The Relative Accessibility of Semantic and Deep Structure Syntactic Concepts

Alice F. Healy+ and Andrea G. Levitt++

ABSTRACT

Three experiments were conducted to determine the relative accessibility of semantic and deep structure syntactic concepts. In Experiment I, which employed a concept formation task, subjects learned the concept "deep structure subject" more slowly than the case concept "experiencer." In Experiments II and III, which employed a new recognition memory procedure, subjects performed more poorly when the sentences to be remembered were differentiated on the basis of deep structure syntactic relations than when they were differentiated on the basis of semantic relations. These results favor Fillmore's case grammar, or another semantically-based theory, rather than the "standard theory" of Chomsky in a model of linguistic behavior.

INTRODUCTION

A number of different versions of transformational generative grammar have been proposed in recent years, including, among others, "generative semantics" (McCawley, 1968; Lakoff, 1971), case grammar (Fillmore, 1968, 1970,

*Portions of this paper concerning Experiments I and II were read at the 18th Annual Meeting of the Psychonomic Society, Washington, D. C., November, 1977. This paper also appears in Memory & Cognition 6, 518-526.

+Also Yale University.

++Also Wellesley College.

Acknowledgment: This research was supported in part by PHS Grants MH26573 and RR07015, NSF Grant BNS77-00077 to Yale University and NICHD Grant HD01994 and BRSG Grant RR05596 to Haskins Laboratories. The senior author was supported by a Junior Faculty Fellowship from Yale University during the preparation of this manuscript. The authors are deeply indebted to Guy Carden for many informative discussions, Robert G. Crowder for a careful reading of an earlier version of this manuscript, George A. Miller for inspiring us to perform Experiment III and for other valuable suggestions, Jonathan Livingston for help with the design of Experiment I and help with the construction of the sentences used in Experiments I and II, Wendy Brown for help with the conduct and analyses of Experiments I and II, Oren Pataaschnik for help with the conduct and analyses of Experiment I, and Loretta Polka for help with the conduct and analyses of Experiment III.

Surface Structure Representation:
"The doctor gave the books to John."

The Standard Theory

```
S
  NP
  |   VP
  |   |
  DET N  V
  |
the doctor gave
  |   NP
  |   PREP PHRASE
  |   |
  |   DET N
  |   |
  |  the books to
  |
  |  John
```

Case Grammar

```
S
  NP
  |   Proposition
  |   |
  DET N  V
  |
the doctor gave
  |   NP
  |   Goal
  |   |
  |   DET N
  |   |
K
  |
NP
```

Figure 1: Surface structure representation of sentence "The doctor gave the books to John" according to the standard theory (top panel) and case grammar (bottom panel). Case grammar representation is based on Fillmore (1968).
Deep Structure Representation:
"The doctor gave the books to John."

The Standard Theory

Case Grammar

Figure 2: Deep structure representation of sentence "The doctor gave the books to John" according to the standard theory (top panel) and case grammar (bottom panel). Case grammar representation is based on Fillmore (1968).
1971, 1977), and the "standard theory" of Chomsky (1965). One aspect of the standard theory that is attacked by the proponents of both case grammar and generative semantics is syntactic deep structure as a level of linguistic description. Syntactic deep structure plays a prominent role in the standard theory. In fact, the deep syntactic level of representation is central. The deep syntactic representation is mapped into a semantic representation, on one hand, and into a surface syntactic representation, on the other hand. A system of semantic projection rules (Katz and Fodor, 1963) is posited to link the deep syntactic and semantic levels, and a system of syntactic transformations is posited to link the deep syntactic and surface syntactic levels. In contrast, according to generative semantics, the level of syntactic deep structure is not necessary. Rather, the semantic representation is mapped directly into a surface syntactic representation. A single system of transformational rules is envisioned to link the semantic and surface syntactic representations. Similarly, a purely syntactic level of deep structure is not included in case grammar and, in fact, was deemed "an artificial intermediate level" by Fillmore (1968, p. 88). According to case grammar, case relations, which are semantic as well as syntactic, replace the purely syntactic deep structure relations, such as deep structure (logical) subject.\(^1\) Although the standard theory and case grammar posit essentially identical surface structure representations (see Figure 1), the deep structure representations differ substantially (see Figure 2).\(^2\) Note in particular that whereas the relation deep structure subject can be defined in terms of the deep structure of the standard theory (technically, the subject is the noun phrase immediately dominated by the node labeled S), the relation deep structure subject cannot be simply defined in terms of the deep structure of case grammar. Rather, in case grammar the subject of the sentence is treated in an entirely parallel manner to the other cases in the deep structure, and the relation subject is "seen as exclusively a surface-structure phenomenon" (Fillmore, 1968, p. 17). Only as a result of subject selection and transformational rules is a subject created and placed in its proper location in the surface structure of the sentence.

