Visual Components of the Reading Process

Ralph Norman Haber
Lyn R. Haber

Since reading must begin with the pickup of information on the page, we have examined in some detail the principal arrangements of that information, and considered what is known about the pickup process. One source of information is the arrangements of the print on the page: arrangements which supply information about word, sentence, and paragraph boundaries and even substantial syntactic and semantic information. Evidence is reviewed to show that reading performance is hurt when these conventional arrangements are violated. The second source is from the letter shapes alone, in which most work has been done using a feature description. While four different procedures for generating features are reviewed, none of them either meet adequate tests of validity, nor supply convincing evidence that readers actually use feature tests to identify letters. Many suggestions for future work are offered, but the current state of feature testing theories is pretty dismal.

The final source concerns higher order visual information, especially of the configuration of groups of letters in a word, or of the shape of the entire word itself. Evidence is reviewed to show that such information is substantial, especially for some types of words, and that if pushed, readers can use that information to identify words and to comprehend text. For all three sources we have shown that readers can use the information. Much less is known about how readers routinely use these sources.

When we read, we combine information picked up visually from the page of print with different forms of knowledge we already possess. Our prior knowledge includes the spelling rules for specifying permissible and likely letter sequences; perhaps phonological rules that translate spelling to sound; syntactic rules for the permissible and likely sequences of words within sentences; world knowledge rules for the likely contents that may occur in the story; and even stylistic rules for the likely sequences of words and sentences, given the form and author of the passage. Differences between the principal models and theories of reading have focused on the description of interactions

Visible Language, XV 2, pp. 147-182.
Habers' address: Department of Psychology, University of Illinois, Chicago, IL 60680.

among these different rule systems and how a reader juggles them in order to extract meaning from printed text. But all theories agree that information is present on the page of print that must be used in the reading process, and that, at some point, the reader must have visual access to this information. The present paper examines the different sources of this visual information available to a reader, and considers how these are normally used by skilled readers.

Three general categories of visual information are described: the arrangements of the printed characters on the page of print; the visual features of the printed alphabetic characters; and the visual shapes of whole word configurations. While English is the assumed language for these analyses, the principles apply equally for all other printed languages, even though the content of particular rules may differ from language to language.

Arrangements of print on the page

Every printed language has rules which specify how and where the marks on the page are to be placed. For English the characters must be printed in horizontal rows, in which the sequence runs from left to right across the page. There is always a left justified margin, and usually one on the right as well. While often overlooked by theorists, the rules which specify how the printed characters (and non-characters such as spaces) must be arranged, provide substantial information about the content of the text – information that is potentially available to readers. Readers do routinely make use of this information, since disruption in reading results when these printing rules are altered or violated.

Table I lists a few of these print arrangement rules, and the information specified by each rule. For example, in print (and in typing and handwriting) a single space character must separate words (unless something else does so, such as punctuation). This rule means that words are visually identified by the surrounding white space, and anything with surrounding white space functions as a word in English. This definition of a word in print is far more general and easily applied than current definitions of a word in speech (see Cole & Jakimik, 1980). It is also probably more quickly learned, too, though we know of no evidence. As a second example of a print arrangement rule, sentences are visually marked according to three inviolate arrangement rules: sentences begin with the initial word capitalized; sentences end with the final word followed by one of a specifiable class of terminal punctuation marks; and that terminal punctuation is followed by an extra space. Table II illustrates several examples in which some of these printing conventions are violated in otherwise normal text.

Table I. Some examples of printing conventions (from Haber & Haber, 1981).

End of line conventions	*Transformation conventions*
Right and left justified	Question has question mark at end (some languages at beginning).
Hyphen indicates word division	Imperatives and exclamations have exclamation mark.
Paragraph conventions	Statements have periods.
Indentation	Emphatics have italics or boldface.
Skip a line	*Direct speech conventions*
Bold face, initial character enlarged, etc.	Quotation marks
Sentence boundary conventions	Paragraphing for change of speaker
Initial capitalization	*Word conventions*
Final punctuation marks	Space before and after
Extra space	Initial capitalization for proper nouns
Phrase boundary conventions	
Punctuation	
Extra space	

Table II. Examples of passages in which some printing rules have been violated (from LeCarre, 1975: 238-9).

A. WhenSmileyhadlefttheIslayforGrosvenorSquarethatmorning,thestreetshadbeen bathedinharshsunshineandtheskywasblue.NowashedrovethehiredRoverpasttheun lovablefacadesoftheEdgwareRoad,thewindhaddropped,theskywasblackwithwait ingrain,andallthatremainedofthesunwasalingeringrednessonthetarmac.

B. he parked in st johns wood road in the forecourt of a new tower block with a glass porch but he did not enter by the porch passing a large sculpture describing as it seemed to him nothing but a sort of cosmic muddle he made his way through icy drizzle to a descending staircase marked exit only

C. The first flight was of terazzo tile and had a bannister of African teak. Below that, the contractor's generosity ceased. Rough-rendered plaster replaced the earlier luxury and a stench of uncollected refuse crammed the air.

D. HiS mAnNeR wAs CaUtIoUs RaThEr ThAn FuRtIvE, bUt WhEn He ReAcHeD tHe IrOn DoOr He PaUsEd BeFoRe PuTtInG bOtH hAnDs To ThE lOnG hAnDlE aNd DrEw HiMsElF tOgEtHeR aS iF fOr An OrDeAl.

Print arrangement rules serve a number of functions in addition to specifying word and sentence boundaries. Final punctuation marks, such as question marks and exclamation points, provide not only syntactic but affective intent. The print arrangement conventions for dialogue delimit each direct quote and keep separate the utterances of each speaker. Print conventions for the portrayal of conversational dialogue include: (1) direct quotes begin with a capital letter, (2) end with a limited set of punctuation markers, (3) are en-

Table III. Examples of violations of print arrangement rules for dialogue (from Fowles, 1969).

A. "And since?"

"Not a sign."

