# Short-Term Order and Item Retention[1]

ELIZABETH L. BJORK

*University of Michigan*

AND

ALICE F. HEALY[2]

*Rockefeller University*

This paper evaluates two hypotheses: (a) that transposition errors made in the recall of letter strings occur as a by-product of acoustic confusion errors and do not represent the loss of order information, and (b) that order and item information are independently retained in short-term memory. An experiment was conducted in which four-consonant strings containing exactly zero or two acoustically confusable items were recalled in order after retention intervals of 3, 8, or 18 intervening digits, all characters being successively presented at a rate of 400 msec per item and read aloud by the subject. An analysis of errors in relation to intra- and extrastimulus sources of acoustic confusion, retention interval, and serial position produced results that refute hypothesis (a) and support hypothesis (b). The implications of the present results for an adequate theory of the short-term retention of ordered strings are indicated.

Of the possible information about a stimulus that could be stored in memory at the time of input, the information actually stored can be inferred in part from the type of errors subjects make in recall. For example, Conrad (1964, 1965) has observed that when subjects are required to recall strings of letters, a disproportionate number of their overt errors are intrusions of letters that share acoustic features with the correct letter. Not only do such errors of acoustic confusion constitute a large pro-

portion of the errors subjects make in recalling letter strings, but the proportion changes dramatically as a function of retention interval: At short retention intervals subjects make many more acoustic confusion errors than would be expected by chance, while at long retention intervals they make only slightly more acoustic confusion errors than would be expected by chance (Conrad, 1967). A similar shift in the ratio of confusion to nonconfusion errors as a function of retention interval has been reported by Estes (1970, 1972). Primarily on the basis of these and similar observations regarding the frequency and distribution of acoustic confusion errors, the basic format of short-term memory is argued to be auditory or acoustic. In particular, it is argued that the phonemic structure of letters is somehow represented, stored, and rehearsed in short-term memory (Conrad, 1967; Estes, 1970, 1972; Sperling, 1968).

In addition to making acoustic confusion errors when required to recall strings of letters, subjects make frequent transposition errors

(Conrad, 1965; Estes, 1970); that is, rather than recall the letter actually presented in a particular position of the stimulus string, subjects recall another letter that was presented in the string, but in a different position. Whereas there is general agreement that the frequent occurrence of acoustic confusion errors in such experiments implies that the constituent phonemes of a letter are somehow represented in short-term memory, the implications of transposition errors as to the representation of order information in memory remain debatable. Conrad (1965) has argued that such errors of transposition occur only at the time of output and actually reflect acoustic confusability. That is, transposition errors occur as a by-product of acoustic confusion errors and, thus, only seem to be order errors; in fact, they are item errors. The occurrence of transposition errors, therefore, does not require a mechanism or process for retaining position or order information in memory independently of content or item information.

Conrad's (1965) notion that acoustic confusion errors and transposition errors do not reflect different memory processes is based primarily on two observations: (a) Of the recall errors he classified as transpositions, many more involved acoustically confusable letters than involved nonconfusable letters, and (b) of the recall errors he classified as "other" or nontransposition errors, many more involved nonconfusable letters than involved confusable letters.

Although Conrad's observations are, on the surface, compelling, they cannot, in fact, be interpreted as ruling out the possibility that item and order information are retained separately. Firstly, the correlation Conrad obtained between transposition errors and acoustic confusability is based on an analysis of only those six-letter strings for which exactly two letters were incorrectly recalled, and such letter strings comprised only a small subset (approximately 17%) of the total number presented. Thus, the subset of letter strings involved in Conrad's analysis may not have

had the same properties as the total set. For the total set of six-letter strings, the chance probabilities that a transposition error would involve two acoustically similar letters versus two acoustically dissimilar letters were equal; for the subset of strings for which exactly two letters were incorrectly recalled, these chance probabilities may not have been equal. It is possible, for example, that many of the strings in this subset are ones that contain four or five letters from only one of the two subsets of five acoustically similar letters that Conrad used in constructing his stimuli. For such letter strings, a transposition error would be more likely to involve two confusable letters by chance than two nonconfusable letters.

A second problem involves the way in which Conrad classified transposition versus "other" errors. The classification problem inherent in Conrad's scoring scheme is demonstrated in the following examples. (a) Suppose that the letter string BTFCSM was presented, and BTXVSM was recalled. This recall would contain exactly two errors, one in the third and one in the fourth positions, and this error pair would be classified by Conrad as a nontransposition or "other" error involving nonconfusable letters. It is true that the correct letters for positions three and four, F and C, are not acoustically similar letters. However, each of them has been replaced in the subject's recall by a letter to which it is acoustically similar but which was not presented in the same stimulus string. Thus, if each position for which an error has occurred is considered separately, two "other" errors involving acoustically similar letters have been made. (b) Suppose now that for the same presented string, BTFCSM, the subject recalled BSFNSM, making errors in the second and fourth positions. The error pair in this recall would be classified by Conrad as an "other" error involving acoustically similar letters, although the correct letter for position two, T, has been replaced by a letter from within the same presented string and to which T is not acoustically similar, and the correct letter for

position four, C, has been replaced by a letter from outside the presented string and to which C is not acoustically similar. Considering each error individually, a transposition error involving dissimilar letters was made in position two and an "other" error involving dissimilar letters was made in position four.

The types of ambiguous and misleading error classifications demonstrated in the above examples occur under Conrad's classification scheme because errors are labeled as transpositions or not, and as involving acoustically similar letters or not, on the basis of the relationship between the two letters that were presented but incorrectly recalled, rather than on the basis of the relationship between the letter presented in a certain position and the letter intruded into that position. In the second example given above, one letter was repeated in the incorrect recall—an event which might seem unlikely. However, in Conrad's experiment, repetitions were very frequent mistakes. Conrad reports that of 982 recalls in which only one letter was incorrectly recalled, more than half of the errors were repetitions.

