Chapter 6

Error Patterns in Novice and Skilled Transcription Typing

Jonathan T. Grudin

Errors have long been viewed as an important source of insight into the organization underlying performance. In this study, the general patterns of errors made by novice and expert typists suggest how skill in this complex motor task is organized and developed.

Early accounts of typing errors were largely descriptive (e.g., Lessenberry, 1928; Dvorak, Merrick, Dealey, & Ford, 1936). Lashley's (1951) suggestion that they are a potentially valuable source for inferring the processes in skilled performance was picked up in the 1960s and thereafter by MacNeilage (1964), Shaffer and Hardwick (1968, 1969), Long (1976), and Rabbitt (1978). Sophisticated process models for typing have been proposed (e.g., Shaffer, 1978; Sternberg, Monsell, Knoll, & Wright, 1978; Terzuolo & Viviani, 1980; Rumelhart & Norman, 1982). The increased availability of computer and video systems, which are particularly suited for analysis of typewriting, makes it possible to correct and extend previous accounts. This study goes beyond initial descriptive categorizations to suggest functional classifications that support the divisions of Chapter 2 and to support constraints on a model of typing explored elsewhere (Grudin, 1981).

Lessenberry (1928) compiled letter confusion matrices in which 60,000 typing errors are categorized according to the letter intended and the latter actually struck. I have extended the analyses of Lessenberry's data and compiled two additional confusion matrices to allow a more detailed comparison of novice and expert performance.

One of my tables was constructed from all substitution errors found in a large corpus of text transcribed by expert typists. The second consisted of all substitution errors from a practice exercise by about 70 beginning high school typists. These confusion matrices constitute the Appendix to this chapter.

The Lessenberry Confusion Matrix

Important contextual information is lost when only the key intended and the key actually struck are considered. As Dvorak et al. (1936) mentioned, the context usually supplies information necessary for determining the cause of a given error. Nevertheless, some patterns emerge from Lessenberry's large corpus.

The correct character was replaced by a character immediately adjacent and in the same row in 43% of the errors. (These are referred to as row errors.) Substitutions of a neighboring letter from the same column (column errors) accounted for 15%.

Most common after neighboring letters is the substitution of the homologous (mirror-image) letter, typed by the same finger in the same position but the wrong hand. This error is typically the second or third most frequent substitution. It accounts for 10% of the errors overall, even though several letters have no corresponding homologous letter. (The Lessenberry data were restricted to letters, with no information on substitutions involving punctuation.)

Note that the keyboard is not quite symmetric. Because of mechanical constraints on early typewriter design, the vertical columns are actually on a diagonal. This complicates the determination of homology for keys in the bottom row. It also permits us to contrast purely spatial symmetries with "movement symmetries."

For example, relative to the positioning of the hands at the keyboard, the letters *m* and *c* are in homologous positions, but the *m* is typed by the index finger and the *c* is typed by the middle finger. There were relatively few substitutions of *m* for *c* or *c* for *m*.

At What Level Do Homologous Errors Occur?

The confusion that leads to a homologous intrusion could conceivably occur at any of a number of levels: (a) in the selection of the motor program (the set of commands to muscle groups); (b) in the specification of the hand, finger, and finger position that determine the key to be typed; (c) within a more abstract representation of the keyboard (e.g., a spatial representation). These are clarified below.

A confusion of motor programs would result from a possible association between symmetric movements. The special relationship between symmetric motions is manifest in the relative difficulty of making different motions with the hands—it is more difficult to pat your head and rub your stomach simultaneously than to pat or rub both.

A confusion at the "movement component" level could occur if keys are at some point specified in terms of hand, finger, and finger position, and if one of these components, in this case hand, is specified incorrectly. This differs from a confusion at the motor program level in that, for example, it could lead to the specification of three components not typically associated, such as "right middle finger down and inward." Because the middle finger does not normally make such a movement, presumably no motor program has been formed for that particular movement, so such a confusion could not occur at the motor program level.

A confusion at the abstract representational level can be pictured by specifying keys with Cartesian coordinates on a grid that has one axis down the middle of the keyboard. A homology would occur if the sign of the x coordinate were reversed.

I argue that (a) is responsible for some, although not all, of the errors, and that (b) probably accounts for most. Although (c) could be made to account for some, there is no evidence that requires positing such a level of abstraction.

First, consider (a). If learned motor programs for letters are likely to be confused when they result in mirror-image or homologous movements, the failure to get *m-c* homologies is explained—there are no learned mirror-image programs for these letters. A down-and-inward movement exists for the middle finger of the left hand and results in typing *c*, but there is no learned down-and-inward pattern for the right middle finger. So no confusion occurs. This also explains the *v-n* homology, but it runs into trouble with the relatively numerous *v-m* confusions. The motor programs for these should be quite different, one being inward and the other outward. There is a similar problem in explaining the *b-n* confusions, which outnumber even the *v-n* substitutions.

Explanation (b), a confusion of movement components, can explain the *v-m* and *b-n* confusions if we assume finger position specifications like "down" for *m* and *v*, and "down and inward" for *n* and *b*. This also explains the absence of *m-c* substitutions (different finger assignments). But it does not predict *n-v* homologies, which are even more common than those of *m-v*. Since *n-v* are motorically homologous, (b) and (a) together cover all of the errors.

The failure to find an *m-c* homology would eliminate (c), the confusion at the level of an abstract mental representation, if the typist's representation of the keyboard were an undistorted version of the actual keyboard. But imagine a representation in which the rows have been aligned, as in Figure 6.1. This alignment is a natural one, in that the keys in each column are typed by the same finger. In fact, when I gave skilled typists a set of loose keys and asked them to arrange them as on a keyboard, they invariably produced this pattern. Given this distortion, (c) is effectively indistinguishable from (b). (Below I show that it is also necessary to increase vertical separations relative to horizontal separations and to place a larger separation between the center alphabetic keys than between other keys to make the abstract representation useful. None of these modifications is really implausible, but such a representation is not needed to account for any data.)

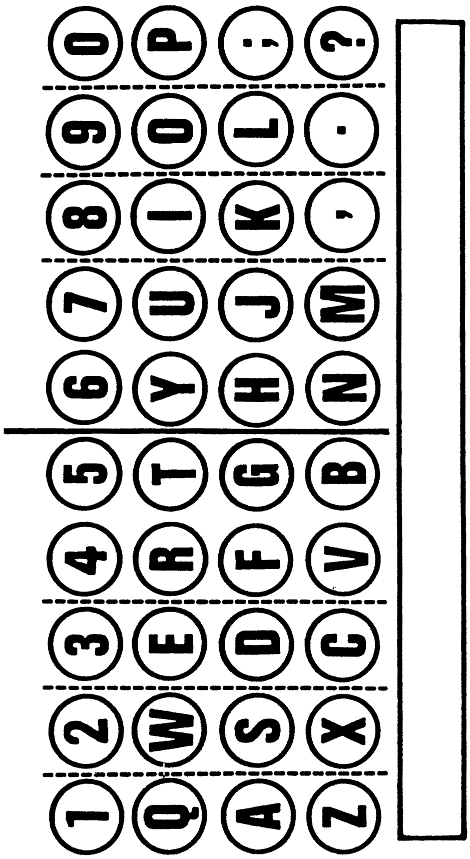

Fig. 6.1. A symmetric alignment of the keyboard.

