Visual and acoustic confusability in a visual search task¹

G. A. KAPLAN, A. YONAS AND A. SHURCLIFF²

CORNELL UNIVERSITY

Visual and acoustic confusability between a target item and background items was varied in a visual search task. Visual confusability was a highly significant source of difficulty while acoustic confusability had no effect. The results do not seem to be interpretable within a theory which assumes compulsory auditory encoding of visual information.

Several models have recently been postulated which assume a stage of acoustic or auditory transformation of visual input. The models have been derived from studies of such varied phenomena as perception of figure and number strings (Glanzer & Clark, 1962, 1963), latency periods in the reading of words aloud (Levin & Biemiller, 1965; Biemiller & Levin, 1966), amount of recall in a visual-memory task (Sperling, 1963), amount of recall of a written message for a given exposure duration (Mackworth, 1962, 1963), correlation between errors made in immediate recall of visually presented material and those made in the recall of auditorily presented material (Conrad, 1964), and immediate recall of visually presented letter strings (Conrad, Freeman, & Hull, 1965). The above authors have argued that incoming visual information is eventually encoded and stored in a form related to its acoustic representation. However, with the exception of some discussion of the time parameters of this process of encoding, there has been little critical discussion of the generalizability of this process across or within visual tasks.

The present study investigates the relative importance of visual and acoustic factors in a scanning task developed by Neisser (1963), performance on which was shown by Gibson and Yonas (1966a) to be highly affected by the amount of visual confusability between the target and the background. The nature of the task allowed acoustic confusability and visual confusability between the target and the background to be independently manipulated.

METHOD AND PROCEDURE

The apparatus was a modified version of Neisser's (1963) as described by Gibson and Yonas (1966a). A list was composed of 30 lines, four letters wide, typed on a sign typewriter. A line consisted of four letters randomly sampled (with replacement) from the background set of four letters, with the constraint that no letter appear more than twice per line. Each list was 5/8 in. wide with 1/4 in. lines spaced 1/8 in. apart, and the entire list was 10 in. long. For each experimental group, a single target letter (E or K) could appear once in each list on one of 16 lines (1, 3, 5, . . . ,

29, 30). The relationship between the target letter and the background letters was varied by manipulating the amount and type of confusability between the two. Background-target combinations of high acoustic confusability/high visual confusability (H_a-H_v), high acoustic/low visual (H_a-L_v), low acoustic/high visual (L_a-H_v), and low acoustic/low visual (L_a-L_v) were generated using confusion matrices for acoustically presented letter names and visually presented letters obtained by Conrad (1964) and Gibson, Osser, Schiff, and Smith (1963), respectively. For each target letter the remaining letters of the alphabet (excluding Z) were ranked both acoustically and visually with respect to the number of times they were confused with the target letter. Conditions H_a-H_v , H_a-L_v , L_a-H_v , and L_a-L_v were then generated by combining letters which were ranked appropriately for a given condition. The target-background combinations together with the rankings for each member of the set are shown in Table 1.

The Ss were 24 college students fulfilling a course requirement. Half were randomly assigned to each target letter and received the four confusability conditions in a random order, the order being counterbalanced across target conditions. Each S searched through 16 lists per condition for a total of 64 lists. The line on which the target was presented and the position in the line where it appeared were balanced across conditions and randomized within conditions.

Ss were instructed to fixate a small spot of light which

Table 1. Background-Target Confusability

		Target - E							
		H_a				L_a			
H_v	R_a	D	P	G	T	F	H	L	J
	R_v	5.0	1.0	9.0	4.0	22.0	11.0	22.0	24.0
L_v	R_a	3.0	5.5	11.0	11.0	1.0	2.0	5.5	5.5
	R_v	6.0	2.0	3.0	10.0	25.0	15.0	20.0	19.0
		Target - K							
		H_a				L_a			
H_v	R_a	N	L	F	H	X	V	Y	R
	R_v	3.0	7.0	10.0	11.0	18.0	15.0	12.0	14.0
L_v	R_a	2.0	7.5	7.5	4.5	1.0	4.5	3.0	7.0
	R_v	1.0	2.0	4.0	8.0	25.0	24.0	22.5	17.0
		Target - K							
		H_a				L_a			
H_v	R_a	A	B	O	P	C	G	I	T
	R_v	1.0	2.0	4.0	8.0	25.0	24.0	22.5	17.0
		Target - K							
		H_a				L_a			
H_v	R_a	N	L	F	H	X	V	Y	R
	R_v	3.0	7.0	10.0	11.0	18.0	15.0	12.0	14.0
L_v	R_a	2.0	7.5	7.5	4.5	1.0	4.5	3.0	7.0
	R_v	1.0	2.0	4.0	8.0	25.0	24.0	22.5	17.0

Caption—The visual (R_v) and acoustic (R_a) confusability rankings for each background letter with its target. A rank of 1 indicates highest confusability.

