The present experiment examined recall for sentences admitting optional surface transformations. Such sentences were consistently more difficult to recall correctly than were control sentences which did not admit optional transformations. There were strong biases in the degree to which the two surface forms of an optional-transform sentence were correctly recalled. The recall biases showed little relation to the transformational status of the sentences, but showed a strong relation to speech production preference ratings, supporting a theory of memory which postulates that sentences are remembered in terms of their underlying meanings and that in recall the speech output system is used to reconstruct new surface structures for these retained meanings.

In one of the earliest experiments on human memory, Binet and Henri (1894) outlined a theory of reconstructive memory. They suggested that in the recall of prose material, idea memory (memory for meaning) is far superior to verbal memory (memory for surface information), so that when faced with the task of recalling prose, Ss remember the underlying meaning and reconstruct the surface form of their responses from the meaning. Binet and Henri pointed out that for children a process of "verbal assimilation" occurs during the reconstructive process. By verbal assimilation they meant that in reconstructing a new surface structure for a given meaning the children tended to replace the original surface structure with lexical items and syntactic constructions more characteristic of their own speech. Thus, according to Binet and Henri, recall of linguistic material involves the use of the speech production system to reconstruct a new surface form to express a retained meaning. Binet and Henri's (1894) reconstructive hypothesis was adopted by Bartlett (1932), but many recent theories of sentence recall have taken other approaches. A number of investigators have developed theories that assume that sentences are retained in terms of their surface structure (Johnson, 1968; Rosenberg, 1965), while others have postulated that sentences are remembered in terms of underlying syntactic entities (Mehler, 1963; Rohrman, 1968).

The present experiment used the recall of sentences that allow an optional transformation (Langendoen, 1969) to explore the reconstructive hypothesis and to compare the reconstructive approach with surface structure theories and syntactic theories of sentence recall. If recall of sentences involves the reconstruction of syntactic relations and lexical items from an abstract semantic representation, then the recall of sentences allowing optional transformations should produce particular difficulties. Sentences of this type have the property of providing two different surface realizations for the same semantic content (e.g., "The hi-fi fanatic turned up the volume," and "The hi-fi fanatic turned the volume up."). Thus, when attempting to reconstruct the original sentence from the retained meaning, two linguistic realizations will be available. When scored for verbatim recall, such sentences should be much more difficult than sentences with only one readily available surface realiza-
TABLE 1
EXAMPLES OF OPTIONAL-TRANSFORM SENTENCE PAIRS

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Item type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Particle movement</td>
<td>U</td>
<td>The hi-fi fanatic turned up the volume.</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>The hi-fi fanatic turned the volume up.</td>
</tr>
<tr>
<td>Genitive</td>
<td>U</td>
<td>The girlfriend of the policeman was a shoplifter.</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>The policeman's girlfriend was a shoplifter.</td>
</tr>
<tr>
<td>Pronominalization</td>
<td>U</td>
<td>The magician touched the girl and the girl disappeared.</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>The magician touched the girl and she disappeared.</td>
</tr>
<tr>
<td>Dative</td>
<td>U</td>
<td>Billy Graham told the president a dirty joke.</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>Billy Graham told a dirty joke to the president.</td>
</tr>
<tr>
<td>That deletion</td>
<td>U</td>
<td>The painting that the gallery sold was a forgery.</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>The painting the gallery sold was a forgery.</td>
</tr>
<tr>
<td>Adverb preposing</td>
<td>U</td>
<td>Tarzan heard the jungle drums at sunset.</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>At sunset Tarzan heard the jungle drums.</td>
</tr>
</tbody>
</table>

Note. Abbreviations: U = untransformed version, T = transformed version.

tion. Examination of the surface forms chosen by Ss to express a particular meaning should give information about the reconstructive process that is hypothesized to underlie recall.

The use of optional transformations also allows a test of syntactic theories of sentence memory. If sentences are remembered in terms of underlying syntactic entities, the untransformed versions should be recalled correctly more frequently than are the transformed, and the transformed versions should shift to the untransformed versions in recall more frequently than the reverse.

