JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 16, 77–90 (1977)

Interpreting Anaphoric Relations: The Integration of Semantic Information while Reading

SIMON GARROD AND ANTHONY SANFORD

Glasgow University

To understand fully the pair of sentences "*A bus came roaring round the corner*; *The vehicle narrowly missed a pedestrian*," the reader has to deduce that the vehicle in question is a bus which came roaring round the corner. In Experiment 1 we show that the reading time for the second sentence in such a pair is in part determined by the semantic distance between the two items to be integrated (vehicle and bus in this case). This result suggests that the information from the two sentences is integrated at the time of reading. Experiment 2 replicates the semantic distance effect in a situation where the two sentences are separated in the text. In Experiment–3 and 4 it is shown that the effect can be abolished under conditions where the two items appear in unrelated phrases. On the basis of these results a model of textual comprehension is proposed.

In order to understand a passage of prose fully it is necessary to be able to integrate information from several different sentences in the passage. This integration usually involves appreciating relationships between various objects, people, or events mentioned in the text. For instance, if a passage were to contain the two sentences,

(1) A bus came roaring round the corner;
(2) The vehicle nearly flattened a pedestrian,

it would be necessary to know that the vehicle in question was a bus which came roaring round the corner. In other words the phrase "the vehicle" does not refer to a vehicle in the usual sense of description but rather serves the function of identifying a particular vehicle from the preceding context (Strawson, 1964). The way in which this identifying function is signaled is through the definite article.

The comprehension of anaphoric relations of this sort raises two types of problems (Stenning, Note 5). The first might be described as a rhetorical problem: a problem of deciding which pairs of phrases are to be related as antecedent and anaphor. In the example above this reduces to the problem of identifying the phrase "A bus" in sentence (1) as being antecedent to the phrase "The vehicle" in (2). The second problem is one of representation: of how the information from the two sentences is represented semantically once these antecedent–anaphor pairs have been identified.

In this paper we will be concerned primarily with how the process of identifying antecedent–anaphor relations enters into the comprehension of sentences in text. It could be argued that antecedent–anaphor relations are ignored until after the sentence has been interpreted. For instance, on encountering a sentence such as (2) the reader might retrieve information about vehicles and incorporate this into his semantic representation of the sentence without appreciating that the vehicle is in any way related to the bus mentioned in sentence (1). Any integration of the material from both sentences would only occur after the sentences had been interpreted. A more

Part of this work was supported by a grant from the British Social Science Research Council.

Send reprint requests to Simon Garrod, Department of Psychology, Adam Smith Building, Glasgow University, Glasgow, G12 8QQ, Scotland.

ISSN 0022-5371

plausible assumption is that the antecedent–anaphor relation is identified at the time the sentence is interpreted, thus making it possible for the reader to create an integrated representation of the text directly. One way in which this might come about is by treating anaphoric phrases like bound variables (see Harman, Note 2). In order to come up with a semantic representation of a sentence such as (2) it would be necessary to assign some variable to the phrase "the vehicle" which would be bound by the quantifier in the antecedent phrase. Under most conditions this variable could be determined by finding an antecedent phrase related to the same object in the text. However, when no such antecedent is found some default variable would have to be set up before any representation could be assigned to the sentence. As a result of such a process sentences (1) and (2) might be represented as:

A vehicle (a bus which came roaring round the corner) nearly flattened a pedestrian.

In this account the process of identifying the antecedent–anaphor relation would be obligatory for interpreting the sentence, since the outcome of that interpretation is an integrated representation of both sentences.

There is already some evidence that people form an integrated representation of sentences read in text. For instance, Bransford and Franks (1971) have shown that under memorization conditions subjects tend to integrate information from several previously presented sentences. However, in this study there was a long delay between the subjects' reading the sentences and their being tested for memory, so that it is not clear whether integration occured at the time of reading or resulted from the recall process. In order to determine this, it is necessary to discover whether the identification of antecedent–anaphor relations actually enters into the reading process.

The rationale behind the experiments reported here is that if integration occurs at the time of reading one would expect any manipulation which affects the identification time to have a comparable effect on reading time for that sentence. For identification to occur in the sentences given above we need to be able to say that the vehicle and the bus are the same thing, and in order to say this, we need to access the fact that buses are vehicles.

There is evidence from a variety of sources that the time taken to retrieve or utilize such class-membership information can be systematically manipulated. Battig and Montague (1969) produced norms for a situation in which subjects were asked to generate short lists of exemplars of prespecified classes. Some words were listed by many subjects (High conjoint frequency) and others by few (Low conjoint frequency exemplars). Low and high conjoint frequency items differ in the speed with which they are accepted as members of their class in a reaction time (RT) task, high conjoint frequency exemplars being accepted more rapidly (Rosch, 1973; Wilkins, 1971). Thus, in situations where an *overt evaluation* of class membership is required, the time taken to perform the task shows an effect of conjoint frequency.

We suggest that the presence of such an effect might be used as an *indicator* to show when the identification takes place in the two-sentence situations outlined above. Low conjoint frequency exemplars of the category should be identified as possible antecedents less rapidly than high conjoint frequency examples (for instance, *bus* may be identified rapidly, while *tank* may take longer). By comparing suitably constructed sets of sentences, it should be possible to examine this point. Should a conjoint frequency effect emerge, we would assume identification was taking place.

