JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 9, 499–508 (1970)

Encoding and Retention of Semantic and Phonemic Information in Short-Term Memory[1]

HARVEY G. SHULMAN[2]

University of Michigan, Ann Arbor, Michigan 48104

A probe recognition task was used to evaluate the relative effectiveness of semantic and phonemic encoding in STM. On each trial a list of 10 words was presented at a rate of either 350, 700, or 1400 msec per word. Recognition was tested with a probe which could be a homonym, a synonym, or identical to one of the words in the list. The retention functions for all three probe types were similar in shape, supporting the hypothesis that semantic encoding is possible in STM. An interaction between type of encoding and rate occurred, indicating that encoding is a time-dependent serial process.

Because phonemic similarity is apparently a more effective variable than semantic similarity in short-term memory (STM) tasks, it is often assumed that the theoretical short-term storage (STS) system, which presumably supports STM performance, is primarily limited to the storage of phonemic information. However, Shulman (1970) argued that the differential effectiveness of the two types of similarity on STM can also be accounted for by the hypothesis that the encoding of an item takes place over time, and that features most closely related to the sensory input, e.g., phonemic features, are encoded more rapidly than semantic features. In order to maximize the time available for rehearsal and under the pressure of relatively fast presentation rates, *S*s may tend to encode incoming information as quickly as possible, which implies that encoding will be based primarily on sensory attributes of the input. A subsidiary hypothesis was that semantic encoding is possible in STS when required by task demands or when slow presentation rates are used.

The purpose of the present experiment was to investigate the hypotheses that semantic encoding is possible in STS and that the encoding process is time dependent. A further object was to study the retention functions for semantic and phonemic information separately. Kintsch and Buschke (1969) have suggested that the effects of semantic similarity are confined to the asymptotic portion of the short-term retention function while recently presented information is encoded phonemically. Their experimental approach was to look for confusion errors based on the two types of similarity; hence, the task was not designed to require semantic encoding, just to detect it if it occurred. If it is true that the effects of semantic similarity in an STM task are mediated entirely by the contributions of long-term store (LTS) to performance, and if it is also true, as current hypotheses about the nature of this contribution have it, that LTS contributes about equally at all except the

[1] This research was supported by the Advanced Research Projects Agency, Department of Defense, and was monitored by the Air Force Office of Scientific Research under Contract AF 49(638)-1736 with the Human Performance Center, Department of Psychology. During part of the preparation of this document, the author was supported by a United States Public Health Service Predoctoral Traineeship under Grant No. 571MH6338-11 to the Department of Psychology. This paper is based in part on a dissertation submitted by the author in partial fulfillment of the requirements for the degree of Doctor of Philosophy from the University of Michigan, 1969. Arthur W. Melton served as chairman of the doctoral committee.

[2] Now at the Human Performance Center, Ohio State University, 404-B West 17th Avenue, Columbus, Ohio 43210.

shortest (1 or 2 sec) retention intervals, then when *S*s are forced to use semantic encoding in an STM task the resulting retention function should show no recency effect and should be relatively flat over a wide range of retention intervals. Furthermore, if STS is primarily phonemic and contributes to performance as a rapidly decreasing function of retention interval, then the retention function for phonemic information should show a marked recency effect. If traces in STS contain both semantic and phonemic information, then the shape of the retention functions should be similar for both types of information.

In order to force *S*s to encode items both phonemically and semantically, a probe recognition task was used in which the probe item might be either a synonym or a homonym of a presented word. In three conditions of the experiment *S*s were required to make "yes" or "no" decisions of three types. These were: (*a*) Is the probe word *identical* to any of the presented words? (*b*) Is the probe word a *homonym* of any presented word? (*c*) Is the probe word a *synonym* of any presented word? If semantic encoding does take longer than phonemic encoding, then an interaction between presentation rate and probe type should be predicted. In particular, decreasing the presentation rate should facilitate performance in the semantic probe condition more than in the phonemic.

METHOD

The method chosen was intended to maximize the likelihood that each word presented was encoded as fully as possible. This was done by requiring judgments of similarity according to three different criteria and by informing *S* about which criterion to use only after presentation of the 10-word list on each trial. In addition, a system of monetary payoffs was used which heavily stressed accuracy and also placed a lesser but positive emphasis on speed.

