GRAMMATICAL INTRUSIONS IN THE RECALL OF STRUCTURED LETTER PAIRS: MEDIATED TRANSFER OR POSITION LEARNING? 

KIRK H. SMITH
Walter Reed Army Institute of Research

Sets of letter pairs were constructed from 4 classes of letters, M, N, P, and Q to form MN and PQ sequences. In free recall, Ss produced more intrusions of the form MQ and PN than would be expected if intrusions were produced by randomly combining the letters appearing in the presented pairs. An adequate account of the intrusion data is provided by a theory of position learning proposed by Braine. An alternative proposal by Jenkins and Palermo (1964) based upon mediated transfer fails to predict the high proportion of MQ and PN intrusions.

When Ss attempt to recall sets of letter pairs which have been constructed from two classes of letters, A and B, with A letters occurring first in all pairs and B letters second, Smith (1963) has shown that the most prominent class of intrusions produced during recall consists of AB combinations which were omitted during presentation. That is, Ss “recall” pairs they have not heard but which are formally consistent with the pairs to which they have listened. Smith argued that these rule-governed intrusions were, on a limited scale, grammatically appropriate generalizations of the presented pairs and illustrated the phenomenon of “rule-governed creativity” which Chomsky (1964) characterized as the essential feature of a command of the grammar of a natural language. The recall of grammatically structured nonsense material was proposed as a technique for investigating the variables relevant to grammatical competence. Braine (1965a) has recently proposed the label “verbal reconstructive memory” for the technique.

Two explanations of rule-governed intrusions are presently available. Braine (1963) in an account of the development of grammatical skill proposed that what is learned is the position occupied by the morphemes and phrase units of the language. Braine’s theory is elaborated in such a way that the definition of position is readily articulated with a phrase-structure model of grammar. When applied to the letter-pair recall experiment, Braine’s proposal asserts that Ss will learn that the A letters come first in the pair and B letters second. Learning here would be a type of perceptual learning in which the A letters come to sound familiar in the first temporal position and B letters, in the second temporal position. Hence, any AB combination will sound familiar whe-
ther the particular letters have occurred together or not.

An alternative explanation, also offered as an account of the development of grammatical competence in natural language, is suggested by Jenkins and Palermo (1964), who proposed that the formation of word classes could be explained as acquired stimulus and response equivalence based upon implicit mediating responses assumed to intervene between explicit verbal responses. Their explanation focused upon the occurrence of rule-governed novelties observed by Esper (1925) in the learning of two-unit utterances from a miniature linguistic system. Although the account is not completely explicit (cf. Smith, 1963), it appears that the following logic would apply to the letter-pair recall experiment. When the utterances, $ab$, $cb$, and $ad$, are learned in that order, mediation theory assumes that a contingency between an implicit $b$ response and an explicit $d$ response is learned during the acquisition of $ad$, because $b$ continues to occur implicitly in the presence of $a$; therefore, the utterance $cd$ is expected to have some strength due to the strength of the two contingencies, $c \rightarrow b$ and $b$ (implicit) $\rightarrow d$, which can be run off sequentially to form a chain between $c$ and $d$.

In the present experiments, Ss were presented letter pairs constructed from four classes of letters, M, N, P, and Q, and exposed only to MN and PQ sequences. Mediation theory predicts that if some of the MN and PQ combinations are omitted, Ss will produce them as intrusions so long as a triplet $(m_i n_j, m_i n_j, m_i n_j)$ can be constructed for any omitted pair, $m_i n_j$. Similar predictions are made for any omitted pair $p_i q_j$, mutatis mutandis; however, so long as no MQ and PN combinations are presented, S cannot acquire any contingencies such that the pairs $m_i q_j$ or $p_i n_j$ will be mediated. Mediation theory therefore predicts that mediated generalization cannot occur on the basis of position alone. Rule-governed intrusions are expected to be confined to MN and PQ combinations. On the other hand, Braine's position-learning proposals predict that since both M letters and P letters occur in the first position and N letters and Q letters in the second, Ss should learn only a grammar XY such that the X class includes M and P and the Y class, N and Q. Because MN, PQ, MQ, and PN intrusions are all XY intrusions, Braine would predict that all four types of pairs will be highly probable intrusions.