Letter Frequency Effects

Consider the symmetry of homologous errors. Is a typist more likely to replace a left-hand letter with its right-hand homologue than to replace a right-hand letter with its left-hand homologue? Is a person more likely to replace a lower frequency letter with its higher frequency homologue than a higher frequency letter with its lower frequency counterpart? I found no effect of hand, but a strong effect of frequency. For example, *d* has a higher frequency than its homologue *k*. The letter *d* was struck for *k* 484 times, whereas *k* was struck for *d* only 287 times. However, the absolute frequency of error is not the best test. For example, although *o* is more frequent than *w*, *o* was substituted for *w* only 104 times while *w* was struck for *o* 122 times. This is misleading, because the typists had fewer opportunities to err on *w*, since *w* was encountered less often. We need to control for letter frequency; that is, to examine the *proportion* of the occurrences of a letter that leads to a substitution. To allow such controlled comparisons, I constructed a normalized confusion matrix for each raw data matrix (see Appendix). After normalization for letter frequency, it is always the case that higher frequency letters are more likely to replace lower frequency letters in homologous substitutions.

Summary of Lessenberry Data Analyses

The most common error is the striking of a key immediately adjacent, either horizontally or vertically, to the intended key. Also highly frequent is the striking of the homologous or mirror-image key. The best explanation for most of the homologous errors is that the representation of a keystroke includes a specification of the hand to be used, and that an error in this specification leads to the homologous intrusion. Finally, a higher frequency letter is more likely to be typed for a given lower frequency homologous letter than vice versa.

Although the size of the corpus makes Lessenberry's data useful, we know little of the circumstances in which it was gathered. Russon and Wanous (1973, p. 205) describe the typists as "students." For this reason, I constructed confusion matrices from the errors of novice and expert typists to corroborate and extend the forgoing analyses.

Novice and Expert Confusion Matrices

Method

Six professional typists transcribed magazine articles totaling approximately 60,000 characters. In some sessions, typing was done on a computer keyboard with which the typists were familiar; in the others, on a Microswitch keyboard designed to look and feel identical to an IBM Selectric typewriter keyboard. The text was presented as double-spaced typed copy on individual sheets of paper. After a 10-minute warm-up with another text, the typists were given an article and asked to transcribe it. They were told not to worry about errors and to type for speed. Keypresses and the corresponding times were recorded by a microcomputer.

In addition, novice typists were recruited from the beginning typing classes at a local high school. Eight students spent 1 hour each week transcribing text on the Microswitch keyboard, using the procedure just described. In addition, class papers were collected 6 weeks into the class, when the students had been acquainted with all letters of the keyboard for 2 weeks, and a three-paragraph exercise was scored for errors.

Some sessions of novice and expert typing were videotaped with a rotary shutter camera aimed down at the keyboard from above. Two views of the fingers were obtained by placing a mirror behind the keyboard at a 45-degree angle. By forming its image in less than 2 msec, the rotary shutter yields an image quite free of blur. The video fields were serially numbered with an electronic video counter and analyzed by using a Sony video motion analyzer.

I undertook separate analyses of novice and expert data. Table 6.1 gives the typing speeds, error rates, and a categorization by purely descriptive error type for each skilled typist and for the novices as a group. (The typing speed for students is a good approximation; all rates are based on a five-character word and make no adjustment for errors.)

Substitution, insertion, and omission errors all refer to single-letter errors in otherwise correctly typed words. Many of the miscellaneous ("other") novice errors are, in fact, words in which two substitutions appear to have been made. Thus, substitution errors completely dominate the errors of students. These are substantially reduced in experts, where they are the second most common error. Insertion errors, the most frequent for experts, are not necessarily the most interesting: The overwhelming majority of a randomly selected subset examined on videotape

Table 6.1. Typing Speeds and Error Rates

Typist	Speed (wpm)	Error Rate (%)	Error Type (% of all errors)				
			Ins.	Omi.	Sub.	Tra.	Oth.
1	90.4	1.1	28	21	27	10	13
2	65.6	0.5	17	11	69	3	0
3	76.1	1.9	75	5	6	4	10
4	74.9	1.0	47	15	4	4	29
5	61.3	0.4	53	19	9	11	9
6	81.9	0.8	35	12	21	11	20
Novices	20	3.2	9	4	75	4	11

Note. Error types listed are Ins., insertion; Omi., omission; Sub., substitution; Tra., transposition, Oth., other.

are misstrokes, two keys struck by one finger. For this reason, only substitution errors were used to generate confusion matrices. These included some 3,300 and 500 substitutions for novices and experts, respectively.

Results

Immediately adjacent keys of the same row accounted for 59% of the novice substitutions and 31% for the experts (compared with 43% for Lessenberry). Errors of the same finger in the same column were 8% of novice substitutions, 16% of experts, and 15% in Lessenberry's data. The proportion of substitutions that fit the description of homologous error were 4% for the experts and 16% for the novices. Lessenberry's data showed 10% of this type. Possibly Lessenberry examined intermediate typists or, more likely, a range of skill levels.

Homologous Substitutions

Chance level for producing homologous errors by random substitution of letters is about 3%. However, substitution errors are not random. Seventy-eight percent of skilled typist substitutions and 72% of novice typist substitutions are within hand. If we restrict the analysis to 2H errors, homologies account for 17% of expert and 62% of novice errors (chance being 7%). Thus, there may be homologous substitution by skilled typists, although infrequently compared with novices. (See also Chapter 10 for further evidence of homologous errors in typing.)

Lessenberry's data were for letters only, leaving open the question whether novices homologously confuse letters and punctuation—the *a* and the semicolon, for example, or the *c* and comma. In my corpus there is no example of such an error. Punctuation is struck for adjacent letters —the semicolon for the *l*, the comma for the *m*—but never for the homologous key. Possible reasons for this are discussed in the next section.

Homologous errors by novices can be found in sequence or with an intervening key typed correctly, as in *learn* → *siarn* and *think* → *thend.*

Twenty-three homologous errors by eight novices were examined on videotape. In 18 cases only one finger moved. Twice, fingers moved toward the homologous keys simultaneously. In three cases, the correct finger moved to the correct key, withdrew, and then the error was made. This indicates that the error is often, but not always, made early, more likely in the specification of the components of action than at the level of the motor program.

Frequency Effects in Homologous Errors

As we have seen, Lessenberry's data show a frequency effect. For each homologous letter pair, the higher frequency letter is more likely to intrude in place of its mirror-image than vice versa. My study of novices confirmed the effect. Nine of the 10 pairs showed the pattern. The set of substitution errors by skilled typists contained only 19 homologous mistakes and did not show the same pattern, with 10 intrusions by higher frequency keys and nine by lower frequency keys.

Some of the homologous errors in their context indicate that multicharacter response units may be represented during performance. In particular, homologous errors rarely create low-frequency digraphs. For example, the letter *k* appeared in three words in the novice typing exercise: *stroking*, *think*, and *know*. Students mistyped it as *d* 46 times. In the first two words, the substitution of a *d* produces no unusual letter combinations. These accounted for 45 of the 46 homologous errors. Only once did a typist type *dnow*, which includes the unusual digraph *dn*. Similarly, the word *sequences* was typed on different occasions with homologous errors in the second, third, fifth, and sixth positions, but not in the fourth, which would have produced the illegal digraph *qr*. In fact, substituting *q* for *p*, which would usually form an illegal digraph, is proportionally the rarest homologous error in both Lessenberry's corpus and my own.