METHOD

Subjects

The Ss were 40 undergraduate students at the University of Illinois who participated in partial fulfillment of the course requirement for introductory psychology.

Materials

Presentation and test lists were constructed from a pool of 140 sentences. Ninety-six of the sentences were the paired surface forms derived from 48 sentences written to allow an optional transformation. In the framework of transformational linguistics, one member of each of the pairs would be considered the untransformed sentence, and the other the transform (i.e., derived by application of an optional transformation). Six different types of optional transformations were included, each represented by 8 different sentence pairs. The transformations were Dative (Burt, 1971); Pronominalization (Langacker, 1967); Adverb Preposing (Jacobs & Rosenbaum, 1968); Particle Movement (Jacobs & Rosenbaum, 1968); that Deletion (Rosenbaum, 1967); and Genitive (Langendoen, 1969). Examples of pairs employing each of these transformations are given in Table 1. This table also indicates which form has been postulated by linguists to be the untransformed version and which the transformed version.

Of the other 44 sentences in the original pool, 28 were control items, randomly arranged into 14 pairs. The control sentences were simple declaratives, e.g., "The boulder blocked the entrance to the cave." The remaining 16 sentences consisted of 8 pairs of filler items.

Two 70-item master lists were constructed by randomly assigning one member of each of the 70 sentence pairs to each master list. The two master lists were each subdivided into seven lists of 10 sentences. Sentences were randomly assigned to these 10-item lists, with the constraint that each list contain two control items and no list contain more than two instances of a given optional-transform type. Ordering of sentences within lists was random, with the restriction that 2 sentences of the same type not appear consecutively. Ordering of pairs within and across lists was constant, so that pair mates always appeared in parallel positions. Eight additional items were used to construct a practice list.

Procedure

Recall. Two groups of 11 Ss were run in the recall condition, each receiving an 8-item practice list and seven 10-item experimental lists derived from one of the two master lists. The E read each list of sentences aloud with normal intonation, allowing 3 sec. between sentences. After the last sentence in each list, E read four strings of seven- and eight-digit numbers. After each string of digits, Ss wrote...
these numbers on the first page of their answer booklets. Immediately following the digit-recall task, Ss were asked to recall the sentences. Approximately 1 min. elapsed between the reading of the last sentence and the beginning of the memory test. The Ss were given 4 min. to write their responses for each list of sentences. The Ss were informed that they were participating in an experiment on memory for sentences and that written recall would be required. The instructions were neutral with respect to rote- or gist-memory strategies.

Cues for each sentence, in the same order in which those sentences had been read, were provided on the answer sheets. The cues given for the optional-transform pairs were the same, and consisted of a single noun phrase that appeared in both sentences. Cues were selected to avoid biasing recall toward either the untransformed or transformed member of the pair. Pairs in the Adverb Preposing and Genitive classes were cued with noun phrases that were noninitial in both versions of the sentence. Cues for the remaining sentences were chosen so that over all sentences approximately half were sentence-initial and half were noninitial.

Rating. The 48 experimental sentence pairs were rated by 18 Ss for speech preference. Rating booklets contained each of the sentence pairs in a two-alternative multiple-choice format. Order of pairs and order of sentences within a pair were random. The Ss were instructed to indicate the member of each pair which they thought "sounded better" or was "more natural."

Scoring. The sentences from the recall protocols were scored as correct, shift to pair mate, error, or omission. In order to be scored correct, a sentence had to retain the exact words of the original sentence in their presented order, with the exception of changes in tense, number (singular/plural), articles (definite/indefinite), spelling, contractions, and abbreviations. A shift was scored if the recalled sentence was the optional-transform pair mate of the presented sentence, with the same exceptions as in the correct category. An omission was scored if nothing was written. All other responses were scored as errors.

