Journal of Experimental Psychology
1966, Vol. 71, No. 3, 396–404

PHONEMIC SIMILARITY AND INTERFERENCE IN SHORT-TERM MEMORY FOR SINGLE LETTERS [1]

WAYNE A. WICKELGREN

Massachusetts Institute of Technology

172 Ss copied a list of PI letters, then copied a single letter to be recalled later, then copied a list of RI letters, and then attempted recall of the single letter. The length (0, 4, 8, or 16 letters) and phonemic similarity (0, 25, 50, 75, or 100% similar letters) of the PI and RI lists were varied systematically. Both PI and RI were demonstrated in STM for single letters ($p < .001$). RI continued to increase with increasing length of RI list; PI did not increase appreciably beyond 4 letters. Both PI and RI increased with increasing phonemic similarity of the PI and RI lists for low and medium degrees of similarity of the other interference list, RI or PI list, respectively ($p < .001$). The findings suggest a 2-factor theory of forgetting in STM, involving retrieval interference and decay or storage interference.

It is well established that the similarity of an interpolated list to the original list affects the amount of retroactive interference (RI) in "long-term memory" (LTM) studies of serial and paired-associate learning (McGeoch and McDonald, 1931; Osgood, 1949, 1953). Attempts to demonstrate an analogous effect in short-term memory (STM) have apparently been unsuccessful (Broadbent, 1963). Thus, Broadbent (1963) has argued that, in STM, interference from activity interpolated between presentation and recall is essentially independent of the nature of that activity so long as the activity prevents rehearsal for the same period of time. If true, this would be an important difference between STM and LTM. On the contrary, if interference in STM can be shown to depend upon the nature of the interfering activity in the same way as in LTM, then it will be plausible to assume that STM and LTM are performed by the *same* system operating in a *quantitatively* different manner under different degrees of learning (or, less parsimoniously, different systems operating on the same principles).

[1] This work was supported primarily by Grant MH 08890-01 from the National Institute of Mental Health, United States Public Health Service. Further aid was received from a National Aeronautics and Space Administration grant, NsG 496, to Hans-Lukas Teuber.

The similarity of two lists can be varied in many different ways, and failure to obtain a relationship between similarity and RI may only mean that one has chosen dimensions of similarity that are inappropriate for the task. A study by Conrad (1964) of the types of intrusion errors occurring in immediate recall of a list of six letters suggested that the important dimensions of similarity in STM may be *auditory* dimensions, not dimensions of meaningful similarity. Conrad found that intrusion errors tend to have a vowel phoneme in common with the correct letter, even though the letters were presented visually at the rate of .75 sec/letter under conditions where the probability of perceptual error was known to be negligible. This finding has since been replicated for a wide variety of consonant and vowel similarities among both letters and numbers (Wickelgren, 1965a).

Studies of LTM have repeatedly demonstrated that forgetting of a list is more rapid when this list is preceded by other lists than when it is the first list to be learned. Murdock (1961) has demonstrated this proactive interference (PI) in STM in a completely definitive manner. Murdock presented 0, 3, 6, 9, or 12 prior words that did not not have to be remembered, followed by a single word to be recalled later, followed by backward counting. Recall of the single word declined from 0 to 3 prior words and remained approximately constant thereafter. Since *S*s were tested on 60 trials a day for 4 days, it is clear that this is short-term PI, in the sense that it can be built up on any trial of a long experiment. Short-term PI does not reach a maximum on the second or third trial of an experiment and remain constant thereafter, which suggests that Keppel and Underwood (1962) were studying a different kind of PI than Murdock (1961).

The present study was designed to replicate Murdock's demonstration of short-term PI and study: (*a*) the effect on STM of the length of the PI and RI lists, and (*b*) the effect on STM of the phonemic similarity of the PI and RI lists to the item to be recalled.

Method

Procedure.—The *S*s listened to and copied a list of PI letters presented at the rate of .5 sec/letter, followed by a 1-sec. pause and a .5-sec. tone, followed by a single letter to be recalled later which was also to be copied as it was presented, followed by a .25-sec. tone, followed by a list of RI letters which were to be copied as presented, followed by 14 sec. in which *S*s were to attempt recall of the single letter. The *S*s covered what they had copied, so this was not available at the time of recall.