The effect of letter frequency—that in the more common novice errors, higher frequency letters are more likely to intrude—may explain the failure to find homologous errors involving punctuation keys. Both punctuation keys and the letters in positions homologous to them are low in frequency (with one exception). Low-frequency keys, by definition, seldom occur, so there is little chance for substitution. The exception, *a*, is homologous to the semicolon, which has extremely low frequency. Thus, semicolon is unlikely to replace *a*, and there are few opportunities for *a* to replace semicolon. Of course, other factors may be at work. For example, most homologies involving punctuation would create low- or zero-frequency digraphs, which, as we have seen, rarely occur in substitution errors.

Adjacent-Letter Substitutions

Most substitution errors in both unskilled and skilled typing occur when an immediate neighbor of the target key is struck in its place. Every researcher investigating typing errors has noted the prevalence of these errors, particularly for horizontally adjacent keys. It is reasonable to suppose that these represent aiming or trajectory errors. On videotapes of skilled and novice typists, I located instances of substitution by a horizontally adjacent letter. I restricted the set to those in which the two keys would normally be typed by different fingers (e.g., *small* → *smsll*, *each* → *wach*, *golf* → *gold*). In each case the intended and typed keys are adjacent on the keyboard but are struck by different fingers.

In these substitutions the question is, Which finger strikes the key? For 22 of the 25 skilled typist errors and 42 of the 44 novice errors examined on videotape, the key was struck by the finger that usually strikes it. The errors could not be attributed to errant finger trajectories.

Thus, typists are more accurate in the execution of keystrokes than might have been supposed, but occasionally err in their specification of finger. This is the more interesting explanation, since it indicates that the finger to be used is explicitly represented during execution.

Column substitutions are a moderate source of error, accounting for 8% of novice and 16% of expert errors. Analysis of the videotapes to determine whether a finger strikes a key squarely or not is more difficult than determining which finger strikes a key, but in most cases it is clear. Of the 14 examples examined, in 8 cases there was a clean motion to the wrong key; 3 were misstrokes, landing between the keys; and the remaining 3 were difficult to judge. Thus it is likely that most vertical errors are also specification errors.

Immediate neighbors are much more likely to be substituted than distant letters in the same row or column. The data in Table 6.2 support this observation. Each number represents the median number of substitution errors within the pairs in the specified category. For example, the top row indicates that for row substitution errors involving one finger (1F) and a distance between the correct and struck letter of one (i.e., they are immediately adjacent), Lessenberry's (normalized) data show the median number of substitutions to be 1,521. This would be such pairs as *r–t*.

The next three rows of the table indicate that the likelihood of substitution falls off quickly with distance across the keyboard. These rows show pairs typed by two fingers of the same hand (2F) with separations of

Table 6.2. Substitution Errors: Median Number for Pairs in Each Category

Error type	H/F	Distance	Lessenberry	Novice	Expert
Row substitution errors	1F	1	1521	145	10
	2F	1	974	60	9
	2F	2	149	4	0
	2F	3	60	1	1
	2H	1	562	21	3
	2H	2	81	6	1
	2H	3	122	3	1
Column substitution errors	1F	1	623	18	5
	1F	2	152	1	0
Diagonal substitution errors	1F	1	232	13	2
	2F	1	163	1	0
	2H	1	91	1	0

one, two, and three letters. Examples of these would be *a–s*, *a–d*, and *a–f*, respectively. The column errors show a similar effect of proximity.

Across-hand (2H) and diagonally adjacent pairs are included for comparison. For example, a 2H error of distance 1 is a substitution by an immediately adjacent key typed by the other hand, such as a *t* for a *y*. For 2H substitution errors of distances 1 and 3, the numbers are high because several pairs are homologous. Diagonal errors are notably fewer than either row or column adjacent errors, even in 1F diagonal confusions such as *f–t*, where hand and finger are constant. Perhaps diagonality is not a position equivalent to the vertical or horizontal. Alternatively, the paucity of errors could arise from the lack of diagonal movements of the other fingers, and a consequent reduced probability of a confusion in this specification.

Given the relatively few 2H errors, physical proximity alone is not adequate to explain the predominance of substitutions of immediately adjacent neighbors. Neighboring fingers share musculature, and postural compensations for finger movements may be similar for neighboring fingers. As was argued earlier from the pattern of homologous errors, the confusion probably occurs at the level of the movement components of hand, finger, and finger position, or at the lower level of the motor program itself.

Frequency Effects in Adjacent Errors

As with homologous errors, there are large frequency effects in adjacent errors in Lessenberry's data. Once again, after normalizing for frequency in the language, a typist is more likely to substitute a higher frequency letter for a neighboring low-frequency letter than vice versa. This asymmetry held for every pair of row adjacent keys and 15 of 16 pairs of column adjacent keys.

We are now in a position to use the disproportionate likelihood of immediately adjacent letter intrusions to explain why MacNeilage (1964) found a preference for substituting home-row keys. Home-row keys have two adjacent vertical neighbors, whereas keys on the upper and lower rows have only one. Thus, there is no intrinsic preference for the home row. The perference is for higher frequency letters. (Other data of MacNeilage can be explained by this frequency effect as well.)

The frequency effect was confirmed in my novice study, with 31 row or column adjacent pairs favoring the higher frequency letter and 7 pairs favoring the lower frequency letter: chi-square(1) = 22.04, $p < .01$. Because the skilled typists made proportionally fewer row and column errors, as well as fewer errors overall, their data are noisier, with many empty cells. Of those pairs with one or more substitutions in each direction, 18 favor the higher frequency letter (after normalization) and only 6 favor the lower frequency letter, a significant difference: chi-square(1) = 5.04, $p < .025$.

We can eliminate the possibility that the frequency effects result from the typist's more careful scrutiny of relatively unfamiliar low-frequency keys during performance. Such scrutiny could lead to early detection of potential errors in which low-frequency letters are about to be typed. However, the relatively high proportion of substitutions within pairs such as *z–x* and *j–k* suggests that low-frequency letters do not get particularly careful inspection before they are typed. Thus it is more likely that intrusion by the higher frequency letter brings about the substitution error.

Permutation Errors

Novices and experts show different patterns of letter transpositions. Experts average around 80% across-hand transpositions, and almost always exchange two successive letters from the text. Only twice in the corpus did a letter migrate across more than one position (*deliberately* → *deliberatyel*, *outweigh* → *ouweight*), and only once did letter components appear to switch (*simple* → *simo;e*). Although a world champion typist (Owen, 1919) reported that she eventually began to transpose words rather than letters, our typists did this only twice. More common were "interchanges," the switching of two separated letters, such as *big* → *gib*, *figuring* → *firuging*, and the more complicated examples *denomination* → *demonimation* and *and more* → *amd nore*. As the examples indicate, these were almost all within-finger errors.

Novices are almost as likely to make within-hand as across-hand transpositions (40% vs. 60%). They are more likely than experts to transpose homologous letters (25% vs. 12% of 2H transpositions). Novices, more than the experts, move a key past two intervening keys (e.g., *sequences* → *sequeences*, *lower* → *leowr*). Every novice permutation falls into one of three categories: transpositions, interchanges, and migrations across two positions. (These constitute fewer than half of the theoretically possible permutations.)

The data suggest that migrations and most transpositions involve one mechanism and interchanges a different mechanism. Most transpositions involve nonhomologous 2H letter sequences. In contrast, interchanges (e.g., *major* → *jamor*, *also* → *aosl*) by experts and novices are generally 1F: The two keys involved (but not the intervening key) are typed by the same finger. Hand and finger, two of the three principal components determining a keystroke, are shared. Most of the remaining interchanges are 2F and involve row adjacent keys, and again two components—hand and position—are shared. Finally, of the few 2H interchanges, most involve homologous keys, sharing finger and position.

