The relative accessibility of semantic and deep-structure syntactic concepts

ALICE F. HEALY
Yale University, New Haven, Connecticut 06520
and Haskins Laboratories, New Haven, Connecticut 06511

and

ANDREA G. LEVITT
Wellesley College, Wellesley, Massachusetts 02181
and Haskins Laboratories, New Haven, Connecticut 06511

Three experiments were conducted to determine the relative accessibility of semantic and deep-structure syntactic concepts. In Experiment 1, which employed a concept-formation task, subjects learned the concept "deep-structure subject" more slowly than the case concept "experiencer." In Experiments 2 and 3, 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.

A number of different versions of transformational generative grammar have been proposed in recent years, including, among others, "generative semantics" (Lakoff, 1971; McCawley, 1968), case grammar (Fillmore, 1968, 1970, 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 & 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. Likewise, 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 into its proper location in the surface structure of the sentence.

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

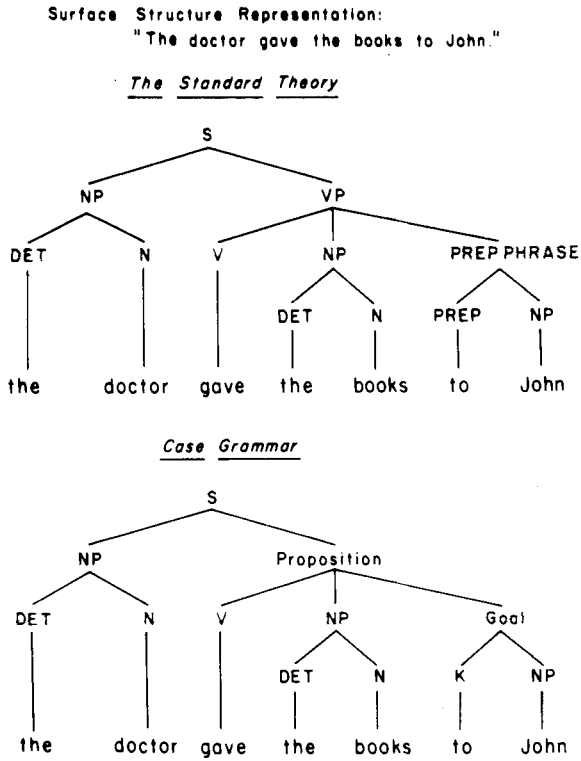*The Standard Theory*

Case Grammar

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).

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 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 used a concept-formation task, which enabled 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 1 of the present study, we compared the ease of learning a case relation and a deep-structure syntactic relation. Two baseline conditions were 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 learned a surface-structure syntactic relation, expected to be relatively trivial, and in the second baseline condition, subjects learned an arbitrary relation, defined in a manner analogous to that of the other three concepts.

It should be noted that, although case relations were specifically manipulated in this investigation, this study does not allow us to discriminate among various

Deep Structure Representation:
"The doctor gave the books to John."
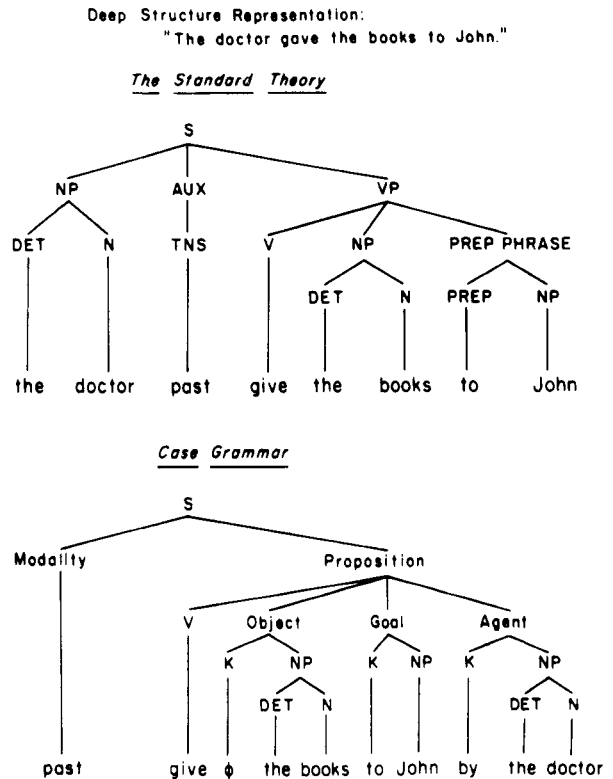
The Standard Theory