Design.—Two "types" of spoken letters were used in the experiment: letters whose pronunciation contains the vowel sound "ē" (B, C, D, E, G, P, T, V, Z) and letters whose pronunciation contains the vowel sound "ĕ" (F, L, M, N, S, X). Conrad (1964) has shown that confusions in auditory perception are much more common within these classes than across them, when allowances are made for differences between British and American pronunciation. The pronunciation of these two types of spoken letters can be indicated as follows: (bee, see, dee, ee, jee, pee, tee, vee, zee) and (eff, ell, emm, enn, ess, eks). The spoken names of these letters are morphemes composed of a syllable nucleus (vowel) and zero, one, or two consonant phonemes. In a standard linguistic notation these two sets of letter names would be written as follows: (biy, siy, diy, iy, jiy, piy, tiy, viy, ziy) and (ef, el, em, en, es, eks). All the letter names in the first set contain the syllable nucleus /iy/; all the letter names in the second set contain the syllable nucleus /e/. It makes no difference to the present study whether /iy/ is considered to be two phonemes or one. For simplicity we shall speak of "phonemic similarity" and refer to the two types of letters as letters containing the phoneme "ē" ("ē" letters) and letters containing the phoneme "ĕ" ("ĕ" letters).

It is important to note that phonemic similarity (or syllable-nucleus similarity) is on a different linguistic level from distinctive feature similarity. The fact that some pairs of consonants are more similar in terms of distinctive features than other pairs of consonants simply adds unavoidable variance to the present study. Fortunately, this variance should be small in relation to the effect of phonemic similarity since possession of a common phoneme implies possession in common of the entire set of features composing the phoneme.

Thus, the letter to be recalled could be either an "ē" letter or an "ĕ" letter. The length of the PI list and the length of the RI list were varied, the different levels for each being 0, 4, 8, and 16 letters. The phonemic similarity of the PI list and the phonemic similarity of the RI list to the letter to be recalled were also varied, the different levels for each being 0, 25, 50, 75, and 100% of the letters being "similar" letters, the rest "different" letters. A similar letter is a letter having a phoneme in common with the letter to be recalled; in this case either "ē" or "ĕ." For example, if the letter to be recalled was "D," the length of PI list was 4, the similarity of the PI list was 75%, the length of the RI list was 8, and the similarity of the RI list was 50%, the following would have been an acceptable sequence: GPXT D FXGPBSEM.

There were three further restrictions on the construction of sequences: (*a*) the letter to be recalled never appeared in the PI or RI lists for that trial, (*b*) the same letter was never repeated as the letter to be recalled in any sequence of three trials, (*c*) the same letter was never repeated in any block of four letters in the PI and RI lists.

There were six groups in the experiment. In one group the number of RI letters was held constant at 8 and the similarity of the RI letters was held constant at 0% for all trials. Call this group RI8—0%. For this group the number of PI letters was varied over four levels (0, 4, 8, and 16 letters), and the similarity of the PI was varied over five levels (0, 25, 50, 75, and 100%) for trials in which the length of the PI list was 4, 8, or 16. This yields $5 \times 3 + 1 = 16$ conditions for each type of presented letter (ē or ĕ), or 32 conditions altogether. The order of presenting the conditions was randomized in blocks of 32 trials, and there were four blocks in the experiment, for a total of 128 trials. The other five groups in the experiment also had either the RI or the PI held constant at length 8 and similarity 0, 50, or 100%. If the RI was controlled, the PI was varied in the above manner, and vice versa. The six groups were: RI8—0%, RI8—50%, RI8—100%, PI8—0%, PI8—50%, and PI8—100%. Everything was identical for the six groups, except the controlled interference list. This means the same random order of conditions, the same letter to be recalled on the corresponding trials, and the same variable interference list in the corresponding trials across groups.

It is desirable to eliminate differences in the frequency of repeating items in PI or RI lists of the same length, but different similarity. In this experiment the variable interference lists (whether PI or RI in that group) were always constructed so that within any block of 32 trials the repetition structure was identical for all lists of the same length but different similarity. For example, the repetition structure of XFNSLXNM is identical to the repetition structure of BTZCPBZG.

Subjects.—The *S*s were 172 Massachusetts Institute of Technology undergraduates taking psychology courses, who participated in the experiment as a part of their course requirements. The number of *S*s in each group was approximately equal, varying from 23 to 33.