Although the use of other indicators of the same order may be entertained, we feel the conjoint frequency effect is the most useful. Formal measures of semantic distance based on nested categories (Collins & Quillian, 1970), or varying membership acceptance time in terms of category size (Landauer & Meyer, 1972) both depend upon the control of the conjoint frequency of the items for their use, and may even be due to the conjoint frequency effect itself (e.g., Smith, Shoben, & Rips, 1974; also Conrad, 1972).

A potential problem with this method is that the RT differences between high and low conjoint frequency exemplars have only been demonstrated in situations where class membership decisions have to be made *overtly*. What evidence there is suggests that such RT differences vary in magnitude with the task in which they are measured (Sanford & Seymour, 1974; Sanford & Garrod, Note 3). For this reason it is of interest to determine whether the conjoint frequency effect can be obtained under natural conditions of reading to understand, a necessary prerequisite to using the effect to draw conclusions about the identification of antecedent–anaphor relations.

Experiment 1

This experiment was primarily designed to test whether conjoint frequency had any effect on the reading time for a sentence requiring identification. However, two types of antecedent–anaphor relations were considered. In the first, the anaphoric phrase contains the category of which the antecedent is an exemplar. An example of this type is shown below, with both high and low conjoint frequency exemplars in the first sentence.

(3) A (Robin, HCF) / (Goose, LCF) would sometimes wander into the house.

(4) The *bird* was attracted by the larder.

In the second, the anaphoric phrase contains the exemplar, and the antecedent the category as in (5) and (6).

(5) A *bird* would sometimes wander into the house

(6) The (Robin, HCF) / (Goose, LCF) was attracted by the larder.

These two types will be referred to as category-last and instance-last, respectively.

The reason for considering both the category-last and instance-last cases was that the conditions for identification in the two cases are different. In the more usual category-last case the category in the second sentence is already entailed by the instance in the first, i.e., the reader knows that the robin in sentence (3), is a bird before reading sentence (4). This situation does not hold for the instance-last case: the reader has to deduce from the text that the bird in sentence (5) is the robin in sentence (6).

Method

Materials and design. The sentences used were constructed around a set of categories and exemplars derived from the Battig and Montague (1969) norms. There were 16 categories, each having a high and a low conjoint frequency exemplar associated with it which were matched as closely as possible in terms of word length (letters and syllables) and word frequency. The high conjoint frequency exemplars were drawn from the most common examples possible from the norms, and the low ones were as rare as possible, while allowing the matching procedure to be carried out. All of the exemplars were true members of the prespecified class. (This is not true of all cases in the norms, for example, "worm" appears as an instance of insect. Such cases were not used in the present study.) The categories and instances had been used in a variety of previous studies involving an overt class membership decision, and gave different acceptance times in such a task, high conjoint frequency being reliably faster by the *min F′* statistic (Clark, 1973). The magnitude of the conjoint frequency effect obtained in this way was about 100 milliseconds upwards (Sanford & Garrod, Note 3) and so it was thought the materials should prove reasonably sensitive.

Two groups of sentences were constructed using thses 16 categories and 32 instances, i.e., as two sets of 32 sentence-pairs. In one group the *instances* appeared in the second sentence, and in the other, the *categories* appeared second. After each pair of sentences a question was added, to check that the subjects had understood the sentences. Half of the questions were arranged so that they required integration of the facts in the statement sentences in order to be answered. The remainder did not, but could be answered on the basis of one or the other sentence alone. Whereas half of the questions referred to the instance, the other half referred to the category. The questions were kept simple, a typical combination being:

(7) The pear was a new variety.
(8) The fruit was sweeter than ever.
(9) Was the fruit a new variety?

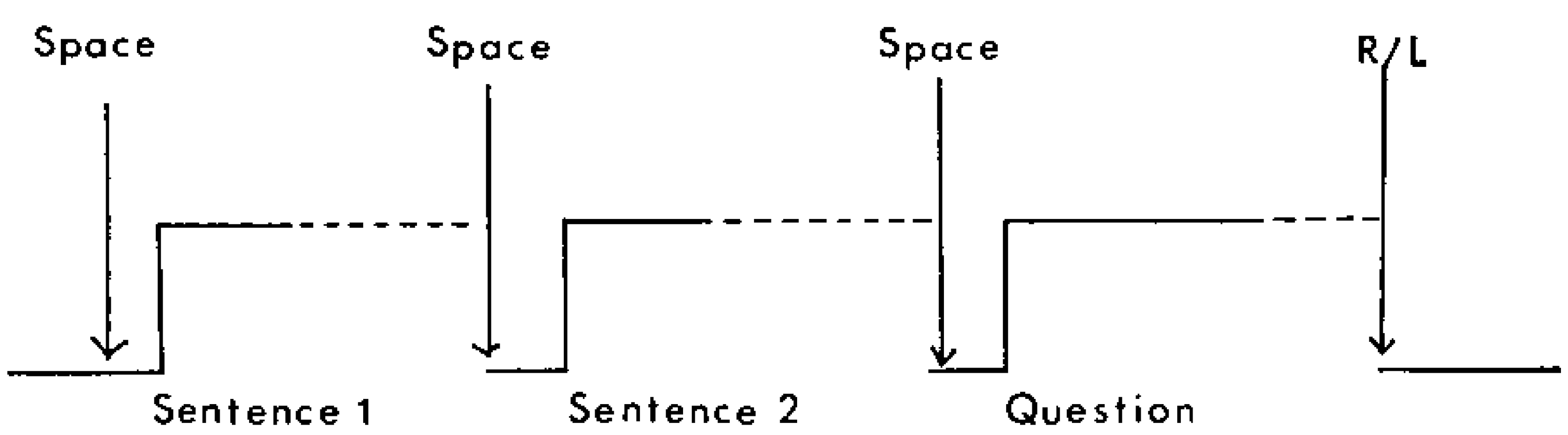
The experiment was run with two independent groups of 12 subjects each. One group saw only the *instance-last* combinations, the other only the *category-last.*

Procedure. In order to measure reading times for each of the sentences presentation of materials was through a teletype coupled to a NOVA 2/10 computer. The structure of any given trial follows the outlines of Figure 1.

