SPATIAL AND VERBAL COMPONENTS OF THE ACT OF RECALL*

LEE R. BROOKS†
McMaster University

ABSTRACT

Research presented in this paper shows that while a person is recalling a line diagram he can more readily signal information about that diagram by speaking than by spatially monitored output (e.g., pointing to correct items in a column of symbols). When recalling a sentence, he can more readily signal information about that sentence by spatially monitored output than by speaking. These results suggest that spatial and verbal information is recalled and processed in a modality-specific manner. Recall of verbal information is most readily disrupted by concurrent vocal activity; recall of spatial information is most readily disrupted by concurrent spatially monitored activity. This differential conflict occurs even though the concurrent activity is a recoding of the information that is being recalled.

WHEN A PERSON is asked to describe from memory a diagram such as a map or floor plan, he is likely to say that he generated a mental representation of the diagram and then derived his description from that. Even in the absence of vivid mental imagery, there is a clear impression that some underlying visual or spatial process is involved in this type of performance. In contrast, the process involved in recalling a specific sentence seems to have more to do with speech than with vision or spatial movements. If there is a visualized component in sentence recall, it appears to be less crucial than in the recall of spatial relationships.

This paper will present performance data to support the subjectively plausible notion that verbal and spatial information are handled in distinct, modality-specific manners. These data are obtained from experiments which induce conflict between overt responding and the act of recall. Subjects are asked to recall memorized material (sentences or line diagrams) and to simultaneously signal information about that material. If making signals in one modality (for example, speaking) uniquely disrupts recall of one of these types of material, then it will be assumed that the recall of that material is accomplished in a modality-specific manner. If a different modality of response (for example, pointing to a sequence of symbols) provides the strongest conflict when the

*This research was supported by N.R.C. grant A.P.A. 210. I would like to thank Mrs. Wilma Clarke and Miss Rosemary Squire for their assistance in running subjects and tabulating and analyzing the data of these experiments.

†Present address: Department of Psychology, McMaster University, Hamilton, Ontario.

recall material is changed, then this type of material will be assumed to be specific to a different modality.

Experiment I

One of the tasks in this experiment was to categorize each word in a recently presented sentence. For example, a subject listened to the sentence "a bird in the hand is not in the bush," and then successively categorized each word as a noun or a non-noun. In this instance he would produce the sequence, "no, yes, no, no, yes, no, no, no, no, yes." However, the subject was given three different ways of signalling this sequence: (*a*) saying "yes" and "no" as above, (*b*) tapping with the left hand for each noun, and the right hand for each non-noun, and (*c*) pointing to a "y" for each noun and an "n" for each non-noun as in Figure 1 (to produce the sequence given above, the subject would point to the top "n," the second "y," the third and fourth "ns," etc.). If the sentence is recalled

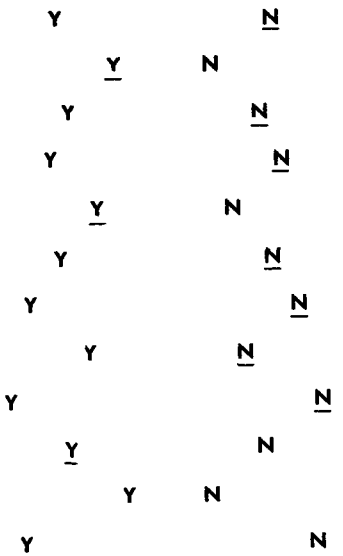

Figure 1. A sample output sheet for the pointing condition of Experiment I. The underlined letters are those which would be pointed to in categorizing the sentence "a bird in the hand is not in the bush." The letters are staggered to force close visual monitoring of pointing.

in a specifically articulatory manner, then concurrently *saying* something different should provide difficulties not present when the same information is signalled by a different type of response. In this case, vocal responses should take longer than either tapping or pointing.

The other task in this experiment was to categorize each corner in a line diagram. For example, a subject looked at a block letter such as that in Figure 2 and then, from memory, categorized each dot as a point on the extreme top or bottom or as a point in between. In this example, starting from the asterisk and proceeding in the direction shown by the arrow, he would produce the sequence "yes, yes, yes, no, no, no, no, no, no, yes." If the "F" is recalled in a specifically visual or spatial manner, then concurrently looking at a different spatial array should lead to difficulties not present when the information is signalled in a different manner. In this case, pointing to the "ys" and "ns" should take longer than either speaking or tapping.

When these two tasks are combined, an experiment with six conditions is obtained. The referent being categorized is either a sentence or a line drawing; for each of the referents there are three ways of communicating the categorizations: vocal, tapping, and pointing. If the expected conflicts are present, speaking should be the slowest form of output when categorizing sentences, and pointing should be the slowest form of output when categorizing the line drawings.

