*Behavior Research Methods & Instrumentation*
*1983, 15 (4), 406-412*

# Compact representations of positional knowledge in short and long words for letters and features

IAN MORRISON
*Queen's University, Kingston, Ontario, Canada*

**Single-letter statistical measures providing values for each letter-position and word-length combination are too unwieldy for use in working letter-recognition models; more compact tables are needed. Compact tables collapsing word length to short or long words and letter positions to first, last, or middle letters are presented for the frequency and versatility of single letters. Letter-position and word-size differences are preserved in this reduced format. To test awareness of these values, subjects rate the commonness of letters in each letter position. Their responses indicate high intersubject agreement and correlate highly with the frequency and versatility measures. A LISP program that translates the letter knowledge for each letter into a corresponding knowledge for each feature in a feature set is described. Distinctiveness values for each feature (see Shimron & Navon, 1981) are computed.**

Single-letter statistical measures such as frequency and versatility have been tabulated for various word-length and letter-position combinations ( e.g., Mayzner & Tresselt, 1965; Solso, 1979; Solso & King, 1976). These tables provide a powerful source of knowledge but may be too unwieldy, or too detailed, for working models of letter recognition. Although it is not difficult to store large tables with a machine, it does seem incredible that humans would retain information in such detail. Rawlinson (1976) expressed a similar reservation regarding bigram frequency tables and provided a new, more compact table of bigram frequency with only three letter-pair positions: first, last, and other. Here, the same distinction is applied to single-letter statistics. Word-length divisions are limited to short (three-five letters) and long (six-eight letters) words. Source values before collapsing were taken from Solso (1979) and Solso and King (1976) for four- to eight-letter words. Three-letter word values were compiled from the Kučera and Francis (1967) word-frequency count. Two statistical measures were compiled in this collapsed format: the frequency and the versatility of letters. Whereas frequency is the number of times a letter appears in a given position, versatility is the number of times it appears in that position in different words.

Conspicuously absent from most papers tabulating letter-frequency information is some indication of how well the measures reflect people's actual knowledge. Although the primary purpose was the production of more compact, yet still precise, tables of positional letter information, people's knowledge is still of interest. A demonstration of knowledge of statistics would not indicate when, how, or even if, people use the knowledge to aid letter recognition, but it would show that the knowledge is available for use. The absence of demonstrated knowledge may indicate only covert awareness, not ignorance. Still, the presence of subjects' knowledge of the statistical properties of letters would be reassuring to most letter-recognition theorists. Consequently, the subjects here were asked to rate the commonness of letters for the three letter positions in both short and long words.

## METHOD

The basic tables in this paper list the frequency and versatility measures for short and long words in the first, middle, and last letter positions. The middle-position value is an average for the middle positions in a word. For example, the middle frequency value in a five-letter word is the average of the second, third, and fourth positions. The average is taken so that longer words do not have artificially larger middle values. Similarly, the values for short and long words are averages for their constituent sizes: three-, four-, and five-letter words for short words and six-, seven-, and eight-letter words for long words.

The source values for frequency and versatility for four- to eight-letter words were taken from the tables provided by Solso (1979) and Solso and King (1976). The three-letter-word values were compiled from the Kučera and Francis (1967) count, since that was the source for the Solso and King tables. Words and abbreviations not listed in *Webster's New Collegiate Dictionary* (1976) were not used in the frequency and versatility calculations. Once the middle positions had been averaged, the values for each position (i.e., first, middle, and last) were averaged for each set of three word sizes to obtain a value for each letter position in both short and long words.

Computing equipment used in compiling the tables and producing the program in this paper was supported by Natural Sciences and Engineering Research Council of Canada Grant A-9581 to Brian E. Butler of Queen's University. The author was supported by a postgraduate grant by the same institution while the paper was in preparation. Requests for reprints should be sent to Ian Morrison, Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada.