Table 2. Mean Scanning Rates for E & K Groups (in sec.)

		E Acoustic confusability		K Acoustic confusability	
		hi	lo	hi	lo
Visual Confusability	hi	.2675	.2674	hi	.3123
	lo	.1760	.1461	lo	.2147

indicated the top of the list and allowed them to converge properly. When ready they were told to press a small hand-held switch which would expose the list and start a clock. Instructions were to scan from the top down as quickly as possible without missing the target. When they reached the target they pressed the switch which removed the list from view and stopped the clock. Ss were told the target could appear on any line, and at the end of each trial they reported the approximate position of the target. If S missed the target, the list was rerun at a later time. Scanning time was recorded by a Hunter "Klockounter" to the nearest msec.

RESULTS

The data consist of 64 scanning times per S, divided equally across the four conditions. Since what is of interest is the time taken to process each line, the time taken to scan each list was transformed by dividing by the line on which the target appeared on that list. This gives a measure of the rate of scanning adjusted for the target position. In general, a linear relationship would be expected between target position and scanning rate. The data confirm this expectation with the exception of the instances where the target item appeared close to the beginning of the list. This initial non-linearity was also found by Gibson and Yonas (1966a) and is, presumably, due to contamination of the rate measure by factors peculiar to the beginning of search. Thus, in order to get as pure a measure of the rate as possible, the data for the trials where the target appeared on any of the first four possible lines (lines 1, 3, 5, & 7) were excluded from the analysis. Table 2 presents the means for the two target groups, averaged across the Ss in each group and across the remaining 12 target positions.

The data for each target group was treated as a single experiment and analyzed by an appropriate analysis of variance. A descriptive summary of the analysis is presented in Table 3. The most striking result is that while the visual factor is highly significant in both groups ($F=56.37, 49.42; df=1/11; p<.001$, for E and K, respectively), the acoustic factor is negligible and non-significant ($F=1.25, 0.00; df=1/11; n.s.$, for E and K, respectively). The target line effect is significant in both analyses ($F=8.32, 8.30; df=11/121; p<.001$, for E and K, respectively). The $A \times V$, $A \times L$, $V \times L$, and $A \times V \times L$ interactions are not significant in either analysis. However, the interactions of A and/or V with Ss are all significant beyond the $p=.001$ level.

Because the experiments were analyzed separately it is not possible to test directly the possibility of an

interaction of the main effects with the two different target groups. This was done indirectly by using a t test for independent groups to compare the distribution of differences due to the level of acoustic or visual confusability in the E group with the same distribution in the K group. Neither analysis indicated any significant differences. Visual inspection of the data indicated that a test for an interaction between target groups and target position was not warranted.

DISCUSSION

The results of this study support those of Gibson and Yonas (1966b) who found that a potentially distracting auditory context (played through earphones) did not interfere with a scanning task identical to the present one. Likewise, Glucksberg (1965) found no effect due to acoustic similarity of items in a short term visual memory task. Chase (1965) compared a visual recognition task, a visual comparison task, and a memory search task and found no effect of acoustic confusability of items in any of the three tasks. He did, however, find a strong effect due to visual confusability (defined by an overlap measure) in his visual comparison task, a result which is supportive of our own.

It should be pointed out that the acoustic and visual interactions with Ss were all highly significant. We interpret this to indicate that different strategies are used by our Ss. However, there were no indications in the data of any S which would indicate a larger effect due to acoustic confusability than that due to visual confusability. In addition, we find no support for the speculation (Kaplan & Carvellas, 1965) that Ss rehearse the target letter; both the reports of our Ss and the results would seem to preclude such an interpretation.