Another design consideration is important to obtain these conflicts. The subject must be uncertain about the exact categorization that he is to make. If he knew that he had to categorize every sentence for nouns, he could simply produce the appropriate sequence of yesses and noes while the sentence was first being presented to him. When asked for his categorization he could then recall the pattern of yesses and noes rather than the actual sentence. Since the anticipated effect of the experiment

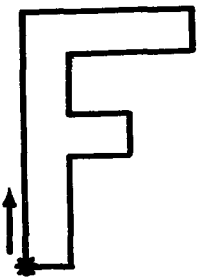

FIGURE 2. A sample of the simple block diagrams used. The asterisk and arrow showed the subject the starting point and direction for both reproduction and categorization.

depends on the attempt to induce conflict by concurrently recalling a different set of words than those being spoken, this strategy would eliminate any conflict. To avoid this situation, the subject was not told until the sentence had been presented whether he was to categorize it for nouns or for grammatical articles (categorization of the example given above for articles would produce "yes, no, no, yes, no, no, no, no, yes, no"). For analogous reasons, subjects were not told until after presentation of each line figure whether they would have to categorize the figure for "outside points" or "top/bottom points."

Method

Half the *Ss* were run on sentences first and half on line diagrams first. The procedure will be described for an *S* who was given sentences first. *S* was told that he would be asked to categorize each word in a series of sentences. Sometimes he would be asked to decide whether each word was a concrete noun and sometimes whether each word was a grammatical article (a, the, an). To illustrate these instructions, a sentence was read, *S* repeated it from memory, and then *E* produced a string of yesses and noes corresponding to the concrete nouns. This yes-no sequence was repeated until *S* clearly understood the relationship between the sequence and the sentence. *S* was then asked to produce a yes-no sequence which corresponded to the grammatical articles in the same sentence. When this had been done correctly, the other two forms of output were explained. To demonstrate comprehension, *S* tapped to categorize nouns in the sentence and pointed to categorize the articles in the same sentence. For pointing, *Ss* were asked to actually touch each "y" and "n," rather than simply gesture in the general vicinity.

A block of three sentences was run for data collection. The procedure for each of the three sentences was the following. The *S* listened to the sentence, repeated it, and then listened to a list of the nouns and articles in the sentence (this list of nouns and articles was given to eliminate difficulty in deciding whether a word was a noun; the main interest in the experiment lies in the conflict between recall and output, not in the actual decision process). *S* was told which form of output would be required and prepared himself either by placing his hands on the table for tapping, or by placing a pointer (a pencil) at the top of the y/n page for pointing. As soon as *S* had repeated the sentence, *E* said either "nouns," or "articles," and started a stopwatch. After completing the output, one of the other two modes of output was named, and the procedure was repeated with the same sentence. Finally, the third form of output was run, again using the same sentence. The order in which the three forms of output were run for a given sentence was counterbalanced, as was the order in which *S* categorized for nouns or for articles. If *S* reported that he "lost" the sentence, the trial was terminated and repeated one trial later.

After all three forms of output had been run for each of the three sentences, *S* was run on line diagrams. The instructions and procedure were analogous to those used for the sentences. With the figures, however, *Ss* were categorizing for top/bottom points, as given above, or for "outside" points (points on the extreme right or the extreme left of the figure). The top-bottom and the outside points were explicitly pointed out to *S* for each new figure. On the initial presentation of each figure, *S* demonstrated his retention by drawing it from memory in the order indicated by the asterisk and arrow. Immediately before each performance, *S* was asked to mentally recall the

TABLE I

MEAN OUTPUT TIME IN SECONDS, BETWEEN-SUBJECTS STANDARD DEVIATION IN PARENTHESES

	Output		
Referent	Pointing	Tapping	Vocal
Sentences	9.8 (2.6)	7.8 (2.1)	13.8 (3.0)
Diagrams	28.2 (12.1)	14.1 (5.4)	11.3 (3.5)

figure for approximately two sec. *E* then said "top-bottom" or "outside," and started the timing. As with the sentences, each *S* performed each of the three modes of output on a practice diagram and on the three diagrams used for data.

The four sentences used were: rivers from the hills bring fresh water to the cities; a bird in the hand is not in the bush; there is the low fiend who stole the child's candy; no man who has a wife is still a bachelor. Each sentence contained ten words and was chosen to provide variety of grammatical form. Each of the four line diagrams had ten points and was a block letter: F (shown in Figure 2), N, G, and Z. Each letter had an asterisk next to it indicating the starting point and an arrow indicating that the points were to be taken in clockwise direction from the starting point. The "ys" and "ns" were distributed in staggered columns down and 8½ × 11 in. piece of paper; the columns were staggered to force closer visual monitoring of the pointing. Ss were eight McMaster undergraduates who served in the experiment to fulfil a course requirement.

Results and Discussion

The average time to complete output for each of the six conditions is shown in Table I; each number is the average of three trials for each of the eight subjects. The data for the sentences indicate that all subjects showed longer average time for vocal output than for the other two forms of output. Inspection of output times for individual sentences showed the same consistency of effect; only one sentence for one subject showed a shorter vocal output time than times for the other two modes of output. This result supports the idea of conflict between vocal output and recalling the sentences. A highly significant difference also existed between pointing and tapping for sentences ($p = .015$), seven of the eight subjects taking longer with pointing output than with tapping output.

