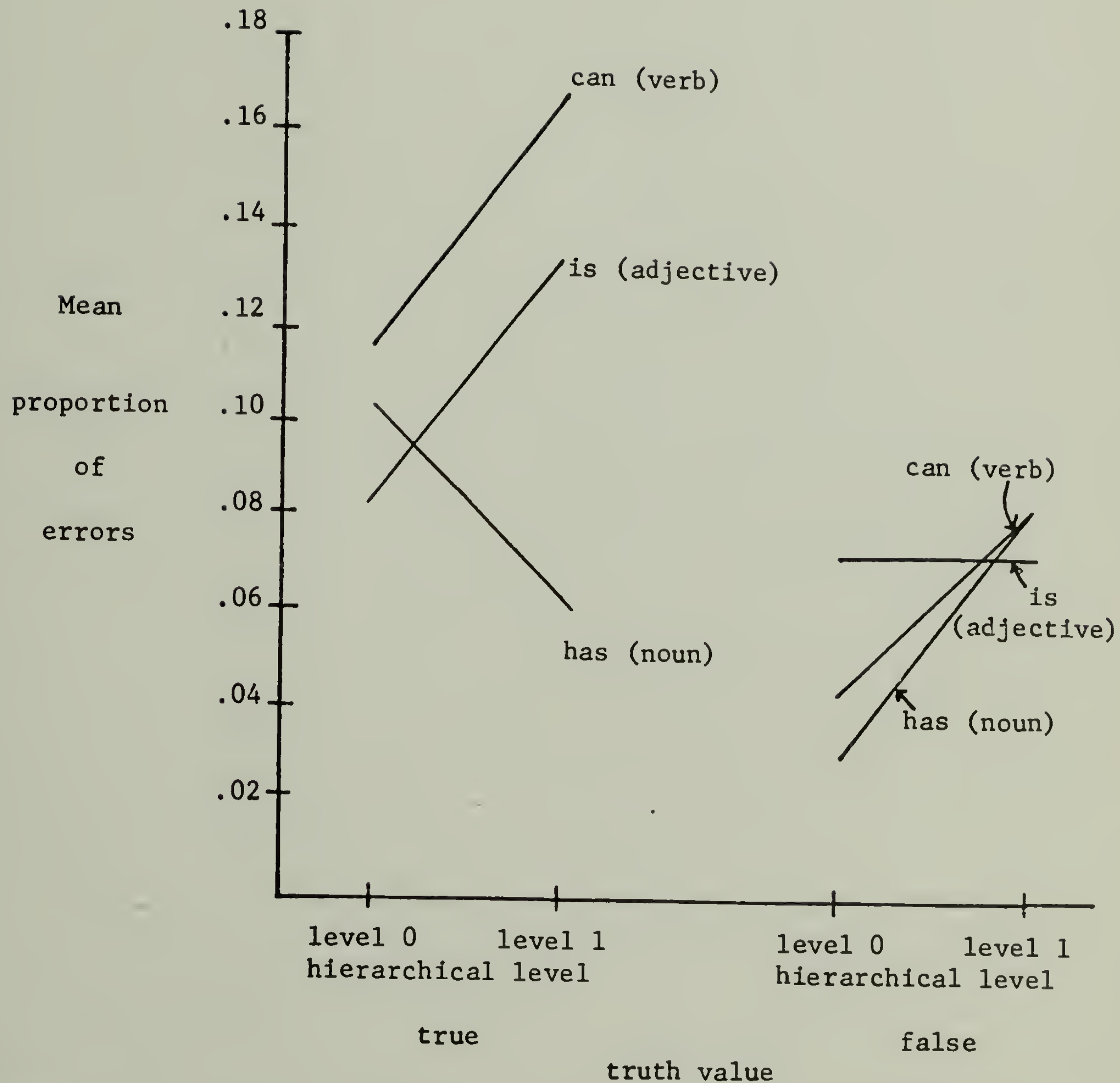

FIG. 9. Mean proportion of errors for true and false property relation sentences using the three different predicates from experiment one.

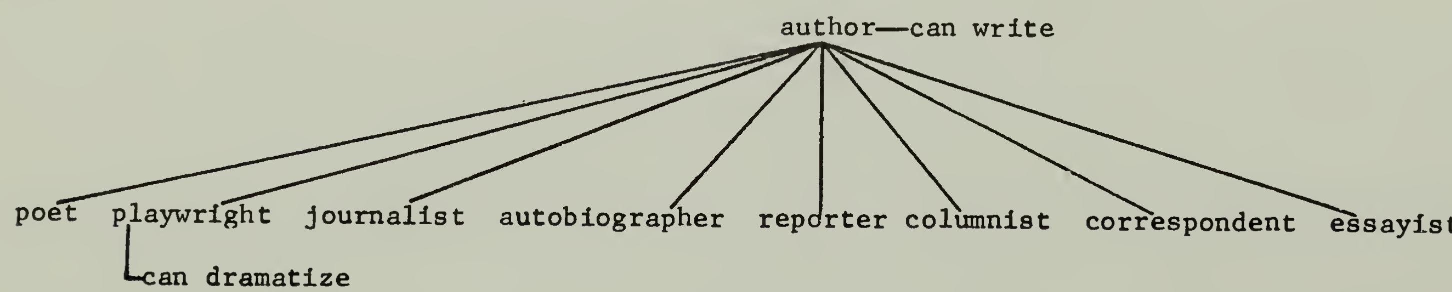

FIG. 10. A two-level semantic hierarchy about writers.

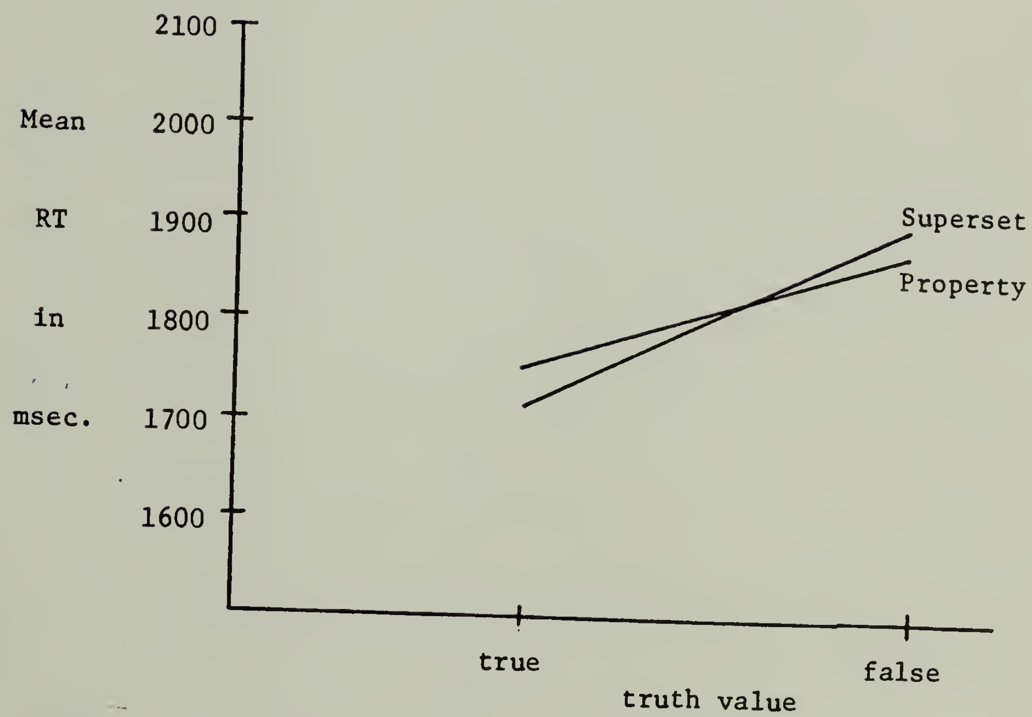

FIG. 11. Mean RT to true and false, property and superset sentences in experiment two.

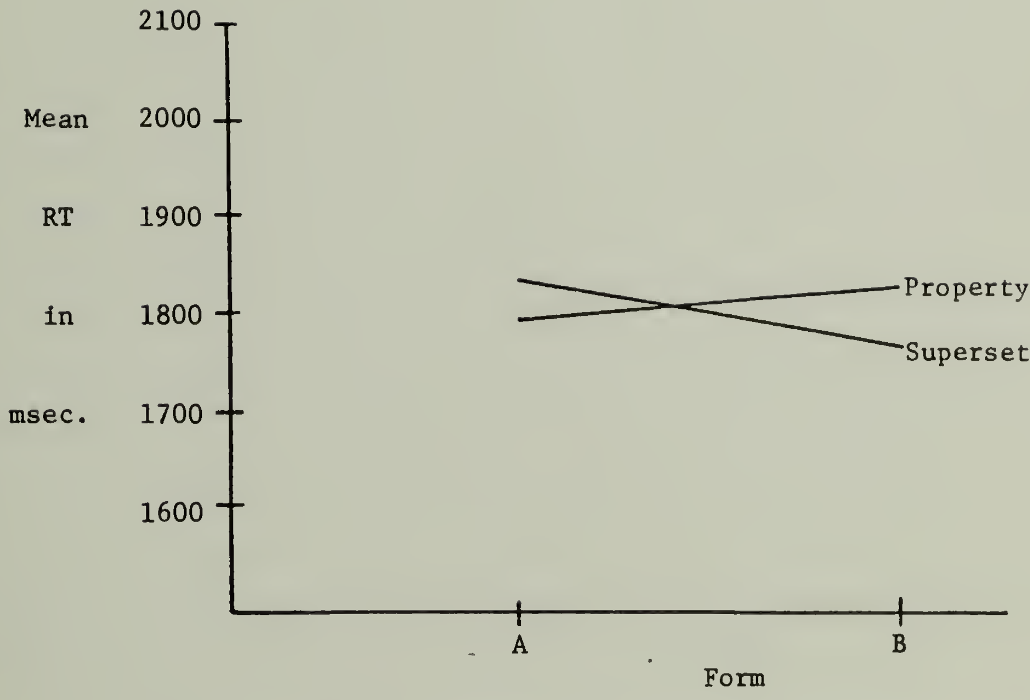

FIG. 12. Mean RT to property and superset sentences in forms A and B in experiment two.