"You saw the vicar?"

"No, but Miss Trimble assures me he went to Marlborough House this forenoon. He was told Mrs. Poulteney was unwell. He spoke to Mrs. Fairley. All she knew was that some disgraceful matter had come to Mrs. Poulteney's knowledge, that she was deeply shocked and upset . . .". The good Mrs. Tranter broke off, apparently almost as distressed at her ignorance as at Sarah's disappearance. She sought her niece's and Charles's eyes. "What can it be – what *can* it be?"

B. That is the depot for the coaches, you know. The Dorchester to Exeter omnibuses did not descend the steep hill to Lyme, but had to be picked up at a crossroads some four miles inland on the main road to the west. But Mrs. Hunnicott spoke to the man. He is most positive that Miss Woodruff was not there.

C. I have no one to turn to. I hoped I had made it clear that Mrs. Tranter – Has the kindest heart. But I do not need kindness. There was a silence. He still stood parting the ivy. I am told the vicar is an excellently sensible man. It was he who introduced me to Mrs. Poulteney.

closed within double quotation marks; (4) if a new speaker intrudes, his utterance begins on a new line. Table III shows several examples of conversational dialogue. The first passage is written according to normal print conventions. In the second passage, the quotation marks have been removed. In the third, the new line convention is violated as well. The difficulty we experience in sorting out the utterances of the two speakers in the third passage shows in an informal way how useful these print arrangement rules for dialogue are.

Print arrangement rules also specify special modes of communication. For example, lists are usually presented in lines, either vertically or horizontally; outlines use these conventions to display the overall hierarchical organization of the text; formulas are set off; and poetry has a number of uses and conventions concerning line arrangement. A few of these are illustrated in Table IV.

In the more conventional forms of poetry, line length is specified by metric count. The final words of the lines rhyme either with adjacent or more remote line-final words, depending upon the specific stanza form employed. The most straight forward use of line length in poetry allows line end and final rhyme to coincide with major syntactic units, as in example A in Table IV. More frequently, syntactic units are played off against the line arrangement demanded by the conventional form, resulting in a tension between rhyme, and meter, and syntax. This is illustrated in example B. Where no rhyme is used, poets still employ this tension between syntax and the line length demand-

Table IV. Examples of violations of print arrangement rules in poetry.

A. Line length determined by metric count; rhyme at end of line coincides with the end of the syntactic unit; tetrameter; Poe (1950).

At midnight in the month of June,
I stand beneath the mystic moon.

B. Line length determined by metric count; end of line rhyme in conflict with syntactic unit; pentameter; Browning (1952).

That's my last Duchess painted on the wall,
Looking as if she were alive. I call
That piece a wonder, now: Fra Pandolf's hands
Worked busily a day, and there she stands.
Will't please you sit and look at her? I said
"Fra Pandolf" by design, for never read
Strangers like you that pictured countenance,
The depth and passion of its earnest glance . . .

C. Line length determined by metric count; blank verse, no rhyme; tension between syntactic units and line units; Shakespeare (1959).

Tomorrow, and tomorrow, and tomorrow,
Creeps in this petty pace from day to day,
To the last syllable of recorded time;
And all our yesterdays have lighted fools;
The way to dusty death. Out, out, brief candle;
Life's but a walking shadow; a poor player,
That struts and frets his hour upon the stage,
And then is heard no more; it is a tale
Told by an idiot, full of sound and fury,
Signifying nothing.

D. Line arrangement used to mirror meaning; e. e. cummings (1950).

Buffalo Bill's
defunct
                who used to
                ride a watersmooth – silver
                                                            stallion
and break onetwothreefourfive   pigeonsjustlikethat
                                                                                Jesus
he was a handsome man
                                        and what i want to know is
how do you like your blueeyed boy
Mister Death

ed by the metric convention. The passage from *MacBeth* (example C) is one example. A very different use of line arrangement in poetry is reflected in the final example. Here, Cummings has used line arrangement (in conjunction with violations of punctuation and word spacing rules) to reflect meaning.

In each of these examples the poetic form involves a violation of normal line arrangement and the imposition of new conventions (rhyme, meter, meaning).

A number of experiments have shown that when the print arrangement rules described in Table I are violated normal reading is disrupted. Such data are evidence that readers routinely use printing arrangement as a source of visual information in the reading process. Consider for example, how difficult it is to read some of the brief passages in Table II. Hochberg, Levin, and Frail (1966) found that simply omitting the space character reduces reading speed by up to 50%, compared to reading comparable unadulterated text. More importantly, they found the percentage reduction greatest for the better readers, suggesting that better readers use this source of visual information more than do poorer readers.

As another example of the disruptive effect when print arrangement rules are violated, both Smith (1969) and Fisher (1975) had college subjects read text in which the letters alternated in case (as in the last passage of Table II). Fisher also deleted the space character. Both studies showed massive disruption in reading when performance was compared to reading unadulterated passages. No interactions were reported with reading ability in these studies, presumably because of the homogeneity of the sample with respect to ability.

Less research has been reported on other aspects of printing conventions. McNamara, Patterson, and Tinker (1953) review the evidence for the role of print size, but find that except for very small sizes or very large sizes, varying the overall size of the letters does not affect reading performance. Similarly, they conclude that line length, except for extremely long or short lines, is unimportant. We expect this conclusion to be more limited than McNamara et al. suggest, since short line lengths, such as the typical 35-45 character newspaper column, should deprive the reader of substantial peripheral vision (see also Tinker, 1956). Rayner (1975) showed that access to peripheral vision up to 15 characters to the right of fixation is useful, and for much of the time in reading newspaper columns, readers have less than this.

