Can simultaneous speech stimuli be classified in parallel?*

ANNE M. TREISMAN and STEPHEN FEARNLEY

Department of Experimental Psychology, South Parks Road, Oxford, England

Reaction times to detect a known or unknown digit in paired or single auditory test stimuli were measured. The results suggest that in classification or matching tasks with stimuli belonging to separate verbal classes, parallel or selective processing may be possible. There was no interaction of type of task (classify vs match) with either dichotic vs mixed monaural presentation, or pairs vs single stimuli, or negative vs positive responses. An attempt was made to suggest the separate processing stages underlying performance in this task.

There is now considerable evidence that Ss cannot identify two simultaneous strings of words at once when the presentation rates are fast. Moray and O'Brien (1967) showed that Ss detect considerably fewer target letters embedded in each of two strings of digits when their attention is divided than they do in one string when their attention is focused on it. In particular, they seldom detect both of a pair of targets when these occur simultaneously. Treisman and Riley (1969) controlled for any possibility of response competition or of memory decay in a similar monitoring task and confirmed that perception of the verbal content of a second auditory message is severely reduced while another is being identified. Finally, Sternberg (unpublished) has obtained results suggesting that when Ss are asked to detect the presence or absence of any of a variable number of target digits in a simultaneous pair of auditory test digits, the comparisons of each test item to the set of targets are carried out serially rather than in parallel. The slope of reaction time against target set was steeper for negatives and (in one of three Ss) for left ear positives compared to right ear positives.

The suggestion that simultaneous speech stimuli must be analyzed serially was first put forward by Broadbent (1954, 1958) to explain his finding that Ss presented with three pairs of dichotic digits prefer to recall them grouped ear by ear rather than alternating between the ears to approximate their temporal order. With the added assumption that it takes time to "shift attention" between simultaneous items to identify them serially, this explains why at fast presentation rates it is more efficient to identify first the three

from one ear, then the three from the other, than to alternate between the ears to identify serially each member of the simultaneous pairs.

A difficulty for this account was the finding by Gray and Wedderburn (1960) and by Yntema and Trask (1963) that when verbal cues conflict with the ear grouping they may be dominant in determining the recall order. For example, if the successive words of a phrase like "Mice eat cheese" or three color words, red, green, and blue, are interleaved with three digits on alternate ears, recall order often follows the verbal cues rather than the ear grouping. However, verbal cues do not make alternation *within* a pair easier than successive recall of two triplets (Bryden, 1964). The problem is to reconcile Ss' ability to select the correct member of a pair (where these differ only in verbal content) with the idea that parallel perception of the verbal content of simultaneous stimuli is impossible. One explanation for the apparent selection of verbal classes (Broadbent & Gregory, 1964) is that this is *not* in fact possible at the perceptual level; that mixed verbal classes make the ear grouping more difficult (because it takes time to shift between conceptual categories) rather than positively facilitating recall over the level obtained with the ear grouping and homogeneous verbal items.

An alternative explanation, which retains but qualifies the idea that parallel analysis of simultaneous stimuli is impossible, is that Ss can set themselves to identify not only a particular word but a particular subset of words, such as the class of digits. They then test each successive stimulus pair for the presence of members of this target class, without, in the process, identifying any nonmember. In conditions where only one member of the stimulus pair belongs to the relevant class of words, this would result in identification of the relevant item without the need for serial identification of both stimuli. This

selective testing might be possible only with well-learned verbal classes, and not with arbitrary subsets of items within a verbal class (like Sternberg's subset of target digits). In an experiment on visual search, Brand (1971) showed that Ss could be as quick to locate an unknown digit in a background of letters as to locate a particular specified digit, but were much slower to locate a specific digit in a background of irrelevant digits. Equivalent results were obtained in searching for letters in backgrounds of digits or letters. Clearly, there is no simple physical characteristic differentiating all digits from all letters, so Brand suggests that tests for all 10 digits can be carried out more rapidly than tests to distinguish any single digit from the other digits. There are two points of interest in these results: one is that search for a specified digit is faster in a background of letters than in a background of digits, and the other is that search for any digit in a background of letters can be as fast as search for a specific digit. The former finding suggests that classification of background items is faster than discrimination within a class, while the latter finding suggests either that Ss use the same classification strategy of rejecting any nondigit, even when told the identity of the target digit, *or* that tests for all 10 digits are made in parallel at the same rate as tests for a specific digit.