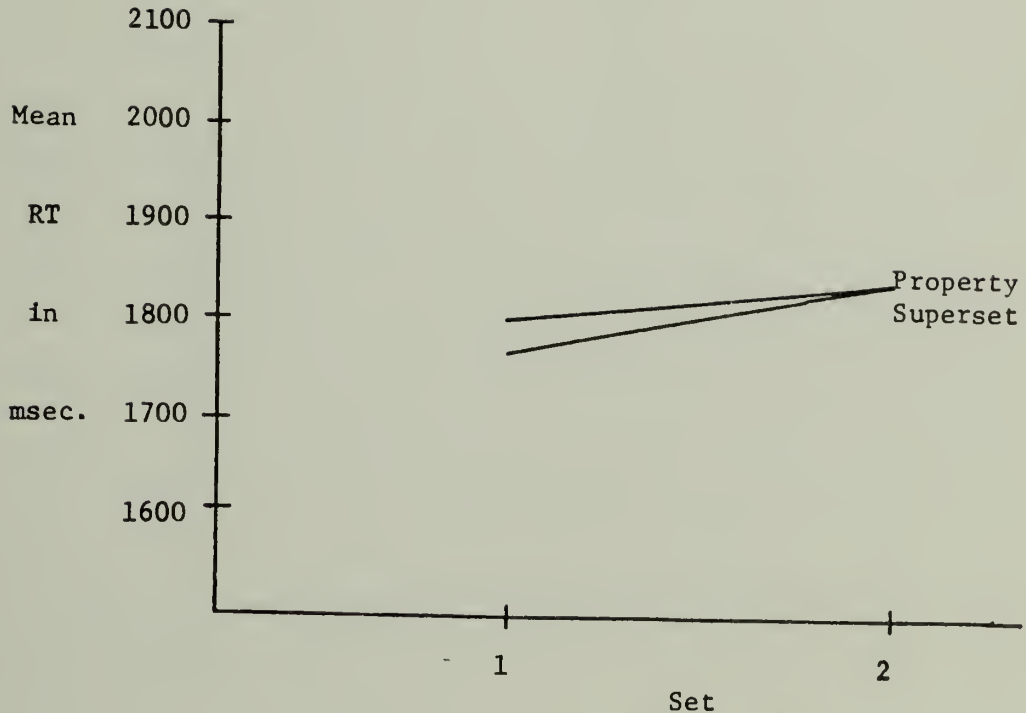

FIG. 13. Mean RT to property and superset sentences in sets 1 and 2 in experiment two. Means shown are same as those in Figure 12, except that they are grouped by sets rather than by form. Set 1 consists of property sentences in form A and superset sentences in form B. Set 2 consists of property sentences in form B and superset sentences in form A.

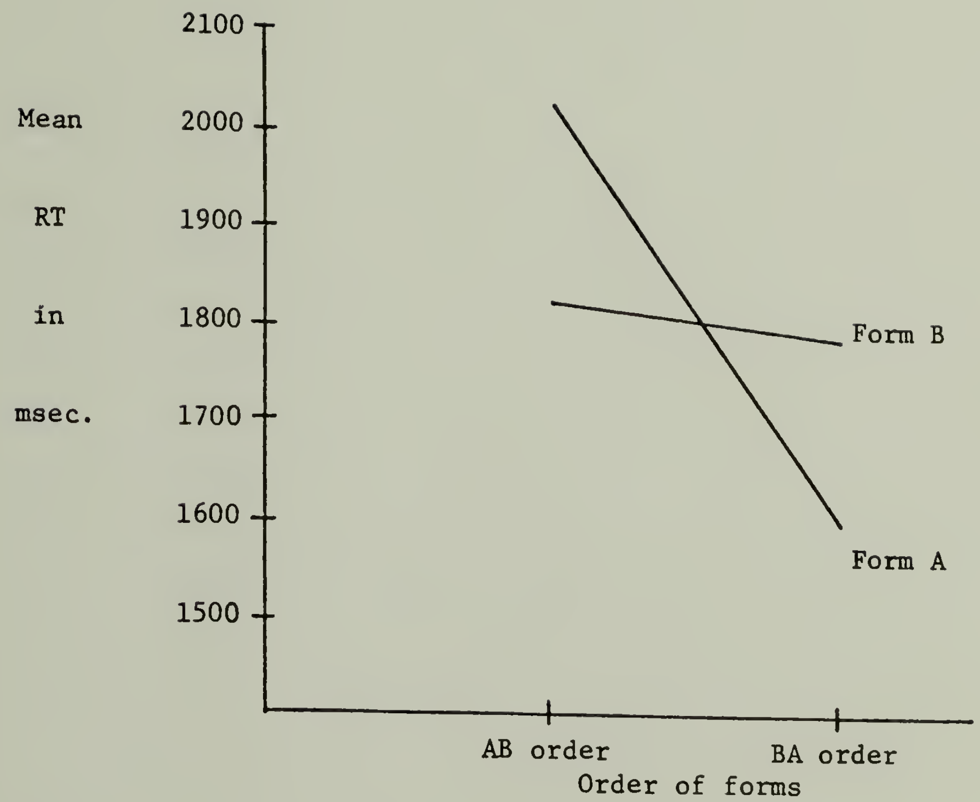

FIG. 14. Mean RT to sentences in form A and form B of subjects in AB order group and subjects in BA order group in experiment two.

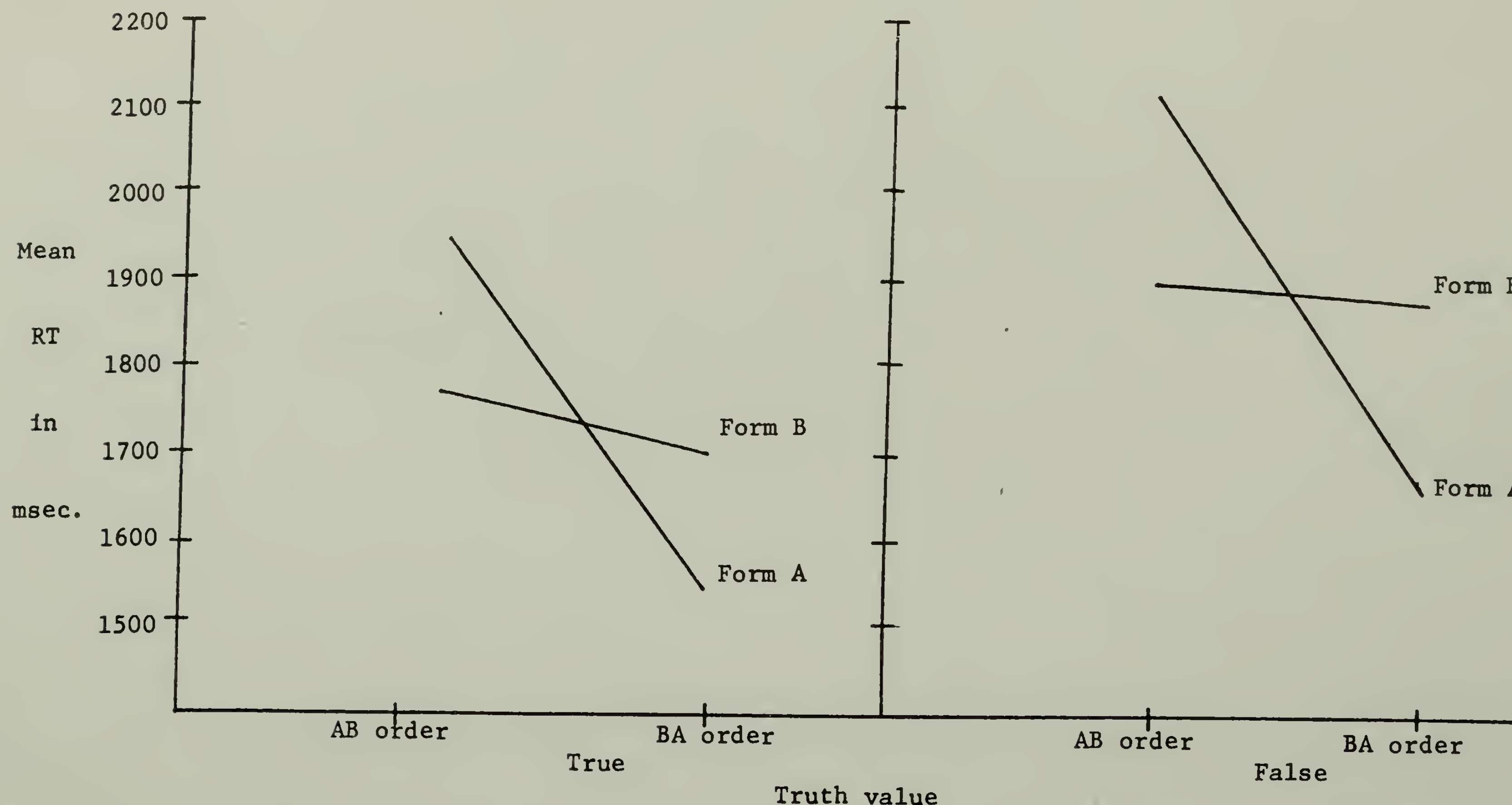

FIG. 15. Mean RT to true and false sentences in forms A and B of subjects in AB order group and subjects in BA order group in experiment two.

