INTERFERENCE BETWEEN A VOCAL AND A MANUAL RESPONSE TO THE SAME STIMULUS

Daniel Holender

Département de Psychobiologie expérimentale
Institut de Neurophysiologie et Psychophysiology du C.N.R.S., Marseille, France

Three experiments were reported in which subjects were instructed to make both a naming and a key-pressing response to a single stimulus which was a letter. Compared with single tasks in which each response was performed alone, it was found that, in the dual-task, there was no slowing down of the manual response but the naming response was considerably delayed. The same pattern of results was also observed in a task requiring to synchronize both responses. Finally, instructions to give full priority to the naming response resulted in a severe impairment of the two response latencies. The results were discussed in the framework of Theios' (1973) model.

I. INTRODUCTION

The background for the present study has to be found in the information processing model proposed by Theios in 1973 to account for a large variety of results obtained with reaction time tasks. One of the purpose of that model was to explain the difference between key-pressing and naming responses to alpha-numerical stimuli in situations involving a one-to-one stimulus-response mapping. When key-pressing responses are required, reaction time (RT) is an increasing function of the number of equally likely stimulus alternatives or, with a constant number of stimulus alternatives, RT is a decreasing function of the increasing presentation probabilities of the stimulus alternatives. When naming responses are required, RT is generally not affected by either the number of stimulus alternatives or by the relative stimulus frequency. Though this last proposition is certainly an overgeneralization, Theios (1973) is not the only one to use the difference between naming and key-pressing responses to alpha-numerical stimuli to infer that different information processing mechanisms should be used in each case (e.g., Keele, 1973; Sanders, 1967).

Theios argued that, since the whole naming process is unaffected by either the number of stimuli or stimulus frequency, whereas key-pressing responses are strongly affected, it implies that the effect of those variables has to be located at the level of response determination and selection but not at all at the level of stimulus identification. Identification of the stimulus is carried out by a parallel search into long term memory. For overlearned responses such as naming of letters or digits, response determination presumably requires only a small transformation of the iden-
tification code, which is also unaffected by the number of stimulus alternatives or by their relative frequencies of occurrence. Newly learned stimulus-response associations such as key-pressing for letters or digits can be retrieved either by a slow parallel long term memory search or by a fast serial scanning of a limited capacity short term buffer into which some of the stimulus-response associations are temporarily stored. Serial self terminating scanning, and the number and position of the stimulus-response associations included into the buffer are responsible for the effects of stimulus set size as well as the effects of the frequency of occurrence of the stimuli.

Theios' further assumption that the long term memory response retrieval and the short term scanning can be carried out in parallel was at the origin of the present work. If this assumption is correct, there is at least the possibility for two simultaneous RT tasks, each involving only one of the two different processing mechanisms, to be carried out simultaneously without interference. The first experiment was designed to test the more restrictive hypothesis that a naming and a key-pressing response to the same stimulus, which was a letter, can be performed concurrently without impairment, that is with no increase in the latency of each response in the dual-task relative to their latency when performed in isolation.

A non rejection of the hypothesis of no interference between the vocal and manual responses would constitute a further argument in favor of the assumption that both kind of responses are mediated by different mechanisms. A rejection of that hypothesis would be compatible with two different interpretations. The first is that both responses share some stages of processing so that they are not mediated by independent mechanisms. The second is that mechanisms are in fact independent but that they compete for a common limited processing capacity. It should be stressed that the fact that naming responses are unaffected by the number or the frequency of the stimuli does not imply no capacity demand at all, but rather that the same amount of capacity is necessary whatever the level of these variables. Since the problem of mechanism independence is orthogonal to the question of capacity independence, one cannot chose between the two interpretations on the basis of the data alone. A reappraisal of the arguments favouring the independent mechanism model would be welcome at this point.

There is no doubt that manual response latencies are influenced by set size and by stimulus frequency. On the other hand, Theios' contention that naming responses are not affected at all by those variables has to be considered with caution. Evidence for the contrary can be found in the literature.