**Results**

All analyses of variance were carried out with both Ss and sentences as random effects (Clark, 1973). Since the basic data consisted of 0s and 1s, the lower bound of the appropriate quasi F ratio ($F'_{\text{min}}$) was used to test each fixed effect. Following Clark's notation, $F_1$ refers to the test statistic with Ss as a random effect, and $F_2$ to the test statistic with items as a random-effect.

The percentage of responses in each scoring category for the optional-transform sentences was 41% correct, 16% shift, 33% error, and 10% omit. For the control sentences the corresponding findings were 57% correct, 28% error, and 15% omit. The difference between the optional-transform sentences and the control sentences in percent correct is significant, $F_1 (1, 21) = 34.58$, $F_2 (1, 74) = 13.95$, $F'_{\text{min}} (1, 93) = 9.94$, $p < .005$. However, when the correct and shift categories for the optional-transform sentences are collapsed, the overall rate of meaning-preserving responses becomes 57%, which is the same as the percentage of corrects for the control sentences.

For the purpose of comparing the data with the predictions of the syntactic models of sentence memory, Table 2 gives the percentages of responses in each scoring category for each of the six types of sentences, with items of each type classified as untransformed or transformed. Analysis of the correct responses showed that there is a significant interaction of transform status and sentence type, $F_1 (5, 105) = 13.13$, $F_2 (5, 42) = 6.36$, $F'_{\text{min}} (5, 85) = 4.29$, $p < .005$. Examination of Table 2 shows that for three of the six sentence types the untransformed sentences were correctly recalled more frequently than the transformed sentences, and for the other three sentence types the reverse held. Follow-up
tests of the simple effects showed that the difference between the transformed and untransformed sentences was significant only for Particle Movement, $F_1 (1, 21) = 45.00$, $F_2 (1, 7) = 18.98$, $F'_{\text{min}} (1, 13) = 13.35$, $p < .005$, with the untransformed sentences more frequently correct.

A similar analysis was carried out on shifts. There was a significant interaction of transform status and sentence type, $F_1 (5, 105) = 12.47$, $F_2 (5, 42) = 9.01$, $F'_{\text{min}} (5, 103) = 5.23$, $p < .0005$. Examination of Table 2 shows that for three of the six sentence types the transformed sentences were more frequently recalled as untransformed sentences than the reverse, and follow-up tests showed these differences to be significant for Particle Movement, $F_1 (1, 21) = 26.38$, $F_2 (1, 7) = 9.80$, $F'_{\text{min}} (1, 13) = 7.15$, $p < .05$, and Adverb Preposing, $F_1 (1, 21) = 7.51$, $F_2 (1, 7) = 15.91$, $F'_{\text{min}} (1, 27) = 5.10$, $p < .05$. For the other three sentence types the untransformed sentences were more frequently recalled in the transformed version than the reverse. Follow-up tests showed these differences to be significant for Genitive, $F_1 (1, 21) = 19.61$, $F_2 (1, 7) = 12.21$, $F'_{\text{min}} (1, 16) = 7.52$, $p < .05$, and for Pronominalization, $F_1 (1, 21) = 21.72$, $F_2 (1, 7) = 16.66$, $F'_{\text{min}} (1, 18) = 9.43$, $p < .01$.

In order to examine the relation between the preference ratings and the recall data, each member of each optional-transform pair was classified as either preferred or nonpreferred. The preferred items were the members of each pair chosen by the largest percentage of Ss in the rating condition. For four items the preference scores (the percentage of Ss choosing a given member of an optional transform pair) were identical, and members of these pairs were assigned randomly to the preference–nonpreference categories. Table 3 gives the recall data for each of the six sentence types for preferred and nonpreferred forms, and the preference scores for each category.

Analysis of corrects showed that there was no significant interaction of preference–nonpreference with sentence type. Examination of Table 3 shows that for each sentence type the preferred forms are recalled correctly more frequently than are the nonpreferred forms. The main effect of
preference–nonpreference was significant, $F_1 (1, 21) = 43.68$, $F_2 (1, 42) = 28.97$, $F_{\text{min}}' (1, 62) = 17.42$, $p < .0005$. A similar analysis on shifts showed that there was no significant interaction of preference–nonpreference with sentence type. Examination of Table 3 shows that for each sentence type the nonpreferred forms shift in memory to the preferred forms more often than the reverse. The main effect of preference–nonpreference was significant, $F_1 (1, 21) = 84.70$, $F_2 (1, 42) = 42.98$, $F_{\text{min}}' (1, 63) = 28.51$, $p < .0005$.