**Table 1**
**Frequency**

| | Short | | | Long | | |
|---|---|---|---|---|---|---|
| | First | Middle | Last | First | Middle | Last |
| A | 16796 | 20646 | 755 | 5431 | 5940 | 680 |
| B | 6315 | 501 | 253 | 4285 | 747 | 27 |
| C | 4643 | 2373 | 265 | 6708 | 2592 | 834 |
| D | 3188 | 1278 | 19941 | 3344 | 1542 | 11045 |
| E | 2258 | 15803 | 45948 | 3339 | 10828 | 10092 |
| F | 9161 | 761 | 578 | 3403 | 725 | 516 |
| G | 3001 | 1089 | 1892 | 1763 | 1534 | 5623 |
| H | 10817 | 32842 | 6534 | 2462 | 2043 | 1823 |
| I | 1713 | 12332 | 93 | 1942 | 6749 | 125 |
| J | 1093 | 36 | 0 | 496 | 68 | 2 |
| K | 1092 | 857 | 2697 | 386 | 590 | 245 |
| L | 4670 | 5889 | 5260 | 2501 | 4100 | 3349 |
| M | 6375 | 1793 | 3143 | 4032 | 1732 | 1048 |
| N | 4143 | 15548 | 8732 | 1651 | 5357 | 5517 |
| O | 5457 | 17424 | 2967 | 1552 | 5803 | 295 |
| P | 3188 | 740 | 1026 | 6371 | 1510 | 126 |
| Q | 197 | 14 | 1 | 287 | 97 | 0 |
| R | 2420 | 8076 | 11136 | 4205 | 5914 | 5804 |
| S | 9736 | 3363 | 16400 | 9195 | 2901 | 11995 |
| T | 39403 | 4697 | 17506 | 3572 | 5036 | 5718 |
| U | 1349 | 6647 | 1134 | 773 | 3247 | 49 |
| V | 826 | 2013 | 4 | 1032 | 927 | 14 |
| W | 16429 | 1387 | 2659 | 2410 | 551 | 330 |
| X | 4 | 156 | 317 | 0 | 234 | 88 |
| Y | 2493 | 459 | 7507 | 140 | 444 | 6053 |
| Z | 43 | 87 | 59 | 18 | 97 | 23 |

**Table 2**
**Versatility**

| | Short | | | Long | | |
|---|---|---|---|---|---|---|
| | First | Middle | Last | First | Middle | Last |
| A | 127 | 310 | 96 | 353 | 577 | 161 |
| B | 178 | 29 | 24 | 436 | 88 | 8 |
| C | 163 | 56 | 21 | 555 | 187 | 75 |
| D | 122 | 53 | 145 | 335 | 160 | 878 |
| E | 63 | 283 | 333 | 225 | 857 | 690 |
| F | 121 | 17 | 20 | 285 | 67 | 17 |
| G | 99 | 41 | 47 | 242 | 121 | 491 |
| H | 112 | 51 | 77 | 242 | 144 | 107 |
| I | 35 | 221 | 35 | 147 | 590 | 51 |
| J | 46 | 2 | 0 | 67 | 6 | 0 |
| K | 50 | 31 | 79 | 72 | 81 | 64 |
| L | 125 | 149 | 103 | 222 | 394 | 225 |
| M | 128 | 50 | 51 | 323 | 152 | 181 |
| N | 54 | 129 | 160 | 119 | 462 | 464 |
| O | 52 | 254 | 68 | 130 | 455 | 85 |
| P | 139 | 41 | 56 | 448 | 133 | 23 |
| Q | 9 | 1 | 1 | 25 | 10 | 0 |
| R | 107 | 170 | 116 | 352 | 489 | 413 |
| S | 250 | 74 | 431 | 752 | 253 | 1307 |
| T | 139 | 91 | 179 | 318 | 365 | 371 |
| U | 21 | 141 | 13 | 91 | 270 | 14 |
| V | 41 | 29 | 3 | 108 | 72 | 7 |
| W | 85 | 29 | 24 | 193 | 54 | 29 |
| X | 2 | 8 | 17 | 0 | 18 | 16 |
| Y | 20 | 25 | 186 | 22 | 48 | 478 |
| Z | 9 | 10 | 11 | 11 | 23 | 11 |

