FINGERSPELLING BY COMPUTER

by

Stephen A. Weyer

TECHNICAL REPORT NO. 212

August 17, 1973

PSYCHOLOGY AND EDUCATION SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

FINGERSPELLING BY COMPUTER*

Stephen A. Weyer

Stanford University

In this paper, I describe two experiments using computer graphics to represent the alphabet used for manual communication by deaf persons. The first experiment measured subjects' ability to read fingerspelled sentences at different rates of presentation. The second experiment used scaling techniques to measure similarities between fingerspelled characters by examining the confusions caused when the characters were rapidly presented to subjects.

The fingerspelling alphabet, which consists of 26 hand positions, is shown in Figure 1. Each character was coded as a sequence of graphics commands for an Imlac Corporation PDS-1 graphic display. The display model used in the experiments

Insert Figure 1 about here

---

*This research was supported by OE Grant OEG-0-70-4797 (607). I appreciate the assistance of Patrick Suppes, Dexter Fletcher, Adele Goldberg, and Marian Beard.

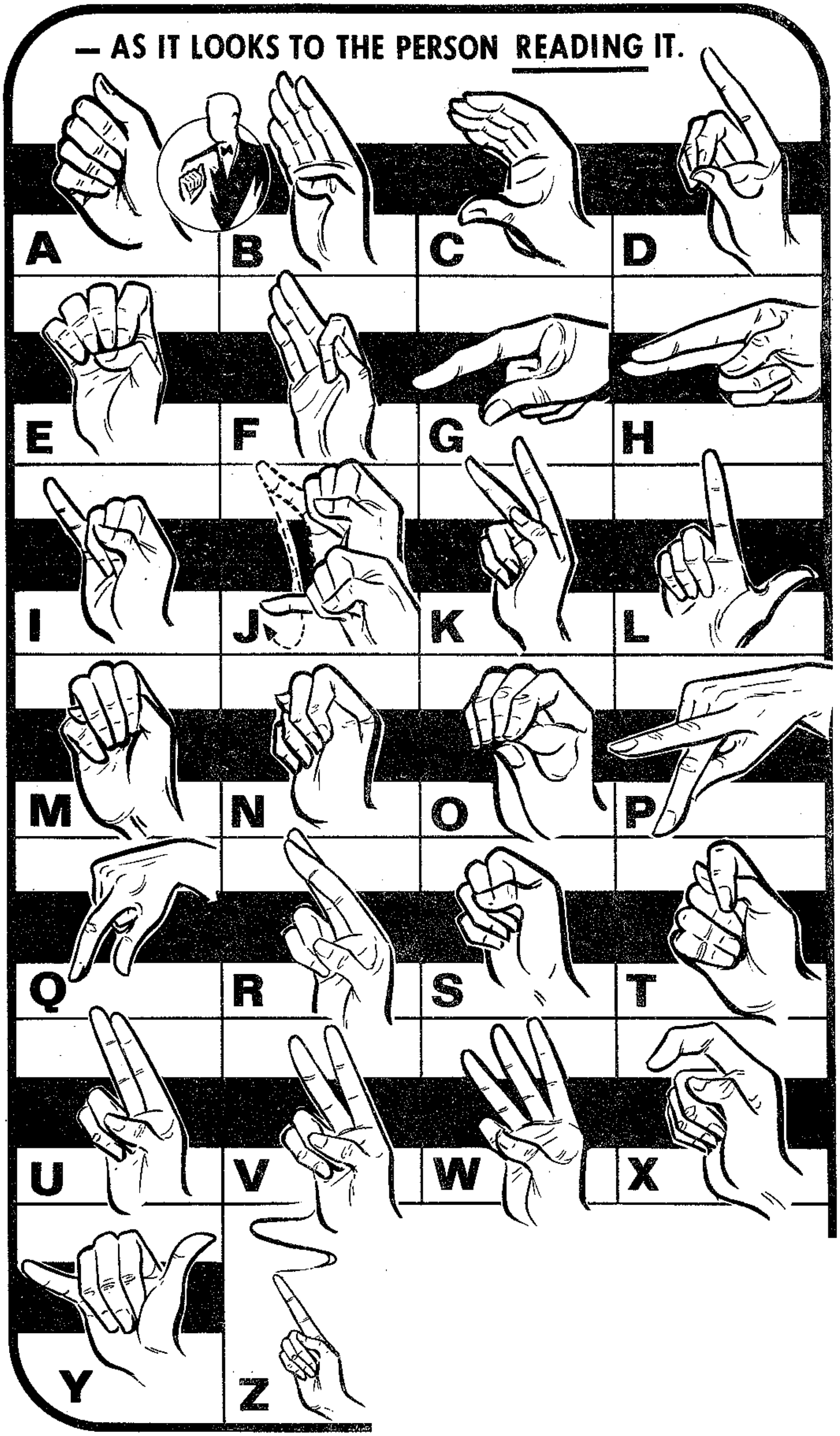

Fig. 1. The manual alphabet.

has a 4096-word memory and the capability to refresh the display screen 40 times each second. Although the fingerspelled characters represented on the Imlac were small and individually contained within a 5/8-inch square area, they were easily readable. The Imlac display communicated with the PDP-10 computer at the Institute for Mathematical Studies in the Social Sciences (IMSSS), Stanford University. A program running on the PDP-10 monitored the presentation of items in each experiment and recorded subjects' response data.

1. THE FINGEX EXPERIMENT

FINGEX, the first experiment, attempted to increase receptive manual communication skills by training subjects to read fingerspelled sentences presented at different display rates. Learning to read fingerspelling is perhaps the most difficult task in learning manual communication. Six hearing subjects, who had already memorized the manual alphabet, participated in the experiment. Each subject completed 21 FINGEX sessions.

1.1. Procedure

Each item in FINGEX consisted of an incomplete sentence that was fingerspelled on the Imlac display. Breaks between words were indicated by a blank character. Each sentence was followed by a list of four words displayed as ordinary orthographic characters. Subjects were to choose the one word from among the four displayed that best completed the fingerspelled sentence.

For example, the FINGEX program fingerspelled the incomplete sentence: A very small piece of bread is called a .... Subjects then saw the following four words displayed in orthographic characters.

1 cake
2 ball
3 cut
4 crumb