At the onset of each trial, the subject pressed the space-bar on the teletype. At this point the three sentences (two statements and a question) were printed out on three separate lines, one above the other. This output could not be seen by the subject, since the upper part of the teletype carried a cover obscuring the paper. When the subject was ready, he pressed the space-bar, and by way of a line-feed, the first sentence became immediately visible. The subject read this, and then pressed the space-bar to request the next sentence, which was put into the subject's line of vision by means of another line-feed. Again, after reading it, the space-bar was pressed by the subject, and the question was then shown in the same way. The subject pressed a key of his choice on the right hand side of the teletype keyboard if the answer to the question was "Yes". He pressed a key of his choice on the left if the answer was "No". The reading times for *each* sentence and the response time for the question were recorded automatically. Every effort was made so that the subject would be comfortable, with his thumbs resting readily on the space-bar and his two forefingers resting on suitably comfortable keys.

In summary, the subject began by requesting the three sentences (space-bar). After they had been printed out (out of his view), he called them up one at a time (space-bar), and answered the question (left or right keys). A few minutes of practice on this enables subjects to read the sentences in a comfortable, smooth flow, which was obviously required if reliable measurements were to be made.

After 12 practice trials each subject was given two blocks of 16 experimental trials. Each block consisted of 8 low and 8 high conjoint frequency exemplars; if a sentence pair contained a high exemplar in the first block it would contain the low exemplar in the



SENTENCE 1: A carnation had won a prize

SENTENCE 2: The flower was the biggest in the show

QUESTION : Was the carnation very small?

FIG. 1. Sequence of events in a typical trial in Experiment 1.

second. The order of presentation of the two blocks was counterbalanced across subjects.

Results and Discussion

The mean reading times for the second sentences in the two main conditions and the two conjoint frequency levels are shown in Table 1.

The mean times for each subject in each condition were calculated, and an analysis of variance was carried out on this data with subjects treated as a random factor. A similar analysis of variance was also carried out on the mean time for each set of materials, with materials treated as a random factor, enabling the calculation of the *min F′* statistic.

The general findings from this analysis confirm the reliability of trends revealed in Table 1. There was a main effect of conjoint frequency, with $min\,F'(1, 27) = 7.601, p < .025$, low conjoint frequency sentences requiring longer reading times than high sentences by about 80 milliseconds. A main effect of order (instance- or category-last) was obtained also, with $min\ F'(1,25) = 4.253$, $p < .05$. The second sentence in the category-last case is read more rapidly than in the instance-last case. The interaction of order and conjoint frequency did not approach significance with *min F′* and the individual *F*s by subjects and by words all having a value of less than 1 (standard error (SE) for the interaction = 14.5 milliseconds).

It may be objected that the faster reading time for the category-last case results from the fact that subjects saw the identical second sentence twice in this condition, as compared to only once in the instance-last case. To eliminate this possibility, the mean reading times for the second sentence in the two order conditions were calculated on the basis of the first block of the trials only. Under these circumstances, the measure is based only on the subject's first encounter with the second sentence. Again, the category-last order gave the faster reading time, by some 305 milliseconds, an effect which was statistically reliable, with $min\ F'(1, 26) = 6.264, p < .025$.

TABLE 1

READING TIMES FROM EXPERIMENT 1 FOR THE SECOND SENTENCES AS A FUNCTION OF ORDER (INSTANCE-LAST OR CATEGORY-LAST) AND CONJOINT FREQUENCY[a]

	Conjoint frequency			
	High	Low	Difference	S.E.
Instance-last	1554	1623	*69*	18
Category-last	1320	1401	*81*	18.5

[a] Reaction time in milliseconds.

The presence of the conjoint frequency effect in the reading times for the second sentence is consistent with the claim that identification occurs at the time of reading. In turn this would suggest that the information from the two sentences is integrated at the time of reading the second sentence. Further support for the integration argument comes from an analysis of the question times. To answer half of the questions it was necessary to integrate information from both sentences. For instance to answer question (9) it is necessary to integrate the information about the pear *and* the fruit. On half of the trials this would involve coordinating a high conjoint frequency exemplar with its category and on the other half a low conjoint frequency one. It is therefore possible to compare the times for answering both types of questions. Such a comparison provides no evidence of a conjoint frequency effect (mean for low conjoint frequency questions = 1.874 seconds; mean for high conjoint frequency = 1.914 seconds), indicating that any integration of the information has already occured before the question is encountered.

The presence of the main effect for the instance-last versus category-last comparisons is in accord with the different requirements for identification in these two conditions. Whereas in the category-last condition the category gives no further information about the nature of the antecedent, in the instance-last condition the instance does give extra information. It is, therefore, suggested that the difference between the two conditions is a reflection of the fact that more information has to be incorporated into the representation that the subject has of the text in the instance-last case than in the category-last. This point will be taken up in the final discussion.