RESULTS

Only letters to be recalled that were correctly copied at the time of presentation were scored for correct recall, to eliminate perceptual errors. Figure 1 shows the frequency of correct recall in every condition in all six groups plotted as a function of the percentage of similar letters in the interference list that was varied within the group. Length of the variable interference list is a parameter within each of the six boxes. The six boxes correspond to the six groups of *S*s, and the similarity of the controlled interference list is indicated in each box. The length of the controlled interference list was always eight letters. Every point represents the average of the corresponding conditions for "ē" and "ĕ" letters to be recalled. Frequency of correct recall was 81.2% for "ĕ" letters and 75.1% for "ē" letters.

Figure 2 shows the same data as Fig. 1, but here the "independent variable" is the similarity of the *controlled* interference list. Thus, the comparisons in each box are across groups of *S*s. The parameter in each box is again the length of the "variable" interference list, but here the similarity of the "variable" interference list is *not* the "independent variable." The symbols, R0, R4, R8, R16, P0, P4, P8, P16, are used to represent RI and PI lists of length 0, 4, 8, and 16, respectively.

The comparisons in Fig. 2 are subject to variance due to individual differences, but with the size of the groups in the present study this is of little concern. The comparisons in Fig. 2 have a compensating advantage over the comparisons in Fig. 1, namely, that there is no variance due to random differences in the letter to be recalled and the other interference list.

In almost all cases a straight line

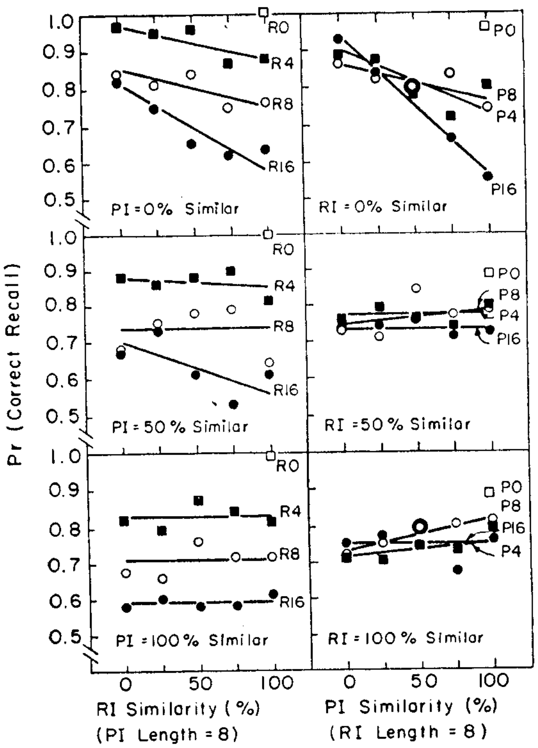

FIG. 1. Frequency of correct recall of the single letter as a function of the length and phonemic similarity of the variable interference list in each group of *S*s. (Comparisons in each box are across different conditions for the same *S*s.)

fits the points well enough to be a convenient description of the data, but determination of the exact form of the functions is beyond the scope of the study. The lines were fitted by eye and are only meant to indicate whether the trend is up, down, or straight across. When there is a separation of the points for different lengths, three separate best-fitting lines are drawn. When there is no consistent separation of the points for different lengths, only one best-fitting straight line is drawn.

Length of the RI list. Figures 1 and 2 demonstrate that RI continues to increase with increases in the length of the RI list from 0 to 4 to 8 to 16, under all conditions of similarity of the PI and RI lists. This finding is in complete agreement with earlier findings of Peterson and Peterson (1959), Murdock (1961), and Hellyer (1962). Since there are no reversals in this relationship in any comparison in Fig. 1 and 2, the difference is statistically significant even by a sign test, $p < .001$.

Length of the PI list.—Figure 1 demonstrates that four PI letters produce a substantial amount of PI compared to zero PI letters ($p < .001$ for each box using a sign test with N equal to the number of individuals in the group). However, the differences between 4, 8, and 16 PI letters are significant only in the condition of 0% similar RI and then only for the higher degrees of similarity of the PI list. Three-way analyses of variance (Length × Similarity × *S*s) for the RI8-0% group yielded a significant effect of PI length, $F\ (2,52) = 5.33$, $p < .01$. The interaction between PI length and PI similarity was also significant, $F\ (8,208) = 6.75$, $p < .001$. Thus, PI in STM appears to "build up rapidly" from 0 to 4 letters. Further increases in PI with increases in

