COGNITIVE PSYCHOLOGY 2, 331–350 (1971)

The Abstraction of Linguistic Ideas[1]

JOHN D. BRANSFORD[2] AND JEFFERY J. FRANKS[3]
University of Minnesota

The phenomenon of "idea acquisition and retention" is demonstrated experimentally and contrasted with an "individual sentence memory" point of view. Results indicate that during an acquisition phase of the experiments, Ss spontaneously integrate the information expressed by a number of nonconsecutively experienced (but semantically related) sentences into wholistic, semantic ideas, where these ideas encompass more information than any acquisition sentence contained. Ss' subsequent attempts to recognize those exact sentences heard during acquisition are shown to be a function of the complete ideas acquired. Thus, Ss are most confident of "recognizing" sentences expressing all the semantic relations characteristic of a complete idea, in spite of the fact that such sentences expressed more information than was communicated by any single sentence on the acquisition list. Ss become less confident of having heard particular sentences as a function of the degree to which a sentence fails to exhaust all the semantic relations characteristic of a complete idea.

In recent years many psycholinguistic studies dealing with the relation between language and memory have been conducted. Most of these have dealt with memory for sentences and have looked at the effect of various aspects of sentence structure on what is learned and stored in memory. Questions about the relation between syntactic structure and memory have received the most attention. Johnson (1965), for example, has demonstrated some effects of phrase structure on the recall of sentences. Savin and Perchonock (1965) suggest that increasing syntactic complexity produces increasing strain on short-term memory, and long-term memory studies like that of Mehler (1963) indicate a trend towards

[1] Preparation of this paper was supported in part by grants to the University of Minnesota, Center for Research in Human Learning, from the National Science Foundation (GB-17590), The National Institute of Child Health and Human Development (HD-01136 and HD-00098), and the Graduate School of the University of Minnesota. The authors thank J. J. Jenkins, J. R. Barclay, J. Glick, J. Yellott, and A. Yonas for their advice and criticisms of previous versions of this paper. Also, the authors are indebted to the editors of *Cognitive Psychology* for their helpful advice.

[2] Present address: State University of New York at Stony Brook, Stony Brook, N. Y. 11790.

[3] Present address: Vanderbilt University, Nashville, Tenn.

syntactic simplicity when sentences which have been stored for some time are recalled.

Other memory studies have stressed semantic rather than syntactic variables. Sachs (1967), for example, shows that information about the particular syntactic form of a sentence is quickly forgotten, while its semantic content is very well retained. The work by Kolers (1966) with bilinguals also demonstrates the primacy of semantic encodings, although in his study the effect is demonstrated at the level of individual words. Studies dealing with semantic variables have been much less frequent than those dealing with syntactic variables, presumably because linguistic theories provide a much better description of the syntactic than the semantic domain.

Irrespective of whether emphasis is placed on syntactic or semantic variables, all of the above studies have an important aspect in common. They all deal with memory for individual items. That is, they all study memory for individual sentences or individual words. The primary concern of the present paper is not with memory for individual words or sentences; rather it is with memory for wholistic, semantic ideas. Wholistic ideas need not be communicated by single sentences. They may result from the integration of information expressed by many different sentences experienced successively and often nonconsecutively in time. Emphasis on the acquisition and retention of wholistic ideas thus focuses on memory for sets of sentences expressing common semantic content. The purpose of the present paper is to discuss a methodology for studying the phenomenon of idea acquisition and to demonstrate the psychological reality of "inter-sententially defined" ideas.

THE EXPERIMENTS

The studies to be presented below were designed to communicate four different ideas to each subject, where each idea could be exhaustively characterized as those semantic relations contained in a single complex sentence (e.g., *The rock which rolled down the mountain crushed the tiny hut at the edge of the woods*). During an acquisition phase of the experiments, Ss were never presented with sentences expressing the complete complex ideas, however, but only with sentences encompassing various subsets of the four different semantic domains (*The rock crushed the tiny hut; the hut was at the edge of the woods, etc.*). Idea acquisition would be demonstrated to the extent that such an acquisition procedure resulted in Ss acquiring the complete ideas defined by the integration of the information contained in related sentences.

The experiments were designed to demonstrate the fact of idea acquisition and retention in as strong a manner as possible. Thus they sought

to demonstrate that Ss not only could acquire the complete ideas from exposure only to partial ideas, but also that the acquisition of ideas is so natural and compelling that Ss would actually think they had heard sentences expressing the complete ideas during acquisition when in fact they had not.

In order to test the hypothesis of idea acquisition, a recognition test was administered immediately following the acquisition procedure. Ss were told that they would be read a set of sentences, all of which were very related to those just heard during acquisition. Their task was to decide which exact sentences they had heard during acquisition, which ones they had not, and how confident they felt about their answers. Recognition sentences included sentences actually heard during acquisition (OLD sentences), sentences not actually heard during acquisition but which were consonant with the general ideas expressed there (NEW sentences), and sentences neither heard during acquisition nor consonant with the ideas presumably acquired (NONCASE sentences).

To the extent that Ss acquired the complete ideas during acquisition, the following results were expected: First, Ss should show evidence of productivity. That is, they should think they recognize novel examples of the ideas acquired during acquisition in spite of the fact that they had not heard these sentences before. Some of these novel sentences contain combinations of relations never expressed by any single sentence presented during acquisition. It is especially important to see if Ss think they recognize these, since such recognition could not be accounted for by memory for any single previously experienced sentence. It is also important to see whether Ss can differentiate novel sentences from those actually heard during acquisition.

