

Recognition memory for syntactic and semantic aspects of connected discourse¹

JACQUELINE STRUNK SACHS

BELL TELEPHONE LABORATORIES, MURRAY HILL, N. J.

This study investigates the pattern of retention of syntactic and semantic information shortly after comprehension of connected discourse. Ninety-six Ss listened to 24 taped passages and, after each passage, heard one recognition test sentence which was either identical to a sentence that had occurred in the passage, or was changed in some slight way. The Ss responded "identical" or "changed," rated their confidence, and classified changes as "meaning" or "form." Two independent variables were manipulated: (1) The relationship between the original sentence in the passage and the test sentence. The test sentence was (a) semantically changed, (b) changed from active to passive voice or vice versa, (c) formally changed in other ways that did not affect the meaning, or (d) unchanged. Each sentence appeared in all change types. (2) The amount of interpolated material between the original and test sentences was zero, 80, or 160 syllables of connected discourse which was a continuation of the passage. Each S heard passages representing all levels of each variable. All combinations of particular passages, relationship of original and test sentence, and amount of interpolated material were tested.

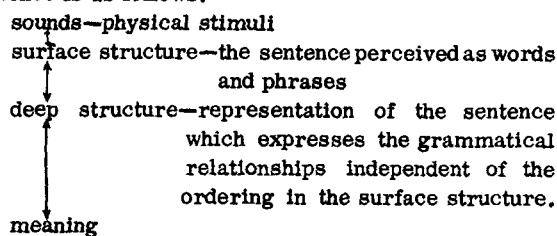
When the test sentence was heard immediately after the original, retention was high for all test types. But after 80-160 syllables, recognition for syntactic changes had dropped to near chance levels while remaining high for semantic changes. Even when the meaning of a sentence was remembered, formal properties that were not necessary for that meaning were forgotten very quickly. The results suggest that the original form of the sentence is stored only for the short time necessary for comprehension to occur. When a semantic interpretation has been made, the meaning is stored. Thus the memory of the meaning is not dependent on memory of the original form of the sentence.

When language is comprehended, it seems that the meaning of what was heard or read is remembered to some extent, but unless special attention is given to the style or other characteristics of the words, the exact wording is forgotten. This phenomenon can be viewed in the framework of a model of comprehension derived from linguistic theory (Chomsky, 1965).

A naive view of comprehension might assume that the semantic interpretation of a sentence is made directly from the perceived string of words by combining their individual meanings. However, if one heard three isolated words in sequence—monkeys, wrestling, entertaining—various associations and images might be called up, but the hearer would not be likely to reconstruct what the speaker had in mind. Understanding is not simply adding up a sequence of meanings of

individual words. If these words are placed into two sentences, "Wrestling monkeys are entertaining" and "Wrestling monkeys is entertaining" we know the sentences have quite different meanings. At first glance, one might suppose that the difference between the two sentences lies solely in the difference between "are" and "is" since these are the only two words that differ. However, even one sentence may have two different meanings, if there are two different grammatical structures underlying it. "Wrestling monkeys can be entertaining" is essentially two sentences, and may be understood as "It can be entertaining to wrestle with monkeys" if the subject is taken to be "Wrestling," or "When monkeys wrestle it can be fun to watch them" if "monkeys" is taken as the subject. Therefore, in both examples the semantic interpretation depends both on the meanings of the words and on the syntactic structure that interrelates them.

A representation of the sentence which expresses these syntactic relationships is called the "deep structure" (Chomsky, 1965). Words that are closely related in the deep structure may not be those that are physically close in the actual sentence as a perceptual stimulus. The linguistic model of the analysis of a sentence is as follows:



Our model of comprehension states, similarly, that the meaning of a sentence must depend on the deep structure representation which expresses grammatical relations. For comprehension to occur, it seems clear that one must know, at some level, which part of the sentence is the subject, object, and so forth. However, we cannot yet specify the psychological processes involved in ascertaining grammatical relations that are not apparent from the physical ordering of the words in the sentence. According to the model, characteristics of the surface structure that are not uniquely related to the meaning of the sentence may not be remembered after comprehension. In many instances a sentence will not be recalled in its exact wording, but in another wording that expresses the same meaning or deep structure.