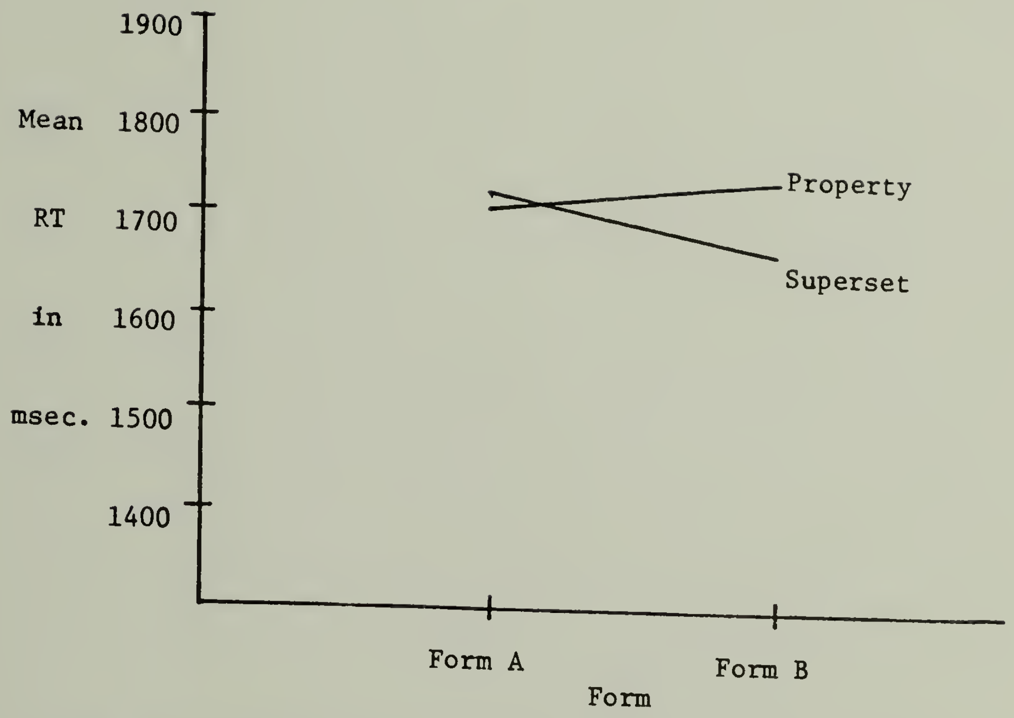

FIG. 16. Mean RT to true property and superset sentences in forms A and B in experiment two.

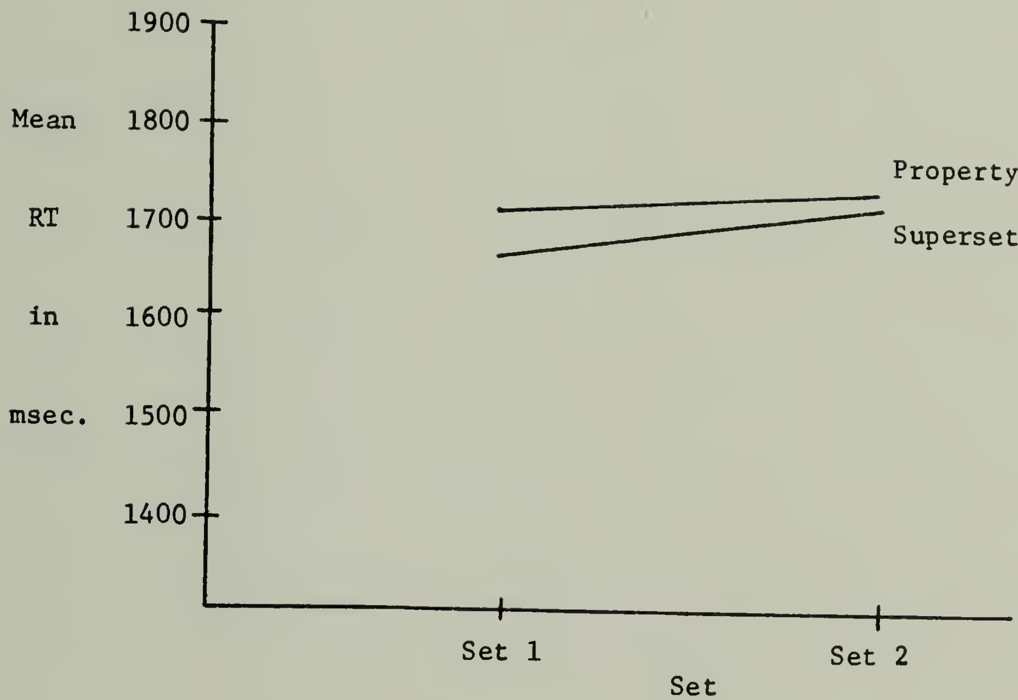

FIG. 17. Mean RT to true property and superset sentences in sets 1 and 2 from experiment two.

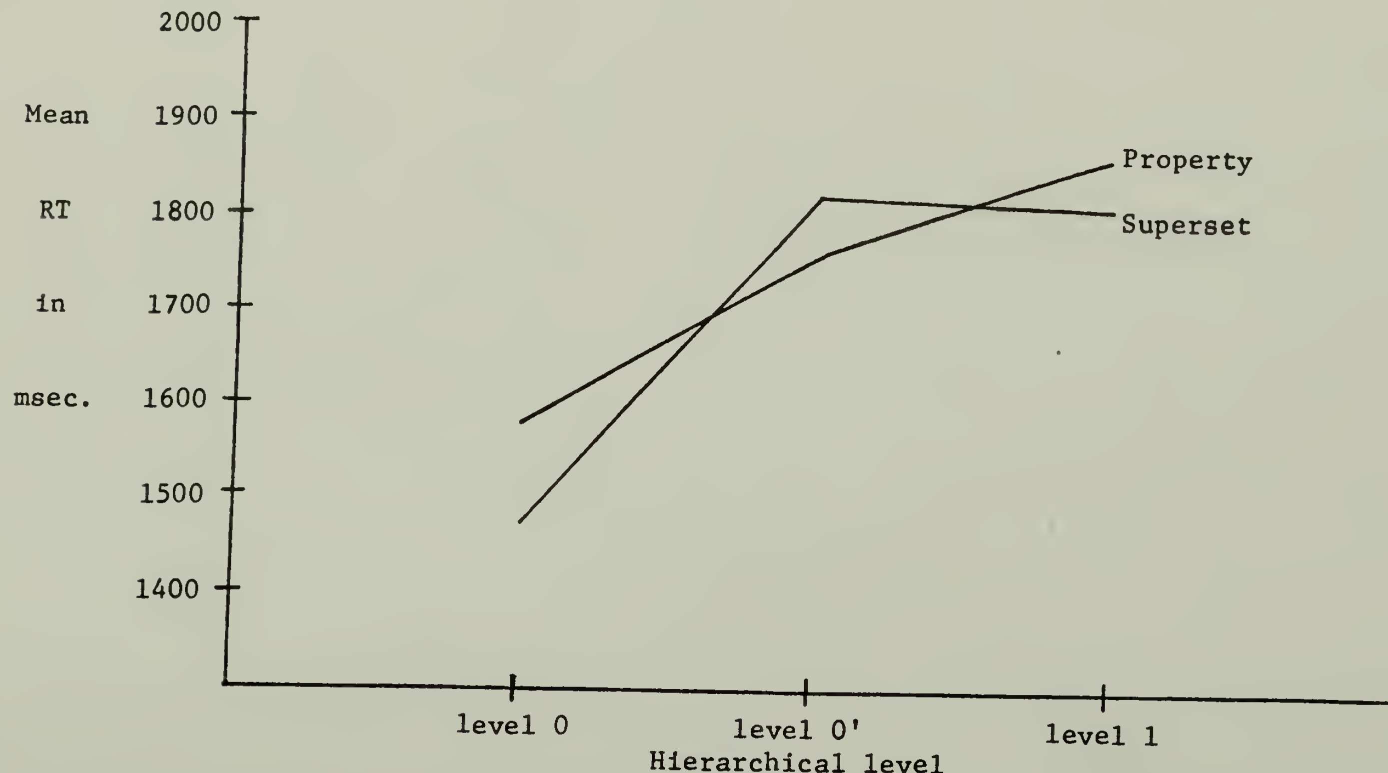

FIG. 18. Mean RT to true property and superset sentences at hierarchical levels 0, 0', and 1 in experiment two.

