R E P O R T R E S U M E S

ED 016 958 48 AL 000 928

ASSOCIATIVE CLUSTERING IN THE PRODUCTION OF CONNECTED DISCOURSE.

BY- ROSENBERG, SHELDON
MICHIGAN UNIV., ANN ARBOR,CTR.FOR RES.LANG.AND BEH

REPORT NUMBER BR-6-1784 PUB DATE 1 SEP 67
CONTRACT OEC-3-6-061784-0508
EDRS PRICE MF-$0.25 HC-$0.36 7P.

DESCRIPTORS- *LANGUAGE RESEARCH, *ASSOCIATION (PSYCHOLOGICAL), VERBAL STIMULI, *NOMINALS, SENTENCES, CONNECTED DISCOURSE, ASSOCIATIVE CLUSTERING,

ONE GROUP OF 16 SS (SUBJECTS) WAS GIVEN A RANDOMIZED LIST OF WORDS THAT CONTAINED EIGHT SERIES OF FOUR ASSOCIATIVELY RELATED NOUNS, WHILE ANOTHER GROUP OF 16 SS WAS GIVEN A RANDOMIZED LIST OF WORDS THAT CONTAINED EIGHT SERIES OF FOUR ASSOCIATIVELY UNRELATED NOUNS, AND THE TASK FOR BOTH GROUPS WAS TO WRITE A STORY THAT CONTAINED THE NOUNS FROM THE EXPERIMENTAL LISTS. FREE ASSOCIATION NORMS WERE USED TO SELECT THE NOUNS. THE MAIN FINDINGS WERE--(A) THE CLUSTERS OF KEY ITEMS THAT APPEARED IN THE SENTENCES OF THE HIGH ASSOCIATION (HA) PRODUCTIONS WERE LARGER THAN THE CLUSTERS THAT APPEARED IN THE SENTENCES OF THE LOW ASSOCIATION (LA) PRODUCTIONS, (B) A KNOWLEDGE OF ASSOCIATIVE RELATIONSHIPS MAKES IT POSSIBLE TO PREDICT WHICH NOUNS WILL OCCUR TOGETHER IN THE SAME SENTENCE, (C) ASSOCIATIVELY RELATED NOUNS THAT OCCUR IN THE SAME SENTENCE ARE MORE LIKELY TO APPEAR AS A SUBJECT OR PREDICATE-NOUN COMPOUND THAN ASSOCIATIVELY UNRELATED NOUNS (ASSOCIATIVEL" UNRELATED NOUNS TEND TO APPEAR IN DIFFERENT UNDERLYING SENTENCES), AND (D) THE HA PRODUCTIONS TENDED TO BE SHORTER THAN THE LA PRODUCTIONS. THIS REPORT APPEARS IN "STUDIES IN LANGUAGE AND LANGUAGE BEHAVIOR, PROGRESS REPORT V," SEPTEMBER 1, 1967. (AUTHOR/AMM)

BR 6-1784
PA 48

ED016958

Associative Clustering in the Production of Connected Discourse[1]

Sheldon Rosenberg

Center for Research on Language and Language Behavior
University of Michigan

One group of 16 Ss was given a randomized list of words that contained 8 series of 4 associatively related nouns, while another group of 16 Ss was given a randomized list of words that contained 8 series of 4 associatively unrelated nouns, and the task for both groups was to write a story that contained the nouns from the experimental lists. Free association norms were used to select the nouns. The main findings were: (a) the clusters of key items that appeared in the sentences of the high association (HA) productions were larger than the clusters that appeared in the sentences of the low association (LA) productions; (b) a knowledge of associative relationships makes it possible to predict which nouns will occur together in the same sentence; (c) associatively related nouns that occur in the same sentence are more likely to appear as a subject or predicate-noun compound than associatively unrelated nouns (associatively unrelated nouns tend to appear in different underlying sentences); and (d) the HA productions tended to be shorter than the LA productions.