Kolers and Perkins (1969) have explored the importance of the print convention of left-to-right sequence. They found that rearranging all the sequences in a right-to-left order still permits English readers to understand the text, but initially it slows them down tremendously. (We would again expect this effect to obtain more strongly for good than for poor readers.) In another experiment concerning this print convention, Pollatsek, Bolozky, Well, and Rayner (1980) showed that the advantage of the left-to-right order for English was not due to any asymmetries in visual sensitivities or to any laterality differences in the brain, but presumably only to practice with the particular order. They used a computer generated display of text in which the part of the text that was readable depended upon where the reader's eye was fixated. In one condition the reader could see text well to the right of fixation but not to the left (presumably the normal situation in reading English – see Rayner, 1975); whereas, in the other, peripheral vision was possible only to the left of fixation, and not to the right. The subjects were American college students reading English, and bilingual Israeli students reading both English and Hebrew. Reading was impaired only when peripheral vision was blocked out in the direction of reading. Thus, the Israeli readers needed to see to the left for Hebrew and to the right for English. Pollatsek et al. concluded that the preferred direction of reading was entirely the result of practice and did not depend upon any structural characteristic of the brain or the visual system.

Several researchers (see Hartley, 1980, for a review) have suggested that novel spatial arrangements of print can be used to increase comprehension or reading speed. The last issue of this journal (volume XV number 1) was devoted to the spatial arrangement of text, with the individual articles arranged typographically to illustrate the theories advocated by the various authors. Guest editor Hartley suggests in his introduction that the debate is with the details rather than with the general proposition that typography can be manipulated in order to improve comprehension. It is possible, however, that such effects apply to technical material and/or only to certain kinds of search activities, rather than more generally to all materials being read for all purposes. Thus, we need more research on whether there is anything to be gained by adding more printing convention rules which will still further specify the content of the text.

These few examples suggest the general importance of the printing conventions as sources of visual information in reading. However, we can read even when some of the conventions are violated. While every study that tested violations showed disruption in reading, readers were still able to proceed, even if at reduced efficiency. Further, the few studies that provided enough practice with the altered materials showed that performance continued to improve, and with enough practice, might even return to normal (e.g., Kolers and Perkins, 1975). Thus printing conventions provide information that readers use, so that reading suffers when this information is not available. Yet we know print conventions are not the only sources of information, and readers can learn to read even when some of the printing convention information is absent. An interesting question to explore is whether even extensive practice could permit normal fluent reading for skilled readers, if the more important print arrangement rules were simultaneously violated.

Visual characteristics of individual letters

Whereas research on printing conventions has been scanty, so these rules are often even ignored in discussions of visual processes in reading, considerable attention has been devoted to the visual processing of letters. Since words are composed of letters, each of which has a distinct identity and name, it seems self-evident – and has been very generally assumed – that part of the visual information in reading must include the visual shapes of the individual letters of the alphabet. Yet there has been remarkably little agreement on the description of that information. In the absence of agreement on even the basic description, the lack of unanimity on how readers extract and use letter shape information comes as no surprise.

The problem is graphically illustrated in Figure 1, in which the 26-letter lowercase English alphabet is printed in 24 different typefaces. Not only do we need a system that uniquely distinguishes each letter from every other letter on each line, but one which tells us how the reader knows that all the first letters on each line have the same name, as do the second letters on each line, and so forth.

Two general categories of descriptions of the visual information in letters have been proposed, known as template or prototype models on the one hand, and feature models on the other. The template description of a letter is wholistic and gestalt-like, corresponding to what might be called the concept or essence of the letter. In contrast, a feature description of a letter is a list of the individual elements of a letter that make it uniquely different from all other letters. Neisser (1967) provided an early but still relevant discussion of the relative virtues of and problems with each approach; and his conclusion still reflects the current literature 15 years later. He argued that template descriptions are untenable for all but unrealistically circumscribed alphabets used in highly controlled machine-like reading tasks; and that, therefore, we have to look to descriptions based in some way on separate features. But while he went on to point out theoretical and practical problems with a feature approach, this remains the major approach today.

Unfortunately, there are no agreed upon procedures to identify the visual features of a set of letters. In fact, few complete descriptions of a set of features have been offered in the literature, based upon any procedure. Four rather different procedures have been used, or at best mentioned as ways to determine the distinctive visual features of letters. The first is an armchair