The present study does not attempt to resolve the linguistic issue concerning the existence of a level of deep structure syntax. Rather, the aim of the present study is to provide a test of the relative psychological accessibility of the deep structure syntactic relations included in the standard theory. Deep structure syntactic relations are compared to case relations. Early experiments by Blumenthal (1967) and Blumenthal and Boakes (1967) used a cued recall technique to demonstrate the salience of the deep structure subject of the sentence. The deep structure subject was the best cue to recall a given sentence. However, although deep structure syntactic and surface structure syntactic relations were unconfounded in these experiments, deep structure syntactic and semantic relations were left

\(^1\)In a more recent paper, Fillmore (1977) has recognized the need for purely syntactic deep structure relations as well as case relations.

\(^2\)Although Fillmore (1968) originally represented deep structures in terms of tree structures, as in Figure 2, more recently Fillmore (1971) has announced a preference for a different type of notation.
confounded. (In other words, cue words that differed in their deep structure syntactic categories also differed in their semantic roles.) In contrast, in a more recent study also employing cued recall of sentences, Perfetti (1973) unconfounded deep structure syntactic and semantic relations but left confounded deep structure syntactic and surface structure syntactic relations. (In other words, cue words that differed in their deep structure syntactic categories also differed in their surface structure syntactic categories.) The present study successfully unconfounds for the first time all three types of relations—surface structure syntactic, deep structure syntactic, and semantic.

Instead of employing the cued recall technique used in the studies reviewed above, in the present study we use a concept formation task, which enables us to assess the extent to which subjects are able to learn by example various syntactic and semantic concepts. This technique was employed by Baker, Prideaux and Derwing (1973) to study surface structure syntactic concepts and by Shafto (1973) to study the semantic concepts of case grammar. Although Shafto successfully studied the ease of learning various case relations (and found "agent" easiest, followed by "experiencer," followed by "instrument" and "object"), he did not compare case concepts to any other linguistic concepts. In Experiment I we compare the ease of learning a case relation and a deep structure syntactic relation. Two baseline conditions are also included in this experiment, one to provide information about the upper limit of performance and the other to provide information about the lower limit of performance. In the first baseline condition subjects learn a surface structure syntactic relation, expected to be relatively trivial, and in the second baseline condition, subjects learn an arbitrary relation, defined in a manner analogous to that of the other three concepts.

It should be noted that although case relations are specifically manipulated in this investigation, this study will not allow us to discriminate among various semantically-based grammars, since differences in case relations are necessarily confounded with differences in other semantic variables. Likewise, this investigation will not allow us to discriminate among different models of sentence memory and comprehension that assert structures of a semantic variety (for example, Rumelhart, Lindsay and Norman, 1972; Schank, 1972; Anderson and Bower, 1973; Kintsch, 1974). However, this study will enable us to discriminate between such models based primarily on semantic relations and any plausible alternative models based primarily on deep structure syntactic relations.

**EXPERIMENT I**

This experiment employed lists of simple sentences, each of which included the word John in one of two semantic roles ("experiencer" or "goal")\(^3\), one of two deep structure syntactic categories (deep structure

---

\(^3\)We employed the definitions for these cases given by Fillmore (1971) and restricted ourselves to the benefactive meaning of the goal case: "Where there is a genuine psychological event or mental state verb, we have the Experiencer;...where there is a transfer or movement of something to a person, the receiver as destination is taken as the Goal." (p. 42)
<table>
<thead>
<tr>
<th>Surface Structure Role</th>
<th>Subject</th>
<th>Object of Preposition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiencer</strong></td>
<td>1. John was sleepy near the fire.</td>
<td>7. John was assured misery.</td>
</tr>
<tr>
<td><strong>Goal</strong></td>
<td>4. John was the recipient of the grant.</td>
<td>5. John was given the book.</td>
</tr>
</tbody>
</table>

**Object of Preposition**

<table>
<thead>
<tr>
<th>Experiencer</th>
<th>2. The accident was imagined by John.</th>
<th>3. The roar was deafening to John.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal</td>
<td>8. The fruit was obtained by John.</td>
<td>6. The property was leased to John.</td>
</tr>
</tbody>
</table>
subject or deep structure object of the preposition), and one of two surface structure syntactic categories (surface structure subject or surface structure object of the preposition). Although there are eight possible combinations of these three kinds of relations, only six of them (types 1-6) were employed here (see Table 1). Two combinations (types 7 and 8)—deep structure object of preposition/surface structure subject/experiencer, and deep structure subject/surface structure object of preposition/goal—were not employed since fully satisfactory examples of these types could not be generated (see the introduction to Experiment III below for a more complete discussion of this problem). Since the two missing combinations included one of each of the deep structure syntactic concepts, one of each of the semantic concepts, and one of each of the surface structure syntactic concepts, their exclusion should not bias the learning of any of these concepts. For each of the concepts learned by the subjects, three of the sentence types were instances of the given concept and three were not. For the semantic concept, the three positive types were those with experiencer (types 1, 2, 3, Table 1) and the three negative types were those with goal (4, 5, 6); for the deep structure syntactic concept, the three positive types were those with deep structure subject (1, 2, 4) and the three negative types were those with deep structure object of the preposition (3, 5, 6); for the surface structure syntactic concept, the three positive types were those with surface structure subject (1, 4, 5) and the three negative types were those with surface structure object of the preposition (2, 3, 6); for the arbitrary concept the three positive types had no regular relationship to each other (2, 4, 6) and the three negative types also were not related in any regular way (1, 3, 5).