Assuming that Ss actually acquire the complete ideas, some additional results might be expected with respect to Ss' confidence ratings for having heard certain sentences. It seems reasonable to expect that these ratings will reflect the degree to which a sentence represents what was learned during acquisition. If Ss did indeed acquire the ideas, those complex sentences expressing the complete ideas might be expected to receive the highest confidence ratings. Confidence ratings might then decrease with the degree to which particular sentences fail to exhaust all the semantic relations characteristic of a complete idea. Of course, Ss should be confident that they have not heard NONCASE sentences, since these express meanings which differ from those presumably acquired.

An additional factor was incorporated into the acquisition procedure of the present experiments in order to rule out pure contiguity of sentences as an explanation of the results. Sentences related to each of the four ideas were presented randomly during acquisition, with the

constraint that no sentences related to the same idea occurred consecutively on the acquisition list. Therefore, if evidence can be found that Ss acquired the complete ideas, they must have done this by integrating successive but nonconsecutive instances of the ideas. In addition, the acquisition procedure was in the form of a short-term memory task requiring Ss to remember a sentence long enough to answer a question about it after a 5-sec delay. Throughout acquisition, Ss were not told that they would be asked to perform on a later recognition task. If information from various nonconsecutive sentences is integrated, it will be done without explicit instructions telling Ss what to do.

Within the limits of the present experimental procedures, the strongest demonstration of idea acquisition and retention would be as follows: Ss would "recognize" novel examples of previously acquired ideas and would not be able to discriminate these novel sentences from ones previously heard. The only criterion affecting recognition confidence ratings for a given sentence would be the degree to which it exhausted all the relations of a complete idea as a whole.

EXPERIMENT I

Since the first two studies are closely related, we shall present the methodologies for both of them before discussing any results.

Method

Subjects

The Ss were 15 University of Minnesota undergraduates enrolled in introductory psychology courses.

Materials

Materials consisted of a set of English sentences constructed in the following manner: (1) Four complex sentences were constructed, each of which exhaustively represented the semantic information in one of the four ideas to be acquired. (2) Each complex sentence (complete idea) was constructed to represent the relations among four simple declarative sentences. These simple sentences were chosen intuitively without special regard for their theoretical status in existing linguistic theories. (3) Each of the four complex sentences was broken down into its four simple declaratives. These simple sentences were then recombined in a variety of ways. Thus, the complete set of sentences used consisted of the following: (*a*) the four complex sentences (FOURS); (*b*) the four simple sentences of each of these complex sentences (ONES); (*c*) sentences constructed by combining (embedding) two simple sentences from a particular complex sentence (TWOS); and (*d*) sentences constructed by combining (embedding) three simple sentences from a particular complex sentence (THREES). The four complex sentences used were: Idea A—*The ants in the kitchen ate the sweet jelly which was on the table;* Idea B—*The warm breeze blowing from the sea stirred the heavy evening air;* Idea C—*The rock which rolled down the mountain crushed the tiny hut at the edge of the woods;* Idea D—*The old man resting on the couch read the story in the newspaper.*

An example of a complete set of sentences defining one particular idea is given in Table 1. The sets of sentences for the other three ideas were constructed in an analogous manner. Thus, the complete set of sentences for each idea contains one FOUR, three THREES, four TWOS, and four ONES.

Materials also included a set of cards, each of which contained four colors arranged in various orders. These were used in an intervening color-naming task.

TABLE 1
Sentences Comprising Idea-Set A

FOUR:	The ants in the kitchen ate the sweet jelly which was on the table. (On Recognition Only)
THREES:	The ants ate the sweet jelly which was on the table. (On Acquisition Only) The ants in the kitchen ate the jelly which was on the table. (On Acquisition Only) The ants in the kitchen ate the sweet jelly. (On Recognition Only)
TWOS:	The ants in the kitchen ate the jelly. (On Acquisition Only) The ants ate the sweet jelly. (On Both Acquisition and Recognition) The sweet jelly was on the table. (On Recognition Only) The ants ate the jelly which was on the table. (On Recognition Only)
ONES:	The ants were in the kitchen. (On Acquisition Only) The jelly was on the table. (On Acquisition Only) The jelly was sweet. (On Recognition Only) The ants ate the jelly. (On Recognition Only)

Procedure

The Ss were divided into two groups for purposes of counterbalancing presentation order. There were eight Ss in Group I and seven in Group II. Each experimental session consisted of acquisition and recognition trials.

Acquisition. The Ss were told that their task would be to answer questions about sentences which would be read by the *E*. This acquisition procedure was as follows: (1) *E* read a sentence; (2) All Ss, in unison, named four colors in the order in which they appeared on the card held up by the *E*; (3) The *E* read an elliptical question concerning the sentence just read; (4) Ss wrote down the answer to the questions. This procedure continued for all the sentences on the acquisition list. The intervening task of color naming was imposed so that Ss would be required to hold each sentence in memory for a short time (color naming took about 4 sec).

Examples of possible elliptical questions (for the example sentence *The rock rolled down the mountain*) are as follows: *Did what?*, *What did?*, and *Where?*. However, only one question was asked for each sentence. Questions were chosen so that each constituent of each idea was questioned about as often as each other constituent.

The acquisition list consisted of 24 sentences, six from each of the four different idea sets. The acquisition sentences from each set consisted of two ONES, two TWOS, and two THREES. (For example, the acquisition sentences from Idea Set A can be found in Table 1.) Acquisition sentences were chosen so that, as a group, they exhausted the information characteristic of each idea.

The order of presentation of the 24 acquisition sentences was arranged so that in each successive sequence of four sentences there was one sentence from each of the four different idea sets. Sentences were randomized within each block of four sentences with the constraint that no two sentences from the same idea set occurred consecutively on the list. The ONES, TWOS, and THREES from each idea set were randomly distributed across the full acquisition list.