The high conflict conditions in these experiments showed much higher variance than the low conflict conditions, which made parametric statistics inappropriate for these data. The randomization test (Seigel, 1956) was used throughout. Since the largest comparison made in any of these experiments is 2 × 2, this test could be applied to an interaction by testing the differences between pairs of conditions. Except for Experiment VII all comparisons were made within subjects; the between-subjects variance is always several times greater than the within-subjects variance.

In the line diagrams, all subjects showed longer average time for

pointing output than for the other two forms of output. Again, there was high consistency within subjects; only one subject showed one trial on which pointing output time was shorter than for the other two modes of output. The difference between tapping and vocal output for the diagrams only bordered on significance ($p = .06$).

Both major predictions, then, were strongly confirmed; vocal output was slowest for categorizing the sentences, and pointing (visually monitored) output was slowest for categorizing the diagrams. The subjects reported that they "could say the sentence to themselves" while tapping or pointing, but not while saying "yes" and "no." The diagrams could be "pictured" while the subjects were tapping or saying "yes" and "no," but not while they were trying to point. These conflicts reportedly had the effect of making it easier to lose track of where one was in the sentences or the diagrams.

Experiment II

In Experiment I, the difficulty of classifying sentences vocally was attributed to the conflict between saying one set of words while recalling another set of words. This assumption would be more plausible if it could be demonstrated that the degree of conflict depends upon the complexity of the verbal output. Within limits, longer or more difficult words in the vocal output should provide more effective interference for recalling the sentence. If the conflict is articulatory, then the effect should be responsive to the articulatory characteristics of the interfering words.

In this experiment, the subjects are given exactly the same means of signalling in all conditions: a word is spoken if it is a positive instance of the category; the word is whispered if it is a negative instance. The actual words spoken or whispered, however, vary. In the "sentence" condition, the actual sentence forms the output words. To categorize for nouns, for example, the subject would recite the sentence, speaking all nouns and whispering all non-nouns. This condition would not be expected to provide much difficulty, since the vocal output would certainly be consistent with the sentence itself. In the la condition, the subject would be asked to whisper or speak the monosyllable "la." To categorize for nouns in this case, the subject would produce a la for every word in the sentence, speaking every la which corresponds to a noun and whispering every la which corresponds to a non-noun. The final output condition, january, signals in the same way, but the longer and more complex word january is used. The three output conditions—sentence, la, and january—all use whispering and speaking as the signalling device; but hopefully these conditions will produce different amounts of conflict by providing dif-

ferent degrees of compatibility between the sentence being categorized and the words in the output.

These three conditions may differ, however, for reasons other than conflict with a verbal referent. For example, it may simply take longer to produce a repetitious chain of polysyllabic words such as january than it would to produce the sentence itself. To control for response-execution speed (under the same conditions of decision as the experimental conditions), the subjects used the same three types of output on written sentences present in front of them throughout performance.

Another possibility to eliminate is that producing the repetitious string of januarys, or to a lesser extent, las might simply interact with the requirement of working from memory, rather than working from a specifically verbal memory. To control for this possibility, the same three output conditions were run using diagrams instead of sentences as the referents. Since each output condition was vocal, no interaction would be expected because of conflicting recall and output. If memory load were the crucial variable, however, the interaction should be present for diagrams as well as sentences.

Method

Part A. This procedure had two phases: in the first phase, Ss worked from memorized sentences; in the second, Ss worked from written sentences present during performance. Each of the ten undergraduate Ss was told the general nature of the task, and was given a demonstration of each output condition (in the order of: sentence, la, and january). The procedure on each trial was run as in Experiment I; all three output conditions were run in counterbalanced order on each sentence. Each S was given three practice sentences and six data sentences in the memory condition. Ss were informed of all errors and were encouraged to minimize them. All experimental sessions were tape-recorded to allow an independent measurement of response times. At the conclusion of these trials, each S was run on six more sentences under identical conditions; however, the sentence remained present throughout each performance.

Part B. A different set of nine undergraduates served in this phase of the experiment. As in Part A, each S worked first from memorized drawings and then from drawings that remained present during performance. Each S was introduced to the diagrams and the general task, as in Experiment I. Ss were then given a demonstration of the la and the january output procedures. Each S was given two practice trials, four memory trials, and four visual trials. All other aspects of the procedure were as previously described.

Results and Discussion

Independent measurement from the tape recordings showed that 95 per cent of the responses agreed ±.3″ with those recorded in the session.

The only data included in the analyses of this and subsequent experiments are the response times from the first instance in which a subject

TABLE II

MEAN OUTPUT TIME IN SECONDS, BETWEEN-SUBJECTS STANDARD DEVIATION IN PARENTHESES

	Output		
Referent	Sentence	la	january
Sentences			
Memory	6.7 (1.8)	10.7 (3.2)	13.6 (2.6)
Visual	4.6 (0.5)	5.3 (1.0)	6.2 (1.0)
Diagrams			
Memory		7.6 (1.1)	8.3 (1.4)
Visual		5.9 (1.1)	6.4 (0.9)

performed with a given sentence or diagram. The other output conditions were immediately run with the same sentence or diagram to give the subject rapid practice on the task. All analyses to be described resulted in the same statistical decisions when these additional trials with the same material were included.

