

Clauses and the semantic representation of words

BARBARA VON ECKARDT
Wellesley College, Wellesley, Massachusetts

and

MARY C. POTTER
Massachusetts Institute of Technology, Cambridge, Massachusetts

When a sentence with more than one clause is processed, words of the first clause become less available for recall or recognition once the clause boundary has been passed. One common interpretation of this observation is that the representation of a given word shifts from a predominantly surface form (e.g., phonological or lexical) to a semantic form, after the clause boundary. Two experiments that test that interpretation are reported. In the first experiment, two-clause spoken sentences were followed by spoken word or picture probes. Pictures were intended to provide a fast semantic match to probed words, but a slow surface match. Although response times exhibited a robust main effect of the clause position of the probe, no interaction with probe type was found. Similar results were obtained in a second experiment that compared pictures with written word probes. The faster response to picture than to word probes in the second experiment indicates that subjects did not covertly name the picture, but made a semantic match (as intended). These observations suggest the following reinterpretation of earlier sentence memory experiments: Words in the most recent clause of a sentence are more available than words in an earlier clause because their semantic representations are more active, not because their surface representations are more active.

The most recent clause of a sentence seems to have a special status in memory. It is more accurately recalled than earlier clauses (Jarvella, 1971; Marslen-Wilson & Tyler, 1976), and a recognition response to a probe word is faster when the word appeared in the more recent of two clauses (Caplan, 1972; Chang, 1980).¹ Because semantic information about sentences and text (their meaning or gist) is retained after syntactic or lexical specifics have been forgotten (e.g., Anderson, 1974; Garrod & Trabasso, 1973; Sachs, 1967), the clause effect has been explained as a shift from a verbatim surface representation (acoustic, phonological, lexical, or syntactic) to an interpreted semantic representation. The following passages from three textbooks and a review chapter illustrate this view.

As the sentence is received, it is assigned to a short-term store where the fragments that constitute each of its sentoids are collected together. Material is dismissed from this storage as soon as it can be assigned to a completed sentoid. It is because each sentoid is dismissed from this store en bloc that the clause functions as a unit of speech perception [After describing Jarvella's (1971) results] Jarvella concludes that when a clause is dismissed from im-

mediate memory it is "recoded" into some semantic representation. It is this representation that is stored for long-term recall (Fodor, Bever, & Garrett, 1974, pp. 342-343).

When a sentence is heard, the words are held in immediate memory until an interpretable segment is complete, then while the next sequence of words comes in, the first segment is processed. When the interpretative process is completed, the verbatim information is discarded to make room for the subsequent incoming material (Glucksberg & Danks, 1975, p. 88).

After a clause is assigned a semantic analysis, apparently, detailed information about the clause such as the individual lexical items that occurred in it is no longer as readily available as when it is being processed It is tempting to speculate that the lexical item from the most recently heard clause is more available because that clause is still at the syntactic-processing stage, while the earlier clause has already been assigned a meaning and has progressed to the memory storage stage. Since the assignment of meaning represents a recoding of information, the clause must be returned to an earlier processing stage before lexical recognition [recognition that a probe word was in the sentence] can take place (Cairns & Cairns, 1976, p. 165).

We have hypothesized that an important perceptual and processing unit within a long sentence is the clause. Information about the surface linguistic structure would be retained only for the time necessary to process one clause. Once this clause has been isolated, interpreted, and its meaning stored in memory, surface syntactic information would undergo a rapid decay (Flores d'Arcais, 1978, p. 157).

This work was supported by an NINCDS Public Health Service Research Fellowship (1 F32 NS 05263) to the first author and by NIMH Grant MN-27536 and NSF Grant BNS-8015597 to the second author. We thank Jerry Fodor for discussion and Barbara Faulconer for assistance. Requests for reprints should be addressed to the first author at 73 Hamilton Street, Hamilton, NY 13346, or to the second author at E10-032, M.I.T., Cambridge, MA 02139.