Method

Subjects. Forty young men and women, who were recruited by posters on the Yale University campus, participated as subjects and were paid $1 for their participation, which lasted approximately 20 minutes. No subject had any formal training in linguistics. There were four conditions—Arbitrary, Deep Structure Syntactic, Semantic and Surface—with ten subjects in each condition. The assignment of subjects to conditions was determined by time of arrival for testing according to a fixed rotation of conditions.

Materials. Sixty sentences, ten of each of the six sentence types, were employed as stimuli. These sentences are shown in the Appendix. Note that for all the sentences where John was the surface subject, John was the first word in the sentence, and for all sentences where John was the surface object of the preposition, John was the last word in the sentence. Hence surface location of the word John was perfectly confounded with its surface structure category but unconfounded with its deep structure category and semantic role. Each of the sixty sentences was typed in the center of four 4 X 6 cards. Four decks of cards were constructed, one for each of the four conditions. Each deck included all sixty sentences; only the order of the sentences varied across decks. In each deck the order of the sentences was pseudo-random with the constraint that each 12-sentence block included two sentences of each type. The sentences within a given one of the five blocks were the same in the four decks, but the order of sentences within a block differed across decks. The order of sentences in a given block in the Semantic condition was the same as that in the Deep Structure Syntactic condition, except for the placement of four of the sentences, including two of each of two types—
sentences that were positive instances of the semantic concept but negative instances of the deep structure syntactic concept (type 3), and sentences that were negative instances of the semantic concept but positive instances of the deep structure syntactic concept (type 4). The two sentences of the first type in the deep structure syntactic deck were replaced by the two sentences of the second type, and vice versa, to form the semantic deck. However, the two sentences of a given type maintained their position relative to each other. Similarly, the order of sentences in a given block in the Surface condition was the same as that in the Deep Structure Syntactic condition except for the placement of four of the sentences, including two of each of two types—sentences that were positive instances of the surface concept but negative instances of the deep structure syntactic concept (type 5) and sentences that were negative instances of the surface concept but positive instances of the deep structure syntactic concept (type 2). The two sentences of the first type in the deep structure syntactic deck were replaced by the two sentences of the second type, and vice versa, to form the surface deck. The deck for the Arbitrary condition was analogously related to that of the Deep Structure Syntactic condition. These relationships among the four decks insured that the sequences of correct responses (positive and negative instances of the given concept) were the same in all four conditions.

Procedure. Each subject was tested individually on one of the four concepts. The experimenter, who sat across a table from the subject, showed the subject all the cards from the appropriate deck, one at a time in the prescribed order. The subject was allowed to view only the sentence currently under test at any given instant. The subject was to respond orally "yes" or "no" to each sentence depending on whether he thought it was a positive or negative instance of the concept. After the subject responded, the experimenter supplied immediate oral feedback by telling him whether he was correct and what the correct answer was. The experimenter recorded the subject's responses on an answer sheet, with an indication of whether a given response was an error. Sentence presentation was experimenter paced, dependent on the subject's speed of responding. The following instructions were read to the subject at the start of the experiment:

"You will be presented with a series of cards on each of which is printed a simple sentence with the word John in it. The word John has some kind of relation to the other words in the sentence. Your task is to determine what the relationship is.

When you see each card you are to judge whether the sentence on it illustrates the test relation between John and the other words in the sentence. Say 'yes' if you think that it does or 'no' if you think that it does not. The experimenter will tell you if your answer is right or wrong. You may not look back at cards that you have already seen."

Results and Discussion

The results are summarized in Table 2 in terms of mean percentages of errors per 12-sentence block as a function of block position and condition. As expected, subjects performed best in the Surface condition, where the concept to be learned was assumed to be trivial, and worst in the Arbitrary
### TABLE 2: Mean percentage of errors in Experiment I as a function of Condition and Block.

