

# Recognition and confusion of the lowercase alphabet

L. H. GEYER

*Department of Industrial Engineering, Northeastern University, Boston, Massachusetts 02115*

The confusion matrix for the full lowercase English alphabet is estimated, based upon 25 trials per letter for each of seven subjects. Average correct recognition was controlled to 0.5 by limiting brightness and duration of displays to individually determined levels. Comparison of the obtained data to that reported by Bouma (1971) for eccentric vision supports the conclusion that limited energy foveal recognition is qualitatively different from eccentric vision recognition. Comparison of the obtained data to the uppercase confusion matrix reported by Townsend (1971) supports the inference that recognition performance has more between-letter variability for both recognizability and confusion pairings for the lowercase alphabet than for the upper.

The confusion matrix for the full English alphabet is a 26 by 26 array of stimulus-response probabilities. Each row represents an individual stimulus letter, the main-diagonal cell is the correct response probability for that letter, and incorrect responses generate off-diagonal confusion probabilities. Each row must sum to 1.0; column sums may vary from unity, and the variation is sometimes taken as an estimate of response bias effects.

Most modern theories of recognition of familiar symbols assume the use of unique sets of distinctive attributes, i.e., features, as the basis of recognition processes: Geyer (1977), Geyer and DeWald (1973), Gibson (1969), Gibson, Gibson, Pick, and Osser (1962), Neisser (1967) Townsend (1971), Gibson, Osser, Schiff, and Smith (Note 1), Yonas and Gibson (Note 2). As many of these studies make clear, confusion data and especially full alphabet confusion matrices are among the best sources of data for inferring the identity of the specific features comprising the set used in recognition processes.

One problem in the development of alphabetic confusion matrix data is limiting correct performance. Under normal viewing conditions, correct recognition is highly probable and a resultant confusion matrix is uninteresting. It has become quite usual to employ some visual technique to attempt to control correct recognition, averaged over all stimuli, to 0.50. Thereby, differences in main-diagonal cells are interpreted as estimates of differential recognizability of stimuli and the off-diagonal cells as estimates of confusability of individual letter pairs.

A very good confusion matrix data base for the

uppercase English alphabet has been reported by Townsend (1971). He restricted average correct recognition to 0.50 by limiting stimulus duration to an individually predetermined time for each of six subjects. Each subject saw each letter 25 times, generating an average performance matrix based upon 150 presentations of each letter. Importantly, for users of these data, he required a response to each trial. Individually adjusted average correct recognition and forced choice response eliminates possible confounding of confusion processes with individual uncertainty criteria, i.e., one subject's guess might otherwise be internally identical to another's "I don't know," if this latter response is permitted.

The data base for confusion performance of the lowercase alphabet is not as assuring. Bouma (1971) reports the results of two studies using different techniques to control average correct performance to 0.50—one a long viewing distance and the other eccentric (nonfoveal) presentation. In the first, he did not use a forced choice procedure. In the second, he used a forced choice procedure, but the letter "y" was not part of the stimulus set. Perhaps more importantly, it should be regarded as an empirical question whether or not performance limitations due to eccentric presentation tap the same sources of recognition errors as do limitations due to foveal presentations at a controlled energy level. The critical role of feature processing in theories of recognition/confusion has been noted.

At 7° eccentric vision angle, the density of rod endings is less than 1% the density of cone endings in the central region of the fovea, and rod endings do not have unique retinal location to optical nerve fibre mappings as do cones (Geldard, 1953). There would appear to be no reason for assuming that recognition processes under these circumstances tap the full range of features that may be available upon

I appreciate the efforts of Mr. F. Pirzada in conducting the laboratory sessions reported.

foveal viewing. The data from Bouma's (1971) study will be compared to the results obtained in this study, using a limited energy foveal presentation, in an attempt to explore this question.

There is considerable variability in the actual configuration of the various letters of the lowercase alphabet which is a function of the particular type style. Much of this variability is in degree of ornateness. This study used an extremely simple, unornamented letter style in order to minimize the influence of any particular type style idiosyncrasy upon obtained performance, and thereby hopefully to maximize the aspects of recognition/confusion which are common across type styles. The alphabet used, Tactype Futura demi 5424, is shown in Figure 1.