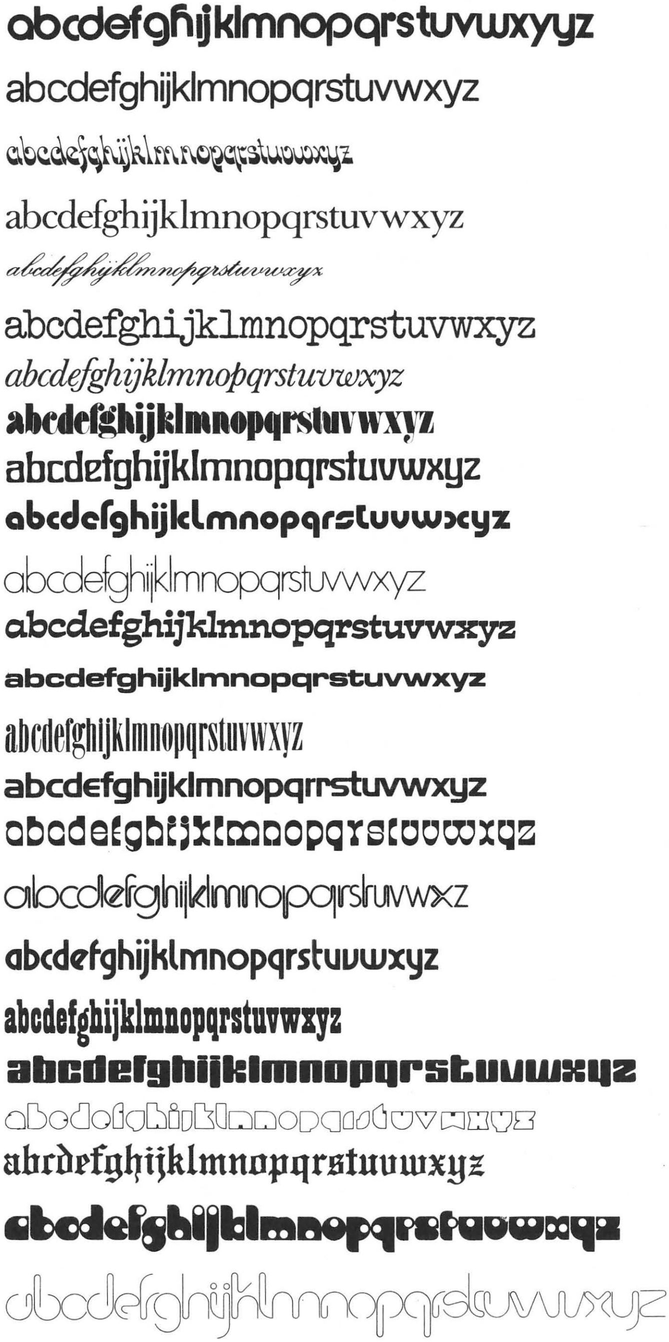


Figure 1. Twenty-four examples of different English alphabet type styles.

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

Figure 2. The typeface Helvetica (Figure 1, line 2) on which the distinctive features presented here are based.

approach, in which the comfortably ensconsed scholar devises a feature set which accounts for the typeface in hand. The second approach is to create new alphabetic characters and attempt to discover salient differences among them, which are then identified as features. The third approach uses real alphabets and real subjects, and tries to group letters as a function of similarity judgements, where presumably groups of letters share one or several features; and the fourth approach uses errors subjects make during an identification task to group confusable letters – these groups presumably share one or more features. We will discuss each of these four approaches in some detail.

It is important to keep in mind from the outset that the current literature described here deals with the visual characteristics of individual letters considered in isolation. Yet in reading, letters occur with other letters, in meaningful words embedded in an even larger meaningful context of phrase, sentence, paragraph, entire text. While a skilled reader may process single letters, it seems unlikely that he does so independently of their context. Thus the existing literature on feature extraction may have little application to the reading process.

In the armchair approach a judge attempts to group the alphabetic characters by pairs, in which the two members of each pair are visually identical in all respects except for one difference. Thus, in the typeface Helvetica (Figure 2) p differs from q only by the right versus left placement of a vertical descender: therefore, right versus left must be one minimal distinctive feature. Similarly, b differs from p by whether the leftward vertical line extends above or below the round part, so the direction of the extension (up vs. down) is another feature. This process, if continued, isolates all of the pairs that differ by only one distinctive feature and represents the beginning of the feature list. Table V shows this process continued for a number of letter pairs.

Table VI and Figure 3 show a complete distinctive feature list for Helvetica. We know of no other exhaustive distinctive feature list for lowercase print in the literature. This is unfortunate, because other armchair authors of feature lists might well have selected other features: this list works but others could, too. For example, half versus whole diagonal line segments might be used to differentiate x and w, where x has only half segments and w has both half and whole.

Table V. Several sets of minimal pairs for the typeface Helvetica (Figure 2) using the distinctive feature system shown in Figure 3.

Feature	*Minimal Pairs*								
	pq	bp	bd	gq	cs	mn	ce	ae	nr
Extender									
Direction	DD	UD	UU	DD	−−	−−	−−	−−	−−
Side	LR	LL	LR	RR	−−	−−	−−	−−	−−
Hook	−−	−−	−−	+−	−−	−−	−−	−−	−−
Central Part									
Enclosed	++	++	++	++	−−	−−	−+	++	−−
Round	++	++	++	++	++	++	++	++	++
Side Gap	−−	−−	−−	−−	RL	−−	RR	LR	−R
Up-Down Gap	−−	−−	−−	−−	−−	BB	−−	−−	BB
Doubling	−−	−−	−−	−−	−−	+−	−−	−−	−−

Irrespective of how such lists are worked out, it is critical to note that most pairs of letters differ by more than one feature. For example, b differs from q by two features: vertical line left vs. right; vertical line ascending vs. descending. For b vs. g, a third feature is needed, the presence/absence of a hook.

There are some printed elements of letters that are not distinctive features at all, even though they are printed. For example, using the same typeface, the dot over the i and the j does not serve to distinguish either of those letters from any non-dotted ones, so the omission of the dot would not change the letter, though we might be unaccustomed to its looks. Similarly for the horizontal crossbar on the f and the t. Those letters differ from other letters by several different features, so the familiar crossbar is not distinctive.

This armchair exercise is straight forward, and follows the model used by linguists for isolating the distinctive features for any aspect of language (for example, see Jakobson, Fant, & Halle, 1965). Three tests are available to decide if such a procedure is successful in generating a reasonable feature list. First, can all differences between all letters be reduced to such a set of independent features? If there are many letters which differ from all others in so many ways that the entire letter has to be a feature, then the feature list is not very useful. Second, is the number of different features substantially less than the number of letters and other symbols to be described? Presumably, if the number of different features exceeds the number of letters (plus numbers, punctuation marks, and other printed symbols a reader might expect to encounter on a page of print), such a feature system would be less interesting in theory than if the 50-100 symbols could be described by no more than 10-20 features. A related test of adequacy requires that the number of features needed increases only slightly as new typefaces are added. Finally, our confidence in any feature list is increased if the features predict actual performance with the letters. Thus, letters that differ by only a few features would be more confusable, or slower to name than letters that differ by many features.

While the feature list in Table VI was designed to differentiate fully among the 26 characters in Figure 2, and was thus specific to a particular typeface, the feature list in Table VI accounts for nearly all the typefaces in Figure 1. Usually, not the features themselves but the specific feature description for

Table VI. A set of distinctive features for Helvetica (Figure 2).

The features are divided into two sets: those concerned with a possible extended element above or below the central portion of the letter, and those concerned with just the central portion. No assumption is made about the order of the tests nor whether they are done sequentially or all in parallel (except that if the central portion has a gap on both the left and the right side [e.g., s or x] it is classified in the table for simplicity as left gap only; similarly for a top and bottom gap).