It is not always true that set size exerts no effect at all on naming alpha-numerical stimuli. Letters and digits are to be considered as ordered sets of elements. Following Fitts and Switzer (1962) a familiar subset consists of consecutive elements taken from the beginning of the natural ordered set. Compared with an eight-choice naming condition, Fitts and Switzer (1962) observed faster RT for a two-choice condition using a familiar subset but no advantage when an unfamiliar subset was used. A strong effect of the number of alternative stimuli was also found by Hawkins and Underhill (1971) with 2, 4 and 8 - choice tasks using the familiar subsets: A, B, A, B, C, D, and A, B, ..., H. With unfamiliar subsets, subjects (Ss) behave as if elements from the entire set might be presented so that, except for the slight effect found by Forrin and Morin (1966), naming latencies are generally unaffected by stimulus set size (Brainard, Irby, Fitts & Alluisi, 1962; Morin, Konick, Troxell & McPherson, 1965; Mowbray, 1960; Theios, 1973).

When observed, set size effects on naming responses are generally weak
and cannot be considered as very damaging for Theios' model. The problem of stimulus relative frequency is more serious. This time, the absence of effect has to be considered as the exception and reliable strong effects as the rule. Aside from Theios (1973, 1975), most of the studies have shown relative frequency effects on naming latencies (Fitts, Peterson & Wolpe, 1963; Hawkins & Underhill, 1971; Krinchik, 1974; Miller & Pachella, 1973). A study by Sanders (1970) gave a possible clue to the understanding of those discrepant results. With the letter E more frequent than the letters A, I and O, Sanders (1970) observed a strong frequency effect when Ss were required to respond by saying ES, AS, IS, and OS and no frequency effect with the responses SES, SAS, SIS and SOS. Sanders (1970) speculated that this frequency effect reflected a verbal motor preparatory adjustment. Preparation would mean preshaping of the vocal tract in order to be ready to pronounce a particular response. That this would logically be possible only when the responses begin with different phonemes is actually supported by Sanders' (1970) results. In each of the above mentioned studies showing an effect of stimulus frequency on naming responses, the effect could have been achieved by motor preparation. On the other hand, Theios' (1973, 1975) use of the digits four and five as stimuli prevented the use of motor preparation and no frequency effect was observed.

The assumption that, with naming responses, the relative frequency effect is to be located at the level of response execution rather than at the level of response retrieval is of critical importance for the acceptance of the model proposed by Theios (1973) in the form outlined above. This assumption will be further documented in Experiment 1.

II. EXPERIMENT 1

The main aim of Experiment 1 was to test the hypothesis that naming and key-pressing in response to a letter could be carried out simultaneously without interference, that is without increase in the RT for each response in the dual-task relative to their RT when performed in isolation. The second aim of Experiment I was to know whether stimulus frequency effects on the vocal response would be observed with letters beginning by the same phoneme, so that no motor preparation would possibly take place.

I. Method

Ss were eight students from the Université Libre de Bruxelles. They were given one training and four one-hour experimental sessions each comprising three different four-choice tasks. In the vocal single task, Ss were required to name the presented letter; in the manual single task, Ss had to press the key assigned to the presented letter; in the dual-task, Ss performed both the vocal and the manual responses to the same stimulus. Each task was performed with two different blocks of trials. The first block consisted of a random sequence of 20 stimuli with equal frequencies of occurrence of each of the four letters. The second block consisted of 64 trials with unequal frequencies of the letters. There were four different letter frequency assignments each of which was used during one experimental session. The four frequency assignments were respectively for the letters L, N, R, and S: .50/.125/.125/.25, .25/.125/.125/.50, .125/.25/.50/.125 and .125/.50/.25/.125.
The letters I, N, R and S were chosen as stimuli because, when pronounced in French, they all begin with the same phoneme. This was assumed to be a sufficient condition to avoid different triggering delays of the voice-key and to preclude verbal motor preparation (Sanders, 1970).