For 96 of 163 transpositions by experts, the result of the error is to place a lower frequency letter in front of a higher frequency letter. The effect is significant chi-square(1) = 5.52, $p < .02$. Five typists are in accord, with Typist 4 running strongly counter to the others: 18 of her 25 errors move a higher-frequency letter into the first position.

In a videotape study of 66 transposition errors, I found that for both novices and experts, the finger that made the first motion toward the key was usually the finger that first struck a key. Thus, the reversal may typically occur early in execution.

Learning. To determine the effects of practice, I examined a sample of novice typing taken 1 month later. In this extra month, the novices had 67% more experience with typing and up to three times the experience with the last keys introduced in the last class.

As expected, performance was generally closer to but not equal to expert performance. There are fewer errors overall. Of all transpositions, 28% are within hand. Homologous reversals account for 29% of 2H transpositions, which is still more than skilled typists. The misplacement of a letter by more than one position is rarer than before, bringing the students in line with expert performance in this regard. Interchanges still occur, and as before usually involve keys sharing two of the three movement components of hand, finger, and finger position.

Other Errors

Norman (1981) uses the term *activation error* to describe an error in which a similar but more common or more recent performance is substituted for the behavior one had intended. A number of errors made by skilled typists seem to be of this sort. In the error *Even experts* → *Even Experts* the (space)-*e* sequence in *experts* may have reactivated the (space)-(shift)-*e* sequence. Another example is the mistyping of "chew everything carefully, never gulp" into "chew everything carefully, nevery gulp."

Substitution errors not yet accounted for include a large number of vowel-for-vowel substitutions (even after normalizing for their high frequency of occurrence), and confusions of the letters *c, g, s, t, w*. Both may represent an activation of a multiletter response unit, perhaps one recently active, by the presence of component letters. In the case of the five consonants, the units could be digraphs ending with the letter *h*. Thus, most substitutions of the five consonants (except *s* with *w*, those being adjacent keys) are in such cases as *Ruth* → *Ruch, Rugh; three* → *chree; show* → *whow; check* → *sheck.*

Misstrokes, which are not a major source of substitution errors, are more evident in instances where an extra letter is inserted into the text—a finger strikes two keys simultaneously. Well over half the insertion errors by experts are potentially such errors, and most examined on videotape are two keys struck by the same finger. In other cases, the finger adjacent to the key being struck moves along with it and hits a key, causing an insertion. These two mechanisms account for all the examined insertions by Typist 3, who alone accounts for over half of all insertion errors by skilled typists. One skilled typist (Typist 5), whose retraction from striking the space bar normally causes her right middle finger to skim close to the *k*, actually strikes it occasionally—but only when the letter about to be typed is a *p*: The motion toward the *p* on top of the retraction from the space bar causes the middle finger to hit the key. Her only insertions of the letter *k* thus come before words beginning with *p*.

Most insertion errors of the misstroke description occur when an index finger is reaching inward or diagonally for one of the six center keys. Many insertions that do not appear to be misstrokes consist of typing a letter that appears elsewhere in the word, usually later—for example, *fiber* → *bfiber*, *crash* → *cracsh*. However, this is true for only two of the six typists. These two typists are shown elsewhere (Gentner, Note 1) to have particularly independent finger movements, leading to more overlapping finger motion. There are several possible explanations. With a number of fingers approaching keys simultaneously, anticipatory keystrokes might be more likely. Also, these errors could be transpositions, interchanges, or migrations that were subsequently corrected. (Examples are *notion* → *ntotion*, *ravenous* → *vravenous*.)

Errors of omission by skilled typists follow the general serial position pattern described by MacNeilage (1964) for errors of this type: They are rare in the first-letter position and most common in the next few positions. Videotape analysis indicates that for approximately half there is no motion toward the omitted key. When there is motion toward the key, it varies in its degree of completion: At times the finger seems to contact the key, while at other times the finger simply moves over the key but never strikes at it. We note in Gentner, Grudin, and Conway (Note 2) that there are often two clearly delineated parts to a keystroke: the motion toward the key and a rapid downstroke. On still other occasions, the finger positions itself over the key and makes a very weak thrust in the direction of the key, coming nowhere near it. I see no pattern to these different responses.

Omitted letters are likely to appear, typed correctly, in the word preceding or following the word being typed, or elsewhere in that word itself. (For example, three typists omitted the third *i* in *artificial*.) This was true for over 60% of all omissions. In the video study the letter is significantly more likely to precede the omission; in another study the letter is almost equally likely to follow the omission. Omissions of one of a double-letter pair occurred only 17 times, but this is 20% over chance based on the percentage of doubles in the text.

The low incidence of omissions in the first-letter position in a word suggests that that letter is particularly strongly activated, and for that reason possibly subject to less noise. Therefore, I looked particularly carefully at the 31 omissions that did occur in the first position. In 42% of them the omitted letter was also one of the three preceding letters, usually the immediately preceding letter, as when *the entire* was typed *the ntire* or *keep putting* was typed *keep utting*.

In many of the remaining omissions, the finger previously used was the finger that should have typed the omitted letter. For example, three words dropped an initial *p* or *l* when that letter was immediately preceded by a carriage return, which uses the same finger. This suggests that the "deactivation" of a letter for motor program following a keystroke has, in these cases, interfered with the typing of a subsequent key.

There were a small number of doubling errors, in which the wrong letter is doubled (e.g., *well* → *weel*), and just one error fitting the description of an alternation error (*where* → *whrer*), by Typist 3. Of the 124 alternating sequences in the text, *ere* is among those typed most quickly by Typist 3, despite being within hand.

Lashley (1951), Shaffer and Hardwick (1968), and Rumelhart and Norman (1982) argue that doubling and alternation errors indicate the use of special markers for such sequences—when the marker is applied to the wrong letter, an error occurs. The omission errors suggest why such special measures may be necessary—without them, the deactivation process following a keystroke would interfere with the quick retyping of the same key.

The Development of Skilled Typing

In this section, I summarize the results of the investigations of typing errors and discuss the implications as they bear on three developmental changes in error patterns: (a) the disappearance of homologous errors with the acquisition of skill; (b) the reduction in the proportion of adjacent substitution errors; (c) the marked increase in the percentage of across-hand transpositions.

The major categories of substitution error are row, column, and homologous errors. In most cases, the error is due to a deliberate stroke by the finger appropriate for the key actually struck, with no motion toward the correct key. Thus, these errors are best explained as occurring prior to the active involvement of the motor program, when the keystroke is specified in terms of hand, finger, and finger position. For touch typists, this mapping of finger to key is particularly orderly. In those errors where two fingers are in motion simultaneously, the confusion may have occurred later, possibly among motor programs.

Although novices make more errors than experts, their errors are orderly. The majority are substitution errors, of which 75% are substitutions of immediately adjacent keys. Fifty-one percent of the remainder are homologous errors. Experts make proportionally fewer substitution errors, and fewer of them are adjacent keys. Skilled typists make very few homologous errors.