For Group I, acquisition sentences were presented in the order 1–24; for Group II the order was 24–1. Each group went through the acquisition list once.

During acquisition, Ss were not told that there would be a second part to the experiment (i.e., recognition).

Recognition. Following acquisition, Ss were given a 4–5 min. break. They were then told that the *E* was now going to read a new set of sentences, all of which were closely related to the set of sentences they had just heard. Their task was to indicate which of the sentences in the new set they had actually heard before and which ones they had not. In addition to making "yes" or "no" ratings (indicating whether or not they felt they had heard a particular sentence before), Ss were asked to indicate how confident they were about their answer. A 5-point confidence scale was provided for this purpose which ranged from "very low" to "very high" confidence.

The recognition list consisted of 28 sentences. All 28 sentences were from the original four idea sets. Twenty-four of these sentences were NEW sentences; that is, they had not been presented in acquisition. There were six of these sentences from each of the four idea sets. Each of these groups of six sentences contained two ONES, two TWOS, one THREE, and one FOUR (the only four). For example, those sentences marked Recognition in Table 1 are the sentences used in recognition for Idea Set A.

In addition to the 24 new sentences, four sentences from the acquisition list were included in recognition (OLD sentences). These include two Ones (*The rock rolled down the mountain* and *The breeze was blowing from the sea*), one TWO (*The ants ate the sweet jelly*), and one THREE (*The old man resting on the couch read the story*).

The order of presentation of the 28 sentences was similar to that of acquisition. Thus, each block of four sentences contained one sentence from each of the four different idea sets. The four OLD sentences were randomly assigned positions in this list. Again the list was constrained so that no two sentences from the same idea set were consecutive. ONES, TWOS, THREES, and FOURS were randomly distributed throughout the list.

Each group was given the recognition list twice with no break between the two presentations. For Group I, the order of presentation was 1–28; 1–28. For Group II, the order was 28–1; 28–1.

EXPERIMENT II

Experiment II is essentially a replication of Experiment I, except that this study contains certain NONCASE sentences, the importance of which will be discussed below.

Method

Subjects

The Ss were 16 University of Minnesota undergraduates enrolled in introductory psychology courses.

Materials

The construction of the sentences paralleled that of Experiment I. The four complex sentences or complete ideas (FOURS) used were: Idea E—*The scared cat running from the barking dog jumped on the table;* Idea F—*The old car pulling the trailer climbed the steep hill;* Idea G—*The tall tree in the front yard shaded the man who was smoking his pipe;* and Idea H—*The girl who lives next door broke the large window on the porch.*

The ONES, TWOS, and THREES were constructed as in Experiment I. In addition to all these sentences, six NONCASES were constructed. These contained information present in the four idea sets, but their composition violated relationships represented in the ideas to be learned. There were two types of violation: (1) One NONCASE had the same "units" as one of the FOURS, but the relations were changed (i.e., *The scared cat ran from the barking dog which jumped on the table*). (2) The other five NONCASES were constructed by combining information *across* rather than within idea sets, thus grossly changing the relationships involved. These NONCASES are as follows: *The old man who was smoking his pipe climbed the steep hill; The tall tree in the front yard shaded the old car; The barking dog jumped on the old car in the front yard; The scared cat which broke the window on the porch climbed the tree; The man who lives next door broke the large window on the porch.* Each NONCASE sentence represents four simple sentences, precisely the same number that are represented in the four complex FOURS.

Procedure

The procedure again included acquisition and recognition. Two groups were used. There were seven Ss in Group I and nine in Group II.

Acquisition. As in Experiment I, the acquisition list contained 24 sentences, six from each idea set. These were chosen, randomized, and presented exactly as in Experiment I (including the use of the intervening color-naming task).

Recognition. The recognition procedure paralleled that of Experiment I. The list consisted of 24 new sentences, six from each of the four idea sets. As in Experiment I, the list contained two ONES, two TWOS, one THREE, and one FOUR from each idea set. The order of presentation and randomization procedure paralleled Experiment I. One of the six NONCASES was presented within each block of four "clearcase" sentences (i.e., those that actually belong to one of the four idea sets). No sentences from acquisition appeared on recognition. Thus, there were a total of 30 sentences on the recognition list. Each group was presented

with this list twice. The presentation order for Group I was 1–15, 16–30; 15–1; 30–16. For Group II, the order was 30–16, 15–1; 16–30, 1–15.

Results

Data were analyzed as follows: Ss' ratings were converted into numerical values. A "yes" response received a "plus"; a "no" response received a "minus." A "very high" confidence rating received a 5, a "high" confidence rating received a 4, and so on down to a 1 for "very low" confidence. Thus, a 10-point rating scale emerged ranging from plus 5 to minus 5 (excluding zero).

Ratings for each sentence by each S were summed algebraically for recognition trials I and II. The mean rating for each sentence per trial was then computed, as well as the mean rating per sentence summed over both trials I and II. All data are reported in terms of means (which, of course, must fall within the range of +5.0 to −5.0).

Figure 1 illustrates data which are representative of results from both experiments. This table contains recognition sentences presented for Idea Set A. The three numerical values above each sentence represent mean recognition ratings for (*a*) trials I and II, (*b*) trial I, and (*c*) trial II, respectively. Sentence number 4 is an OLD sentence; all others in Figure 1 are NEW.