<table>
<thead>
<tr>
<th>Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface</td>
<td>18</td>
<td>9</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Semantic</td>
<td>42</td>
<td>27</td>
<td>22</td>
<td>13</td>
<td>9</td>
<td>23</td>
</tr>
<tr>
<td>Deep Structure</td>
<td>33</td>
<td>38</td>
<td>28</td>
<td>24</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Syntactic</td>
<td>49</td>
<td>43</td>
<td>44</td>
<td>35</td>
<td>34</td>
<td>41</td>
</tr>
<tr>
<td>Arbitrary</td>
<td>49</td>
<td>43</td>
<td>44</td>
<td>35</td>
<td>34</td>
<td>41</td>
</tr>
</tbody>
</table>

### TABLE 3: Mean percentage of errors in Experiment I as a function of Condition and Sentence Type.

<table>
<thead>
<tr>
<th>Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface</td>
<td>8</td>
<td>10</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Semantic</td>
<td>19</td>
<td>29</td>
<td>26</td>
<td>32</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Deep Structure</td>
<td>11</td>
<td>44</td>
<td>42</td>
<td>30</td>
<td>31</td>
<td>20</td>
</tr>
<tr>
<td>Syntactic</td>
<td>26</td>
<td>48</td>
<td>56</td>
<td>22</td>
<td>68</td>
<td>27</td>
</tr>
<tr>
<td>Arbitrary</td>
<td>26</td>
<td>48</td>
<td>56</td>
<td>22</td>
<td>68</td>
<td>27</td>
</tr>
</tbody>
</table>
condition, where the concept to be learned was not meaningful (except as a
disjunction of sentence types). In addition, performance was better in the
Semantic condition than in the Deep Structure Syntactic condition, but this
difference was not statistically significant. Two analyses of variance were
performed on these data, one with subjects ($F_1$) and one with sentences ($F_2$) as
the random effect. The statistic $\min F'$ (Clark, 1973) was computed on the
basis of these analyses. These analyses yielded a significant main effect of
condition, $\min F' (3,46) = 9.8, p < .001 \ [F_1 (3,36) = 11.1, MS_e = 5190, p < .001; F_2 (3,90) = 82.4, MS_e = 140, p < .001]$. Newman-Keuls tests, based on
the analysis with subjects as the random effect, revealed significant differen-
tces between the Deep Structure Syntactic condition and the Surface condition
and between the Arbitrary condition and the Semantic and Surface conditions,
all at the .01 level, as well as a significant difference between the Semantic
and Surface conditions at the .05 level. No other differences among condi-
tions were statistically significant. In particular, these analyses did not
allow us to distinguish between the critical Semantic and Deep Structure
Syntactic conditions or between the Deep Structure Syntactic and Arbitrary
conditions.

Learning was evident across the five 12-sentence blocks; the main effect
of blocks was significant, $\min F' (4,84) = 6.7, p < .001 \ [F_1 (4,144) = 15.5,
MS_e = 818, p < .001; F_2 (4,30) = 11.9, MS_e = 212, p < .001]$. Furthermore,
there was more learning evident across the five blocks in the Semantic
condition than in the Deep Structure Syntactic condition. Although the
overall analyses of variance did not reveal a significant interaction of
condition by blocks, $\min F' (12,128) < 1 \ [F_1 (12,144) = 1.6, MS_e = 818, p = .105; F_2 (12,90) = 1.8, MS_e = 140, p = .054]$, planned analyses of variance
with only the critical Semantic and Deep Structure Syntactic conditions did
reveal a significant interaction of condition by blocks in both the test with
subjects as the random effect, $F_1 (4,72) = 3.9, MS_e = 715, p = .006$, and the
test with sentences as the random effect, $F_2 (4,30) = 2.9, MS_e = 192, p = .037$, but not in the more conservative test combining them, $\min F' (4,74) = 1.7, p = .163$. In addition, the planned analysis for the two critical
conditions yielded a significant main effect of condition with sentences as
the random effect, $F_2 (1,30) = 8.0, MS_e = 192, p = .008$.

The analyses of variance further revealed a significant effect of
sentence type, $\min F' (5,131) = 4.6, p < .001 \ [F_1 (5,180) = 7.9, MS_e = 1500, p < .001; F_2 (5,30) = 11.1, MS_e = 212, p < .001]$, and a significant interaction
of condition by sentence type, $\min F' (15,270) = 2.6, p = .001 \ [F_1 (15,180) = 3.8, MS_e = 1500, p < .001; F_2 (15,90) = 8.1, MS_e = 140, p < .001]$. Table 3
presents the mean percentages of errors as a function of condition and
sentence type. Clearly certain sentence types caused more trouble for
learning some concepts than others, but the nature of the interaction of
condition and sentence type did not appear to be completely comprehensible
and, as will be shown below, was not entirely consistent across experiments.

These results suggest that semantic case relations are indeed learned
more rapidly than deep structure syntactic relations. Since the largest
difference between the Semantic and Deep Structure Syntactic conditions was at
the last block of training, where learning was greatest, a more sensitive test
comparing these two conditions seems to be one where all testing is conducted
after training has been completed. For that reason a new recognition memory
paradigm was devised for Experiment II to compare the learning of the semantic and deep structure syntactic relations with all testing conducted after the completion of training. This paradigm, like the concept formation task, was designed to test the psychological accessibility of various linguistic concepts. Whereas the concept formation task allowed us to determine whether subjects could learn the given concepts, the recognition memory task allows us to determine whether the given concepts will be discovered by subjects in their attempts to learn a list of sentences for a subsequent memory test.