The mean output times, averaged over two items per subject per condition, are shown in Table II. First, let us consider the two sentence and la conditions for the sentences. The pattern of results is as anticipated; requiring the subject to whisper and speak la increases the difficulty of the task much more when performance is from memory than from a written sentence. This interaction between presentation and output was highly significant ($p = .002$) and was reflected in the data of every subject. In addition, all subjects for both types of presentation had longer average response times using la than using the sentence output. These findings would suggest that word compatibility, not simply vocal activity, is important for producing the conflict between recall and output.

The same pattern of results was obtained for the la and january outputs for the sentences. The presentation by output interaction was significant ($p = .03$) and was reflected in the data of all subjects except one. In addition, for both types of presentation, all subjects except one had higher average response times using january than la. The complexity of the output word, then, differentially affects the memory condition.

This pattern did not hold for the diagrams. The interaction between presentation and output did not approach significance. The effect of form of presentation was significant for both types of output ($p = .015$; $p = .015$), and the effect of output was marginal ($p = .064$; $p = .046$). Although the january output tended to take longer than the la output, this difference did not tend to be greater when working from memory than when the diagrams were present.

In general, it took longer to complete la output than sentence output, and longer for january than for la. These differences were much larger, however, when the subject was categorizing a memorized sentence than

when working with any other kind of material. This special status for memorized sentences cannot be accounted for either by the difficulty of categorizing or the necessity of working from memory per se, since there was no comparable interaction with the diagrams. It would also be difficult to explain the greater difficulty of january than la on grounds of formal or semantic similarity between the output word and the words of the sentence.

Experiment III

When line diagrams were being categorized in Experiment I, the most difficult type of output was pointing to the "ys" and "ns." This difficulty was attributed to conflict between simultaneously recalling the diagram and monitoring the pointing. However, the evidence provided by Experiment I does not rule out the possibility that the conflict was due to disruptive effects of movement, quite independently of any monitoring of the movement with respect to spatial locations or patterns.

This experiment is designed to assess the hypothesis that spatially monitored movement can be particularly destructive of spatial recall. To accomplish this assessment, three types of output will be examined for both sentences and line diagrams. As in Experiment II, all three types of output will use basically the same device for signalling; a √ will signal a positive instance of the concept, and an X will signal a negative instance. In no-movement output, the Xs and √s will all be made one on top of the other. In unmonitored-movement output, the subject will be asked to place the Xs and √s one under the other down the page without looking at them. In the monitored-movement output, the subjects will be asked to place the Xs and √s exactly in the appropriate boxes in a column of small boxes printed on a piece of paper. If movement per se is disruptive of diagram recall, then the movement output condition should take longer than the no-movement condition. If spatial monitoring of movements is a major disrupting element, then the monitored-movement condition should take longer than the movement condition.

To show that these effects are unique to spatial referents, sentences will also be run. Assuming that sentences are processed independently, the differences between these output conditions should make no difference that cannot be accounted for by differences in inherent response-execution speed.

Method

Ten undergraduates served in the diagrams phase of the experiment, and a different group of nine undergraduates served in the sentences phase. Both groups performed on two practice items and nine data collection items. Ss in both groups were asked to perform on each of three output conditions for each item. (*a*) No movement:

instructions were given to make √s or Xs roughly on top of one another. The Ss were asked not to look at the page while performing and not to worry about small deviations in location of the marks. (*b*) Unmonitored movement: Ss were asked to place their marks one under the other to form a rough column. Once again, they were told not to look at the page while performing and not to worry about making a neat column. (*c*) Monitored movement: Ss were given an 8½ × 11 in. piece of paper with a set of 15 small boxes (¼ in. side) evenly place in an exact column down the page. Ss were asked to work from top to bottom and to make sure that their marks at least started in the box. If their marks started to stray from the boxes during the course of the experiment, Ss were reminded of the original instructions. Ss placed their pencils on the appropriate starting point before their pre-performance repetition of the sentence or recall of the diagram. All other aspects of procedure were handled as in Experiments I and II.

Results and Discussion

Average response times for each of the six conditions are presented in Table III; each number is the average of three trials for each of 9 subjects for the sentences, and 10 subjects for the diagrams.

There was no over-all difference among the three conditions for the sentences ($p > .10$ for all pairs). In contrast, every subject for every item

TABLE III

MEAN OUTPUT TIME IN SECONDS, BETWEEN-SUBJECTS STANDARD DEVIATION IN PARENTHESES

	Output		
Referent	No movement	Unmonitored movement	Visually monitored movement
Sentences	6.6 (0.9)	6.5 (1.2)	6.8 (1.2)
Diagrams	10.1 (3.3)	11.5 (3.9)	16.4 (5.0)

performed faster in the unmonitored-movement than in the monitored-movement condition when working with the line diagrams ($p = .004$). The difference between the unmonitored-movement and the no-movement conditions for diagrams is significant, but not nearly as consistent ($p = .04$).