The first important point about the data is that many NEW sentences (i.e., sentences not on acquisition) received positive recognition ratings, indicating that Ss actually thought they had heard these sentences during the acquisition task (e.g., see sentences 1, 2, and 5 in Fig. 1). In both Experiments I and II, NEW FOURS, THREES, and TWOS generally

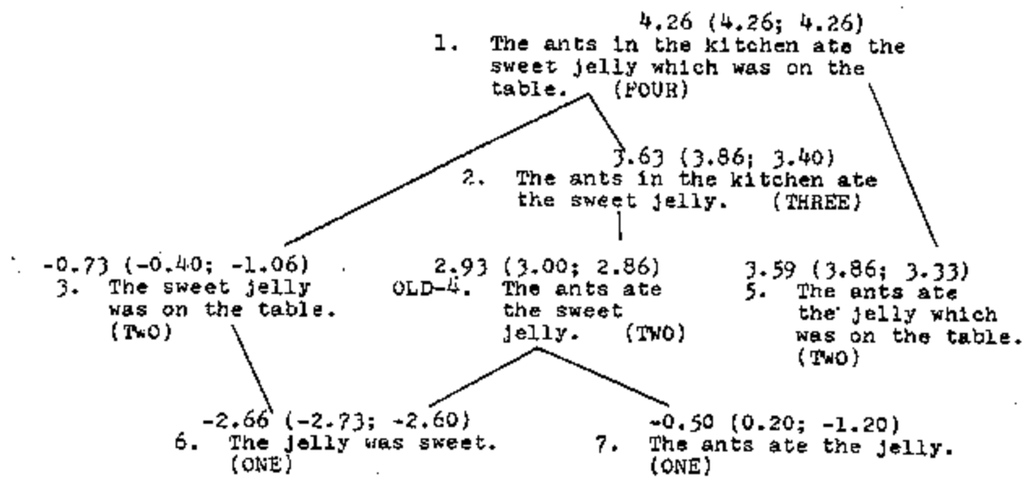

FIGURE 1

received such positive ratings. ONES tended to vacillate between the low positive and negative range. All NEW FOURS and THREES received positive recognition ratings. This latter result is especially important, since those sentences contained combinations of relations which Ss had never experienced in any single acquisition sentence. It is especially impressive that Ss thought they recognized FOURS.

The second point to consider about the data concerns ratings for OLD sentences. In Experiment I, four OLD sentences were presented during recognition; an example is sentence 4 in Figure 1. Note that NEW sentences 1, 2, and 5 (in Fig. 1) received higher recognition ratings than the OLD sentence. This result is typical of the other idea sets. Many NEW sentences received higher recognition ratings than OLD sentences Ss had actually heard before. In fact, 7 (of 24) NEWS received higher recognition ratings than the highest ranking OLD and 15 NEWS were higher than the lowest OLD. If Ss remembered those sentences heard during acquisition, OLD sentences should have received higher confidence ratings than all NEW sentences. Data clearly indicated, however, that OLD sentences did not receive the highest ratings on the recognition list. (Additional OLD–NEW data will be discussed in Experiment III below.)

The third point to consider about the data concerns the relationship between the number of semantic propositions comprising a sentence and recognition confidence ratings; that is, consider the relative recognition ratings for FOURS, THREES, TWOS, and ONES. Results of both experiments were very similar, and their combined averages are summarized in Fig. 2 (the means are represented independently for trials I and II). Recognition ratings clearly ordered FOURS > THREES > TWOS > ONES, and the NONCASE sentences from Experiment II received the lowest ratings of all. It is possible to analyze the data in a much more sensitive manner than is represented by overall mean ratings, however. This more sensitive method is as follows.

Figure 1 illustrates a sample of the data from the two experiments. It shows a sentence-by-sentence analysis of the recognition ratings received by all sentences in Idea Set A. Such an analysis can be applied to all the data for both experiments. It allows one to look at the effects of the number of semantic propositions comprising a sentence under conditions where differences in semantic content are controlled.

The sentence-by-sentence analysis examines ordinal recognition ratings among various FOURS, THREES, TWOS, and ONES. The ordinal comparisons of interest are those among sentences within each idea set. Hence no comparisons are made between sentences from Idea Sets A and B (for example). In addition, valid comparisons can only be made between two