The stimuli were displayed on a Nixie indicator located, at eye level, at a distance of approximately 80 cm from the seated S. The response keys for the manual responses consisted of four vertical strips of steel resting on microswitches. Their tips were flush with the distal edge of a weakly inclined board on which S rested his hands and forearms. From left to right, the keys respectively corresponded to the letters I, N, R, and S and were respectively activated by the left middle, left index, right index, and right middle fingers. The verbal response activated a microphone located in the center of the response board, approximately 30 cm from the mouth of the S.

An interval of ten sec separated each successive presentation of a new stimulus as a function of a preprogrammed random sequence. One second before the letter onset, a small neon bulb, located above the Nixie, was flashed for half a second in order to provide a warning signal.

In the single tasks, Ss were required to react as fast as possible while keeping a high accuracy level. Ss were encouraged to do so by rewarding each fast response by a monetary bonus and by deleting the equivalent of two fast responses for each error. For each single task, the mean RT in the training session was used as a cutoff for judging fast responses. Ss were given knowledge of their gain after each block.

In the dual-task, Ss were told that the criterions for fast responses remained the same as in the single tasks. They were instructed to try to go as fast as for each response in the dual-task as for the corresponding responses in the single tasks. Ss were given no trial by trial information about their performance. After each block, Ss were only informed about their gain but not about the way it was attained.

2. Results and Discussion

Figure 1 shows the mean RT, averaged over Ss and sessions, for the vocal and the manual responses in the single and in the dual-task as a function of relative stimulus frequency. The data were submitted to various analysis of variance and to subsequent post-hoc comparisons when necessary. Since the results generally speak for themselves, no detailed report of the analysis will be given. A level of .05 was required for statistical significance.

In the single tasks, the vocal RT is always faster than the manual one. There is a slight effect of stimulus relative frequency for the vocal response and a strong effect for the manual one leading to the usual interaction between relative frequency and response mode. Subsequent post-hoc Duncan comparisons carried out on the naming latencies showed the .50 response to be significantly faster than the two other ones as well as faster than the control condition. No other difference was significant.

In the dual-task, the vocal RT is slower than the manual one by an amount of time that is constant over the different stimulus frequencies. As a consequence, there is no more interaction between relative frequency and response mode. As can be seen in Figure 1, both the absolute level of performance and the relative frequency effect are exactly the same for the manual response either performed alone or concurrently with the vocal one. The dependence of the naming latency upon the key-pressing latency not only holds at the level of the mean RT but also at the level of the whole RT.
distribution. Within each block of trials, the Bravet-Pearson correlation coefficient between the manual and the vocal latencies was very high, ranging from .74 to .88 with a mean of .83.

Like in Theios (1973, 1975), the possibility to use a strategy of verbal motor preparation was probably preclude in the present experiment since each response began with the same phoneme. Nevertheless, contrary to Theios (1973, 1975), the naming response was still slightly affected by stimulus relative frequency. It is important to note that this effect differs from the one observed with key-pressing response in at least two ways: i. only the most frequent response is concerned, ii. even in that case, the effect is much more weaker than with manual responses. In my opinion, a reasonable conclusion is that the present results cannot be considered as strong arguments against the assumption that naming and key-pressing responses to letters are mediated by different processing mechanisms.

The assumption that a naming and a key-pressing response to a letter could be carried out concurrently without interference is not supported by the results of Experiment 1. There seems to be some stage or stages of analysis, necessary subsequent to stimulus identification which is common for both responses, that cannot be carried out in parallel. In view of that limitation, it is surprising that both responses were not impaired. Instead, the dual-task exerts a detrimental effect on the vocal response only, without any decrement in the manual performance. Whatever the reasons for which Ss are so consistently giving priority to the manual response, it is clear from the results that the vocal response cannot be released before a minimum delay after manual response initiation has elapsed. It is not possible to know which stage (stages) of the key-pressing response is (are) responsible for the delay in the naming response. All the operations after stimulus identification are possible candidates. On the other hand, one can spe-
culate that the only operation that is prevented to occur is vocal response initiation. This assumption is supported by the fact that response determination and selection are assumed to be achieved by a small transformation of the identification code. Moreover, the whole naming process is normally carried out faster than the key-pressing one, as it is clearly the case in the single tasks.