SERIAL VS PARALLEL ANALYSIS IN DIFFERENT PERCEPTUAL TASKS

Suppose there are three main types of analysis, one resulting in classification ("nonsense syllable" or "digit"), one resulting in identification ("Taz" or "8"), and one resulting in matching to a previously presented template ("same" or "different"). These could logically occur either in series, in parallel, or selectively when two test stimuli are presented simultaneously. (Selective analysis means restriction of the perceptual tests carried out on the input to those features which will identify a particular item or a particular subset of items.) Figure 1 shows the combinations of these possibilities. In tasks requiring identification of *both* stimuli, the evidence suggests that serial analysis is necessary. On the other hand, Yntema and Trask's results suggest that with different verbal classes, selective identification, or parallel classification followed serially by identification of the items from the relevant class, may be possible. Sternberg's results suggest that matching must be serial at least when members within the same class must be discriminated. Possibly when different classes are used, it could be preceded by a parallel classification stage, after which

*This research was supported by a grant from the Medical Research Council. We are grateful to S. Sternberg for his comments on an earlier draft of the paper and to Bell Telephone Laboratories for the use of their facilities in generating the stimulus tapes.

Fig. 1. Possible combinations of perceptual task and mode of processing.

items from an irrelevant class could be discarded from further matching operations. For example, if the negative set consisted of nonsense syllables as well as or instead of irrelevant digits, one might get different results from the apparently serial process Sternberg obtained.

The present experiment explores the classification and two types of matching-to-sample tasks within the same paradigm that Sternberg used. In one matching task a physical match was possible and in the other a name match was necessary. The task is conceptually similar

to Brand's visual search task except that (1) the display of many irrelevant background items is replaced by a single irrelevant item present on half the trials; (2) the focus of interest is on serial vs parallel analysis of the *stimulus* items (those listened *to*) rather than the *target* items (those listened *for*). Brand's use of rate of search as her measure assumes serial analysis of stimulus items, which [as results by Sternberg & Scarborough (1969) show] may not be justified. In the present experiment, RT to detect a digit in a pair of simultaneous test words is compared to

the same task with a single test word, where the nondigits are members of a totally different verbal class, i.e., nonsense syllables. In the classification condition, the target is any digit, and in the matching conditions it is a particular digit specified in advance. If the two items in a simultaneous pair are always analyzed serially, one would predict that the change in RT from classification to matching should be doubled for a pair of negative test stimuli compared to a single negative test stimulus (see Fig. 2). For positive pairs it would also be doubled if analysis were

Fig. 2. Relative processing times under different assumptions. * There is no reason to assume equal ease of selection in the classification and matching tasks. The increase for single items *could* be longer than that for pairs.

exhaustive, and it would increase by 50% if analysis were self-terminating. If simultaneous stimuli can be tested in parallel, either for classification or for matching, the effect on RT should not be doubled, although it might still differ for the pair and for the single stimuli. The same would apply to selective analysis and to any combination of the three tasks and methods (e.g., serial classification with parallel matching). In no case other than that in which both tasks are serial would one expect the difference to be doubled for pairs of test stimuli. It could even, in certain cases, be less for pairs than for singles, e.g., where the longer analysis of single items can occur in parallel for pairs, while the shorter analysis of single items must be serial with pairs.