A serious objection that could be made to our interpretation of the conjoint frequency effect for the instance-last condition is that we were comparing reading times for sentences that were in fact different and perhaps it is due to this difference that the effect emerged. For instance, it might be that when the sentences contained low conjoint frequency items they were less plausible than when they contained high conjoint frequency items. The low conjoint frequency sentence: *The tank nearly flattened the pedestrian*, is perhaps less plausible than the matching high sentence: *The bus nearly flattened the pedestrian*, and perhaps it is this which leads to the reading time advantage for high conjoint frequency sentences in this condition.

A comparable objection can be raised about the reading time difference in the two order conditions. It might simply take longer to read the sentences containing the instances than the sentences containing the categories irrespective of the context. In order to take account of these problems it was therefore decided to run an extra control condition to assess the reading time for the second sentences when presented in isolation. The same procedure and order of presentation was used as in the main experiment. However, the subjects were only presented with the *individual* second sentences, each followed by an appropriate question. Twelve subjects were given the two blocks of instance sentences (16 in each) as in Experiment 1 and another 12 were given one block of category sentences (all 16 categories). It was thus possible to obtain an estimate of reading time for the high versus low conjoint frequency sentences in isolation directly from the first group of subjects and an estimate of the instance versus category sentence times by comparison of the first block times for the first group with the times for the second group.

An estimate of reading time from the first control group yielded a mean reading time for high conjoint frequency sentences of 1.750 seconds versus a mean time of 1.744 seconds for the low conjoint frequency sentences. Therefore, there was an overall slight advantage of 6 milliseconds in favor of the low sentences (SE = 35 milliseconds). In fact, of the 12 subjects run in this control 7 were faster, on the average, reading the low conjoint frequency sentences, whereas 5 were faster on the high sentences.

Comparing the results for the first block of trials in this group with the matched block for the second group yielded a mean instance sentence reading time of 1.805 seconds versus a mean category reading time of 1.933 seconds. When read in isolation the sentences containing categories took *longer* by 188 milliseconds (SE = 80 milliseconds). In neither case can the effects observed in the main experiment be attributed to any peculiarity of the sentences used; these effects could only be due to the particular anaphoric conditions set up in the experiment.

Taken together the results have bearing on both the rhetorical and representational problems brought up in the introduction. The presence of the conjoint frequency suggests that identification of antecedent–anaphor pairs occurs at the time that the sentence is interpreted and involves making an implicit class membership judgement, whereas the difference between the instance-last and category-last conditions suggests that subsequent to identification new information is being incorporated into some general representation of the text.

Experiment 2

In the first experiment we observed that integration of two sentences took place when the sentences were presented consecutively. Do the same processes occur when there are intervening sentences? It could be argued that the two sentence task is a special case, firstly because the subject does not have very much information to integrate, and secondly because the anaphoric sentence follows on immediately from its antecedent.

The second experiment was desinged with a view to overcoming these two objections. Consider the following section of text;

(10) A *vehicle* came roaring round the corner.
(11) It had had a brake failure.
(12) The *bus* nearly flattened a pedestrian.

With these sentences, (11) adds no further information about the nature of the vehicle. We know it had a brake failure, but we do not know anything else about it. Not until the third sentence do we find out that it is a bus.

If integration occurred in this situation we would expect to find a comparable conjoint frequency effect for sentence (12) whether it was put in the second or third position in the sequence. Since the process of identifying the relationship between antecedent and anaphor should involve an implicit check of the instance being a member of the specified category the conjoint frequency effect should obtain whatever the spacing. Indeed, it may be even *more* difficult to discover the relation of the category to the instance with a greater number of intervening sentences and this could conceivably lead to an *increased* conjoint frequency effect, on the grounds that task difficulty increases the effect.

We might also expect to find an overall time difference for reading the sentence in the two positions. This could either be because the greater distance between the two sentences increased overall identification time or because by the time the third sentence is encountered the reader is having to hold more information in memory. If such a difference emerged it would be possible to test each explanation by comparing the reading times for the "It" sentences in either of the positions. Although the pronoun "It" in sentence (11) itself identifies an antecedent phrase this would always be from the previous sentence. If (11) occured in second position it would refer to the vehicle, in third position it would refer to the bus. Thus any position effect for the "It" sentences could not be attributed to separation from the antecedent sentence.

Method

Stimulus materials. These were essentially the same as those used in the previous study. The same pairs of sentences were made up, but only the category-first order was used. To each pair an additional "It" sentence was added which could either refer to the instance or the category depending on what position it held in the sequence. The "It" sentence did not, however, give any additional information about the nature of the referent. A new set of questions was made up which included questions about the "It" sentences.

Design and procedure. Each subject was presented with all the materials. For half of the categories, a particular subject received the "It" sentence in the third position, e.g.,

A vehicle came trundling down the hill.
The bus nearly flattened a pedestrian.
It had had a brake failure.
Did the vehicle have a brake failure?

For the other half, the "It" sentence was presented in the second position, so that the category and instance sentences were separated.

The categories allocated for each condition were counterbalanced across subjects. Apparatus, practice, instructions, etc., were identical with those in the previous study.

Subjects. The subjects were 12 undergraduates and nonacademic university staff, paid at the rate of £0.50 per session.