In the long tradition of research with verbal materials, many studies support the notion that exact wording is not remembered, but few have been designed specifically to investigate the memory for the structure of sentences after comprehension. Mehler (1963) found that when Ss learned a set of unrelated sentences of various syntactic types, the bulk of the errors occurred because Ss altered the syntactic form of the sentence in recall. Clifton, Kurcz, and Jenkins (1965) found that even with the more sensitive recognition procedure, Ss confused sentences they had attempted to learn with sentences of syntactically related form.

Mehler (1963, p. 350) has suggested that when S tries to learn a sentence, he "presumably codes it as an underlying kernel plus some 'mental tag' that indicates that [a] transformation must be applied for recall." Our interest, however, lies in the process of normal comprehension and storage of information, rather than in sentence memorizing. We hypothesize that sentences are encoded only with respect to their meaning, and that no "mental tag" indicating form is stored if the task is comprehension instead of learning. In the present study, it was desirable to keep the experimental task as similar to the natural comprehension situation as possible, and to do everything possible to insure that language was responded to *as language*. Therefore the stimulus material was connected discourse, not isolated sentences, with auditory presentation, and the required responses forced Ss to attend to the meaning of what was heard.

The questions asked in the present study are: Is the syntactic information contained in a sentence stored at all after comprehension? Are some types stored better than others? What is the time course of loss of this information? Answers were sought by finding out how well Ss could recognize syntactic changes that did not change the meaning of the sentences they had comprehended.

To investigate these questions, we must, of course, have a clear distinction between the semantic and syntactic information contained in a sentence. If every syntactic difference also signals a semantic difference, the exact wording of a sentence could not be changed without changing the original meaning. Present linguistic theory, however, classifies certain formal properties of language as "purely stylistic." That is, two sentences can have different forms but express the same meaning. For example, "He looked the number up" can be changed to "He looked up the number," thus changing the surface structure while maintaining the deep structure (Chomsky, 1965). Changes of this type will be referred to as "formal" changes. In other cases, however, it is not so clear whether a change in the syntax signals a semantic difference. For example, the active and passive of a sentence seem to express much the same denotative meaning, but may express a difference in emphasis. In Katz

and Postal's (1964) analysis, the difference between active and passive is represented in the deep structure. If this linguistic analysis is correct, and if one remembers sentences in terms of their "deep structure," active and passive sentences should be remembered differently. On the other hand, if memory is something even more abstract and less language-specific, the difference between active and passive sentences should not be noted.

Let us review the argument. We suggested a theory of comprehension in which a deep structure is derived from the surface structure of a sentence by syntactic processing, and a semantic interpretation is made from that deep structure. We have also suggested that the meaning of a message is retained without the exact form because the surface structure has been lost during the process of interpretation. Form which is not relevant to the meaning is normally not retained. Previous studies have supplied little information about the time course of memory for grammatical form.

On this basis we propose the following hypothesis: If Ss are asked to recognize sentences that they have comprehended, they should be able to do so better on the basis of the meaning than of the form. Sentences which differ only in their form, but not in their meaning, will be difficult to recognize as different sentences after a brief interval.

METHOD

Subjects

The Ss were 96 undergraduate students (51 female, 45 male) who participated to fulfill requirements for psychology courses.

Materials

Twenty-eight passages of connected discourse were constructed. A sample passage is given here and will be referred to in the explanation of the methodology:

There is an interesting story about the telescope. In Holland, a man named Lippershey was an eye-glass maker. One day his children were playing with some lenses. They discovered that things seemed very close if two lenses were held about a foot apart. Lippershey began experimenting and his "spyglass" attracted much attention. He sent a letter about it to Galileo, the great Italian scientist. Galileo at once realized the importance of the discovery and set about to build an instrument of his own. He used an old organ pipe with one lens curved out and the other in. On the first clear night he pointed the glass toward the sky. He was amazed to find the empty dark spaces filled with brightly gleaming stars! [80 syllables] Night after night Galileo climbed to a high tower, sweeping the sky with his telescope. One night he saw Jupiter, and to his great surprise discovered near it three bright stars, two to the east and one to the west. On the next night, however, all were to the

west. A few nights later there were four little stars. [160 syllables] [Bell rings]

Set of related sentences (one of which would be the test sentence):

Base—He sent a letter about it to Galileo, the great Italian scientist.

