Recognition memory as influenced by differential attention to semantic and acoustic properties of words[1]

GREG CERMAK, JOHN SCHNORR,[2]
HERMAN BUSCHKE,[3] and R. C.
ATKINSON, Stanford University,
Stanford, Calif. 94305

Lists of 264 words were studied by 75 Ss. The Ss were divided into three groups instructed to remember either meaning or sound or both meaning and sound of the study words. During recognition, S was required to indicate which word of each test pair had been presented earlier. The incorrect alternative for each test pair was a homophone, synonym, or a word unrelated to the correct choice. Recognition increased as instructional emphasis on remembering word meaning increased. Overall recognition was best when the incorrect choice was unrelated to the study word, and worst when the incorrect choice was a synonym of the study word. The high performance levels obtained under some conditions suggest that word features other than semantic and phonetic information are important for word recognition.

Research has shown that word recognition in a forced-choice test depends upon the relationship between the distractor word (i.e., incorrect alternative) and the target word (i.e., correct alternative). Recognition is better when the distractor is unrelated to the target word than when the distractor is an associate (Underwood & Freund, 1968), a synonym (Marshall, Rouse, & Tarpy, 1969), or an acoustically similar word (Buschke & Lenon, 1969). These findings suggest that words are represented in long-term memory (LTM) as composites of semantic, associational, and phonetic information (Anisfeld & Knapp, 1968). Results have generally indicated that semantic and associational word features are of greater importance for recognition than are phonetic features (Marshall et al, 1969; Underwood & Freund, 1968), leading to the conclusion that LTM is primarily semantic in nature (e.g., Adams, 1967).

Results of the Buschke & Lenon (1969) study suggest that Ss attended selectively to particular qualities of information during word presentation. In their experiment, target words were presented for study either singly or as the underlined member of a pair. In the latter case, the nonunderlined word was the distractor

item used in the subsequent recognition test. For target words studied singly, recognition was better when the distractor was unrelated to the target than when it was either a synonym or a homophone, a result consistent with previous studies. However, when both the target and its distractor were presented together during study, error rates did not differ for the three types of distractor words. Thus, the typical result that recognition is decreased by homophone and synonym distractors did not hold when both the distractor and target words were presented together during study. One implication of this finding is that Ss compared the underlined and nonunderlined words of each study pair and concentrated on those word features that best differentiated the target from its distractor, thus enhancing their chances for correct recognition. The purpose of the present experiment was to test this hypothesis further.

SUBJECTS

The Ss were 75 male undergraduate students at Stanford University whose participation fulfilled a course requirement in introductory psychology.

MATERIALS

From the words used in the Buschke and Lenon study, 594 were selected for use in this experiment. There was no restriction with respect to parts of speech, and all the words had Thorndike-Lorge (1944) frequencies of 1-100 per million. The chosen words formed 99 homophone pairs (H pairs) and 99 synonym pairs (S pairs). The remaining words were randomly paired, with the restrictions that the words in each pair be unrelated to each other in sound, meaning, and association (U pairs).

PROCEDURE

The experiment consisted of both a study and a test phase. In the study phase, a list of words was presented at a 5-sec-per-word rate on a Teletype controlled by an on-line PDP-1 computer. Each S was presented a unique list of 264 words. One word was selected randomly for study from each of the 99 H pairs, each of the 99 S pairs, and from 66 of the 99 U pairs, the 264 words being randomly permuted.

The experimental sessions were conducted with groups of from four to eight Ss. Each S was assigned to one of eight on-line Teletypes located in a single room. The Ss run in a given session were

assigned to one of three instructional groups. The instructions for each group were identical in their explanation of the study and test procedures and differed only in reference to how the Ss were to treat the study words. For each group, the instructions contained the phrase: ". . . concentrate on the _____ of each word." For the phonetic group, the word "sound" appeared in the blank; for the semantic group, the words "dictionary meaning" appeared; for the semantic-phonetic group, the phrase "sound and dictionary meaning" appeared in the blank. These phrases appeared three times in the instructions for each group, and were read with marked emphasis. The instructions were presented to Ss by a tape recorder, and questions were answered by paraphrasing the instructions.