SINGLE VS SEPARATE CHANNEL PRESENTATION OF TEST STIMULI

There is another distinction which can be investigated more easily with auditory than with visual stimuli and which is of some theoretical interest. If there is a limit to perceptual analysis forcing serial rather than parallel processing, there are two forms this limit might take: (1) It might require different sensory inputs to be handled sequentially (for example, the left-ear item followed by the right-ear item); (2) it might require different decisions or outputs of analysis to be reached sequentially (for example, the number "two" followed by the nonsense syllable "TAZ"). This distinction between separate sensory inputs to the analyzing systems and separate outputs of analysis is discussed by Treisman (1969). While the second limit would affect both dichotic and mixed binaural or monaural inputs, the first could arise only with dichotic inputs (or with inputs which differed on some other physical dimension). The auditory system allows the possibility of receiving a single "sensory input" (two words matched in all other physical characteristics and presented monaurally or binaurally) which can give rise to two different outputs from a single analyzing system, where the two outputs of speech recognition are the same as would be given by two separate sensory inputs (the same words presented dichotically or in different voices or intensities, etc.).

Mixed monaural or binaural presentation, as well as suffering from increased peripheral masking, would clearly be disadvantageous where one of the items is known in advance to be irrelevant. However, it might be an advantage where either item could be relevant, since it might then allow Ss to test the composite input for the presence of a specified item instead of testing each dichotic input in turn.

Alternatively, it might also be possible under some conditions to ignore their separate localization and to treat dichotic inputs as a single composite stimulus, identical to a mixed binaural or monaural input except that peripheral masking is reduced or eliminated. One may be able to choose whether to handle dichotic inputs separately or as a single compound input, while with a mixed binaural or monaural presentation, Ss are forced to attend to the composite of both words in order to identify either.

This point might be important in explaining the results quoted earlier on simultaneous monitoring tasks for letters embedded in digits (Moray & O'Brien, 1967; Treisman & Riley, 1969). These would seem to refute the idea that even *classification* tasks can be carried out in parallel for simultaneous auditory inputs. However, in both cases Ss were required not only to monitor both inputs but to distinguish in their responses *which* ear received which items. Thus, it could have been the necessity to label the two inputs with their ears of arrival which made parallel testing impossible, in that it prevented Ss from handling dichotic stimuli as a single composite input.

To summarize, the main question raised here is whether analysis of two simultaneous speech stimuli must always be serial, or whether they can under some conditions be analyzed either selectively or in parallel. Two factors that may be relevant in determining which mode of processing occurs are: (1) the nature of the perceptual analysis required (classification, name or physical matching vs identification or discrimination within a verbal class) and (2) single- vs separate-channel presentation of the test stimuli.

METHOD

Stimuli

Four types of test stimuli were used: pairs of two different nonsense syllables (NN), digit/nonsense-syllable pairs (DN), single nonsense syllables (N), and single digits (D).

The digits and nonsense syllables used in the experiment were all recorded from a digital master tape (made jointly with Sternberg at Bell Telephone Laboratories). Each item was spoken by the same male speaker, who attempted to keep their durations as constant as possible. The nonsense syllables used were the following: DEK, DIF, DUR, FET, GIM, HAF, JAX, JEN, JOK, KEP, LAN, LIB, MUL, NEB, NOP, POS, PUD, SEM, SIV, TAV, VIK, WUD, WUZ, ZAP. Together with the digits 1 to 10, they were recorded at a sampling rate of 10 kc, filtered at 4 kc, and approximately equated in loudness. Their