Semantic—Galileo, the great Italian scientist, sent him a letter about it.

Passive—A letter about it was sent to Galileo, the great Italian scientist.

Formal—He sent Galileo, the great Italian scientist, a letter about it.

For each passage there was a different set of four related sentences, any one of which could appear in the passage. The sentence that actually appeared in the passage will be referred to as the "original sentence." The original sentence is italicized in the sample passage. A "test sentence" tested S's recognition memory of the original sentence. The characteristics of the sets of sentences used as original and test sentences are described below.

The passage length preceding the original sentence ranged from 27 to 180 syllables with a mean of 98 syllables. Three aspects of the construction of the passage and test sentence were varied systematically: (1) amount of interpolated material between original and test sentence, (2) relationship between original and test sentence, and (3) order of presentation of the two related sentences.

(1) Amount of Interpolated Material. To test the strength of recognition memory for the original sentence at various intervals after comprehension, three amounts of interpolated material (IM) occurred between the original and test sentences: zero syllables, 80 syllables (approximately 27 sec), and 160 syllables (approximately 46 sec). The IM was a continuation of the passage, with no break of any kind at the original sentence, so S never knew what sentence would be heard in the recognition test. At the end of the passage, a bell rang softly to signal the onset of the test sentence.

(2) Relationship between the Original and Test Sentences. The test sentence could be identical to the original, or changed from it in some respect. For each passage there was a set of four related interchangeable sentences. Any one could be inserted into the passage as the original sentence. The first sentence in each set will be referred to as the "base" sentence. The other three were derived by changing the base sentence in some respect. Changes were of three types:

(a) Semantic (Sem). The meaning of the base sentence was altered. To insure against Ss recognizing the change on the basis of a word that had not occurred in the passage, new content words were not introduced in the changed version. Semantic changes were achieved by changing subject and object in the sentence or phrase, by negation, or by substitution of a word that occurred

elsewhere in the passage. The passage was constructed to be neutral with respect to the semantic content in the original sentence. That is, no information in the rest of the passage either supported or denied the meaning of the sentence that would be heard in the recognition test. Either the base or its semantically changed version, Sem, could appear as the original sentence.

(b) Passive/Active (P/A). The base sentence was changed from active to passive voice or vice versa. In many cases the change from active to passive was accompanied by dropping the agent in order to preserve the ordinary usage of the passive.

(c) Formal (Form). The form of the base sentence was changed without changing its meaning. These changes were based on examples from various linguistic books. The two versions of the sentence were completely synonymous and presumably had identical deep structures.

Therefore, there were four relationships between the original and test sentences: semantic, passive/active, formal changes, and identical.

(3) Order of Presentation of the Two Related Sentences. It is possible that one form from any of the pairs of related sentences (Base-Sem, Base-P/A, Base-Form) might be a more common form, or that it might be preferred by Ss for some reason. For example, in the following pair, one of the sentences might sound more normal to S.

"He sent a letter about it to Galileo, the great Italian scientist."

"A letter about it was sent to Galileo, the great Italian scientist."

If S did not remember the form of the sentence from the original, he would be more likely to accept a normal sounding sentence as having occurred, but reject the less normal one. If the preferred form were always the test sentence, the recognition scores would be higher than they actually should be. That is, we could not tell if S remembered the sentence or whether it simply sounded right to him. To avoid this problem, the base sentence was used in the passage as the original half the time, and a changed version (Sem, P/A, or Form) was used as the original half the time.

Five judges read the passages before the experiment and, having read the instructions that Ss would receive in the experiment, guessed in each passage which sentence would be the recognition test sentence. They were not able to pick the original sentence better than chance ($t = .50, p > .30$).