Design

Each *S* had a total of 50 practice and 240 experimental trials in a single session lasting 1.5 hr. Each experimental trial consisted of the sequential visual presentation of a 1.3-sec warning signal, 10 words, a 1.3-sec instructional cue, and a recognition test word or probe. Blank periods of 100 msec separated the warning signal and the words, the words from each other, the last word from the instructional cue, and the instructional cue from the onset of the test word. The independent variables were presentation rate, serial position of the tested word, and instructional cue type, all varied factorially within *S*s. On a given trial each of the 10 words was exposed for the same duration, either 250, 600, or 1300 msec, followed by the 100-msec blank interstimulus interval, resulting in presentation rates of 350, 700, and 1400 msec per word. There were 80 trials at each presentation rate. The tested word was drawn from each of the 10 serial positions with equal frequency, resulting in a total of 18 trials at each of 10 retention intervals. The remaining 60 trials were catch trials, on which a test word requiring a "no" response was presented. A third variable was the nature of the recognition test on a given trial. Thus, *S* was instructed that he might be asked to respond "yes" or "no" to one of the three types of questions on any trial. On any particular trial, *S* was informed as to which of the similarity criteria to use by the presentation of a single-letter cue for 1300 msec after the 10th word. The cues were the letters I (identical), H (homonym), or M (means the same). Each of these cues was used on 60 experimental and 20 catch trials, twice with each of the 30 combinations of serial position and presentation rate and either six or seven times on catch trials at each presentation rate. Two observations per *S* were made in each of the 90 combinations of cue type, presentation rate, and serial position where a "yes" response was correct.

The sequence of trials was divided into two 120-trial blocks, each containing one observation at each of the 90 combinations of cue type, rate, and serial position, and either three or four catch observations at each of the nine possible combinations of cue type and presentation rate. The order of occurrence of conditions within 60-trial subblocks was randomized with constraints imposed on runs of cue types, presentation rates, and serial positions.

Materials

The stimuli were 2580 monosyllabic or disyllabic words from three to six letters in length chosen from the list of words occurring at least once per million in the Thorndike–Lorge General Count (Thorndike & Lorge, 1944). Sixty synonym pairs were chosen from Riegel (1965), Jenkins and Palermo (1965), and by inspection of words in the Thorndike–Lorge list. Synonyms were chosen to be of maximum similarity, with preference given responses showing unimodal, peaked frequency of response distributions. Some examples are: talk-speak, leap-jump, angry-mad, nation-country, and empty-vacant. Sixty homonym pairs were chosen from Whitford (1966). Examples of these are: ball-bawl, pray-prey, board-bored, whole-hole, cereal-serial. The distributions of word lengths and Thorndike–Lorge frequencies for homonym and synonym pairs were made as nearly equal as possible. Sixty single words were chosen for the identical and catch trials which also approximated the homonyms and synonyms with respect to word frequency and length. Complete counterbalancing of the assignment of words to Rate and Serial Position conditions, within each cue condition, was achieved with 30 *S*s.

Each word pair was assigned to a string of 10 words such that the nine other words used were nominally unrelated to each other and to the test pair. No more than two words per string had the same first letter, and all word lengths were represented in every string. Probe words ranged from two to eight letters in length, with 2-, 7-, and 8-letter words used rarely and equally often as correct and catch words.

Apparatus

The *S*s sat in a dimly lit soundproof room containing the display and a two-key response panel. The display consisted of a linear array of 10 Burroughs Nixie tubes, measuring 31.8 cm across, placed at eye level, 137.2 cm from *S*. Each tube subsumed 1°20′ of visual angle; hence, a 6-letter stimulus subsumed 8° of visual angle. Preprogrammed paper tape provided a control unit with information concerning presentation rate and character set prior to each stimulus presentation. The control unit also turned on an electronic clock simultaneously with the onset of each test word. Depression of a response key stopped this clock and caused a record of *S*'s reaction time and response to be punched on paper tape. The response panel consisted of two piano-like keys. The right key was labeled "yes," the left "no." The *S*s were instructed to rest their right index finger lightly on the "yes" key and their left index finger on the "no" key.

Procedure

Prior to the 240 experimental trials, *S*s were given a total of 50 practice trials. The first 20 of these involved number recognition. The next 30 trials were practice at the experimental task, using only the 350- and 700-msec rates. Strings of four words were presented and followed by either the letter I, H, or M as the cue for a particular type of test.