Results and Discussion

Reading times for the "instance" sentences and the "It" sentences are shown in Table 2. Analyses of variance carried out on the mean times for the "instance" sentences revealed that the conjoint frequency effect was significant by subjects ($F(1, 11) = 21.109, p < .001$) and by words ($F(1, 15) = 4.949, p < .05$). This is marginally reliable in terms of *min F′*, with *min* $F'(1, 21) = 4.00$ (critical for $p = .05$ is 4.3). Although the sentence in the third position is read a little slower this effect is only reliable by subjects ($F(1, 11) = 4.77, p < .05$) and not by materials or *min F′*. The interaction of conjoint frequency with position is totally unreliable by all measures. (SE = 32 milliseconds).

The presence of a comparable conjoint frequency effect for sentences in both second and third positions indicates that identification is occuring in the three sentence situation. It also runs against the view that the effect is due to a generalized associative priming of the sort demonstrated by Meyer, Schvaneveldt and Ruddy, (1975). If this had been the case, we should have observed a reduced conjoint frequency effect for sentences in position three. In fact, the effect is numerically greater for sentences in the third position, though not significantly so.

To interpret the weak main effect for position we need to look at the reading times for "It" sentences in both positions. Analysis of variance on these times reveals a reliable position effect (*min* $F'(1, 25) = 12.5, p < 0.005$). This would indicate that it is not the separation of the anaphoric phrase from its antecedent that causes an increase in reading time but rather the position of the sentence within the sequence. Although it is tempting to attribute this to some overall increase in memory load as the amount of text encountered goes up, it could just as well be due to the fact that the third sentence immediately precedes the question and subjects might hesitate momentarily before calling it up.

Experiment 3

In the first two experiments a conjoint frequency effect was present in a situation where a category and instance both refer to the same thing in the text. It was suggested that the effect was the result of a process whereby the

TABLE 2

Reading Times from Experiment 2 for the Instance Sentence in Position Two or Three as a Function of Conjoint Frequency[a]

	Conjoint frequency				
Position taken	"It"[b]	High	Low	Difference	Standard Error
Sentence (3)	1743	1551	1611	*60*	34
Sentence (2)	1493	1682	1781	*99*	40
Combined		1616	1696	*80*	18

[a] Reaction time in milliseconds.
[b] Corresponding "It" sentences.

reader checks the relationship between category and instance in order to assign some referent to the anaphoric phrase. Is it possible that a similar checking procedure will occur when the category and instance do not actually refer to the same thing in the text? Any predictions that one might make for the non-coreferential situation would depend upon the origin of the checking procedure.

In the introduction we assumed that the checking procedure was triggered by the presence of syntactic cues which indicated the presence of an antecedent phrase which was retrievable from the text. On discovering the cues the reader would then search back through his representation of the text to try and find something which fit this. In the particular examples that we have considered it would be the definite article which serves as a cue for the presence of an antecedent. There are, however, a wide variety of different syntactic devices for indicating this, which have been discussed in detail elsewhere (Halliday, 1967, Stenning, Note 4). If the checking procedure were triggered in this fashion we would not expect to see a conjoint frequency effect in a situation where a category mentioned in one sentence could not refer back to an instance in a previous one, since under these conditions the syntactic cues would not be present.

There are, however, alternatives to the process outlined above. For instance, it is also possible that the reader is actively searching subsequent sentences for information that could be related to what has already been mentioned in the text. We might expect such a system to be looking for information relevant to the current topic of discourse for instance. If this sort of process were operating we could well observe a conjoint frequency effect even in a non-coreferential situation.

The difference between the two systems is a difference of priority for semantic and syntactic processes. If the syntactic analysis is prior to any semantic look-up we would expect the first system, whereas if semantic and syntactic analysis occur concurrently we might expect something similar to the second. Accordingly, an experiment was run to investigate conjoint frequency in a non-coreferential situation, using a similar procedure to the experiments already mentioned.

Method

Two sets of materials were generated. In one set 16 pairs of sentences were made up which could contain either a high or low conjoint frequency instance in the first sentence and the category in the second. However, they were written in such a way that the category *could* not refer back to the instance. This was done in two ways. First each of the second sentences began with a pronoun which referred back to the previous instance, and, second the category was always in an indefinite noun phrase. An example of such a pair is shown below:

(13) A (Bus)/(Tank) came roaring round the corner.

(14) It nearly smashed some *vehicles*.

(15) Did the bus smash some vehicles?

For the other set there were 16 second sentences in which the category did refer back to the instance. To match the non-coreferential sentences the category always occurred in grammatical object postion. The corresponding pair of coreferential sentences is shown below.

(16) A (Bus)/(Tank) came roaring round the corner.

(17) A pedestrian was nearly killed by the *vehicle*.

(18) Was the pedestrian nearly killed by the bus?

For both sets of sentences a set of questions was generated as with the previous studies.

The experiment was run with two independent groups of 12 subjects each. One group saw only non-coreferential sentence pairs, the other only coreferential pairs. The 24 subjects were all undergraduates or nonacademic staff at Glasgow University and each was paid

£0.50 per session. The apparatus, procedure, etc., were the same as those used in the previous studies.

Results and Discussion

The mean reading times for both experimental conditions are shown in Table 3. The times for each condition were analyzed separately because the sentences in the two conditions were not comparable. An analysis of variance on the reading times for coreferential sentences revealed a significant conjoint frequency effect, with *min* $F'(1, 25) = 6.95$, $p < 0.025$ (SE = 18.8 milliseconds) as was anticipated. However, analysis of the non-coreferential times revealed that the reduced conjoint frequency effect was unreliable by subjects with $F(1, 11) = 2.55$, $p < 0.1$ (SE = 21.5 milliseconds) but significant by materials, with $F(1, 16) = 4.71$, $p < .05$ (SE = 15.02 milliseconds). This leads to an insignificant *min* F'. Whereas all 12 subjects in the coreferential situation showed an overall positive conjoint frequency effect, only 9 of the 12 showed it in the non-coreferential condition.