durations were measured from an oscillograph (they ranged from 225 to 350 msec) and their onsets were determined as accurately as possible. A stimulus-ordering program (Sternberg & Pruzansky, in preparation) was then used to generate the test tapes. This program allowed subsequent accurate synchronization of the onset of paired stimuli on two tracks of an analogue tape recording by recording alternate digital samples of the two items to be synchronized. Pairs whose durations differed by more than 70 msec were avoided, and, as a result, the mean difference in duration of the nonsense-syllable pairs was 22 msec and of the digit/nonsense-syllable pairs, 27 msec. Each test item or test pair was preceded by a 250-msec binaural warning tone followed by a 1-sec interval. Six different tapes were made with different randomizations of these test stimuli, each consisting of 80 stimuli (20 of each type in random order). Two of these tapes differed from the other four in that each test stimulus was preceded by a binaural recording of a digit. If the test stimulus contained or was a digit, this preceding digit was identical to it; otherwise it was chosen randomly. These two tapes were made without the computer by splicing tape after the author had left Bell Laboratories, so the interval between the target digit and the test item or pair was not exactly specified; the onset-to-onset time was, however, subsequently measured with a voice key and timer, and found to average 2,670 msec, with a range of 2,450 to 2,840 msec and a standard deviation of 115 msec. Equal numbers of single stimuli and of digits in DN pairs were recorded on each track of the tape recorder, and the order of these was also randomized.

Subjects

The six Ss were volunteers, paid 5/- an hour. There were three men and three women, all students either at university or at the College of Further Education. All were clearly right-handed, and, on an auditory threshold measure with a 1,000-Hz tone, their ears were matched within 2 dB or less.

Procedure

The task was to press one key if the test stimulus was or contained a digit and the other key if it did not. Speed was encouraged, but not at the expense of accuracy. Each key was pressed with a separate hand, and the key used for positives was allotted to the right hand for three Ss and to the left for the other three.

Three conditions were tested: (1) *Visual cue (name matching)*. At least 5 sec before

Table 1
Mean RTs to Detect Presence or Absence of a Digit. (Standard deviations over Ss in brackets.)

	Pair		Single	
	Negative	Positive	Negative	Positive
Dichotic				
No Cue	732 (85)	613 (39)	607 (54)	532 (46)
Visual Cue	614 (90)	511 (69)	503 (34)	420 (33)
Auditory Cue	630 (72)	504 (60)	527 (59)	441 (20)
Mixed				
No Cue	883 (210)	760 (109)	696 (103)	564 (69)
Visual Cue	763 (166)	641 (109)	572 (65)	434 (40)
Auditory Cue	747 (133)	612 (116)	582 (85)	479 (81)

each trial, Ss were shown a card on which the positive digit, if any, was typed. If the test stimulus was negative (NN or N) a random digit was shown. (2) *Auditory cue (physical matching)*. In this condition the tapes with the target digits were used, so Ss heard the target digit about 2,670 msec before the test item or pair. (3) *No cue (classification)*. In this condition, Ss were told to respond positively to any digit and no information was given as to which particular digit would occur. Each S heard two tapes in each condition, giving 40 RTs for each type of test stimulus in each condition (20 to the right ear and 20 to the left for the single stimuli and the digits in DN pairs). The pairs and single test stimuli were presented in random order so that Ss could not know in advance which they would hear. This whole set of conditions was run through once with the paired stimuli presented dichotically and once with them presented mixed together monaurally. The single stimuli were always presented monaurally. In the mixed condition, the warning tone was also given monaurally on the ear on which the test stimulus would occur. In the dichotic condition the warning tone was binaural.

Ss each had a practice session of 40 RTs with dichotic and 40 with mixed presentation in each of the three cueing conditions. They then had four test sessions lasting about 1½ h each, two dichotic and two mixed, each including each type of cueing condition, with the order of presentation type and the order of cueing conditions balanced across Ss.

In a control condition a further four Ss went through two of the tapes in the mixed no-cue condition, with instructions that accuracy was all-important; they were to press the key for "digit" or "no digit" only when they were satisfied that their judgment was as accurate as they could make it.