In all of the investigator's research on associative facilitation in the recall of connected discourse (e.g., Rosenberg, 1966b, 1967, in press), it was necessary for control purposes to create passages which were comparable in all respects except in the strength of association (as determined from free association norms) between certain words, most often nouns. Thus, a high association (HA) passage might contain the following:

After they ordered, there was time to look at a newspaper, where there was a story about a king and a queen. It concerned a ruler who had just been given a new crown.

The key words in this excerpt are king, queen, ruler and crown. In the low association (LA) counterpart, queen, ruler and crown might be replaced by such items as nurse, leader and plane.

While the approach that has been taken to the construction of experimental passages appears to be fully justified in terms of the need to control certain critical variables of importance in any verbal learning study (e.g., word frequency, length, contextual constraints), an interesting question arises as to the normative characteristics of passages that are constructed by naive Ss using lists of nouns that vary in inter-item associative strength. For example, are associatively unrelated nouns as likely to occur

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

AL 000 928

in the same sentence as associatively related nouns? If not, is it not likely that HA passages will be shorter, on the average, than LA passages? The task of trying to integrate associatively unrelated nouns into a story is likely to involve the inclusion of more additional content than would be the case with associatively related nouns.

The answers to such questions are likely not only to be useful to the experimenter trying to create materials for use in studies of the role of association in complex verbal learning, but they may also contribute to our understanding of the general problem of the relationship between semantics and syntax.

In the present study, Ss were given either a list of 32 nouns that contained eight groups of four associatively related items, or a list of 32 nouns that contained eight groups of four associatively unrelated items. The lists were arranged so that associatively related items (and their LA counterparts) never occurred contiguously. The S's task was to write a story that included the nouns from his list, the only restriction being that he was not permitted to use any key word more than once. It was anticipated that in examining his list, an HA S would (among other things) tend to rearrange the nouns into clusters of associatively related items. Although the task is different, this expectation is consistent with the results of research on associative clustering in free recall (e.g., Jenkins & Russell, 1952; Rosenberg, 1966a).

As far as the present task is concerned, the result of associative clustering should be to increase the likelihood that associatively related items will occur in the same complex sentence (it was assumed that Ss would produce sentences that were combinations of simpler sentences) as a subject or predicate-noun compound (i.e., as identical syntactic constituents of the same underlying sentence), since associatively related nouns tend to be items that not only contrast minimally with each other (Deese, 1965), but, from a linguistic standpoint, share syntactic privileges of occurrence as well. As for LA nouns, while it is possible to select them so that they share syntactic privileges of occurrence (i.e., so that each could appear as a subject or as a predicate-noun), the absence of minimal contrast is likely to result in their not occurring in the same underlying sentence. Data to support these expectations can be found in the research of McNeill (1963) on the origin of associations in the same grammatical class.

Method

Subjects. The Ss were 32 paid volunteers from introductory undergraduate psychology courses who were assigned alternately to two groups of 16 Ss each.

Materials. Eight stimulus nouns were selected from the Palermo and Jenkins (1964) free association norms for college students, and for each, three high-strength response nouns and three low-strength response nouns. The HA list contained the eight groups of four HA items, and the LA list the eight groups of four LA items. An example of the kinds of items used was given earlier; i.e., king, queen, ruler, crown, vs. king, nurse, leader, plane. The HA responses were selected from among the most frequently occurring items in the norms, whereas the LA counterparts consisted of idiosyncratic responses and items that did not occur at all as responses to the stimulus nouns. The LA response nouns were selected so that they would be comparable to the HA response nouns in Thorndike-Lorge (1944) frequency and in length (most of the items used were either AA or A words). In addition, HA and LA items were matched closely on such linguistic markers as animate-inanimate, human-nonhuman, mass-count and abstract-concrete, and each group of four HA and LA items contained no less than three nouns which shared the same syntactic privileges of occurrence.

Four different orders of each list were constructed to control for possible serial effects, with the exception that associatively related items (and their LA counterparts) never occurred contiguously. The experimental lists did not appear to be differentially weighted with respect to intra-list associations between the items from the different groups of four. Each S in the experiment received a booklet that contained a sheet of printed instructions, a sheet on which was printed one of the lists of 32 nouns, and three sheets of blank lined paper.