The two theories may be said to predict the learning of two slightly different grammars, only one of which characterizes MQ and PN combinations as "grammatical." The present experiments provide a direct test of the adequacy of both theories in explaining the occurrence of rule-governed intrusions in retention data.

**Experiment I**

**Method**

The letter pairs were constructed from the following classes of letters: M: v, s, r; N: g, k, l; P: z, x, d; and Q: m, f, j. Of the 18 possible MN and PQ pairs, vg, sk, rl, zm, xf, and dj were omitted, leaving a set of 12 combinations. Randomly ordered lists of these pairs were read to groups of Ss at the rate of 20 per min. with a 1-sec pause between pairs. The Ss were told they would be required to recall the pairs at the end of the list. The reading of the list was followed by 3 min. of written free recall. The Ss were instructed to write the pairs "... in any order that they occur to you."

Two presentation lists were used. List A was 60 pairs long and contained five appearances of each of the 12 letter combinations. List A was presented to a group of 18 college students (Group A). A second list (B), 180 pairs long with 15 appearances of
TABLE 1
MEAN NUMBER OF LETTER COMBINATIONS PRODUCED IN EXP. I

<table>
<thead>
<tr>
<th>Type of Letter Combination</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Pairs</td>
<td>A</td>
</tr>
<tr>
<td>MN and PQ Intrusions</td>
<td>4.44</td>
</tr>
<tr>
<td>MQ and PN Intrusions</td>
<td>1.78</td>
</tr>
<tr>
<td>Other Intrusions with Correct Vocabulary</td>
<td>2.89</td>
</tr>
<tr>
<td>Other Intrusions with New Vocabulary</td>
<td>3.72</td>
</tr>
<tr>
<td>All Combinations (excludes repetitions)</td>
<td>15.39</td>
</tr>
<tr>
<td>SD*</td>
<td>5.45</td>
</tr>
</tbody>
</table>

* SD of the number of combinations of all types produced.

each combination, was administered to three groups of college students. The three replications will be designated Group B1 (N = 40), Group B2 (N = 38), and Group B3 (N = 40). Group B1 had previously participated in an experiment involving the recall of words.

Results

The mean number of letter combinations produced by Ss in five different classes is shown in Table 1. There appears to be considerable variation among the groups. Group A recalled a little more than half as many presented pairs as the remaining three groups while producing many more intrusions, particularly those introducing a letter which did not appear in the presented pairs, i.e., new vocabulary. Among the three groups receiving 15 exposures to the pairs, Group B1 produced significantly more different letter combinations 2 (p < .05), and Groups B2 and B3 differed significantly in mean presented pairs recalled (p < .05). These differences suggested that statistical analysis should be performed on each group independently. Since this leads to identical conclusions for all groups, the question is not crucial.

The t test was used to determine whether the mean number of MQ and PN intrusions was greater than would be expected if Ss were producing intrusions by randomly combining the letters which appeared in the presented pairs. Under the latter hypothesis, it is expected that 18/120 (0.15) of all intrusions would be MQ and PN combinations. (Twelve letters can form 144 combinations of which 12 are combinations of a letter with itself. Since the latter were never produced in recall, it seemed reasonable to exclude them from the set of potential intrusions. Twelve other combinations are, of course, presented to Ss, leaving 120 potential intrusion combinations.) An expected number of MQ and PN intrusions was computed for each S by multiplying 0.15 by the total number of intrusions involving two letters from the correct vocabulary. The mean differences between obtained and expected numbers of MQ and PN intrusions are shown in Table 2.