Evidence for multicharacter response units in skilled typing, digraph units in particular, is presented elsewhere (Grudin, 1981). These units may help optimize postural and positional adjustments across a series of movements. As digraphs are relied on more heavily, substitutions based primarily on errors in hand, finger, and finger position specification may decline, since they often produce sequences of low or zero frequency; and movement components governing hand, wrist, and arm could come into play with the development of digraph response units. The greatest violation of such global preparations is a keystroke by the wrong hand, so homologous errors drop off quickly. Vertical movements require wrist and arm motion, so column errors would conflict with such preparation more than row errors. Among row errors, those of the same finger are most compatible, those of adjacent fingers reasonably compatible, but those of distant fingers might require different postural adjustments, and thus be less likely to occur. This parallels the distribution of these errors.

The pattern of substitution errors is marked by strong frequency effects. A higher frequency letter is more likely to substitute for a lower frequency neighbor or homologue than vice versa. This appears to result from an intrusion of the more common letter, which in an activation model could yield to either a recency or frequency explanation. A long-lasting residual activation could follow the typing of a key; keys active more recently or more often would have more residual activation. Alternatively, the activations for higher frequency elements could have higher resting levels or lower thresholds for initiating action.

Residual effects of deactivation are indicated by the pattern of omission errors. A letter is more likely to be omitted if the same letter was recently active (and thus recently deactivated). This suggests that deactivation may be a source of variability in typing: When a letter is insufficiently deactivated, a substitution error may follow, whereas when a letter is too strongly deactivated, the same letter appearing soon afterward may fail to be typed.

In addition, omissions are strongly influenced by serial position within the word being typed, with initial letters least likely to be omitted and medial letters most likely. This matches other determinations of the relative strengths of letters, and suggests an initial profile of letter activation.

For novices and experts, all errors in which the correct letters are permuted result from either the misplacement of a single letter or the switching of two letters. The great majority of these are either transpositions or interchanges, with the former more frequent. Interchanged letters almost always share components, most often being two letters typed by the same hand and finger, while transpositions are typically typed by different hands. A possible reason is that transposed letters are part of a multicharacter response unit, whereas interchanged letters belong to different response units.

Elsewhere I argue from the timing of transposition errors the presence of centrally issued "trigger" pulses for the two letters involved (Grudin, 1981). Thus, within a unit, subordinate letters may be transposed because keystroke times have little flexibility. If one finger is out of position, another finger might come in early. Transpositions are likely to be 2H for several reasons: In such a sequence, a movement or mispositioning of one hand can influence one finger while not affecting the other; the second letter has more freedom to reach its key early if it is on a different hand; interstroke intervals are shorter for 2H sequences, and the activation levels guiding the keystrokes may accordingly be more equal. If interchanges, by contrast, involve two response units, they are free of these timing constraints, and are more likely to be affected by response similarities of the letters involved.

The different patterns of transposition errors in the typing of novices and experts may be due to the greater reliance of skilled typists on multicharacter response units. Novices transpose two letters typed by the same hand twice as often as experts. This is consistent with the argument that the prevalence of 2H transpositions in skilled typing is due to constraints on timing within multicharacter sequences (Grudin, 1981). Some novice transpositions may be interchanges involving single-character response units, interchanges with no intervening letter. Just as interchanges involve keys sharing components, 69% of 1H novice transpositions are horizontally or vertically adjacent letters, and 25% of the 2H cases are reversals of homologous keys. Thus, 41% of novice (and only 16% of expert) transpositions share two of the hand, finger, and finger position specifications.

The remaining novice 2H transpositions may, in fact, represent multicharacter sequences already being learned. Over half of them are in function words—*the*, *that*, *than*, *for*, *to*, and *and*. The average length of a word containing a 2H transposition is under four letters, while the length of a word containing a 1H transposition averages over six letters. A similar effect holds for our experts. Words containing 2H transpositions average 6.1 letters while words with 1H transpositions average 7.8 letters. This significant difference results from the absence of short words containing 1H transpositions: there are 21 across-hand and no 1H errors in words of two and three letters (there are many 1H function words, such as *are*, *as*, *at*, *be*, *in*, *on*, and *was*). Short words may be executed as units, less susceptible to the errors based on shared components (which includes most 1H errors) that I associate with errors across response units. This analysis also suggests that longer words are not executed as units.

Are these multicharacter response units "syllables"? Shaffer (1975) reported that most transpositions occur within a syllable. This was true for 91% of our transpositions of letters. However, 87% of all letter-letter transitions are within syllable in the text. The difference is not significant, so transpositions provide no evidence for syllable representation in typing.

There is further evidence that novices have begun abstracting patterns. The special treatment of double letters, inferred by Lashley (1951), Shaffer and Hardwick (1968), and Rumelhart and Norman

(1982) from the existence of errors such as *ill* → *iil*, is indicated by such novice errors as *speed* → *spped*, *spiid*, and *letter* → *leteer*, *lettee*, *lettrr*. But novices also produce errors such as *speed* → *spede* and *letter* → *lerter*, which suggest that they do not always handle double letters as units.

When novices move a letter past two intervening letters, as in *that* → *atht*, the skipped letters are usually digraphs with high frequency or high transitional probability—*th*, *or*, *at*, *qu*. This raises the question why I found no transpositions of multicharacter sequences in skilled typing. Two-letter insertions, omissions, and even substitutions occur (though much less frequently than single-letter errors), but not multicharacter transpositions. Possibly typists strongly inhibit the activation of distant multicharacter units. Most likely, typists detect potentially multicharacter errors and stop typing after only one or two incorrect keystrokes. Typists do detect most errors during execution, virtually always within one or two letters of the error (Long, 1976; Rabbitt, 1978). If detected following one keystroke, a partly executed digraph transposition would appear as an anticipatory insertion. If detected following two keystrokes, it might appear as a two-letter omission.

Summary

Studies of the errors made during transcription by novice and skilled typists allow the correction and extension of previous analyses of typing errors, with implications for representation at various levels of the motor system during performance. Videotape records suggest that a keystroke is explicitly represented in terms of the hand, finger, and finger position that uniquely specifies it, and that a common source of error is the incorrect assignment of one of these three components. Further analyses provide support for previous indications that multicharacter response units, notably digraphs, are represented during execution, and that certain errors occur within, and other errors across, such units. The formation of such multicharacter units could explain differences in the patterns of novice and expert errors. Finally, the special problems arising from "deactivation" of representations in order to avoid perseveration may explain other errors, as well as mechanisms developed to avoid them.

Acknowledgments

I thank Craig Will and Don Gentner for their assistance and many discussions, Don Norman for his advice and contributions, and the LNR Research Group for its encouragement and support of this work. I also thank the typists who contributed to these studies, including students in typing classes conducted by James Henderson at La Jolla High School. Mark Wallen developed software; Patty Haden assisted with data analysis, and Eileen Conway with manuscript preparation.

The research reported here was conducted under Contract N00014-79-C-0323, NR 667-437 with the Personnel and Training Research Programs of the Office of Naval Research, and was sponsored by the Office of Naval Research and the Air Force Office of Scientific Research.

Reference Notes

1. Gentner, D. R. *Skilled finger movements in typing* (Tech. Rep. 104) La Jolla, Calif.: University of California at San Diego, Center for Human Information Processing, July 1981.
2. Gentner, D. R., Grudin, J. T., & Conway, E. *Finger movements in transcription typing* (Tech. Rep. 8001). La Jolla, Calif.: University of California at San Diego, Center for Human Information Processing, May 1980.

References

Dvorak, A., Merrick, N. L., Dealey, W. L., & Ford, G. C. *Typewriting behavior.* New York: American Book Company, 1936.

Grudin, J. T. *The organization of serial order in typing.* Unpublished doctoral dissertation, University of California at San Diego, 1981.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley, 1951.