It seems clear that no process which relies on the acoustic encoding of visual information can account for our results. Visual confusability was the major source of difficulty. The scanning task therefore is at least one instance where explanatory mechanisms that assume compulsory auditory representation are of no predictive value. There will, of course, be situations which require some form of acoustic or auditory transformation of visual information; e.g., perhaps those which require rehearsal or verbal identification. However, the con-

Table 3. Percentage of total variance accounted for:

Variable	% of Variance		d.f.	
	E	K	E	K
Subjects (Ss)	26.82	42.86	11	11
Acoustic (A)	.64	.00	1	1
Visual (V)	32.03	22.48	1	1
Target line (L)	3.95	2.83	11	11
$A \times V$	.62	.56	1	1
$A \times L$	.76	.48	11	11
$V \times L$	.20	.53	11	11
$A \times V \times L$	.91	.51	11	11
Ss-interactions	34.06	29.74	517	517
Total	100.00%	100.00%	575	575

textual sensitivity of the perceptual mechanisms must be emphasized: Gibson (1966) and Bower (1965) have suggested that perceptual learning and visual selection, respectively, can be thought of as processes in which there is a filtering of information according to its utility. Such a view is consistent with that taken here; given a visual scanning task in which a visual comparison strategy is sufficient for maximal performance, it would be surprising if acoustic confusability had any effect.

References

- Biemiller, A. J., & Levin, H. Studies of oral reading II. Pronounceability. Preliminary draft, 1966.
- Bower, T. G. R. Visual selection: scanning or filtering. *Psychon. Sci.*, 1965, 3, 561-562.
- Chase, W. G. *The effect of auditory and visual confusability on visual and memory search tasks*. M. A. Thesis, University of Wisconsin, 1965.
- Conrad, R. Acoustic confusions in immediate memory. *Brit. J. Psychol.*, 1965, 55, 75-84.
- Conrad, R., Freeman, P. R., & Hull, A. J. Acoustic factors versus language factors in short-term memory. *Psychon. Sci.*, 1965, 3, 57-58.
- Gibson, Eleanor J. *Perceptual development and the reduction of uncertainty*. Paper to be read at the International Congress of Psychology, Moscow, 1966.
- Gibson, Eleanor J., & Yonas, A. A developmental study of visual search behavior. *Percept. & Psychophys.*, 1966a, 1, 169-171.
- Gibson, Eleanor J., & Yonas, A. A developmental study of the effects of visual and auditory interference on a visual scanning task. *Psychon. Sci.*, 1966b, 5, 169-170.
- Gibson, Eleanor J., Osser, H., Schiff, W., & Smith, J. An analysis of critical features of letters, tested by a confusion matrix. In *A basic research program on reading*. Final Report on Cooperative Research Project No. 639 to Office of Education, Department of Health, Education, and Welfare.
- Glanzer, M., & Clark, W. H. Accuracy of perceptual recall: an analysis of organization. *J. verbal Learn. verbal Behav.*, 1963, 1, 289-299.
- Glanzer, M., & Clark, W. H. The verbal loop hypothesis: binary numbers. *J. verbal Learn. verbal Behav.*, 1963, 2, 301-309.
- Glucksberg, S. *Decay and interference in short-term memory: comments on experiments by Steffy and Eriksen*. Paper read at Psychonomic Society meetings, Chicago, 1965.
- Kaplan, I. T., & Carvellas, T. Scanning for multiple targets. *Percept. mot. Skills*, 1965, 21, 239-243.
- Levin, H., & Biemiller, A. J. Studies of Oral Reading III. Contingent versus non-contingent spelling patterns. Preliminary draft, January, 1966.
- Mackworth, Jane F. The relation between the visual image and post-perceptual immediate memory. *J. verbal Learn. verbal Behav.*, 1963, 2, 75-85.
- Mackworth, Jane F. The visual image and the memory trace. *Canad. J. Psychol.*, 1962, 16, 55-59.
- Neisser, U. Decision-time without reaction-time: experiments in visual scanning. *Amer. J. Psychol.*, 1963, 76, 376-385.
- Sperling, G. A model for visual memory tasks. *Hum. Factors*, 1963, 5, 19-31.

Notes

1. This research was partially supported by grant OE 6-10-156 from the United States Office of Education.
2. The authors would like to thank Dr. E. J. Gibson for her valuable advice and support.

(Received in the Editorial Office May 9, 1966.)