All mean differences for the MQ and PN intrusions were significant at the .01 level. Comparable values for MN and PQ intrusions, predicted by both mediation and position-learning theory, are also shown in Table 2. A direct
comparison of MQ and PN intrusions with MN and PQ intrusions is also presented. The number of MN and PQ intrusions was multiplied by three in the latter comparison to compensate for the fact that possible MQ and PN combinations outnumbered possible MN and PQ combinations 18 to 6. None of the comparisons of the latter type were significant.

**Discussion**

The differences between Group A and the three groups receiving 15 exposures seem readily attributable to differential exposure. Not only does the number of presented pairs recalled increase with exposure, but intrusions with new vocabulary appear to decrease in frequency. The data in Table 1 suggest that while increased exposure does not reduce the number of intrusions incorporating correct vocabulary, intrusions of this type become more concentrated in the MN, PQ, MQ, and PN classes.

The somewhat increased production of all types of combinations in Group B1 is readily explained by the fact that this group had received practice in the recall procedure prior to participation in the present experiment. Group B1 consistently produced more intrusions than Groups B2 and B3 in every classification, but the relative frequencies do not appear to differ.

Mediation theory, as developed by Jenkins and Palermo, also suggests that the learning of AB combinations should lead to the strengthening of AA and BB combinations. This conclusion is a consequence of applying three-stage mediation paradigms (see Jenkins, 1963, for details) to the structure of the letter pairs. When any two combinations $a_1b_1$ and $a_2b_2$ are learned, the three-stage paradigms predict transfer to $b_1b_2$. However, such a hypothesis not only fails to account for the high proportion of AB-type intrusions in the earlier studies and for the similarly high relative frequency of MN, PQ, MQ, and PN intrusions in the present study; it also fails to account for the other types of intrusions involving correct vocabulary. Among the 537 intrusions produced from letters appearing in the presented pairs by Ss in the three groups given 15 exposures, 77% were MN, PQ, MQ, or PN intrusions. Only 11% were of the types MM, PP, NN, or QQ. Eight percent were MP, PM, NQ, or QN intrusions, which for reasons similar to those discussed for the four-stage paradigms are excluded by mediation theory. On the basis of the number of potential intrusions, 20% (24/120) of the intrusions would be expected to be MM, PP, NN, and QQ intrusions, if Ss were randomly combining correct vocabulary to produce intrusions. Another 30% (36/120) would be expected in the MP, PM, NQ, and QN classes. Clearly, the intrusion data do not support the mediation-theory account of how Ss produce grammatical intrusions in the recall of structured letter pairs.

**TABLE 2**

**MEAN DIFFERENCE BETWEEN VARIOUS OBTAINED AND EXPECTED NUMBERS OF INTRUSIONS**

<table>
<thead>
<tr>
<th>Difference</th>
<th>A</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obtained vs. Expected MQ + PN Intrusions</td>
<td>0.99***</td>
<td>2.05***</td>
<td>1.61***</td>
<td>1.83***</td>
</tr>
<tr>
<td>MN + PQ Intrusions</td>
<td>0.29</td>
<td>0.91***</td>
<td>0.70***</td>
<td>0.69***</td>
</tr>
<tr>
<td>MQ + PN vs. 3(MN + PQ)</td>
<td>-0.11</td>
<td>0.68</td>
<td>0.47</td>
<td>0.23</td>
</tr>
</tbody>
</table>

**p < .01.**

**p < .001.**
TABLE 3
LETTERS USED IN EXP. II BY CLASSES

<table>
<thead>
<tr>
<th>Letter Set</th>
<th>Letter Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

EXPERIMENT II

The second experiment was undertaken to answer a variety of methodological objections concerning the first experiment and to investigate the generality of the results. The motivations for the various procedural modifications were as follows:

1. The opportunities for mediation in the first experiment are severely limited. Among the presented letter pairs there are no triplets of pairs \((m_1n_1, m_2n_2, m_3n_3)\) such that \(m_3n_3\), the pair to which mediated transfer is predicted, is another presented pair. That is to say, mediation processes cannot assist \(S\) in producing correct pairs, or to use Jenkins’ (1963) phrasing, \(S\) will not be reinforced for mediating. This limitation is due to the poverty of the MN and PQ systems. In Exp. II the membership of the four classes was increased from three to four items with the result that every presented pair could be mediated on the basis of two different triplets of presented pairs. Concretely, if \(M\) consists of \(m_1, m_2, m_3, \) and \(m_4\), and \(N\), of \(n_1, n_2, n_3, \) and \(n_4\), and the presented set includes all MN combinations except \(m_1n_1, m_2n_2, m_3n_3, \) and \(m_4n_4\), then the two triplets \((m_1n_1, m_2n_2, m_3n_3)\) and \((m_4n_4, m_2n_2, m_4n_4)\) each predict mediated transfer to the presented pair \(m_4n_4\).

2. The failure to demonstrate mediated transfer might conceivably reflect characteristics of the particular letter pairs, especially since both presented pairs and intrusions have preexperimental histories of exposure. In Exp. II, four sets of letter pairs were constructed in such a way that any specific combination was used as a presented pair in only one set and would be treated as an MQ or PN intrusion in another set and as a completely ungrammatical pair (e.g., QM type) in the remaining two sets.

3. The first experiment provided only fragmentary information on the changes in recall and intrusion measures with increasing exposure to the material. The second experiment measured retention at four different levels of exposure.

4. To obtain a different, and possibly more sensitive measure of retention, Exp. II included three recognition tests in addition to two recall tests for each \(S\). The advantage of the recognition test is that a controlled set of potential intrusions of selected types may be presented and a response to this set obtained from every \(S\).

Method

Materials.—Four sets of 24 letter pairs were constructed from the 16 letters shown in Table 3. The presented pairs for each set consisted of all combinations of Class M followed by Class N and of Class P followed by Class Q except pairs formed from letters sharing the same ordinal position within a class. For example, in Set 1 the presented MN pairs consist of all combinations of \(v, s, r, \) or \(h\) followed by \(g, k, l, \) or \(w\) except the Pairs vg, sk, rl, and hw.

The composition of the recognition tests can be described in terms of the four classes of letters, as follows: Eight presented pairs—four MN and four PQ; eight potential grammatical intrusions—two each, MN, PQ, MQ, PN; eight other potential intrusions—one each, MM, MP, PM, PP; NN, NQ, QN, QQ. Recognition Tests No. 1 and 2 had no specific pairs in common. Because the design of these two tests exhausted the potential MN and PQ intrusions, it was necessary
to reuse four combinations of this type in Test No. 3. For all other types of pairs used in Test No. 3, an equal number of pairs which had and had not been previously used was selected.

Procedure.—The Ss, 28 enlisted men stationed at Walter Reed Army Medical Center, were run individually. All materials for the experimental session were tape recorded on a single tape, including all instructions, presentation lists, recognition tests, and a 2-min. period of silence during the recall tests.

The 24 presented pairs of each letter set were randomly arranged in 16 different orders in such a way that a pair did not occupy the same ordinal position more than once. Each random order constituted an "exposure" of the set. There were four "trials" consisting of four complete exposures each. Pairs were read onto the tape at the rate of one pair every 3 sec. with a pause of 6–8 sec. between each of the four exposures in the trial. Recognition Tests No. 1 and 2 followed Trials 1 and 3, respectively. Two-minute recall periods followed Trials 2 and 4. The second recall test (Trial 4, 16 exposures) was followed by Recognition Test No. 3. Pairs in the recognition tests were presented in a random order at the rate of one every 10 sec. The S was instructed to say "yes" or "no" on the basis of whether he had heard the pair read in the presentation list. He was informed in advance that "about half" of the pairs would be ones he had actually heard. During recall periods S was required to say back aloud as many as he could remember. He was not informed of the correctness of any pair he produced. No important changes were made in the instructions of Exp. I concerning recall. Some additional procedural information was supplied including the nature of the test to be conducted at the end of each presentation period. The order in which the pairs were presented and recalled was described as unimportant.