Lessenberry, D. D. *Analysis of errors.* Syracuse, N.Y.: L. C. Smith and Corona Typewriters, School Department, 1928. (Reprinted in A. Dvorak, N. Merrick, W. Dealey, & G. Ford, *Typewriting behavior.* New York: American Book Company, 1936.)

Long, J. Visual feedback and skilled keying: Differential effects of masking the printed copy and the keyboard. *Ergonomics*, 1976, *19*, 93-110.

MacNeilage, P. F. Typing errors as clues to serial ordering mechanisms in language behaviour. *Language and Speech*, 1964, 7, 144-159.

Norman, D. A. Categorization of action slips. *Psychological Review*, 1981, *88*, 1-15.

Owen, Margaret B. *The secret of typewriting speed.* Chicago: Forbes, 1919.

Rabbitt, P. Detection of errors by skilled typists. *Ergonomics*, 1978, *21*, 945-958.

Rumelhart, D. E., & Norman, D. A. Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, 1982, *6*, 1-36.

Russon, A. R., & Wanous, S. J. *Philosophy and psychology of teaching typewriting* (2nd ed.). Cincinnati: South-Western Publishing, 1973.

Shaffer, L. H. Control processes in typing. *Quarterly Journal of Experimental Psychology*, 1975, *27*, 419-432.

Shaffer, L. H. Timing in the motor programming of typing. *Quarterly Journal of Experimental Psychology*, 1978, *30*, 333-345.

Shaffer, L. H., & Hardwick, J. Typing performance as a function of text. *Quarterly Journal of Experimental Psychology*, 1968, *20*, 360-369.

Shaffer, L. H., & Hardwick, J. Errors and error detection in typing. *Quarterly Journal of Experimental Psychology*, 1969, *21*, 209-213.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning.* New York: Academic Press, 1978.

Terzuolo, C. A., & Viviani, P. Determinants and characteristics of motor patterns used for typing. *Neuroscience*, 1980, *5*, 1085-1103.

Appendix

NOVICE ERRORS

(read: column item 'a' was struck for row item 'c' 1 time, etc.)

*	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	-	0	1	0	35	0	1	5	0	0	0	0	0	4	29	1	1	1	72	2	2	0	0	1	0	2
b	0	-	0	1	0	0	0	0	0	0	0	0	1	19	0	0	0	0	0	1	0	22	0	0	0	0
c	0	1	0	3	3	0	0	0	0	0	0	0	0	3	0	0	0	0	3	0	0	7	0	0	0	0
d	1	1	16	-	18	4	3	0	0	0	46	9	0	2	0	0	0	1	70	1	0	0	0	0	1	0
e	26	0	0	36	-	0	0	3	53	0	1	1	0	1	19	1	1	131	6	9	3	0	54	0	4	1
f	0	1	1	11	0	-	11	2	0	2	1	0	0	0	0	0	0	7	0	3	0	4	1	0	0	0
g	0	3	0	0	0	10	-	21	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
h	0	0	0	0	0	4	1	-	0	16	1	0	0	4	4	0	0	1	3	1	0	0	1	0	9	0
i	1	0	0	0	138	0	0	0	-	0	2	1	0	1	84	0	0	2	3	0	36	0	1	0	1	0
j	0	0	0	0	0	8	0	14	0	-	5	0	9	1	1	1	0	1	2	3	4	0	0	0	0	0
k	0	0	0	105	0	0	1	8	1	2	-	22	0	0	1	0	0	0	2	0	1	0	0	0	0	0
l	4	0	0	9	0	0	0	6	0	1	20	-	0	0	7	0	0	1	7	6	2	0	0	0	0	0
m	0	2	0	0	0	0	0	0	0	0	0	0	-	140	0	0	0	0	1	0	0	0	1	0	0	0
n	0	2	0	2	1	0	1	6	0	2	0	0	111	-	1	2	0	4	5	1	0	2	4	0	3	4
o	14	0	0	0	44	0	0	3	55	0	0	21	0	3	-	38	0	2	1	6	9	0	4	0	4	0
p	5	0	0	0	3	0	0	0	0	0	0	0	0	0	65	-	4	3	0	0	0	0	1	0	3	0
q	2	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2	-	0	1	1	0	0	5	0	1	0
r	1	0	0	0	88	5	0	0	0	0	0	0	0	5	1	2	1	-	12	73	12	0	3	0	3	0
s	71	0	9	112	25	1	0	0	0	0	6	10	0	7	3	0	0	9	-	4	1	0	28	8	0	1
t	1	0	0	5	22	26	2	2	1	0	0	1	0	14	7	8	0	198	2	-	2	0	1	0	62	0
u	1	0	1	0	2	0	0	0	4	3	0	0	0	0	6	1	0	19	0	4	-	0	1	0	131	0
v	1	39	7	0	1	0	0	0	0	0	0	0	2	3	0	0	0	2	0	0	0	-	0	0	0	0
w	2	0	0	2	68	0	0	0	0	0	0	0	0	1	10	0	9	0	33	5	0	0	-	0	3	0
x	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	2	-	0	10
y	1	0	0	0	1	1	0	5	0	0	0	2	0	0	0	1	0	5	0	28	47	0	0	0	-	0
z	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2	0	-
,	0	0	0	1	0	0	0	0	0	0	1	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0
.	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
;	0	0	0	0	0	0	0	0	0	0	0	5	0	0	1	5	0	1	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
-	1	0	0	0	1	0	0	0	0	0	0	0	2	3	0	0	0	0	0	0	0	0	0	0	0	0
1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0

NOVICE ERRORS, NORMALIZED FOR LETTER FREQUENCY

(reflects probability of substitution)

*	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	-	0	2	0	8	0	5	3	0	0	0	0	0	2	8	1	10	0	29	1	2	0	0	10	0	20
b	0	-	0	1	0	0	0	0	0	0	0	0	2	11	0	0	0	0	0	0	0	110	0	0	0	0
c	0	5	0	2	1	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	35	0	0	0	0
d	0	6	32	-	4	7	15	0	0	0	153	9	0	1	0	0	0	0	28	0	0	0	0	0	1	0
e	9	0	0	24	-	0	0	2	66	0	3	1	0	1	5	1	10	47	2	3	2	0	54	0	2	10
f	0	5	2	7	0	-	55	1	0	20	3	0	0	0	0	0	0	2	0	1	0	20	1	0	0	0
g	0	15	0	0	0	17	-	13	0	0	3	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0
h	0	0	0	0	0	7	5	-	0	160	3	0	0	2	1	0	0	0	1	0	0	0	1	0	6	0
i	0	0	0	0	32	0	0	0	-	0	7	1	0	1	24	0	0	1	1	0	30	0	1	0	1	0
j	0	0	0	0	0	13	0	9	0	-	17	0	15	1	0	1	0	0	1	1	3	0	0	0	0	0
k	0	0	0	70	0	0	5	5	1	20	-	22	0	0	0	0	0	0	1	0	1	0	0	0	0	0
l	1	0	0	6	0	0	0	4	0	10	67	-	0	0	2	0	0	0	3	2	2	0	0	0	0	0
m	0	10	0	0	0	0	0	0	0	0	0	0	-	82	0	0	0	0	0	0	0	0	1	0	0	0
n	0	10	0	1	0	0	5	4	0	20	0	0	185	-	0	2	0	1	2	0	0	10	4	0	2	40
o	5	0	0	0	10	0	0	2	69	0	0	21	0	2	-	35	0	1	0	2	8	0	4	0	2	0
p	2	0	0	0	1	0	0	0	0	0	0	0	0	0	19	-	40	1	0	0	0	0	1	0	2	0
q	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	-	0	0	0	0	0	5	0	1	0
r	0	0	0	0	20	8	0	0	0	0	0	0	0	3	0	2	10	-	5	21	10	0	3	0	2	0
s	24	0	18	75	6	2	0	0	0	0	20	10	0	4	1	0	0	3	-	1	1	0	28	80	0	10
t	0	0	0	5	3	43	10	1	1	0	0	1	0	8	2	7	0	71	1	-	2	0	1	0	39	0
u	0	0	2	0	0	0	0	0	5	30	0	0	0	0	2	1	0	7	0	1	-	0	1	0	82	0
v	0	195	14	0	0	0	0	0	0	0	0	0	3	2	0	0	0	1	0	0	0	-	0	0	0	0
w	1	0	0	2	11	0	0	0	0	0	0	0	0	1	3	0	90	0	13	1	0	0	-	0	2	0
x	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2	-	0	100
y	0	0	0	0	0	2	0	3	0	0	0	2	0	0	0	1	0	2	0	8	39	0	0	0	-	0
z	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	10	0	0	0	0	0	0	20	0	-
,	0	0	0	1	0	0	0	0	0	0	3	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0
.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
;	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	5	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
-	0	0	0	0	0	0	0	0	0	0	0	0	3	2	0	0	0	0	0	0	0	0	0	0	0	0
1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0