These tables are more compact, but less precise, representations of the same information resident in the Solso and King (1976) tables. They are more compact because there are fewer letter-position and word-length combinations, and they are less precise because of the averaging. The term "same information" is used because averaging for the middle position took place at the level of individual word size before the data were collapsed to short- or long-word categories. The word-size categories are, then, somewhat deceptive. First, the sizes included in each category resulted from an arbitrary decision. Second, a true two-category distinction would not average before collapsing to obtain the category values. Therefore, although there are fewer letter-position and word-length combinations, the more detailed information from the larger number of combinations is retained, but in an admittedly less precise form.

Eight volunteers ranked the letters of the alphabet in descending order of commonness for each letter-position and word-length combination as found in the tables described above.

## RESULTS

The basic tables for the statistical measures in reduced letter-position and word-length form are presented in Tables 1 and 2. The average correlation over positions between frequency and versatility is .914. The near identity is present for individual letters in long words but not for those in short words. Anomalies exist between frequency and versatility for individual letters in short words. The mismatches do not involve the same letters over all letter positions. For example, the first position shows mismatches for B and C, but the middle and and last positions show identity between frequency and versatility for the two letters. The identity of the most frequent and versatile letters also differs over letter positions for both short and long words. For example, the most frequent letters for short words are T, H, and E for the first, the middle, and the last positions, respectively. Similar differences exist for the versatility of letters in words. Figure 1 shows the similarities and differences between the two statistics; letters have been graphed according to their standing within the alphabet in standard deviation units for the frequency and versatility measures.

The frequency-versatility mismatches in short words involve mainly letters that are highly versatile but not very frequent (e.g., B, C, D, L, M, P, and S in the first letter position). Some mismatches involving a very frequent but not very versatile letter (i.e., the reverse relationship) can be removed by discounting the effects of only a few words. For example, discounting the word "THE" greatly narrows the gap between frequency and versatility for T, H, and E, since the word's frequency count is by far the highest in the corpus (see Kučera & Francis, 1967), at 69,000 occurrences. Other frequent words, or frequent prefixes and suffixes, could have the same effect on their constituent letters. Drewnowski and Healy (e.g., Drewnowski, 1978, 1981; Drewnowski & Healy, 1977, 1980; Healy, 1976, 1980; Healy & Drewnowski, 1983) have shown that high-frequency short words and suffixes are treated differently from the way words or letter groups are treated.

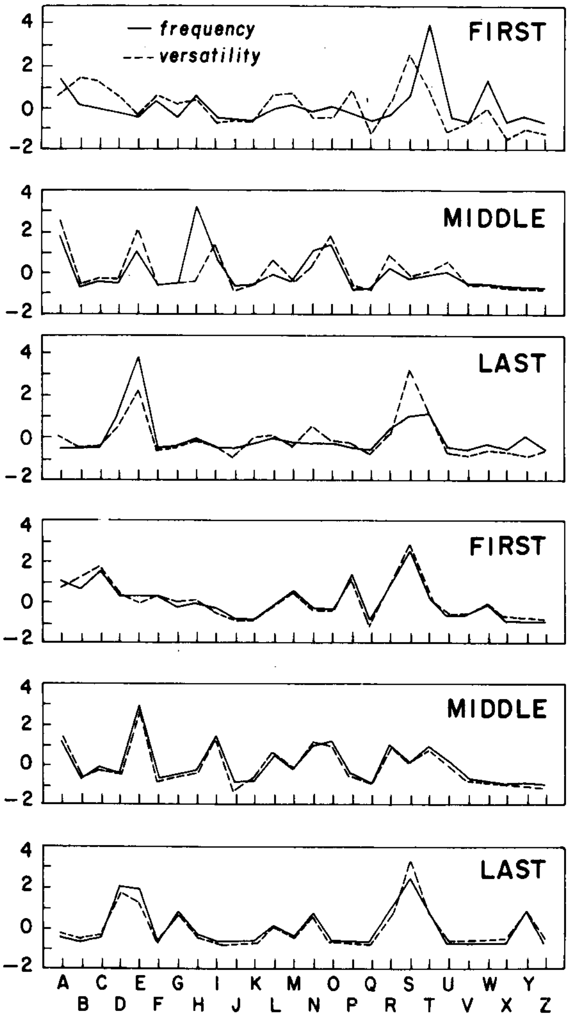

**Figure 1. Z-scores of frequency and versatility for short words (top three panels) and long words (bottom three panels.**