Subjects were then required to type the number corresponding to the word that best completed the fingerspelled sentence.

Forty items were presented during each FINGEX session of about 20 minutes. The first 10 items were fingerspelled at the rate of one character per second; in the three successive groups of 10 items each, the characters were displayed at presentation rates of 1.3, 2, and 4 characters per second, respectively. Depending on response time of the computer system, these times may occasionally have been slightly longer. The 200 items used in FINGEX were selected from Primary and Intermediate forms of the Stanford Achievement Test. The following items are typical of those used.

One who is honest tells the

1 cause
2 truth
3 news
4 time

When a girl grows up, she becomes a

1 father
2 sister
3 son
4 woman

To drive a nail into a piece of wood, you should have a

1 hammer
2 bottle
3 boat
4 ladder

The items for each subject were randomly selected from the pool of 200 items. Sessions 1, 6, 11, 16, and 21 were tests made up of items not previously presented; the intermediate training sessions presented only those items used in the immediately preceding test. Thus, forty items were drawn at random and without replacement from the item pool for sessions 1-5; forty more items were drawn for sessions 6-10; etc.

Times between sessions varied because students were permitted to schedule their sessions at their convenience. Although subjects took from two weeks to two months to complete all sessions, time lapses between individual sessions were not considered in analyzing the data for FINGEX.

1.2. Results

The average number of correct responses for each presentation rate in the five test sessions is presented in Figure 2. At 1, 1.3, and 2 character per second

Insert Figure 2 about here

presentation rates, subjects reached the highest level of performance between sessions 6 and 11. Subjects described the 1-second presentations as being too slow; the slow presentation rate made it difficult to remember each character and to form the words. The data indicate that, generally, the subjects performed better at the 1.3 character per second presentation rate than at the 1 character per second rate. Performance at the 2 character per second rate was almost as good as at the slower presentation rates.

The large gains for items presented at the 4 character per second rate contrast with the minor gains at slower speeds. At 4 characters per second, subjects answered about 3 items correctly on the first test and about 8.5 items correctly on the last test. The slope of the middle portion of this curve, however, does not accurately reflect the relatively large between-subject variance observed in the data. This variance might be due to delays between sessions and the differences between subjects based on fingerspelling ability.