In the test phase of the experiment, which immediately followed study, 198 word pairs were presented at an 8-sec-per-pair rate. On each test pair, the S chose the word he thought had been presented for study, and indicated his choice by pressing an appropriate key on the Teletype.

The disparity between the number of words studied and the number tested was due to the fact that the experimental session could not exceed 1 h. The procedure used here was designed to insure that a sufficient number of study words was given so that recognition performance would be less than perfect, and that an adequate number of test trials would be given so that the data would be reliable. Since time limitations did not permit testing of all study words, a randomization procedure was used in constructing the test lists so that the probability that a given study word was tested was the same for all words.

The average test list contained 49.5 S pairs, 49.5 H pairs, 49.9 U pairs, and 49.1 "mixed pairs." Each mixed pair consisted of a synonym or homophone word that had been studied paired with an unrelated word that had not appeared previously. The purpose of testing on the mixed pairs was to determine if there were overall differences in recognition among the sets of homophones, synonyms, and unrelated words, given that the distractors came from the same pool of words. Performance on the mixed pairs was compared with performance on U pairs in making this determination. Analysis of variance of the mixed pairs and U pairs revealed no difference in recognition among the three types of target words (synonyms, homophones, unrelated words) [F(2,144) = 0.9], and the spread in performance among the three kinds of target words averaged only 4% for each

group. Therefore, any differences in recognition among the H pairs, S pairs, and U pairs can reasonably be attributed to confusions between the target item and its distractor rather than to an artifact of materials.

RESULTS

Figure 1 presents the probability of a correct recognition response, adjusted for guessing, for all groups and distractor types.[4] The direction of the type-of-distractor effect on recognition [F(2,144) = 30.28, p < .001] is consistent with the results of previous research. Recognition was better when the distractor word was unrelated to the target (U pairs) than when the distractor was either a homophone (H pairs) or synonym (S pairs) of the target word. Synonym distractors decreased recognition more than homophone distractors [t(74) = 2.38, p < .02], which supports the conclusion from previous studies that LTM is generally semantic in nature.

Many results suggest that Ss were able to attend selectively to particular word qualities during study. Assuming that LTM is primarily semantic and that Ss were able to follow instructions, recognition should have been directly related to the amount of attention devoted to semantic word features during study. This hypothesis is supported by the result that recognition improved with added instructional emphasis on remembering word meaning [F(2,72) = 28.48, p < .001]. Recognition was best for the semantic group; the lower performance levels for the other two groups suggest that when Ss are asked to attend to the sound of a word, attention to word meaning decreases, with the result that later recognition performance is impaired.

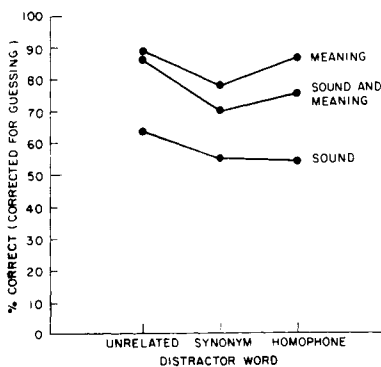The interaction between instruction groups and distractor type [F(4,144) = 3.55, p < .01] means that the

Fig. 1. Probability of a correct response (corrected for guessing) for the three experimental groups and the three types of distractors.

relative recognition performance with homophone, synonym, and unrelated distractors varied with instructions. Pairwise t tests were computed to analyze the nature of this interaction. Recognition was better on U pairs than on H pairs for both the phonetic [t(74) = 3.59, p < .001] and the semantic-phonetic [t(74) = 4.49, p < .001] groups. If these groups attended to phonetic information, as instructed, this result would be expected since phonetic information is of no value in discriminating a word from its homophone, but is valuable in discriminating words unrelated in sound. For the semantic group, however, U pairs and H pairs did not differ in recognition. The implication is that Ss in the semantic group did not use phonetic information in selecting a response during test since, in the absence of phonetic information, homophones and words unrelated in sound to the target are functionally equivalent as distractors. This result is not surprising since Ss in the semantic group were not instructed to attend to phonetic information.