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).

semantically based grammars, since differences in case relations are necessarily confounded with differences in other semantic variables. Likewise, this investigation does not allow us to discriminate among different models of sentence memory and comprehension that assert structures of a semantic variety (e.g., Anderson & Bower, 1973; Kintsch, 1974; Rumelhart, Lindsay, & Norman, 1972; Schank, 1972). However, this study enables us to discriminate between such models based primarily on semantic relations and any plausible alternative models based primarily on deep-structure syntactic relations.

EXPERIMENT 1

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 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 the preposition, surface-structure subject, experiencer, and deep-structure subject, surface-structure object of the preposition, goal—were not employed, since fully satisfactory examples of these types could not be generated (see the introduction to Experiment 3 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 min. No subject had any formal training in linguistics. There were four conditions—arbitrary, deep-structure syntactic, semantic, and surface—with 10 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, 10 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 60 sentences was typed in the center of four 4 x 6 in. cards. Four decks of cards were constructed, one for each of the four conditions. Each deck included all 60 sentences; only the order of the sentences varied across decks. In each deck, the order of the sentences was pseudorandom with the constraint that each 12-sentence block included 2 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

Table 1
Eight Sentence Types with Examples

		Deep-Structure Category	
Surface-Structure Category		Subject	Object of Preposition
Subject	Experiencer	1. John was sleepy near the fire.	7. John was assured misery.
	Goal	4. John was the recipient of the grant.	5. John was given the book.
Object of Preposition	Experiencer	2. The accident was imagined by John.	3. The roar was deafening to John.
	Goal	8. The fruit was obtained by John.	6. The property was leased to John.

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 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 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 $< .001$;

Table 2
Mean Percentage of Errors in Experiment 1 as a Function of Condition and Block

Condition	Block					
	1	2	3	4	5	Mean
Surface	18	9	5	4	3	8
Semantic	42	27	22	13	9	23
Deep-Structure Syntactic	33	38	28	24	25	30
Arbitrary	49	43	44	35	34	41

Table 3
Mean Percentage of Errors in Experiment 1 as a Function of Condition and Sentence Type

Condition	Sentence Type					
	1	2	3	4	5	6
Surface	8	10	7	9	9	5
Semantic	19	29	26	32	13	16
Deep-Structure Syntactic	11	44	42	30	31	20
Arbitrary	26	48	56	22	68	27

$F_1(3,36) = 11.1$,   MSe $= 5,190$,   p $< .001$; $F_2(3,90) = 82.4$, MSe $= 140$, p $< .001$]. Newman-Keuls tests, based on the analysis with subjects as the random effect, revealed significant differences 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 conditions 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 $< .001$; $F_1(4,144) = 15.5$, MSe $= 818$, p $< .001$; $F_2(4,30) = 11.9$, MSe $= 212$, 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 $F'(12,128) < 1$; $F_1(12,144) = 1.6$, MSe $= 818$, p $= .105$; $F_2(12,90) = 1.8$, MSe $= 140$, 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 [$F_1(4,72) = 3.9$, MSe $= 715$, p $= .006$] and the test with sentences as the random effect [$F_2(4,30) = 2.9$, MSe $= 192$, 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 [$F_2(1,30) = 8.0$, MSe $= 192$, p $= .008$].

The analyses of variance further revealed a significant effect of sentence type [min $F'(5,131) = 4.6$, p $< .001$; $F_1(5,180) = 7.9$, MSe $= 1,500$, p $< .001$; $F_2(5,30) = 11.1$, MSe $= 212$, p $< .001$] and a significant interaction of Condition by Sentence Type [min $F'(15,270) = 2.6$, p $= .001$; $F_1(15,180) = 3.8$, MSe $= 1,500$, p $< .001$; $F_2(15,90) = 8.1$, MSe $= 140$, 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 2 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 are discovered by subjects in their attempts to learn a list of sentences for a subsequent memory test.