1.3. Discussion

A person's manual receptive skills might be improved by specifying a learning model that would determine the speed and difficulty of the next item to be presented. Thus, the choice of

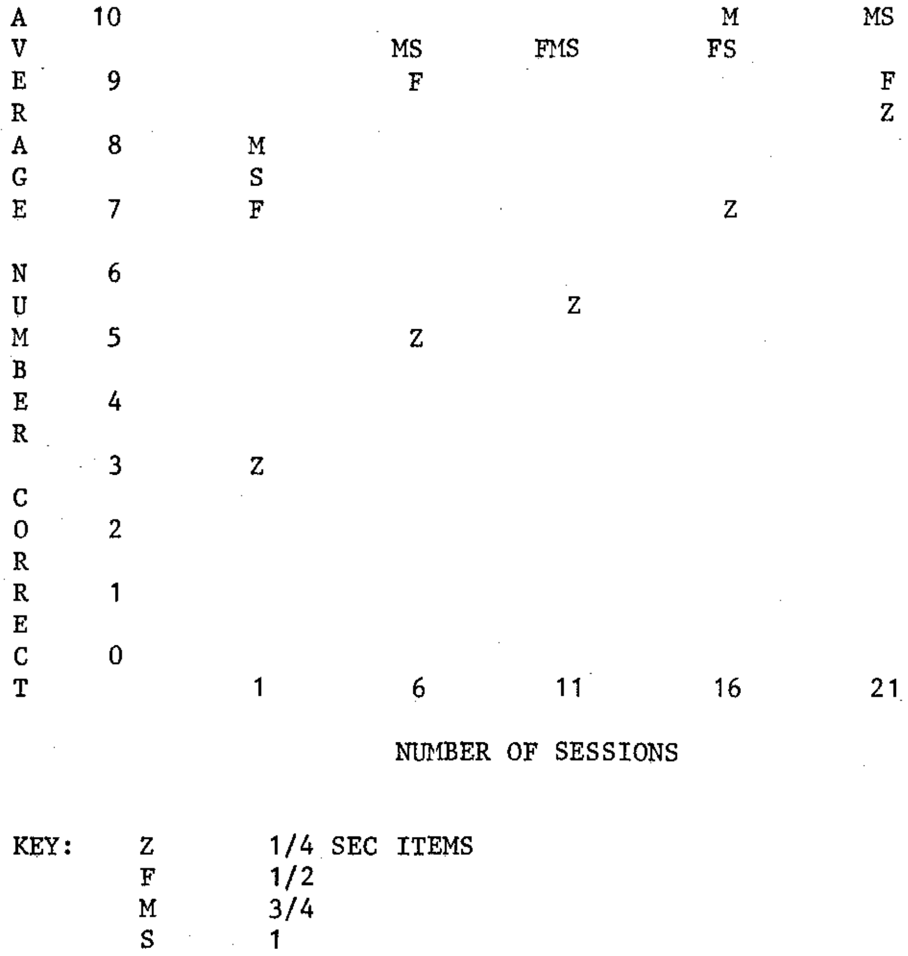

Fig. 2. Five tests of fingerspelling comprehension at four speeds for six subjects.

(Fig. 2, continued.)

	SECONDS PER CHARACTER			
SESSIONS	0.25	0.50	0.75	1.00
1	3.17	7.17	8.17	7.50
6	5.00	9.00	9.50	9.50
11	5.67	9.33	9.33	9.50
16	6.83	9.50	9.83	9.50
21	8.50	9.17	9.83	9.83

AVERAGE NUMBER CORRECT

the next item would depend on the past history of the student's responses, speeds at which items were presented, item difficulty, and the desired percentage of correct responses.

The FINGEX teaching strategy was simple and the intent of the experiment was to describe students' progress. Because the total number of items was small, subjects were trained on items from the previous test before receiving new items on the next test. A less monotonous sequence of items might have been more motivating to the subjects.

In addition, the format of items presented was too restricted. In particular the items were plagued by a high frequency of catch phrases such as "... is called a ...," "when you ...," "to ... is to ...." Too often the answer depended on one key word. An alternate approach would be to vary the length of the items and the mode of response. For example, we could present a paragraph followed by several multiple-choice questions, or we could spell a single word to the subject and require him to transliterate it to traditional orthography.