The relationship between item and transposition errors has also been examined by Murdock and vom Saal (1967). They presented word trigrams composed of either three same-category or three different-category items and tested recall after 3, 9, and 18 sec of interpolated activity. Their experiment was specifically designed to determine whether transposition errors occur during output, as Conrad (1965) suggests, or during storage, and whether they would occur when acoustic confusability was minimized. In addition, they examined the relationship between errors (any noncorrect response) and transpositions, and concluded that item and order information are interdependent. However, there is a problem with the analysis on which this conclusion was based. The authors calculated the probability of an error given adjacent, remote, or no transpositions and found that error rates increased with transpositions. The problem is

that this analysis was not done separately for each retention interval. Since the authors also found transpositions to increase as a function of retention interval, their analysis could actually be reflecting the probability of an error as a function of retention interval, and thus an increasing rate would be expected. In fact, the Murdock and vom Saal study can be considered to lend some support to the concept of independent retention of order and item information since their subjects made more order errors in recalling same-category trigrams than different-category trigrams, although, in terms of item errors, they retained same-category trigrams better than different-category trigrams.

An analogous result was obtained by Wickelgren (1965) in the immediate, ordered recall of auditorily presented nine-letter strings. Item recall did not differ for strings containing acoustically similar letters and strings containing dissimilar letters, while ordered recall was significantly poorer for the similar strings. Similarly, for strings of seven consonant–vowel digrams, ordered recall was significantly poorer for strings in which all the digrams had a common vowel than for dissimilar strings, while item recall was significantly better for the similar strings.

In summary, none of the above studies provides conclusive evidence to support or negate the concept of independent retention of item and order information in short-term memory, although the Conrad study and the Murdock and vom Saal study are frequently cited as evidence for the nonindependence hypothesis. The present experiment is an attempt to remedy this situation by providing a more accurate picture of the relationship between item and order errors as they occur in the retention of letter strings. To this end, recall errors are classified as acoustic confusion errors or as transposition errors solely on the basis of the relationships existing between the letter that was presented in a particular position of the stimulus string and the letter that was intruded into that position in the subject's

recall. The classification of errors into transposition errors and confusion errors according to this rule creates four nonoverlapping error categories: confusion–transposition, confusion–nontransposition, nonconfusion–transposition, and nonconfusion–nontransposition errors. The relationships obtained among the occurrences of these error types are used to assess both the specific hypothesis of the interdependence of acoustic confusion errors and transposition errors and the more general hypothesis that order and item information are independently or separately retained in short-term memory. In particular, the occurrence of these types of errors are examined in relation to intra- and extrastimulus sources of acoustic confusion, retention interval, and serial position.

The present experiment is similar in design to Conrad's (1967) study in that consonant strings constructed from a population of letters containing subsets of acoustically confusable letters are used as stimuli, and retention is tested after varying intervals of digit shadowing. There are three major differences. Firstly, Conrad used three subsets of acoustically similar letters that varied in size and, thus, the probability that a given letter would occur in a stimulus with one or more acoustically similar letters was much greater for some letters than for others. Consequently, there was no way to systematically estimate the probability of an acoustic confusion error in relation to the absence or presence of acoustically similar letters within the same to-be-remembered unit. The present experiment avoids this problem by using subsets of equal size and systematically controlling the context or nature of the sequence in which any particular consonant appears. This control assures that the obtained frequencies of transposition errors that do or do not involve acoustically similar letters can be evaluated in terms of what would be expected on the basis of chance.

Secondly, in Conrad's study retention interval was a between-subjects variable, while in the present experiment all subjects are tested

at each retention interval, and three intervals, instead of two, are examined.

Finally, the present experiment yokes each member of an acoustically confusable set of letters to a corresponding member of a set of letters that are not acoustically confusable. Thus, for each set of acoustically confusable letters there is a control set of nonconfusable letters that is treated in identical fashion with respect to all independent variables.

## METHOD

### Apparatus

All stimuli were visually presented on one cell of a four-cell Iconix Bina-View display device that was operated by means of a Digitronics Paper Tape Reader. Three clocks were used as part of the system, one clock automatically timing the stimulus duration at 400 msec, a second clock automatically timing the intertrial interval at 16 sec, and a third automatically timing the interitem interval at approximately 2 msec. Once the paper-tape reader was started, all trials proceeded automatically.

The subjects sat approximately 1 m in front of the display unit which was located at about eye level on a table. Each character appeared individually on the Bina-View screen and had a height of approximately 3.5 cm and a base of 2.2 cm.

### Subjects

Three male and three female young adults, who had responded to advertisements in local newspapers and posters, served as subjects in the experiment and were paid for their participation at the rate of $2.00 per hour. Each subject served in a single 1-hr session on each of 2 days. After a span of several months, three more male and three more female young adults, obtained in the same way, served as subjects in a replication of the experiment.

### Materials

Three different lists, each containing 144 trials, were constructed for the experiment. A

trial consisted of a four-consonant stimulus and 3, 8, or 18 intervening digits, all characters being presented successively.

The four consonants in each stimulus were chosen from a population of 12 consonants containing four distinct subsets of three letters each. Two subsets, BPV and FSX, were subsets of acoustically confusable letters. The other two subsets, KMR and HLQ, were constructed so that no member of either subset was acoustically confusable with any other member of the 12 consonant population and so that no members of the same subset were adjacent letters in the alphabet. These latter subsets served as artificial confusion sets or as control sets. All four subsets of letters will be referred to as confusion sets with the sets BPV and FSX called the acoustic confusion sets and the sets KMR and HLQ called the control confusion sets whenever a distinction is necessary.

From these four subsets, two types of stimuli were constructed: Paired-Context stimuli and All-Different-Context stimuli. In the Paired-Context stimuli, two of the four consonants are from one of the two acoustic confusion sets and two are from one of the two control confusion sets. In the All-Different-Context stimuli, each of the four consonants is from a different one of the four subsets of letters. Thus in the Paired-Context stimuli, exactly two letters are acoustically similar and two letters are not acoustically similar, and in the All-Different-Context stimuli, no two letters are acoustically similar.

With four subsets from which to select letters and four stimulus positions to put them in, there are 24 possible permutations each for the Paired-Context stimuli and the All-Different-Context stimuli. Each possible permutation for the two context types was used once for each of the three retention intervals, making a total of 144 stimuli, and each subject received a complete set of these 144 stimuli across the two experimental sessions.

The particular letters comprising each of the 144 stimuli were chosen randomly from within the four subsets except for the constraint that all letters appear equally often in each of the two context types and for each of the three retention intervals. A master list was constructed from these 144 stimuli by randomly ordering them except for the following constraints: (a) In each block of 18 successive trials, the two context types and the three retention intervals appeared equally often, and (b) stimuli having the same consonant in the same serial position did not appear on immediately successive trials.