## METHOD

### Subjects

Seven paid volunteers, all males, ranging in age from 20 to 25 and with normal or corrected-to-normal vision, served as subjects.

### Stimuli

The lowercase alphabet was projected upon a screen from transparent slides with a single opaque letter centered on each slide. Letters were Tactype Futura demi 5424. The visual angle subtended by stimuli varied from letter to letter. The angle subtended vertically by the "tall" letters was 0.2 deg. The stimulus appeared centered upon a dim fixation dot.<sup>1</sup>

### Apparatus

Stimuli were presented via a slide projector with Gerbrands G9/66 electronic shutter control mounted. The shutter control was equipped with carrier for neutral density filters. The interval from the subject's initiation of a trial to stimulus onset and the stimulus duration were controlled by millisecond timers (Lafayette Type 50013).

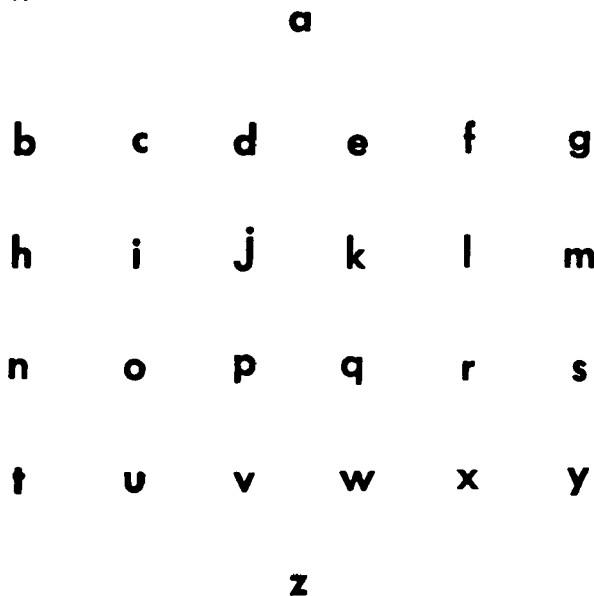


Figure 1. The Tactype Futura demi 5424 alphabet.

### Procedure

The presentation sequence for a trial was: (1) blank white prestimulus field with fixation dot, (2) stimulus field containing a randomly selected letter (3) return to the blank prestimulus field. Trials were subject-paced; when the subject pressed the stimulus-initiation lever, the stimulus appeared after a 0.4-sec delay. The subject then responded with his best opinion. Each subject had a different predetermined random sequence of letters, constrained such that in the first data-taking session there were 12 presentations of each letter and in the second, 13.

Each subject went through four sessions. The first two were preparatory sessions, during which the stimulus brightness and duration was adjusted individually to achieve average probability correct values of 0.50. In these sessions, all letters were presented equally often. The third and fourth sessions were comprised of 312 and 338 data trials, respectively, i.e., 12 and 13 presentations of the alphabet. Data gathering in Sessions 3 and 4 was preceded by 40 warm-up trials.

## RESULTS

Correct performance, averaged across stimuli and subjects, was 0.50. Highest percent correct for any subject was 55.7 and lowest, 40.8. Table 1 presents the obtained composite confusion matrix.

Comparison of Table 1 to the similar matrix for uppercase letters obtained by Townsend (1971) indicates that variations both in recognition (main-diagonal) performance and confusion (off-diagonal) errors are larger for the lowercase data. The sum squares of deviation from the mean (0.5) for the 26 main-diagonal elements for Table 1 and Townsend's data are 1.04 and 0.25, respectively [ $F(25,25) = 4.16, p < .01$ ]. Similarly, sum square variation of the 650 off-diagonal cells from their mean (.02) is .80 for the lowercase data and .51 for the upper [ $F(623,623) = 1.57, p < .01$ ]. Maximum correct recognition in Townsend's study was .66 (for Z); in this study, eight letters had correct recognition greater than .66 (d, g, h, j, k, m, w, and y). Minimum correct recognition in Townsend's study was .26 (for B); in this study, five letters had lower recognition (c, e, s, t, and z).