Zakia and Haber (1971) compared the processing of both orthographic and fingerspelled letter sequences by deaf and hearing subjects. A PDP-8 computer controlled the tachistoscopic presentation of orthographic characters and a deaf person fingerspelled the characters to deaf subjects. The usual rate for sending fingerspelled words in context to a proficient reader is about 200 milliseconds per letter, but in the Zakia and Haber data the rate varied from 162 to 527 milliseconds per letter. Measures of word length, presentation rate, and word familiarity, i.e., low- and high-imagery words versus nonwords, were correlated with the mean number of letters correct. In FINGEX, however, no data were collected for individual letters, and subjects had to perceive the letter sequences as words or sentences. Zakia and Haber noted that experienced fingerspellers did not attend to single hand positions. Instead they concentrated on the overall pattern of finger configurations. Thus, another possible modification of FINGEX would take account of the patterns of finger configurations presented.

The computer-generated fingerspelling presented in FINGEX appeared to be readable and significantly useful in increasing the fingerspelling skills of the subjects. With more flexible graphics systems, it should be possible to display signs for whole words, e.g., signs that would display the face, body, and both arms.

2. THE CONFUS EXPERIMENT

2.1. Procedure

CONFUS, the second experiment, measured similarities between the 26 characters of the manual alphabet. Three deaf subjects and 12 hearing subjects completed a total of 31 CONFUS sessions. A session lasted approximately 10 minutes. For more accurate timing than that used in FINGEX, a routine residing in the Imlac memory controlled the display duration for each character. The main program, which ran on the PDP-10, allowed subjects to display any desired fingerspelled character before each CONFUS session so that they could familiarize themselves with the Imlac keyboard and the computer representation of the manual alphabet before beginning the sessions.

During an individual session, there were 5 presentations of each fingerspelled character or 130 presentations in all. Sequencing of the presentations was randomly ordered except that no character occurred twice in succession. After the character was displayed for 50 milliseconds, a noise pattern masked the disappearing image. The subject was then required to type the orthographic character corresponding to the fingerspelled character displayed. Response latencies were measured as the number of milliseconds that elapsed between display of the noise pattern and receipt of the subject's typed response.

Two matrices of data, one for confusion frequencies and the other for latencies, were collected from each subject. The matrices were of the form:

$$\begin{array}{llll} f(A,A) & f(A,B) & \ldots & f(A,Z) \\ f(B,A) & f(B,B) & \ldots & f(B,Z) \\ \vdots & & & \\ f(Z,A) & f(B,Z) & \ldots & f(Z,Z) \end{array} \qquad \begin{array}{llll} l(A,A) & l(A,B) & \ldots & l(A,Z) \\ l(B,A) & l(B,B) & \ldots & l(B,Z) \\ \vdots & & & \\ l(Z,A) & l(Z,B) & \ldots & l(Z,Z). \end{array}$$

The matrix indices are the occurrences of each character in the manual and orthographic alphabets, f(i,j) is the frequency with which a fingerspelled character i is said to be a j, and l(i,j) is the total latency of these responses.

These data were used to locate the 26 hand positions as points in a space. The distance between each pair of points in the space depended on subjects' tendency to confuse the two fingerspelled characers represented by the points. The distance measure used was the Euclidean distance metric.

Nonmetric multidimensional scaling is a technique originally developed by Shepard (1962, 1972) and Kruskal (1964a, 1964b) to represent the structure and dimensionality underlying proximity data such as that obtained in the CONFUS experiment. The ranked ordering of n*(n+1)/2 measures of similarity between pairs of n objects is monotonically related to distances among n points in some underlying coordinate space. The assumptions for this model are the following metric distance axioms and corresponding similarity constraints:

Distance Axioms	Similarity Constraints
1. $0 \leq d(i,i) < d(i,j)\ i \neq j$ and $d(i,i) < d(j,i)\ i \neq j$	1. $s(i,i) \geq s(i,j)$ and $s(i,i) \geq s(j,i)$
2. $d(i,j) = d(j,i)$	2. $s(i,j) \cong s(j,i)$
3. $d(i,j) + d(i,k) \geq d(j,k)$	3. if $s(i,j)$ and $s(i,k)$ are both large, then $s(j,k)$ should be at least moderately large.