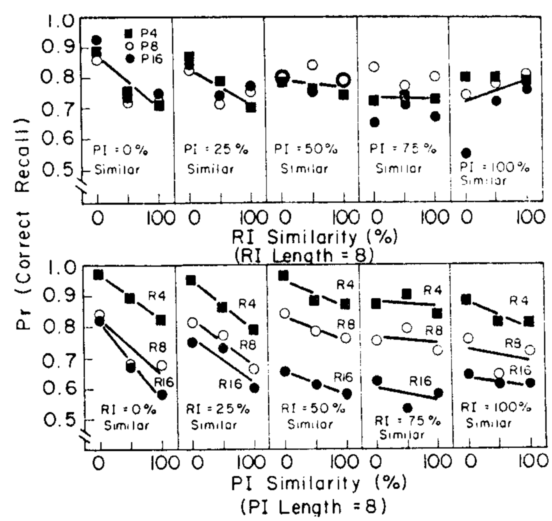

FIG. 2. Frequency of correct recall of the single letter as a function of the phonemic similarity of the controlled interference list. (Length of the variable interference list is the parameter in each box. Comparisons in each box are across groups of *S*s for the same letters to be recalled.)

length of the PI list from 4 to 8 to 16 items are small on the average and closely linked to the conditions of RI and PI list similarity. The rapid build-up of PI from 0 to 4 items is in complete agreement with the findings of Murdock (1961).

Phonemic similarity of the RI list.—Figures 1 and 2 show that greater RI is obtained with greater similarity of the RI list under conditions of 0, 25, and 50% similarity of the PI list. Under conditions of 75% and 100% similarity of the PI list, there is no consistent effect of RI similarity on recall. A sign test of the slopes of the best-fitting lines in Fig. 1 and 2 is significant at the .001 level in the direction of greater RI with greater phonemic similarity of the RI list. In addition, three-way analyses of variance on the PI8–0% and PI8–50% groups showed that the effect of RI similarity was significant at well beyond the .001 level in each case, $F\ (4,116) = 19.22$ and $F\ (4,88) = 32.94$. The consistency of the relationship between RI similarity and recall in Fig. 1 and 2 is particularly convincing. Even the interaction with PI similarity is found in both figures. It should be noted that the relationships between RI similarity and recall in Fig. 2 are based on data from entirely different groups than the groups used to determine the same relationship in Fig. 1. Thus, while Fig. 1 and 2, as a whole, present exactly the same data, the substantively similar halves of each figure are statistically independent tests of the same hypothesis.

Phonemic similarity of the PI list.—Figures 1 and 2 show that PI also increases with increasing phonemic similarity of the PI list, on the average. An overall sign test on the slopes of the best-fitting lines is again significant, $p < .001$. Again there is an obvious weakening of this relationship with increasing similarity of the RI list. For example, three-way analyses of variance on the RI8–0% group showed a very significant effect of PI similarity, $F\ (4,104) = 19.57$, $p < .001$, but there was no significant effect in the RI8–50% and RI8–100% groups in Fig. 1. The comparisons in Fig. 2 which may contain less variance than those in Fig. 1, show a consistent effect of PI similarity under conditions of 0, 25, and 50% RI similarity and small, more variable, effects in the same direction under conditions of 75 and 100% similarity.

Phonemic similarity of intrusions.—Averaging over all groups and conditions in the experiment, the present experiment replicates the finding of Conrad (1964) and Wickelgren (1965a) that intrusions tend to share a phoneme with the correct letter. In this case, "ē" letters are more likely to be substituted for "ē" letters and "ĕ" letters for "ĕ" letters than "ē" letters for "ĕ" letters and "ĕ" letters for "ē" letters. Figure 3 presents the relative frequency of similar intrusions (similar intrusions/total intrusions) for every condition in the experiment. The length of the RI or PI list has no effect on the relative frequency of similar vs. different intrusions. The similarity of the RI list appears to have either no effect on the frequency of similar intrusions or a curvilinear effect. The similarity of the PI list, however, has a pronounced and consistent effect on the phonemic similarity of the intrusions, in the obvious direction, namely, greater frequency of similar intrusions with greater similarity of the PI list. The most plausible explanation of this difference between RI and PI is that the last few letters in the RI list are remembered as being in the RI list and are, therefore, excluded from consideration in recall. This would tend to counteract the relation between interference list similarity and relative frequency of similar intrusions, producing no relation, the opposite relation, or perhaps a curvilinear relation for RI lists.

