JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 8, 732-736 (1969)

The Recall of Verbal Material Accompanying Semantically Well-Integrated and Semantically Poorly-Integrated Sentences[1]

SHELDON ROSENBERG[2]

University of Michigan, Ann Arbor, Michigan 48108

This study tested the hypothesis that the recall of verbal material (critical material) accompanying semantically well-integrated (SWI) sentences will be superior to the recall of verbal material accompanying semantically poorly integrated (SPI) sentences. Complex sentences were constructed which contained two underlying sentences: a matrix sentence and an embedded sentence. Under the SWI condition, one underlying sentence was an SWI string, while under the SPI condition one was an SPI string. The critical material (identical for both levels of semantic integration) was contained in the second underlying sentence. The location of the critical material (i.e., whether it was the matrix or the embedded underlying sentence) was varied. A list-learning study-test procedure was used with 5 trials. The results indicated superior recall for the critical material under the SWI condition, and were interpreted in terms of a storage hypothesis.

Previous research (Rosenberg, 1968b) indicates that semantically well-integrated (SWI) sentences (as determined from norms of associative dependencies in sentences) are recalled better than semantically poorly integrated (SPI) sentences and that this difference in recall is due to a tendency during learning to recode the words in SWI sentences into larger chunks than the words in SPI sentences. Further evidence for a strong storage factor in the superior recall of SWI material comes from a study by Rosenberg (1968a) of the recall and recognition of associatively related and associatively unrelated nouns embedded in sentences in connected discourse. The importance of chunking in learning, as Miller (1956a,b) has pointed out, is that it reduces the load on memory by reducing the number of units to be stored. Thus, from this standpoint, it can be concluded that *S*s exposed to SWI sentences have less to remember than *S*s exposed to SPI sentences.

The present study was designed to test an implication of this conclusion, namely, that the recall of verbal material accompanying SWI sentences will be superior to the recall of verbal material accompanying SPI sentences. This was accomplished by combining two underlying sentences through embedding. One of these sentences—the critical sentence—was an SPI string, while the other was either an SWI string or another SPI string. Thus, in the sentences *The doctor who fired the janitor cured the patient* and *The doctor who fired the janitor shook the author*, the critical sentence is *The doctor fired the janitor*, while the SWI and SPI context sentences are, respectively, *The doctor cured the patient* and *The doctor shook the author*. However, for purposes of comparison, the location of the critical and context predicates was varied, i.e., the critical predicate was part of either the embedded sentence or the matrix sentence.

The hypothesis tested in this experiment, then, was that the recall of the critical verbal material would be enhanced by the presence

[1] The research reported herein was performed in part pursuant to Contract OEC-3-3-061784-0508 with the U.S. Department of Health, Education and Welfare under the provisions of P.L. 85-531, Cooperative Research, and the provisions of Title VI, P.L. 85-864, as amended.

[2] Now at the University of Illinois at Chicago Circle.

of an SWI context. The simple underlying sentences were constructed with the assistance of norms of sequential associative dependencies in active declarative sentences (Rosenberg & Koen, 1968). To produce these norms, *S*s were given sentence frames that contained a subject noun, and spaces for them to associate (to the subject noun) a verb and another noun (e.g., "The dog —— the ——."). The norms consist of frequency counts of the verb-object combinations that accompanied each of the subject nouns.

In a related investigation, Savin and Perchonock (1965) studied the problem of the amount of storage capacity remaining in memory after sentences varying in syntactic complexity (and, presumably, in the amount of "space" required in memory) had been stored. However, more relevant to the present problem was a test by Bower (1969) of the chunking hypothesis which involved the use of context and critical material. Twelve unrelated (critical) words were accompanied in a list by twelve SWI triplets (clichés), by twelve individual words, or by twelve SPI triplets, and *S*s were required to recall freely as much of each list as they could. As anticipated, "recall of the critical words was about the same whether the other twelve units were one-word or three-word (cliché) units, but recall was reduced if the other words comprised more (than 12) chunks." Thus, recall of the critical words varied as a function of the amount of storage space needed for the context material.

METHOD

Subjects

The *S*s for this study were 80 paid undergraduate volunteers who were assigned in rotation to four groups of 20 *S*s each as they appeared for the experiment. The *S*s were tested in groups of from 3 to 12 with a mean of 6 *S*s per session.

Materials

Two lists of 10 sentences were constructed for each condition to increase the generality of the results. All sentences were of the form *The thief who delivered the tape stole the money*. Each sentence in the SWI condition had a counterpart in the SPI condition that contained the same subject noun and an identical predicate. For example, the SPI counterpart of the SWI sentence *The thief who delivered the tape stole the money* was *The thief who delivered the tape passed the wagon.* In this example, *stole the money* and *passed the wagon* represent (along with the subject noun), respectively, the SWI and SPI contexts, while the predicate common to each of these sentences (*delivered the tape*) represents what was referred to earlier as the critical material. The predicates of the SWI and SPI contexts were made comparable in length (average number of letters) and in the Thorndike and Lorge (1944) frequency of their content words. The majority of the words used in the sentences of this study were listed as A or AA in the Thorndike and Lorge norms.