RESULTS

Table 1 and Fig. 3 show the mean RTs in each condition after errors were discarded. The data (means for each S) were subjected to a five-way analysis of variance with the factors of

dichotic-mixed, cueing condition, pair-single, negative-positive, and Ss. All four experimental factors had effects significant at least at $p < .025$, but the only significant interaction was that between dichotic-mixed and pair-single [$F(1,5) = 15.1$, $p < .025$]. Separate analyses were also made on the dichotic and mixed results, adding as fourth factors: (1) right vs left ear (half the dichotic NN pairs, randomly selected, were allocated to each ear), and (2) practice—Sessions 1 and 2 in each condition.

In no case was there a significant effect or interaction involving a difference between right and left ears. However, there was an interesting suggestion in the results of an interaction in the mixed condition between right vs left ears and pairs vs single items. (This almost reached significance [$F(1,5) = 5.35$, where 6.61 would give $p < .05$]. The difference was in an unexpected direction: in every condition with the mixed pairs, the mean RT was actually shorter for the left than for the right ear, while in general there was no difference or the reverse for the single items and for the dichotic conditions. If this finding were more reliably replicated, it would suggest that interference between two simultaneous right-ear items is more damaging than interference between two simultaneous left-ear items, as well as the more general finding that left-ear items are more vulnerable to simultaneous right-ear interference than right-ear items are to left-ear interference. In other words, right-ear interference is more damaging to relevant items on *either* ear rather than

Fig. 3. Mean reaction times in each condition.

Table 2
Mean Percent Errors in Responding to Presence or Absence of a Digit

	Pair		Single	
	Nega- tive	Posi- tive	Nega- tive	Posi- tive
Dichotic				
No Cue	3	4	3	1
Visual Cue	2	4	2	2
Auditory Cue	1	2	1	1
Mixed				
No Cue	13	25	3	1
Visual Cue	7	20	1	1
Auditory Cue	8	21	0	1
Control Ss with Unlimited Time: Mixed				
No Cue	17	32	4	0

left-ear items being more vulnerable to interference from either ear.

The effect of practice did not reach significance (averaging 7% for dichotic and 3% for mixed presentations), and the only significant interaction here was that for mixed presentation between cueing condition, negative-positive and practice [$F(2,10) = 10.07$, $p < .001$]. This interaction reflects the fact that with the visual cue practice helped the positives more, while with no cue it helped the negatives. With the auditory cue it had little effect.

One other significant interaction did emerge in these analyses on the dichotic condition only: this was the interaction between pair-single and negative-positive [$F(1,5) = 13.2$, $p < .025$]. This reflects the fact that the difference between pairs and single stimuli in the dichotic condition is greater for negatives than for positives, averaging 118 msec for negatives and 86 for positives.

One further analysis was done separately for dichotic and mixed data on the differences between the no-cue, or classification, condition and each of the two matching conditions with a cue, and one further significant interaction emerged here. Although the effect of the visual and auditory cues did not differ significantly overall, the type of cue did interact significantly with the pair-single factor, in the mixed condition only [$F(1,5) = 10.2$, $p < .025$]. This interaction reflects the fact that, while the visual cue had the same effect on the pairs as on the single stimuli, the auditory cue helped the pairs more and the single items less than the visual cue.

Table 2 shows the mean percent errors in each condition. They are very high in the mixed condition with positive pairs, where Ss failed to detect over 20% of the digits. However, this may not invalidate the comparison of the RTs in the mixed conditions, since it appears to have been due not to a difference in response

criterion (e.g., Ss responding too soon) but to a genuine inability to detect the digit. The evidence for this is the result that the control Ss, given unlimited time to make their decisions, actually made even more errors in this condition.

