Four-channel effects in short-term memory*

BENNET B. MURDOCK, JR.

University of Toronto, Toronto 181, Ont., Canada

A four-channel task was studied in two experiments. Following simultaneous presentation of four words, two auditory and two visual, Ss tried to report the four words and then were probed for the attribute (male or female voice, capital or small type) of one of them. In general, Ss can identify all four words under simultaneous presentation conditions and, usually, identify the attribute as well. Order of report was organized by mode of presentation, even when high associates (e.g., CAT-DOG) were split across mode. These data provide support for the hypothesis of separate short-term auditory and visual stores; they also suggest that these stores are prelinguistic and have persistence of the order of several seconds.

The fact that there are modality effects in short-term memory is no longer particularly newsworthy. However, their proper interpretation is not yet a settled issue. One possibility (Margrain, 1967; Murdock & Walker, 1969) is that there are separate auditory and visual short-term stores and that these modality effects tell us something about the characteristics of these stores. Evidence consonant with this point of view has been reported recently by Fisher & Karsh (1971), Kroll, Parks, Parkinson, Bieber, & Johnson (1970), and Parkinson, Parks, & Kroll (1971).

To date, most modality effects that have been reported have been obtained in standard short-term memory tasks (e.g., memory span, probe tasks, single-trial free recall, keeping-track tasks, shadowing tasks). The to-be-remembered material has generally been processed at the time of presentation, and the obtained results necessarily reflect the joint effects of the initial sensory registration, linguistic processing, and changes during the retention interval. It might be advantageous to try to examine modality effects at an earlier stage of information processing.

To this end, we would like to present some results with a task involving four-channel presentation. On each trial, four words are simultaneously presented to S. Of these four, two are auditory (in the present case, dichotic) and two are visual. Further, the words have sensory-specific attributes (in the present case, male or female voice for the auditory words, upper- or lowercase type for the visual words). The S's task is twofold. First, he must report what the four words were, and then he is probed for an attribute (e.g., asked to report the word which had

been spoken in a male voice).

The purpose of this four-channel task is to look at the interface between perception and memory. In a way, it is the bisensory analog of Moray's four-eared man (Moray, Bates, & Barnett, 1965), which in turn is an extension of Broadbent's dichotic task (Broadbent, 1958). The initial processing goes on under conditions of severe information overload. If the processing is serial, all unrecognized words must be held in store while their number is being reduced. If the processing is in parallel, then the central processor must keep separate feature-extraction records or template matches for each of the four words. Further, linguistic processing is necessary to report what the four words were, but the sensory input must be preserved in some form to identify the probed attribute.

Two experiments will be reported here. To anticipate, it will be shown that under these conditions (1) identification of the four words is quite accurate, (2) order of report is organized by mode of presentation, and (3) information about the sensory characteristics of the words persists for approximately 5 sec. Finally, the second experiment will report some evidence to suggest that these modality-specific stores are, in fact, prelinguistic.

METHOD

The auditory presentation was dichotic, using headphones, while in the visual presentation, one word was shown above the other in the display. The presentation was simultaneous in that the onsets were aligned and the durations approximately equal. In Experiment 1, all four words started within ± 25 msec of each other, but probably this precision is not necessary to assure perceptual simultaneity. The exposure duration was 500 msec for visual presentation and slightly less, on the average, for auditory presentation.

To synchronize auditory and visual presentation, a Uher Diapilot was used

in conjunction with a four-track stereophonic tape recorder. The dichotic words were recorded on Tracks 1 and 3, while the control signals for the Diapilot were recorded on Track 4. The Diapilot activated a Hunter timer, which in turn controlled the closed-circuit TV apparatus for the visual presentation. The display was shown on a 9-in. Shibaden Monitor using a Shibaden CCTV camera. To obtain sharp onset of the picture, which locked in immediately, it was also necessary to use an Ampex Video Control Center as a switching device.

The attributes (capital or small type and male or female voice) were counterbalanced across channels. With equal probability, the male voice was either in the left or right headphone and the word in capital letters was either the upper or lower word in the display. To probe, E spoke one word (e.g., "capital") and S was to report the target (i.e., the word which had been typed in capital letters). The probed attribute was randomized over trials. On each trial, S had to report the four words, then respond to the probe.

EXPERIMENT 1

The Ss were 12 students of both sexes from an advanced undergraduate psychology course at the University of Toronto. After brief practice with dichotic listening, each S was given 48 trials in a single session. The words were common two-syllable words, not more than eight letters long, randomly selected (without replacement) from the Toronto word pool. The experimental session was tape-recorded for subsequent analysis.

EXPERIMENT 2

There were five Ss (three graduate students, one postdoctoral fellow, and one technician), and each S had 32 trials. The words were common monosyllables, and each set of four words contained one pair of highly related words. They were pairs such as CAT-DOG, GOOD-BAD, NIGHT-DAY, and BOY-GIRL. One member of each pair was auditory, the other visual. Which auditory or visual channel each associate appeared on was counterbalanced. The purpose of this manipulation was to see if associative relationships would cut across organization on the basis of sensory modality.

The Ss were not forewarned about the presence of the associatively related words. As in Experiment 1, they were told simply to recall the four words in any order they wished, then to respond to the probe. A voice key was used which started with onset of stimulus presentation and stopped when S started his recall.

RESULTS

The proportion of cases where all four words were correctly recalled was

*This research was supported by Grants APA-146 from the National Research Council of Canada and OMHF-164 from the Ontario Mental Health Foundation.