Features	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
Extender																										
Direction	–	U	–	U	–	U	D	U	–	D	U	U	–	–	–	D	D	–	–	U	–	–	–	–	D	–
Side	–	L	–	R	–	C	R	L	–	C	L	C	–	–	–	L	R	–	–	C	–	–	–	–	R	–
Hook	–	–	–	–	–	+	+	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–
Central Portion																										
Enclosed	+	+	–	+	+	–	+	–	–	–	–	–	–	–	+	+	+	–	–	–	–	–	–	–	–	–
Round	+	+	+	+	+	+	+	+	–	+	–	–	+	+	+	+	+	+	+	+	+	–	–	–	+	–
Side Gap	L	–	R	–	R	R	–	–	–	–	R	–	–	–	–	–	–	R	L	R	–	–	–	L	–	L
Top/Bottom Gap	–	–	–	–	–	B	–	B	–	–	T	–	B	B	–	–	–	B	–	–	T	T	T	T	T	–
Doubling	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	+	–	–	–

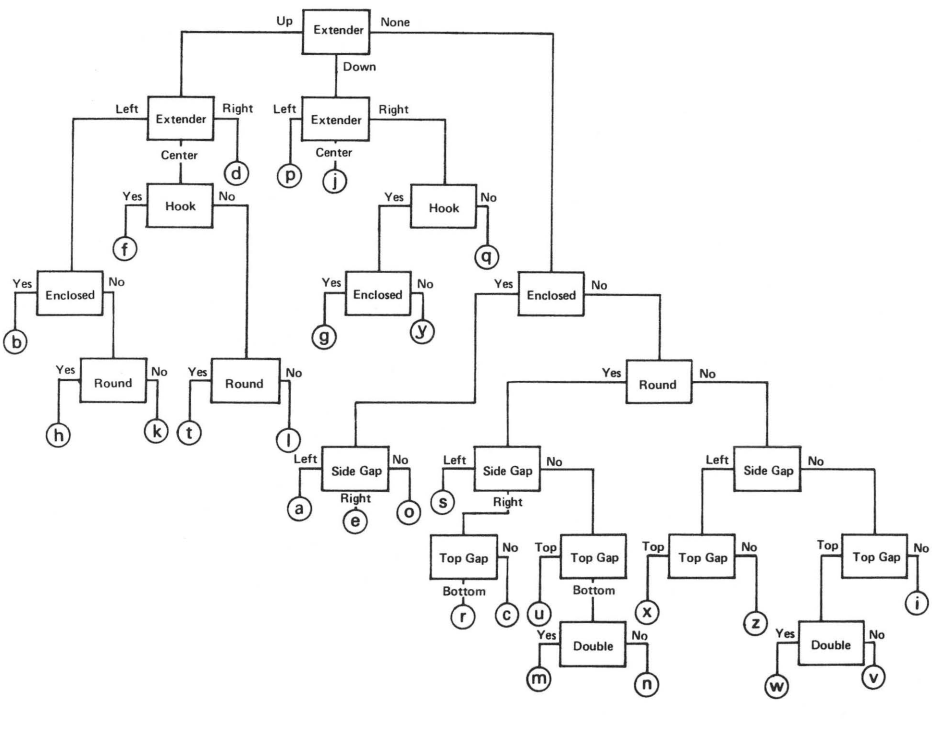
(1) *Extender Direction:* Does any part of the letter extend above, or extend below the central part (trinary test)? (2) *Extender Side:* If extended, is extended part to the left, in the center or to the right of the letter center (trinary test)? (3) *Hook:* If extended, is there a curved hook at extended end of extender? (4) *Enclosure:* Is any of the central part of letter completely enclosed? (5) *Round:* Is any of the central part of letter curved? (6) *Side Gap:* Is there a gap in the central portion facing to the left, to the right, or none (trinary test)? (7) *Top-Bottom Gap:* Is there a gap in the central portion facing up, facing down, or none (trinary test)? And (8) *Doubling:* For a pair of letters is one letter the double of the other?

Several arbitrary classifications are made: The letters i and j are treated as having no extended part above the center (with j of course, having an extension below), and the horizontal segment on the r is treated as rounded. If these decisions are reversed, slightly different tests are needed but the general shape of the flow chart remains.

single letters varies slightly from typeface to typeface. For example (see Figure 1), the h in line 1 differs from the h in line 2 in that the former adds a hook (top); the y in line 4 differs from the y in line 2 by the absence of a hook (bottom). In line 7, where u and v are both rounds (v is linear in line 2), u has a hook (bottom). Only where letters have two widely different but conventionally accepted shapes – such as the a's in the first two lines and the g's in lines 1 and 6 – may a feature need to be added. The serifs, which make the typeface in line 12 appear so different, are in fact irrelevant, and the same features which specify the letters in line 2 specify the letters in line 12 very well.

The feature set listed in Table VI is not the only one that could be constructed to discriminate among the letters in this particular typeface. If one started from a different pair of letters, it is possible that the resulting feature set would be somewhat different, in the description of each specific feature, and perhaps even in the minimal number of features needed. Therefore, no claims can be made for the set that it is unique, or even better than any other

Figure 3. A flow chart version of the distinctive feature tests illustrated in Table VI. While the flow begins with extended tests and then goes to central portion tests, the order could be reversed, or all done in parallel without changing the features. For ease of exposition, once a test results in a unique classification of a letter, no further tests are illustrated in the flow chart drawing.



set. It is a minimal set, however, in that if any feature is deleted from the set, some pairs of letters would be indiscriminable. Obviously, additional features could be added to the set, but these would be redundant in one of two senses: they would pertain to non-distinctive features (such as the dot on an i, or the crossbar on a t), or to features that are perfectly correlated with features already present in the set.