From this master list, two other lists were constructed by mappings that substituted letters from within the four subsets in such a manner that, with the completion of all three lists, each individual letter occurred an equal number of times in all serial positions as well as in each of the two context types and for each of the three retention intervals. Furthermore, for each subset of acoustically similar letters there was a control subset of acoustically dissimilar letters that was identically treated with respect to all independent variables. The three lists thus comprised a completely counterbalanced design for which one replication required three subjects, each subject receiving one list. Four replications were completed in the present experiment.

The digits appearing on each trial were chosen randomly from the digits "1" through "9" with the constraint that no digit occurred twice in a row.

*Procedure*

On each trial of the experiment, the subject read aloud the consonants and digits as they appeared, one at a time, on the Bina-View screen. After the last digit of a trial was presented, the Bina-View screen went blank and the subject was free to recall the four-consonant stimulus. The subject wrote his recall on a 3 × 5 card containing four printed boxes. He was instructed to write each recalled consonant in the box that corresponded to the input position the letter had occupied in the stimulus string. However, he was allowed to do the actual writing in any order.

Beneath each box on the response card, the digits "1," "2," and "3" were printed. The subject was instructed to circle one digit for each consonant recalled depending upon how confident he was that the consonant recalled in a given position was the consonant that had appeared in that position of the stimulus. The subject was to circle the digit "3" if he was certain that the consonant recalled was correct for that position, the digit "1" if he was just guessing, and the digit "2" if he was neither certain of his response nor just guessing. The subject recalled all four consonants before circling any of the confidence digits.

Halfway through the 16-sec intertrial interval, the paper-tape reader clicked twice signaling the subject to finish his response and get ready for the next trial. Just before the start of the next trial, the reader again clicked twice, the Bina-View screen blinked with each click, and then the series of consonants and digits to be read aloud began.

At the start of each experimental session, the subject received two trials consisting of only digits to provide practice on the reading task. Within each experimental session the subject received 72 experimental trials. At the end of the first 39 experimental trials, the subject was given a short rest period.

## RESULTS AND DISCUSSION

The performance of the second group of subjects essentially replicated that of the first. Hence, in the sections to follow, the data from all 12 subjects were combined.

### Scoring Procedure

In analyzing the data, each of the four letters recalled on a given trial was scored separately by comparing the letter presented in a particular position of the stimulus string to the letter recalled for that particular position. If the letter recalled matched the letter presented, the response for that position was scored as correct; otherwise, it was scored as an error. Errors were then classified as follows: If the

letter incorrectly recalled was from the same confusion set as the letter presented, the error was called a confusion error. If the letter incorrectly recalled was not from the same confusion set as the letter presented, the error was called a nonconfusion error. Note that in the context of the present experiment, the label of confusion error does not necessarily denote that the letter presented and the letter incorrectly recalled for it were acoustically confusable since two of the confusion sets in the experiment were artificial or control confusion sets. When it is necessary to distinguish between the two types of confusion errors, the terms acoustic confusion error and control confusion error are used.

If the letter incorrectly recalled for a certain position had also been presented in the stimulus but in a different position, the error was called a transposition error. If the letter incorrectly recalled for a certain position of the stimulus had not appeared in any other position of the stimulus on that trial, the error was called a nontransposition error.

Any overt error can thus be unambiguously classified as a confusion or nonconfusion error, as a transposition or nontransposition error, and as a member of one of the four nonoverlapping error categories created by the orthogonal definitions of confusion and transposition errors.

### Overall Error Percentages

In Table 1 are shown the percentages of total responses that were errors at each retention interval for the Paired-Context and All-Different-Context stimuli. The letters recalled at each retention interval are the same; the only difference is the context or the nature of the letter string in which they appeared. While the differences in recall performance are small, they are all in the direction of poorer performance on the Paired-Context stimuli.

To verify the direction of the results shown in Table 1, a similar table was constructed for each of the 12 subjects from their individual recall data. Summing across the number of

TABLE 1

ERROR PERCENTAGES FOR EACH STIMULUS CONTEXT

|  | Number of intervening digits | | |
|---|---|---|---|
| Context type | 3 | 8 | 18 |
| Paired-Context | 17 | 41 | 63 |
| All-Different-Context | 14 | 35 | 60 |

intervening digits, the error percentages for the Paired-Context stimuli were compared to the error percentages for the All-Different-Context stimuli for each subject by a Wilcoxon Matched Pairs Signed Ranks Test (see Siegel, 1956, pp. 75–83). The difference in performance for the two types of context was found to be statistically reliable, $\sum$ (negative ranks) $= -11, p < .05$.

In Table 2, the error percentages obtained for each context type at each retention interval are separated into errors made on the acoustically confusable letters and errors made on the control letters. From Table 2, it can be seen that the differences shown in Table 1 between the Paired-Context stimuli and the All-

TABLE 2

ERROR PERCENTAGES AND TOTAL NUMBER OF ERRORS (N) FOR EACH STIMULUS CONTEXT AND LETTER SET

| Context type Letter set | Number of intervening digits | | |
|---|---|---|---|
|  | 3 | 8 | 18 |
| *Paired-Context* | | | |
| Confusable letters | 23 | 44 | 65 |
|  | N = 130 | N = 252 | N = 374 |
| Control letters | 10 | 38 | 62 |
|  | N = 59 | N = 218 | N = 353 |
| *All-Different-Context* | | | |
| Confusable letters | 14 | 36 | 60 |
|  | N = 78 | N = 207 | N = 345 |
| Control letters | 15 · | 35 | 60 |
|  | N = 84 | N = 200 | N = 345 |

Different-Context stimuli result from the poorer recall performance at the shorter retention intervals on the acoustically confusable letters in the Paired-Context stimuli. The control letters in the Paired-Context stimuli appear to be as well recalled as the control letters in the All-Different-Context stimuli. Furthermore, there does not appear to be a real difference between recall performance on acoustically confusable letters and on control letters in the All-Different-Context stimuli.