The general expectations were confirmed. Varying the movement and guidance of movement during output did not affect verbal recall, but had a strong effect on recalling diagrams. An effect due to movement *per se* exists when recalling diagrams, but it is minor compared to the effect of making movements to precise locations not arranged in the same pattern as the diagram to be recalled.

Experiment IV

In Experiment II, it was shown that a complex verbal response conflicted with verbal recall more than did a less complex verbal response. In Experiment III, it was shown that visual monitoring of a movement interfered with spatial recall more than did movement *per se*. However, in both of these experiments, the effects could be due to the over-all difficulty of the tasks. It is possible that once conflict between recall and output was induced, *any* increase in difficulty of the response would lead to a disproportionate increase in the demands of the task. If this were true, it would not be correct to attribute the difference between la and january to an increase in specifically *verbal* difficulty. Similarly, it would not be correct to attribute the difference between unmonitored movement and visually monitored movement to an increase in specifically *spatial* difficulty. In both cases, it might be that simply one more complication was being added to the lot of a subject who was already coping with a conflictful task.

To meet this interpretation it is necessary to show a much smaller effect due to adding a difficulty that is not specific to the modality in which the conflict is occurring. This is accomplished by combining the two previous experiments and requiring subjects to concurrently signal in *both* a verbal and a spatial mode. For instance, subjects were asked to signal nouns by simultaneously whispering or speaking la and making a column of Xs and √s. In this instance, response difficulty could be added by either requiring visual monitoring or by changing the response word to january. If increased response difficulty is more potent in the vocal mode, then changing to january should produce a larger effect than adding visual monitoring. However, if a diagram is being categorized, then the reverse should occur.

Method

There were four output conditions for both sentences and diagrams; each condition required categorizing by whispering and speaking either la or january and simultaneously by either the unmonitored-movement or the visually monitored-movement condition of the preceding experiment. Each S was run for five days. On the first day each S was given three practice trials with each of the four output conditions for the sentences, and three practice trials for each of the four output conditions for the diagrams. The instructions and material used were analogous to those used in Experiments II and III. Half the Ss worked with diagrams on Days 2 and 3 and then with sentences for Days 4 and 5; the other Ss worked with sentences on Days 2 and 3 and diagrams on Days 4 and 5. On each of the last four days, each of the four output conditions was run in blocks of five trials each; only one output condition was required for each sentence or diagram used. The order in which the output conditions were run was counterbalanced between Ss and days.

Since practice was required before Ss could perform smoothly, a larger stock of

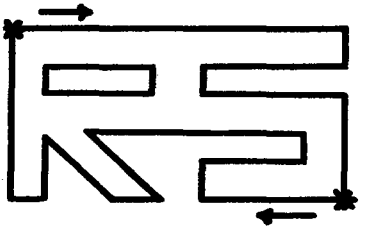

FIGURE 3. A sample of the diagrams used on Days 2 through 5 in Experiment IV. Subjects were told which of the two starting points to use for categorization only after they had successfully reproduced the figure.

easily memorized diagrams was needed than could be supplied by the block letters used in the other experiments in this paper. As a consequence, for this experiment alone, combined letters such as that shown in Figure 3 were used for the last four days. Each of these diagrams had either 16 or 17 points to be categorized; the sentences used in the last four days were increased to the same length. All other aspects of the procedure were the same as in Experiments II and III. Ss were paid undergraduate volunteers.

Results and Discussion

The mean time to complete output is shown in Table IV; each number is the average of five trials for each of the eight subjects. Only the data for the second day of performing with each type of material are presented (Days 3 and 5), although the same trends were apparent in the preceding days. First the data for the diagrams will be considered. If simply adding response difficulty while the person is coping with conflict is sufficient to consistently increase task difficulty, then there should be a difference

TABLE IV

MEAN OUTPUT TIME IN SECONDS, BETWEEN-SUBJECTS STANDARD DEVIATION IN PARENTHESES

		Verbal output	
Referent	Spatial output	la	january
Sentences	Unmonitored movement	15.1 (5.2)	15.6 (4.9)
	Visually monitored movement	15.2 (4.9)	16.2 (4.4)
Diagrams	Unmonitored movement	21.2 (7.3)	22.2 (8.6)
	Visually monitored movement	27.5 (8.2)	26.3 (9.6)

between la and january. In fact, the difference between la and january when combined with unmonitored movement does not approach significance ($p > .10$). When combined with visually monitored movement, where the increased spatial conflict would suggest a larger la-january effect, the difference is even in the wrong direction. In contrast to the lack of a la-january effect, every subject showed greater average response times in the visually monitored output than for unmonitored output (both in combination with la and in combination with january).

For the sentences, the non-specific response difficulty explanation would suggest a significant difference between unmonitored and visually monitored movement, especially when in combination with the high conflict, january condition. The obtained means are in the correct direction, but in neither case do they approach significance ($p > .10$). However, it is also true that the difference between la and january was not as strong as it was in Experiment II. In both cases the difference barely reached significance ($p < .05$). It is possible that with sufficient practice, no difference would be found.