There is another way to interpret the results of Experiment 1. Namely, the results are compatible with the hypothesis that Ss make a grouped response to the stimulus. Since the vocal response is normally faster than the manual one, it means that Ss have to delay the vocal response in order to synchronize it with the manual one. The fact that the responses are actually not simultaneous does not refute that hypothesis because we do not know what Ss are eventually trying to synchronize. Since both responses result from partially different processes and are executed by different effectors, it could well be that a successful covert synchronization of two events still leads to asynchronous overt responses.

Since both interpretations of the delay in the naming responses predicts a high correlation between the RT for the two responses, which is indeed the case in Experiment 1, it is not possible to choose between them on the basis of the observed results.

III. EXPERIMENT 2

The first purpose of Experiment 2 was to know whether it is possible to choose between the two interpretations proposed above by asking Ss either to realize a double performance or a synchronization of the responses in the dual-task. The second aim of this experiment was to study what happens if Ss are required to give priority to the vocal response instead of the spontaneous manual priority they showed in Experiment 1.

1. Method

The main characteristics of the procedure were the same as in Experiment 1. The frequency bias was suppressed and Ss were only tested with four-choice tasks with equally likely alternatives. A basic experimental unit consisted of three blocks of 40 trials, one of which was used for the single manual task, the other for the single verbal task, and the last one for the dual-task. This unit was repeated twice in each of two sessions. The first unit was considered as practice and discarded from the analysis except that the mean RT in the single tasks provided cutoffs to judge for fast responses during the rest of the experiment.

Three groups of six Ss each were tested. Groups differed only by the way Ss were instructed to perform the dual-task. Here are the instructions for each group:

Group 1: Ss were given the same dual RT task and the same instructions as in Experiment 1. They were required to go as fast for each response in the dual-task as for the corresponding responses in the single tasks. They were rewarded accordingly.

Group 2: Ss were required to synchronize both responses. They were told that the main aim of the task was to keep the interval between the two responses at a minimum. A bonus was given each time the two responses were separated by less than 40 msec. Ss were not informed about that criterion. Ss were also asked not to respond to slowly but this was a secondary goal
for which they were not rewarded. Ss were given no trial by trial information about their performance, they were only told their total gain after the block was completed.

Group 3: Ss were instructed to give full priority to the vocal response while doing their best for the manual one. They were rewarded only for naming responses faster than the mean RT in the single naming task. Fast key-pressing responses were not rewarded at all.

2. Results and Discussion

The mean RT for correct responses are shown in Figure 2. For the single tasks, which were the same in each condition, the three groups display very similar performances. This is true both at the absolute level and for the average 80 to 90 msec advantage of the naming response over the key-pressing one.

As can be seen in the left panel of Figure 2, the results of the dual RT task are very similar to the one observed in Experiment 1. The manual mean RT is not affected at all by the concurrent vocal response whereas the vocal response is significantly slower than the manual one. Though the delay is smaller than with the control trials in Experiment 1 (45 vs. 83 msec), it is clear that the naming response has to wait for the completion of the key-pressing one in order to be elicited. This is confirmed by the fact that, again, there is a high correlation between the RT for the two responses (r = .79 on the average across Ss).

Except for a slight increase in the manual RT, which is not astonishing since speed was not emphasized in that condition, Ss in the synchro-

![Figure 2. Mean RT for the vocal and the manual responses as a function of the type of task (S: single task, D: dual-task) in each condition of Experiment 2](image-url)
nization task behave much like the ones in the dual RT task. As can be seen in the central panel of Figure 2, the failure to synchronize the responses give place to a delayed vocal response very similar to the one observed in group 1, except that the delay was somewhat bigger in group 2 (69 msec) than in group 1 (45 msec). The correlation between the two RT is still higher, but not significantly so, than in group 1. The average of the individual correlation coefficients is .90.

