Bulletin of the Psychonomic Society
1978, Vol. 11 (3), 185-187

The role of visual and acoustic coding in retrieval from very short-term memory

JEFFREY W. JANATA, JOHN M. JOELSON,
KIRBY A. JOSS, and DOUGLAS J. HERRMANN
Hamilton College, Clinton, New York 13323

Several studies have implicated acoustic processing in short-term memory despite stimuli being visual. It is commonly assumed that very short-term memory (VSTM) for visual stimuli, in contrast, operates entirely in the modality of input. The present study tested this assumption by investigating whether an acoustic code is used in VSTM. An experiment required subjects to detect whether a critical letter was present or absent in a brief display of 16 letters. Latency and accuracy of detection were not affected by acoustic similarity but were affected by visual similarity between the critical letter and the other letters of the display. Thus, the results support the assumption that retrieval from very short-term memory (visual) does not involve acoustic information.

It is well known that people are limited in the number of letters they can report from brief visual displays, for example, about 4 or 5 letters out of a 16-letter display presented for 50 msec. Some have proposed that this limit is due to the inability of short-term memory (STM) to accommodate all of the letters as they are read from the trace of the briefly presented display (Estes & Taylor, 1964; Sperling, 1960). The interpretation that STM limits the retrieval from very short-term memory (VSTM) has been challenged recently by Wolford and Hollingsworth (1974). They noted that several investigators have concluded that encoding in STM is predominantly acoustic (Conrad, 1964; Sperling, 1963; Wickelgren, 1965). Wolford and Hollingsworth reasoned that, if items retrieved from VSTM must pass through STM, then acoustic encoding should be implicated in VSTM tasks, as it has been in STM tasks. They found, however, that errors in recall for a briefly presented display of letters were those letters which were visually rather than acoustically confusable with display letters. Wolford and Hollingsworth concluded that acoustic coding was not involved in recall from VSTM.

While no evidence of acoustic confusion in recall from VSTM was apparent in Wolford and Hollingsworth's (1974) data, their study does not preclude the possibility that a VSTM paradigm designed to induce acoustic confusion might produce support for acoustic coding in VSTM retrieval. For example, a subject might be required to detect the presence of a letter in an array of letters (cf. Estes & Taylor, 1964) wherein the critical letter (e.g., B) would be acoustically similar or dissimilar to background letters (e.g., V and J, respectively). If detection involves transforming letters in a display into an acoustic code and matching the code against the critical letter in STM, then acoustic similarity should affect detection accuracy and latency. For example, acoustic matching might follow initial visual processing of stimulus letters, which determines the acoustic code called for. According to this view, both visual and acoustic characteristics of the stimulus would affect processing. Alternatively, if VSTM processing is iconic (Neisser, 1967), that is, solely in the modality of input, then acoustic similarity should not influence detection for visual displays.

The present experiment examined the effect of acoustic similarity between a critical letter and background letters in a task like that described above. To provide a basis of comparison for any possible effects from acoustic similarity, background letters also varied in visual similarity with the critical letter (cf. Krumhansl & Thomas, 1977; McIntyre, Fox, & Neale, 1970). The manipulation of acoustic and visual similarity was done in an orthogonal manner in order to assess the influence, separate and joint, of the two variables.

Requests for reprints should be sent to Douglas J. Herrmann, Department of Psychology, Hamilton College, Clinton, New York 13323. The authors thank Sam Hollingsworth for comments on the manuscript and Edgar S. Catlin, Paul R. Edick, and James R. Harwood for assisting in the data collection and analysis.

METHOD

Subjects

Eight students from Hamilton College served in the experiment and were paid $1.25 for participation.

Apparatus

Stimuli were presented in a three-field tachistoscope (Iconix, Model 6137). One channel presented a blank adaptation field, a second channel presented a fixation field, and a third channel presented the test display. The illumination of all three fields was approximately 3 fc. The fixation field consisted of two horizontal lines, separated by a visual angle of 2.35 deg, which fell outside the area occupied by the test display in the subject's visual field. The fixation field was presented for 600 msec prior to the appearance of the test display.

Responses were made on two microswitches, one held in each hand. The latency of responses was measured by a counter accurate to ±1 msec (Iconix 6246 response buffer).

Stimulus Materials

Six practice displays and 64 test displays were each composed of 16 letters in a 4 by 4 matrix. Letters were typed in uppercase bulletin-style typeface on white cards. The average visual angle of a display letter was .08 deg on the horizontal and .127 deg on the vertical. The matrix subtended a .987 deg visual angle on its diagonal.