To verify the differences or lack thereof indicated in Table 2, a separate Table 2 was constructed for each subject. Summing across the number of intervening digits, the error percentages obtained in the four conditions of Table 2 were compared by means of four Wilcoxon Matched Pairs Signed Ranks Tests with the following results: A comparison of the error percentages obtained on the acoustically confusable letters versus the control letters within the Paired-Context stimuli showed this difference to be statistically reliable, $\sum$ (negative ranks) $= 0, p < .01$. A comparison of the error percentages obtained on the acoustically confusable letters versus the control letters within the All-Different-Context stimuli showed this difference not to be statistically reliable, $\sum$ (positive ranks) $= 39.5, \sum$ (negative ranks) $= 38.5, p > .05$. A comparison of the error percentages obtained on the acoustically confusable letters presented in the Paired-Context stimuli versus the error percentages obtained on the same letters presented in the All-Different-Context stimuli showed this difference to be statistically reliable, $\sum$ (negative ranks) $= 0, p < .01$. A comparison of the error percentages obtained on the control letters presented in the Paired-Context stimuli versus the error percentages obtained on the same letters presented in the All-Different-Context stimuli showed this difference not to be statistically reliable, $\sum$ (positive ranks) $= 40, \sum$ (negative ranks) $= 38, p > .05$.

On the basis of these results, it would seem that if acoustic similarity is to reduce recall

performance, the acoustically similar letters must be present in the same to-be-remembered unit. The mere presence of acoustically similar letters within the experimental population appears not to be sufficient to decrease recall performance. Furthermore, the presence of acoustically similar letters within the same to-be-remembered stimulus does not affect the level of recall performance for the other acoustically dissimilar letters that are present in the same stimulus.

*Error Analyses*

In the error analyses to follow, the obtained proportions of total errors that were errors of a certain type are compared to the proportions of such errors that would be expected to occur by chance. These chance expectations were obtained by calculating the likelihood of each type of response if the subject were to make his response by picking a letter at random from the total stimulus population of 12 letters. (Responses from outside the experimental population of letters did not occur.) Therefore, if a subject were performing on the basis of chance, the probability of a correct response would equal 1/12 or .08, and the probability of an error would equal 11/12 or .92. In addition, since there were three letters in each confusion set and four letters were presented on each trial, the chance probability of a confusion error was 2/12 or .17, and the chance probability of a transposition error was 3/12 or .25. Further, given that an error was made, the chance probability that the error was a confusion error was 2/11 or .18 and the chance probability that it was a transposition error was 3/11 or .27.

In the Paired-Context stimuli, for each letter presented on a given trial, one of the other three letters presented on that same trial was from the same confusion set. Thus, for Paired-Context stimuli, the chance probability of a confusion error, given that a transposition error has occurred, was equal to 1/3 or .33. In the All-Different-Context stimuli, for each letter presented on a given trial, none of the other three letters presented on that same trial were from the same confusion set. Thus, for All-Different-Context stimuli, the chance probability of a confusion error, given that a transposition error occurred, was always equal to zero. For the Paired-Context stimuli and the All-Different-Context stimuli, the chance probability of a confusion error, given that a nontransposition error occurred, was 1/8 or .12 and 2/8 or .25, respectively.

*Confusion errors.* In Table 3 are shown the mean proportions across subjects of total errors that are confusion errors at each retention interval for the two types of context and for the two types of letters within each context. The mean proportion and the standard error of the mean proportion shown in each cell were calculated by converting each subject's proportions into angles using the arcsin transformation (angle = arcsin (proportion)$^{\frac{1}{2}}$, Snedecor & Cochran, 1967, pp. 569–571), computing the mean and the standard error of the mean of these transformed scores, and reconverting the mean and the standard error of the mean obtained on the transformed scores back into proportions.

Given the two types of stimuli in the present experiment, Paired-Context stimuli and All-Different-Context stimuli, it is possible to assess the likelihood of an acoustic confusion error in relation to the presence or absence of other acoustically confusable letters within the same stimulus. The data needed to make this assessment are contained in Rows 1 and 3 of Table 3. In Row 1 are presented the mean proportions of the errors made on the acoustically confusable letters that are confusion errors, when these letters appeared in the Paired-Context stimuli. These proportions of .71, .37, and .24 for retention intervals of 3, 8, and 18 intervening digits, respectively, replicate in general the findings of Conrad (1967) and Estes (1970), being much greater than the proportion of acoustic confusion errors expected by chance (.18) at the shortest retention interval and only slightly greater at the longest retention interval.

TABLE 3

MEAN PROPORTION (X̄) OF TOTAL ERRORS (N) THAT WERE CONFUSION ERRORS
TOGETHER WITH STANDARD ERROR (SE) OF THE MEAN PROPORTION
FOR EACH CONTEXT TYPE AND LETTER SET

| | Number of intervening digits | | |
|---|---|---|---|
| Letter set | 3 | 8 | 18 |
| *Paired-Context* | | | |
| Confusable letters | X̄ = .71<br>SE = .10<br>N = 130 | X̄ = .37<br>SE = .05<br>N = 252 | X̄ = .24<br>SE = .03<br>N = 374 |
| Control letters | X̄ = .31<br>SE = .14<br>N = 59 | X̄ = .25<br>SE = .05<br>N = 218 | X̄ = .18<br>SE = .03<br>N = 353 |
| *All-Different-Context* | | | |
| Confusable letters | X̄ = .13<br>SE = .06<br>N = 78 | X̄ = .30<br>SE = .06<br>N = 207 | X̄ = .15<br>SE = .02<br>N = 345 |
| Control letters | X̄ = .02<br>SE = .02<br>N = 84 | X̄ = .11<br>SE = .03<br>N = 200 | X̄ = .12<br>SE = .03<br>N = 345 |

Note. On the basis of chance, one would expect .18 of the total errors made to be confusion errors.

However, from the data presented in Row 3, a very different picture emerges for the acoustically confusable letters when they appeared in the All-Different-Context stimuli. The proportions of .13, .30, and .15 for retention intervals of 3, 8, and 18 intervening digits, respectively, are substantially different from the proportions of confusion errors obtained for these same letters in the Paired-Context stimuli (Row 1) and are essentially the proportions of acoustic confusion errors that would be expected by chance, particularly at the shortest and longest retention intervals.

Considering these results and those shown in Table 2, there is apparently no effect of acoustic similarity, either on the level of recall performance or on the likelihood of acoustic confusion errors, when the source of acoustic confusion lies outside a given to-be-remembered stimulus. This observation holds even when there are letters within the stimulus population as a whole to which a given to-be-remembered letter is acoustically confusable. Indeed, the immediately preceding stimulus frequently contains such letters.