We assign the following meaning to each symbol.

<	less than or equal
>	greater than or equal
≠	not equal
=	approximately equal
d(i,j)	distance of point i to point j
s(i,j)	similarity of object i and object j

Because data collected from each subject were sparse, all data for all subjects were combined into one frequency matrix and one latency matrix. In the resulting frequency matrix, shown in Table 1, the diagonal entries

---

Insert Table 1 about here

---

are larger than the off-diagonal entries, and differences between symmetric entries are generally small. Because the definitions of 'large' and 'moderately large' are relative and do not seem to describe many of the off-diagonal entries, the triangle inequality, corresponding to similarity constraint 3, is harder to check. Thus, the data were triangularized to ensure symmetry for

TABLE 1
Frequency Confusion Matrix for 15 Subjects.

RESPONSES TYPED (A - M)

		A	B	C	D	E	F	G	H	I	J	K	L	M
	A	152	1											
	B		138				11							
	C			153				1						
	D				137		4			1		1	1	
LETTER PRESENTED	E	2	1			113							1	23
	F		5		3		136	2		2	1	1		
	G							132	19			1		
	H							28	122					
	I				2		2			136	11			
	J				1		1	1		4	131	1		
	K	1					1					88		
	L												154	
	M	6				26		1			1			108
	N	2				6								20
	O	1		1	1	5				1				14
	P		1		1		1					7		
	Q								1					
	R				2		2		2					
	S	6			4	12	1	1	1					12
	T	12	2		1									2
	U		3		2		7	1						
	V							1				27		
	W					1	1	2				8		
	X				1									
	Y						1							
	Z	1			1		2		1		2	2		

(TABLE 1, continued.)

RESPONSES TYPED (N - Z)

LETTER PRESENTED	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A							1						
B								2					
C					1								
D					10						1		
E	5					6	3	1					
F							1	2					
G				1									
H			2										
I								1					
J												1	10
K	1				1			2	55				
L													
M	5					2	4						
N	82	1			2	11	26				1		
O	3	125									2		
P			143										
Q			2	147		1				1	1		2
R			1		144								
S	20	8			1	74	12				1		
T	3					4	124				1	1	
U					8			131					
V			1			1		5	111	6			
W								1	12	127		1	
X								1			146		6
Y								1				152	
Z					6			1					127

multidimensional scaling. The arithmetic average of symmetric entries was computed for frequencies and total latencies using the obvious computations:

$$[f(i,j) + f(j,i)]/2 \text{ and } [l(i,j) + l(j,i)]/2.$$

To obtain 'normalized' latencies, I divided the sum of symmetric latencies by the sum of symmetric frequencies:

$$[l(i,j) + l(j,i)]/[f(i,j) + f(j,i)].$$

An inspection of the frequency matrix revealed few violations of the similarity constraints.

In multidimensional scaling, 'stress' denotes goodness of fit or departure from monotonicity. What is considered a 'good' or 'poor' stress value often depends on how complete the data are and how well they satisfy the metric axioms. In interpreting results, one may increase the number of dimensions until some acceptable level of stress is achieved and then attach a meaning to the coordinates. However, increased dimensionality obscures the model, and additional coordinates may merely fit errors in the data.

2.2. Results

MDSCAL (a computer program written by J. B. Kruskal, 1964a, 1964b, version 5M) yielded stress values of.2354 for three dimensions and .3107 for two dimensions of the frequency matrix. Figure 3 shows the spatial