SKILLED TYPIST ERRORS

(read: column item 'a' was struck for row item 'e' 9 times, etc.)

*	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	,	.	;	sp
a	-	0	0	0	9	1	0	1	1	0	0	1	0	0	1	0	1	3	4	3	1	0	1	0	0	0	0	0	0	0
b	1	-	0	0	0	1	0	3	0	0	0	0	1	1	1	1	0	0	0	2	0	5	0	0	0	1	0	0	0	0
c	1	1	-	6	2	0	1	0	0	0	0	0	0	0	1	0	0	0	4	7	0	3	0	3	0	0	0	0	0	0
d	0	0	2	-	11	9	2	0	0	0	3	1	1	1	0	0	0	0	4	2	0	0	0	0	0	0	0	0	0	0
e	7	1	1	2	-	0	3	0	2	0	2	2	0	0	0	0	0	3	3	4	1	3	6	0	1	0	0	0	0	1
f	0	1	0	3	0	-	5	0	1	1	0	0	1	1	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0
g	0	4	1	0	0	7	-	1	1	0	0	0	0	0	0	0	1	0	1	3	0	0	0	0	0	0	0	0	0	0
h	1	0	0	0	0	0	2	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
i	3	0	0	0	3	0	0	0	-	0	0	0	0	0	9	0	0	1	0	2	6	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	1	0	1	0	-	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
k	0	0	0	2	2	0	0	0	0	0	-	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	1	0	1	0	2	0	0	0	1	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1
m	2	1	0	0	1	0	1	0	0	0	0	0	-	6	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
n	3	1	0	0	1	0	0	0	1	0	0	1	7	-	0	0	0	3	0	1	0	0	0	0	0	0	0	0	0	0
o	6	0	1	0	0	1	0	0	9	0	0	2	0	1	-	2	0	1	0	1	2	0	0	0	1	0	0	0	0	0
p	0	0	0	0	2	0	0	0	0	0	0	0	0	0	3	-	0	2	0	0	0	0	0	0	0	0	0	0	0	0
q	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0	0
r	4	1	0	0	6	1	0	0	0	0	0	0	1	0	1	0	0	-	1	6	1	1	0	0	0	0	0	0	0	0
s	14	0	9	4	4	0	0	1	1	0	0	0	0	2	1	1	0	2	-	3	0	0	1	1	0	0	0	0	0	0
t	1	4	1	1	5	4	4	0	2	0	2	1	0	2	3	0	0	4	3	-	1	0	3	0	1	0	0	0	0	0
u	0	0	1	0	0	0	0	0	4	0	0	0	0	1	5	2	0	0	0	1	-	0	0	0	0	0	0	0	0	0
v	1	3	2	0	0	4	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0
w	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	2	0	6	4	0	0	-	0	0	0	0	0	0	0
x	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	-	0	1	0	0	0	0
y	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	1	1	5	0	0	0	-	0	0	0	0	0
z	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	-	0	0	0	0
,	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-	12	0	0
.	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	-	0	0
;	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0
sp	0	0	6	0	4	0	0	0	0	0	0	1	1	5	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	-
/	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0
3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

SKILLED TYPIST ERRORS, NORMALIZED FOR LETTER FREQUENCY

(reflects probablility of substitution)

*	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	,	.	;	sp
a	-	0	0	0	3	1	0	1	1	0	0	1	0	0	0	0	44	2	2	1	1	0	2	0	0	0	0	0	0	0
b	0	-	0	0	0	1	0	2	0	0	0	0	2	1	0	2	0	0	0	1	0	17	0	0	0	53	0	0	0	0
c	0	2	-	6	1	0	2	0	0	0	0	0	0	0	0	0	0	0	2	3	0	10	0	37	0	0	0	0	0	0
d	0	0	2	-	3	13	4	0	0	0	14	1	2	1	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0
e	3	2	1	2	-	0	5	0	1	0	9	2	0	0	0	0	0	2	2	2	1	10	14	0	2	0	0	0	0	0
f	0	2	0	3	0	-	9	0	1	37	0	0	2	1	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0
g	0	10	1	0	0	10	-	1	1	0	0	0	0	0	0	0	44	0	1	1	0	0	0	0	0	0	0	0	0	0
h	0	0	0	0	0	0	4	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0	0	0	0
i	1	0	0	0	1	0	0	0	-	0	0	0	0	0	4	0	0	1	0	1	7	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	1	0	1	0	-	0	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
k	0	0	0	2	1	0	0	0	0	0	-	2	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	0	0	1	0	1	0	0	0	1	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0
m	1	2	0	0	0	0	2	0	0	0	0	0	-	3	0	0	0	0	0	0	0	0	0	12	0	0	2	0	0	0
n	1	2	0	0	0	0	0	0	1	0	0	1	11	-	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
o	3	0	1	0	0	1	0	0	5	0	0	2	0	1	-	4	0	1	0	0	2	0	0	0	2	0	0	0	0	0
p	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	-	0	1	0	0	0	0	0	0	0	0	0	0	0	0
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0	0
r	2	2	0	0	2	1	0	0	0	0	0	0	2	0	0	0	0	-	1	2	1	3	0	0	0	0	0	0	0	0
s	6	0	10	4	1	0	0	1	1	0	0	0	0	1	0	2	0	1	-	1	0	0	2	12	0	0	0	0	0	0
t	0	10	1	1	1	6	7	0	1	0	9	1	0	1	1	0	0	2	2	-	1	0	7	0	2	0	0	0	0	0
u	0	0	1	0	0	0	0	0	2	0	0	0	0	1	2	4	0	0	0	0	-	0	0	0	0	0	0	0	0	0
v	0	7	2	0	0	6	2	0	1	0	0	0	0	1	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0
w	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	89	0	3	2	0	0	-	0	0	0	0	0	0	0
x	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	-	0	53	0	0	0	0
y	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	6	0	0	0	-	0	0	0	0	0
z	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	12	0	-	0	0	0	0
,	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-	35	0	0
.	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	-	0	0
;	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0
sp	0	0	7	0	1	0	0	0	0	0	0	1	2	3	0	0	0	0	1	0	0	3	0	0	0	0	2	0	0	-
/	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0

LESSENBERRY ERROR MATRIX

(read: column item 'a' was struck for row item 'b' 69 times, etc.)