Note, however, the proportions of errors that are confusion errors for the control letters shown in Rows 2 and 4 of Table 3. If the subject's guessing behavior is occurring according to the scheme described above, the proportions of confusion errors obtained for the control letters should be in line with the proportion expected by chance (.18). However, at the shortest retention interval for both context types, the obtained proportions of confusion errors for the control letters obviously deviate from this expected chance proportion, being too large in the Paired-Context stimuli and too small in the All-Different-Context stimuli. An explanation of these deviations will emerge in the analysis of the transposition errors to follow—an explanation that, in turn,

indicates a different interpretation of the effects of extrastimulus sources of acoustic confusion from the one discussed above.

*Transposition errors.* In Table 4 are shown the mean proportions across subjects of total errors that were transposition errors at each retention interval for the two types of context and for the two types of letters within each context. The mean proportion and the standard error of the mean proportion shown in each cell were calculated on the arcsin transformations of each subject's proportions and were then reconverted back into proportions as described above for Table 3.

The proportions of total errors that are transposition errors steadily decrease with increasing retention interval. Unlike acoustic confusion errors, however, the obtained proportions are much greater at all retention intervals than the proportions expected by chance (.27), indicating both that the two

types of errors are not reflecting loss of the same type of information and that order information is lost at a different rate from acoustic-item information.

A remarkable feature of the data shown in Table 4 is the small variation in the observed proportions of total errors that are transpositions as a function of letter set for either context type. In the Paired-Context condition, more than twice as many errors are made at the shortest retention interval on the acoustically confusable letters as on the control letters (130 versus 59, respectively). Yet the proportions of errors attributable to transpositions are about the same (.81 versus .86, respectively). Similarly, in the All-Different-Context stimuli, the proportions of errors attributable to transpositions are essentially the same for the acoustically confusable and the control letters. The data of Table 4 are not, therefore, consistent with a hypothesis that transposition

TABLE 4

MEAN PROPORTION ($\bar{X}$) OF TOTAL ERRORS (N) THAT WERE TRANSPOSITION ERRORS
TOGETHER WITH STANDARD ERROR (SE) OF THE MEAN PROPORTION FOR
EACH CONTEXT TYPE AND LETTER SET

| | Number of intervening digits | | |
|---|---|---|---|
| Letter set | 3 | 8 | 18 |
| *Paired-Context* | | | |
| Confusable letters | $\bar{X} = .81$ | $\bar{X} = .65$ | $\bar{X} = .46$ |
| | SE = .06 | SE = .06 | SE = .04 |
| | N = 130 | N = 252 | N = 374 |
| Control letters | $\bar{X} = .86$ | $\bar{X} = .63$ | $\bar{X} = .41$ |
| | SE = .07 | SE = .06 | SE = .04 |
| | N = 59 | N = 218 | N = 353 |
| *All-Different-Context* | | | |
| Confusable letters | $\bar{X} = .76$ | $\bar{X} = .50$ | $\bar{X} = .49$ |
| | SE = .12 | SE = .06 | SE = .03 |
| | N = 78 | N = 207 | N = 345 |
| Control letters | $\bar{X} = .78$ | $\bar{X} = .56$ | $\bar{X} = .47$ |
| | SE = .12 | SE = .04 | SE = .04 |
| | N = 84 | N = 200 | N = 345 |

Note. On the basis of chance, one would expect .27 of the total errors made to be transposition errors and .73 to be nontransposition errors.

errors occur only as a result of acoustically similar letters being confused in recall.

*Conditional analysis.* Data concerning the relationship between acoustic confusion errors and transposition errors are presented in Table 5. The top section of Table 5 shows the proportions of transposition errors that are confusion errors for the two letter types within the Paired-Context stimuli.

For the control letters presented in the Paired-Context stimuli, the proportions of transposition errors that are also confusion errors are essentially what would be expected by chance (.33) at all retention intervals. For the acoustically confusable letters presented in the Paired-Context stimuli, however, the proportion of transposition errors that are also confusion errors assumes a value more than twice what would be expected by chance at the shortest retention interval, and then

TABLE 5

MEAN PROPORTION ($\bar{X}$) OF TRANSPOSITION ERRORS (N) OR NONTRANSPOSITION ERRORS (N')
THAT WERE CONFUSION ERRORS TOGETHER WITH THE STANDARD ERROR (SE) OF THE
MEAN PROPORTION FOR EACH CONTEXT TYPE AND LETTER SET;
THE PROPORTIONS EXPECTED BY CHANCE ARE SHOWN IN PARENTHESES

| Letter set | Number of intervening digits | | |
| --- | --- | --- | --- |
| | 3 | 8 | 18 |
| P (Confusion\|Transposition error) | | | |
| *Paired-Context* | | | |
| Confusable letters | $\bar{X}$ = .73 (.33) SE = .09 N = 99 | $\bar{X}$ = .39 (.33) SE = .06 N = 152 | $\bar{X}$ = .34 (.33) SE = .06 N = 170 |
| Control letters | $\bar{X}$ = .32 (.33) SE = .14 N = 44 | $\bar{X}$ = .32 (.33) SE = .09 N = 128 | $\bar{X}$ = .35 (.33) SE = .04 N = 140 |
| P (Confusion\|Nontransposition error) | | | |
| *Paired-Context* | | | |
| Confusable letters | $\bar{X}$ = .61 (.12) SE = .16 N' = 31 | $\bar{X}$ = .34 (.12) SE = .04 N' = 100 | $\bar{X}$ = .11 (.12) SE = .04 N' = 204 |
| Control letters | $\bar{X}$ = .17 (.12) SE = .20 N' = 15 | $\bar{X}$ = .12 (.12) SE = .05 N' = 90 | $\bar{X}$ = .06 (.12) SE = .03 N' = 213 |
| *All-Different-Context* | | | |
| Confusable letters | $\bar{X}$ = .72 (.25) SE = .11 N' = 24 | $\bar{X}$ = .66 (.25) SE = .08 N' = 101 | $\bar{X}$ = .31 (.25) SE = .06 N' = 178 |
| Control letters | $\bar{X}$ = .23 (.25) SE = .18 N' = 21 | $\bar{X}$ = .29 (.25) SE = .10 N' = 92 | $\bar{X}$ = .22 (.25) SE = .04 N' = 186 |

Note. There are no entries under P (Confusion\|Transposition Error) for the All-Different-Context stimuli since confusion-transposition errors cannot occur for such stimuli.

drops to the value expected by chance at the middle and longest retention intervals.