**EXPERIMENT II**

The same relations were tested in this experiment as in Experiment I. Furthermore, the same sentences were employed in the two experiments with the following important exception: There were two versions of each of the sixty sentences, with one version containing the word John, as earlier, and one version containing the word Sam instead of John. Every subject was shown each of the sixty sentences, half of which were in the version with John and half in the version with Sam. As in Experiment I, there were four groups of subjects, the groups in this case differing in the rule used to assign John or Sam to each sentence. The assignment was made on the basis of either the deep structure syntactic category of the word John or Sam, the semantic role, the surface structure syntactic category, or the arbitrarily defined rule employed in Experiment I. The subjects' task in this experiment was first to study the given sentences and later, on a subsequent recognition memory test, to decide whether a given sentence which had been studied earlier included John or Sam. In contrast to Experiment I, subjects were not specifically told about the existence of a consistent relation between John (or Sam) and the rest of the words in each sentence. Therefore this experiment allowed us to determine how readily the given relations were discovered by subjects in the course of memorizing a list of sentences, rather than whether the subjects could learn the given relations when required to do so.

**Method**

**Subjects.** Forty young men and women, who were recruited by posters on the Yale University campus, participated as subjects and were paid $1 for their participation, which lasted approximately 20 minutes. There were four conditions--Arbitrary, Deep Structure Syntactic, Semantic, and Surface--with ten subjects in each condition. For each condition, there were two subgroups of subjects (A and B) with five subjects in each subgroup. The assignment of subjects to conditions and subgroups was determined by time of arrival for testing according to a fixed rotation of conditions and subgroups.

**Materials.** Eight decks of cards, each card containing one sentence, were constructed for training, two decks for each of the four groups of subjects. For a given group, the sentences in each deck were identical to those employed in Experiment I except that in one deck (Deck A) all the sentences that were positive instances of the concept tested in Experiment I included the word John and all the sentences that were negative instances of the concept included the word Sam, and in the other deck (Deck B) the opposite assignment of John and Sam was employed. In each condition, one subgroup of subjects (A) was given Deck A and the other subgroup (B) was given Deck B. This method of counterbalancing assured that across subjects the words John and Sam would not
be confounded with the positive and negative instances of the concept. The order of the sentences in a given deck varied across subjects and was determined by the experimenter's thoroughly shuffling the deck of cards before handing it to the given subject.

Four typewritten lists of sentences were constructed for the recognition memory test, one list for each of the four conditions. The list for a given condition included the same sentences as in Experiment I, in the same order. The only differences between the form of the sentences as they appeared on the cards in Experiment I and as they appeared on the test lists in Experiment II were that the sentences were numbered (from 1-60) in Experiment II but not in Experiment I, and the word John in each sentence in Experiment I was replaced by the pair of words John/Sam in Experiment II. As a result of these constraints, the order of correct answers (John or Sam) was the same for subjects in all four subgroups A, and was the same for subjects in all four subgroups B, but the correct answers for subjects in subgroups A were the reverse of those for subjects in subgroups B.

Procedure. Each subject was tested individually with one of the eight study decks of cards. The first eight subjects run, one in each subgroup, were given eight minutes to study the deck of sentences (timed by the experimenter with a stopwatch). The remaining 32 subjects were given five minutes to study the deck of sentences, since the performance of the first subjects seemed to approach the ceiling. Subjects were in no way restricted in their method of studying the sentences. They were allowed to sort the sentences into piles, and they were allowed to look at a given sentence any number of times. The subjects were not encouraged to use any particular strategy in studying the sentences; however, they were given written instructions describing exactly what their task would be during the recognition memory test: "You will be presented with a stack of cards. On each card is a sentence which involves either John or Sam. You are to study these sentences for five (eight) minutes. At the end of that time you will be given two sheets of paper which include each of the sentences on the cards with the words John and Sam replaced by John/Sam. Your task will be to recall for each sentence whether John or Sam was involved in that sentence as it appeared on the card. You are to indicate your response by circling one of the two words John or Sam in the given sentence on the sheet of paper." After studying the sentences on the cards, subjects were reminded of their task on the recognition memory test. Subjects were then given the appropriate test list of sentences and responded by circling the word John or Sam in each sentence, depending on which word they thought occurred in the sentence when it appeared on the card. Subjects were required to respond to every test sentence; they were not allowed to leave blanks. Subjects were given as much time as they needed to complete the recognition memory test.