Each practice and experimental trial was preceded by a warning signal consisting of four asterisks which flashed at the rate appropriate for that trial. The total time taken by the presentation of this warning signal was 1300 msec in all conditions. Each test word remained visible until a response was made, and 3 sec then elapsed before the warning signal for the next trial.

Following the practice trials, the experimental sequence was administered in four segments of about 15 min each with a 2–3-min rest interval between each segment. The *S*s were instructed that this would be the case, and a payoff matrix emphasizing accuracy over speed was described to them. No feedback was given *S* during the experiment.

Subjects

The *S*s were 60 right-handed male students at the University of Michigan. They were drawn from the volunteer *S* pool of the Human Performance Center and were paid $1.50 per hour plus a bonus based on performance.

Instructions

Instructions emphasized the unpredictable nature of the series of trials with respect to the similarity criterion, and clear instructions were given to be prepared for any of the three types of questions (I, H, or M) on every trial, in order to maximize performance. The *S*s were told that there was nothing subtle about the test, that the homonyms were words which sound identical, that the synonyms used were as close in meaning as is possible, and that when a particular instructional cue was shown, the test word would either satisfy the similarity criterion indicated by that cue or would satisfy none of the three possible criteria. Instructions were also given to refrain from speaking words out loud during their presentation.

RESULTS

Separate analyses were done for reaction times (RT) and recognition performance and for trials where "yes" and "no" were the required responses. In order to carry out an analysis of variance on the RT data, the mean RTs over groups of 10 *S*s were used as the basic data. This method was chosen because at the level of individual *S*s there were missing data at certain treatment combinations since the RT data were conditionalized either on correct or incorrect responses.

Figure 1 shows the proportion of correct "yes" responses as a function of probe type and serial position, averaged over presentation rates. The main effects of probe type and serial

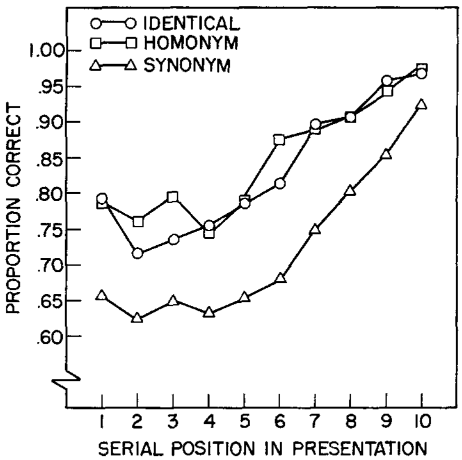

FIG. 1. Proportion correct "yes" responses as a function of serial position (10 is most recent item) with probe type the parameter.

position were significant, $F(2, 116) = 50.83$, $p < .001$, and $F(9, 522) = 60.22$, $p < .001$, respectively, and a marginal interaction between probe type and serial position was present, $F(18, 1044) = 2.86$, $.05 < p < .10$. It is apparent that the effect of probe type reflects the poorer level of performance with synonym probes since homonyms and identical words did not differ from each other. Both this fact and the similarity in shape of the three functions are of interest. Since serial position and retention interval were confounded, the curves in Figure 1 and elsewhere show the retention functions in the various conditions of the experiment. In all such figures the word in serial position 10 was the most recently presented.

The poorer level of performance with synonyms might reflect either a true difference in *S*s' ability to encode semantic information, a scalar difference between the similarity of synonyms to each other as compared to identical words or homonyms, or a combination of both factors. However, the fact that synonyms were correctly recognized 72.4% of the time over all serial positions and rates, and 92.8% of the time at the most recent serial position demonstrates the salience of semantic information in STM. The similarity in shape of the retention functions for the three probe types argues that all of the types of information involved are stored in functionally identical memory systems, and the strong recency effect makes plausible the identification of this system as STS.