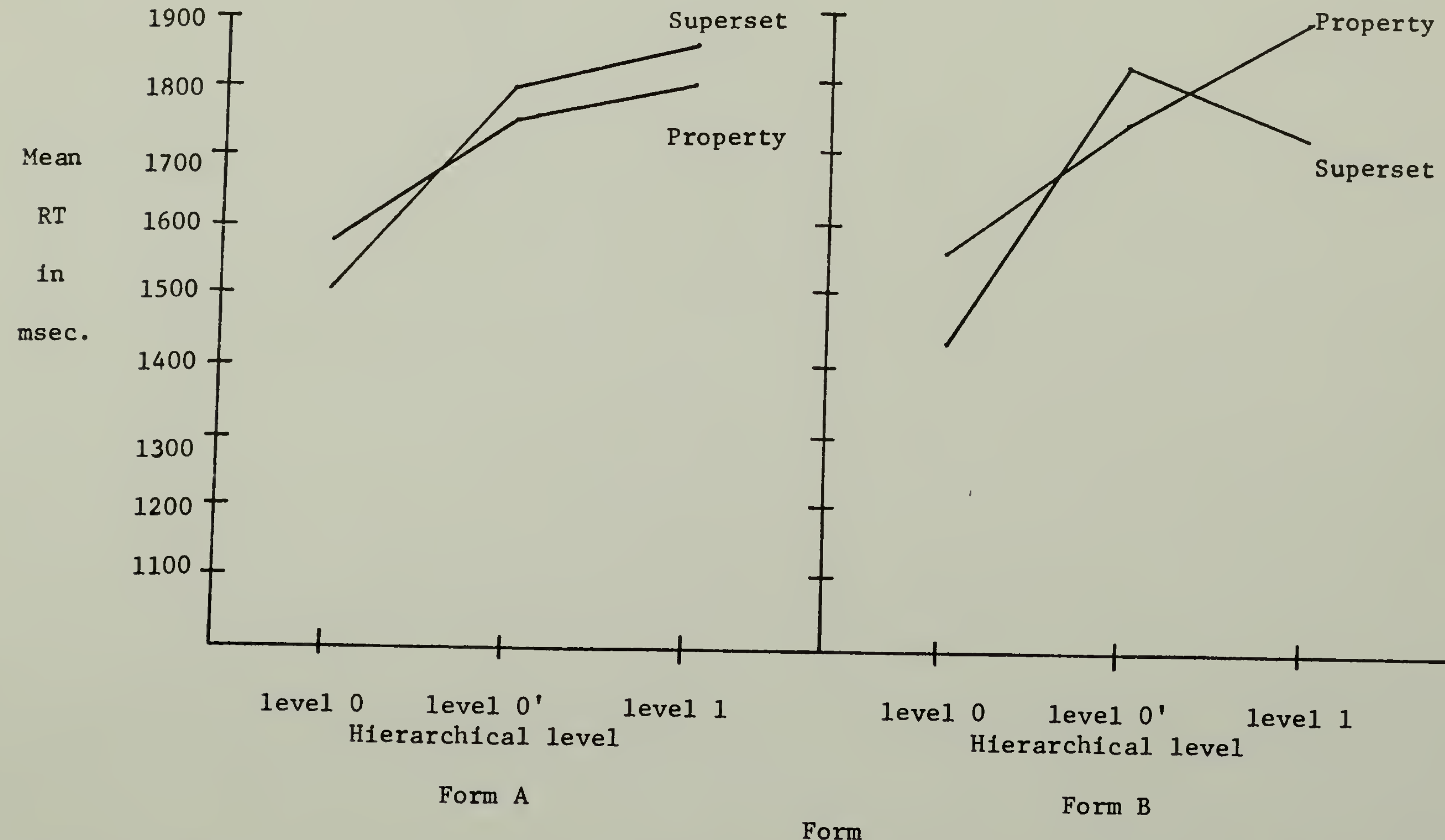

FIG. 19. Mean RT to true property and superset sentences at each hierarchical level in forms A and B from experiment two.

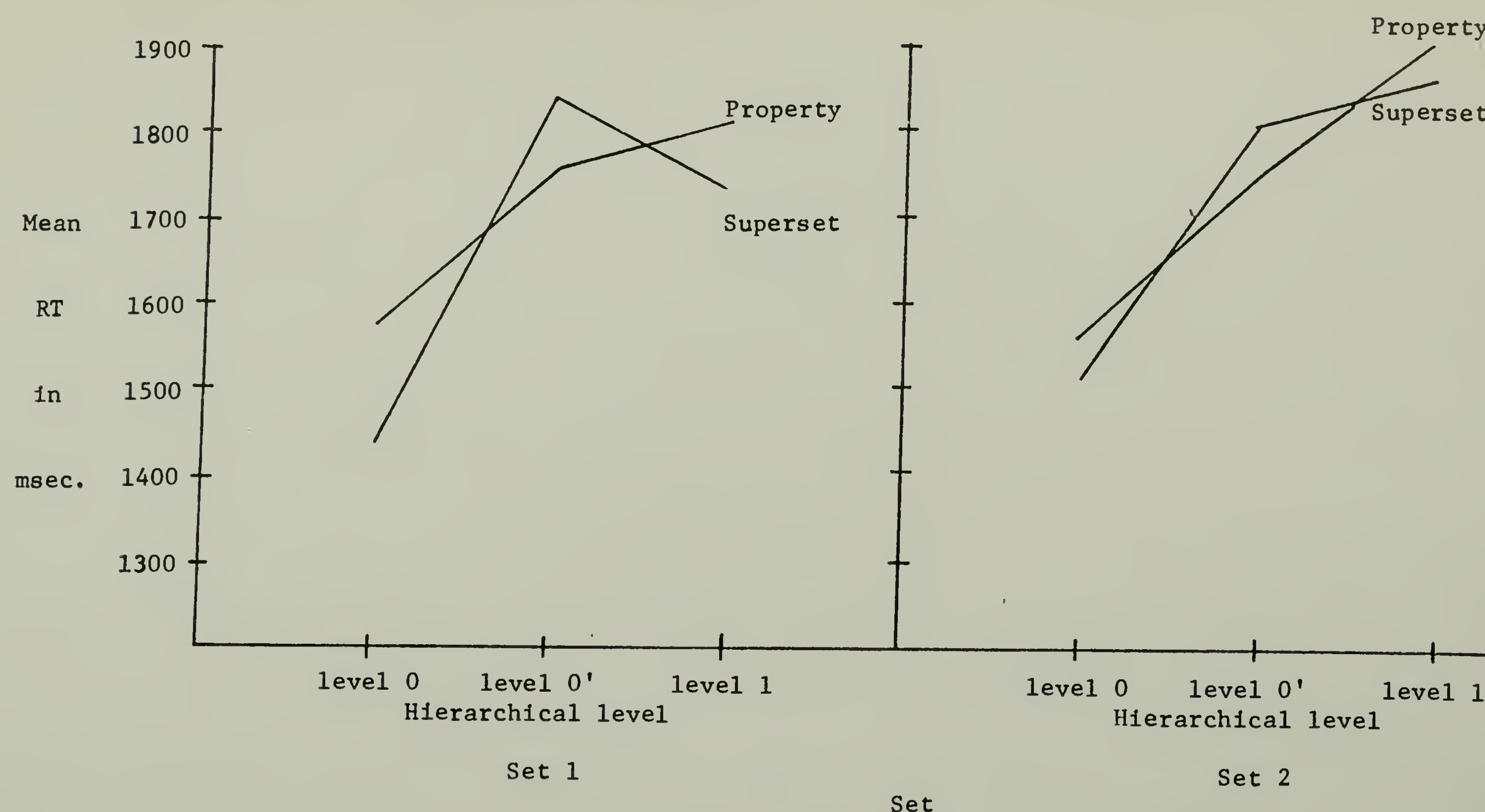

FIG. 20. Mean RT to true property and superset sentences at each hierarchical level in sets 1 and 2 from experiment two. The figure shows the same means as in Figure 19 except that means are grouped by sets rather than by form. Set 1 consists of property sentences in form A and their semantic counterparts, which are the superset sentences in form B. Set 2 consists of property sentences in form B and their semantic counterparts, the superset sentences in form A.

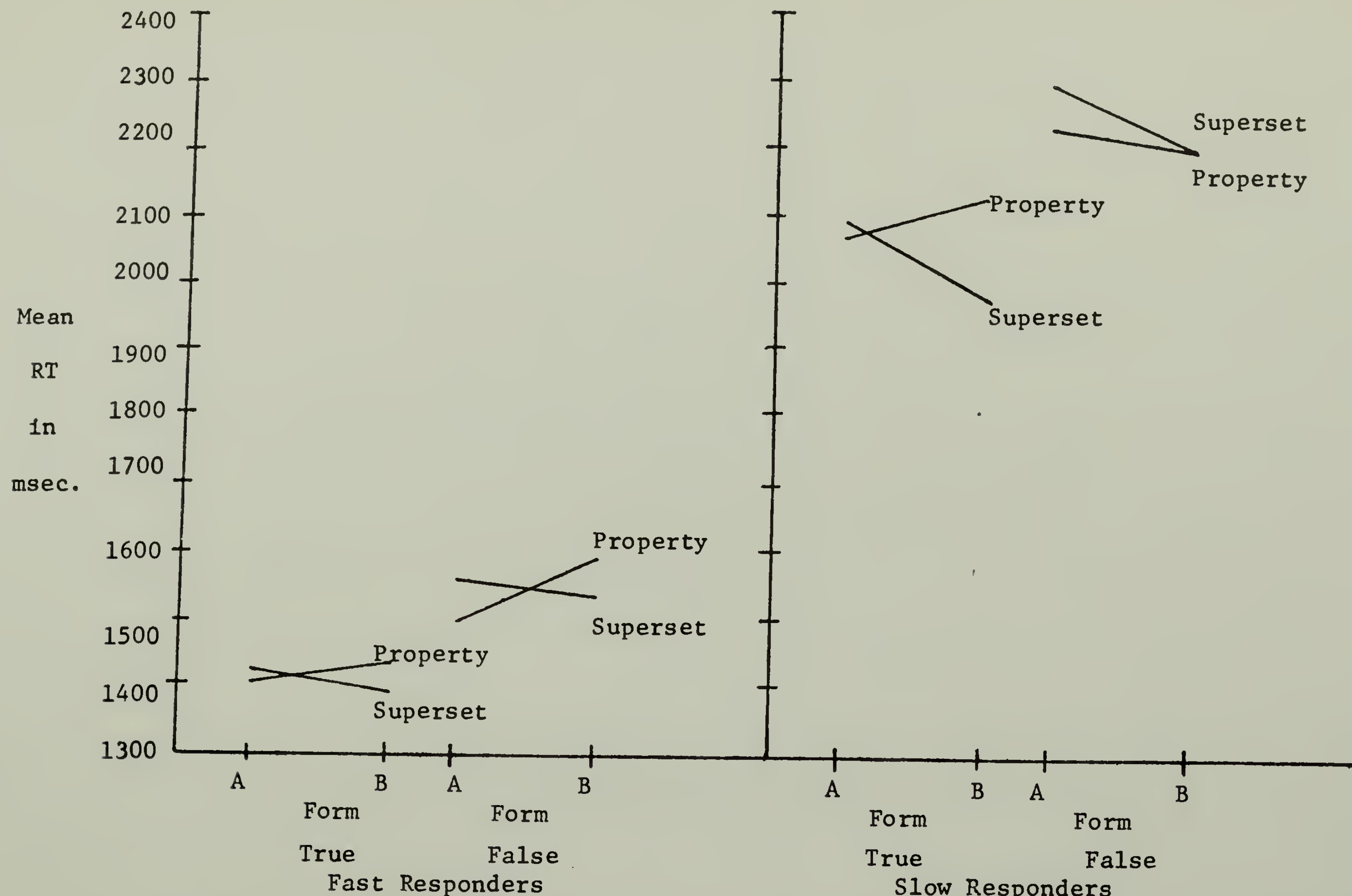

FIG. 21. Mean RTs to true and false, property and superset sentences in forms A and B for fast and slow responders in experiment two.