*	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	-	69	106	71	560	102	24	139	148	26	15	136	58	61	126	60	89	104	1012	116	46	96	100	33	46	139
b	36	-	37	43	33	44	163	42	12	8	3	25	57	262	43	13	1	7	28	40	21	773	7	6	15	9
c	59	43	-	486	98	67	22	19	36	7	27	14	20	84	25	1	7	30	174	28	14	466	22	133	8	9
d	65	43	436	-	755	394	117	23	73	27	484	58	7	80	27	6	2	72	862	119	22	21	30	22	11	1
e	494	28	149	651	-	72	76	114	1019	20	25	141	66	75	174	27	5	813	276	273	95	44	638	17	77	7
f	43	80	92	524	72	-	586	56	66	176	24	38	6	27	56	47	3	419	71	153	12	159	12	1	8	1
g	35	241	34	166	75	541	-	461	24	54	41	17	9	59	8	5	8	37	30	267	18	51	9	3	8	3
h	40	44	43	22	42	47	389	-	100	207	111	21	19	275	46	15	6	26	36	117	44	11	9	4	99	1
i	159	10	17	73	907	28	25	84	-	30	211	81	50	90	915	21	13	59	78	150	560	54	20	2	27	1
j	11	6	10	11	21	140	62	489	46	-	217	49	242	188	28	7	1	6	4	40	128	5	1	2	10	2
k	17	10	34	287	29	11	16	66	172	101	-	521	37	44	56	10	2	17	23	16	7	2	2	4	3	0
l	119	44	16	59	116	54	24	42	106	54	413	-	62	97	567	92	7	40	150	77	29	3	8	7	14	3
m	52	16	11	8	41	6	15	13	27	60	56	21	-	1577	77	24	5	17	25	14	36	39	6	3	2	2
n	93	135	33	44	87	19	68	151	90	28	17	50	1249	-	68	10	6	27	74	67	25	63	8	6	22	3
o	210	18	23	30	134	82	20	68	1290	26	52	671	54	63	-	346	36	154	78	112	154	22	104	8	22	3
p	33	11	9	6	56	11	10	14	28	3	6	46	15	24	405	-	36	23	14	43	19	6	13	3	15	1
q	232	6	5	3	21	11	12	7	9	3	0	4	10	3	10	22	-	6	16	66	20	17	109	3	6	6
r	111	15	40	139	1043	334	36	24	92	5	7	27	14	70	140	33	8	-	186	1488	398	58	67	19	28	7
s	954	23	260	1061	300	106	23	16	69	6	24	162	27	91	68	12	8	156	-	150	57	59	410	157	15	19
t	97	38	24	105	212	185	254	129	157	26	20	65	15	58	79	23	55	1407	168	-	58	7	55	45	327	5
u	33	59	10	28	64	17	15	49	519	124	23	40	13	35	123	21	40	217	81	99	-	22	25	52	419	12
v	45	1001	430	39	71	137	38	28	35	1	2	10	34	101	17	2	18	44	24	11	40	-	11	11	20	1
w	89	9	10	76	927	26	6	19	15	3	1	17	16	22	122	30	96	136	382	95	18	26	-	10	12	26
x	33	12	292	35	24	1	3	11	8	5	1	2	9	8	11	5	5	21	225	23	23	10	12	-	16	191
y	44	10	5	18	57	11	26	109	33	13	3	12	18	31	12	11	5	32	138	508	502	90	43	13	-	8
z	288	7	6	7	6	2	3	2	10	1	3	3	1	5	4	4	6	13	34	6	2	9	30	108	6	-

LESSENBERRY ERROR MATRIX, NORMALIZED FOR LETTER FREQUENCY

(reflects probability of substitution)

*	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	-	115	95	44	110	106	30	63	51	433	58	83	56	21	40	83	2747	46	402	30	43	245	119	458	61	5110
b	11	-	33	27	6	46	204	19	4	133	12	15	55	91	14	18	31	3	11	10	19	1972	8	83	20	331
c	19	72	-	304	19	70	28	9	13	117	104	9	19	29	8	1	206	13	69	7	13	1189	26	1847	11	331
d	21	72	389	-	149	410	146	10	25	450	1862	35	7	28	9	8	62	32	342	31	20	54	36	306	14	37
e	156	47	133	407	-	75	95	52	354	333	96	86	63	26	56	38	154	357	110	71	88	112	760	236	101	257
f	14	133	82	328	14	-	733	25	23	2933	92	23	6	9	18	65	93	184	28	40	11	406	14	14	11	37
g	11	402	30	104	15	560	-	210	8	900	158	10	9	20	3	7	247	16	12	70	17	130	11	42	11	110
h	13	73	38	14	8	49	486	-	35	3450	427	13	18	95	15	21	185	11	14	30	41	28	11	56	130	37
i	50	17	15	46	179	29	31	38	-	500	812	49	48	31	293	29	401	26	31	42	519	138	24	28	36	37
j	3	10	9	7	4	146	78	222	16	-	835	30	233	65	9	10	31	3	2	10	119	13	1	28	13	74
k	5	3	30	179	6	11	20	30	46	1683	-	318	36	15	18	14	62	7	9	4	6	5	2	56	4	0
l	38	73	14	37	23	56	30	19	37	900	1588	-	60	34	182	128	216	18	60	20	27	8	10	97	18	110
m	16	27	10	5	8	6	19	6	9	1000	215	13	-	548	25	33	154	7	10	4	33	99	7	42	3	74
n	29	225	29	28	17	20	85	69	31	467	65	30	1201	-	22	14	185	12	29	17	23	161	10	83	29	110
o	66	30	21	19	26	85	25	31	448	433	200	409	52	22	-	481	1111	68	31	29	143	56	124	111	29	110
p	10	18	8	4	11	11	13	6	10	50	23	28	14	8	130	-	1111	10	6	11	18	15	15	42	20	37
q	73	10	4	2	4	11	15	3	3	50	0	2	10	1	3	31	-	3	6	17	19	43	130	42	8	221
r	35	25	36	87	205	348	45	11	32	83	27	16	13	24	45	46	247	-	74	388	369	148	80	264	37	257
s	302	38	232	663	59	110	29	7	24	100	92	99	26	32	22	17	247	68	-	39	53	151	488	2181	20	699
t	31	63	21	66	42	193	318	59	55	433	77	40	14	20	25	32	1698	617	67	-	54	18	65	625	430	184
u	10	98	9	18	13	18	19	22	180	2067	88	24	13	12	39	29	1235	95	32	26	-	56	30	722	551	441
v	14	1668	384	24	14	143	48	13	12	17	8	6	33	35	5	3	556	19	10	3	37	-	13	153	26	37
w	28	15	9	48	182	27	8	9	5	50	4	10	15	8	39	42	2963	60	152	25	17	66	-	139	16	956
x	10	20	261	22	5	1	4	5	3	83	4	1	9	3	4	7	154	9	89	6	21	26	14	-	21	7022
y	14	17	4	11	11	11	33	50	11	217	12	7	17	11	4	15	154	14	55	132	465	230	51	181	-	294
z	91	12	6	4	1	2	4	1	3	17	12	2	1	2	1	6	185	6	13	2	2	23	36	1500	8	-