The various conditions of the experiment resulted in variation in false-recognition rates which are shown in Table 1. These values were

TABLE 1

PROPORTION OF FALSE-POSITIVE RESPONSES AS A FUNCTION OF PROBE TYPE AND RATE

Probe	Rate, msec 350	700	1400
Identical	.074	.086	.081
Homonym	.133	.155	.183
Synonym	.271	.206	.198

used to calculate estimates of P', the true probability of a correct "yes" response, as dictated by a high threshold model of recognition performance (Luce, 1963). Thus if $P(Y|Y)$ is the observed proportion of "yes" responses given that "yes" was the required response, and $P(Y|N)$ is the observed proportion of "yes" responses given that "no" was required, $P' = [P(Y|Y) - P(Y|N)]/[1 - P(Y|N)]$. This correction procedure produced no change in the pattern of results. Other correction procedures are of course dictated by other models of recognition performance, such as low threshold theory (Luce, 1963) or the theory of signal detectability (Green & Swets, 1966). There is no obvious method of estimating the parameters of either theory from the present data, and no reason to assume that either is the correct model in any case. In order to deal with the present recognition data, it is therefore necessary to assume that the available estimates of P' bear some simple monotonic relationship to true sensitivity and are therefore not misleading.

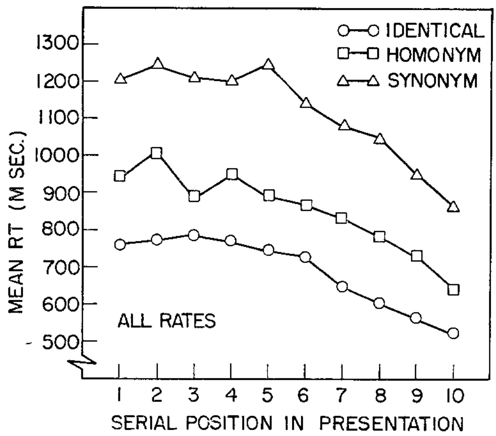

FIG. 2. Mean correct "yes" RT as a function of serial position with probe type the parameter.

Figure 2 shows the mean RT for correct "yes" responses as a function of serial position and probe type. In order to minimize variances, RTs greater than 3.0 sec were discarded prior to this and all subsequent analyses to be reported. This procedure resulted in the loss of 355 responses (2.5% of the data) and lowered mean RTs somewhat without causing any changes in the pattern of results. The effects of probe type and serial position were significant, $F(2, 10) = 611.20$, $p < .001$, and $F(9, 45) = 100.08$, $p < .001$, as was the Probe × Serial Position interaction, $F(18, 90) = 1.72$, $p < .05$. This interaction was also significant when data from Serial Positions 5–10 only were analyzed, $F(10, 50) = 2.12$, $p < .05$, reflecting the fact that the serial position curve for synonyms had a greater slope than the other serial position curves.

An important feature of these data is the sizable difference in RT between identical and homonym probes, which did not differ in terms of recognition accuracy. This makes unreasonable any theory that predicts a high correlation between accuracy and latency by making them both direct measures of a single hypothetical construct such as trace strength. If the assumption is made that accuracy is some simple function of trace strength, then latency must be assumed to involve at least one other factor. The concept of memory search has been useful in explaining latency data (e.g., Sternberg, 1966) and can be used to predict latency differences where no differences in accuracy exist.

Two features of the data in Figure 2 are relevant to a search hypothesis. First, there is a large difference in the right-intercept values of the RT functions, which represent estimates of the time taken to process the probe word, and secondly the three functions are roughly the same shape. In order to precisely determine the slope and intercept values, a straight line was fitted to the mean RTs for Serial Positions 5–10. The linear correlation between serial position from the end of the list (retention interval) and mean RT was $r = .98$ for homonym probes and $r = .99$ for identical and synonym probes. The slope of the synonym function (72 msec per item) was greater than those of the homonym functions (50 and 47 msec per item) which differed only slightly. The intercept values were 472, 617, and 905 msec for identical, homonym, and synonym probes.

Smith (1968) proposed that the intercept of an RT function reflects the durations of three processes: stimulus encoding, response selection, and response execution. The slope is indicative of the time taken by the operations of searching memory and making comparisons between memory traces and the encoded test stimulus. In the data shown in Figure 2, differences in intercept must reflect differences in stimulus encoding time since the response selected and executed is the same for all of the data. Because the cue telling *S*s which similarity criterion to use followed the string of words and lasted 1.3 sec, it is reasonable to assume that the encoding of the word string was completed prior to the onset of the probe word. For this reason and because of the large differences in intercept and small differences in slope between probe conditions, the differences in RT must reflect variation in the time taken to process the probe word. The data in Figure 2 also suggest that the search process for identity or homonymity proceeds at the same rate and is slightly faster than the search for semantic identity.