In this way the results from the non-coreferential group are not clear cut. They indicate that there is some sort of checking process going on (producing a conjoint frequency effect) but that perhaps this is optional, and does not always occur, or does not occur as often as it does in the coreferential case. By optionality we simply mean that when checking goes on the subject decides whether or not the individual in sentence one is a member of the class specified in sentence two. For instance, with sentences (13) and (14) the subject may decide that *the bus* is a *vehicle* but not the "some vehicles" being referred to. This involves following up the check. There is an alternative psychological explanation of the result. It could be argued that the conjoint frequency effect does not completely originate in a checking procedure at all, but at least to a degree results from some "priming" effect at the word level. There is now ample evidence that it is easier to read a single word if it has been preceded by an associate (Meyer et al., 1975) and that preceding sentences made it easier to read words related to concepts which have been introduced already (Kennedy, 1975). Since high conjoint frequency exemplars may prime their categories more than low exemplars (which may be thought of as weaker associates) the conjoint frequency effect could conceivably emerge from differential priming (Sanford & Garrod, Note 3).

We may suggest that the results of Experiment 2 run against a very simple version of this argument, since there is evidence (Meyer et al., 1975) that with an increase in the delay between two associates, or with an intervening activity, the resultant priming effect diminishes. This did not happen in our experiment, even when subjects were processing another sentence during the delay. However, it would be naive to suppose this to be the only type of word priming which might occur. It may be the case that when a concept is currently being used in working memory, other associated

TABLE 3

Reading Times for the Second Sentence as a Function of Coreference and Conjoint Frequency in Experiment 3

	Conjoint frequency			
	High	Low	Difference	Standard Error
Non-coreferential	1348	1394	*46*	28
Coreferential	1383	1480	*97*	25

concepts have words relating to them maintained in a primed state, ready for use.

There is a clear difference between word level priming as a possible origin for the conjoint frequency effect and implicit class-membership evaluation: They imply the involvement of totally different mechanisms. The experiment below was designed to enable a choice to be made between these possible theories. In practice the situation to this point is that we have encountered one case of an unreliable and numerically reduced conjoint frequency effect in a condition where we expect to find no such effect. It was decided to try to find a situation in which one would expect to find an effect of conjoint frequency with word priming but not with class-membership evaluation.

Experiment 4

An attempt was made to find materials in which the results of optional-checking versus word-priming could be maximized. Such materials are found in sentence pairs of the following type:

(19) A (Tank)/(Bus) came roaring round the corner.

(20) It nearly hit a horse-drawn *vehicle*.

Sentence (20) contains the *word* vehicle, which is the only requirement for testing a word-priming model of the conjoint frequency effect. The essential aspect of word-priming is that it is based on a principle of excitation, so that when a word precedes a second related word, the second word has to be facilitated. There is nothing in the model which allows a selective principle to operate; in this way whether *vehicle* is preceded by *horse-drawn* or not it will not affect the degree to which it is facilitated by the word *tank* or *bus*. On the other hand, when considering mechanisms which might produce the postulated checking operations it seems reasonable to suppose that only highly related concepts would be considered against one another. We suggest that "tank" or "bus" are highly related to "vehicle," or "some vehicles" but not to "horse-drawn" vehicle. For this reason it was hoped that by using materials like (19) and (20) above, the option of implicit checking for identity would not take place, and the conjoint frequency effect should be absent.

Method

The stimulus materials were related to those used in the non-coreferential condition for Experiment 3. However each category in sentence two was qualified with an adjective in such a way that the noun phrase no longer served as superordinate to either of the instances in sentence one. Apart from the modification in materials, the design, procedure, and apparatus were the same as those used in the previous studies. The 12 subjects were all undergraduates or non-academic university staff and were paid at the rate of £0.50 per session.

Results and Discussion

Analysis of the mean reading times for the second sentences reveal a small but totally unreliable conjoint frequency effect of 19 milliseconds. The mean reading time for the high conjoint frequency sentences was 1.320 seconds; for the low sentences it was 1.339 seconds. *Min F'* and *F*s by both subjects and materials were all less than 1 (SE = 13.4 milliseconds).

This result goes against a simple word priming account for the effect of conjoint frequency and together with those of Experiment 3 points to the possibility of an active search process aimed at discovering relationships between key concepts that have already been mentioned in the text and material currently under interpretation.

General Discussion

The experiments reported above have illustrated various factors which affect the time it takes us to read sentences presented in text.

We have assumed that the reading time reflects the underlying processes involved in the interpretation of the sentences, and a few of them have been discussed. We have not, however, presented any general account of how such processes enter into the comprehension of the text as a whole.

Our analysis of textual comprehension originates from the assumption that the reader's task is to extract some sort of coherent and integrated representation of the text, to be stored in memory. A text presents the reader with information about individual objects, events, people, etc., and his task is to organize this information appropriately. We will suggest, along with Anderson and Hastie (1974), that each of these "individuals" is represented at only one location in the memory structure irrespective of how many times it is mentioned in the text. Information about the "individual" can then be stored at this location.

