

RETROACTIVE INTERFERENCE IN THE RECALL OF SIMPLE SENTENCES

BY JACQUES MEHLER AND GEORGE A. MILLER

Harvard University

A list of English sentences was learned by the method of free recall. Performance was scored both for sentences recalled precisely and for sentences recalled in content regardless of their syntactic form. Interpolated lists of sentences differing from the original list in syntactic details or in semantic form were used to produce retroactive inhibition (RI). The RI for semantic content was negligible; when the interpolated list differed only in syntax and preserved the semantic content of the original list the effect was facilitatory rather than inhibitory. When performance was scored for syntactic as well as semantic aspects RI was appreciable. The results are interpreted as additional support for the hypothesis that the two aspects, semantic content and syntactic form, pose more or less distinct tasks for the learner.

In previous articles (Miller, 1962; Mehler, 1963) it has been argued, on the basis of errors observed in the free recall of sentences varying in both content and grammatical form, that when a person tries to memorize a sentence he deals with it in (at least) two separate parts—one carrying its meaning, or semantic content, the other carrying information about its syntactic structure. The present paper reports a further attempt to test the validity of that general distinction.

Since retroactive inhibition (RI) is ordinarily easier to demonstrate with nonsense syllables than with meaningful materials, and since, under the experimental conditions employed in these studies, the syntactic structure is relatively arbitrary, whereas the content of the sentence is meaningful, it seemed plausible that RI might affect these two aspects of the learner's task differentially.

The terminology of transformational grammar (Chomsky, 1957) will be used to describe the syntactic information that has to be remembered. The assumption is that most sentences are derived from a more fundamental type of sentence by certain syntactic rules called 'optional transformations'. The fundamental sentences, or *k-sentences* (derived from underlying kernel strings using only obligatory transformations), are simple, active, affirmative, declarative sentences, e.g., *The boy has hit the ball*. Three singular optional transformations are used in this experiment: the negative (*N*), the passive (*P*), and the interrogative (*Q*). Application of these three transformations in all possible combinations can produce seven additional sentences from each *k-sentence* (*K*).

The relations among these eight sentences can be represented by a cube, as in Fig. 1. Each vertex of the cube represents a separate syntactic form; each edge corresponds to a pair of sentences that differ only by a single transformation.

EXPERIMENTAL PROCEDURE

A familiar RI paradigm was used; original learning (OL) was followed by controlled intervening activity (IL), then by a retest (T) on the original material. The interpolated tasks differed for different groups of subjects, but the original learning and final test were the same for all.

Sets of eight sentences were presented to be learned by the method of free recall. Each sentence was typed on a separate 3 in. × 5 in. card and one card was presented manually every 4 sec.

Subjects were instructed to remember each sentence as precisely as they could, because a recall test would be given when the deck of cards was finished. After all eight sentences had been presented, the subject was asked to write down, in any order, as many sentences as he could remember. This procedure was repeated for a total of either two or four learning trials, depending on the experimental condition to which the subject had been assigned; each learning trial was followed immediately by a test of recall. The order of presentation of the sentences was varied for each trial.

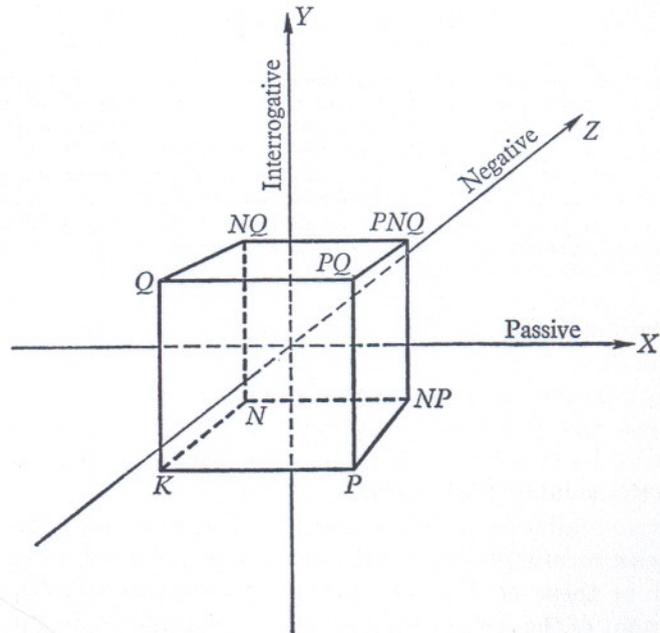


Fig. 1. Representation of the eight syntactic forms in terms of the corners of a cube, when the co-ordinates are labelled with the names of grammatical transformations.