Townsend's (1971) AON model was used to generate a predicted confusion matrix for the lowercase alphabet. This model assumes that either the stimulus is perceived correctly or there is a complete state of uncertainty. Each letter as stimulus has its own recognizability parameter, and each letter as response has its own response bias parameter. This study used a computerized iterative gradient search procedure, varying these two free parameter arrays, so as to minimize the sum square error between model predictions and data matrix. This solution procedure was preferred to Townsend's (1971, p. 43) because it avoids the computational ambiguity he noted as associated with zero valued matrix cells. In fact, when applied to his uppercase data, as a check, the results were essentially identical to his report. Tables 2 and 3 contrast the unexplained sum squares

**Table 1**  
Confusion Matrix for the Lowercase Letters of the English Alphabet

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
a	.54	.01	.01	.06	.05	.00	.03	.01	.00	.00	.00	.02	.05	.14	.01	.03	.00	.00	.01	.03	.00	.00	.00	.01	.00	.00
b	.03	.53	.01	.02	.01	.00	.06	.21	.01	.00	.07	.00	.00	.02	.01	.00	.01	.00	.00	.00	.02	.01	.00	.01	.00	.00
c	.01	.03	.23	.02	.02	.03	.02	.01	.09	.02	.09	.00	.01	.03	.01	.00	.00	.18	.04	.02	.02	.02	.01	.10	.01	.01
d	.01	.03	.01	.80	.01	.00	.01	.00	.01	.06	.00	.01	.01	.01	.01	.00	.00	.00	.00	.00	.01	.01	.00	.01	.00	.01
e	.18	.03	.00	.05	.21	.01	.05	.02	.01	.02	.00	.00	.01	.06	.27	.02	.03	.00	.01	.00	.01	.00	.02	.00	.00	.00
f	.00	.00	.01	.03	.00	.27	.01	.01	.06	.15	.01	.26	.00	.00	.00	.01	.01	.11	.01	.02	.02	.00	.00	.01	.01	
g	.03	.00	.00	.01	.02	.00	.75	.01	.00	.01	.00	.00	.00	.03	.02	.01	.07	.00	.01	.00	.02	.00	.01	.00	.01	
h	.01	.05	.00	.01	.00	.00	.02	.69	.01	.00	.09	.01	.01	.05	.01	.01	.00	.00	.00	.02	.01	.00	.02	.00	.00	
i	.01	.01	.03	.02	.00	.03	.00	.01	.34	.09	.01	.18	.01	.00	.01	.01	.01	.01	.00	.07	.01	.00	.00	.03	.01	
j	.00	.00	.00	.01	.00	.03	.01	.00	.03	.80	.00	.10	.00	.00	.00	.00	.01	.00	.01	.00	.01	.00	.00	.00	.01	
k	.01	.01	.00	.00	.01	.02	.01	.07	.01	.00	.71	.03	.00	.02	.01	.00	.00	.01	.01	.02	.01	.00	.01	.03	.01	
l	.00	.00	.00	.01	.00	.07	.00	.00	.12	.09	.02	.50	.01	.00	.00	.01	.01	.06	.01	.09	.00	.01	.00	.01	.01	
m	.03	.00	.01	.00	.00	.00	.01	.01	.01	.02	.01	.01	.73	.06	.00	.01	.01	.02	.01	.00	.01	.01	.02	.02	.01	
n	.05	.01	.01	.02	.01	.01	.02	.03	.02	.01	.00	.01	.02	.64	.03	.01	.02	.03	.02	.00	.01	.01	.01	.01	.00	
o	.14	.02	.03	.01	.10	.01	.02	.01	.01	.00	.00	.00	.01	.07	.42	.03	.02	.02	.02	.01	.05	.01	.01	.01	.01	
p	.03	.03	.01	.02	.03	.00	.03	.02	.00	.00	.02	.00	.02	.14	.06	.50	.01	.01	.01	.01	.02	.00	.02	.01	.01	
q	.15	.00	.00	.02	.01	.02	.10	.01	.01	.01	.01	.03	.04	.02	.05	.44	.01	.01	.01	.02	.00	.01	.00	.00	.00	
r	.02	.01	.02	.01	.01	.03	.02	.01	.07	.02	.03	.01	.00	.02	.01	.02	.00	.37	.04	.05	.01	.03	.03	.09	.04	
s	.06	.03	.02	.00	.07	.00	.02	.04	.02	.01	.06	.01	.01	.13	.07	.01	.02	.05	.14	.01	.01	.06	.05	.07	.03	
t	.00	.00	.03	.02	.00	.04	.01	.01	.14	.02	.09	.15	.00	.01	.00	.00	.01	.10	.03	.24	.02	.02	.00	.02	.03	
u	.08	.04	.01	.03	.02	.00	.02	.01	.01	.01	.02	.01	.02	.04	.05	.00	.00	.01	.01	.00	.50	.07	.03	.01	.01	
v	.00	.03	.00	.00	.01	.00	.00	.01	.02	.00	.03	.01	.00	.03	.03	.00	.00	.03	.02	.01	.04	.51	.15	.02	.05	
w	.01	.00	.01	.01	.01	.01	.00	.00	.00	.00	.01	.00	.03	.03	.00	.00	.01	.00	.00	.01	.00	.01	.06	.73	.02	
x	.01	.01	.02	.01	.01	.00	.01	.01	.04	.01	.08	.01	.01	.04	.01	.01	.00	.05	.03	.01	.01	.04	.02	.53	.02	
y	.01	.00	.00	.00	.00	.02	.01	.00	.01	.00	.02	.00	.00	.01	.01	.01	.00	.01	.00	.01	.02	.13	.02	.02	.67	
z	.01	.00	.04	.02	.01	.02	.03	.00	.03	.02	.03	.00	.01	.01	.02	.01	.00	.10	.05	.01	.03	.07	.04	.16	.19	