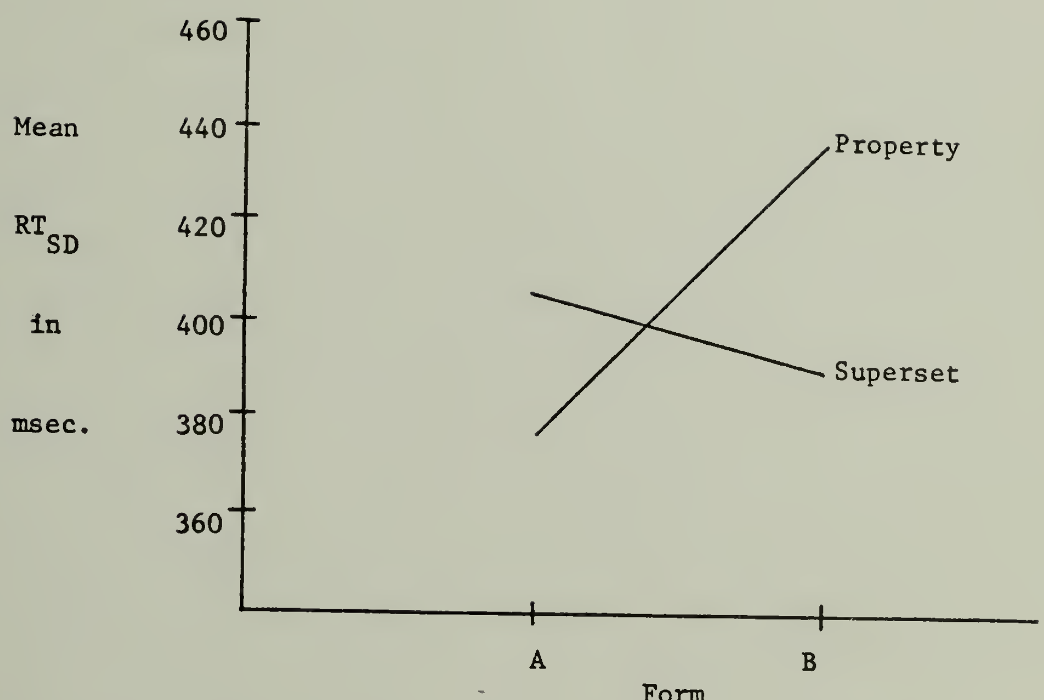

FIG. 22. Standard deviation of mean RT to property and superset sentences in forms A and B from experiment two.

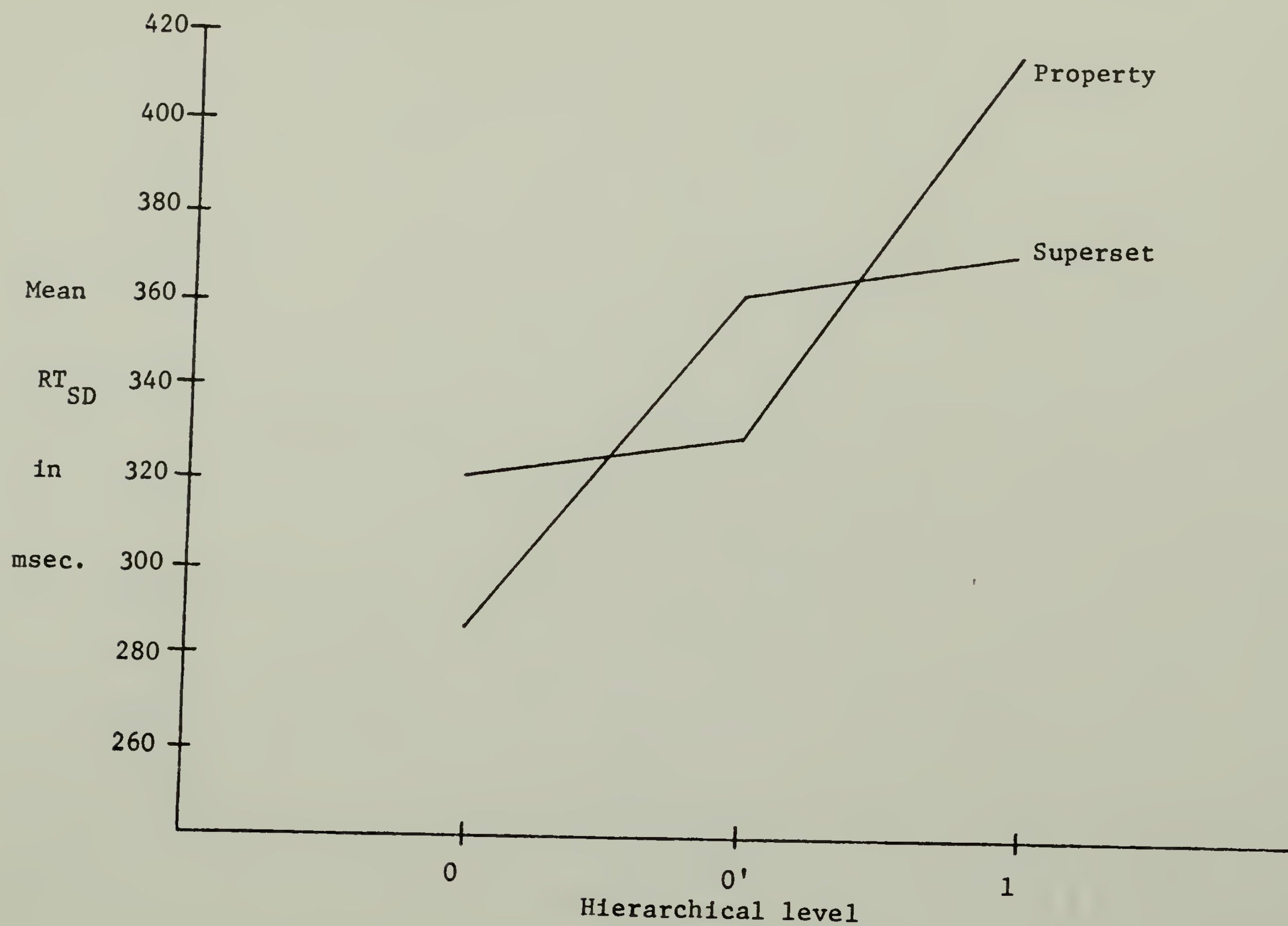

FIG. 23. Standard deviation of mean RT to true property and superset sentences at each hierarchical level from experiment two.

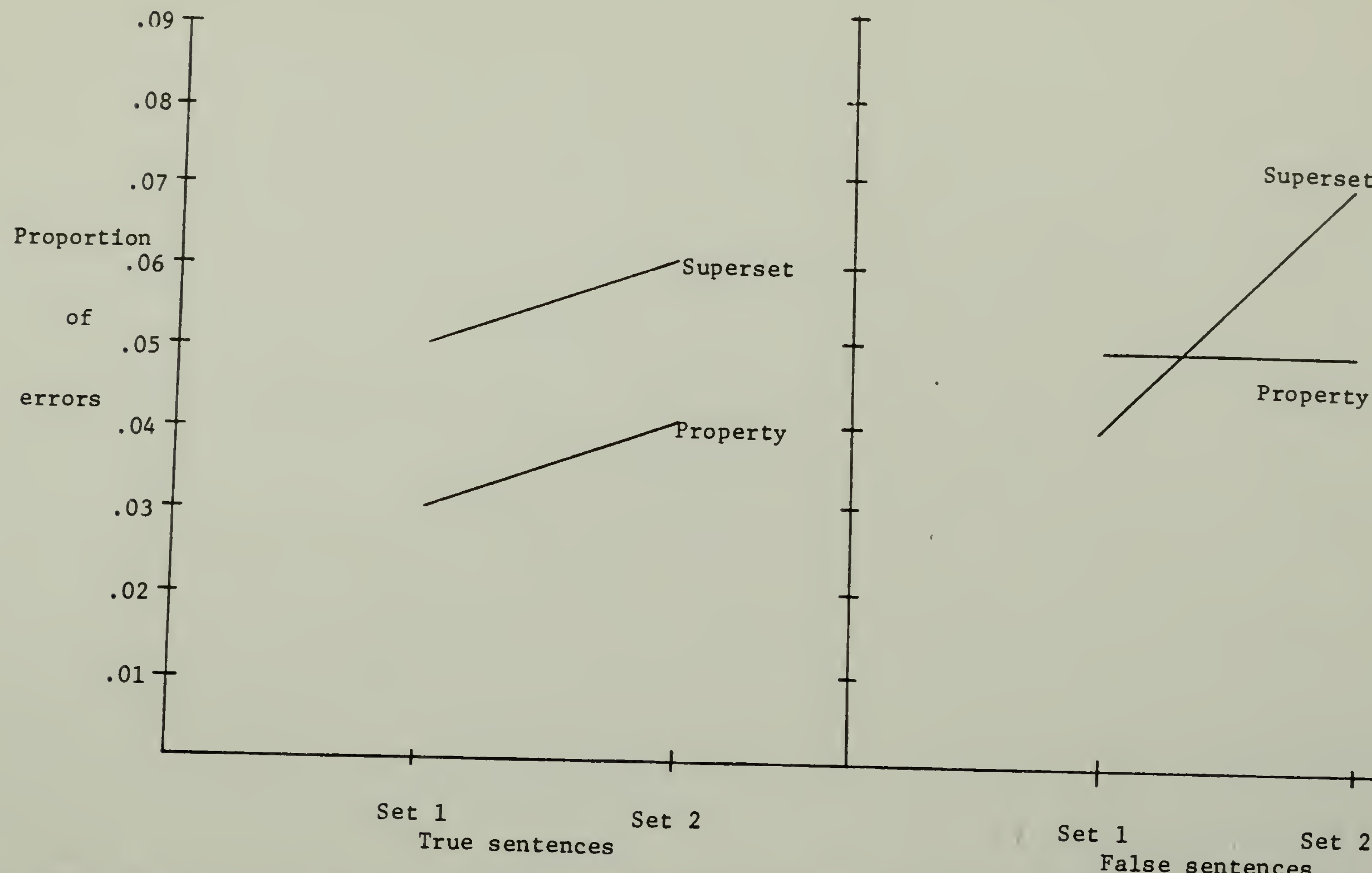

FIG. 24. Proportion of errors for true and false sentences in set 1 (P_A and S_B) and set 2 (P_B and S_A) from experiment two.