The data of Table 4 show that for the acoustically confusable letters in the Paired-Context stimuli, transposition errors constitute .65 and .46 of all errors made at the middle and longest retention intervals, respectively. Yet according to Table 5, the proportions of these transposition errors that are acoustic confusion errors are no greater than what would be expected by chance. One cannot, therefore, explain the preponderance of transposition errors at these retention intervals as a by-product of acoustic confusion errors.

What then is the relationship between transposition errors and acoustic confusion errors or acoustic similarity? Clearly, transposition errors do not only occur as a by-product of acoustic confusion errors. Is it also true that their occurrence is completely unaffected by the acoustic similarity of the letters within a stimulus string? Looking at Table 4, it can be seen that at the shortest retention interval the proportions of errors attributable to transpositions are both quite high and not different for the acoustically confusable letters and the control letters. This finding would indicate that transposition errors or the retention of order information is not affected by the presence of intrastimulus acoustic similarity. However, there are two reasons why this conclusion is not warranted.

Firstly, since N is small (59) in the case of the control letters, the finding of no difference may just be the result of sampling error. Secondly, in looking at Table 4, it can be seen that at the shortest retention interval about twice as many errors are made on the acoustically confusable letters as on the control letters in the Paired-Context stimuli. If one now examines the sources of this increase in errors (from 59 to 130) between control letters and acoustically confusable letters, it is clear that the vast majority of the additional errors on the acoustically confusable letters are confusion–transposition errors. Specifically, the frequency of confusion–transposition errors

increases from 16 for control letters to 60 for acoustically confusable letters. Thus, it appears that although transposition errors constitute a majority of the errors made at the shortest retention interval, even when there are no acoustically confusable letters being confused in recall, the presence of intrastimulus acoustic similarity can serve as an additional source of transposition errors. Furthermore, since no more transposition errors are made on control letters in the Paired-Context stimuli than in the All-Different-Context stimuli (Table 4), it appears that the presence of two acoustically similar items in the same to-be-remembered stimulus does not increase the loss of order information for all letters in the stimulus string but rather produces rapid loss of order information specific to the two similar letters, and that the two similar items do not suffer an absolute loss of order information but rather a loss of order information relative to each other.

The bottom section of Table 5 presents the proportions of nontransposition errors that are confusion errors for the two types of context and for the two types of letters within each context. For the control letters in both context types and at all retention intervals, the proportions of nontransposition errors that are confusion errors are essentially what would be expected by chance. For the acoustically confusable letters in both the Paired-Context stimuli and the All-Different-Context stimuli, the picture is quite different. Of the few nontransposition errors made on the acoustically confusable letters at the shortest and middle retention intervals for both context types, many more are confusion errors than would be expected by chance. Indeed, at the shortest retention interval, the great majority of nontransposition errors are acoustic confusion errors.

*Reinterpretation of Confusion Errors*

In the discussion of the confusion-error analysis presented above, it was noted that at the shortest retention interval, the obtained

proportions of confusion errors for the control letters in both context types obviously deviated from the expected chance proportion. The rarity of nontransposition errors at the shortest retention interval, demonstrated in Table 4, suggests an explanation for these seemingly anomalous entries of Table 3. The scheme used to calculate the proportions referred to as chance throughout this section assumed that a subject, when he did not remember the correct item presented in a certain position, selected his response at random from the entire population of experimental letters. One assumption made by such a scheme for estimating chance proportions is that order and item information are lost from memory simultaneously. However, the data indicate otherwise. Namely, item information is available to the subject longer than order information and can be used, particularly at the shortest retention interval, to restrict the set of letters from which the subject selects his response. At this retention interval, the subject seldom intrudes letters into his recall from outside the set of letters presented on that trial.

In Table 3 the proportion of errors that are confusion errors for the control letters in the Paired-Context stimuli appears too large at the shortest retention interval when compared to the proportion expected by chance. However, since at the shortest retention interval nearly all recall errors are transpositions, one would expect the observed proportion of errors that are confusions to be greater than the proportion expected by chance, as chance was originally defined. One would, in fact, expect the obtained proportion to be close to .33, the probability of a confusion error given that a transposition error has occurred in a Paired-Context stimulus.

On the other hand, in the All-Different-Context stimuli, confusion errors are necessarily nontransposition errors. Since, as indicated by Table 4, only one out of every four or five errors is a nontransposition error, one would expect the chance occurrence of confusion errors in the All-Different-Context

condition to be not .18, but more in the neighborhood of .05 to .06. The proportion of confusion errors obtained at the shortest retention interval for the control letters in the All-Different-Context stimuli is in line with this expectation.

The data of Tables 4 and 5 together show that the proportions in Table 3, particularly those obtained at the shortest retention interval, should not be evaluated against the proportions of such errors expected by chance, as chance was originally defined, and that interpretations of Table 3 based on such comparisons are unjustified. To interpret the data of Table 3 correctly, the confusion-error proportions for the acoustically confusable letters should be evaluated against the confusion-error proportions for the control letters since these latter proportions represent a more accurate measure of the chance occurrence of confusion errors in relation to context type and retention interval in this situation. On the basis of this comparison, it is obvious that the original interpretation of the acoustic-confusion error proportions obtained at the shortest retention interval in the All-Different-Context stimuli is incorrect. Extrastimulus sources of acoustic confusion do affect the likelihood of acoustic confusion errors. At the shortest retention interval, over six times as many confusion–nontransposition errors occurred for the acoustically confusable letters than for the control letters in the All-Different-Context stimuli.

Finally, from Table 5, it can be seen that of the very few nontransposition errors that occur at the shortest retention interval for the acoustically confusable letters in both context types, the great majority are errors in which the intruded letter belongs to the same acoustic confusion set as the correct letter. This finding supports the notion that item information is lost from memory in a partial manner. That is, these item errors occur because the information as to the unique identity of an item has been lost, while the information concerning its set membership is still available. For a

further discussion of this aspect of the data, the reader is referred to Healy (1970) in which a Markov model including partial loss of item information and independent loss of order and item information is applied to the data of the present experiment.