The stimulus materials were tape-recorded by a male actor. To achieve balancing of the design, 24 separate tapes of 28 passages each were made, each 1/2 h long. Between the test sentence and the beginning of the next passage there were 15 sec for Ss' responses. Each passage was introduced by a number to signal the onset of the next trial and to be sure S

RELATIONSHIP BETWEEN ORIGINAL AND TEST SENTENCES	SEMANTIC		PASSIVE/ACTIVE		FORMAL		IDENTICAL					
	BASE SEM	SEM BASE	BASE P/A	P/A BASE	BASE FORM	FORM BASE	BASE ID	ID BASE				
AMOUNT OF INTERPOLATED MATERIAL	0	80	160	0	80	160	0	80	160	0	80	160
PASSAGE NO.	1	2	3	4	24							

Fig. 1. Balancing of the experimental design. The shaded cells give an example of the make-up of one of the 24 tapes prior to randomization of the order of passages.

responded in the correct place in his answer booklet. The recordings were made at the University of California Language Laboratory under recording studio conditions to avoid background noise or unclearness in the tapes, and also to edit out all errors or hesitations made during the reading. Playback was on Wollensack tape recorders.

Design

Four of the 28 passages were designated as warm-up trials. These were the same for all Ss. The balance of the design is diagrammed in Fig. 1. There were four relationships between original and test sentences, two orders of presentation of the related sentences, and three amounts of interpolated material. These combine to form 24 possible treatments, each represented by a column in Fig. 1. Each passage is represented by a row. Twenty-four tape recordings of the stimulus materials were prepared. On each tape, all the rows and columns were represented, but none was represented more than once. The shaded cells give an example of the makeup of one of the 24 tapes, prior to randomization of the order of passages. The order of passages was randomized by experimental blocks, so that several instances of the same type of treatment never occurred together. Each tape was heard by four Ss; therefore, four Ss are represented in each cell of the matrix. The assignment of Ss to tapes was randomized.

Procedure

Each S was tested individually in one 45-min session. The S heard a series of 28 passages on tape. Each passage was interrupted by a bell, and a test sentence was heard. If S thought the test sentence was identical to one in the preceding passage, he marked "identical" in an answer booklet. If he perceived any change at all he marked "changed," and then attempted to classify the type of change as in the meaning or in the form of the sentence. He also rated his confidence in his judgment of "identical" or "changed" on a 5-point scale. The instructions were as follows:

"The purpose of this experiment is to find out how well people can remember what they have just listened to. The experiment will take about 3/4 hour.

"You will hear, on tape, a series of short passages—some folk tales, history, biography, and so on—but each passage will be interrupted. A bell will sound, and then one sentence from somewhere in the passage will be repeated. Sometimes it will be repeated with exactly the same words as you heard in the original. But sometimes it will be changed in some small way. Either the meaning of the sentence will be changed, or only the grammatical form.

"For example, if you had heard 'He gave her a doll' and then 'She gave him a doll' this would be a change in meaning. If you had heard 'He looked the number up' and then 'He looked up the number' this would be a change in form. In English we often have alternative ways of saying the same thing. That is, the meaning is the same but the form is different.

"If the words are just as they were in the original sentence in the passage, mark *identical* in your answer booklet. If there is any change at all, you will mark *changed*. Look at the right side of the booklet. If you know that the change was in the meaning or in the form, mark one of these. If you think it has changed but don't know how—either because it just sounds different to you or because you don't know how to classify the change—mark under *changed, don't know*."

In order to be sure that S understood the task and the classification system, and to familiarize S with the system for recording answers, S had to classify seven sample pairs of sentences that were printed on the instruction sheet. If any error were made, it was pointed out and explained by E. The instructions continued:

"Of course in the experiment you won't see these written out and have to try to remember what the original sentence was. Listen normally to the passages. The stories go much too fast to try to memorize anything, and anyway, you must attend to the meaning fully as well as the form."