Similar views about the transition between verbatim memory for the current clause or sentence and interpreted memory for earlier material are expressed in the psycholinguistics textbooks of Clark and Clark (1977, pp. 135, 139, 141) and Foss and Hakes (1978, p. 112). Although some authors, such as Clark and Clark, have pointed out that a partial interpretation of the most recent segment of text may be available together with surface information, it is assumed that the surface representation accounts for the high accuracy of recall of the most recent segment. As Clark and Clark (1977, p. 141) summarize, "Both verbatim wording and semantic interpretations are retained in short-term memory, but 'normally' only semantic interpretations in long-term memory The more time listeners have had to work on an utterance, the less they have of its form and the more they have of its substance."

The purpose of the present study was to compare two explanations of the clausal shift in the accuracy and availability of information about a sentence. According to the hypothesis implicitly or explicitly espoused in the various quotations cited above, the clausal shift is due to a shift from reliance on a surface representation to reliance on a semantic representation. We will call this the representation-shift hypothesis. According to the second hypothesis we considered, the clausal shift simply reflects a change in the availability of semantic or conceptual information. There is a decrease in the availability of such information about the first clause, once a second clause has been processed. We will call this the semantic hypothesis. This hypothesis is consistent with evidence that semantic information bearing on the interpretation of a sentence becomes available as soon as words are heard (e.g., Cole & Jakimik, 1980; Marslen-Wilson, Tyler, & Seidenberg, 1978; Potter & Faulconer, 1979; Swinney, 1979).

The method used to test the two hypotheses was a variation of Caplan's probe recognition procedure. In Caplan's (1972) experiments, subjects listened to two-clause sentences followed after 150 msec by a spoken or written probe word. The task was to judge whether or not the probe word had appeared in the sentence. Even with the serial position of the target word held constant, responses to words that had appeared in the second clause were 67 msec faster than responses to words from the first clause (averaging the effects from four experiments). Chang (1980) replicated these results, using self-paced reading.²

According to the representation-shift hypothesis, prior to clause segmentation the semantic representation of words is either not available at all, for matching to a probe, or is less available than a surface representation. Conversely, after clause segmentation, the semantic representation is fully available, whereas surface representations are unavailable or less available. Because the probe word is itself a surface representation of a lexical item, matching is faster to a target in the most recent clause than to one in an earlier clause. (This explanation assumes that the most recent clause is still dominated by the surface representation at the time the probe is presented, 150 msec after the last word of the clause.)

The semantic hypothesis, in contrast, attributes the clause effect to a straightforward reduction in the availability of a semantic representation of words in the prior clause by the time a second clause is processed. The key difference between the two hypotheses concerns the semantic representation of a word. Therefore, in the present experiments, pictures as well as words (spoken or written) were used as probes in an adaptation of Caplan's procedure.

The use of pictures as probes of words in a sentence is a way of getting at the question of the semantic representation of a word, unconfounded by its surface representation. Pictured objects can be understood as rapidly as, or more rapidly than, written words in tasks such as superordinate category matching (Potter & Faulconer, 1975; Rosch, 1975) or judging the relevance of the picture or word to a preceding sentence (Potter, Valian, & Faulconer, 1977). But if the task is to name the item aloud, to make a rhyme judgment, or to perform some other task that requires retrieval of the surface form of a word, response to a picture is some 260 msec slower than response to the corresponding written word (e.g., Potter & Faulconer, 1975). The disparity between understanding and naming pictures indicates that a picture makes contact with a semantic representation well before the appropriate surface representation in the lexicon.³ For words, however, retrieval of surface information presumably precedes (or is coincident with) semantic retrieval (e.g., Smith & Magee, 1980; Snodgrass, 1984).

In a word-probe task such as Caplan's (1972), then, the surface representation of a target word could be compared to a picture probe only by naming the picture or by recoding the target's surface form into a semantic form; in contrast, a semantically encoded target could be compared more directly to the picture. For a word probe, however, the surface representation should be available at least as rapidly as a semantic representation, so that a direct surface match should be possible. Thus, according to the standard interpretation of the Caplan experiment—the representation-shift hypothesis—an interaction would be expected between the clause position of the target and the type of probe. In particular, the time to respond to the picture probe should be relatively faster when the target word appears in the first of two clauses, whereas the time to respond to the word probe should be relatively faster when the target appears in the second clause.