variance for the upper- and lowercase alphabets, using the AON model.

The sum square deviation of the 676 cells of the complete confusion matrices obtained in this study and Bouma's (1971) eccentric vision matrix is 2.08. The sum square deviation between the full matrix of this study and a matrix uniformly composed of 0.5 along main-diagonal cells and 0.02 elsewhere is 1.79. Thus, there is less agreement between these data and Bouma's than there is between these data and a uniform S/R matrix totally devoid of psychological information except that average correct performance was constrained to 0.5.

**DISCUSSION**

The fact that the confusion matrix reported herein differs from that obtained by Bouma (1971) more than it does from a uniform matrix of .5 on the main diagonal and .02 elsewhere would appear to support the idea that foveal and nonfoveal recognition processes differ qualitatively. Perhaps, therefore, these data and Bouma's should be used selectively, depending upon the perceptual information processing situation involved.

Despite the higher variability of lowercase recognition/confusion data than of upper case, Townsend's (1971) rather simple AON model predicts both matrices about equally well. It is true that the model's main-diagonal performance appears notably better for lowercase data and off-diagonal performance better for uppercase. However, probably this just reflects the four times higher main-diagonal variability of the lowercase data. In this situation, the AON model, with its great parametric concentration on main-diagonal performance, could and would achieve best overall results by minimizing large main-diagonal errors and tolerating some increase in the less potent off-diagonal predictions. Note that this is only a matter of degree; the AON model does quite well with the main diagonal in both cases.

Clearly, there are limitations to what can be expected from a model which does not deal with similar-

**Table 2**  
Lowercase

AON vs. Data	SS <sub>error</sub>
Along Main Diagonal	.0002
Off Diagonal	.5828
Total	.5830

**Table 3**  
Uppercase

AON vs. Data	SS <sub>error</sub>
Along Main Diagonal	.02
Off Diagonal	.47
Total	.49

ity of individual letter pairings. Townsend (1971) presented two models which postulated individual pairwise similarity parameters for each letter pair. He showed these models to yield about one-half the residual variance of the AON model. Most probably, a similar value could be obtained for these lowercase data. However, these models necessarily postulate a very large number of free parameters (350 and 351, respectively). It is difficult to relate such results to specific information or perceptual processes.