Results and Discussion

The results are summarized in Table 4 in terms of mean percentages of errors on the recognition test as a function of condition and sentence type. The data were averaged over subgroups (A and B) since that factor was not found to be significant. The difference between the Deep Structure Syntactic and Semantic conditions in this experiment was striking. Performance on the deep structure syntactic relations was considerably worse than on the semantic
and surface relations and, in fact, somewhat worse than on the arbitrary relations. Two analyses of variance were performed on these data, one with subjects (F₁) and one with sentences (F₂) as the random effect. According to these analyses, there was a significant effect of condition, min F' (3,46) = 5.5, p = .003 [F₁ (3,36) = 6.2, MSₑ = 598, p = .002; F₂ (3,162) = 48.5, MSₑ = 77, p < .001], an effect of sentence type which approached significance, min F' (5,205) = 2.1, p = .069 [F₁ (5,180) = 3.4, MSₑ = 174, p = .006; F₂ (5,54) = 5.3, MSₑ = 112, p < .001], as well as a significant interaction of these two factors, min F' (15,307) = 1.8, p = .029 [F₁ (15,180) = 2.7, MSₑ = 174, p = .001; F₂ (15,162) = 6.0, MSₑ = 77, p < .001]. As in Experiment I, certain sentence types caused more trouble for some conditions than for others. However, the nature of the interaction of condition and sentence type was somewhat different from that found in Experiment I.

Separate planned analyses of variance were conducted with the data from just the two critical conditions (Deep Structure Syntactic and Semantic). These analyses also revealed a significant effect of condition, min F' (1,21) = 6.8, p = .016 [F₁ (1,18) = 7.4, MSₑ = 724, p = .014; F₂ (1,54) = 87.8, MSₑ = 61, p < .001], but the effect of sentence type, min F' (5,144) = 1.2, p = .337 [F₁ (5,90) = 1.8, MSₑ = 158, p = .129; F₂ (5,54) = 3.3, MSₑ = 84, p = .011], and the interaction of condition and sentence type, min F' (5,138) = 1.8, p = .113 [F₁ (5,90) = 2.5, MSₑ = 158, p = .035; F₂ (5,54) = 6.5, MSₑ = 61, p < .001], were not significant. Furthermore, Newman-Keuls tests, based on the overall analysis conducted with subjects as the random effect, revealed significant differences (p < .05) between the Deep Structure Syntactic and Semantic conditions, between the Arbitrary and Semantic conditions, and between the Arbitrary and Surface conditions, a significant difference (p < .01) between the Deep Structure Syntactic and Surface conditions, and nonsignificant differences between the Semantic and Surface conditions and between the Deep Structure Syntactic and Arbitrary conditions. It is clear from these data that subjects easily discovered the semantic case relations discriminating the Sam and John sentences when studying the sentences for a subsequent recognition memory test, and the subjects were not able to discover the deep structure syntactic relations so easily. These latter concepts were discovered no more easily than purely arbitrarily defined relations.

**EXPERIMENT III**

Only six sentence types were employed in Experiments I and II although there are eight possible combinations of the three kinds of relations. The two missing sentence types (7 and 8) had been excluded because it was difficult to find satisfactory examples of them. However, examples of these sentence types do exist and one of each of these two types is shown in Table 1. The problem with sentences of these types is that the case role of John seems to be ambiguous. Specifically, in type 7 sentences John seems to be in both the roles of experiencer and goal, although the experiencer role does seem more salient. In some type 8 sentences, it is not entirely clear whether John is in the role of goal or agent, a problem which exists for several sentences of other types as well. Despite these difficulties, all eight sentence types were employed in Experiment III, which was otherwise a replication of Experiment II. (Type 7 was selected to be a positive instance of the arbitrary concept, and type 8 was selected to be a negative instance of the arbitrary concept.) This experiment enabled us to test our contention
### TABLE 4: Mean percentage of errors in Experiment II as a function of Condition and Sentence Type.

<table>
<thead>
<tr>
<th>Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>2.7</td>
</tr>
<tr>
<td>Semantic</td>
<td>5</td>
<td>4</td>
<td>9</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>4.8</td>
</tr>
<tr>
<td>Deep Structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic</td>
<td>6</td>
<td>28</td>
<td>14</td>
<td>19</td>
<td>25</td>
<td>17</td>
<td>18.2</td>
</tr>
<tr>
<td>Arbitrary</td>
<td>4</td>
<td>17</td>
<td>7</td>
<td>11</td>
<td>26</td>
<td>33</td>
<td>16.3</td>
</tr>
</tbody>
</table>

### TABLE 5: Mean percentage of errors in Experiment III as a function of Condition and Sentence Type.

<table>
<thead>
<tr>
<th>Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface</td>
<td>3.75</td>
<td>5.00</td>
<td>1.25</td>
<td>2.50</td>
<td>1.25</td>
<td>3.75</td>
<td>2.50</td>
<td>6.25</td>
<td>3.28</td>
</tr>
<tr>
<td>Semantic</td>
<td>5.00</td>
<td>7.50</td>
<td>7.50</td>
<td>7.50</td>
<td>6.25</td>
<td>10.00</td>
<td>6.25</td>
<td>11.25</td>
<td>7.66</td>
</tr>
<tr>
<td>Deep Structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic</td>
<td>21.25</td>
<td>26.25</td>
<td>20.00</td>
<td>30.00</td>
<td>23.75</td>
<td>22.50</td>
<td>16.25</td>
<td>28.75</td>
<td>23.59</td>
</tr>
<tr>
<td>Arbitrary</td>
<td>13.75</td>
<td>31.25</td>
<td>37.50</td>
<td>30.00</td>
<td>31.25</td>
<td>37.50</td>
<td>15.00</td>
<td>30.00</td>
<td>28.28</td>
</tr>
</tbody>
</table>
that the previous results were not due to any bias created by employing only six sentence types.