DISCUSSION

Serial or Parallel Processing

The most striking finding is the absence of any interaction between pairs vs single items and no cue vs the visual cue. The visual cue reduced RT, compared to no cue, by about the same amount in every case. In the introduction we argued that if Ss analyzed each member of the pair serially, the effect of the change of task should be doubled for negative pairs compared to single negatives. This clearly did not occur. With the auditory cue the effect on pairs is slightly larger than the effect on single stimuli, but it is still far from double and, at least with the mixed presentation, the difference is larger for positives than for negatives. These results, therefore, give no support to the idea that the two items in a pair must be examined serially in both tasks. They differ markedly from Sternberg's preliminary data in which Ss searched for a varying number of positive digits in a test pair where the nontargets were also digits, and gave results much more consistent with serial analysis of the test items. The relatively small effect of the cue in the present experiment, although it reduced the target ensemble from 10 to 1, is also of some interest. If a test item were compared serially to each digit in the no-cue condition, this would imply a comparison rate of 12 msec/item, which is three times faster than that typically found by Sternberg (1966) for search through arbitrary subsets of digits. This difference strengthens the conclusion that a different form of processing is involved in the present classification task and in a task which requires discrimination between items *within* a familiar class.

On the other hand, there is one aspect of the present results which could be taken as evidence for serial processing. This is the interaction with dichotic presentation between pairs vs single items and negatives vs positives. If analysis were serial and self-terminating, one would expect the increase in positive RTs to be half the increase in negatives when a second test item is added: In fact, it is 70%, which is rather more than consistently self-terminating comparisons would predict. The interaction is present only with dichotic presentation and not with mixed, which would be consistent with the hypothesis suggested earlier that serial analysis might be forced only when the test items arrive on separate channels. However,

an alternative account is suggested by the surprising discrepancy between the results on the single negative items in the dichotic and in the mixed conditions. The stimuli and the task here were identical (except that the warning tone was given monaurally in the mixed condition and so gave 1 sec advance warning of which ear would receive the test item. Any help this might have given was clearly outweighed by some other disadvantage). Perhaps the simplest explanation for the discrepancy is a difference in Ss' criterion for negative responses in the mixed and dichotic conditions, either making them abnormally fast with the dichotic negatives or abnormally slow with the mixed negatives. If the former were correct, this evidence for serial processing would disappear; but if the latter were correct, this would leave open the possibility of serial self-terminating analysis in the mixed condition too. The latter, in fact, seems more likely. It seems plausible that Ss should be less confident in the mixed than in the dichotic condition that no digit has been presented with, but masked by, the single nonsense syllable. They would thus be more likely to recheck the single negatives in the mixed than in the dichotic condition. This criterion difference would not be expected with the positives. Having heard and identified one item (the digit) there would be no need to check for a nonexistent second one, since Ss knew that no test items contained more than one digit. If this is the correct explanation for the difference between the conditions with these identical negative stimuli, it might be concealing or canceling out the interaction between pairs vs singles and negatives vs positives in the mixed condition, which is revealed in the dichotic condition.

To summarize, strong evidence against the idea of serial analysis of the simultaneous test items is provided by the identical effect of the visual cue on RTs to pairs and to single items. On the other hand, mild support for the idea of serial, self-terminating analysis comes from the steeper increase from single to paired items for negatives than for positives in the dichotic condition; its absence in the mixed condition could be due to an exceptional increase in criterion for the single negatives.

However, the striking absence of *any* interaction, at least with the visual cue, is surprising and may have implications going beyond the question of serial vs parallel analysis. Sternberg (1969) has argued that additivity of effects on RTs suggests the existence of separate stages, each affected by one of the factors which contributes to the total RT. What might these different stages be in the present task? Two possible

Fig. 4. Two hypotheses about the possible stages of processing.

arrangements are shown in Fig. 4.

Possible Processing Stages

In Version 1 the first stage is a preprocessing one at which the two test items are separated or disentangled when a pair is presented. This precedes the stage at which the classification into digit vs nondigit is made. The separation would therefore be based on phonetic continuity within each spoken word, or their separate localization when the items are presented dichotically. This separation stage takes longer with mixed pairs, either because of greater mutual masking between the paired items or because the additional localization cue is absent. At the next stage, the test items are searched for the presence of a digit. This analysis is on average 115 msec faster in the matching than in the classification task, i.e., when the S knows which particular digit he must test for. The fact that the reduction in latency is the same for pairs and single items suggests that this operation is carried out in parallel on the two test items, or that analysis of the pair is selective.