In summary, it does not seem plausible that the results of either Experiment II or Experiment III can be explained by response difficulty that is not specific to the mode in which the conflict occurs.

Experiment V

One might ask whether the conflict between vocal output and sentence recall would be reduced by presenting the sentences in written form. If a person could remember how the sentence looked, he might be able to use a visual or spatial means of recalling the sentence, thereby reducing the conflict with speaking.

In the following experiment, the sentences were presented in either written or spoken form. The categorizing responses were chosen to maximize either the visual conflict or the articulatory conflict. The vocal response was to whisper and speak the word january, as in the previous experiment; the visual-spatial response was to make a written mark in a series of small squares arranged in the same way as the "ys" and "ns" in Figure 1. The subjects were asked to actually mark the squares, rather than merely point, to insure close monitoring of the movements and to thereby possibly produce greater spatial conflict.

Method

Each S was introduced to the sentences and the noun and article categorizations in the same manner as in Experiments I, II, and III. The two types of response, january and checking, were demonstrated by *E*; S was asked to produce each response using the same sentence. For the checking response, S was cautioned to place his check *in* the appropriate box rather than just in the general vicinity.

For the visual presentation trials, S was shown a 5 in. × 8 in. card with the sentence typed on it; the card was shown for the length of time required for E to read it aloud. S repeated the sentence from memory and was shown the card again for the same length of time. He was then informed of the type of response for that trial and, in the case of the checking response, took up a pencil and prepared to begin responding. Immediately before the instruction "noun" or "article" was given, he was shown the sentence again and was requested to read it aloud. The same procedure was followed for the auditory presentation trials; however, no written sentence was shown, and S recited the sentence immediately before the categorization instruction.

Two practice sentences were run, followed by 16 data sentences. The visual and auditory conditions were run for each sentence although, as in all experiments except the first, the data are reported for only the first response condition for each sentence. The four response combinations, january-noun, january-article, checking-noun, checking-article, were run in four scrambled series of four each.

The Ss were 10 McMaster undergraduates, serving to fulfil a course requirement.

Results and Discussion

The mean time to complete output is shown in Table V; each number is the average of four sentences for each of 10 subjects. In the auditory sentences, all 10 subjects produced longer times for the january response than for the checking response. In the visual sentences, eight subjects produced the same result, and another subject tied ($p = .01$). Both presentation conditions, then, produced evidence consistent with the notion of an articulatory recall-output conflict. It is difficult to control for inherent differences in response-execution speeds, as in the last experiment. Strictly speaking, it would be necessary to control for response speed to make the above conclusion, since it might simply take longer to produce the string of januarys than to check the boxes. However, both the reports of the subjects and the occasional errors support the notion of an articulatory conflict.

However, the difference between the two response conditions tended to be smaller after visual than after auditory presentation ($p = .11$). If this result were accepted as reliable, it would suggest that some consequence of having seen the written sentence reduced the articulatory conflict, and possibly even increased the difficulty of the checking response.

TABLE V

MEAN OUTPUT TIME IN SECONDS, BETWEEN-SUBJECTS STANDARD DEVIATION IN PARENTHESES

	Output	
Referent	january	Checking
Auditory sentences	12.5 (2.6)	8.7 (1.9)
Written sentences	11.5 (2.2)	8.8 (1.5)

The present experiment is typical of four experiments performed using this auditory-visual presentation variable. All experiments have produced an interaction in the same direction as the present experiment, but, in spite of a number of minor changes in procedure, all showed the same marginal degree of consistency both within and between subjects. Although the four experiments contain a suggestion of visual processing, the recall of both written and spoken sentences clearly conflicts mainly with the articulatory responses.

A similar result was obtained by Conrad (1964) and Baddeley (1966). Both found that acoustic confusions accounted for a substantial portion of the systematic errors in short-term verbal memory, even though the stimuli were visual. The processing of Conrad's written material must have depended more heavily on acoustic than on visual categories, since formal visual similarity did not produce a detectable effect.

Posner (1967), however, has reported data which show that acoustic effects with written material are not universal. Chase and Posner presented subjects with a 10-second visual array of up to four letters. After a 10-second delay, a comparison letter was shown which was to be classified as having been in the array or not. The subjects showed a strong effect due to visual confusion, but no effect due to acoustic confusion. Their experiment, however, used a smaller amount of material than any of the others cited. The visual recall of three or four letters seems much more feasible than the visual recall of an entire sentence.

Experiment VI

Another question one might ask about the articulatory conflict is whether it is dependent on auditory feedback from the subject's own vocal responses. If the subject were asked only to mouth the vocal output responses, it is possible that the conflict with verbal recall would be reduced.

In this experiment the subjects were asked to categorize sentences with both audible and mouthed yesses and noes. These two response conditions were also run with diagrams to show whether any effect of mouthing was specific to the recall of sentences.