The average intercorrelations of the eight subjects are shown in Table 3. The level of agreement between subjects is high, but long words tend to produce weaker correlations, as do the first positions for both word lengths. Each subject's responses were correlated with the two statistical measures. The averages over subjects are presented in Table 3. In long words, there is no difference between frequency and versatility, but in short words, a difference does exist for the first and middle positions.

In summary, long words show no mismatches between frequency and versatility, although the particular identity of high-scoring letters does vary over letter positions. Short words show a similar variation in letter identity over letter positions but also a high number of frequency-versatility mismatches. These results are compatible with a short- and long-word distinction and show that the reduction of letter-position categories to first, middle, and last letter positions has preserved positional

Table 3
Average Correlations

| | | First | Middle | Last |
|---|---|---|---|---|
| | | Between Subjects | | |
| | Short | 0.72 | 0.80 | 0.81 |
| | Long | 0.67 | 0.77 | 0.78 |
| | | Between Subjects and Measures | | |
| Short Words | Frequency | 0.66 | 0.75 | 0.83 |
| | Versatility | 0.77 | 0.86 | 0.81 |
| Long Words | Frequency | 0.74 | 0.85 | 0.86 |
| | Versatility | 0.72 | 0.83 | 0.79 |

variability. The subjects show a high degree of agreement with each other, with the level of agreement at the first letter position being slightly lower. Again, short words show a difference between the frequency and versatility measures.

## DISCUSSION

More compact tables of positional letter knowledge are now available for use in working letter-recognition models. The word-length distinction is useful, especially in consideration of frequency-versatility mismatches. In short words, both the frequency and versatility of letters should be considered. The subjects' responses support this conclusion.

Some recent papers investigating the source of familiarity effect in letter recognition (e.g., Appelman & Mayzner, 1981; Wandmacher, Shapiro, & Mohr, 1981) have concluded that letter familiarity is not utilized in the identification, per se, of at least single letters or in feature extraction. These studies fail to show that letter familiarity is not used in identification when some confusion might exist, as in a multielement display, or in the localization of letters in a display. Butler (1980a, 1980b, 1981) suggested that identification and localization are independent processes but that recognition depends on both (see also Bridgeman, Lewis, Heit, & Nagle, 1979; Kent, 1981; Krumhansl & Thomas, 1976; Milner, 1974). These tables are useful, then, even in the face of such recent findings.

Shimron and Navon's (1981) study of feature information within letters indicates that all features are not equally informative. Specifically, distinctiveness and uniqueness of features is important for letter recognition. The frequency and versatility of features may also be important and may contribute to the unevenness. Using Tables 1 and 2, or tables of another letter statistic in the same format, a simple computer program could produce corresponding feature information for any feature set desired and thus supply useful knowledge bases for feature extraction.

## A COMPUTER PROGRAM TO TRANSLATE THE LETTER DATA BASE TO A FEATURE DATA BASE

A LISP program has been developed to translate the frequency and versatility values of letters from Tables 1 and 2 into similar values for features. Suppose we had an alphabet of four letters (e.g., A, B, C, and D) and a feature set of four features that describe the letters (e.g., $f_1$, $f_2$, $f_3$, and $f_4$). Table 4 shows the feature definition for each of our hypothetical letters and also shows a set of letter values for a hypothetical statistic. In Table 4, the letter A has the features $f_1$ and $f_2$, B has $f_1$ and $f_3$, and so forth for C and D. Table 4 also has values for first, middle, and last character positions of 2, 4, and 6 for short words and 3, 4, and 5 for long words (similarly for B, C, and D). In addition, Table 4 lists the values calculated by the computer program, which translates the hypothetical values for the letter statistic into corresponding values for features. The program simply sums the letter values for each letter that contains the feature being calculated. For example, $f_1$ belongs to A, B, and C. The values for $f_1$ are the sums for the A, B, and C values. If the frequency values from Table 1 had been used, the values for $f_1$ would have been 27,754, 23,520, 1,273, 16,424, 9,279, and 1,541.