The particular feature tests used for this analysis can be done hierarchically, that is, in the order specified, so that some features can be considered of a higher order than others. But this particular feature set also works if tested in parallel or in any other order. What we do not know, of course, is how perceivers carry out feature tests. We expect that serial tests better characterize inexperienced or unpracticed readers of letters, and that skilled readers are both more likely to use parallel testing, or to abandon a feature analysis altogether for each letter, switching instead to much higher order combinations (see below).

The fact that another armchair crystal gazer might draw forth a different but adequate set of features characterizes one problem with this approach.

Furthermore, since redundancy is so prevalent in every aspect of language, there is every reason to expect that readers use a larger feature set than a minimal one. This means that there would be more than one set of feature tests that would result in the identification of each letter. Thus, individual readers may differ as to which configurations of specifying features they select, both between and within readers, as a function of typeface, immediate letter context, degree of attention, or mere whim. So even a "correct" minimal feature description of the letters may not correspond to what readers do; given the many possible combinations of features and contexts, readers may be highly variable in what they do.

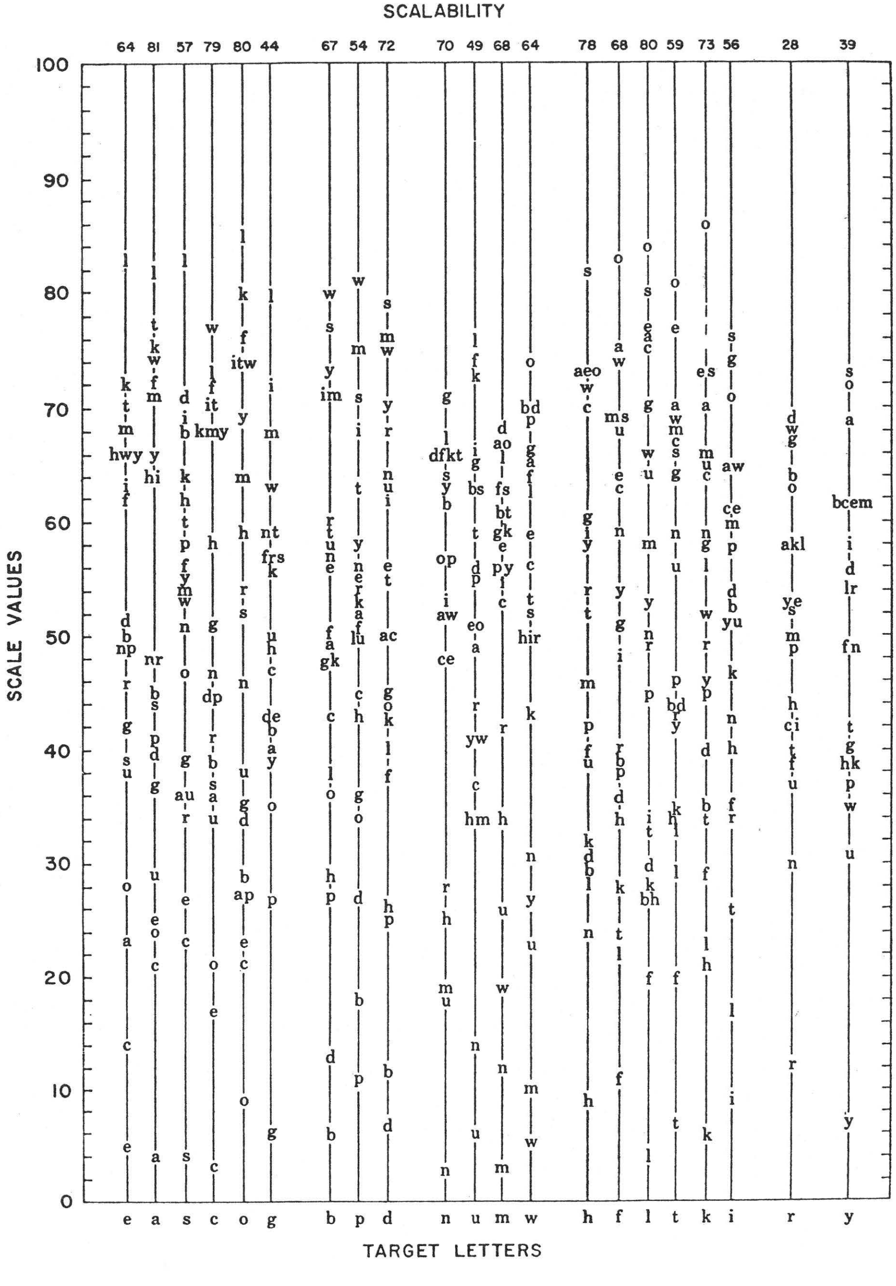
This kind of variation has been studied at the level of syntax and morphology in spoken language in contexts, where, as with letter features, the language permits the application of more than one rule. For example, in English, when forming the past tense of a verb, there are a number of possible rules besides adding the regular "ed." The verb "blink," past tense "blinked," exemplifies the regular rule. But "drink," past tense "drank," involves the application of a different rule, one also well-rooted in the language and used by speakers when creating new past forms. And "think," past tense "thought," represents another possibility from which the English speaker might generalize. Thus, if a new word is presented, such as "bink," speakers resort to at least the three rules just described. In fact, Haber (1975) found substantial variation within and among speakers in these circumstances, and used the term *muzzy* to relate the presence of multiple available rules to subsequent variation in usage.

The muzziness in the feature sets can result in substantially different performance rates across and within subjects, and this may make it more difficult to demonstrate that any particular feature system is the one being used at any one moment.

A second technique for studying letter features has been to create new alphabets and manipulate differences between letters to determine which differences are most salient. Gibson, Pick, & Osser, (1962) began this approach, using capital letter-like forms. While this work supports the notion that distinctive features might be a useful description of letters, it does not tell us much about which features are important for particular actual typefaces. The recent work by Shillman (e.g., Naus & Shillman, 1976) is an approach that may overcome this objection.