Discussion

The present findings clearly demonstrate: (*a*) that both PI and RI exist in STM, (*b*) that RI continues to increase with increasing length of the RI list, (*c*) that PI does not increase appreciably with increases in the length of the PI list beyond four letters, and (*d*) that both PI and RI increase with increasing phonemic similarity of the PI and RI lists for low degrees of similarity of the other interference list, RI or PI list, respectively. These interference effects in STM are completely compatible with what is known about interference effects in LTM.

The difference in the relation between length of interference list and amount of interference for PI and RI lists suggests that forgetting in STM involves two factors: decay (or storage interference) and associative interference. RI results from both decay and associative interference; PI results only from associative interference. Decay refers to a decrease with time in the strength of association between the internal representative of the cue (whatever the cue is) and the internal representative of the correct letter. Associative interference refers to the loss of information that occurs when remote (competing) associations are established between the cue representative (cue rep) and representatives of nearby letters in the PI and RI lists.

It is reasonable to assume that the correct letter is recalled when the association from the cue rep to the correct letter rep has a strength above some absolute threshold and is stronger than the association from the cue rep to any competing letter rep. Strength of association may be presumed to be a random variable. Therefore, the more PI or RI items, the greater the probability that one of the in-

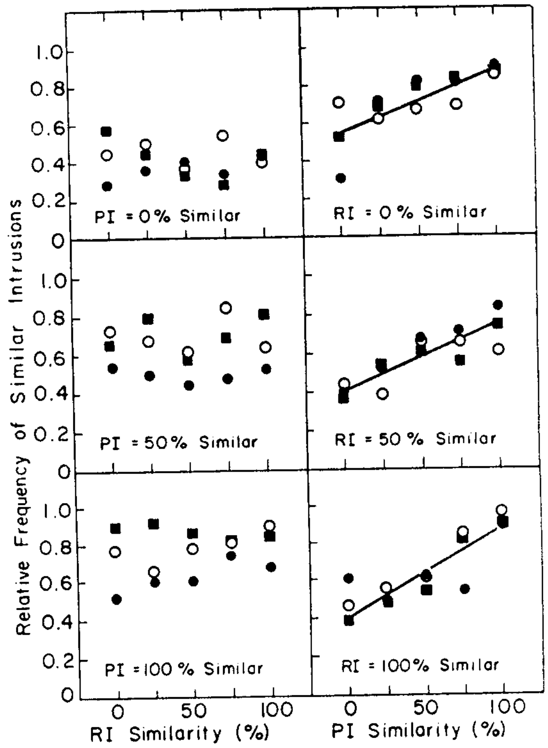

Fig. 3. Relative frequency of phonemically similar intrusions in recall (similar intrusions/total intrusions). (Except for the dependent variable, the plot is identical to Fig. 1.)

terfering item reps will have the greatest strength of association to the cue rep and therefore be (incorrectly) recalled. Also, the more RI items, the greater time the correct association has had to decay.

The two-factor theory obviously explains why an RI item produces greater interference than a PI item. It also explains why PI increases only from 0 to 4 letters while RI is still increasing from 8 to 16 letters in this study and from 54 to 81 digits in Hellyer's (1962) study.

Pure interference theory can explain these asymmetries between PI and RI by simply substituting storage interference for decay. So far we have been careful to limit discussion to *associative interference*, which has its interfering effect by establishing competing associations, not by reducing the strength of the correct association. Thus, we have been discussing a type of *retrieval interference*, not *storage interference.* Retrieval interference is necessary in order to explain PI, and there is much direct evidence for it in the findings concerning remote associates in serial learning (e.g., Lepley, 1934; McGeoch, 1936; Wilson, 1943). Thus, in any event one probably needs a two-factor theory of forgetting in STM. Whether the two factors are associative interference and decay or associative interference and storage interference makes very little difference, since at present there is no testable difference between decay and storage interference.

The phonemic similarity findings suggest a phonemic-associative theory of STM. The theory makes four assumptions: (*a*) The internal representative of a verbal item is an ordered set of the internal representatives of the phonemes of the item. (*b*) There is only one internal representative of any phoneme. (*c*) Short-term associations are formed between two internal representatives when the internal representatives are activated in close temporal contiguity. (*d*) Short-term associations decay in time.