Table 2 shows the mean RTs for correct and incorrect responses in each of the nine combinations of probe type and rate. Correct "no" responses may be seen to vary as a function of probe type in the same manner as correct "yes" responses. Cohen (1968) has reported a similar difference in correct "no" latencies under the same three similarity criteria. This finding provides additional support for the interpretation of the intercept differences in Figure 2 in terms of variations in the time taken to encode the probe word, since this process should be independent of *S*'s eventual response.

TABLE 2

MEAN RT AS A FUNCTION OF RATE AND PROBE TYPE IN FOUR RESPONSE CATEGORIES[a]

Category	Rate, msec 350	700	1400	Weighted mean
Identical				
Correct Yes	613 (520)	699 (610)	740 (664)	683
Correct No	794	891	841	842
Incorrect No	899	955	1033	967
Incorrect Yes	846	1074	1090	1001
Homonym				
Correct Yes	771 (661)	860 (732)	909 (793)	846
Correct No	1249	1162	1261	1222
Incorrect No	1286	1368	1368	1343
Incorrect Yes	947	996	1048	1005
Synonym				
Correct Yes	1044 (944)	1105 (1026)	1158 (1081)	1103
Correct No	1368	1387	1513	1427
Incorrect No	1355	1420	1542	1436
Incorrect Yes	1169	1158	1436	1249

[a] Figures in parentheses are median correct "yes" RTs.

The effects of manipulating presentation rate provide further information about encoding and search processes. Figure 3 shows the effect of rate on the probability of a correct "yes" recognition response, before and after the correction for guessing. The interaction between rate and probe type is significant, $F(4, 232) = 2.86$, $p < .025$, by an analysis of variance done before the correction for guessing. The effect of Rate on mean correct "yes" RT in each probe condition is shown in Table 2. There is clearly no interaction between rate and probe type in these data, $F(4, 20) < 1$. Table 2 also shows median correct "yes" RT in each condition.

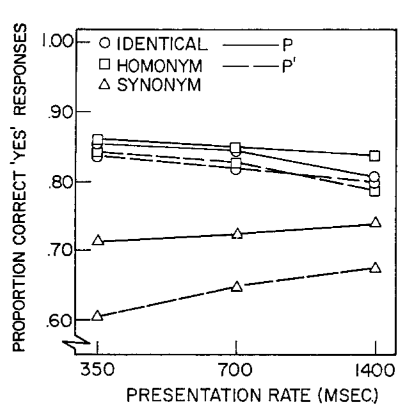

FIG. 3. Proportion correct "yes" responses as a function of rate with probe type as the parameter. Solid lines are the raw data; dashed lines are data corrected for guessing.

Correct "yes" RTs may be of two types, either true "yes" responses or false "yes" responses which are the result of response biases and false detections. Because there were differences between conditions in false-positive rates and RTs, as measured by incorrect "yes" responses on catch trials, the data in Table 2 may give a misleading picture of changes in the true correct "yes" RT. The median RTs in Table 2 are offered as an estimate of the variation in true correct "yes" RT as a function of probe type and rate. Medians are less sensitive than means to distortion caused by mixing false-positive RTs with true "yes" RTs and have the additional advantages of being less sensitive to varying numbers of observations and of giving a model-free estimate of the true correct "yes" RT.[3] Since the pattern of results did not change when medians were used, a reasonable conclusion is that the means in Table 2 accurately reflect changes in true "yes" RT. These data and the data in Figure 3 indicate a need to postulate mechanisms that will produce different effects of presentation rate on RT and correct recognition rates. Two processes that account for the data are time-dependent trace decay and time-dependent encoding. The arguments necessary to justify these assumptions will be presented in the Discussion section of the present paper.