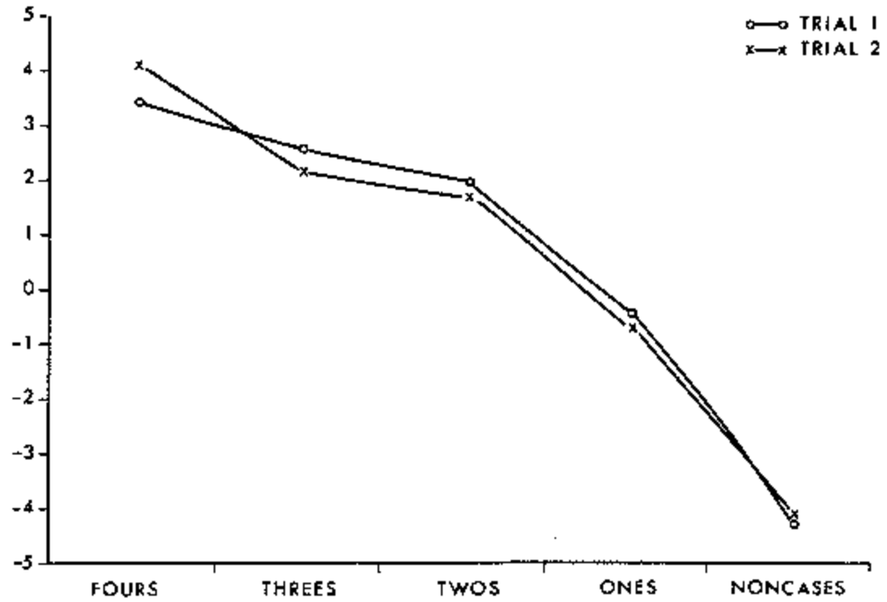

FIGURE 2

sentences with equivalent qualitative content. That is, they must be identical except for the absence of one or more embedded semantic propositions. For example, a sentence like *The ants in the kitchen ate the sweet jelly* (Fig. 1, sentence 2) can be compared with an identical sentence minus the phrase *in the kitchen* (Fig. 1, sentence 4). It cannot, however, be compared with a sentence like *The ants ate the jelly which was on the table* (Fig. 1, sentence 5) even though the first sentence is a THREE and the latter a TWO. The reason for this constraint is that the more complex sentence contains no phrase *which was on the table,* whereas the shorter sentence does (but is missing portions like *in the kitchen* as well). Hence the two sentences are not adequately equivalent as defined above. This a priori constraint on equivalence is very important for valid interpretation of the data. Thus, if the phrase *which was on the table* was not acceptable to Ss, the TWO sentence above might receive lower ratings than the THREE for reasons other than the fact that the two sentences differed in the degree to which they exhausted all the relations characteristic of the complete idea. The general formula for defining valid comparisons is as follows: a sentence X is predicted to receive a higher recognition rating than sentence Y if X contains all the basic propositions in Y plus one or more additional propositions (assuming, of course, that sentence X is not a NONCASE). The valid predictions among sentences in Idea Set A allowed by these constraints are presented immediately below Fig. 1.

A set of predictions such as $a > b > c > d$ is considered to be transitive. Thus, the prediction includes comparisons between $a > c$, $a > d$, and

$b > d$ as well. In Experiment I, 47 of 49 such ordinal predictions were in the predicted direction (for means summed over trials I and II). Experiment II yielded 39 out of 41. Table 2 summarizes the ordering results. Due to the lack of independence inherent in such transitivity predictions, plus the fact that the degree of interdependence differs for each idea set, there are no statistical tests applicable to the present results. Therefore, a Monte Carlo technique was used to analyze the data. The Monte Carlo program randomly assigned a set of ranks to all recognition sentences and checked to see how many of the appropriate predictions that FOUR > THREE > TWO > ONE came out by chance. This procedure was carried out 1000 times. The result was a frequency distribution specifying how many times the number of predictions confirmed by the data (or more) came out by chance. For both studies I and II, the probability values were $<< .001$.

The final set of data to be considered are ratings for the NONCASES presented in Experiment II. These contained combinations of relations which were not consonant with any of the ideas presumably acquired during acquisition. Data indicated that Ss were quite confident that they had not heard these NONCASE sentences before as can be seen in Fig. 2. Recognition ratings for the six NONCASES ranged from —2.37 to —4.87, and their overall mean was —4.19. With one exception, the distributions for individual NONCASE and CLEARCASE sentences were nonoverlapping (for means summed over trials I and II). The exception was that a ONE received a slightly lower recognition rating than the highest ranking NONCASE.

Throughout the present Results section the overall mean ratings for each sentence have been considered to be the most important data. Each sentence was presented twice during recognition in order to achieve a less noisy picture of the results. This procedure eliminated, to some extent, the effect of successive sentences upon one another, and it helped to control for momentary quirks in *Ss'* ratings. Most important, it allowed Ss to become better acquainted with the complex scaling procedures and makeup of the recognition list. When one examines the

TABLE 2
Number of Ordering Predictions Confirmed

	Trials 1 and 2	Trial 1	Trial 2
Expt. 1	47 of 49	43 of 49	47 of 49
Expt. 2	39 of 41	35 of 41	39 of 41
Expt. 3	—	135 of 140 (G I) 134 of 140 (G II)	—

means for the two recognition trials computed separately, the same basic patterns of results nevertheless emerge. Product-moment correlation coefficients between mean values on trials I and II were .87 for Experiment I and .95 for Experiment II. In each study, the first trial means contained six reversals in ordinal predictions compared to two reversals when means were computed across both recognition trials (see Table 2). For each study the first trial results were still highly significant by the Monte Carlo test ($p < .001$). For the second trials in both studies, Table 2 shows that the means ordered just as well as the overall means (i.e., only two reversals per study). Evidently Ss were able to discriminate between having heard a sentence before during *recognition* and thinking they had heard it during acquisition, otherwise one might expect second trial ordering results to be worse than those of the first trial. For example, NONCASES might be expected to receive very high ratings on the second trial because Ss would actually have heard them before; similarily for ONES, TWOS, etc. In actuality, however, ratings for both trials were generally comparable. Average NONCASE values, for example, were —4.20 for trial I and —4.17 for trial II.

Discussion

Results of Experiments I and II are very congruent. Both show evidence of productivity, in that Ss thought they recognized many NEW sentences that were never actually presented during the acquisition task. In addition, recognition confidence ratings covaried with the number of semantic propositions comprising a sentence: for appropriate comparisons, FOURS > THREES > TWOS > ONES. NONCASE sentences received very low recognition ratings; Ss were quite confident that they had not heard them before.

So far there has been only a small amount of data comparing OLD and NEW sentences. Experiment III investigated this variable in more detail.

EXPERIMENT III

Method

Subjects

Ss were 50 University of Minnesota undergraduates enrolled in introductory psychology. They were divided into two groups (G I and G II). G I was composed of three subgroups a, b, and c, of 10, 7, and 9 Ss, respectively. G II was composed of subgroups d, e, and f of 9, 9, and 6 Ss, respectively.