"In the booklet is also a place to mark how confident you are of your judgment. This confidence rating applies only to whether you think the test sentence is identical or changed. It does not apply to the classification of change—that is why the rating follows the identical-changed response. Rating 1 is very low confidence—if you are making a complete

guess; 5 is for very high—if you are absolutely sure you are right. Circle one of the confidence ratings each time you make a judgment. Sometimes you will have to guess, but be sure to always mark one alternative under *identical* or *changed* and mark low confidence if you must guess. Do not be disturbed at the difficult ones. Just listen normally and judge the sentence, if you can't remember, on whether it sounds right to you."

The S was encouraged to ask questions. After the first two test passages had been played, the tape was stopped and S again had an opportunity to ask questions. Then he was left alone in the experimental cubicle to avoid distractions while he listened to the passages. He did not know how many passages he would hear, and the answer booklet contained extra pages to avoid an "end spurt" effect. The E could be contacted in the next room if a problem arose, but no S ever interrupted the test session.

RESULTS

The percentage of correct judgments for each test type is illustrated in Fig. 2. Ss recognized Semantic changes at a high level at all three amounts of IM tested, whereas recognition of identical sentences as well as of the Passive/Active and Formal changes is close to Semantic at IM-0, but drops off considerably before IM-80. Ss also rated their confidence in each of their "identical" or "changed" judgments. A score for each response was based on the correctness and the confidence rating. Thus a score for a response could be from +5 to +1 when correct, and from -1 to -5 when incorrect. The mean score for each test type is shown in Table 1. Using these scores, a Friedman rank test (which is less sensitive than analysis of variance to violations of normality) showed no

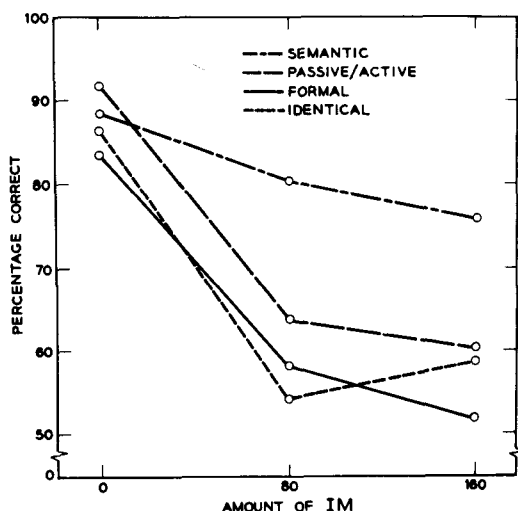


Fig. 2. Percentage of judgments *identical* and *changed* that were correct for each test type.

Table 1

		Mean score for each test type		
		Syllables of Interpolated Material		
Relationship		0	60	180
	Sem	3.72	2.71	2.47
	P/A	3.87	.96	.78
	Form	3.22	.55	.17
Ident	3.44	.44	.95	

significant differences among Sem, P/A, Form, and Id at IM-0 ($F = .97$, $df = 2.98/283.10$, $p > .25$). Analysis of variance for P/A, Form, and Id relationships at IM-80 and IM-160 revealed no significant effect due either to the relationship between the original and test sentences ($F = 2.09$, $df = 2/190$, $p > .10$), or to the amount of interpolated material ($F = .02$, $df = 1/95$, $p > .25$). However, there are clear differences between the recognition memory for semantic changes as compared with the other test types at IM-80 and IM-160 ($t = 5.74$ and 5.22 , respectively, $p < .001$ for both tests).

After 80 and 160 syllables of IM, Ss were still somewhat able to recognize syntactic changes at above chance levels. At IM-80, the difference between "changed" responses to formally changed sentences and "changed" responses to identical sentences is small, but significant ($t = 2.68$, $p < .01$). At the point of lowest recognition (Form-IM-160), the difference is significant only at the .05 level ($t = 2.31$).

The ability to classify the type of change correctly also drops off sharply for P/A and Form changes. With immediate repetition, 83% of the changes were classified as formal for P/A and 82% for Form. For these test types at IM80 and 160, the percentage of "changed" judgments that were also assigned to the correct classification is approximately 50%. Even if only the judgments with the highest confidence rating (5) are considered, Ss were only guessing at the classification of the change. When Ss could recognize that a sentence had been changed, if the change was syntactic they could not identify it as such after a brief interval.