Table 4
Hypothetical Example

| | Short | | | Long | | |
|---|---|---|---|---|---|---|
| | First | Middle | Last | First | Middle | Last |
| | | | Positional Values for the Letters | | | |
| A | 2 | 4 | 6 | 3 | 4 | 5 |
| B | 3 | 4 | 5 | 3 | 4 | 5 |
| C | 4 | 4 | 4 | 6 | 4 | 8 |
| D | 8 | 2 | 6 | 4 | 2 | 4 |
| | | | Positional Values for the Features | | | |
| $f_1$ | 9 | 12 | 15 | 12 | 12 | 18 |
| $f_2$ | 10 | 6 | 12 | 7 | 6 | 9 |
| $f_3$ | 11 | 6 | 11 | 7 | 6 | 9 |
| $f_4$ | 4 | 4 | 4 | 6 | 4 | 8 |

*Note*–*Feature definitions:* $A = f_1f_2$; $B = f_1f_3$; $C = f_1f_4$; $D = f_2f_3$. *Overall distinctiveness values:* $f_1 = 0.267$; $f_2 = 0.600$; $f_3 = 0.660$; $f_4 = 0.800$.

The user can select one of four feature sets supplied with the program (e.g., Briggs & Hovecar, 1975; Geyer & Dewald, 1973; Keren & Baggen, 1981; Lindsay & Norman, 1972), and either the frequency or the versatility statistic can be chosen. User-defined feature sets and statistics can be employed, but the user must produce the files in the same format as that of the default files supplied. The user can build his own alphabet as well. Positional values for the statistic selected are computed for each feature in the feature set. A nonpositional statistic is calculated by averaging the positional values, and a total is computed over word-length distinctions, one for each character position.

In addition to calculating the positional and nonpositional statistics for each feature, the program accomplishes two other goals. First, it calculates distinctiveness values (Shimron & Navon, 1981) for each feature. Second, all values calculated by the program are placed under appropriate indicators, or name tags, on property lists in the LISP environment.

The distinctiveness (Shimron & Navon, 1981) of a feature of two letters is the degree of feature overlap between two letters after the feature has been removed. The distinctiveness is, then, the overlap of the fragments. The range is from zero, or minimal distinctive value, to one, or maximal distinctive value. The program calculates the overall average distinctiveness value for each feature. For a particular feature, say $f_1$ in Table 4, the distinctiveness values would be calculated for every letter that contained $f_1$ compared with all other letters in the alphabet. For our hypothetical alphabet values would be calculated for: A compared with B, C, and D; B compared with A, C, and D; and C compared with A, B, and D. Values would not be calculated for D, since it does not contain $f_1$. The average of all these values is placed on the property list. Although the individual values for each letter comparison, and the average for each feature for each letter that it belongs to, are not retained in the LISP environment, they are available through the TRACE function in LISP as the values are being calculated. Table 4 shows the distinctiveness values calculated for the hypothetical statistic and alphabet.