Procedure. The data were collected in a group-testing situation, and E read the instructions (which were the same for the HA and LA conditions) outloud while the Ss followed them in their booklets. The instructions that follow describe in detail the S's task in this study.

Instructions: We would like you to help us prepare some materials for use in studies of verbal behavior. On the attached sheet of paper you will find a list of 32 nouns. Your task is to make up a story that includes these nouns. Please use the following procedure in carrying out your task.

1. Read through the entire list of nouns before beginning.

2. Use each noun once only. As soon as you have used a noun in your story put a line through it so that you can keep track of what you've used.

3. Whatever else besides these nouns that you include in your story is entirely up to you.

4. The number of nouns that you include in any given sentence is entirely up to you.

5. The nouns may appear in your story in any order that you like.

6. The length of the sentences, the length of the paragraphs and the number of paragraphs in your story is entirely up to you. However, we would like you to try very hard to complete your task by the end of the hour.

7. Do all of your writing on the lined sheets that accompany the list of nouns. Please write neatly.

8. You can mark up the list of nouns any way that you feel might be helpful to you.

As soon as you have finished, check over your papers to be sure that you have used all of the nouns in the list. Do not look at your neighbor's paper, since he or she may have a different list of nouns. What is more important, however, is that you do your own work. If you finish early, please remain seated until the others have finished.

You should not be too concerned about literary style. Just do the best you can in the time you have. You may refer back to these instructions at any time.

Results

Each S's story production was scored initially for the total number of sentences that contained two or more key items, without reference to the HA and LA quadruplets (Q's), and it was found that the two groups did not differ on this measure. The mean for Group HA was 8.38, and for Group LA, 8.44. However, Group HA ($\bar{X}$ = 25.44) was superior to Group LA ($\bar{X}$ = 22.44) in the number of key words that appeared in sentences that contained two or more key items (t (30) = 2.13, $p < .025$, one-tailed). What this finding means is that the clusters of key items that appeared in the sentences of the HA productions were larger than the clusters that appeared in the sentences of the LA productions.

The means for the number of clusters that were found in sentences that contained two or more words from associative Q's were, for Groups HA and LA,

respectively, 7.13 and 3.13. The value of $t(30)$ for this comparison was found to be 8.33, $p < .001$, one-tailed. Thus, our ability to predict which of the key items will occur in the same sentence appears to be better for associatively related items than for associatively unrelated items.

Each sentence from the HA productions that contained two or more items (clusters) from an associative Q, and their LA counterparts, were examined in order to identify the syntactic relationships within which the items in question appeared. These sentences were all complex sentences, i.e., sentences that had as their constituents two or more simple sentences. It was necessary, therefore, in performing this analysis, to first reduce each complex sentence to its basic underlying strings. The syntactic relationships that accounted for all of the HA and LA clusters were, (a) subject/predicate-noun, (b) subject or predicate-noun compound, (c) different underlying sentences (the nouns from a cluster did not share the same subject or predicate), and (d) direct-indirect object. The present discussion, however, will be limited to the first three dependent variables, since instances of the direct-indirect object relationship were very rare in both groups.

The mean per cent occurrence of subject/predicate-noun relationships was greater in Group HA ($\bar{X} = 23.50$) than in Group LA ($\bar{X} = 15.63$), but not significantly so, $t(30) = 1.33$, $p > .05$, one-tailed. Group HA also surpassed Group LA in the mean per cent occurrence of compounds. The means in this case were 55.63 and 40.06. This difference, however, was significant, $t(15) = 2.05$, $p < .05$, one-tailed. The degrees of freedom was reduced by half for this comparison (Edwards, 1960) because of heterogeneity of variance. The mean per cent occurrence of clusters in which the items appeared in different underlying sentences was 19.81 in Group HA and 39.56 in Group LA. This difference was highly significant, $t(15) = 3.00$, $p < .005$, one-tailed. Here, also, it was necessary, because of heterogeneity of variance, to reduce the degrees of freedom by half. From these results, it appears that associatively related nouns tend to have shared privileges of occurrence in the same underlying sentence, while associatively unrelated nouns tend to appear in different underlying sentences.