The semantic hypothesis, however, predicts that there will be no interaction between clause position and type of probe, because the match is in every case based on a semantic representation. The usual clause effect should be obtained, but it should be equal in magnitude for picture and word probes.

EXPERIMENT 1

Method

Subjects. The subjects were 40 college students who were paid for their participation. All were right-handed and native speakers of English.

Materials and Design. Twenty-four positively probed sentences were presented, each consisting of two coordinate clauses. For example, "The carrot was badly overcooked, and the stew had too much pepper in it."

Each clause was nine syllables long (an average of 6.4 words) and contained at least one concrete noun: a noun in one of the two clauses functioned as a target. The same target—in this example, "carrot"—was always used for a given sentence. The target words appeared in all serial positions except the first; the average position was 3.3 words before the end of the clause. The probe item was presented at the end of each sentence. (The 24 experimental sentences are shown in the Appendix.)

Positive probes were of two sorts, a spoken word matching the target word or a picture representing the object named by the target word. Across subjects, the clause order of a given sentence was reversed (by simply permuting the two clauses) and the type of probe was varied, resulting in four conditions of presentation for each sentence: first clause target/word probe, first clause target/picture probe, second clause target/word probe, second clause target/picture probe. Note that this procedure preserves the local environment of the target but does not control for its serial position within the sentence, only within the clause (see the final discussion). Each subject was presented with 6 sentences of each type, randomly interspersed with 30 filler sentences (5 with positive probes and 25 with negative probes). The filler sentences included sentences of the same grammatical form as the experimental sentences but with a negative probe, and sentences with one adverbial clause in place of a coordinate clause, presented with both positive and negative probes. Some probes of filler sentences were not concrete nouns, thus lessening the possibility that the subject could adopt a strategy of attending only to concrete nouns.

Sentences were tape-recorded with normal intonation and speed. At the end of each sentence, before presentation of the probe, a 50-msec high-frequency tone appeared on a second channel to signal the subject that the sentence was over; the probe was presented 100 msec later (150 msec after the sentence). A millisecond timer was triggered at the onset of the probe. The timer was stopped when the subject pressed one of two response buttons, designating 'yes' or 'no.' Word probes were spoken by the same person who recorded the sentences and were tape-recorded on the same channel as the sentences. Picture probes were presented by means of a slide projector equipped with a shutter. The pictures, taken from a set of materials developed by Potter and Faulconer (1975), consisted of line drawings of objects.

Procedure. Subjects were told that they were participating in an experiment on sentence perception. They were instructed to listen (on earphones) to the sentence presented to them and to indicate as quickly as possible whether the probe was represented in the sentence. Reaction times and errors were recorded by the experimenter. To encourage subjects to listen normally, they were presented with several simple comprehension questions after approximately every 10 sentences.

Results and Discussion

For the most part, subjects had no difficulty in carrying out the task. But four of the sentences had high error rates and produced atypically long response times, so they were dropped from the analysis.⁴ Altogether, 2.4% of the positive probes from the remaining 20 sentences were missed, and 1.8% of the negative probes were called positive. There was no evident difference in the error rates for picture and word probes, and the number of errors (a total of 19, on positive trials) was considered too small for further analysis.

Response times were analyzed after averaging each subject's correct, positive responses to each of the four types

Table 1
Experiment 1: Mean Time (in Milliseconds) to Respond to Positive Probes of Targets in Each Clause of the Sentence, and Percent Error

Clause	Spoken Words		Pictures	
	Mean	PE	Mean	PE
First	881	2.5	903	3.5
Second	824	2.0	829	1.5
Clause effect	57		74	

of sentence. Subjects were nested in four counterbalanced groups. (Since there was not a significant overall difference between groups, this variable will be ignored.) Probe type and clause position were within-subject variables. Similarly, the item analysis (F2) combined the scores of all subjects who saw a given item in a given condition.