There were three different interpolated tasks, one designed to produce syntactic interference, a second designed to produce semantic interference, and a third (control) intended to interfere minimally with the eight sentences originally learned. Each subject had only one kind of interpolated task. All of the experimental groups had four trials of IL.

Immediately after the interpolated task was completed, each subject was asked to recall all he could remember of the original eight sentences. When the subject reported that he had recalled as much as he could, he was told that if he thought that with more time he could do better, he could take two or three minutes more.

The 100 subjects used in this study were all undergraduates at Radcliffe College or Harvard University; they were paid \$1.50 per hour for their services.

Materials

The following list of eight sentences was used as the original learning task for all groups:

- The boy has taken the photograph.*
- The man hasn't bought the house.*
- The passenger has been carried by the airplane.*
- Has the girl worn the jewel?*
- The discovery hasn't been made by the biologist.*
- Hasn't the secretary typed the letter?*
- Has the tree been hit by the car?*
- Hasn't the essay been written by the student?*

The syntactic form of these sentences will be referred to respectively as *K*, *N*, *P*, *Q*, *PN*, *NQ*, *PQ* and *PNQ*.

Syntactic interference. Ten of the subjects in this group had two trials to learn the original material; thirty subjects had four trials. The interpolated material was derived by modifying the syntactic structure of the sentences learned originally.

The thirty subjects having four trials of OL were divided into three subgroups of ten each; the subgroups differed only with respect to the degree of syntactic similarity between the original and interpolated materials. The interfering materials were derived from the original list by rotating the cube in Fig. 1. For the first subgroup, each sentence in the intervening list differed by only one transformation from its corresponding sentence on the original list. Of the three lists that it is possible to derive in this way, only the one obtained by rotation around the *Y*-axis was used. For the second subgroup, each sentence in the intervening list differed by exactly two transformations from its corresponding sentence in the original list; the derived list was obtained by rotation of the original two steps around the *Y*-axis. And in the third subgroup, each sentence differed in all three transformations. The syntactic changes introduced in order to produce the interpolated lists are given explicitly in Table 1. These transformed sentences were presented for four trials in the same way as the original sentences had been.

Table 1. *Syntactic changes used to produce three types of interpolated lists differing only in syntactic form from the original list*

Original list	Interpolated list		
	1 step	2 step	3 step
<i>K</i>	<i>P</i>	<i>PN</i>	<i>PNQ</i>
<i>N</i>	<i>K</i>	<i>P</i>	<i>PQ</i>
<i>P</i>	<i>PN</i>	<i>N</i>	<i>NQ</i>
<i>Q</i>	<i>PQ</i>	<i>PNQ</i>	<i>PN</i>
<i>PN</i>	<i>N</i>	<i>K</i>	<i>Q</i>
<i>NQ</i>	<i>Q</i>	<i>PQ</i>	<i>P</i>
<i>PQ</i>	<i>PNQ</i>	<i>NQ</i>	<i>N</i>
<i>PNQ</i>	<i>NQ</i>	<i>Q</i>	<i>K</i>

Semantic interference. Ten of the subjects in this group had two trials to learn the original material; twenty subjects had four trials. All thirty subjects were then asked to learn a new set of eight sentences, completely unrelated to their original learning. These sentences, presented for four trials, in exactly the same way as the sentences that were learned originally, were:

- The doctor has treated the patient.*
- The child hasn't eaten the pie.*
- The money has been stolen by the burglar.*
- Has the candidate won the election?*
- The pipe hasn't been corroded by the water.*
- Hasn't the horse jumped the fence?*
- Has the deer been shot by the hunter?*
- Hasn't the card been punched by the machine?*

It will be noted that these eight sentences, like the original ones, included one representative of each of the eight syntactic structures represented in the cube in Fig. 1, but their semantic content is, of course, different.