Method

Subjects. Forty male and female undergraduate students of Yale College who were taking a course in introductory psychology participated as subjects, receiving course credit. As in Experiment II, there were eight subgroups of subjects with five subjects in each subgroup. The assignment of subjects to subgroups was determined by time of arrival for testing according to a fixed rotation of subgroups.

Materials. Eight decks of cards were constructed for training, one deck for each of the eight subgroups of subjects. The decks were constructed in a manner strictly analogous to that employed for Experiment II except that there were 64, rather than 60, sentences in each deck, including 8, rather than 10, of each sentence type. The sentences employed are shown in the Appendix. The four sentences of each of the original six types included in Experiment II but not in Experiment III have an asterisk beside them in the Appendix.

Four typewritten lists of sentences were constructed for the recognition memory test in an analogous manner to the lists constructed for Experiment II. On each list the order of the sentences was pseudo-random, with the constraint that each 16-sentence block included two sentences of each type. The sentences within a given one of the four blocks were the same on the four lists, but the order of the sentences within a block differed across lists. In particular, the order of the sentences in a given block in the Semantic condition was the same as that in the Deep Structure Syntactic condition, except for the placement of eight of the sentences, including two of each of four types—sentences that were positive instances of the semantic concept but negative instances of the deep structure syntactic concept (types 3 and 7), and sentences that were negative instances of the semantic concept but positive instances of the deep structure syntactic concept (types 4 and 8). The two sentences of type 3 in the deep structure syntactic deck were replaced by the two sentences of type 4, and the two sentences of type 7 in the deep structure syntactic deck were replaced by the two sentences of type 8, and vice versa, to form the semantic deck. However, the two sentences of a given type maintained their position relative to each other. The lists for the Surface and Arbitrary conditions were analogously related to the list for the Deep Structure Syntactic condition. As in the earlier experiments, the relationships among the four lists insured that the sequence of correct answers (John or Sam) was the same for subjects in all four subgroups A, and was the same for subjects in all four subgroups B.

Procedure. The procedure was the same as in Experiment II except that all subjects were given five minutes to study the deck of sentences.

Results and Discussion

The results are summarized in Table 5 in terms of mean percentages of errors on the recognition test as a function of condition and sentence type. Despite the difference in the sentence types included, the pattern of results was strikingly similar to that found in Experiment II. Two analyses of
variance were performed on these data, one with subjects (F₁) and one with sentences (F₂) as the random effect. These analyses yielded a significant main effect of condition, min F' (3, 46) = 10.1, p < .001 [F₁ (3, 36) = 11.4, MSE = 1024, p < .001; F₂ (3, 168) = 84.6, MSE = 111, p < .001], but neither the main effect of sentence type, min F' (7, 149) = 1.1, p = .373 [F₁ (7, 252) = 2.6, MSE = 179, p = .013; F₂ (7, 56) = 1.9, MSE = 198, p = .090], nor the interaction of condition and sentence type, min F' (21, 418) < 1 [F₁ (21, 252) = 1.2, MSE = 179, p = .238; F₂ (21, 168) = 1.6, MSE = 111, p = .063], was significant.

Separate analyses of variance were conducted with the data from just the Semantic and Deep Structure Syntactic conditions. These analyses also yielded a significant effect of condition, min F' (1, 21) = 7.1, p = .015 [F₁ (1, 18) = 7.6, MSE = 1337, p = .013; F₂ (1, 56) = 98.7, MSE = 82, p < .001]. In addition, Newman-Keuls tests, based on the overall analysis conducted with subjects as the random effect, revealed significant differences (p < .01) between the Deep Structure Syntactic and Semantic conditions, between the Arbitrary and Semantic conditions, between the Arbitrary and Surface conditions, and between the Deep Structure Syntactic and Surface conditions; but the differences between the Deep Structure Syntactic and Arbitrary conditions and between the Semantic and Surface conditions were not significant by these tests. The conclusions reached on the basis of Experiment II are clearly supported by this pattern of results.

**SUMMARY AND CONCLUSIONS**

These experiments indicate that the deep structure syntactic relations we studied were both learned more slowly and discovered less readily than the semantic case concepts we studied. These results suggest that the deep structure syntactic relations included in the standard theory are less accessible than semantic relations and, in fact, are no more accessible than arbitrarily defined relations. An implication of these findings is that the cued recall studies purporting to demonstrate the importance to sentence memory of the relation deep structure subject (Blumenthal, 1967; Blumenthal and Boakes, 1967) were misleading because of their confounding deep structure syntactic and semantic relations. The present study suggests that deep structure syntactic relations, when unconfounded from semantic relations, do not play a major role in sentence memory.