The subject means are shown in Table 1, together with the percentage of misses. The main effect of probe type (a difference of 14 msec in favor of words) was not significant; both F1 and F2 were less than 1.0. As Caplan (1972) and others had found, responses to the second, more recent, clause were faster (by 65 msec) than responses to the first clause [$\min F' (1,42) = 4.65, p < .05$].

The result of greatest theoretical importance was the absence of an interaction between clause position and probe type; F1 and F2 were each less than 1.0. As discussed earlier, if the standard interpretation of the Caplan results (the representation-shift hypothesis) is correct, then one would expect to see such an interaction. Hence, the absence of an interaction between clause position and probe type is evidence that this hypothesis is incorrect. Note that the numerically larger clause effect for pictures than for words is the opposite of what would be expected on the basis of this hypothesis.

In contrast, the result was exactly as predicted by the semantic hypothesis: there was a main clause effect, but no interaction with probe type. There is, however, one difficulty. If subjects were (silently) naming the pictures before comparing them to the sentence, one would not expect to find an interaction between probe type and clause position, even if the representation-shift hypothesis were correct, for its prediction of an interaction is based on the assumption that pictures are matched to the target directly at the semantic level. Our second experiment was designed to examine the possibility that subjects were silently naming the pictures before matching them to the sentences.

EXPERIMENT 2

Experiment 2 was identical to Experiment 1 except that written word probes were used instead of auditory word probes. As already noted, a written word's surface representation is available substantially faster than that of a picture. For the 20 target items analyzed in Experiment 1, an earlier experiment (Potter & Faulconer, 1975) had shown a 234-msec word advantage in naming latency, which is a measure of relative access to surface form. If subjects are covertly naming the picture probes, responses

should be about 234 msec slower than responses to written words. But if picture probes can be compared with a semantic representation of the words in the sentence, response to pictures should be as fast as or faster than responses to written words. For example, in the earlier study, these 20 pictures were matched to a superordinate category name 33 msec faster than the words.

The fact that responses to pictures and to auditory probes in Experiment 1 were equally fast, when measured from the onset of each probe, does not decide the issue. The information in a spoken word comes in over several hundred milliseconds, whereas the information in a visual stimulus is presented instantaneously. Furthermore, experiments comparing time to identify written and spoken words have given conflicting results. Murdock and Walker (1969) found that spoken words took 189 msec longer to categorize than written words, whereas Kirsner and Smith (1974) reported a 46-msec advantage for spoken words in lexical decision. Kirsner and Craik (1971) probed a spoken list of eight words and found that a spoken probe was about 120 msec faster than a written probe. Caplan (1972) reported an average advantage of 364 msec for auditory over visual word probes, although the size of the clause effect was not consistently different for the two modalities. Given these inconsistent results, it is difficult to tell from the overall latencies of Experiment 1 whether subjects were covertly naming the pictures. A comparison between written words and pictures is required.

Method

Subjects. The subjects were 36 college students who were paid for their participation. All were right-handed and native speakers of English.

Design and Procedure. Sentences and picture probe materials were identical to those used in Experiment 1. Instead of auditory probes, visual word probes were used. These consisted of slides of written names presented in the same manner as the pictures. The procedure was identical to that of Experiment 1.

Results and Discussion

The error rate was again low: 3.8% on positive trials (omitting four sentences; see Footnote 4).⁵ The response times were analyzed as in Experiment 1,⁶ and the results were much the same; the means and error rates are shown in Table 2. The 50-msec picture advantage was significant in the subjects analysis [$F(1,32) = 13.19, p < .001$], but not in the items analysis [$F(1,19) = 1.22, p < .30$]. Plainly, the pictures were not named before being compared with the sentence.