Control. Ten of the subjects in the control group had two trials to learn the original material and twenty had four trials. Then, instead of learning a new list, these subjects added columns of 4-digit numbers (taken from a table of random numbers) for a period of 13 min. each. Subjects were told that they were being tested for mathematical ability, and were asked to make as few errors as possible. They were also required to keep a running total, since they were told that they might be stopped at any point and that their score would be the number of digits they had correctly added to that point.

SCORING PROCEDURES

Performance in this situation can be scored in several ways; two of them have been used and are described here.

Consider first the *overall sentence score*. A response sentence can, of course, be scored as completely correct only if it is a verbatim reproduction of a stimulus sentence. Any difference between the stimulus sentence and the recalled sentence can be counted as an error; the only exception (which occurred infrequently) was that the substitution of an obvious synonym was not scored as an error. These overall sentence scores constitute the basic data from the experiment.

A second type of score was used in order to distinguish between semantic and syntactic sources of error. We shall say that two sentences are derived from the same kernel strings if we can obtain one from the other by applying one or more optional transformations (or their inverses) to one member of the pair. (As before, if one word in one sentence is a synonym of the equivalent in the other sentence, the pair is still considered to have the same sentence content if the above definition holds.) Two sentences that are derived from the same kernel are said to share the same semantic content and to differ only in their syntactic features.

This definition enables us to separate the errors into two classes, semantic and syntactic, and so to define a *sentence-content* score. In this method of scoring, a response is considered correct if it shares the semantic content of the stimulus sentence, i.e., if it is correct overall or if it can be obtained from the stimulus sentence by the singularly optional transformations *N*, *P*, and *Q* or their inverses (or by synonym substitution where the case might arise).

Percentage RI was computed by the formula $100(T_c - T_e)/T_c$, where T_c and T_e are the retest scores for the control and experimental groups, respectively.

RESULTS

Before considering the results for the main experiment, let us first dispose of the three subgroups subjected to alternative types of syntactic interference. In the present experiment one might assume that the three rotations of Table 1 define equal intervals along some scale of syntactic similarity; since it is well known that the similarity of the original and interpolated materials is an important variable affecting RI, we might expect to find that RI is a function of syntactic similarity, so defined.

From the top half of Table 2, however, it can be seen that although the amount of overall RI is fairly large for all three subgroups, the differences among them are relatively small and not statistically significant ($\chi^2 = 2.7$). The differences that do exist may be attributable to differences in the levels of mastery of OL. In any case, there seemed to be no significant trend, so the results for the three subgroups were pooled into a single group of thirty subjects, hereafter called the syntactic interference group.

Now return to the main experiment, which was designed to test the hypothesis that, in the context of this experiment, sentence content is more resistant to RI than is syntactic form. If we consider first the overall scores in the lower half of Table 2 we see that syntactic interference produced more RI (48%) than did semantic inter

ference (18%), presumably because of the greater similarity between OL and IL for the syntactic group. In part, however, this difference is attributable to the difference in the level of mastery on the final (fourth) trial of OL; when we apply a constant of proportionality to equate the syntactic group to the other two groups on OL, the estimate of RI is reduced somewhat for the syntactic group, but is still larger than for the group exposed to semantic interference.

Table 2. Average overall and content scores (maximum score = eight sentences) on the final (fourth) trial of OL (4 OL) and on the retest (T) after IL, and percentage RI, for groups having syntactic, semantic, and neutral IL

Condition	(N)	Overall score			Content score		
		4 OL	T	% RI	4 OL	T	% RI
Syntactic interference							
1 step	10	4.4	2.5	55	7.8	7.8	2
2 step	10	4.8	3.9	30	7.3	7.4	7
3 step	10	4.4	2.4	57	7.7	7.5	6
Syntactic interference combined							
(Corrected*)	30	4.5 (5.9)	2.9 (3.8)	48 (32)	7.6	7.6	5
Semantic interference	20	5.8	4.6	18	7.5	7.0	13
Control group	20	6.0	5.6	—	7.5	7.9	—

* Corrected scores were obtained by multiplying observed scores by a constant of proportionality to equate this group with the others on OL.