Insert Figure 3 about here

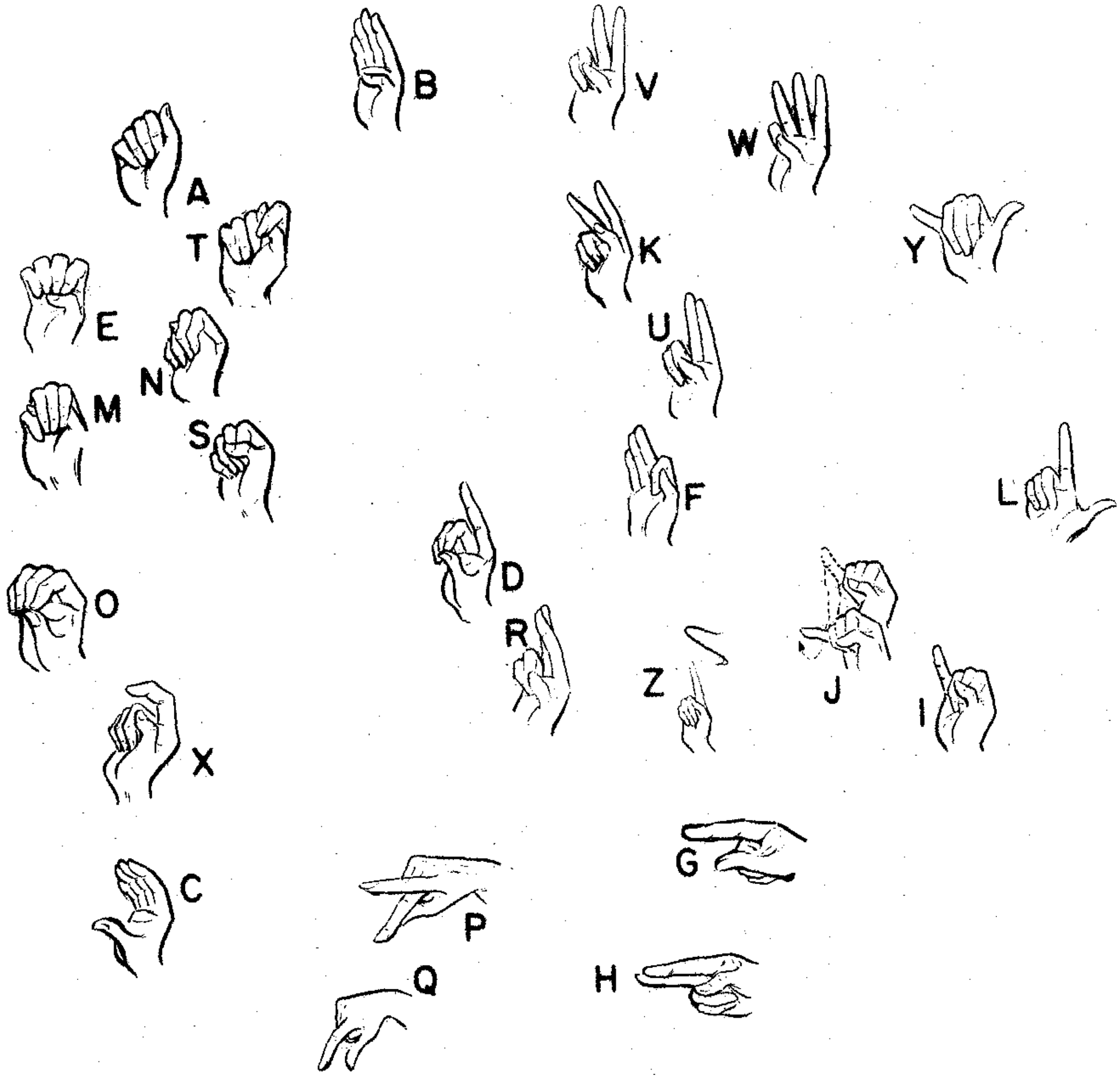

Fig. 3. Two-dimensional spatial configuration obtained by applying multidimensional scaling to the similarity data of Table 1.

configuration of the manual alphabet in two dimensions, and depicts hand symbols adjacent to their corresponding names and coordinates. To further aid in interpretation, the HICLUS program (written by S. C. Johnson, 1967) used the similarity measures to derive a hierarchical clustering (diameter method). This clustering was then superimposed on the set of objects separated by derived MDSCAL distances (Figure 4).