The third technique, like the fourth, uses real alphabets, and involves data on actual letters. The relative similarity judgement technique, best exemplified by Dunn-Rankin (1968, 1978), presents subjects with all possible triplets of letters. For each triplet, the subject is asked whether the first letter is visually more similar to the second or to the third letter of the triplet. From these judgements over a large number of subjects, a scale of similarity can be generated for each letter. Figure 4, taken from Dunn-Rankin (1968), is an example of scales for 21 letters (j, q, x, y, and f were not included), in which the lower the scale score, the greater the similarity. Thus, the letter e is most similar to e (fortunately), then to c, then to a. At the other extreme, e is least similar to l, k, and t. The rankings of the similarities for each letter can be intercorrelated, and that correlation matrix factor-analysed to determine if any groups of letters have related patterns of similarities. Dunn-Rankin (1968) found six factors, which have been grouped together along the bottom of Figure 4. Thus, this analysis suggests that the letters b, p, and d are very similar to each other, and quite different from all others. Similarly, n, u, m, and w form a cluster, and so forth. Dunn-Rankin did not name his clusters, nor did he speculate on what the underlying distinctive feature structure might be. In fact, trying to name the features from his results is no more objective than doing so from an armchair. Knowing that 315 subjects generated scales that place e next to a in the same cluster does not say anything about how many features they differ by, or what these features might be. Nor does it say that because e is in one cluster and l is in another, that e differs from l by more or more different features than does e from a. On the other hand, it is possible to correlate Dunn-Rankin's clusters with an armchair-derived set of features. If letters sharing clusters also tend to share more features, such a result would lend credence to the distinctive feature model. No one seems to have done this.

Figure 4. Similarity scales for 21 lowercase letters indicated along the bottom (omitting j, q, v, x, and z), in which the lower the scale score, the greater the similarity of a letter on each scale. The letters along the bottom have been grouped according to a factor analysis, indicating the six independent factors (data from Dunn-Rankin, 1968).



The last technique for generating a set of distinctive features uses the pattern of errors subjects make in a letter identification experiment to generate a clustering of similar and dissimilar letters. As one example that did not work, Coffin (1978) tried to determine if a spatial frequency analysis of letters would predict recognition accuracy and confusion. However, he found little support for spatial frequencies as the unit of visual information for letters.

In a more important study Bouma (1971) presented a brief flash of a single letter which the subject was to identify. Parameters were adjusted so that the average performance was 50% correct (chance being 4% if the subjects guessed randomly). Each response was entered into a stimulus-response confusion matrix, in which a correct response falls into the diagonal. The entries are the percent of responses given to each stimulus letter. Table VII is an example from Bouma in which the letters have already been grouped according to whether they are small or have extensions which ascend or descend.

Bouma used several procedures to extract feature information from these data. In one, he did a simple clustering, in which he took the reciprocal of the percent confusion between every pair of letters as an estimate of the perceptual distance between them, with small perceptual distances corresponding to highly similar letters. Figure 5 is an example of such a clustering, showing 12 groupings: aszx, eoc, mn, vw, u, rf, til, j, hkb, d, gq, and yp. These clusters do not appear to be very different from those of Dunn-Rankin, though neither offers a test of similarity to the other.

As a second procedure Bouma arbitrarily selected certain features on which the letters differed and looked to see if the results suggested that the subjects were using those features. For example, one feature he tested was the height of the letter, since all of the small letters were each less than 2 mm, while all of the tall letters were over 2.5 mm. If height is a feature, then when a small letter is presented, either the responses should be correct, or the erroneous response should be another small letter. The results showed that 46% of all small letters were correctly recognized, with another 38% of the responses being other small letters. Thus, small size has a discrimination value of 84%, since 84% of all responses to small letters were small letters. Similarly, tall letters should produce responses which are tall letters, and did so 92% of the time. Table VIII lists those feature tests that Bouma reported. It is possible that other features would have yielded values equally high if he had tried them, but since there is no guarantee in this method that all important features will be identified, it is only as good as the input provided.

The internal features of small letters were apparently not very useful (discriminate index = .38). This result may be due to Bouma's procedure. In order to cause the subjects to make errors, the letters were either presented at a great distance, or well off center of the fovea. Had he used a large foveal presentation, and shortened the duration, perhaps the internal feature differences of small letters would be more salient in his data. However, in order to construct a confusion matrix, subjects have to be induced to make errors, so any such procedure may bias the results in favor of some features over others.

Table VII. A stimulus-response confusion matrix of responses made to brief flashes of single letters (Bouma, 1971).

Each entry in a row is the percentage of times a particular response was given for the stimulus letter for that row – the sum of the entries in each row is 100%. The entries in the diagonal of the table are correct responses. The last column are responses of "don't know." Bouma arranged the table according to several groups of small letters, ascending letters and descending letters.

	Stimulus \ Response	a	s	z	x	e	o	c	n	m	u	r	v	w	d	h	k	b	t	i	l	f	g	p	j	y	q	-	
S S1	a	37	5	3	2	3	7	1	3	1	3	1			8	2			1	1								22	a
	s	15	10	1	2	10	10	3	4			1	2		4				1	3		1	1	1	1			30	s
	z	7	7	6	6	7	2	2			1	10	3		3				2	15		1		3	2			23	z
	x	5	7	6	35	3	5	2	5		4	3	2	1	2	1	1			1								17	x
S2	e	7	3	2	1	24	23	5	4		3		3	2	2		2						1				1	17	e
	o	2				3	58	4	12	2	6		1	1	2			1						3				5	o
	c	2		3		19	11	43	4	1	3	2	2	1									2	1			1	5	c
S3	n	6			2	2	4	2	64	6	2					6	1	1										4	n
	m	3		1		3	5		32	45	1			3	1		2											4	m
	u	6				2	7	2	2		54		4	9	2			1					1				5	5	u
S4	r			3	1	4		3				54	1						7	3		15		3	1	1		4	r
	v				1							5	86	2										2		1		3	v
	w				1					1	2		25	70														1	w
A A1	d	3	1					1							93			1							1				d
	h	2							2					1		72	5	17										1	h
	k	2	1	2	4				2		2		1			10	56	7	3	4								6	k
	b								1						1	16	3	79											b
A2	t	2						8				2	2						59	15	3	2		2			1	4	t
	i							1							1				5	66	22	2			3				i
	l											1							1	48	43				7				l
	f			1				1				6							9	4		79							f
D	g	1		1			2	1			1	3	14	1	1	1			1			2	21	2	1	4	40	3	g
	p											1							1					98					p
	j																			3	4				93				j
	y											10	4						8			7		10	4	57			y
	q														2				1				7			1	89		q
		a	s	z	x	e	o	c	n	m	u	r	v	w	d	h	k	b	t	i	l	f	g	p	j	y	q	-	