Materials

The construction of experimental materials paralleled that of Experiments I and II. The four complex FOURS were: Idea A—*The rock which rolled down the mountain crushed the tiny hut at the edge of the woods;* Idea B—*The warm breeze blowing from the sea stirred the heavy evening air;* Idea C—*The scared cat running from the barking dog jumped on the table;* Idea D—*The tall tree in the front yard shaded the man who was smoking his pipe.* The ONES, TWOS, and THREES were constructed as in the first experiments. For each of the four idea sets there were 4 ONES, 4 TWOS, 3 THREES, and 1 FOUR.

Procedure

Ss were run in small groups (subgroups a–f). The procedure used was identical to Experiments I and II.

Acquisition. Two sets of acquisition sentences were used. G I received acquisition list 1. G II received acquisition list 2. Acquisition set 1 contained 24 sentences, 6 related to each of the ideas A, B, C, and D. Acquisition set 2, also composed of 24 sentences, contained the other 6 sentences related to each idea. Thus, there was no overlap between the two acquisition sets in terms of the exact sentences used. Table 3 shows the number of each type of sentence (ONE–FOUR) used to construct these acquisition sets. Acquisition sentences were randomized and presented as in Experiments I and II.

Recognition. The recognition set for both G I and G II consisted of all 48 sentences, 12 related to each of four different ideas. Thus, for G I (which received acquisition set 1) half of the recognition sentences were NEW and the other half were OLD. For G II, the opposite halves were NEW and OLD. Recognition sentences were randomized as in the first two experiments and three different recognition orders were used, with one order being presented to subgroups a and d, a second to

TABLE 3
Composition of Acquisition Sets

	Idea Sets							
	A		B		C		D	
	Acq. 1	Acq. 2	Acq. 1	Acq. 2	Acq. 1	Acq. 2	Acq. 1	Acq. 2
FOURS	1	0	0	1	1	0	0	1
THREES	1	2	2	1	1	2	2	1
TWOS	2	2	2	2	2	2	2	2
ONES	2	2	2	2	2	2	2	2

subgroups b and e, and the third to subgroups c and f. Since there was such a large number of recognition sentences, each was presented only once. Ss were asked to recognize which exact sentences they had heard during acquisition and to assign a confidence rating to their answers, exactly as in Experiments I and II.

Results

Recognition confidence ratings were converted into numerical values exactly as in the previous experiments. The mean rating for each of the 48 recognition sentences was computed independently for G I and G II. Since Groups I and II received identical recognition sentences, a number of different comparisons was possible. For example, one could compare ratings for each individual sentence when it was OLD versus NEW.

First consider the OLD versus NEW differences. What effect does actually experiencing a sentence on acquisition have on ratings in the recognition task? Imagine what the recognition effects should be if Ss could clearly discriminate between OLD and NEW sentences. This would occur, for example, if Ss actually stored representations of those exact sentences experienced during acquisition. In this case all OLD sentences should be assigned higher recognition ratings than NEW sentences. That is, Ss should be able to partition the set of recognition sentences into two nonoverlapping subsets defined in terms of OLD versus NEW. Figure 3 shows a hypothetical set of data (averaged over all FOURS, all THREES, all TWOS, and all ONES) illustrating perfect OLD versus NEW discrimination. Note that these hypothetical data take into account

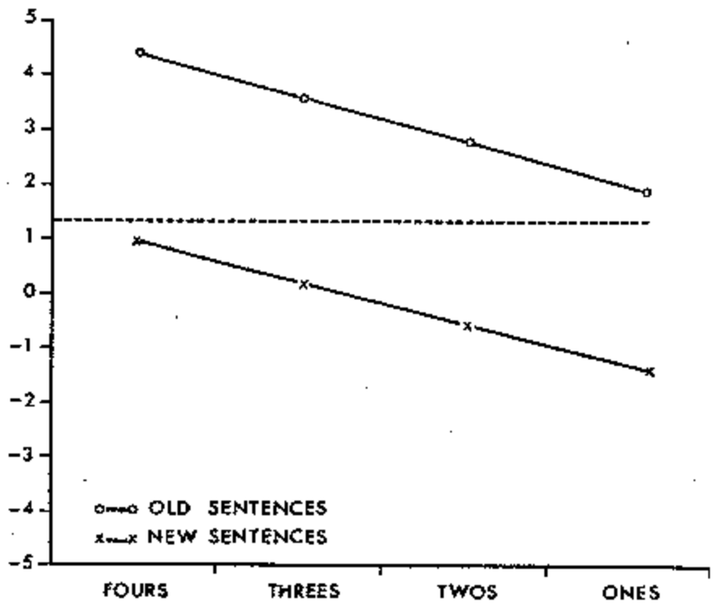

FIGURE 3

the ordering effect found in Experiments I and II (i.e., for the appropriate comparisons, FOURS > THREES > TWOS > ONES), although the particular numerical values represented in the graph are arbitrarily selected. The important point is that, for this hypothetical situation, the lowest rating given an OLD sentence should be higher than the highest rating given a NEW sentence. Recognition ratings for all NEW sentences should fall below the dotted line.

Figure 4 shows the data actually obtained in the present experiment. Mean recognition ratings for FOURS, THREES, TWOS, and ONES were 3.95, 2.53, 1.50, and —1.11 for OLD sentences and 3.89, 2.91, 0.95, and —2.46 for NEWS. It is very clear that Ss did not even come close to being able to partition the recognition set into two nonoverlapping subsets of OLD versus NEW sentences. New sentences frequently received recognition ratings which were as highly positive as (or even higher than) OLD sentences, just as was found (albeit with only a small sample of cases) in Experiment I.