Insert Figure 4 about here

Rather than label axes or attach special significance to the number of dimensions, the investigation examined clusters of objects, as shown in Figures 3 and 4, to see whether they are in fact similar. The largest cluster, composed of S, N, T, and A, are hand positions that involve making a fist or folding all fingers. They differ from one another in thumb position only. The other signs in the group represented by fists and folded fingers, O, M, and E, comprise an adjacent cluster. B, F, and U are represented by 4, 3, and 2 fingers, respectively, extended vertically. The character K, frequently confused with V, looks like a V both on paper and on the computer display. K is included in the group made up of two fingers extended vertically, although from a side perspective one finger appears nearly horizontal. V is represented by two fingers spread like a V; W uses three fingers. R, D, X, and Z all involve the index finger, either

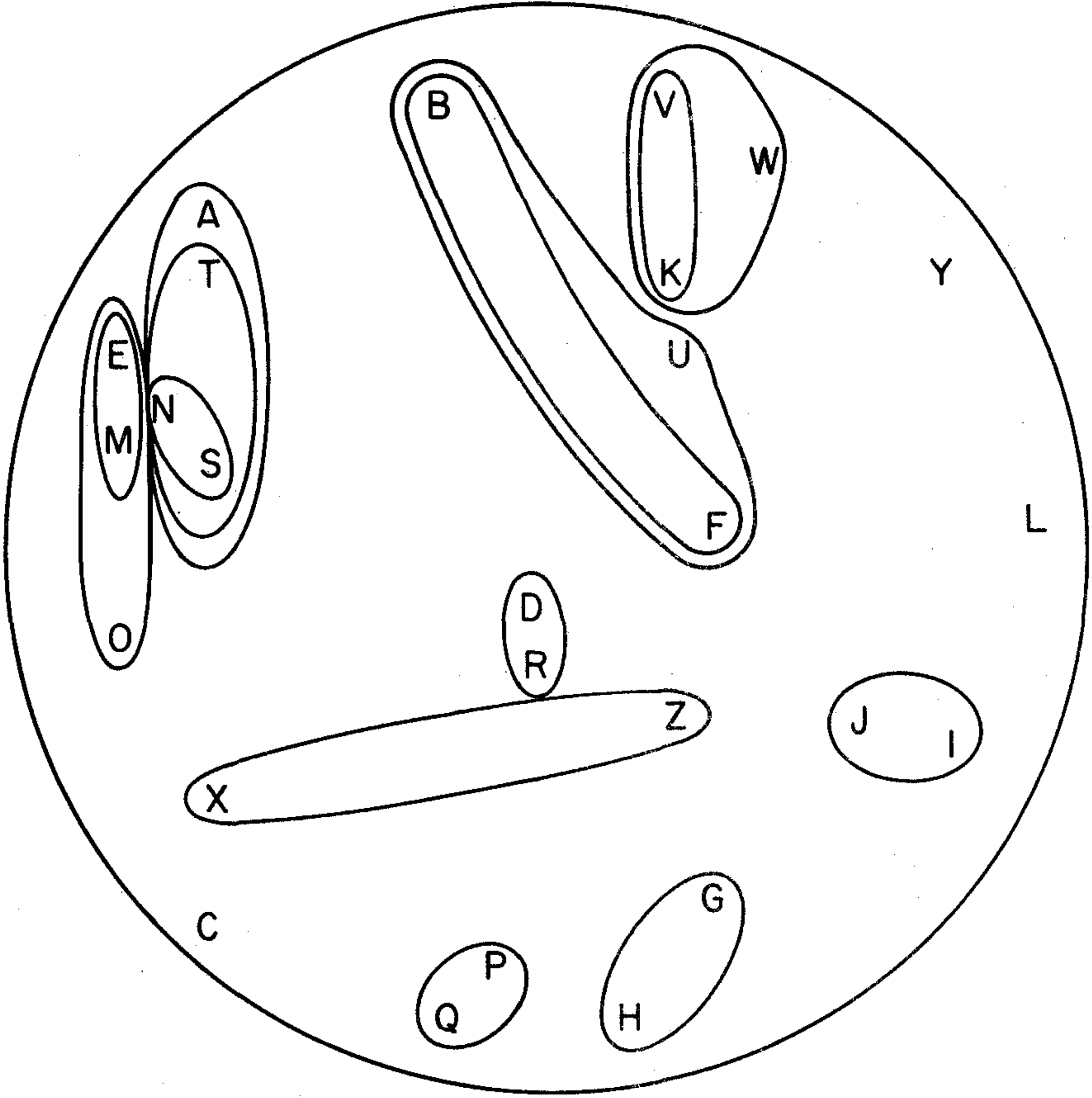

Fig. 4. Hierarchical clustering applied to the data of Table 1 and superimposed on the spatial solution of Figure 3.

crossed with the middle (R), extended (D), bent (X), or moving zigzag (Z). G and H are distinctive and were confused only with one another. I and J appear similar except for their orientation in space. On the computer display the movement of J (as for Z) was indicated by a dotted line. P and Q have the same 'down' orientation, but share little resemblance in shape and were not often confused in the data. C, L, and Y are not included in any cluster. This may be explained by the almost nonexistent confusion of C, L, or Y with any other letter.