EXPERIMENT 2

The same relations were tested in this experiment as in Experiment 1. Furthermore, the same sentences were employed in the two experiments, with the following important exception: There were two versions of each of the 60 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 60 sentences, half of which were in the version with "John" and half in the version with "Sam." As in Experiment 1, 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 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 1. 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 that had been studied earlier included "John" or "Sam." In contrast to Experiment 1, 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 min. There were four conditions—arbitrary, deep-structure syntactic, semantic, and surface—with 10 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 1, except that in one deck (Deck A) all the sentences that were positive instances of the concept tested in Experiment 1 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 1 in the same order. The only differences between the form of the sentences as they appeared on the cards in Experiment 1 and as they appeared on the test lists in Experiment 2 were that the sentences were numbered (from 1 to 60) in Experiment 2 but not in Experiment 1, and the word "John" in each sentence in Experiment 1 was replaced by the pair of words "John/Sam" in Experiment 2. 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 directly opposite to 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 8 min to study the deck of sentences (timed by the experimenter with a stopwatch). The remaining 32 subjects were given 5 min 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. They were, however, 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.

Table 4
Mean Percentage of Errors in Experiment 2 as a Function of Condition and Sentence Type

Condition	Sentence Type						
	1	2	3	4	5	6	Mean
Surface	4	3	2	1	1	5	2.7
Semantic	5	4	9	5	4	2	4.8
Deep-Structure Syntactic	6	28	14	19	25	17	18.2
Arbitrary	4	17	7	11	26	33	16.3

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 (F_1) and one with sentences (F_2) as the random effect. According to these analyses, there was a significant effect of condition [min $F'(3,46) = 5.5$, $p = .003$; $F_1(3,36) = 6.2$, MSe = 598, $p = .002$; $F_2(3,162) = 48.5$, MSe = 77, $p < .001$], an effect of sentence type that approached significance [min $F'(5,205) = 2.1$, $p = .069$; $F_1(5,180) = 3.4$, MSe = 174, $p = .006$; $F_2(5,54) = 5.3$, MSe = 112, $p < .001$], and a significant interaction of these two factors [min $F'(15,307) = 1.8$, $p = .029$; $F_1(15,180) = 2.7$, MSe = 174, $p = .001$; $F_2(15,162) = 6.0$, MSe = 77, $p < .001$]. As in Experiment 1, 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 1.

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 $F'(1,21) = 6.8$, $p = .016$; $F_1(1,18) = 7.4$, MSe = 724, $p = .014$; $F_2(1,54) = 87.8$, MSe = 61, $p < .001$], but the effect of sentence type [min $F'(5,144) = 1.2$, $p = .337$; $F_1(5,90) = 1.8$, MSe = 158, $p = .129$; $F_2(5,54) = 3.3$, MSe = 84, $p = .011$] and the interaction of condition and sentence type [min $F'(5,138) = 1.8$, $p = .113$; $F_1(5,90) = 2.5$, MSe = 158, $p = .035$; $F_2(5,54) = 6.5$, 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. The latter concepts were discovered no more easily than purely arbitrarily defined relations.

EXPERIMENT 3

Only six sentence types were employed in Experiments 1 and 2, 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 that exists for several sentences of other types as well. Despite these difficulties, all eight sentence types were employed in Experiment 3, which was otherwise a replication of Experiment 2. (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 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 University who were taking a course in introductory psychology participated as subjects, receiving course credit. As in Experiment 2, 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 2, 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 two sentences of each of the original six types included in Experiment 2 but not in Experiment 3 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 2. On each list the order of the sentences was pseudorandom with the constraint that each 16-sentence block included 2 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 2, except that all subjects were given 5 min 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 2. Two analyses of variance were performed on these data, one with subjects (F_1) and one with sentences (F_2) as the random effect. These analyses yielded a significant main effect of condition [min $F'(3,46)$ = 10.1, $p < .001$; $F_1(3,36)$ = 11.4, MSe = 1,024, $p < .001$; $F_2(3,168)$ = 84.6, MSe = 111, $p < .001$], but neither the main effect of sentence type [min $F'(7,149)$ = 1.1, $p = .373$; $F_1(7,252)$ = 2.6, MSe = 179, $p = .013$; $F_2(7,56)$ = 1.9, MSe = 198, $p = .090$] nor the interaction of condition and sentence type [min $F'(21,418) < 1$; $F_1(21,252)$ = 1.2, MSe = 179, $p = .238$; $F_2(21,168)$ = 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, $p = .015$; $F_1(1,18)$ = 7.6, MSe = 1,337, $p = .013$; $F_2(1,56)$ = 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 2 are clearly supported by this pattern of results.