We can therefore conclude that Ss did have some ability to recognize the form of the sentence after 80 and 160 syllables of interpolated material, but that it was quite low and contrasted greatly with their memory for the semantic content of the sentence. Furthermore, almost all the loss had occurred before 80 syllables of interpolated material had been heard.

The pattern of recognition of changes from active to passive and vice versa (P/A) was quite similar to that for the other syntactic changes (Form). Analysis of variance revealed no significant differences between P/A and Form at any level of IM. The results of classification of changes also indicated that P/A changes were reacted to as formal changes. Of test sentences correctly judged as changed, 71% of P/A changes and 72% of Form changes were classified as "changes in form."

DISCUSSION

The findings indicate that recognition memory for the form of a sentence declines much more rapidly than recognition memory for the meaning. If the test sentence followed the original immediately, recognition was high for all sentences. After only 80 syllables of connected discourse had been interpolated between the original and test sentences, Ss' recognition of syntactic changes was close to chance. Changes in the meaning of a sentence, on the other hand, were recognized quite well, even though these changes were subtle. For example, the difference between "There he met an archaeologist, Howard Carter, who urged him to join in the search for the tomb of King Tut," and "There he met an archaeologist, Howard Carter, and urged him to join in the search for the tomb of King Tut" was recognized by seven of the eight Ss after 80 syllables of interpolated material. If the same original sentence was changed to "There he met an archaeologist, Howard Carter, who urged that he join in the search for the tomb of King Tut," only one out of eight Ss judged correctly after 80 syllables. Thus, very slight changes in the words of a sentence had vastly different effects on the experimental task, depending on whether or not the change affected the meaning.

The finding that changes which were classified as purely formal (test type Form) were not recognized well implies that "surface structure" is not stored long. In this experiment, changes from active to passive and vice versa were also used as recognition test sentences (test type P/A). The pattern of recognition for these changes was quite similar to that for the other syntactic changes, and the changes were most frequently classified by Ss as "changes in form." If we accept Katz and Postal's (1964) view that passivity is represented in the deep structure, the findings imply that what is in memory does not correspond exactly to linguistic "deep structure" either.

Although the experiment was designed as a comprehension test, Ss were instructed that recognition tests would follow the passages and they were presumably trying to remember the exact wording. Thus the performance achieved indicates an upper limit to the ability to recognize the form after the intervals tested. It is quite probable that in reality this ability is not often called upon. Since the primary goal was to discover what occurs in normal comprehension, the results indicate only that the formal structure is probably not stored for longer than the tested intervals. It may well be that the form is usually stored for a shorter interval.

The findings reported here are consistent with a theory of comprehension which contends that the meaning of the sentence is derived from the original string of words by an active, interpretive process. That original sentence which is perceived is rapidly forgotten, and the memory then is for the information contained in the sentence.

References

- Chomsky, N. *Aspects of the theory of syntax*. Cambridge: M.I.T. Press, 1965.
- Clifton, C., Kurcz, Ida, & Jenkins, J. Grammatical relations as determinants of sentence similarity. *J. verbal Learn. verbal Behav.*, 1965, 4, 112-117.
- Katz, J., & Postal, P. *An integrated theory of linguistic descriptions*. Cambridge: M.I.T. Press, 1964.
- Mehler, J. Some effects of grammatical transformation on the recall of English sentences. *J. verbal Learn. verbal Behav.*, 1963, 2, 346-351.

Note

1. Based upon a dissertation submitted to the University of California at Berkeley in partial fulfillment of the requirements for the Ph.D. degree. The research was supported in part by a National Science Foundation Predoctoral Fellowship. The author is indebted to Dan I. Siobin for suggestions and criticism. Some of the data were presented at the April, 1966 meeting of the Eastern Psychological Association meeting in New York, New York.

(Accepted for publication May 10, 1967.)