While the separate analyses of the recall data in terms of the syntactic hypothesis and in terms of the preference hypothesis clearly show the superiority of the preference hypothesis, it is possible to compare the two hypotheses in a more direct fashion. Table 4 gives the preference ratings for the transformed and untransformed sentences for each of the six types of optional transformations and the number of transformed and untransformed sentences preferred on the rating task. This table shows that in some classes, such as Particle Movement and Adverb Preposing, the untransformed pair members tended to receive higher speech preference ratings, while in other classes, such as Pronominalization and Genitive, the transformed pair members tended to receive the higher preference ratings. Across all six optional-transform classes there were 23 sentence pairs in which the transformed member was preferred in the rating task. These pairs provide a crucial comparison of the syntactic hypothesis vs. the preference hypothesis, since for each sentence pair the syntactic hypothesis predicts that the untransformed sentence will be recalled correctly more frequently than the corresponding transformed sentence, while the preference hypothesis predicts the reverse. Of the 23 pairs where the predictions are opposed, there were 21 where the recall data were not tied for number correct, and for 16 of these pairs the transformed (preferred) member of the pair was recalled correctly more often than the untransformed (non-preferred) member ($p < .01$ by a Wilcoxon matched-pairs test). Thus, on the items for which the two theories make opposing predictions, the recall data clearly support the preference hypothesis.

**TABLE 4**

| Sentence type         | Item type | Percentage of Ss preferring | No. of P items
|-----------------------|-----------|-----------------------------|----------------
| Particle movement     | U         | 71                          | 7              
|                       | T         | 29                          | 1              
| Genitive              | U         | 28                          | 1              
|                       | T         | 72                          | 6              
| Pronominalization     | U         | 13                          | 0              
|                       | T         | 87                          | 8              
| Dative                | U         | 53                          | 3              
|                       | T         | 47                          | 2              
| *That* deletion       | U         | 50                          | 3              
|                       | T         | 50                          | 5              
| Adverb preposing      | U         | 63                          | 7              
|                       | T         | 37                          | 1              
| All types             | U         | 46                          | 21             
|                       | T         | 54                          | 23             

Note. Abbreviations: U = untransformed version, T = transformed version, P = preferred. *Does not include the four items with tied preference scores.

**DISCUSSION**

The results show that sentences that admit an optional surface transformation are more difficult to recall correctly than control sentences with only one readily available surface structure. All six classes of optional transforms sampled showed this difficulty in memory, suggesting that this difficulty will occur in any class of sentences admitting optional surface transformations.

Although the experimental sentences were correctly recalled less frequently than control sentences, the total of *shift* and *correct* recalls was the same as the recall score for control sentences. This suggests that for both types of sentences an abstract representation of the meaning was remembered rather than the exact words, and that in recall the surface structure was reconstructed from this abstract representation. The reconstructive process was not revealed in the recall of the control sentences, but the availability of two surface structure variants for the experimental sentences resulted in the large number of meaning-
preserving shifts during the reconstructive process.

Sachs (1967) and Begg (1971) have shown with a recognition task that surface structure changes in sentences are difficult to detect, while semantic changes are easily identified. Their finding that surface information is remembered less well than the semantic content of the sentence is confirmed by the present experiment with a recall task. The relatively poor recall of surface information compared to the recall of meaning is demonstrated in rather striking fashion by the finding that for the nonpreferred members of the sentence pairs Ss were as likely to write down the alternative surface form as they were to write down the form actually presented.