The sentences in the HA productions ($\bar{X} = 19.25$) were longer than the sentences in the LA productions ($\bar{X} = 18.50$), but not significantly so ($t\ 30) = .41$, $p > .05$), whereas the LA productions ($\bar{X} = 407.44$) contained

significantly more words than the HA productions ($\bar{X}$ = 348.75). The value of $t(30)$ for this last comparison was 1.72, $p < .05$, one-tailed. Since the HA and LA productions did not differ in sentence length, the difference in number of words can be seen to have been the result of a difference in the number of sentences that each contained.

Discussion

The results of the present study suggest strongly that the associative-semantic structure of nouns influences the way in which they will be utilized in the production of connected discourse. At a superficial level, there is the finding that it is possible to predict which nouns are likely to occur together in the same sentence, but, what is more important, associatively related nouns tend to behave differently syntactically than associatively unrelated nouns. From the standpoint of the research (Rosenberg, 1966b, 1967, in press) on associative facilitation in the recall of connected discourse, the results of the present study suggest that in the case of a string from an HA passage, such as there was a story about a king and a queen, as contrasted with its LA counterpart, there was a story about a king and a nurse, nurse is less likely to be recalled than queen not only because of a weak associative dependency between king and nurse, but also because nurse is not likely to have shared privileges of occurrence with king in the same underlying sentence.

There was some support for the expectation that HA productions would be longer than LA productions, although the effect was not a striking one. However, if the effect of association on the length of story productions is a reliable one, it is reasonable to hypothesize that it was the result of the fact that HA Ss were able to include more key items in their sentences than LA Ss. The finding that the two groups did not differ in the number of sentences that contained two or more key items (without reference to the associative Q's) reflects, perhaps, the fact that clustering does take place on the basis of factors other than association (e.g., Tulving, 1962). In addition, even if there is a tendency to include each LA item in a different sentence, the tendency to produce sentences that are combinations of simpler sentences should result in an appreciable amount of within-sentence clustering of LA items.

References

Deese, J. *The structure of associations in language and thought.* Baltimore: Johns Hopkins Press, 1965.

Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1960.

Jenkins, J. J., & Russell, W. A. Associative clustering during recall. *Journal of Abnormal and Social Psychology,* 1952, *47,* 818-821.

McNeill, D. The origin of associations within the same grammatical class. *Journal of Verbal Learning and Verbal Behavior,* 1963, *2,* 250-262.

Palermo, D. S., & Jenkins, J. J. *Word association norms: grade school through college.* Minneapolis: University of Minnesota Press, 1964.

Rosenberg, S. Associative clustering and repeated trials. *Journal of General Psychology,* 1966, *74,* 89-96. (a)

Rosenberg, S. Associative factors in the recall of connected discourse. *Psychonomic Science,* 1966, *4,* 53-54. (b)

Rosenberg, S. Recall of noun pairs embedded in connected discourse as a function of association strength. *Studies in language and language behavior, IV,* February 1, 1967, Center for Research on Language and Language Behavior, University of Michigan, Contract OEC-3-6-061784-0508, U.S. Office of Education.

Rosenberg, S. Associative facilitation in the recall of connected discourse. In E. Z. Rothkopf (Ed.), *Verbal learning research and the technology of written instruction.* Chicago: Rand McNally, in press.

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

Tulving, E. Subjective organization in free recall of "unrelated" words. *Psychological Review,* 1962, *69,* 344-354.

Footnote

[1]The research reported herein was performed pursuant to Contract OEC-3-6-061784-0508 with the U. S. Department of Health, Education and Welfare, Office of Education, under the provisions of P.L. 83-531, Cooperative Research, and the provisions of Title VI, P.L. 85-864, as amended. This research report is one of several which have been submitted to the Office of Education as *Studies in language and language behavior, Progress Report V,* September 1, 1967.