The fact that Ss could not partition the present set of recognition sentences into two nonoverlapping subsets of OLD versus NEW sentences does not necessarily mean that a sentence's presence on acquisition resulted in absolutely no memory effect on recognition. It might merely mean that the difference between recognition ratings for OLD versus NEW sentences was simply not very large. A more sensitive test of OLD–NEW differences is one that takes advantage of the fact that for each individual sentence, either G I or G II received that sentence during acquisition (an OLD sentence for one group) and the other group did

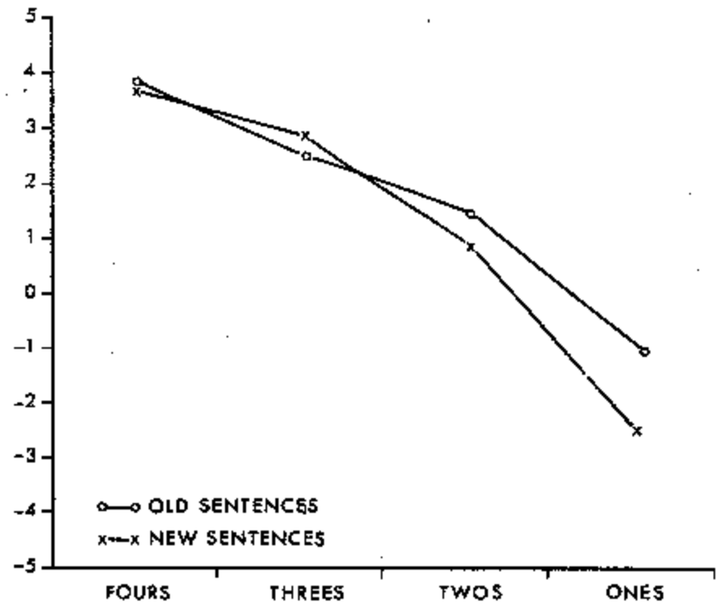

FIGURE 4

not (a NEW sentence for the other group). Thus, one can compute two mean recognition ratings for each sentence, one for when it was NEW, one for when it was OLD.

A comparison of OLD versus NEW ratings was made for each individual sentence. For FOURS, two sentences received higher ratings as OLD than as NEW sentences, but for the other two FOURS these orderings reversed (i.e., NEWS got higher ratings than OLDS). The THREES exhibited a similar trend: in 5 of 12 possible comparisons OLD sentences received higher ratings than NEW sentences. For the other 7 sentences, however, NEWS beat the OLDS. Whether a sentence received a higher rating as an OLD or a NEW entry thus appeared to be entirely a matter of chance. At the level of TWOS, OLD sentences received higher ratings than NEW sentences in 11 of 16 comparisons. The proportion is still within the realm of chance by a sign test ($p > .10$), although it may be indicative of a trend. For ONES, however, OLD sentences received higher ratings than NEW sentences in 14 of 16 comparisons. This value is significant by a sign test, $p < .002$. Thus there is a relatively slight but reliable effect for OLD sentences at the level of ONES.

Even though there was a slight specific memory effect for ONE sentences, the amount of variance this effect accounted for was extremely small. For example, the OLD–NEW variable had little effect on the ordering data that FOURS > THREES > TWOS > ONES, even when comparisons involved a NEW sentence being predicted to receive a higher rating than an OLD. Considering the ordering data for G I, for example (which received acquisition list 1), it was possible to make 140 ordering predictions among the sentences (note that half of these sentences were OLD and the other half NEW). The same predictions could be made for G II, except that the opposite sentences were OLD and NEW. (Note that predictions were made only between sentences with "equivalent semantic content" as defined in Experiments I and II.) For G I the number of predictions confirmed was 135 of 140. The comparable data were 134 of 140 for G II (see Table 2). Of the 11 total reversals for these two groups (5 for G I and 6 for G II) only 4 of these reversals were from predictions that a NEW sentence would receive a higher rating than an OLD sentence. Other reversals occurred between two OLD sentences, two NEW sentences, or when an OLD sentence was predicted to be greater than a NEW sentence. In short, overall ordering data were not adversely affected by the variable OLD versus NEW. Note also that these ordering data replicated Experiments I and II.

At a more specific level of analysis, rank ordering reversals should

have been most likely to occur when NEW TWOS were predicted to receive higher ratings than OLD ONES (since specific memory effects were greatest for ONES). There were 32 valid predictions that NEW TWOS would receive higher ratings than OLD ONES: 30 of these 32 predictions were confirmed. Thus, in spite of the fact that OLD ONES received higher ratings than NEW ONES, OLD ONES nevertheless received lower ratings than NEW TWOS.

GENERAL DISCUSSION

Results of Experiment III were consistent with those of the first two experiments. Data indicate that Ss acquired something more general or abstract than simply a list of those sentences experienced during acquisition. Ss integrated the information communicated by sets of individual sentences to construct wholistic semantic ideas. Memory was a function of those ideas acquired during acquisition. Ss thought that they "recognized" novel sentences (NEWS) consonant with the ideas abstracted but were quite confident that they had not heard NONCASE sentences that were not derivable from the ideas acquired. Ss were most confident of having heard those sentences expressing all the semantic information characteristic of the complete ideas acquired during acquisition; and for appropriate comparisons, confidence ratings ordered FOURS > THREES > TWOS > ONES. The fact that Ss "recognized" NEW THREES and FOURS was especially important, since these sentences contained combinations of semantic relations never expressed in any single acquisition sentence. The information encompassed by NOVEL THREES and FOURS could only have been acquired by integrating information across various acquisition sentences experienced nonconsecutively in time.