The final stage is that at which Ss decide their response. Since this version assumes analysis to be parallel, the faster RT with the dichotic single negatives is best explained here on the assumption of a lower criterion for single stimuli in the relatively easy dichotic condition. This affects chiefly the negatives, perhaps because of a floor effect on the positives. (This is probably the least satisfactory part of this account of the findings.)

In Version 2, the suggestion is that Ss start by presetting their analyzing systems

for the detection task and that this preparation takes longer when the digit is unknown than when it is known. For example, they might set up a set of 10 templates rather than 1, or critical tests for 10 rather than 1 digit, then use all the templates or tests in parallel on any incoming test stimuli. The analysis of test stimuli would therefore take equal time with 1 and with 10 targets, but the 10 would take longer than 1 to prepare initially. The major drawback here is the necessary assumption that the preparation cannot be completed before the test stimuli are presented, for example as soon as the warning tone occurs, particularly in this experiment where the cueing condition was constant within a block of RTs. The next stage is the preprocessing and separation stage on the test stimuli, similar to that in Version 1, and the final stage is the one at which Ss test serially for the presence of a digit, this stage being unaffected by whether or not the digit is cued. The process is self-terminating, at least on a proportion of trials, so that the difference between pairs and singles is greater for negatives than for positives. The criterion for a negative response is higher for single items in the mixed condition.

Version 2 seems rather less plausible than Version 1, mainly because it does not give a satisfactory account of the first stage. If we accept Version 1, the experiment suggests that with simultaneous items belonging to the different verbal classes of digits and nonsense syllables, analysis can be either parallel or selective, in contrast to the serial processing necessary for identification or

discrimination between two subsets of digits (Sternberg, unpublished). If so, this would confirm the paradox mentioned earlier, that Ss can select by verbal class which of two simultaneous items to identify first (Yntema & Trask, 1963), although they are unable to identify both items in parallel. The present results do not distinguish parallel from selective testing. This might be possible if one set Ss the task of responding "same" if two items matched in verbal class and "different" otherwise. One could then compare simultaneous with successive presentation of the items to be compared and see whether this interacted with advance knowledge of which class the items would share if they were the same.

The apparent conflict between the possibility of parallel or selective classification in the present experiment and its absence in Moray and O'Brien's monitoring task (1967) seems best resolved by the fact that their Ss were required to distinguish which ear received a target letter, while ours were not. This suggestion would, of course, be easy to test. Another possibility, which should be tested, is that the crucial factor is our use of unfamiliar nonsense syllables as negative stimuli compared to their use of digits with letters. Finally, it might be that parallel analysis is possible only for a single pair of items, while simultaneous sequences of items in some way induce Ss to focus their attention on one set of consecutive items at the expense of the other.

Physical vs Name Matching

The discussion so far has been based on the results with no cue and with the visual cue. In a series of experiments with visual stimuli (Posner, 1970), physical matches were found to be faster than visual name matches (upper- and lowercase letters) provided the interval was less than 2 sec. However, auditory-visual matches were slower than visual-visual matches only at intervals of less than 1 sec. Posner suggests that Ss used the interval to "generate a visual code" from the auditory stimulus. Similarly, in the present experiment there was no significant difference in mean RT between the visual and the auditory cue conditions. This may be because our Ss also generated an auditory template from the visual cue, or because the auditory trace of the cue had decayed or been disrupted by the warning tone which occurred between the cue and the test stimulus. However, the fact that there was a significant interaction in the mixed condition between type of cue and pairs vs single items suggests that there was a difference in the form of perceptual processing at least on some proportion of