More generally, these results favor case grammar, rather than the standard theory, in a model of linguistic behavior. It should be noted, however, that although these results are clearly difficult for the standard theory, they do not discriminate among different semantically-based grammars.

Our rejection of the standard theory in a model of linguistic behavior may not seem very consequential for two reasons. First, there have been a proliferation of proposed revisions of the standard theory; however, some influential theorists have recently argued (see Bever, Katz and Langendoen, 1976) that the dismissal of the standard theory may have been too rash. It is also relevant to note (see footnote 1) that Fillmore, who initially rejected purely syntactic deep structure relations, now includes them, along with case relations, in his case grammar (Fillmore, 1977). Second, our results follow a number of others that have failed to provide support for the standard theory.
as a basis for linguistic behavior. However, the experimental evidence against the standard theory has been evidence against the grammatical operations, not against the structural descriptions of the theory (Fodor, Bever and Garrett, 1974). The present evidence against the standard theory pertains instead to the structural descriptions. Hence, our demonstration of the inaccessibility of deep structure syntax has definite consequence.

This study also has an important methodological implication. In studying the relative accessibility of various linguistic concepts, the new recognition memory paradigm developed here seems to be more sensitive than the traditional concept formation task. Two factors may have been responsible for the increase in sensitivity: (1) All testing occurred after training was completed in the memory paradigm. (2) The memory paradigm tested whether a given concept would be discovered by the subject rather than whether the concept could be learned by the subject.

APPENDIX

Sentences Used as Stimuli

(1.) DS Subject--SS Subject--Experiencer

1. John was sleepy near the fire.
2. John was warm near the radiator.
3. John was comfortable near the window.
4. John was nervous next to the swimming pool.
5. John was cool near the stream.
6. John was content on the balcony.
7. John was confident at the wheel.
8. John was speechless in the gallery.
9. John was unhappy near the stage.
10. John was at ease in the motor boat.

(2.) DS Subject--SS Object of Prep.--Experiencer

1. The accident was imagined by John.
2. The concept was visualized by John.
3. The concert was enjoyed by John.
4. The odor was savored by John.
5. The director was feared by John.
6. The story was believed by John.
7. The theory was respected by John.
8. The teacher was despised by John.
9. The result was foreseen by John.
10. The earthquake was felt by John.

(3.) DS Object of Prep.--SS Object of Prep.--Experiencer

1. The roar was deafening to John.
2. The nap was refreshing to John.
3. The suggestion was disturbing to John.
4. The wasp was annoying to John.
5. The mask was frightening to John.
6. The play was amusing to John.
7. The conclusion was astonishing to John.
8. The crime was puzzling to John.
9. The voyage was exciting to John.
10. The message was comforting to John.

(4.) DS Subject--SS Subject--Goal

1. John was the recipient of the grant.
2. John was the beneficiary of the allowance.
3. John was the inheritor of the mansion.
4. John was the borrower of the bicycle.
5. John was the thief of the porcelain.
6. John was the buyer of the refrigerator.
7. John was the acquirer of the painting.
8. John was the consignee of the suitcase.
9. John was the receiver of the prize.
10. John was the catcher of the ball.

(5.) DS Object of Prep.--SS Subject--Goal

1. John was given the book.
2. John was bequeathed the inheritance.
3. John was tossed the paper.
4. John was handed the spatula.
5. John was mailed the record.
6. John was awarded the medal.
7. John was dealt the ace.
8. John was paid the bribe.
9. John was sent the instructions.
10. John was assigned the duty.

(6.) DS Object of Prep.--SS Object of Prep.--Goal

1. The property was leased to John.
2. The bottle was passed to John.
3. The football was kicked to John.
4. The jewels were entrusted to John.
5. The materials were supplied to John.
6. The money was allotted to John.
7. The scholarship was granted to John.
8. The reward was presented to John.
9. The book was returned to John.
10. The results were communicated to John.

(7.) DS Object of Prep.--SS Subject--Experiencer

1. John was assured misery.
2. John was permitted remorse.
3. John was authorized exuberance.
4. John was provided serenity.
5. John was offered happiness.
6. John was guaranteed anxiety.
7. John was promised tranquility.
8. John was allowed timidity.

(8.) DS Subject--SS Object of Prep.--Goal

1. The fruit was obtained by John.
2. The merchandise was recovered by John.
3. The taxes were collected by John.
4. The prize was received by John.
5. The crop was gathered by John.
6. The ball was caught by John.
7. The donuts were taken by John.
8. The frisbee was retrieved by John.

*Sentences used in Experiments I and II but not in Experiment III.

REFERENCES