For LISP modelers, the property-list feature is the real utility of the program. The presence of values on property lists leaves an enriched LISP environment ready for use by a working model of letter recognition. For LISPers, the names and descriptions of the property-list indicators follow. The indicators for letters are: (1) features—a list of the letter's component features; (2) 1-knowledge—the average, or nonpositional, statistic; (3) short—a list of the three positional statistics for the letter in short words; (4) long—a list of the three positional statistics for the letter in long words; (5) all—a list of the three positional statistics totaled for short and long words. Each feature in the feature set has the indicators 1-knowledge, short, long, and all, which contain the information calculated from the letter values under

**Table 5**
**Sample Dialogue for a Program Run**

```
.R LISP
(SYSIN 'SETUP.LSP')
(TRANSLATE NIL)


      PROGRAM TRANSLATE
      BY IAN MORRISON
      QUEENS UNIVERSITY
      KINGSTON, CANADA


THE PROGRAM TRANSLATES A LETTER DATA-BASE INTO
A FEATURE DATA-BASE.  MAKE APPROPRIATE SELECTIONS
FROM THE FOLLOWING MENUS.


CHOOSE A LETTER DATA-BASE:
  1. FREQUENCY
  2. VERSATILITY
  3. USER DEFINED STATISTIC
  9. EXAMPLE STATISTIC
9


CHOOSE ONE OF THE FOLLOWING FEATURE SETS
  1. BRIGGS & HOVECAR  1975
  2. GEYER & DEWALD    1973
  3. LINDSAY & NORMAN  1972
  4. KEREN & BAGGEN    1981
  5. USER-DEFINED FEATURE-SET
  9. EXAMPLE FEATURE-SET
9


SELECT DISTINCTIVENESS OPTION (NAVON & SHIMRON, 1981):
  1. NOT NECESSARY
  2. READ FROM FILE
  3. CALCULATE
  4. CALCULATE TRACING ROW-V
  5. CALCULATE TRACING CELL-V
  6. CALCULATE TRACING ROW-V AND CELL-V
3


SELECT PRINT OPTION:
  1. CALCULATE  ONLY
  2. CALCULATE AND PRINT TO FILE
2
```

**Table 5 Continued**
**Sample Dialogue For Program Run**

```
USE DEFAULT ALPHABET <Y/N>? N

  ENTER A LIST OF CHARACTERS:

  --- (A B C D)


FILENAME FOR OUTPUT: 'TESTO.OUT'
```

*Note–Intermediate results and file input echoed to the terminal in LISP has been omitted for clarity.*

**Table 6**
**Program Output From Example in Table 5**

```
FEATURE-BASED EXTRACTION KNOWLEDGE

NONPOSITIONAL VALUES:

(F1 13)

(F2 8)

(F3 8)

(F4 5)


POSITIONAL VALUES:

(F1 9 12 15 12 12 18)

(F2 10 6 12 7 6 9)

(F3 11 6 11 7 6 9)

(F4 4 4 4 6 4 8)


POSITIONAL VALUES FOR BOTH WORD-SIZES:

(F1 21 24 33)

(F2 17 12 21)

(F3 18 12 20)

(F4 10 8 12)


CONFUSION-SETS:

(F1 A B C)

(F2 A D)

(F3 B D)

(F4 C)


DISTINCTIVENESS VALUES:

(F1 0.2666667E-01)

(F2 0.6000000E-01)

(F3 0.6000000E-01)

(F4 0.8000000E-01)
```

those indicators. Two additional indicators are supplied for each feature: (1) confusions–a list of the letters that contain the feature; (2) distinctiveness–the overall distinctiveness value.

Table 5 shows an example of the dialogue for a program run. The example feature set and statistic are the hypothetical ones used earlier. Since the distinctiveness values need considerable processing time for evaluation, they can either be omitted on a run or accessed from a file designated by the user (i.e., they need only be calculated once). The distinctiveness values, positional and nonpositional statistics, and confusion sets for each feature can be printed to a file named by the user. Letter values are retained in the LISP environment but are not printed. Table 6 shows the output for the hypothetical feature set and statistic produced by selecting the example options.

## Availability

The program is implemented in RT-11 LISP on a PDP-11 computer (RT-11 LISP was written by Jeffrey Kodosky in 1977 and is available from the Decus Librarian for copy costs). A limitation of 5,000 words of free space demands a slow algorithm. For speedier use on a larger machine, the program can be translated into another dialect of LISP with a minimum of effort by a LISP programmer. The original version of the program was written in FRANZ LISP on a VAX 11/750. The author will supply a hardcopy listing on request and/or a copy to an 8-in. floppy diskette if supplied. Disks will be returned in double-density format from an RX02 drive.