The fourth assumption has already been discussed. If one is to explain the phonemic similarity of intrusions in STM, there appears to be no obvious alternative to the first assumption that verbal items are coded as sequences of phonemes. Of course, phonemes may be coded as sets of "distinctive features" (Jacobson,

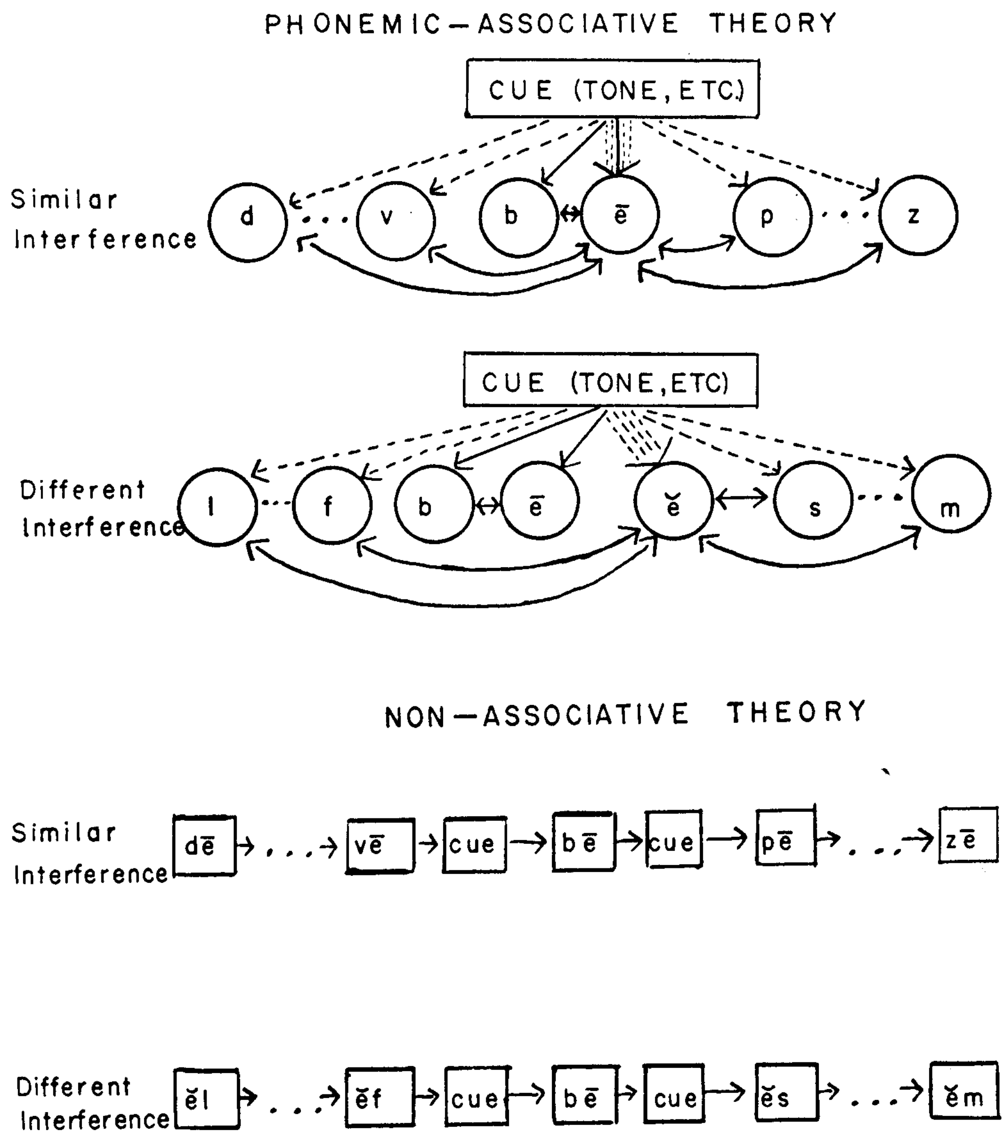

FIG. 4. Diagram of an associative and a nonassociative theory of STM. (Intraitem and direct associations are indicated by solid lines, remote associations by dashed lines.)

Fant, & Halle, 1952; Miller & Nicely, 1955), but that is merely a finer level of analysis, not a formulation contradictory to the phonemic coding hypothesis.