If slower rates have their effects on correct "yes" RT via some such decay process, then the amount of time taken per comparison in memory, that is, the search rate or slope, should become greater as items become faded in memory. Thus, the slope of the serial position function must increase as presentation rate decreases, which will be seen as an interaction of rate and serial position in the correct "yes" RT data. The predicted interaction was present and statistically reliable when all 10 serial positions were analyzed, $F(18, 90) = 2.04$, $p < .025$, and marginally when only Serial Positions 5–10 were considered, $F(10, 50) = 1.89$, $.05 < p < .10$. The data for Serial Positions 5–10 and the best-fitting straight lines are shown in Figure 4. The linear correlations between serial position from the end of the list and mean RT were .973, .998, and .993 for the 350-, 700-, and 1400-msec rates, respectively. There is a regular change in slope with rate and the intercept values of the three functions are quite close. The 350-msec intercept is slightly lower than those for 700 and 1400 msec (618 vs. 635 and 637 msec, respectively), but this seems due to the deviation of the first point on the 350-msec function

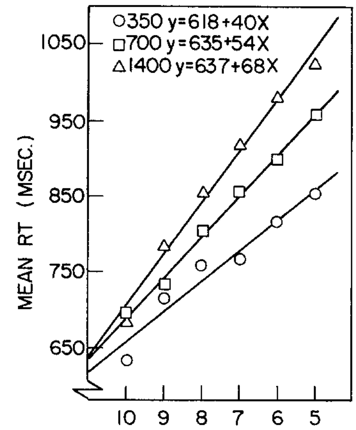

FIG. 4. Mean correct "yes" RT and best-fitting straight lines, for Serial Positions 5–10 only (10 is most recent). Presentation rate is the parameter, and the data are averaged over all probe types.

from the straight line that would best fit the other five points. The increase in slope as rate decreases is roughly the same in each probe condition as evidenced by the lack of a significant Probe × Rate × Serial Position interaction, both in the data for all 10 serial positions, $F(36, 180) = 1.30$, $.10 < p < .25$, and for Serial Positions 5–10 alone, $F(20, 100) = 1.04$. These data support the hypothesis that the decay process is the same for all types of information since if the decay rate were greater in one condition than another the effect of rate on slope would be greatest in that condition.

DISCUSSION

The present experiment constituted an effort to answer three experimental questions. These were: (*a*) Can evidence for semantic encoding in STM be obtained when such encoding is a task demand? (*b*) Are the retention functions for semantic and phonemic information similar in shape? (*c*) Does the manipulation of presentation rate have a different effect on semantic encoding than on phonemic encoding? The experimental results provide affirmative answers to each of these questions.

[3] A correction procedure for mean RT similar to the one used for recognition rate is easily derived. However, in order to use such a procedure, it is necessary to make certain assumptions about the variances of the component RT distributions which were not met in the data.

The present data imply that the memory trace in STS is multidimensional. Because the retention functions for semantic and phonemic information resemble each other closely, and show the strong recency effect which defines STS, it is unreasonable to attribute the evidence for semantic encoding in the present experiment to the contribution of LTS to performance. This is an important result since prior demonstrations of the salience of semantic information in STM have usually used retention intervals of from 10–20 sec; hence, the possibility of LTS being responsible for the effects of semantic variables could not be ruled out.

It also seems unlikely that the recognition of synonymity in the present experiment can be attributed to the recoding of information in a phonemic STS, at the time of presentation of the probe item. Such a mechanism would increase the amount of time taken to retrieve each item in memory, hence would lead to a sizable slope difference between the synonym conditions and the others in the RT data. An estimate of the probable magnitude of this slope difference can be obtained from the intercept difference between the synonym and homonym conditions, which is attributed to a similar difference in encoding processes. Since these intercept differences average 188 msec over all rates, and since the slope differences between the synonym condition and the others in Figure 2 are about one-tenth this size, it appears unreasonable to hold the hypothesis that STS is strictly phonemic with recognition of synonyms based on recoding at the time of retrieval.

An obvious question to raise next is why evidence for semantic encoding in STM is readily found when required by the experimental procedure, but is difficult to obtain in situations that let *S* control the ways in which information is encoded into STS (cf. Kintsch & Buschke, 1969). The hypothesis offered earlier was that semantic encoding requires more time than phonemic encoding, hence is less compatible with rehearsal. Given the limited ability of *S* to engage in time-sharing between different mental tasks (Peterson, 1969) and the utility of rehearsal in an STM task, it is not surprising that phonemic encoding is generally used when semantic encoding is not required. In order to provide support for this hypothesis, evidence to the effect that semantic encoding does require more time than phonemic encoding is desirable. The data shown in Figure 3 provide this evidence. Decreasing the rate of presentation improves recognition in the Semantic Probe condition but not in the Identical or Homonym condition. This result may be attributed to an increase in the probability of encoding a word semantically as more time is provided.