## REFERENCES

Appelman, I. B., & Mayzner, M. S. The letter-frequency effect and the generality of familiarity effects on perception. *Perception & Psychophysics*, 1981, **30**, 436-446.

Bridgeman, B., Lewis, S., Heit, G., & Nagle, M. Relation between cognitive and motor-oriented systems of visual position perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, **5**, 692-700.

Briggs, R., & Hovecar, D. A new distinctive feature theory for upper case letters. *Journal of General Psychology*, 1975, **93**, 87-93.

Butler, B. E. Selective attention and stimulus localization in visual perception. *Canadian Journal of Psychology*, 1980, **34**, 119-133. (a)

Butler, B. E. The category effect in visual search: Identification versus localization factors. *Canadian Journal of Psychology*, 1980, **34**, 238-247. (b)

Butler, B. E. Identification and localization in tachistoscopic recognition: The effects of data and resource limitation. *Canadian Journal of Psychology*, 1981, **35**, 36-51.

Drewnowski, A. Detection errors on the word *the*: Evidence for the acquisition of reading levels. *Memory & Cognition*, 1978, **6**, 403-409.

Drewnowski, A. Missing -ing in reading: Developmental changes in reading units. *Journal of Experimental Child Psychology*, 1981, **31**, 154-168.

Drewnowski, A., & Healy, A. F. Detection errors on *the* and *and*: Evidence for reading units larger than the word. *Memory & Cognition*, 1977, **5**, 636-647.

Drewnowski, A., & Healy, A. F. Missing -ing in reading: Letter detection errors on word endings. *Journal of Verbal Learning and Verbal Behavior*, 1980, **19**, 247-262.

Geyer, L., & Dewald, C. Feature lists and confusion matrices. *Perception & Psychophysics*, 1973, **14**, 471-482.

Healy, A. F. Detection errors on the word *the*: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, **2**, 235-242.

Healy, A. F. Proofreading errors on the word *the*: New evidence on reading units. *Journal of Experimental Psychology: Human Perception and Performance*, 1980, **6**, 45-57.

Healy, A. F., & Drewnowski, A. Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception and Performance*, 1983, **9**, 413-426.

Kent, W. *The brains of men and machines*. New York: McGraw-Hill, 1981.

Keren, G., & Baggen, S. Recognition models of alphanumeric characters. *Perception & Psychophysics*, 1981, **29**, 234-246.

Krumhansl, C. L., & Thomas, E. A. C. Extracting identity and location information from briefly presented letter arrays. *Perception & Psychophysics*, 1976, **20**, 243-258.

Kučera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, R.I: Brown University Press, 1967.

Lindsay, P., & Norman, D. *Human information processing: An introduction to psychology*. New York: Academic Press, 1972.

Mayzner, M. S., & Tresselt, M. E. Tables of single-letter and diagram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements*, 1965, **1**, 13-32.

Milner, P. A model for visual shape recognition. *Psychological Review*, 1974, **81**, 521-535.

Rawlinson, G. E. Bigram frequency counts and anagram lists. *Quarterly Journal of Experimental Psychology*, 1976, **28**, 125-142.

Shimron, J., & Navon, D. The distribution of information within letters. *Perception & Psychophysics*, 1981, **30**, 483-491.

Solso, R. L. Positional frequency and versatility of letters for six-, seven-, and eight-letter English words. *Behavior Research Methods & Instrumentation*, 1979, **11**, 355-358.

Solso, R. L., & King, J. F. Frequency and versatility of letters in the English language. *Behavior Research Methods & Instrumentation*, 1976, **8**, 283-286.

Wandmacher, J., Shapiro, R., & Mohr, W. Letter familiarity does not aid feature extraction. *Psychological Research*, 1981, **43**, 33-43.

*Webster's New Collegiate Dictionary*. Springfield, Mass: Merriam, 1976.

(Manuscript received October 25, 1982;
revision accepted for publication July 7, 1983.)