Geyer and DeWald (1973) analyzed Townsend's (1971) data using a feature processing model and specific sets of features for the capital letters. They obtained residual variance essentially the same as the AON model, but with only 17 free parameters—basically, the retention strength of each feature plus two guessing threshold parameters. To apply that model to these data implies the preliminary step of mapping the letters of the lowercase alphabet into distinct lists of features from a hypothesized feature set. For the lowercase alphabet, the literature does not provide any well-studied starting point.

The investigator attempted to capture the feature mappings implied, at least conceptually, by previous literature. Five trial sets of feature lists explored feature list assumptions dominated by three basic characteristics: (1) upper and lower protrusions (2) curved and straight segments and (3) patterns closed or open in various quadrants. None of these five feature sets when run with the Geyer and DeWald (1973) model produced residual variance values lower than that for the uniform matrix discussed earlier, nor were any illuminating pattern for synergistic combinations detected. While the feature processing model may be less capable of predicting lowercase data, at least, for the moment, the more appealing assumption is that the feature sets tried were very inaccurate.

#### REFERENCE NOTES

1. Gibson, E. 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*, Cooperative Research Project No. 639, U.S. Office of Education, 1963.

2. Yonas, A., & Gibson, E. J. *A developmental study of feature processing strategies in letter discrimination*. Paper presented at the Eastern Psychological Association, Boston, April 1967.

#### REFERENCES

- BOUMA, H. Visual recognition of isolated lower-case letters. *Vision Research*, 1971, 11, 459-474.
- GELDARD, F. A. *The human senses*. New York: Wiley, 1953.
- ERIKSEN, C. S., & COLLINS, J. F. Reinterpretation of one form of backward masking in visual perception. *Journal of Experimental Psychology*, 1965, 70, 343-351.
- GEYER, L. H. Temporal course of the icon: A summary. *Perceptual and Motor Skills*, 1977, 44, 668.
- GEYER, L. H., & DEWALD, C. G. Feature lists and confusion matrices. *Perception & Psychophysics*, 1973, 14, 471-482.
- GIBSON, E. J. *Principles of perceptual learning and development*. New York: Meredith, 1969.
- GIBSON, E. J., GIBSON, J. J., PICK, A. D., & OSSER, H. A developmental study of the discrimination of letter like forms. *Journal of Comparative and Physiological Psychology*, 1962, 55, 897-908.
- KAHNEMAN, D. Methods, findings, and theory in studies of visual masking. *Psychological Bulletin*, 1968, 70, 404-425.
- NEISSER, U. *Cognitive psychology*. New York: Meredith, 1967.
- TOWNSEND, J. T. Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, 1971, 9, 40-50.

#### NOTE

1. A reviewer of an earlier draft of this paper noted that the fixation dot constituted a possible source of forward and backward masking. In general, any immediate pre- or poststimulus presentation display will cause some performance reduction due to luminance-summation/contrast-reduction effects (Eriksen & Collins, 1965; Kahneman, 1968). However, substantial effects due to luminance summation/contrast reduction appear to require mask areas many times that of the stimulus letter (cf. Eriksen & Collins, 1965). The fixation dot was small compared to the stimulus letters, smaller in diameter than their stroke width, as well as being viewed as part of a dimly illuminated pre- and postfield. The contribution of the fixation dot to performance limitation is estimated to be extremely small. The dot could have been located slightly below the stimulus letter, as Townsend (1971) did, but he thereby accepted an angle of almost  $\frac{1}{2}$  deg between fixation point and stimulus display center. This appears to be a tradeoff, but one which seems unlikely to substantially influence the results.

(Received for publication March 30, 1977;  
revision accepted July 14, 1977.)