Several other confusions present in the data are not distinguished by the cluster analysis although they are consistent with the MDSCAL solution. For example, D and F are complementary signs with one finger up and three down or one finger down and three up, respectively; U and R involve two vertical fingers (U) or the same two fingers crossed (R). Another confusion probably resulted from the dotted line shown with J and Z, which indicates movement rather than similarity in shape. The signs, P and K, which are the same except for orientation, were confused, although this confusion did not appear in the MDSCAL or HICLUS solution.

2.3. Discussion

Object confusions within clusters were high while confusions between clusters were low. This clustering indicates a lack of firmness or determinacy in the distances between clusters, which implies that there could be other solutions. This intercluster structure could be revealed by adding more subjects or decreasing the display duration in order to increase errors. However, it is also valuable to explore certain subsets of characters by using latencies and to compare deaf with hearing subjects.

In order to obtain more accurate response times, subjects were not allowed to change an answer once it had been typed, and this requirement led to spurious confusions. On the other hand, group latencies were not useful, because subjects differ in typing skills and because the latencies acted like weights on the frequencies. It is uncertain, however, whether this weighting was in the direction of similarity or of dissimilarity. Normalized latencies were not meaningful because the off-diagonal entries differed both in magnitude and direction from the diagonal entries, in clear violation of the metric axioms. However, subsets of these latencies may be useful to test the hypothesis that the set of confusing alternatives differs for each symbol and subject. Although it is a limited sample, the frequency data in Table 1 indicates, generally, that the cardinality of these sets is no larger than 9 (S is an exception with 13).

Locke (1970) compared data on the kinesthetic similarity judgments of deaf subjects on nine consonant fingerspelled characters with the data of Conrad and Rush (1965) on recall errors made by deaf subjects for the corresponding nine orthographic characters. Conrad and Rush dealt with short-term memory encoding, and they found that deaf subjects do not appear to forget orthographic characters on the basis of phonetic or visual confusions. Locke suggested that covert motor rehearsal might affect the similarity judgments and proposed to measure the 'feel' or kinesthetic similarity of the corresponding fingerspelled characters. I expected the similarity judgments for tactile perception in Locke's experiment to be related to those for visual perception. Multidimensional scaling applied to Locke's data failed to yield interpretable results consistent with the visual confusion data and the spatial solution for confusions reported here.

Given more data, separate representations for different groups of subjects might have been derived for investigating the hypothesis that hearing subjects confuse fingerspelling stimuli on both visual and auditory dimensions in contrast to the visual and possibly kinesthetic confusions of deaf subjects. Another interesting grouping would have compared skilled with novice fingerspellers. Unfortunately, the number of confusions by individual subjects and by deaf subjects overall was too small to permit these analyses.

In conclusion, the computer-generated manual alphabet was found to be a useful tool in teaching fingerspelling and in obtaining empirical measures of similarity.

REFERENCES

Conrad, R., and Rush, M. L. On the nature of short term memory encoding by the deaf. *Journal of Speech and Hearing Disorders*, 1965, *30*, 336-343.

Johnson, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, *32*, 241-254.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, *29*, 1-27.

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, *29*, 115-129.

Locke, J. L. Short-term memory encoding strategies of the deaf. *Psychonomic Science*, 1970, *18*, 233-234.

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. I, II *Psychometrika*, 1962, *27*, 125-140, 219-246.

Shepard, R. N., Romney, A. K., and Nerlove, S. (Eds.) *Multidimensional scaling: Theory and applications in the behavioral sciences*. Volumes 1, 2. New York: Seminary Press, 1972.

Zakia, R. D., and Haber, R. N. Sequential letter and word recognition in deaf and hearing subjects. *Perception and Psychophysics*, 1971, *9*, 110-114.