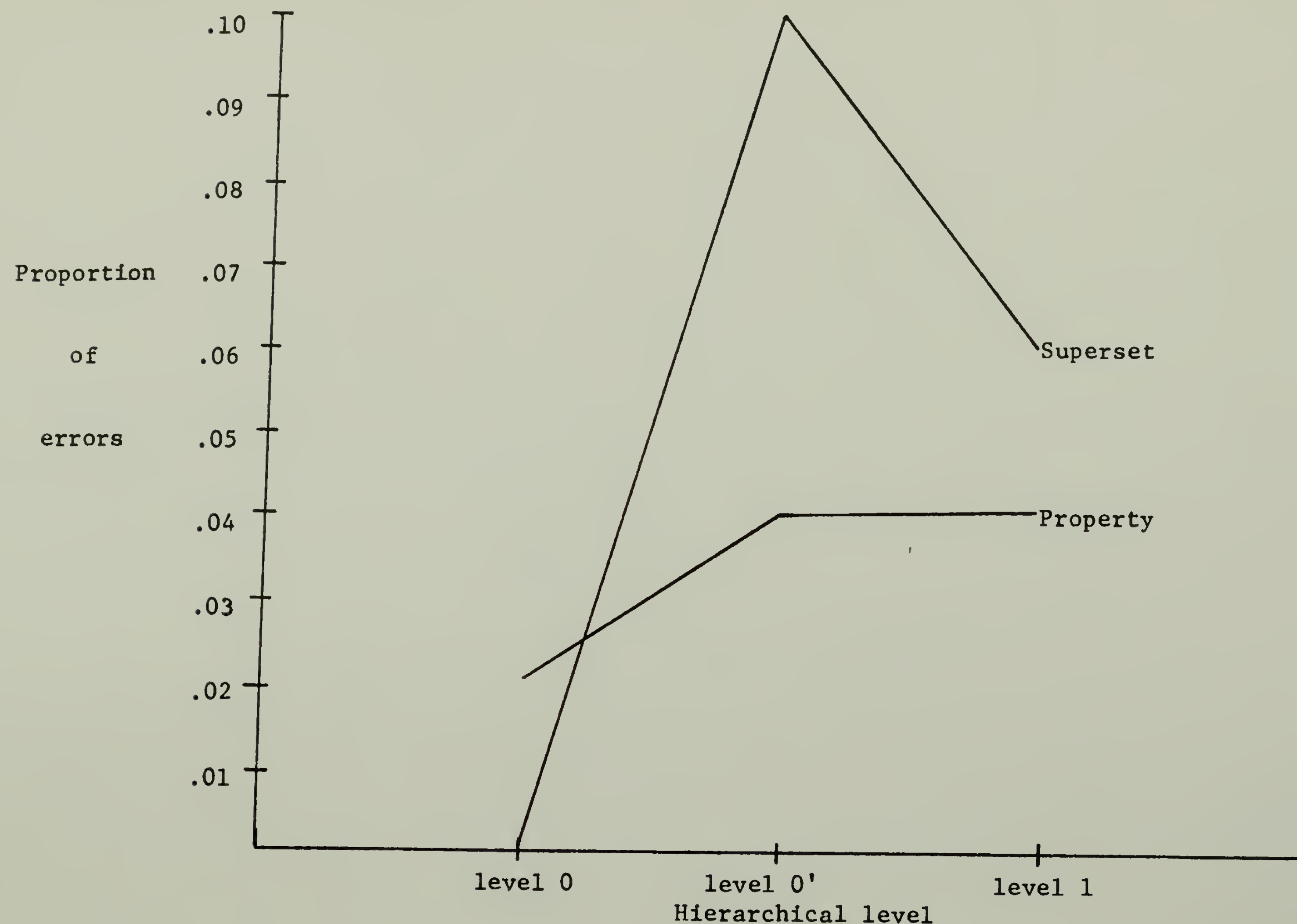

FIG. 25. Proportion of errors for true property and superset sentences at each hierarchical level from experiment two.

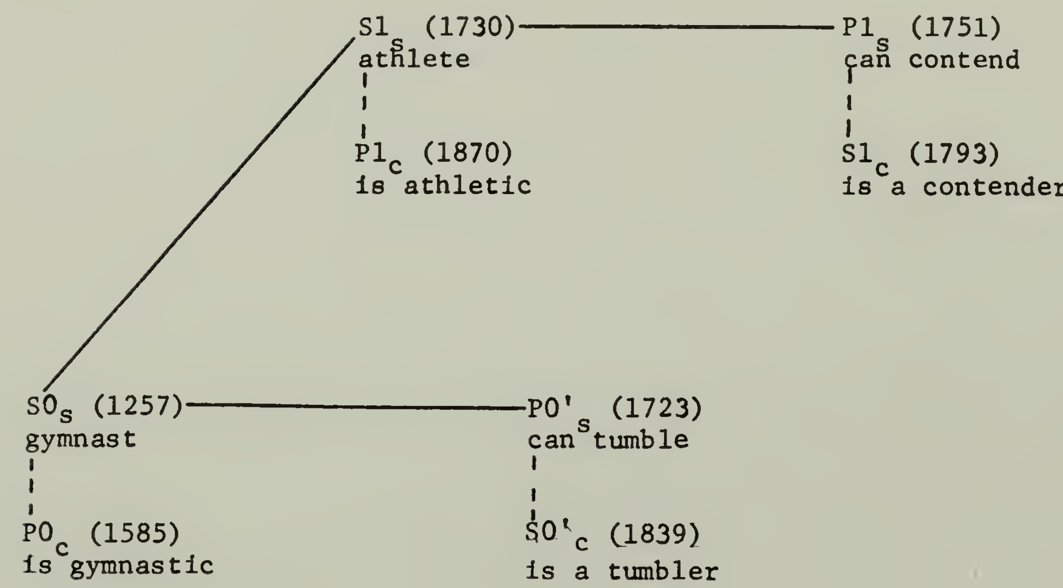

FIG. 26. Memory network showing associative connections and grammatical operations involved in processing different sentence types. "Gymnast" was not used as an instance of the category "athlete" in this experiment. It is shown in the network to illustrate a lexically simple level 0 category. The other instances and properties were used in sentences in experiment two.

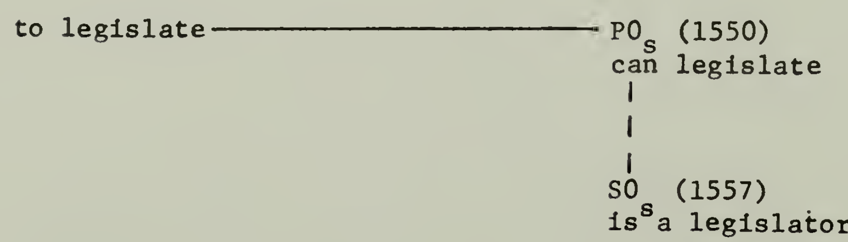

FIG. 27. Elaboration of memory network to incorporate RTs to PO_s and SO_c sentences.

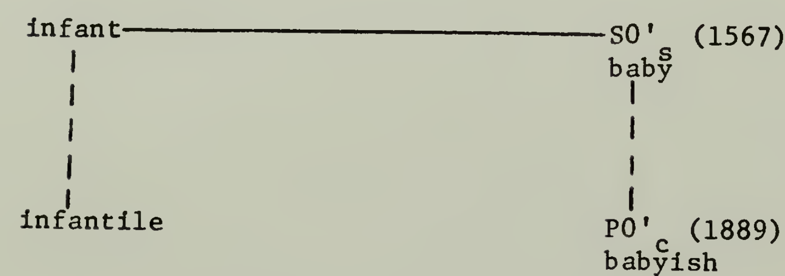

FIG. 28. Memory network to represent the relationship between the synonymous terms "infant" and "baby." The sentences used in experiment two were: "An infant is a "baby," and "An infant is babyish."

Appendix

TABLE A

Sentences Used in Replication of Collins and Quillian (1969)

Sentence Conditions	Narrow Hierarchies	
	True Sentences	False Sentences
P0	BASEBALL HAS INNINGS	SKATING HAS PAWNS
P1	BADMINTON HAS RULES	TENNIS IS RISKY
S0	FISHING IS FISHING	BOWLING IS A RACE
S1	SOCCER IS A SPORT	SWIMMING IS A TOY
P0	A SHARK IS DANGEROUS	A HERRING CAN CLIMB
P1	A SALMON CAN SWIM	A GOLDFISH CAN CRAWL
S0	A SARDINE IS A SARDINE	A LOBSTER IS A TURTLE
S1	A PIKE IS A FISH	A MACKERAL IS A REPTILE
P0	A SUBMARINE CAN SUBMERGE	A CANOE IS SUPERSONIC
P1	A SCHOONER HAS A RUDDER	A ROWBOAT CAN FLY
S0	A SAILBOAT IS A SAILBOAT	A RAFT IS A HELICOPTER
S1	AN OCEAN LINER IS A SHIP	A YACHT IS AN AUTOMOBILE

TABLE A cont'd.