.70 in Experiment 1 and .98 in Experiment 2. The difference between these two values probably reflects S selection. In Experiment 1, the distribution of scores across Ss was negatively skewed, and there were one or two Ss who were a good bit worse than the others. In any case, these data show that Ss can generally identify correctly four simultaneously presented words.

Given that the four words were correctly recalled, the conditional probability that the order of report was one mode followed by the other (i.e., AAVV or VVAA) was .99 in Experiment 1 and .96 in Experiment 2. Thus, for all practical purposes, the order of report was always ordered by mode of presentation in Experiment 1. In Experiment 2, introducing associatively related word pairs did not affect this organization materially. Actually, after the experiment, each S in Experiment 2 said he had been aware of the fact that associates had been presented. Several Ss said that they had tried to report them in that order once or twice but found it easier to report them grouped by mode of presentation.

Given that all four words were correctly recalled, the conditional probability that the probed attribute was correctly identified was .86 in Experiment 1 and .96 in Experiment 2. Thus, not only could Ss recall the four words, but, in general, they could accurately identify the attribute. Experiment 2 shows a ceiling effect, but in Experiment 1, this conditional probability can be broken down by mode. Conditionalizing as before (for otherwise the data would be difficult to interpret), the probability that the visual attribute was correctly identified was .907 and the probability that the auditory attribute was correctly identified was .808. Of the 12 Ss, 9 were better on the visual, 2 on the auditory, and 1 tied. Thus, if anything, the visual attribute (capital or small type) was more memorable than the auditory attribute.

In Experiment 1, the mean time lapsing between stimulus presentation and presentation of the probe was 6.04 sec. This time provides an estimate of the minimum duration of attribute information necessary to permit accurate identification of the probed attribute.

DISCUSSION

The fact that Ss can generally report all words in this four-channel task is of some interest itself. Even under sudden and heavy overload, the word-recognition system does not break down. Information about separate events obviously must persist for the several seconds it takes the central processor to identify (or at least report) four unrelated common words.

Further, it seems likely that attribute information persists in relatively pristine form until after the four words have been reported. For one thing, most of the Ss reported that they tried to report the words as quickly as possible and that they did not rehearse before reporting. While the (binary) attributes could have been encoded by order of report, only one S did so and then only for the auditory attribute. (Such encoding would be reflected in a report order that was constant vis-à-vis attributes, but only this one S was significantly nonrandom in this regard.) Finally, in Experiment 2, the mean latency (as measured by the voice key) between end of stimulus presentation and the start of recall was 780 msec. This period of time seems too short to permit much rehearsal to occur.

A reasonable presumption, then, is that sensory-specific information about words persists, both visually and auditorily, for something of the order of 5 sec. (What persists cannot be discovered from postexperimental inquiry; reports of auditory and visual imagery were as variable across Ss as the data were constant.) Such a finding is contrary to any models of short-term memory that limit the duration of visual sensory information to Sperling-type millisecond ranges (see Kroll et al, 1970, and Parkinson et al, 1971). These experiments suggest, then, that there are conditions under which auditory and visual persistence may be greater than is sometimes assumed.

The conclusion has been suggested that these modality stores are prelinguistic. The argument is simply that if they were postlinguistic, then, at least some of the time, recall would be ordered associatively rather than by mode of presentation. Since associative ordering almost never occurred, a postlinguistic status for these stores seems unreasonable. That is not to say that an associative order

of report would be impossible. However, it should take longer if the words have to be processed by modality and then reordered for output.

An implication of the very strong ordering effects is that processing is generally serial, mode by mode. Subjectively, one has a strong impression of attending first to one mode and then to the other. However, given the difficulty of deciding between serial and parallel processing in other situations, some caution in this regard is indicated. Clearly some converging operations are necessary before the matter can be considered settled.

This four-channel task might be useful in studying word-recognition processes under demanding conditions or persistence of information in the first few seconds before and after recognition has occurred. It is not unrelated to a subsidiary-task technique (e.g., Murdock, 1965) in that central processing must be allocated to competing demands. Finally, it provides a way of testing some of the initial characteristics of short-term auditory and visual stores.

REFERENCES

- BROADBENT, D. E. *Perception and communication*. New York: Pergamon Press, 1958.
- FISHER, D. F., & KARSH, R. Modality effects and storage in sequential short-term memory. *Journal of Experimental Psychology*, 1971, 87, 410-414.
- KROLL, N. E. A., PARKS, T., PARKINSON, S. R., BIEBER, S. L., & JOHNSON, A. L. Short-term memory while shadowing: Recall of visually and aurally presented letters. *Journal of Experimental Psychology*, 1970, 85, 220-224.
- MARGRAIN, S. A. Short-term memory as a function of input modality. *Quarterly Journal of Experimental Psychology*, 1967, 19, 109-114.
- MORAY, N., BATES, A., & BARNETT, T. Experiments on the four-eared man. *Journal of the Acoustical Society of America*, 1965, 38, 196-201.
- MURDOCK, B. B., JR. Effects of a subsidiary task on short-term memory. *British Journal of Psychology*, 1965, 56, 413-419.
- MURDOCK, B. B., JR., & WALKER, K. D. Modality effects in free recall. *Journal of Verbal Learning & Verbal Behavior*, 1969, 8, 665-676.
- PARKINSON, S. R., PARKS, T. E., & KROLL, N. E. A. Visual and auditory short-term memory: Effects of phonemically similar auditory shadow material during the retention interval. *Journal of Experimental Psychology*, 1971, 87, 274-280.