Method

This experiment was run in the same way as Experiment I; however, there were only two response conditions, and only four sentences and four diagrams. On the mouthing trials, *S* was asked to close his eyes and simply mouth yes and no. He was told not to be too elaborate about his mouthing; if *E* could not understand the response, she would tell *S*. *Ss* were 10 McMaster undergraduates serving for course credit.

TABLE VI

MEAN OUTPUT TIME IN SECONDS, BETWEEN-SUBJECTS STANDARD DEVIATION IN PARENTHESES

	Output	
Referent	Overt	Mouthing
Sentences	8.7 (2.0)	9.1 (2.3)
Diagrams	7.1 (2.1)	8.0 (1.8)

Results and Discussion

The mean time to complete output is shown in Table VI; each number is the average of two trials for each of 10 subjects. The difference between overt voicing and mouthing did not approach significance for either the diagrams or the sentences. No tendency toward an interaction occurred. These conclusions are also supported by the error scores and the reports of the subject. As in Experiment I, all measures indicated greater conflict with the sentences than with the diagrams.

Both the size and direction of the effects support the conclusion that the sentence conflict is not critically dependent on auditory feedback.

EXPERIMENT VII

Experiments III and IV demonstrated a conflict between recalling the diagram and making visually monitored movements, but there is no reason to assume that both of these activities are being executed in a specifically "visual" system. It is quite possible that one would suffer comparable impairment from marking in a series of tactually monitored locations. If so, it would be more appropriate to refer to a spatial system. Comparable considerations have led Beritoff (1965) to refer to "image-driven" behaviour, rather than to identify control of spatial location with a specific type of sensory input.

Method

Ten undergraduates served in the diagrams phase of the experiment and a different set of 10 undergraduates served in the sentences phase. Both groups performed on three practice items and six data collection items. *Ss* in both groups were asked to perform on both output conditions for each item. (*a*) Unmonitored movement: This condition is identical to that in Experiment II. (*b*) Tactually monitored movement: an 8½ × 11 in. sheet of cardboard was placed in front of S. A regular vertical column of fourteen holes ¾ × ¾ in. had been cut in the sheet at ⅝-in. intervals. S was asked to proceed down the sheet, placing a √ in the hole if it corresponded to a positive instance of the category, and X if negative. He was told to close his eyes while performing, and to use one hand to locate each hole while the other was marking. All other details of procedure were handled as in Experiment II.

TABLE VII

MEAN OUTPUT TIME IN SECONDS, BETWEEN-SUBJECTS STANDARD DEVIATION IN PARENTHESES

	Output	
Referent	Unmonitored movement	Tactually monitored movement
Sentences	6.6 (0.9)	9.3 (1.6)
Diagrams	14.4 (4.1)	22.2 (7.5)

Results and Discussion

The mean time to complete output is shown in Table VII; each number is the average of three trials for each of 10 subjects. All subjects for both conditions took longer to complete tactually monitored than unmonitored responses. The difference, however, was larger for the diagrams than for the sentences (Mann-Whitney U 20; $p < .05$) (This comparison is made between subjects, which means that two distributions of 10 differences each are involved; since some overlap between the two distributions existed, the randomization test was too unwieldy to use.) Tactual monitoring of output apparently provides particular difficulty when the subject is recalling the diagrams. This suggests that the recall of spatial information conflicts with spatially monitored output, regardless of whether the monitoring is tactual or visual. However, this interpretation must be taken as only tentative since it is still open to the same objection that was rejected in Experiment IV.

GENERAL DISCUSSION

This research has shown that when a person is recalling a sentence he is restricted in the type of articulatory output he can use concurrently to signal information about that sentence. This restriction does not extend to all vocal signals (for example, whispering and speaking the words being recalled), but does extend to at least some verbal signals (for example, concurrently saying la or january). The conflict between incompatible recall and output is not critically dependent on initial auditory presentation or on auditory feedback from the vocal response. Concurrent signalling with spatial movements is relatively unrestricted.

When a person is recalling a diagram, he is restricted in the type of spatially monitored output he can concurrently use. At the level of practice employed in the present experiments, recall of diagrams conflicts mildly with actual spatial movements that are not in the same pattern as the diagram to be recalled. A much stronger conflict is evident when

the movements are either visually or tactually monitored. Concurrent vocal signalling is relatively unaffected. These conclusions are based on research using relatively long sequences of responses with tasks not highly practised and which require sequential, item-by-item responding.

Explanation for Conflict

Three explanations for these conflicts will be considered: similarity, overlapping response mechanisms, modality systems.

Conflict might be a result of semantic or formal similarity between the referent and the output. In this scheme all types of recall and output would be placed on a continuum of similarity; the more similar two verbal or spatial responses were the more conflict they would produce. Verbal recall and verbal output would conflict because they are both verbal and therefore formally similar; verbal recall and spatial output would produce less conflict because they are more distinct. Effects due to formal and semantic similarity can be shown. For example, one might try to categorize the following sentence by saying "three" for each non-noun and "four" for each noun: "now is the time for great prudence and high wisdom." There is usually a hesitation after the categorization of the word "for" which is probably attributable to the similarity between "for" and "four." However, this does not eliminate the question of whether formal and semantic similarity is sufficient to explain the present experiments. The results of Experiment II suggest that it is not. For the similarity interpretation to be sufficient, the word january would have to be taken as more similar to the referent sentences than is the word la.