The second and third assumptions that STM is associative and that the internal representative of a phoneme is unique are suggested by the greater RI and PI produced by phonemically similar letters than by phonemically different letters in the present experiment. If we assume that the cue rep is directly associated to the internal representatives of both the vowel and consonant phonemes in the correct letter, which are associated to each other (intraitem association), then phonemically similar interference letters will produce intraitem associations that compete with the correct intraitem association. Phonemically different interference items will produce intraitem associations that do not compete with the correct intraitem association. Remote associations between the cue rep and the phoneme reps in PI and RI letters should be equally strong for all *consonant* phonemes regardless of phonemic similarity, but stronger for the same *vowel* as in the letter to be recalled than for the different vowel. Considering both remote associations and intraitem associations it is clear that phonemically similar interference items should have stronger associations to the cue rep than phonemically different items and thus produce greater interference. Figure 4 illustrates why phonemically similar letters produce greater retrieval interference in an associative memory.

The uniqueness assumption and the associative memory assumption were discussed together because a nonassociative memory assumption requires that it be possible to have multiple internal representatives of the same item or phoneme (as illustrated in Fig. 4). Nonassociative theories of STM are those in which an ordered set of cells (boxes, locations, registers, etc.) are set aside as temporary (buffer) storage, and any item can be encoded into any cell. Ordered recall is possible because a list of items is generally or always stored in the cells in a fixed order and generally or always retrieved from the cells in the same order. Associative and nonassociative memory assumptions make different predictions about lists with repeated items, the findings being consistent with the associative memory assumption for STM (Wickelgren, 1965b). The most compelling finding is that items following separated repeated items tend to be confused with each other in recall, far above chance.

Notice that a nonassociative theory does not assert that there is no connection (association) between the cue rep and the representatives of the correct and incorrect items. Recall is far above chance so there has to be a functional connection between them. The question is whether the functional connections are established by strengthening associations between fixed internal representatives (associative theory) or by coding items into memory cells with fixed relations to each other, like the successive positions on a magnetic tape.

The present findings regarding phonemic similarity and interference are very difficult to explain with a nonassociative theory. If PI and RI letters can be processed in cells completely separate from the cell(s) used to store a single letter, then one should expect no effect due to the phonemic similarity of the letters stored in those cells. On the other hand, an associative theory provides a ready explanation of the phonemic similarity effects in STM.

REFERENCES

Broadbent, D. E. Flow of information within the organism. *J. verbal Learn. verbal Behav.*, 1963, **2**, 34–39.

Conrad, R. Aoustic confusions in immediate memory. *Brit. J. Psychol.*, 1964, **55**, 75–84.

Hellyer, S. Frequency of stimulus presentation and short-term decrement in recall. *J. exp. Psychol.*, 1962, **64**, 650.

Jakobson, R., Fant, C. G. M., & Halle, M. *Preliminaries to speech analysis.* Cambridge, Mass.: MIT Press, 1952.

Keppel, G., & Underwood, B. J. Proactive inhibition in short-term retention of single items. *J. verbal Learn. verbal Behav.*, 1962, **1**, 153–161.

LEPLEY, W. M. Serial reactions considered as conditioned reactions. *Psychol. Monogr.*, 1934, **46**(1, Whole No. 205).

MILLER, G. A., & NICELY, P. E. An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Amer.*, 1955, 27, 338–352.

MCGEOCH, J. A. The direction and extent of intra-serial associations at recall. *Amer. J. Psychol.*, 1936, **48**, 221–245.

MCGEOCH, J. A., & MCDONALD, W. T. Meaningful relation and retroactive inhibition. *Amer. J. Psychol.*, 1931, **43**, 579–588.

MURDOCK, B. B., JR. The retention of individual items. *J. exp. Psychol.*, 1961, **62**, 618–625.

OSGOOD, C. E. The similarity paradox in human learning: A resolution. *Psychol. Rev.*, 1949, **56**, 132–143.

OSGOOD, C. E. *Method and theory in experimental psychology.* New York: Oxford Univer. Press, 1953.

PETERSON, L. R., & PETERSON, M. J. Short-term retention of individual verbal items. *J. exp. Psychol.*, 1959, **58**, 193–198.

WICKELGREN, W. A. Acoustic similarity and intrusion errors in short-term memory. *J. exp. Psychol.*, 1965, **70**, 102–108. (a)

WICKELGREN, W. A. Short-term memory for repeated and non-repeated items. *Quart. J. exp. Psychol.*, 1965, **17**, 14–25. (b)

WILSON, J. T. Remote associations as a function of the length of interval between learning and recall. *J. exp. Psychol.*, 1943, **33**, 40–49.

(Received October 28, 1964)