For each sentence in the SWI condition, the context predicate was selected from the top of the associative frequency hierarchy for the subject noun of the sentence using the Rosenberg and Koen (1968) norms, while the critical predicate (and the context predicate for the SPI counterpart) were constructed from responses at the bottom of the associative frequency hierarchy for the same subject nouns. In order to control for length and word frequency, sometimes it was necessary to use a word in a critical or SPI predicate that did not occur at all in the norms as a response to the subject noun. In one condition (Condition A), the critical predicate was embedded in the SWI and SPI contexts, while in another condition (Condition B), its location was reversed. In constructing each of the experimental lists, an attempt was made to reduce intralist associative relationships to a minimum. In brief, then, the basic conditions of the experiment can be designated as SWI–A (e.g., *The doctor who fired the janitor cured the patient*), SWI–B (e.g., *The doctor who cured the patient fired the janitor*), SPI–A (e.g., *The doctor who fired the janitor shook the author*), and SPI–B (e.g., *The doctor who shook the author fired the janitor*).

Each list of sentences was printed in a booklet, one sentence to a page, that measured $8\frac{1}{2}$ by $2\frac{1}{4}$ inches. There was a cover sheet on the front of each booklet and a blank lined card attached to the back for use during the written recall task. Each *S* received five booklets, one for each trial.

Procedure

A study-test written recall procedure was employed in this experiment. All *S*s were tested in a sound-insulated research classroom. Detailed instruction in the use of the booklets, and for the learning and written recall tasks, were recorded on magnetic tape and presented to the *S*s after they were seated at tables. The tape also contained verbal signals for turning the pages of the booklets and for the beginning and the

end of the recall period. Since the instructions were identical for all conditions, it was possible to test *S*s from more than one condition simultaneously. The *S*s were told that their task was to learn verbatim as many of the sentences as they could in any order and to record in writing as many of the sentences as they could remember during the recall period. In addition, they were told that it was not necessary to record the sentences they remembered in the order in which the sentences had appeared in the booklet. They were urged to write down everything they could remember and to guess at words they could not recall. There were five different orders of the sentences in each condition and five different arrangements of these orders from trial to trial. Thus, for any given *S*, no order of sentences occurred more than once within the packet of five booklets.

The study period began 8 sec. after a signal to get ready, the exposure interval for each sentence was 8 sec., and the interval between exposure of the last sentence and the beginning of the recall period was 8 sec. The recall period was 4 min., the intertrial interval was 16 sec., and there were five trials.

Results

With the exception of minor spelling errors and changes in the relative pronoun from *who* to *that*, all scoring was for verbatim recall. Each *S*'s written recall protocols were scored for the total number of: (*a*) complete sentences, (*b*) subject nouns, (*c*) words (verbs and nouns) from context predicates, and (*d*) words from critical predicates recalled correctly. The scores for each one of these variables for each *S* were summed over trials. The means for each of these measures are shown in Table 1, where it can be seen that for each of the dependent variables, performance under the SWI conditions was superior to performance under the SPI conditions. In addition, there was a tendency for recall to be better when the critical material was embedded in the matrix sentences (Condition A) than when it was part of the matrix sentences (Condition B).

TABLE 1

Means for Various Measures of Recall

Group	Measure			
	Subject nouns	Context words	Critical words	Complete sentences
SWI–A	42.04	81.55	77.00	33.60
SWI–B	41.50	79.95	75.80	31.90
SPI–A	40.10	72.65	70.20	28.15
SPI–B	39.55	69.60	66.15	26.35

A 2×2 factorial analysis of variance was carried out for each of the dependent variables. In these analyses (with $p = .05$), only the values of $F(1,76)$ for semantic integration reached significance. No other value of F even approached significance. The values of F for Semantic Integration were, for subject nouns, context words, critical words, and complete sentences, respectively, 6.20, 17.68, 10.81, and 11.01. Thus, not only was recall of the SWI material superior to recall of the SPI material, but recall of the critical material was enhanced when it accompanied the SWI material.

Performance increased continuously over trials for the various recall measures, but statistical analysis revealed no evidence of interaction between Trials and the main independent variables. The SWI and SPI groups still differed on the final trials. For example, in the case of complete-sentence recall, proportionally more SWI *S*s (.35) had perfect scores on Trials 4 and 5 than SPI *S*s (.20).

The recall data were also examined with respect to one other measure, the probability of recalling a complete predicate given that its subject noun had been recalled within the same sentence. This recall dependency was used to represent the degree of integration within the simple underlying sentences at the time of recall. It was anticipated that if the critical material (which was poorly integrated semantically before learning) was, indeed, learned better under the SWI condition than under the SPI condition, then this should be reflected in the recall dependencies. Of course, in the case of the underlying context sentences, any differences in subject–predicate recall dependencies should be in favor of the SWI condition, since the predicates in these sentences were selected on the basis of the

strength of their relationship to the subject nouns. To compute this measure, the number of times a subject noun was recalled correctly (summed over sentences and trials) was divided into the number of times it was accompanied by its predicate. This was done for both the context and critical subject–predicate dependencies, and the results have been summarized in Table 2.