A major problem in creating such a memory representation is ensuring that all the information about a particular individual is stored at the appropriate location. As the reader encounters a reference to an "individual" he will have to check this reference against the specification of each of the locations before he can store the new information about that individual. Anderson and Hastie (1974) have argued that proper names have a special status as referential labels here. We would not completely agree with this claim but rather suggest that each individual is represented at the most specific level possible. What we mean by this, is that if the individual is identified through a variety of different terms in the text, the memory location for that individual will be labeled with the most specific of terms. For instance, a reader encounters the following two sentences:

(21) A serviceman was seen staggering down the road.
(22) The sailor had obviously had too much to drink.

After reading sentence (21) he might have represented the information as in Figure 2a. Whereas after reading sentence (22) the information would be represented as in Figure 2b. In order for any subsequent information to be stored in this location it would be necessary to first check if it applied to the sailor.

The inclusion of such a labeling constraint would explain why it is that sentences identifying an individual with a term more specific than that previously mentioned (e.g., the instance-last condition in Experiment 1) take substantially longer to read than those using a more general term (e.g., the category-last condition). In the former case it would be necessary to change the specification on the location allotted to that individual, whereas in the latter case it would not and we would

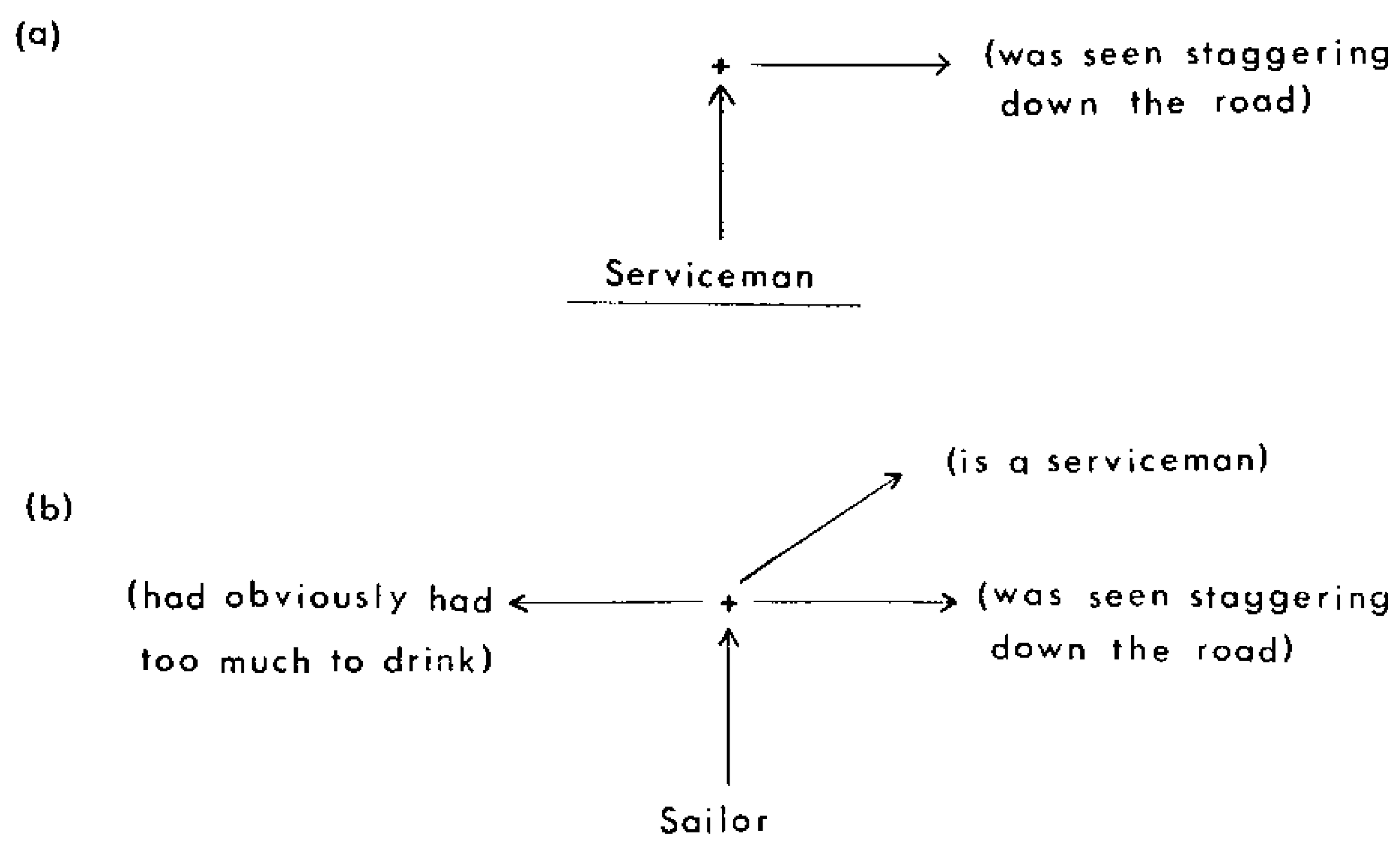


FIG. 2. Proposed memory representations for sentences (21) and (22).

assume that such changes are time consuming. In fact, cases where a more specific term is used in an anaphoric phrase are very rare in normal texts, presumably because of such a constraint.

As we have argued, the reader is able to assign information to the appropriate location by cheching the reference against the label for that location. It was suggested earlier that this checking process could be triggered in two ways. First it might originate from a preliminary analysis of the sentence itself, where a syntactic cue such as the definite article indicates the presence of an antecedent, or alternatively it could originate from an active search for information relevant to the preceding piece of text. The results from the third experiment suggested that some checking goes on in the absence of appropriate syntactic cues, (e.g., non-coreferential condition) which would indicate that both systems might be operating.