the trials, even though this did not result in any overall decrease of response time. The greater effect of the auditory cue on the paired than on the single test stimuli would be consistent with serial comparison of the auditory template with the auditory test stimuli. However, the increase with the pairs was less than 100% and was actually less for the negatives than for the positives, which is the opposite effect to that predicted by a serial self-terminating match. The main effect was to decrease the benefit of the advance information for the single positives compared to the benefit given by the visual cue. This would be consistent with a change from feature to template matching, if the former could start immediately the stimulus began while the latter had to wait until the end of the stimulus to match the whole sound against the auditory template. (This possible handicap to physical template matching of course applies only to auditory stimuli and not to the instantaneously available visual stimuli, as used by Posner.) However, the increase in matching time with single stimuli might be more than offset in the mixed pairs by a greater resistance to masking with this form of physical template matching.

Single- vs Dual-Channel Presentation

A final point of interest in this experiment was the comparison of dichotic with mixed monaural presentation. The

results and discussion so far suggest that the perceptual process is similar in both conditions. There is some suggestion that dichotic separation into two inputs forces serial analysis in this classification task, but other conflicting data makes this conclusion doubtful. RTs are faster and errors reduced with dichotic presentation, but this seems best explained by the decreased amount of mutual masking and the greater ease of the preprocessing stage at which the two words are disentangled. The basic similarity of perception with dichotic and with mixed stimuli is consistent with other findings by Savin (1967) and Treisman (1970).

REFERENCES

- BRAND, J. Classification without identification in visual search. *Quarterly Journal of Experimental Psychology*, 1971 (in press).
- BROADBENT, D. E. Role of auditory localization and attention in memory span. *Journal of Experimental Psychology*, 1954, 47, 191-196.
- BROADBENT, D. E. *Perception and communication*. London: Pergamon Press, 1958.
- BROADBENT, D. E., & GREGORY, M. Stimulus set and response set: The alternation of attention. *Quarterly Journal of Experimental Psychology*, 1964, 16, 309-317.
- BRYDEN, M. The manipulation of strategies of report in dichotic listening. *Canadian Journal of Psychology*, 1964, 18, 126-138.
- GRAY, J. A., & WEDDERBURN, A. A. I. Grouping strategies with simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, 1960, 12, 180-184.
- MORAY, N., & O'BRIEN, T. Signal detection theory applied to selective listening. *Journal of the Acoustical Society of America*, 1967, 42, 765-772.
- POSNER, M. I. Abstraction and the process of recognition. In G. H. Bower and J. T. Spence (Eds.), *The psychology of learning and motivation*, Vol. III. New York: Academic Press, 1970.
- SAVIN, H. B. On the successive perception of simultaneous stimuli. *Perception & Psychophysics*, 1967, 2, 479-482.
- STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, 153, 652-654.
- STERNBERG, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Coster (Ed.), *Attention and performance II*. *Acta Psychologica*, 1969, 30, 276-315.
- STERNBERG, S., & SCARBOROUGH, D. L. Parallel testing of stimuli in visual search. Paper presented at the International Symposium on Visual Information Processing and Control of Motor Activity, Sofia, Bulgaria, 1969.
- TREISMAN, A. M. Strategies and models of selective attention. *Psychological Review*, 1969, 76, 282-299.
- TREISMAN, A. M. Perception and recall of simultaneous speech stimuli. In A. F. Sanders (Ed.), *Attention and performance III*. *Acta Psychologica*, 1970, 33, 132-148.
- TREISMAN, A. M., & RILEY, J. G. A. Is selective attention selective perception or selective response? A further test. *Journal of Experimental Psychology*, 1969, 79, 27-34.
- YNTEMA, D. B., & TRASK, F. P. Recall as a search process. *Journal of Verbal Learning & Verbal Behavior*, 1963, 2, 65-74.

(Accepted for publication October 30, 1970.)