PO	A CANARY CAN SING	A DOVE CAN SPEAK
P1	A WREN CAN FLY	AN EAGLE IS FURRY
SO	A ROBIN IS A ROBIN	A SPARROW IS A COW
S1	A PARROT IS A BIRD	A PIGEON IS AN INSECT
PO	A SURGEON HAS A SCALPEL	A DENTIST CAN BAPTIZE
P1	AN OBSTETRICIAN CAN VACCINATE	A DERMATOLOGIST CAN LEGISLATE
SO	A NEUROLOGIST IS A NEUROLOGIST	A PEDIATRICIAN IS A MINISTER
S1	A GENERAL PRACTITIONER IS A DOCTOR	A PSYCHIATRIST IS A POLITICIAN
PO	A PEA IS GREEN	BROCCOLI HAS A SHELL
P1	A BEAN IS GROWN	SPINACH HAS BONES
SO	ASPARAGUS IS ASPARAGUS	A TURNIP IS A FRANKFURTER
S1	A CARROT IS A VEGETABLE	A CUCUMBER IS A DRINK
PO	A LOCKET CAN OPEN	A CUFF LINK HAS LACES
P1	A NECKLACE IS ORNAMENTAL	A TIE CLASP HAS A SOLE
SO	A RING IS A RING	AN EARRING IS A DERBY
S1	A BRACELET IS JEWELRY	A WATCH IS A HAT

TABLE A cont'd.

PO	AN OAK HAS ACORNS	A WILLOW IS RED
P1	A SPRUCE HAS BRANCHES	A BIRCH HAS PETALS
SO	A MAPLE IS A MAPLE	A SYCAMORE IS A LILY
S1	A CEDAR IS A TREE	A REDWOOD IS A FLOWER
PO	CHAMPAGNE IS SPARKLING	MADEIRA IS GREEN
P1	CHIANTI IS ALCOHOLIC	BRANDY IS CARBONATED
SO	SHERRY IS SHERRY	CLARET IS ROOT BEER
S1	ROSÉ IS WINE	BURGUNDY IS WHISKY
PO	A KNIFE CAN CUT	A STRAINER CAN FRY
P1	A SPOON CAN MIX	A FUNNEL HAS A PLUG
SO	A FORK IS A FORK	A CAN OPENER IS A SAUCER
S1	A SPATULA IS A UTENSIL	A CORKSCREW IS A DISH
PO	SLATE IS GRAY	GRANITE IS GREEN
P1	A PEBBLE IS SOLID	SAND IS METALLIC
SO	LIMESTONE IS LIMESTONE	A COBBLESTONE IS A DIAMOND
S1	GRAVEL IS STONE	SHALE IS A GEM

TABLE A cont'd.

PO	SAUSAGE HAS SKIN	A HAMBURGER HAS DRUMSTICKS
P1	LAMB HAS FAT	HAM CAN SPILL
SO	VEAL IS VEAL	A SALAMI IS LETTUCE
S1	PORK IS MEAT	MUTTON IS FRUIT

Broad Hierarchies

PO	CHESS HAS PAWNS	HOCKEY HAS INNINGS
P1	ROULETTE IS RISKY	A TEDDY BEAR HAS RULES
SO	A RACE IS A RACE	BASKETBALL IS FISHING
S1	A DOLLHOUSE IS A TOY	A PUPPET IS A SPORT
PO	A MONKEY CAN CLIMB	A LAMB IS DANGEROUS
P1	A SNAKE CAN CRAWL	A BUTTERFLY CAN SWIM
SO	A TURTLE IS A TURTLE	A GOAT IS A SARDINE
S1	AN ALLIGATOR IS A REPTILE	A CAMEL IS A FISH
PO	A JET IS SUPERSONIC	A TAXI CAN SUBMERGE
P1	A SPACESHIP CAN FLY	A BUS HAS A RUDDER
SO	A HELICOPTER IS A HELICOPTER	A CABOOSE IS A SAILBOAT
S1	A SEDAN IS AN AUTOMOBILE	A TRUCK IS A SHIP

TABLE A cont'd.

PO	A MAN CAN SPEAK	A TUNA CAN SING
P1	A DOG IS FURRY	A PIG CAN FLY
SO	A COW IS A COW	A GIRAFFE IS A ROBIN
S1	AN ANT IS AN INSECT	A DEER IS A BIRD
PO	A PRIEST CAN BAPTIZE	A PREACHER HAS A SCALPEL
P1	A REPRESENTATIVE CAN LEGISLATE	AN ENGINEER CAN VACCINATE
SO	A MINISTER IS A MINISTER	A REVEREND IS A NEUROLOGIST
S1	A SENATOR IS A POLITICIAN	A GOVERNOR IS A DOCTOR
PO	A CLAM HAS A SHELL	A RADISH IS GREEN
P1	A TROUT HAS BONES	JELLO IS GROWN
SO	A FRANKFURTER IS A FRANKFURTER	A TOMATO IS ASPARAGUS
S1	A SODA IS A DRINK	AN APPLE IS A VEGETABLE
PO	A SNEAKER HAS LACES	A BERET CAN OPEN
P1	A SLIPPER HAS A SOLE	A GALOSHE IS ORNAMENTAL
SO	A DERBY IS A DERBY	A STOCKING IS A RING
S1	A BONNET IS A HAT	A HELMET IS JEWELRY

TABLE A cont'd.

PO	A ROSE IS RED	A GRAPEVINE HAS ACORNS
P1	A DAISY HAS PETALS	A MUSHROOM HAS BRANCHES
SO	A LILY IS A LILY	A CACTUS IS A MAPLE
S1	A BUTTERCUP IS A FLOWER	A ZINNIA IS A TREE
PO	LIME SODA IS GREEN	COFFEE IS SPARKLING
P1	SODA IS CARBONATED	MILK IS ALCOHOLIC
SO	ROOT BEER IS ROOT BEER	TEA IS SHERRY
S1	BOURBON IS WHISKY	SCOTCH IS WINE
PO	A SKILLET CAN FRY	A KETTLE CAN CUT
P1	A TOASTER HAS A PLUG	A REFRIGERATOR CAN MIX
SO	A SAUCER IS A SAUCER	A BOTTLE IS A FORK
S1	A PLATTER IS A DISH	AN OVEN IS A UTENSIL
PO	AN EMERALD IS GREEN	JADE IS GRAY
P1	IRON IS METALLIC	OIL IS SOLID
SO	A DIAMOND IS A DIAMOND	GOLD IS LIMESTONE
S1	A RUBY IS A GEM	ALUMINUM IS A STONE

TABLE A cont'd.

P0	A CHICKEN HAS DRUMSTICKS	A COOKIE HAS SKIN
P1	JUICE CAN SPILL	A PINEAPPLE HAS FAT
S0	LETTUCE IS LETTUCE	A TURKEY IS VEAL
S1	AN APPLE IS A FRUIT	AN ONION IS MEAT

Practice Sentences

Narrow Hierarchies

P0	A LIMOSINE IS LONG	A JEEP CAN ORBIT
P1	A STATION WAGON HAS WHEELS	A CONVERTIBLE HAS WINGS
S0	A CHEVROLET IS A CHEVROLET	A CADILLAC IS A CHARIOT
S1	A CHRYSLER IS A CAR	A SPORTSCAR IS A TRUCK
P0	A CASHEW IS HARD	AN ALMOND IS BLUE
P1	A CHESTNUT HAS A SHELL	A PEANUT IS SOFT
S0	A WALNUT IS A WALNUT	A BRAZIL NUT IS A STRAWBERRY
S1	A PECAN IS A NUT	A HAZEL NUT IS A VINE

TABLE A cont'd.

Broad Hierarchies

P0	A PINE TREE HAS PINE CONES	A BEET IS YELLOW
P1	A WILLOW HAS BARK	A CYPRESS HAS GRAPES
S0	AN ELM IS AN ELM	A PALM TREE IS A GARDENIA
S1	A BIRCH IS A TREE	A TOADSTOOL IS A FLOWER
P0	A LION HAS A MANE	A SPARROW IS LARGE
P1	A CROCODILE IS SCALY	A GOLDFISH HAS FUR
S0	A CROW IS A CROW	A KANGAROO IS A BULLDOG
S1	A CHEETAH IS A CAT	A FROG IS A DOG

TABLE B

Sentences Used in Experiment Two

Experimental Sentences			
Form A		Form B	
T-PO	A SCIENTIST IS SCIENTIFIC	T-SO	A SCIENTIST IS A SCIENTIST
T-P1	A LAWYER IS PROFESSIONAL	T-S1	A LAWYER IS A PROFESSIONAL
T-SO'	A TEACHER IS AN EDUCATOR	T-PO'	A TEACHER CAN EDUCATE
T-S1	A DOCTOR IS A PRACTITIONER	T-P1	A DOCTOR IS PRACTICING
F-P	AN ARCHITECT IS DESTRUCTIVE	F-S	AN ARCHITECT IS A DESTROYER
F-P	A DIPLOMAT IS VULGAR	F-S	A DIPLOMAT IS A VULGARIAN
F-S	AN ENGINEER IS A HYPNOTIST	F-P	AN ENGINEER IS HYPNOTIC
F-S	A MINISTER IS A BRUTE	F-P	A MINISTER IS BRUTISH
T-PO	A LEGISLATOR CAN LEGISLATE	T-SO	A LEGISLATOR IS A LEGISLATOR
T-P1	A SENATOR IS POLITICAL	T-S1	A SENATOR IS A POLITICIAN
T-SO'	A MAGISTRATE IS A JUDGE	T-PO'	A MAGISTRATE CAN JUDGE
T-S1	A PRESIDENT IS A LEADER	T-P1	A PRESIDENT CAN LEAD
F-P	AN ASSEMBLYMAN IS FOREIGN	F-S	AN ASSEMBLYMAN IS A FOREIGNER
F-P	A GOVERNOR IS IDIOTIC	F-S	A GOVERNOR IS AN IDIOT
F-S	A MAYOR IS A PRANKSTER	F-P	A MAYOR IS PRANKISH
F-S	A CONGRESSMAN IS A BEGGAR	F-P	A CONGRESSMAN IS BEGGARLY

TABLE B cont'd.