SUMMARY AND CONCLUSIONS

These experiments indicate that the deep-structure syntactic relations studied were both learned more slowly and discovered less readily than the semantic case concepts 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 & 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 has been a proliferation of proposed revisions of the standard theory. However, some influential theorists have recently argued (see Bever, Katz, & 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,

Table 5
Mean Percentage of Errors in Experiment 3 as a Function of Condition and Sentence Type

| Condition | Sentence Type | | | | | | | | Mean |
	1	2	3	4	5	6	7	8	
Surface	3.75	5.00	1.25	2.50	1.25	3.75	2.50	6.25	3.28
Semantic	5.00	7.50	7.50	7.50	6.25	10.00	6.25	11.25	7.66
Deep-Structure Syntactic	21.25	26.25	20.00	30.00	23.75	22.50	16.25	28.75	23.59
Arbitrary	13.75	31.25	37.50	30.00	31.25	37.50	15.00	30.00	28.28

in his case grammar (Fillmore, 1977). Second, the results of a number of psychological experiments 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, & 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 is not without considerable 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

Type 1: Deep-Structure Subject, Surface-Structure 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.*

Type 2: Deep-Structure Subject, Surface-Structure Object of Preposition, 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.*

Type 3: Deep-Structure Object of Preposition, Surface-Structure Object of Preposition, 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.*

Type 4: Deep-Structure Subject, Surface-Structure 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.*

Type 5: Deep-Structure Object of Preposition, Surface-Structure 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.*

Type 6: Deep-Structure Object of Preposition, Surface-Structure Object of Preposition, 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.*

Type 7: Deep-Structure Object of Preposition, Surface-Structure 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.

Type 8: Deep-Structure Subject, Surface-Structure Object of Preposition, 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 1 and 2 but not in Experiment 3.

REFERENCES

ANDERSON, J. R., & BOWER, G. H. *Human associative memory.* Washington, D.C: Winston, 1973.
BAKER, W. J., PRIDEAUX, G. D., & DERWING, B. L. Grammatical properties of sentences as a basis for concept formation. *Journal of Psycholinguistic Research,* 1973, **2,** 201-220.

BEVER, T. G., KATZ, J. J., & LANGENDOEN, D. T. *An integrated theory of linguistic ability*. New York: Crowell, 1976.

BLUMENTHAL, A. L. Prompted recall of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 203-206.

BLUMENTHAL, A. L., & BOAKES, R. Prompted recall of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 674-676.

CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge, Mass: MIT Press, 1965.

CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 335-359.

FILLMORE, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart, & Winston, 1968. Pp. 1-88.

FILLMORE, C. J. Subjects, speakers, and roles. *Synthese*, 1970, **21**, 251-274.

FILLMORE, C. J. Some problems for case grammar. In R. O'Brien (Ed.), *Report of the twenty-second annual round table meeting on linguistics and language study*. Washington, D.C: Georgetown University Press, 1971. Pp. 35-56.

FILLMORE, C. J. The case for case reopened. In P. Cole & J. M. Sadock (Eds.), *Syntax and semantics*, Vol. 8. New York: Academic Press, 1977.

FODOR, J. A., BEVER, T. G., & GARRETT, M. F. *The psychology of language: An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill, 1974.

KATZ, J. J., & FODOR, J. A. The structure of a semantic theory. *Language*, 1963, **39**, 170-210.

KINTSCH, W. *The representation of meaning in memory*. Hillsdale, N.J: Erlbaum, 1974.

LAKOFF, G. On generative semantics. In D. Steinberg & L. A. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*. Cambridge: Cambridge University Press, 1971. Pp. 232-296.

McCAWLEY, J. D. The role of semantics in a grammar. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart, & Winston, 1968. Pp. 124-169.

PERFETTI, C. A. Retrieval of sentence relations: Semantic vs. syntactic deep structure. *Cognition*, 1973, **2**, 95-105.

RUMELHART, D. E., LINDSAY, P. H., & NORMAN, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972. Pp. 197-246.

SCHANK, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, **3**, 552-631.

SHAFTO, M. The space for case. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 551-562.

NOTES

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.

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" (Fillmore, 1971, p. 42).