A syntactically controlled checking system has already been proposed by Clark (Note 1) as a component to his "Given–New Strategy" and has been discussed in some detail in subsequent papers (e.g., Haviland & Clark, 1974). We would assume that such a system operates as a back-up for a more automatic process which directly checks any potential relationships between a sentence and the information already stored about the text. Presumably in some situations it would be counterproductive to check every location label against every noun phrase encountered while reading. However, it would be worth doing this in the case of key items that serve as the current topic of discourse on the very reasonable assumption that subsequent sentences will refer to this topic in some way.

Such a selective search system could result from two types of location in the representation. If information has just been incorporated into a specific location this could be marked as "open" and subsequent sentences searched for information relating to it. When the topic changes, the location could be closed, and accessed only when the information in a new sentence fails to match that in an open location.

We would think of the identification of information relating to an open location as being reflected in the conjoint frequency effect, as in the sentences considered in this paper. Further, we suggest that one mechanism whereby information is accepted as relevant is the degree of semantic feature overlap between the instance and the category. Thus, overlap between "bus" and "vehicle" is great, but that between "bus" and "horse-drawn vehicle" is slight. Thus, a check on overlap will indicate that a check for relationship (identity) should take place in the former case, but not in the latter case. (The idea of a mechanism checking feature overlap has been used by Smith, Shoben and Rips (1974) to explain certain results obtained in overt class membership evaluation—for example, why it is so difficult to respond "false" to statements like *a robin is a sparrow*.) As we have already suggested, syntax plays a weak role only in preventing checks of words having a high semantic overlap with individuals in open locations when such checks are irrelevant, as in the case of non-coreferential items. Since the semantically driven checks occur only with *open* locations, there is no need to suppose that all nouns and pronouns encountered are checked against all individuals represented in memory. Finally, the advantage of a preliminary system checking feature-overlap is that low overlap items will not normally be checked against memory for inclusion within the extant representations of individuals.

In summary we suggest that the experimental data provide evidence regarding the integration of the material in sentences during reading. When a newly introduced individual is identified as the same as one already present in the memory structure for the text, this is revealed by the presence of a conjoint frequency effect. As we indicate an absence of co-reference between individuals by manipulating the syntax, the conjoint frequency effect begins to break down, possibly because syntactic cues may sometimes be used to prevent the checking operations implicit in establishing co-reference. However, it was only when the semantic overlap between individuals was sufficiently small that the conjoint frequency effect was more-or-less obliterated. This suggests that the semantic overlap check is the most important mechanism governing the process of identification (hence, integration) in the present studies.

References

Anderson, J., & Hastie, R. Individuation and reference in memory: Proper names and indefinite descriptions. *Cognitive Psychology*, 1974, **6**, 495–514.

Battig, W. F., & Montague, W. E. Category norms for verbal items in 56 categories. A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph.* 1969, **80**, 1–46.

Bransford, I. E., & Franks, J. J. The abstractions of linguistic ideas. *Cognitive Psychology*, 1971, **2**, 330–350.

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

Collins, A., & Quillian, M. R. Does category size effect categorization time? *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 432–438.

Conrad, C. Cognitive economy in semantic memory. *Journal of Experimental Psychology*, 1972, **92**, 149–154.

Halliday, M. A. K. Notes on transitivity and theme in English: II. *Journal of Linguistics*, 1967, **3**, 199–244.

Haviland, S. E., & Clark, H. H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 512–521.

Kennedy, R. A. Contextual effects in reading and recognition. In R. A. Kennedy, & A. L. Wilkes (Eds.), *Studies in long-term memory.* London: Wiley, 1975.

Landauer, T. K., & Meyer, D. E. Category size and semantic memory retrieval. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 539–549.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Loci of contextual effect on visual word recognition. In P. M. A. Rabbitt, & S. Dornič (Eds.), *Attention and performance V*, London: Academic Press, 1975.

Rosch, E. On the internal structure of perceptual and semantic categories. In I. E. Moore (Ed), *Cognitive development and acquisition of language*, New York: Academic Press, 1973.

Sanford, A. J., & Seymour, P. H. K. The influence of response incompatability on a semantic classification task. *Acta Psychologica*, 1974, **38**, 405–412.

Smith, E. E., Shoben, E. J., & Rips, L. J. Comparison processes in semantic memory. *Psychological Review*, 1974, **81**, 214–241.

Strawson, P. F. Identifying reference and truth-values, *Theoria*, 1964, **30**, 960118.

Wilkins, A. J. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 382–385.

Reference Notes

1. Clark, H. H. *Comprehension and the given-new contract.* Paper presented at conference on "The role of grammar in interdisciplinary linguistic research", University of Bielefield, Bielefield, Germany, December 11–15, 1973.
2. Harman, G. *Noun phrases derived from variable binding operators.* Unpublished manuscript, Princeton University, 1972.
3. Sanford, A. J., & Garrod, S. *The processing of class membership information directly and indirectly.* Paper presented at a meeting of the Experimental Psychology Society, July 1975, Cambridge, England.
4. Stenning, K. *Understanding English articles and quantifiers.* Unpublished Ph.D. dissertation, Rockefeller University, 1975.
5. Stenning, K. *The text life of noun phrases.* Working paper for the M.I.T. Workshop on New Approaches to a Realisitc Model of Language, 1976.

(Received August 8, 1976)