The clause effect, an overall advantage of 64 msec for targets in the second clause, was again significant [$\min F'(1,41) = 9.17, p < .01$]. Again, the result of greatest importance was the absence of an interaction between probe type and clause position: both F_1 and F_2 were less than 1.0. As Table 1 shows, the clause effect was almost identical for pictures and words. Therefore, contrary to what the standard interpretation of the Caplan results would suggest, a purely semantic or conceptual probe gave the same clause effect as a probe that activated a surface representation of the target.

Table 2
Experiment 2: Mean Time (in Milliseconds) to Respond to Positive Probes of Targets in Each Clause of the Sentence, and Percent Error

Clause	Written Words		Pictures	
	Mean	PE	Mean	PE
First	946	5.7	895	1.7
Second	880	4.0	832	3.9
Clause effect	66		63	

In a comparison of only the word probes in Experiments 1 and 2, the 61-msec advantage of spoken words only approached significance [$F_1(1,72) = 2.63, p < .11$; $F_2(1,19) = 4.61, p < .05$]. There was no significant difference in the size of the clause effect for written versus spoken probes.

GENERAL DISCUSSION

In both experiments, picture probes as well as word probes showed a sizable clause effect: response to a target in the second of two coordinate clauses was faster than response to a target in the first clause. The absence of even a hint of an interaction between the clause effect and probe type (spoken or written words versus pictures) indicates that the standard interpretation of the clause effect, the representation-shift hypothesis, is wrong. This hypothesis claims that the current or just-completed clause is represented only or chiefly in a surface, verbatim code (e.g., phonological, lexical, or syntactic), whereas earlier clauses are represented in a semantic or conceptual code. If this were true, then to the extent that a surface representation of the words in the more recent clause was available and was used in the matching process, words should have shown a larger clause effect than pictures in our experiments. The results suggest instead that, in accordance with the semantic hypothesis, word and picture probes were both matched to a semantic representation rather than to a lexical or phonological one, even when the target appeared in the more recent clause.

The nonsignificant 50-msec picture advantage over written words, a difference close to that observed by Potter and Faulconer (1975) and others for semantic (conceptual) tasks, reinforces the conclusion that surface-representation matching did not play a significant role in the comparison. The failure to find a larger clause effect for spoken than for written word probes, both in the present experiments and in Caplan's (1972) study, also argues against a superficial sensory match of the probe and the words in the more recent clause. The widespread claim that recent material is remembered accurately by means of a surface form is based not only on the accuracy and speed of recent recall, but also on the evidence of Sachs (1967, 1974), Begg and Wickelgren (1974), and others that the specific wording and syntax of a sentence are learned less well or lost more rapidly than the general meaning of the sentence. There is, however, no direct evidence in those studies that the specific wording and syn-

- ROSCH, E., MERVIS, C. B., GRAY, W. D., JOHNSON, D. M., & BOYES-BRAEM, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.
- SACHS, J. D. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, 2, 437-442.
- SACHS, J. D. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition*, 2, 95-100.
- SMITH, M. C., & MAGEE, L. E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General*, 409, 373-392.
- SNODGRASS, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior*, 23, 3-22.
- SWINNEY, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-659.
- TOWNSEND, D. J., & BEVER, T. G. (1978). Interclause relations and clausal processing. *Journal of Verbal Learning and Verbal Behavior*, 17, 509-521.

NOTES

1. This latter result is not due simply to recency, for it can be obtained when the serial position of the tested word is held constant (Caplan, 1972; Chang, 1980).

2. Other investigators have replicated the basic result but have noted that the clause effect is attenuated when the first clause is strongly subordinate to the second clause and cannot be given a complete interpretation until the second clause has been processed (Flores d'Arcais, 1978; Townsend & Bever, 1978). Using a related technique, Jarvella (1971) and Jarvella and Herman (1972) found that recall of just-heard text showed a similar clause effect, although, again, the effect was reduced when the first clause was subordinate. These findings about subordination are compatible with either hypothesis about the basis of the clause effect; in any case, coordinate clauses were used in the present experiments.

3. In principle, a pictured object has many names, at different levels of specificity. In practice, however, most common objects, including those we used, have a single name that is preferred by almost everyone asked to name it (see Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976).