T-PO	A FARMER CAN FARM	T-SO	A FARMER IS A FARMER
T-P1	A BRICKLAYER CAN WORK	T-S1	A BRICKLAYER IS A WORKER
T-SO'	A MECHANIC IS A REPAIRMAN	T-PO'	A MECHANIC CAN REPAIR
T-S1	A PLUMBER IS A LABORER	T-P1	A PLUMBER CAN LABOR
F-P	A LUMBERJACK IS SCHOLARLY	F-S	A LUMBERJACK IS A SCHOLAR
F-P	A MINER IS WEAK	F-S	A MINER IS A WEAKLING
F-S	A TAILOR IS AN OPPRESSOR	F-P	A TAILOR IS OPPRESSIVE
F-S	A CARPENTER IS A QUIBBLER	F-P	A CARPENTER IS QUIBBLING
T-PO'	AN ACROBAT CAN TUMBLE	T-SO'	AN ACROBAT IS A TUMBLER
T-P1	A MAGICIAN CAN ENTERTAIN	T-S1	A MAGICIAN IS AN ENTERTAINER
T-SO	A SINGER IS A SINGER	T-PO	A SINGER CAN SING
T-S1	A JUGGLER IS A PERFORMER	T-P1	A JUGGLER CAN PERFORM
F-P	A DANCER IS FUMBLING	F-S	A DANCER IS A FUMBLER
F-P	AN ACTOR IS MUTE	F-S	AN ACTOR IS A MUTE
F-S	A COMIC IS A GRUMP	F-P	A COMIC IS GRUMPY
F-S	A COMEDIAN IS A BORE	F-P	A COMEDIAN IS BORING
T-PO'	A FENCER CAN DUEL	T-SO'	A FENCER IS A DUELIST
T-P1	A BOXER CAN CONTEND	T-S1	A BOXER IS A CONTENDER
T-SO	A RUNNER IS A RUNNER	T-PO	A RUNNER CAN RUN
T-S1	A WRESTLER IS AN ATHLETE	T-P1	A WRESTLER IS ATHLETIC
F-P	A SWIMMER IS DECEITFUL	F-S	A SWIMMER IS A DECEIVER
F-P	A SURFER IS OAFISH	F-S	A SURFER IS AN OAF
F-S	A SKIER IS AN IDLER	F-P	A SKIER IS IDLE
F-S	A GOLFER IS A SEAFARER	F-P	A GOLFER IS SEAFARING

TABLE B cont'd.

T-PO'	A TRAITOR CAN BETRAY	T-SO'	A TRAITOR IS A BETRAYER
T-P1	A THIEF IS CROOKED	T-S1	A THIEF IS A CROOK
T-SO	A FORGER IS A FORGER	T-PO	A FORGER CAN FORGE
T-S1	A MURDERER IS A CRIMINAL	T-P1	A MURDERER IS CRIMINAL
F-P	A BLACKMAILER IS HEROIC	F-S	A BLACKMAILER IS A HERO
F-P	A BURGLAR IS PROTECTIVE	F-S	A BURGLAR IS A PROTECTOR
F-S	A SMUGGLER IS A SAINT	F-P	A SMUGGLER IS SAINTLY
F-S	A GANGSTER IS A SACRIFICIER	F-P	A GANGSTER IS SACRIFICING
T-PO	A POET IS POETIC	T-SO	A POET IS A POET
T-P1	A JOURNALIST CAN WRITE	T-S1	A JOURNALIST IS A WRITER
T-SO'	A PLAYWRIGHT IS A DRAMATIST	T-PO'	A PLAYWRIGHT CAN DRAMATIZE
T-S1	AN AUTOBIOGRAPHER IS AN AUTHOR	T-P1	AN AUTOBIOGRAPHER CAN AUTHOR
F-P	A REPORTER IS COWARDLY	F-S	A REPORTER IS A COWARD
F-P	A COLUMNIST IS HALLUCINATORY	F-S	A COLUMNIST IS A HALLUCINATOR
F-S	A CORRESPONDENT IS AN ILLITERATE	F-P	A CORRESPONDENT IS ILLITERATE
F-S	AN ESSAYIST IS A FOOL	F-P	AN ESSAYIST IS FOOLISH
T-PO	A CONDUCTOR CAN CONDUCT	T-SO	A CONDUCTOR IS A CONDUCTOR
T-P1	A GUITARIST IS MUSICAL	T-S1	A GUITARIST IS A MUSICIAN
T-SO'	A BANJOIST IS A STRUMMER	T-PO'	A BANJOIST CAN STRUM
T-S1	A PIANIST IS AN ACCOMPANIEST	T-P1	A PIANIST CAN ACCOMPANY
F-P	A FIDDLER IS AGGRESSIVE	F-S	A FIDDLER IS AN AGGRESSOR
F-P	A FLUTIST IS MISERLY	F-S	A FLUTIST IS A MISER
F-S	A VIOLINIST IS A BEAST	F-P	A VIOLINIST IS BEASTLY
F-S	A DRUMMER IS A PRUDE	F-P	A DRUMMER IS PRUDISH

TABLE B cont'd.

T-PO	A GIRL IS GIRLISH	T-SO	A GIRL IS A GIRL
T-P1	A TEENAGER IS YOUTHFUL	T-S1	A TEENAGER IS A YOUTH
T-SO'	AN INFANT IS A BABY	T-PO'	AN INFANT IS BABYISH
T-S1	A CHILD IS A YOUNGSTER	T-P1	A CHILD IS YOUNG
F-P	A BOY IS ELDERLY	F-S	A BOY IS AN ELDER
F-P	A JUVENILE CAN VOTE	F-S	A JUVENILE IS A VOTER
F-S	AN ADOLESCENT IS A BIGOT	F-P	AN ADOLESCENT IS BIGOTED
F-S	A TODDLER IS AN ADULT	F-P	A TODDLER IS ADULT
T-PO'	A SCULPTOR CAN CARVE	T-SO'	A SCULPTOR IS A CARVER
T-P1	A CARTOONIST CAN ILLUSTRATE	T-S1	A CARTOONIST IS AN ILLUSTRATOR
T-SO	AN ENGRAVER IS AN ENGRAVER	T-PO	AN ENGRAVER CAN ENGRAVE
T-S1	A PAINTER IS AN ARTIST	T-P1	A PAINTER IS ARTISTIC
F-P	A DECORATOR IS SINFUL	F-S	A DECORATOR IS A SINNER
F-P	A DESIGNER IS BLUNDERING	F-S	A DESIGNER IS A BLUNDERER
F-S	A PHOTOGRAPHER IS A CORRUPTOR	F-P	A PHOTOGRAPHER IS CORRUPTIVE
F-S	A SKETCHER IS A TREMBLER	F-P	A SKETCHER IS TREMBLING
T-PO'	A KNIGHT CAN CRUSADE	T-SO'	A KNIGHT IS A CRUSADER
T-P1	A PRINCE IS NOBLE	T-S1	A PRINCE IS A NOBLE
T-SO	A KING IS A KING	T-PO	A KING IS KINGLY
T-S1	A COUNT IS AN ARISTOCRAT	T-P1	A COUNT IS ARISTOCRATIC
F-P	A PRINCESS IS SLANDEROUS	F-S	A PRINCESS IS A SLANDERER
F-P	A DUKE IS SAVAGE	F-S	A DUKE IS A SAVAGE
F-S	A QUEEN IS A BARBARIAN	F-P	A QUEEN IS BARBARIC
F-S	A BARON IS A PROLETARIAN	F-P	A BARON IS PROLETARIAN

TABLE B cont'd.

T-PO'	A STOCKHOLDER CAN INVEST	T-SO'	A STOCKHOLDER IS AN INVESTOR
T-P1	A MANUFACTURER IS BUSINESSLIKE	T-S1	A MANUFACTURER IS A BUSINESSMAN
T-SO	AN IMPORTER IS AN IMPORTER	T-PO	AN IMPORTER CAN IMPORT
T-S1	AN INDUSTRIALIST IS AN ENTERPRISER	T-P1	AN INDUSTRIALIST IS ENTERPRISING
F-P	A MERCHANT IS TERRIFYING	F-S	A MERCHANT IS A TERROR
F-P	A BANKER IS GIGGLING	F-S	A BANKER IS A GIGGLER
F-S	AN ADVERTISER IS A REBEL	F-P	AN ADVERTISER IS REBELLIOUS
F-S	A WHOLESALER IS A SENTIMENTALIST	F-P	A WHOLESALER IS SENTIMENTAL
Practice Sentences			
T-PO	AN IMP IS IMPISH	T-SO	AN IMP IS AN IMP
T-P1	AN ANGEL IS SPIRITUAL	T-S1	AN ANGEL IS A SPIRIT
T-SO'	A GHOST IS A SPOOK	T-PO'	A GHOST IS SPOOKY
T-S1	A PHANTOM IS AN APPARITION	T-P1	A PHANTOM IS APPARITIONAL
F-P	A DEVIL IS MORAL	F-S	A DEVIL IS A MORALIST
F-P	A MERMAID CAN FLY	F-S	A MERMAID IS A FLIER
F-S	A PIXIE IS A GIANT	F-P	A PIXIE IS GIGANTIC
F-S	A GOBLIN IS A PATRIOT	F-P	A GOBLIN IS PATRIOTIC
T-PO'	A CHAUFFEUR CAN DRIVE	T-SO'	A CHAUFFEUR IS A DRIVER
T-P1	A MAID IS DOMESTIC	T-S1	A MAID IS A DOMESTIC
T-SO	A COOK IS A COOK	T-PO	A COOK CAN COOK
T-S1	A BUTLER IS A SERVANT	T-P1	A BUTLER CAN SERVE
F-P	A GARDENER IS VILLAINOUS	F-S	A GARDENER IS A VILLAIN
F-P	A HOUSEKEEPER IS INSPIRATIONAL	F-S	A HOUSEKEEPER IS AN INSPIRATION
F-S	A DOORMAN IS A MATRON	F-P	A DOORMAN IS MATRONLY
F-S	A GOVERNESS IS AN IMMORTAL	F-P	A GOVERNESS IS AN IMMORTAL