*Confidence Judgments*

The data concerning confidence judgments were consistent with the frequency data and, by and large, did not add any insights. Those data are not, therefore, reported here. One observation, however, deserves mentioning.

For transposition errors made in the recall of acoustically confusable letters, the average confidence level within any one retention interval was approximately the same regardless of whether the transposition errors were acoustic confusion errors. This was also true of nontransposition errors made at the middle and longest retention intervals. However, for nontransposition errors (extrastimulus intrusions) made at the shortest retention interval, the average confidence level was strongly influenced by whether or not the nontransposition error was an acoustic confusion error. Average confidence was much greater for nontranspositions that were acoustic confusions (2.8 and 2.4 for Paired- and All-Different-Contexts, respectively) than for nontranspositions that were not acoustic confusions (1.8 and 1.5 for Paired- and All-Different-Contexts, respectively).

In fact, in the Paired-Context condition and at the shortest retention interval, subjects did not designate any nontransposition errors that were acoustically similar to the letter replaced as a "guess," whereas they designated nontransposition errors that were not acoustically similar to the letter replaced as "guesses" more frequently than any other type of error. The subjects' high confidence in the correctness of their acoustic confusion–nontransposition errors versus their nonconfusion–nontransposition errors is consistent with the hypothesis that the phonemic structure of letters is represented in short-term memory

and can be lost in parts. Acoustic confusion–nontransposition errors would occur when the information as to the exact identity or the unique phoneme of an item had been lost, while the information concerning its set membership or its common phoneme was still available. Nonconfusion–nontransposition errors would occur when both types of information had been lost.

*Serial Position Data*

In Table 6, the proportions of total responses that were correct, transposition errors, or nontransposition errors are presented as a function of retention interval and input serial position. At each retention interval, the function relating the probability of correctly recalling a letter to the input position it occupied within the stimulus string is clearly symmetrical and bow-shaped. A comparison of the serial position functions for transposition errors versus nontransposition errors clearly

TABLE 6

PROPORTIONS OF TOTAL RESPONSES THAT WERE CORRECT, TRANSPOSITION ERRORS, OR NONTRANSPOSITION ERRORS AT EACH SERIAL POSITION

| Response type | Serial position | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| 3 Digit retention interval | | | | |
| Correct | .91 | .79 | .81 | .88 |
| Transposition | .06 | .15 | .15 | .09 |
| Nontransposition | .03 | .06 | .03 | .03 |
| 8 Digit retention interval | | | | |
| Correct | .72 | .51 | .57 | .68 |
| Transposition | .17 | .28 | .27 | .14 |
| Nontransposition | .11 | .21 | .16 | .18 |
| 18 Digit retention interval | | | | |
| Correct | .45 | .33 | .31 | .45 |
| Transposition | .25 | .33 | .34 | .19 |
| Nontransposition | .31 | .34 | .35 | .36 |

indicates that the bow-shape of the serial position function for correct responses is caused primarily by transposition or order errors. This observation holds at all retention intervals. The functions for both error types tend to flatten with increasing retention interval, but those for transposition errors do so at a much slower rate. At all retention intervals, the functions for transposition errors remain distinctly bow-shaped, while the function for non-transposition errors has become completely flat by the longest retention interval. The occurrence of such different serial position functions for order and item errors as those obtained in this study is apparently not unique to the present experimental situation. At least three other investigators (Aaronson, 1968; Fuchs, 1969; McNichol, 1971) have reported similar results using quite different paradigms.

A measure of the total item information retained independent of order information (i.e., a measure of free recall performance) as a function of input serial position can be calculated from Table 6 by adding together the correct response proportions and the transposition error proportions obtained at each serial position. The resulting functions are distinctly flat in contrast to those for the correct retention of both order and item information (rows labeled "correct").

Whereas most of the results have been concerned with the relation between the occurrence of acoustic confusion errors and transposition errors and the hypothesis that these two types of errors reflect the loss of different types of information, the results of the serial position analysis presented above provide support for the more general hypothesis that order and item information are separately retained in short-term memory. For a further discussion of the serial position data of the present experiment the reader is referred to Estes (1972) in which these data together with other serial position effects are discussed in relation to an associative model for stimulus coding.

## SUMMARY AND CONCLUSIONS

Two principal questions have been examined in the present paper: (a) the question of the interdependence of acoustic confusion and transposition errors, particularly the assertion of Conrad (1965) that transposition errors are caused by acoustic confusion errors and do not represent the loss of order information; and (b) the more general hypothesis that order and item information are separately, or independently, retained in short-term memory.

Three findings of the present study speak against the notion that transposition errors are caused by (or result as a by-product of) acoustic confusion errors. One finding is the two different time courses for the retention of order and acoustic-item information demonstrated in Tables 3 and 4. The proportion of total errors that are acoustic confusions, for the acoustically confusable letters presented in the Paired-Context stimuli, was found to be much greater than what would be expected by chance at the shortest retention interval and only slightly greater at the longest retention interval, replicating previous results (Conrad, 1967; Estes, 1970). The proportions of total errors that are transposition errors, however, were much greater than the proportions expected by chance at all retention intervals, indicating that the two types of errors do not reflect loss of the same type of information and that order information is lost at a different rate from acoustic-item information.

The second finding arises from the data of Tables 4 and 5. The data of Table 4 show that for the acoustically confusable letters presented in the Paired-Context stimuli, transposition errors constitute .65 and .46 of all errors made at the middle and longest retention intervals, respectively. According to Table 5, however, the proportions of these transposition errors that are acoustic confusion errors are no greater than what would be expected by chance. An explanation of these transposition errors as by-products of acoustic confusion errors is clearly not tenable.

Thirdly, the data of Table 4 show that within either the Paired- or All-Different-Context conditions, the proportions of errors attributable to transpositions are substantially greater than would be expected by chance for the control letters as well as the acoustically confusable letters. These data are not, therefore, consistent with a hypothesis that transposition errors occur only as a by-product of acoustically similar letters being confused in recall. Furthermore, the data of Table 4 show that, although very few errors occur at the shortest retention interval, the vast majority of those that do occur are errors of transposition regardless of the stimulus context in which they occur. On the other hand, acoustic confusion errors comprise a majority of the errors made at the shortest retention interval only in the Paired-Context condition.