TABLE 2
MEAN SUBJECT-PREDICATE RECALL DEPENDENCIES

Group	Underlying sentence	
	Context	Critical
SWI–A	.903	.813
SWI–B	.898	.806
SPI–A	.804	.769
SPI–B	.775	.739

The results shown in Table 2 are in general, consistent with expectation. The recall dependencies for the SWI conditions are higher than they are for the SPI conditions for both the context and the critical materials. However, as one would anticipate, the difference is larger for the context material than for the critical material. In addition, for both levels of Semantic Integration, the recall dependencies for the context sentences are higher than they are for the critical sentences. These differences, however, are greatest for the SWI condition. A three-way analysis of variance (with Underlying Sentence as a within variable) of the recall dependencies revealed $F(1,76) = 12.59$, $p < .005$, for Semantic Integration, $F(1,76) = 53.33$, $p < .001$, for Underlying Sentence, and $F(1,76) = 11.00$, $p < .005$, for the interaction between Semantic Integration and Underlying Sentence. None of the other values of F approached significance. The significant interaction appears to be associated with the fact that the effect of Semantic Integration was greater for the context material than for the critical material, and that the differences in recall dependencies between the context and critical material were greater for the SWI condition than they were for the SPI condition. An analysis of the simple effects revealed, however, that the effect of Semantic Integration was significant for both the context ($p < .001$) and the critical material ($p < .05$). An analysis of the simple effects for Underlying Sentence suggests that most, if not all of the main effect of this variable, is to be found within the SWI condition. The effect of Underlying Sentence was highly significant ($p < .001$) for the SWI groups, whereas for the SPI groups, the effect failed to reach significance at the .05 level. This latter finding can be understood in terms of the fact that for the SPI condition, pre-experimental semantic constraints were weak within both the context and the critical underlying sentences.

DISCUSSION

These results offer further support for the hypothesis that SWI sentences are stored more efficiently than SPI sentences. However, there are two possible objections to this conclusion that are likely to be raised, and therefore should be commented upon here.

First of all, is it not possible that the facilitation associated with the SWI condition was the result of construction during recall (correct guessing on the basis of semantic constraints created by what has already been recalled)? The construction hypothesis could account for both the recall and the recall-dependency data for the SWI underlying sentences. The results of previous research (Rosenberg, 1968a,b), however, offer very little, if any, support for this hypothesis. In the case of one study (Rosenberg, 1968a), for example, when opportunities for construction during retrieval of the nouns in SWI strings such as *there were some kittens, cats, dogs, and mice outside as well,* were very drastically reduced, their retrieval was still superior to the retrieval of the nouns in SPI counterparts (*there were some kittens, guns, rocks, and oars outside as well*).

The second objection has to do with the superior recall of the critical material under

the SWI condition. Since more of the words from the SWI contexts were recalled than from the SPI contexts, the *S*s in the SWI condition had more cues available to them during recall for retrieval of the critical material. In other words, the superior recall of the critical material under the SWI condition could have been the result of a retrieval rather than a storage factor. However, while this factor may have made some contribution to the recall of the critical material under the SWI condition, the results for the recall-dependency measure indicate that there must have been a significant storage factor operating as well.

In regard to the recall-dependency measure, it should be pointed out also that the magnitudes of the probabilities shown in Table 2 (the lowest is .739) are such as to suggest strongly (group differences aside) that there was a general tendency to store the information in the underlying strings in memory units larger than the word. These results are consistent with the view (Miller, 1962) that sentence contexts make it possible for us to recode words into higher-order linguistic units. In addition, they evidence what Horowitz and Prytulak (in press) refer to as "redintegrative memory," the criterion for which is whether there is in free recall a high probability that a whole unit will be recalled, given a part of the unit has been recalled.

There was no evidence that the location of the critical or context material in the surface sentence made any difference in recall. This finding suggests that the information in the two underlying strings may have been stored independently of the information relevant to their location in the surface sentence.

References

Bower, G. H. Chunks as interference units in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 610–613.

Horowitz, L. M., & Prytulak, L. S. Redintegrative memory. *Psychological Review*, in press.

Miller, G. A. Information and memory. *Scientific American*, 1956, **195**, 42–46. (a).

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81–96 (b).

Miller, G. A. Some psychological studies of grammar. *American Psychologist*, 1962, **17**, 748–762.

Rosenberg, S. Associative facilitation in the recall and recognition of nouns embedded in connected discourse. *Journal of Experimental Psychology*, 1968, **78**, 254–260 (a).

Rosenberg, S. Association and phrase structure in sentence recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 1077–1081 (b).

Rosenberg, S., & Koen, M. J. Norms of sequential associative dependencies in active declarative sentences. In J. C. Catford (Ed.), *Studies in Language and Language Behavior, Supplement to Progress Report VI.* Ann Arbor: University of Michigan, Center for Research on Language and Language Behavior. U.S.O.E. Contract OEC-3-6-061784-0508, February 1, 1968.

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.

Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Teachers College, 1944.

(Received May 29, 1969)