Half of the practice displays and half of the test displays contained the letter B. The background letters in the B displays and all of the letters in the non-B displays were either P and D, R and H, V and T, or J and X, representing the following similarity conditions with B: visual and acoustic similarity, visual similarity and acoustic dissimilarity, visual dissimilarity and acoustic similarity, and visual and acoustic dissimilarity, respectivly. The selection of background letters for each condition of similarity with the critical letter B was based on visual and acoustic similarity analyses of letters reported by Wolford and Hollingsworth (1974).

Sixteen displays represented each of the four similarity conditions. Eight displays contained a B and eight did not contain a B. Of the eight displays containing the critical letter, the B was randomly positioned in the upper left quadrant of two displays, in the upper right quadrant of two displays, in the lower left quadrant of two displays, and in the lower right quadrant of two displays. However, no two of the eight displays had the critical letter in the same cell of the matrix. Displays with a B had seven of one background letter and eight of the other background letter randomly assigned to the remaining cells of the matrix. Displays without a B had eight of each background letter randomly positioned in the matrix.

Procedure

Each subject was individually presented with 6 practice displays and 64 test displays with the tachistoscope. The subject was told to determine whether or not the briefly presented display contained the letter B. If the display contained a B, the subject was to press the microswitch held in one hand; if the display did not contain a B, the subject was to press the microswitch held in the other hand. Hand dominance was counterbalanced over subjects. If the subject was unsure of whether or not the display contained a B, he was to guess. The directions stressed accuracy over speed in responding.

The test displays were presented to subjects in a partially random order such that each successive string of eight displays contained two displays representing each similarity condition, one containing a B and one with the B absent. Half of the subjects received the same random ordering of displays and half received a reverse ordering of the sequence seen by the first half of the subjects.

The exposure duration of test displays was altered after each successive block of eight displays, if necessary, to maintain the subject's performance at approximately 70% overall. Exposure durations began at 50 msec and at the end of the session ranged from 5 to 30 msec, with the average exposure duration across subjects being 34 msec. After the offset of the test display, the tachistoscope was dark for approximately 3 sec whereupon the adaptation field reappeared. A test session lasted approximately 30 min.

RESULTS

Figure 1 presents the mean correct latency and the percent accuracy (in parentheses) as a function of the critical letter's presence or absence and the visual and acoustic similarity between the critical letter and background letters. As inspection of the figure reveals, latency was greater to visually similar displays [$F(1,7) = 16.76$, $p < .01$] and to displays with no critical letter than to displays with the critical letter [$F(1,7) = 5.33$, $p < .053$]. Latency was not, however, affected overall by acoustic similarity ($F < 1$), and none of the other variables interacted significantly with acoustic similarity or with each other.

Inspection of the accuracy percentages indicated that accuracy varied with conditions in a fashion similar to the latency data. More errors were made to visually similar displays than to visually dissimilar displays [$F(1,7) = 33.56$, $p < .001$], but error rate overall was not affected by acoustic similarity [$F(1,7) = 2.09$, $p > .10$]. Responding was more accurate to displays not containing a B than to displays containing a B [$F(1,7) = 6.35$, $p < .05$]. As with the latency analysis, none of the variables interacted significantly with each other.

DISCUSSION

The results showed no tendency for acoustic similarity to affect detection accuracy or latency, either alone or in interaction with other variables. It should be noted that another, almost identical experiment run in our laboratory also failed to find that detection was impaired by acoustic similarity. Detection from visual VSTM apparently does not involve transforming the letters in our array into an acoustic code and matching them to an accoustically coded critical letter. Thus, the present study provides support for the previously untested assumption that detection from visual VSTM operates entirely with that modality (Estes & Taylor, 1964). Moreover, the present data are consistent with Wolford and Hollingsworth's