The second possible explanation is that recall and overt responding require the same equipment for execution. An instance of this would be the conflict generated if recall necessitated activity of the peripheral vocal apparatus. Verbal recall and verbal signalling would conflict because they would simultaneously require access to the vocal apparatus. However, this explanation predicts too absolute a conflict to be plausible; the only feasible way the conflict could be resolved would be alternation of words in recall and words in signalling. Note that the short-cycle time-sharing mechanism often proposed for attention-splitting tasks would not work here; the sharing mechanism allows successful splitting of computing facilities *provided* that there are separate peripheral stations for input or output, the very condition which is being denied by this explanation. This explanation would be ruled out if a person could recall a sentence at the same time that he was continuously saying an irrelevant word out loud. Preliminary results indicate that this is quite possible, as would be expected from the reports of subjects in the present experiments that they did not alternate recall and output.

The third explanation is that the execution of recall and overt responding share some level of *control* mechanisms within a modality that they do not between modalities. By requiring two different tasks within the same modality, one is forcing a split of attention that is not required if the recall and the signalling were done in different modalities. If this were accurate, one would expect the degree of conflict to be controlled by such familiar variables as degree of practice, complexity of response, and familiarity of material. The interesting thing about the present results is that the same information will produce conflict in one modality but not when recorded into another modality.

A problem with this interpretation arises in attempts to define the boundaries of a "modality." The boundaries of such a hypothetical system obviously cannot be identified with any traditional sensory or motor system. Two examples are shown in the present paper: changing voicing (in the whispering and speaking output) does not conflict with recalling sentences; both tactual and visual or kinaesthetic monitoring conflict with the same type of recall. Another instance is that despite common sensory and motor components, speech and musical information are probably handled relatively independently. One can vary these two types of information independently as shown by the ability to sing familiar lyrics to a novel tune (also, see Kimura, 1967).

Implications

The recent stress on attention-splitting tasks and on recoding in short-term memory has de-emphasized the fact that in a wide variety of situations, people do concurrently handle two different forms of the "same" information: visualizing a verbal description; recalling a picture while thinking of partially descriptive verbal phrases; and recalling instructions while performing a motor task. Even when it is not absolutely required by the task, people often represent the same information in two different forms; for example, imagining some features of the score while listening to music. If this is a general characteristic of cognitive processing, then people must routinely have minor co-ordination problems to solve. For this reason, the notion of modality-specific processing could be useful in explaining why some types of internal housekeeping are handled as they are.

The conflict procedures used in this paper could be useful in another way: to detect *changes* in the internal form of information. For example, when listening to a verbal description of a spatial layout (street directions, furniture arrangements) people often report that they visualize the relationships and later recall them in visualized form. This report is undoubtedly oversimplified, since one also often recalls some of the verbal

description concurrently. However, the major problem of demonstrating a change in form of information remains. An attempt to solve this problem using conflict techniques is reported in a previous paper (Brooks, 1967). Reading, a visual task, was shown to conflict with a task for which visualizing was reported, but not with a task for which visualizing was not reported. Thus, conflict or successive interference techniques (Margrain, 1967) can be used to investigate changes in form of information as well as the co-ordination of simultaneous processing.

Résumé

Quand une personne se rappelle un diagramme linéaire, il lui est plus facile de donner une réponse verbale qu'une réponse d'ordre spatial (v.g. désigner les bons item dans une colonne de symbôles). Quand elle se rappelle une phrase, c'est l'inverse qui s'observe. Ces résultats suggèrent que la mémoire emmagasine l'information verbale et l'information spatiale et ne traite pas cette information de la même façon. Le rappel de l'information verbale est plus facilement troublé par des activités vocales concurrentes; le rappel de l'information spatiale est plus facilement troublé par des activités concurrentes d'ordre spatial. Ce conflit différentiel se produit même si l'activité concurrente consiste et un recodage de l'information mémorisée.

REFERENCES

Baddeley, A. D. Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quart. J. exp. Psychol.*, 1966, 18, 362–5.

Beritoff, J. S. *Neural mechanisms of higher vertebrate behavior.* Boston: Little, Brown, 1965.

Brooks, L. R. The suppression of visualization by reading. *Quart. J. exp. Psychol.*, 1967, 19, 289–99.

Conrad, R. Acoustic confusion and immediate memory. *Brit. J. Psychol.*, 1964, 55, 75–84.

Kimura, D. Functional asymmetry of the brain in dichotic listening. *Cortex*, 1967, 3, 163–78.

Margrain, S. A. Short-term memory as a function of input modality. *Quart. J. exp. Psychol.*, 1967, 19, 109–14.

Posner, M. I. Short-term memory systems in human information processing. *Acta psychol.*, 1967, 27, 267–84.

Siegel, S. *Non-parametric statistics.* New York: McGraw-Hill, 1956.

(First received February 29, 1968)