Figure 3 also shows that the recognition of identical and homonym probes decreases slightly as presentation rate decreases. There is no reason to predict this finding on the basis of a time-dependent encoding process, since the encoding of phonemic or visual information should either improve or remain constant as rate is decreased. The behavior in Figure 3, however, does not reflect the effects of encoding alone. Recognition performance also depends on how much information is retained in memory at the time of test. This suggests that the decrease in recognition of identical and homonym probes as rate decreases is due to an increasing loss of stored information. The forgetting mechanism thus seems time-dependent and may be identified either as time-dependent decay or time-dependent unlearning of critical traces and recovery from unlearning of competing traces. The unlearning hypothesis is difficult to apply in the case of recognition memory. It is not clear how to specify what associations are involved in a recognition task, hence what is unlearned. Furthermore, the concept of time-dependent decay is much simpler than unlearning and can easily be applied to the present situation.

The data in Figure 3 might be taken to suggest different decay rates for phonemic and semantic information. However, this hypothesis is strongly denied by the correct "yes" RT data in Table 2. These data reflect the effect of presentation rate on RTs for recognition of correctly encoded and successfully retained traces. Since there is a uniform decrease in performance in the three Probe conditions as rate decreases, it must be inferred that the rate of decay of information is the same for phonemic and semantic information. The interaction of Rate × Probe Type shown in Figure 3 is therefore taken to be the net result of a uniform decay process and an encoding process that interacts with presentation rate.

The correct "yes" RT serial position functions (Figures 2 and 4) show variations in slope as a function of presentation rate and, to a lesser extent, probe type. These variations indicate changes in the search rate as a function of rate and probe type. The major effect of probe type is on the intercept of the functions in Figure 2, but the synonym function also had a greater slope than the other two. This fact and the poorer recognition performance in the synonym conditions (Figure 1) are likely consequences of a difference between synonym pairs and homonym or identical pairs with respect to the degree of similarity between words in the pair. Thus, while it is virtually impossible to find two words which are truly identical in meaning, it is trivial to find phonemically or formally identical word pairs, and in fact these were used in the homonym and identical conditions. Given that meaningful identity between words is unobtainable, it follows that the decision that two words do or do not have the same meaning will take longer and be more susceptible to error than the corresponding decision with respect to phonemic identity. There also is some indication that the synonym condition results in slightly faster forgetting than the other two probe types, since correct recognition of synonyms decreases faster as a function of retention interval (serial position from the end of the list) in Figure 1. This interaction may be artifactual, since recognition at the recent serial positions in the identical and homonym conditions is constrained by the ceiling on proportion correct. An alternative explanation is that the slower search rate in the synonym condition allows more time for trace strength to decay; hence, each decision in the search and matching process takes slightly longer, producing an interaction between probe type and serial position in the recognition data. Thus, although the rate of forgetting may be greater for semantic information, there is no reason to suspect that the rate of loss of trace strength over time differs for semantic, phonemic, and possibly visual information. If complete semantic identity could be achieved, the present results give reason to believe that semantic information would be retained as well in STS as other information.

References

Cohen, G. A comparison of semantic, acoustic and visual criteria for matching of word pairs. *Perception and Psychophysics*, 1968, **4**, 203–204.

Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.

Jenkins, J. J., & Palermo, D. S. Synonym responses to the stimulus words of the Palermo-Jenkins word association list. The Pennsylvania State University Research Bulletin No. 55, January 1965.

Kintsch, W., & Buschke, H. Homophones and synonyms in short-term memory. *Journal of Experimental Psychology*, 1969, **80**, 403–407.

Luce, D. R. Detection and recognition. In D. R. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*, Vol. 1. New York: Wiley, 1963.

Peterson, L. R. Concurrent verbal activity. *Psychological Review*, 1969, **76**, 376–386.

Riegel, K. F. The Michigan restricted association norms. University of Michigan Research Report. National Institute of Mental Health Grant No. MH07619-OIAIO7098, April 1965.

Shulman, H. G. Similarity effects in short-term memory. *Psychological Bulletin*, 1970, in press.

Smith, E. E. Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 1968, **69**, 77–110.

STERNBERG, S. High speed scanning in human memory. *Science*, 1966, **153**, 652–654.

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

WHITFORD, H. C. *A dictionary of American homophones and homographs*. New York: Teachers College Press, 1966.

(Received March 12, 1970)