4. Each of the four picture probes turned out to misrepresent the sense of the target word. For example, the pictured sandwich was a Swiss cheese sandwich, not the "tuna fish sandwich" mentioned in the sentence. Nonetheless, a subsequent analysis showed that these four sentences followed the same pattern as the other results.

5. The total of 27 errors, although somewhat unevenly distributed over conditions, was again considered too small for regular analysis; a chi-square analysis of the errors in the four conditions was not significant ($p > .25$).

6. The presence of a few very long RTs led us to count all RTs over 2 sec (1.4% of positive responses, evenly divided between pictures and words) as errors, included in the error rate just given. Since there were only two RTs over 2 sec in Experiment 1 (0.3%), and both were less than 2.2 sec, the removal of long RTs in Experiment 2 only does not materially affect the comparison between Experiments 1 and 2.

7. This analysis also raises the question of whether Caplan's design was an appropriate test of the clausal hypothesis: serial position was held constant, but the role and even the syntactic category of the target word were varied.

8. The present experiments do not indicate when a semantic interpretation of each word takes place, except to show that the interpretation was complete by the time the probe was compared with the sentence. One might argue that the semantic interpretation took place in the 150-msec interval between the end of the sentence and the probe onset. But Caplan's experiment (with a 150-msec delay) and the other experiments cited are equally vulnerable to that argument, and it is the claims based on those results that the present experiments tested. Furthermore, other work strongly suggests that word interpretation and higher level processing takes place during the presentation of a sentence (e.g.,

Marslen-Wilson & Welsh, 1978; Potter & Faulconer, 1979; Potter, Kroll, & Harris, 1980). The full meaning and implications of a sentence, however, may continue to unfold after the sentence is complete (cf. Townsend & Bever, 1978).

APPENDIX
Sentences Used in Experiments 1 and 2,
With the Target Word Italicized

1. The *carrot* was badly overcooked, and the stew had too much pepper in it.
2. Jerry wanted a *gun* for Christmas, but Grandmother disapproved of the idea.
3. Last week Michael bought a new *banjo*, and Billy traded in his old trumpet.
4. The kids decorated the *window*, and Herb painted the door beautifully.
5. Harry misplaced the blue *screwdriver*, and the new tool box is falling apart.
6. The *coffeepot* was filled to the brim, and the teacher expected fourteen guests.
7. The *owl* flew back to the cool forest, and the summer rain continued to fall.
8. The young man raced the *motorcycle*, and the three girls waved good-bye tearfully.
9. Molly refinished the antique *chair*, and Hector repapered the living room.
10. The *violin* was extremely old, yet Josephine treated it like a toy.
11. The old farm *horse* won the final race, yet the rider was only nine years old.
12. The young *deer* was clearly visible, and Richard slowed down the car to a halt.
13. The *bus* pulled up to the street corner, and the sailor heaved a sigh of relief.
14. Tom was excited by the *kangaroo*, and Phoebe liked the gray elephants best.
15. The chickens and *cow* were underfed, and the supply of grain was getting low.
16. The grand *piano* was out of tune, and the program was not so exciting.
17. The Jones' *car* was badly damaged, but none of the three passengers were hurt.
18. The *shoe* was missing from the closet, and Sally's stockings were all in the wash.
19. For Christmas Dee wanted a rag *doll*, but for her birthday she wanted a toy train.
20. Ted rode his new *bicycle* to work, and his sister used the new Toyota.
21. Sally's *socks* matched her yellow sweater, and Elaine had on her new jumper.*
22. Brenda forgot the *bread* and seasonings, and the red wine tasted like vinegar.*
23. The hammer and monkey *wrench* were gone, but we found the carton of nails and bolts.*
24. Sandra gave me a tuna *sandwich*, and I found some root beer in the pantry.*

Note—To present the target in the second clause, the two clauses were reversed, keeping the conjunction between them. *These sentences were omitted from the analyses because they resulted in too many errors (see Footnote 4).