Clearly, these findings combine to refute the notion that transposition errors occurring in the recall of letter strings do not reflect loss of order information but rather occur as by-products of acoustic confusion errors and result from the same source.

When the confusion-error proportions obtained for the acoustically confusable letters in the All-Different-Context stimuli were compared against the proportion of such errors expected by chance, there appeared to be no effect of acoustic similarity on the likelihood of acoustic confusion errors, when the source of acoustic confusion lay outside the given to-be-remembered stimulus. These data thus indicated an interdependence between acoustic confusion errors and transposition errors that was directly opposite to that proposed by Conrad (1965). Specifically, these data implied that when transposition errors can involve acoustically similar letters, as they can in the Paired-Context stimuli, then acoustic confusion errors occur more frequently than would be expected by chance; but when transposition errors cannot involve acoustically similar letters, as they cannot in the All-Different-Context stimuli, then acoustic confusion errors occur only by chance. However, this conclu-
4

sion was refuted by the later conditional error analysis which revealed that of the few non-transposition errors made at the shortest retention interval, many more are confusion errors for the acoustically confusable letters than for the control letters in the All-Different-Context stimuli.

Thus, by carefully controlling the context or structure of the stimulus in which any letter appeared and by unambiguously classifying confusion errors and transposition errors on the basis of the relationship existing between the letter presented in a particular position of the stimulus and the letter incorrectly recalled for that position, the present experiment has shown that acoustic confusion errors and transposition errors can occur independently of one another and that they reflect the loss of different types of information.

This is not to say, however, that the retention of order information as reflected by transposition errors is not affected by acoustic similarity. The data of the present experiment show that while the presence of two acoustically similar letters in the same to-be-remembered stimulus does not increase the loss of order information for the other acoustically dissimilar letters in the string, it seems to produce rapid loss of order information specific to the two similar letters. In particular, the similar letters suffer a loss of order information relative to each other. The number of errors made on acoustically confusable letters when they are present in the same stimulus is twice the number of errors made on the non-confusable letters present in the same stimulus. This increase in errors for the acoustically confusable letters is due in large part to an increase in confusion–transposition errors. Thus, it appears that the presence of intra-stimulus acoustic similarity can serve as an additional source of transposition errors.

In addition to refuting the contention that transpositions occur only as a result of acoustic confusions, the results of the present experiment are at variance with Conrad's (1965) observation that "other" errors, his

nontransposition errors, involved more acoustically dissimilar pairs than acoustically similar pairs. According to the error classification scheme used in the present study, the great majority of the nontransposition errors made on the acoustically confusable letters at the shortest retention interval are also confusion errors in both the Paired- and All-Different-Context conditions.

The findings supporting the conclusion that acoustic confusion errors and transposition errors can occur independently of one another and that they measure the retention loss of different types of information also support the more general notion that order and item information are separately or independently retained in short-term memory. In addition, the present experiment provides two other sources of evidence for this notion.

Looking at Table 4 in terms of transposition or order errors versus nontransposition or item errors, one can see that the rate of loss of order information is faster than that of item information. At the shortest retention interval, the subject has seldom lost enough item information about the presented stimulus string to intrude items into his recall that were not actually present in that stimulus, but he frequently has lost information concerning the order in which those items were presented.

Additional evidence for hypothesizing independent retention processes for order and item information comes from the serial position functions obtained for transposition and nontransposition errors. The serial position functions for these measures of order and item information retention are dramatically different in shape and change in different ways as a function of retention interval.

In conclusion, the present study has been productive in showing that acoustic confusion errors and transposition errors can occur independently of one another and that the two error types reflect the retention loss of different types of information. Further, the present study has produced several pieces of evidence in support of the hypothesis that order and item information are independently retained in short-term memory. While the results of the present study cannot be considered conclusive with regard to this latter hypothesis (primarily because of the difficulty of completely separating measures of the retention of order and item information), they are strongly suggestive of such a conclusion and provide several constraints for future theories. In particular, any adequate theory of the short-term retention of an ordered string of items needs to account for the following results of the present experiment. (a) Loss of order information from memory occurs more quickly than does loss of item information, and, furthermore, loss of order information is the dominant source of errors overall. In particular, nearly all of the few errors observed at the shortest retention interval are transposition errors, and the errors at the longest retention interval are about an equal mixture of transposition and nontransposition errors. (b) Retention of order and item information as a function of input position is distinctly different. When transposition and nontransposition errors are plotted as a function of input serial position, both the form of the resulting functions and their manner of interaction with retention interval differs.

REFERENCES

AARONSON, D. Temporal course of perception in an immediate recall task. *Journal of Experimental Psychology*, 1968, **76**, 129–140.

CONRAD, R. Acoustic confusions in immediate memory. *British Journal of Psychology*, 1964, **55**, 75–84.

CONRAD, R. Order error in immediate recall of sequences. *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 161–169.

CONRAD, R. Interference or decay over short retention intervals? *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 49–54.

ESTES, W. K. Evaluation of some models for acoustic confusion effects in short-term memory. In *Communications in Mathematical Psychology*, Rockefeller University Technical Reports, March, 1970.

ESTES, W. K. An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington: Winston, 1972. Pp. 161–190.

FUCHS, A. H. Recall for order and content of serial word lists in short-term memory. *Journal of Experimental Psychology*, 1969, **82**, 14–21.

HEALY, A. F. Two models of confusion and order. In *Communications in Mathematical Psychology*, Rockefeller University Technical Reports, December, 1970.

MCNICHOL, D. The confusion of order in short-term memory. *Australian Journal of Psychology*, 1971, **23**, 77–84.

MURDOCK, B. B., JR., & VOM SAAL, W. Transpositions in short-term memory. *Journal of Experimental Psychology*, 1967, **74**, 137–143.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw Hill, 1956.

SNEDECOR, G. W., & COCHRAN, W. G. *Statistical methods*, 6th Edition. Ames, Iowa: Iowa State University Press, 1967.

SPERLING, G. Phonemic model of short-term auditory memory. Proceedings, 76th Annual Convention, APA, 1968, 63–64.

WICKELGREN, W. A. Short-term memory for phonemically similar lists. *American Journal of Psychology*, 1965, **78**, 567–574.