The Ss were randomly assigned to one of the four letter sets in blocks of four with one S in every block assigned to each letter set and each block completed before a new one was begun.

Results

Recall.—The two recall tests will be identified by the trial on which they were administered. The two tests yielded identical conclusions. The number of different letter combinations produced was 13.43 (Trial 2) and 16.96 (Trial 4). The Ss produced few repetitions of pairs already mentioned within a single recall test (means of 1.54 and 1.64 repetitions for Trials 2 and 4, respectively) and even fewer pairs incorporating a letter not occurring in the presented pairs (means of .11 and .00). The mean number of presented pairs recalled was 10.00 on Trial 2, 14.18 on Trial 4. The two classes of intrusions which are the focus of interest showed no change from the second to the fourth trial (for both trials, MN-PQ intrusions: \( \bar{X} = .54 \); MQ-PN intrusions: \( \bar{X} = 1.50 \)). The mean difference between the number of intrusions of each class produced and the expected value based on the number of intrusions composed of letters occurring in presented pairs was .41 (Trial 2) and .43 (Trial 4) for MN and PQ intrusions and 1.01 (Trial 2) and 1.09 (Trial 4) for MQ and PN intrusions. These values are significantly greater than zero at or beyond the .02 level. The mean differences between the number of MQ-PN intrusions and four times the MN-PQ intrusions were not significantly different from zero—.64 (Trial 2) and .71 (Trial 4). No significant differences were obtained between the four groups exposed to different letter sets for the following measures: number of different combinations produced, number of presented pairs recalled, difference between obtained and expected number of MN-PQ intrusions, the parallel difference for MQ-PN intrusions, and the difference between MQ-PN intrusions and four times the MN-PQ intrusions.

Recognition tests.—Recognition tests will be identified by number, as follows: No. 1 following Trial 1, No. 2 following Trial 3, and No. 3 following the second recall test (Trial 4). Again
identical conclusions were reached in each test. The mean proportions of “yes” responses given to four classes of intrusions are shown in Table 4. There appears to be a trend of increasing proportions for presented pairs as exposure increases and decreasing proportions for all classes of intrusions. With each test treated separately, there was a significant effect due to type of pair tested and neither a significant main effect due to letter sets nor a significant interaction in all three tests. The component of the main effect attributable to the difference between MN-PQ and MQ-PN intrusions did not reach significance. A Student Newman-Keuls range test indicated that for all three recognition tests the proportions of “yes” responses to presented pairs, grammatical intrusions (MN-PQ and MQ-PN intrusions combined), and other intrusions were significantly different from each other.

Discussion

The results of the second experiment confirm those of the first and suggest that the prominence of MQ and PN intrusions holds over a somewhat wider range of conditions. Mediated transfer as an explanation of grammatical intrusions appears to be ruled out completely for the early stages of learning. However, mediation cannot be excluded as a secondary process operative perhaps at some later stage of learning, but prior to complete mastery of the presented pairs. If this were the case, a decline in the number of MQ and PN intrusions would be expected before a decline in MN and PQ intrusions began. The present evidence does not extend beyond 16 exposures, at which point the mean number of presented pairs recalled was 14.18, roughly 60% of the total. Furthermore, no S achieved complete mastery of the presented pairs.