Though Ss in group 3 gave their vocal response before the manual one, they were unable to fulfill the requirement to go as fast as in the single vocal task. The right panel of Figure 2 indicates that, when Ss are urged to behave in a way different from the one they spontaneously adopt, performances of both responses, instead of naming performance only, are severely impaired. There was only one S who was able to perform the task almost perfectly. His vocal response was only 20 msec slower in the dual-task relative to the single task, but his manual response was now 320 msec slower than when performed alone. For that S, it seems that the verbal response processing prevented all manual information processing beyond stimulus identification to take place. Although, at a superficial level, the results of that S look like a mirror image of the results observed in the other groups, there is an important difference: that is, RT for both responses are not correlated \( r = -.14 \). On the average, the correlation between the RT is much smaller in this condition \( r = .38 \) than in the other ones.

The most puzzling observation in Experiment 2 is that, aside from some differences too small to be taken into account, Ss seems to behave in a very similar manner under two different instructions. Namely, the results of the dual RT task are almost the same as the results of the synchronization task. One possible explanation lies in the fact that Ss in the synchronization task were also required to go fast. Even if that requirement was clearly defined as secondary, it could have contributed to reduce the difference between the results of group 1 and group 2. One of the motivation for the third experiment was to test this latter hypothesis.

IV. EXPERIMENT 3

The main purpose of Experiment 3 was to study the synchronization performance in a situation that was not possibly contaminated by RT task requirements. In order to do that, Ss were never submitted to an RT task and speed of responses was never stressed. Ss were only required to synchronize the naming and the key-pressing responses any time they want after stimulus presentation. The second aim of Experiment 2 was to study whether the particular letters used so far have any influence on the observed effects. In addition to L, N, R and S, Ss were also tested with B, D, P and T.

1. Method

Eight Ss, different from the one used in the other experiments, were tested during two 45-minutes sessions. Compared with the synchronization condition of Experiment 2, two procedural differences were introduced: single RT tasks were performed and blocked rather than random presentations of each letter were used. In addition of the letters L, N, R and S that were used during one of the two sessions, the letters B, D, P and T were used during the other session. From left to right, the manual response as-
signments for B, D, P and T was the same as for L, N, R and S. Instructions and payoff were the same as in the synchronization condition of Experiment 2 except that Ss were allowed to make their responses any time they wanted after stimulus onset.

2. Results and Discussion

Table 1 shows the differences between the latencies of the vocal and the manual response for each letter. For the set composed of the letters L, N, R and S, the usual 70 msec delay is observed. This delay is significantly increased with the set including the letters B, D, P and T being 102 msec on the average. The faster S made his first response after a delay of approximately 300 msec whereas the slowest S waited for more than one sec. But the delay in the vocal response was independent of the synchronization latency.

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>N</th>
<th>R</th>
<th>S</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>56</td>
<td>76</td>
<td>83</td>
<td>66</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>D</th>
<th>P</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>91</td>
<td>112</td>
<td>94</td>
<td>109</td>
</tr>
</tbody>
</table>

Table 1. Mean difference (in msec): vocal response latency minus manual response latency. Letters in the same column were responded with the same finger.

The main result of this experiment is that, like Ss in Experiment 2, Ss experiencing no time pressure at all make their naming response 70 msec later than the manual one. The assumption that speed requirements might have played a role in the naming response delay observed in the synchronization condition of Experiment 2 is therefore clearly ruled out.