Similarly, the semantic and semantic-phonetic groups did better on H pairs than on S pairs [t(74) = 3.16, p < .01; t(74) = 1.90, p < .05]. These effects are reasonable since both groups were instructed to attend to semantic information, and semantic information is useful in discriminating homophones but not synonyms. However, the phonetic group showed no difference in this measure [t(74) = −0.81, p > .3). Previous results suggested that these Ss attended less to semantic information than Ss in the other two groups. Hence, the phonetic group would not be expected to show as large a difference in recognition between H pairs and S pairs as was obtained for the other two groups.

However, recognition performance for the phonetic group suggests that these Ss attended to some extent to the semantic content of the study words. If these Ss had attended only to phonetic information, there would have been no basis for recognition on H pairs to exceed the guessing rate, since homophones do not differ phonetically. However, recognition on H pairs was much better than chance for this group, suggesting that nonphonetic information was also attended to during study. The argument receives additional support if the synonym and unrelated distractors do not differ in their phonetic similarity to target words; and there is no reason to expect a difference. Assuming this, total lack of attention to word meaning should have resulted in no difference in performance on S pairs and U pairs. Yet Ss in the phonetic group were poorer on S pairs than on U pairs

[t(74) = 1.90, p < .05], suggesting that they attended to the semantic information of the study words. In summary, the results indicate that the instruction variable influenced the degree of attention given to word meaning during study.

Results also suggest that word features other than semantic and phonetic information are important for word recognition. Assuming LTM is primarily semantic and that only semantic and phonetic word features were attended to during study, performance on S pairs in test should have been poor since (1) phonetic information is not very valuable for LTM tasks according to previous studies, and (2) semantic information does not discriminate synonyms.[5] The point is exemplified by considering the performance level of the semantic group. Results reported earlier suggest that these Ss attended almost exclusively to semantic word content during study. Since semantic information does not discriminate synonyms, recognition performance on S pairs (corrected for guessing) should have been near zero instead of the obtained 80%. Equally perplexing is the result that recognition of S pairs was best for the semantic group, since the other two groups were instructed to attend to phonetic information that should have been of some value in discriminating synonyms. Viewed in this manner, the observed group rank order in recognition of S pairs is the reverse of what would be expected.

One hypothesis to explain these anomalous results is that information other than phonetic and semantic word features is attended to during study, and this information is useful in discriminating a target word from its distractor. During test, possible candidates are the visual information inherent in the printed study words (Norman, 1969, p. 136) or associates unique to each member of a synonym pair (Anisfeld & Knapp, 1969). Attention to visual and/or associative information during study, and use of this information during test could account for the relatively high performance on S pairs for all groups since these word features differ for each member of an S pair. In order to account for the group rank order in recognition of S pairs, it must be assumed that a greater amount of visual and/or associative information is acquired as instructional emphasis on studying word meaning increases, and that this information is more valuable than phonetic information for recognition. Accumulation of visual and/or associative word information during study could also account for the Buschke & Lenon (1969) finding that H pairs, S pairs, and U pairs

did not differ in recognition when both target and distractor words were presented together during study. Since Ss had an opportunity to compare the target and distractor actively during study, they could have attended to those word features (semantic, phonetic, associative, visual) that served to make synonyms and homophones as discriminable as two unrelated words.

REFERENCES

ADAMS, J. A. *Human memory*. New York: McGraw-Hill, 1967.

ANISFELD, M., & KNAPP, M. Association, synonymity, and directionality in false recognition. Journal of Experimental Psychology, 1968, 77, 171-179.

BUSCHKE, H., & LENON, R. Encoding homophones and synonyms for verbal discrimination and recognition. Psychonomic Science, 1969, 14, 269-270.

MARSHALL, J. F., ROUSE, R. O., JR., & TARPY, R. M. Acoustic versus associative models of short-term memory coding. Psychonomic Science, 1969, 14, 54-55.

NORMAN, D. A. *Memory and attention: An introduction to human information processing*. New York: Wiley, 1969. Chap. 7.

THORNDIKE, E. L., & LORGE, I. *The teacher's handbook of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

UNDERWOOD, B. J., & FREUND, J. S. Errors in recognition learning and retention. Journal of Experimental Psychology, 1968, 78, 55-63.

NOTES

4. The correction formula employed was a standard one: $[P(C) - g]/(1 - g)$, where $P(C)$ is the observed proportion correct and g the probability of a correct guess, 0.5 in this case.