The difficulty of optional-transform sentences in recall cannot be adequately explained by theories concerned only with the surface features of language (Johnson, 1968; Rosenberg, 1965). These theories are inadequate because they do not provide an analysis independent of surface structure and therefore make no differential prediction about the recall of sentences admitting two surface realizations and those that do not. The poor recall of the experimental sentences and the occurrence of the predicted shifts in the present study require the postulation of a type of memory that is independent of the surface structure actually employed in the sentences.

The syntactically based theories of memory also fail to explain the results. The kernel-plus-tag theory (Mehler, 1963; Miller, 1962) assumes that sentences are remembered in terms of a simple affirmative active declarative, plus a tag (or tags) containing information about the syntactic form of the sentence. This theory is incapable of accounting for the present data, since the sentences employed require a more detailed analysis of transformational relationships and deep syntactic structures. The theories that grew out of the kernel-plus-tag proposal, however, do make predictions about the data. The hypothesis that linguistic deep structures describe the memory representation (Blumenthal, 1967; Rohrman, 1968) predicts that sentences should be more frequently recalled in their untransformed version, since this is the form closest to the underlying structure. Similarly, the derivational theory of complexity (as reconstructed by Fodor & Garrett, 1967) holds that the greater the number of transformations undergone by a sentence, the more difficult it should be to process and remember; thus, sentences presented in the untransformed version should be correctly recalled more frequently than sentences presented in derived forms. However, the recall data show that there is very little conformity between an item's transform status and the number of times it is recalled in a particular form. Only Particle Movement sentences were recalled correctly significantly more frequently in the untransformed version than in the transformed, and in the shift category, only transformed Particle Movement and transformed Adverb Preposing sentences were recalled significantly more frequently in the untransformed versions. Untransformed Genitive and untransformed Pronominalization sentences, on the other hand, were recalled significantly more frequently in the transformed version, contrary to the predictions of the syntactic theories.

The basic results of this study, however, are compatible with the memory-for-meaning approach (Bartlett, 1932; Binet & Henri, 1894; Bransford & Franks, 1971; Brewer, in press). The Ss appear to have remembered the sentences in terms of their meanings and from these reconstructed the lexical items and syntactic relations.

The asymmetry found in recall for preferred and nonpreferred sentences provides some limited evidence about the nature of the reconstructive processes in memory. In normal speaking there must be a complex speech production system underlying the conversion of the speaker's initial idea into a set of sequentially organized sounds. In speaking it is necessary to establish the particular syntactic realization of the higher level semantic relations ('give the man a book;' vs. "give a book to the man."). The underlying ideas must be realized by particular lexical items (big vs. large). Finally, the resulting structures must be given particular phonological form.

The results of the present study suggest that in attempting to recall sentences a similar reconstructive process occurs. It appears that Ss frequently remembered the meaning of the presented sentences, but not the particular surface realization, and so they had to reconstruct an appropriate surface structure in recall. Since the optional-transform sentences had been designed explicitly to provide two different surface realizations of one underlying meaning, Ss were faced with two alternatives for expressing the retained meaning of such sentences. At the point of reconstructing a particular surface structure, it appears that there were fairly strong speech output biases.
in favor of certain surface forms. The preference-rating task provided an independent assessment of these speech output biases and thus was successful in predicting the large asymmetry in the recall data, with the preferred forms recalled correctly more frequently than the nonpreferred forms and the nonpreferred forms showing more frequent shifts to the preferred forms than vice versa. A similar hypothesis about speech output biases in reconstructive recall has been proposed by James, Thompson, and Baldwin (1973) to account for the recall of active and passive sentences, and it seems likely that many other studies of sentence recall can be reinterpreted to provide additional evidence for the reconstructive hypothesis (e.g., Clark & Clark, 1968; Mehler, 1963).

Thus, the overall results of this study support Binet and Henri's (1894) reconstructive hypothesis, except that the phenomenon of verbal assimilation does not appear to be restricted to children. In the course of recalling sentences, adults also tend to replace the original surface structures with surface structures more compatible with their own speech output system.

REFERENCES


(Received September 1, 1973)