It is interesting to note that the delay in the naming response is not independent of the phonological aspects of the response. This might give some clues about the events Ss try to synchronize. This follows from the fact that the time at which the energy necessary to activate the voice-key become available is not independent of the utterance. For responses starting with a vowel sound, this energy is present from the very beginning of the utterance. On the contrary, for responses starting by a stop consonant, that is by a closure of the vocal tract, there is probably no useful energy before the voicing that follows the explosion. The observation of a greater delay for responses starting by a stop consonant than by a vowel sound (see Table 1) can therefore be used to infer that it is the beginning of the vocal response that Ss try to synchronize with some unknown part of the manual response. Would the responses have been pronounced by native speakers of English, the same argument could have been used to predict a greater delay for unvoiced than for voiced stop consonants. But no such prediction would hold for French speakers because the Voice onset Time is probably close to 0 msec rather than to 30 msec like in English. As a matter of fact, no difference between the voiced and unvoiced responses was observed. Finally, I have no ready explanation for the observation that, for responses starting with a stop consonant, the delay depends on the place of articulation, being greater with bilabial than with dental responses.
As was pointed out in the introduction, Thelos' (1973) contention that naming and key-pressing responses to alpha-numerical stimuli are fully different in nature is perhaps a little exaggerated. The reason is that naming responses are sometimes influenced by stimulus set size as well as stimulus relative frequency. Though the importance of the former effect has been dismissed, the fact that naming responses are sometimes affected by stimulus relative frequency to the same extent as key-pressing responses casts some doubts about the assumption of independent processing mechanisms. But, only studies allowing for vocal motor preparation display naming relative frequency effects (Pitts, Peterson, & Wolpe, 1963, Hawkins & Underhill, 1971, Krinchik, 1974, Miller & Pachalla, 1973). When motor preparation is precluded, either no effect is observed (Thelos, 1973, 1975), or, as in the present Experiment 1, a weak effect, qualitatively different from the one affecting key-pressing responses, is obtained. These results are consistent with the hypothesis that, for naming, relative frequency effects have to be located at the stages of response initiation and execution rather than at the stages of response determination and selection. To conclude, I am inclined to think that there is no strong evidence against the assumption that naming and key-pressing responses are processed by different mechanisms.

With regard to the main purpose of the reported studies, the hypothesis that naming and key-pressing responses to letters can be carried out in parallel is not supported by the results. Dual RT tasks show at least their vocal performance impaired. By accepting the assumption of independent mechanisms for the verbal and manual response, one is left with only one interpretation of the results. That is, at least part of the stages involved in each processing competes for a common limited processing capacity pool. But, this interpretation relies upon the assumption that Ss are actually emphasizing speed for the two responses as they are required to do. In this regard, the fact that Ss behave in a similar manner in the dual RT task and in the synchronization task is troublesome. This, together with the very high correlation between the two RT, is consistent with the idea that Ss are grouping the two responses while performing the dual RT task. Since the vocal latency is shorter than the manual latency when the responses are performed alone, deliberate grouping of the responses should necessarily entail an increase in the vocal RT, but no increase in the manual one. If this is indeed the case, nothing could be said about an hypothetical sharing of a limited processing capacity between the two responses unless one demonstrates that the grouping strategy itself is to be considered as a consequence of capacity limitation. Of great consideration for that matter are the results of the vocal priority condition of Experiment 2. Ss were unable to do the task as required. Ss gave their vocal response prior to the manual response but at the cost of severe decrements in the two latencies relative to the latencies in the single tasks. This demonstrates two things. The first is that most part of the manual response processing cannot be carried out concurrently with the naming processing. The second is that half the dual RT task instructions are impossible to fulfill. In view of these limitations, the grouping strategy appears to be the best possible compromise that allows for a partial fulfillment of the dual RT task instructions while keeping the global performance impairment at a minimum. As a matter of fact, Ss were able to execute the manual response without performance decrement relative to the manual single task.

In summary: the results reported here are consistent with the following assumptions: 1. Naming and key-pressing responses to letters are me-
INTERFERENCE BETWEEN A VOCAL AND A MANUAL RESPONSE

Diated by independent processors, ii. when used together, these processors compete for a common processing capacity pool, and iii. the grouping of the responses is a consequence of processing capacity limitation.

This work has been partially carried out at the Université Libre de Bruxelles while the author was supported as "Aspirant" by the Belgian "Fonds National de la Recherche Scientifique". The author's address: Laboratoire de Psychologie Expérimentale, 117, av. Adolphe Busl, 1050 Bruxelles, Belgium.

REFERENCES