5. It is possible that the words comprising many S pairs were not completely synonymous, so that the semantic content of a study word was slightly different from that of its distractor. In that case, recognition of S pairs would be expected to be somewhat better than that obtained when both words of an S pair are perfect synonyms.

Rated acoustic (articulatory) similarity for word pairs varying in number and ordinal position of common letters[1]

DOUGLAS L. NELSON and LOUISE D. NELSON, University of South Florida, Tampa, Fla. 33620

Ratings of "similarity-in-sound" for pairs of words sharing letters in various ordinal positions were obtained using a 7-category graphic scale on which degrees of similarity were specified by verbal labels. Judged acoustic (articulatory) similarity increased as number of shared letters increased from zero to one, two, and three letters. Pairs overlapping in first letters were rated significantly more similar than pairs sharing middle or last letters. For pairs sharing letters within two ordinal positions, judged similarity increased in the order of first-and-middle, first-and-last, middle-and-last. Implications for the hypothesis that verbal stimuli are coded by pronunciation were discussed.

Manipulations of ordinal position of identical letters within sets of three-letter words have produced relatively consistent findings that have depended upon characteristics of the learning task. When discriminations between stimuli are required, as in paired-associate (Nelson & Rowe, 1969) and serial-recall learning (Nelson, 1969), difficulty of acquisition increased in the order middle (M), last (L), first (F) for stimulus sets sharing letters within single locations and in the order middle-and-last (M + L), first-and-middle, (F + M), first-and-last (F + L) for sets in which letters were shared within two ordinal positions. A similar ordering for the dual-locus conditions also has been reported when paired-associate stimuli were nonsense syllables (Richardson & Chisholm, 1969; Runquist, 1968a). When stimuli could be grouped together as in free recall (Nelson, 1969) and when paired-associate stimuli shared letters with their responses (Nelson & Garland, 1969), orders of difficulty by identity locus were reversed.

One explanation of these results assumes that the stimuli were coded by pronunciation, rendering the codes subject to acoustic (or articulatory) interference or facilitation depending upon the task requirements (Runquist, 1968a). Accordingly, to account for the effects of variations in locus of identical letters in the various learning tasks, this hypothesis predicts that rated acoustic similarity should increase in the order M, L, and F for pairs of words sharing letters within a single locus and in the order M + L, F + M, and F + L for word pairs sharing letters within two ordinal positions. The only available data have been inconsistent with this prediction. Runquist (1968b) found that word pairs sharing first letters were rated as more similar than were pairs overlapping in middle or last letters, which do not differ. No apparent differences were found between pairs sharing letters within two positions. However, the method used for obtaining ratings required estimations of the percentage (0-100) of similarity shared between stimuli of the pair, with the similarity attribute unspecified. To the indeterminate extent that similarity estimates were made on the basis of counting the number of common letters, as is suggested by ratings around 66% for all dual-locus pairs, differences as a function of locus would be attenuated. The purpose of the present study was to scale pairs of words overlapping in various ordinal positions with the similarity attribute specified as "sound similarity." Specification of the acoustic dimension was expected to reduce the likelihood of rating on the basis of letter counting and, perhaps, reveal greater differences within the overlap conditions. Moreover, ratings made on the basis of acoustic similarity should provide evidence relevant to the adequacy of the pronunciation hypothesis as an explanation for the ordinal-position effect.

MATERIALS

The items to be rated consisted of 108 pairs of words. Eighty-four of these pairs were generated by taking all possible pairwise comparisons within each of the following three lists: PAN, PAT, PIN, PIT, FAN, FAT, FIN, FIT; BAD, BAG, BED, BEG, LAD, LAG, LED, LEG; and HAM, HAT, HUM, HUT, RAM, RAT, RUM, RUT. This procedure produced four pairs of items within each list that shared letters in all possible positions, including zero. Thus, there were 12 pairs of words representing each of the seven overlap conditions. The 12 pairs representing complete identity were generated by selecting four single words from each list and by pairing each word with itself. These words were selected so that each different initial, medial, and terminal letter was equally represented (i.e., for the first list, the items chosen for self-pairing were PAN, PIT, FAT, FIN). The remaining 12 pairs were chosen, by E's judgment, so that four pairs represented each of the following three similarity conditions: high