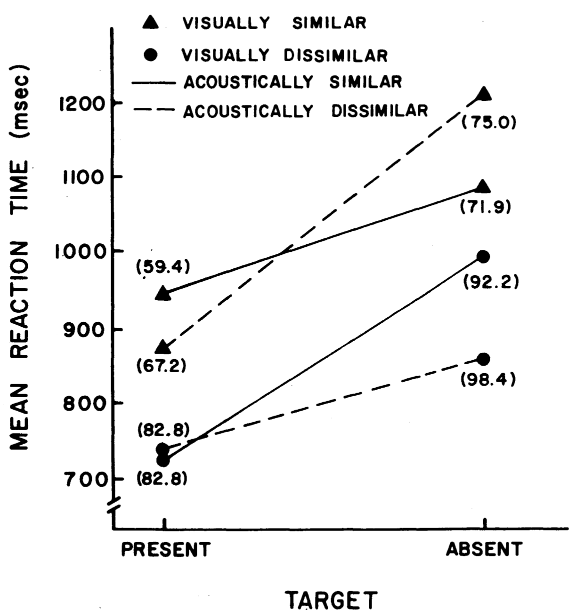

Figure 1. Mean correct reaction time and percent correct (in parentheses) as a function of the presence or absence of the critical letters and of the visual and acoustic similarity between background letters and the critical letter.

(1974) conclusion that an observed failure of acoustic factors to affect VSTM procesisng is inconsistent with the assumption that the information retrieved from VSTM must pass through STM.

Alternatively, visual similarity had clear effects in the present data. Specifically, responses were less accurate and slower when background letters were visually similar to the critical letters. These results for accuracy replicate those of McIntyre et al. (1970). Additionally, the present restults for latency extend the findings of McIntyre et al. on detection accuracy to detection latency. The present accuracy data are also congruent with the effect of visual similarity on recall from VSTM. The effect of visual similarity on detection accuracy and latency is consistent with feature detection models of perception and VSTM (Naus & Shillman, 1976; Neisser, 1967). The quicker latency for visually dissimilar displays and for displays containing the critical letter indicates that the process of searching VSTM for the critical letter (or its features) was in some sense self-terminating. The difference in latency as a function of the critical letter's presence or absence contrasts with previous findings which showed no difference in the latency of detecting one of two alternative letters contained by a display (Wolford, Wessel, & Estes, 1968). Detecting which of a pair of letters is in a display logically involves a self-terminating process. Detecting the absence or presence of just one critical letter, however, permits a self-terminating process for displays containing the letter but requires an exhaustive process for displays not containing the letter. The exhaustive process, of course, would consume more time than the self-terminating process, such as was observed for displays with the critical letter absent and the displays with the letter present in the experiment reported here.

In conclusion, it must be noted that the conflict raised by Wolford and Hollingsworth (1974), and supported by the present work, remains to be settled. One possible explanation for the existence of acoustic coding in STM but not in VSTM could be that VSTM retrieval does not involve use of STM. Another possiblity is that a visual trace is copied from VSTM and placed in STM, after which translation from a visual to an acoustic code occurs when circumstances allow (cf. Craik & Lockhart, 1972). More work is obviously needed to resolve the conflict in the data on encoding effects in VSTM and STM tasks.

REFERENCES

CONRAD, R. Acoustic confusions in immediate memory. *British Journal of Psychology*, 1964, **55**, 75-84.

CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 671-684.

ESTES, W. K., & TAYLOR, H. A. A detection method and probabilistic model for assessing information processing from brief visual displays. *Proceedings of the National Academy of Science*, 1964, **52**, 446-454.

KRUMHANSL, C. L., & THOMAS, E. A. C. Effect of level of confusability on reporting letters from briefly presented visual displays. *Perception & Psychophysics*, 1977, **21**, 269-279.

MCINTYRE, C., FOX, R., & NEALE, J. Effects of noise similarity and redundancy on the information processed from brief visual displays. *Perception & Psychophysics*, 1970, **7**, 328-332.

NAUS, M. J., & SHILLMAN, R. J. Why a Y is not a V: A new look at the distinctive features of letters. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, **2**, 394-400.

NEISSER, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.

SPERLING, G. The information available in a brief visual presentation. *Psychological Monographs*, 1960, **74**(11, Whole No. 498).

SPERLING, G. A model for visual memory tasks. *Human Factors*, 1963, **5**, 19-31.

WICKELGREN, W. A. Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology*, 1965, **70**, 102-108.

WOLFORD, G. W., & HOLLINGSWORTH, S. Evidence that short-term memory is not the limiting factor in tachistoscopic full-report procedure. *Memory & Cognition*, 1974, **2**, 796-800.

WOLFORD, G. L., WESSEL, D. L., & ESTES, W. K. Further evidence concerning scanning and sampling assumptions of visual detection models. *Perception & Psychophysics*, 1968, **3**, 439-444.

(Received for publication December 7, 1977.)