Whether a sentence was OLD or NEW had very little effect on recognition ratings. In general, Ss could not discriminate novel sentences consonant with the ideas acquired during acquisition (NEWS) from sentences that they had actually heard during the acquisition task (OLDS). Only at the level of ONES was there a slight recognition advantage for OLD sentences, yet OLD ONES nevertheless received lower ratings than NEW TWOS. In general, Ss did not store representations of particular sentences. Individual sentences lost their unique status in memory in favor of a more wholistic representation of semantic events.

The results of the three experiments suggest a strong, reliable phenomenon, and the experimental technique appears promising for studying the abstraction of linguistic ideas. The present paper merely scratches the surface of the problem of linguistic abstraction, however, and additional data are needed before more precise claims can be made

about the phenomenon at hand. A very important problem, for example, concerns the question of *what is learned* in the above situations. How can one characterize the nature of the semantic ideas that are acquired? The fact that NONCASE sentences received highly negative recognition ratings indicates that the ideas Ss acquired encompassed a considerable degree of semantic precision. For example, Ss were not simply basing their recognition ratings on identities of individual words. However, there are still a number of alternate characterizations of what is learned that could account for the present results. The problem of precisely specifying what is learned is too complex to be handled in a short discussion section, however. A forthcoming paper will deal with this issue in detail.

Besides the problem of what is learned, many boundary conditions for the present phenomenon are still in question. For example, what effects will different types of acquisition instructions have on the results? In the present experiments we purposely used instructions encouraging semantic processing (as opposed to instructions emphasizing rote memory for individual items, for example), since we wanted to study memory as it is generally manifest in everyday life. In a recent Ph.D. thesis, Curnow (1969) has shown (among other things) the basic phenomenon to be replicable under a variety of different acquisition conditions. There are many different questions about the effect of instructions that Curnow did not have a chance to consider, however, hence additional research still needs to be done.

Other boundary conditions concern the composition of the acquisition list and its effects on whether Ss will actually think they have heard complex sentences. If all acquisition sentences were ONES, for example, one would not expect Ss to think they had heard sentences that were THREES and FOURS. It seems clear that any adequate account of the more general phenomenon of linguistic abstraction will have to postulate at least two relatively independent memory representations: (1) Ss will remember wholistic semantic structures, and (2) Ss will retain information about the general style in which the semantic information was originally expressed. An acquisition list composed entirely of ONES may be sufficient to allow Ss to integrate complex semantic structures, but memory for the general style of the acquisition sentences (i.e., that they were all extremely short and simple) would most likely cause Ss to reject recognition sentences that were THREES and FOURS (and maybe even TWOS). Experiments designed to separate memory for input style from memory for semantic structures are currently being conducted. A subsequent article will examine this research.

Although the experiments reported in this initial paper leave many

questions unanswered, we are hopeful that the experimental techniques will allow investigation of the phenomenon of abstraction in considerable detail. Ultimately we hope to be able to characterize the semantic structures abstracted from exposure to connected discourse, and hence lend some precision to Bartlett's (1932) notions of abstract schemas as *what is learned.*

REFERENCES

BARTLETT, F. C. *Remembering.* Cambridge: Cambridge Univ. Press, 1932.

CURNOW, P. F. Integration of linguistic materials. Unpublished Ph.D. Dissertation, University of Minnesota, 1969.

JOHNSON, N. F. The psychological reality of phrase-structure rules. *Journal of Verbal Learning and Verbal Behavior,* 1965, 4, 469–475.

KOLERS, P. A. Interlingual facilitation of short-term memory. *Journal of Verbal Learning and Verbal Behavior,* 1966, 5, 314–319.

MEHLER, J. Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior,* 1963, 2, 346–351.

SACHS, J. D. S. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics,* 1967, 2, 437–442.

SAVIN, H. B., & PERCHONOCK, E. Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior,* 1965, 4, 348–353.

(Accepted March 1, 1971)

COGNITIVE PSYCHOLOGY 2, 351–368 (1971)

Perceptual Learning and the Theory of Word Perception[1]

ELEANOR J. GIBSON
Cornell University

Perceptual learning involves the learning of distinctive features and higher-order invariants, learning progressing actively toward the most economical features and structure. Features of words are classified as phonological, graphic, semantic and syntactic. Features of these classes are processed independently and sequentially. Ordering of priorities changes with development, and shifts strategically with the demands of the task. Evidence is presented for priority differences for each class of feature depending on task differences.

This paper is the outcome of two long-time endeavors of the author—the development of a theory of perceptual learning, and a program of research on reading. The aim is to try to show how the two are related. First, the theory of perceptual learning will be described, as briefly as possible. It attempts to answer three questions: First, what is learned? Second, how? What are the processes involved? Third, what is the motivation and reinforcement for perceptual learning?

WHAT IS LEARNED?

I believe that what is learned in perceptual learning are *distinctive features* of things, of representations of things, and of symbolic entities like words; also the *invariants* of events that occur over time; and finally the economical *structuring* of both. I think the information for learning these is potentially present in stimulation, to be picked up by the observer given the proper conditions for it.

Consider some examples. Sets of distinctive features characterize objects and entities both natural and artifactual—the furnishings of the world, such as people, dwelling places, things to eat; and, particularly relevant for the present topic, symbols written on pieces of paper, like letters and words. The set of letters of our alphabet is characterized by a set of distinctive features, which in different combinations permit a unique characterization of each one. My students and I have spent

[1] Invited address for Division 3, American Psychological Association, 1970. This work was supported in part by a grant (USOE OEG-2-9-420446-1071-010) to the author from the Office of Education.