When we consider the content scores, however, we see that there is less RI than for the overall scores, which would seem to provide evidence in support of the two-aspect hypothesis that the experiment was designed to test. This conclusion is questionable, however, on the grounds that by the fourth trial all the groups had learned the eight sentence contents almost perfectly, which may somehow have insulated them against the effects of IL. Certainly, if the hypothesis was correct, we should expect to find some facilitation in the content scores for the group with syntactically interfering IL, since the IL simply repeated the same sentence contents that had been presented during OL. Performance was so near perfect, however, that no such facilitation could be observed.

In order to get a clearer picture of the effects of IL on the content scores, therefore, we must consider the data reported in Table 3. Here the conditions are exactly the same, except that only two trials of OL (rather than four) were given. From the content scores in Table 3 we see that there was indeed facilitation, as expected, for the syntactic group. Indeed, the facilitation was sufficient to produce an apparent facilitatory effect in the overall scores for this group, in spite of the similarity of OL and IL. Again, however, there is an unfortunate sampling error that produced a difference in the level of mastery of OL; when the correction is made, as before, in the results for the group with syntactic interference, the magnitude of this facilitation is markedly reduced.

Table 3 also indicates that for the group with semantic interference RI was greater (38%) when both syntactic and semantic features were scored than when only the semantic aspects were considered (14%).

No intrusion of sentence contents from either list was observed with the semantic interference group.

Table 3. *Average overall and content scores (maximum score = eight sentences) on the final (second) trial of OL (2 OL) and on the retest (T) after IL, and percentage RI, for groups having syntactic, semantic, and neutral IL*

Condition	N	Overall score			Content score		
		2 OL	T	% RI	2 OL	T	% RI
Syntactic interference	10	3.7	2.8	-33	6.3	7.3	-49
(Corrected*)		(2.6)	(2.0)	(+5)	(5.2)	(6.0)	(-22)
Semantic interference	10	2.6	1.3	38	5.0	4.2	14
Control group	10	2.6	2.1	—	5.4	4.9	—

* Corrected scores were obtained by multiplying observed scores by a constant of proportionality to equate this group with the others on OL.

DISCUSSION

The results indicate that there is greater RI, under the conditions of these experiments, for the syntactic than for the semantic aspects of a sentence. This conclusion would appear to be consistent with the hypothesis that subjects deal with sentences on two distinct levels; the semantic aspects are coded separately and are relatively free of RI, whereas the syntactic aspects are (in this situation) much more arbitrary, harder to remember, and more subject to RI. The alternative hypothesis, that each sentence is an arbitrary string of symbols to be memorized without regard for syntactic form or semantic content, would be not adequate to predict the differences here obtained between the two methods of scoring.

The principal puzzle posed by these results appears in the overall scores for the group subjected to syntactic interference during IL. Why is the syntactic interference ineffective (or even facilitatory) after two trials of OL, but inhibitory after four trials? If formal similarity between OL and IL were the only factor at work, one would expect strong RI in both experiments.

This puzzle provokes the following speculation. If we accept the two-aspect hypothesis, it is but a short step to a two-stage hypothesis about the process of memorizing sentences. That is to say, all the evidence to date is consistent with the notion that subjects first learn the sentence content, then learn the syntactic details. According to this line of argument, during the first two trials our subjects were concentrating on remembering what the eight sentences were about, so an interpolated task that repeated those sentence contents would scarcely interfere, and might even prove helpful. By the fourth trial, however, subjects had mastered the contents of the sentences and were well into the second stage, trying to straighten out the syntactic details, so an interpolated task that repeated those contents would be of little help and the scrambled syntax could cause considerable confusion. This argument would explain the major discrepancy between the data of Tables 2 and 3.

This investigation was supported by funds granted to Harvard University, Centre for Cognitive Studies, from the National Institutes of Health, United States Public