A pilot study using 24 exposures per pair with frequent recall tests failed to reveal any new trends with greater exposure. In this sample (*N* = 12) only one S produced all 24 pairs following the twenty-fourth exposure. The mean for the other 11 Ss was 11.63 presented pairs. The one atypical S produced five intrusions of the MN-PQ type and one of the MQ-PN type on a test following 12 exposures, at which time he produced only 15 presented pairs. He continued to produce five MN and PQ intrusions (but not always the same five) on every subsequent test (following 16, 20, and 24 exposures); no other MQ or PN intrusions appeared thereafter. His performance may be compared with that of another S in the 24-exposure study who produced 23 presented pairs and no intrusions of any kind on the final recall test. The two best performances in the present experiment appeared on Trial 4 where one S produced 23 presented pairs without intrusions and another S, 22 presented pairs and one intrusion of the MN-PQ type. These individual results suggest that the intrusions predicted by mediated transfer (MN-PQ type) may be the last to be removed from Ss’ recall productions. The practical difficulty in using the verbal reconstructive memory technique to investigate the problem lies in motivating Ss sufficiently to insure complete learning for all Ss without altering some fundamental aspect of the procedure. (For example, the most obvious procedural change, the introduction of specific reinforcement of pair production, has not been investigated to determine whether rule-governed creativity would be affected by this alteration.)
The question of how class formation occurs with classes having identical positional restrictions is an important one because the process occurs in natural language, even if it is not demonstrable in the verbal reconstructive memory situation. Mediated generalization of the kind proposed by Jenkins and Palermo (1964) may find application in explaining word-class formation under such conditions. Braine (1965b) has recently discussed two other mechanisms which would account for the discrimination of two or more classes of morphemes sharing the same position in natural languages. Neither is applicable to the present study, however. In the first, a frequently occurring morpheme may serve as a marker or function word. If each class were associated with a different marker (e.g., a different affix), then the language would provide an overt index to extrapositional regularities. For example, if the structure of the letter pairs had been mN and pQ (where lower-case letters indicate a single entry or “function letter”), the markers would provide a basis for the discrimination of the N and Q classes and combinations such as mq and pn would not be expected to occur, although both N and Q items occupy the second position. However, in the present experiments all letters have an equal frequency of appearance in the presented pairs. Hence, none of them may be expected to form index markers. A second mechanism proposed by Braine involves semantic correlates and obviously cannot apply to semantically empty systems. Semantic support would be available to the MN and PQ pairs if the two types of pairs were associated as names for two different types of stimuli.

Evidence from a quite different experimental technique corroborates the present findings concerning the predictions of the four-stage stimulus and response equivalence paradigms. Braine (1963) taught Ss to complete a set of utterances in a nonsense language. The utterances may be characterized as ab, cb, ad, and cd. The Ss were given training on problems such as a— with response alternatives b and c until all utterances had been mastered. Braine then presented the problem x— and found that Ss chose b and d in preference to a and c. A suitable four-stage mediation paradigm concluding in xb or ad cannot be constructed from the training utterances.

The account of free-recall learning proposed by Garner (1962) is implicitly a position-learning analysis of what Ss learn in a verbal free-recall experiment in that the “variables” used by Garner to describe verbal material are the left-to-right positions occupied by the letters, words, or other units from which the material is composed (cf. Garner & Whitman, 1965). Hence, the production of grammatical intrusions in the recall of pairs having a simple AB structure (as in Smith, 1963) can be explained as the learning of which letters belong to which variables, i.e., positions. However, Garner (1962, p. 168) has also suggested that simple contingencies between variables making up the material should result in a structure which is relatively easy to learn. The MN and PQ letter pairs contained a contingency between classes of letters in the two positions rather than a contingency between individual letters. The evident difficulty of these pairs suggests that a class contingency may be more difficult to learn than a simple contingency between letters. The present data, which do not involve a comparison of the MN and PQ pairs with other types of internal structure, cannot be used to determine whether class contingencies are in fact more difficult to learn.

REFERENCES


Braine, M. D. S. The insufficiency of a finite state model for verbal reconstructive memory. Psychon. Sci., 1965, 2, 291-292. (a)

Braine, M. D. S. On the basis of phrase structure: A reply to Bever, Fodor, and
Weksel. *Psychol. Rev.*, 1965, 72, 483-492. (b)


(Received August 13, 1965)