Distance Vision

Bouma's feature list could be generalized better if he provided either a rationale for deciding which features to test, or how large a discrimination is needed to be significant. He mentions 16 features that he tested, only 6 of which have discriminant indices in excess of .80. If a good armchair feature system were available for this typeface, Bouma's procedure could be used to test each feature in this way.

Until a test such as this is carried out, all of the procedures described are ad hoc, in that there is no guarantee that all potentially useful features have been isolated, nor any evidence that any of those features are actually used by readers. Presentation of a confusion matrix, or a list of similarity scale scores, tells us which letters are confusable or similar, but not in what ways. Therefore, even after a century of research on the visual characteristics of letters, we yet know little about what those characteristics actually are.

Luce (1969, 1973) describes a choice model which takes stimulus-response confusion data and extracts from them one parameter that reflects stimulus similarity and a second that reflects response bias. Townsend (1971) tested Luce's model against capital letter confusion data and found very strong support for the assumptions underlying Luce's model. When there is no multiple letter context, such as would be present with whole words, then the primary source of response bias would be letter frequency and perhaps – as a

Figure 5. Perceptual groupings found by Bouma (1971) made by plotting the perceptual distances between each letter based upon the reciprocal of the entries in Table VII above. Dashed lines have large perceptual distances and are not drawn to scale. All distances not indicated by drawn lines are large. Some arrows are curved only for ease of drawing.

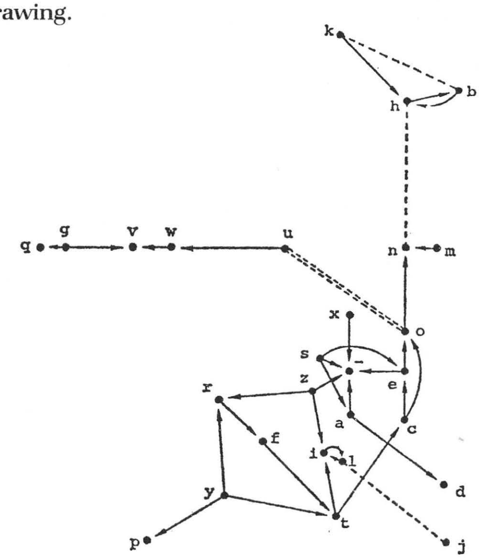


Table VIII. Upper estimates of the cue values of certain letter properties (upper cue value C).

Given the presence of a certain property in letter configurations, its upper cue value is obtained by taking the fraction of responses to which this particular property may have contributed as a perceptual cue. This fraction is the addition of the fraction correct letter responses and the fraction cue confusions, i.e., the fraction of confusions of property letters to cue letters. Data for distance vision and for eccentric vision have been averaged.

Property	Cue	Property letters considered	Cue letters considered	Fraction correct	Fraction cue confusions	Upper cue value C
(1) $H < 2{\cdot}0$ mm or $H/W < 1{\cdot}16$	Short letter	All S letters	All S letters	0·46	0·38	0·84
(2) $H > 2{\cdot}5$ mm or $H/W > 1{\cdot}16$	High letter	All A, D letters	All A, D letters	0·67	0·25	0·92
(3) Rectangular envelope	Rectangular letter (short)	S_1, S_3 letters	S_1, S_3 letters	0·40	0·25	0·65
(4) Presence inner part $H/W < 1{\cdot}16$	Something inside, short letter	/a s z x e m w/	/a s z x e m w/	0·38	0·22	0·60
(5) Circular envelope	Round	/e o c/	/e o c/	0·40	0·23	0·63
(6) Left part envelope circular	Left part round	/e o c/	/e o c d q ɑ g/	0·40	0·34	0·74
(7) Right outer gap + inner part	Right gap	/e/	/e c s z x r t f/	0·29	0·09	0·38
(8) Right outer gap no inner part	Right gap	/c/	/e c s z x r t f/	0·31	0·35	0·66
(9) Two vert. outer parts short letters	Double vert.	/n m u/	/n m u h/	0·59	0·18	0·77
(10) Lower gap 0·4 mm	Lower gap	/n/	/n m/	0·60	0·20	0·80
(11) Upper gap 0·5 mm	Upper gap	/u/	/u v w/	0·55	0·10	0·65
(12) Oblique outer parts	Obliques	/v w/	/v w/	0·75	0·17	0·92
(13) Left upper ext. 0·75 mm	Left upper ext.	/h k b/	/h k b/	0·66	0·23	0·89
(14) Right upper ext.	Right upper ext.	/d/	/d/	0·88	—	0·88
(15) $H/W > 1{\cdot}22$	Slender letter	/t i l f j/	/t i l f j/	0·65	0·25	0·90
(16) Upper dot	Dot	/i/	/i j/	0·67	0·03	0·70

second source of bias – an idiosyncratic tendency to produce certain letters when uncertain, irrespective of the stimulus. Townsend (1971) showed that Luce's model describes such biases. Unfortunately, since Townsend was more interested in Luce's model than he was in letter features, he does not speculate on what features might describe the particular capital letter alphabet he used.

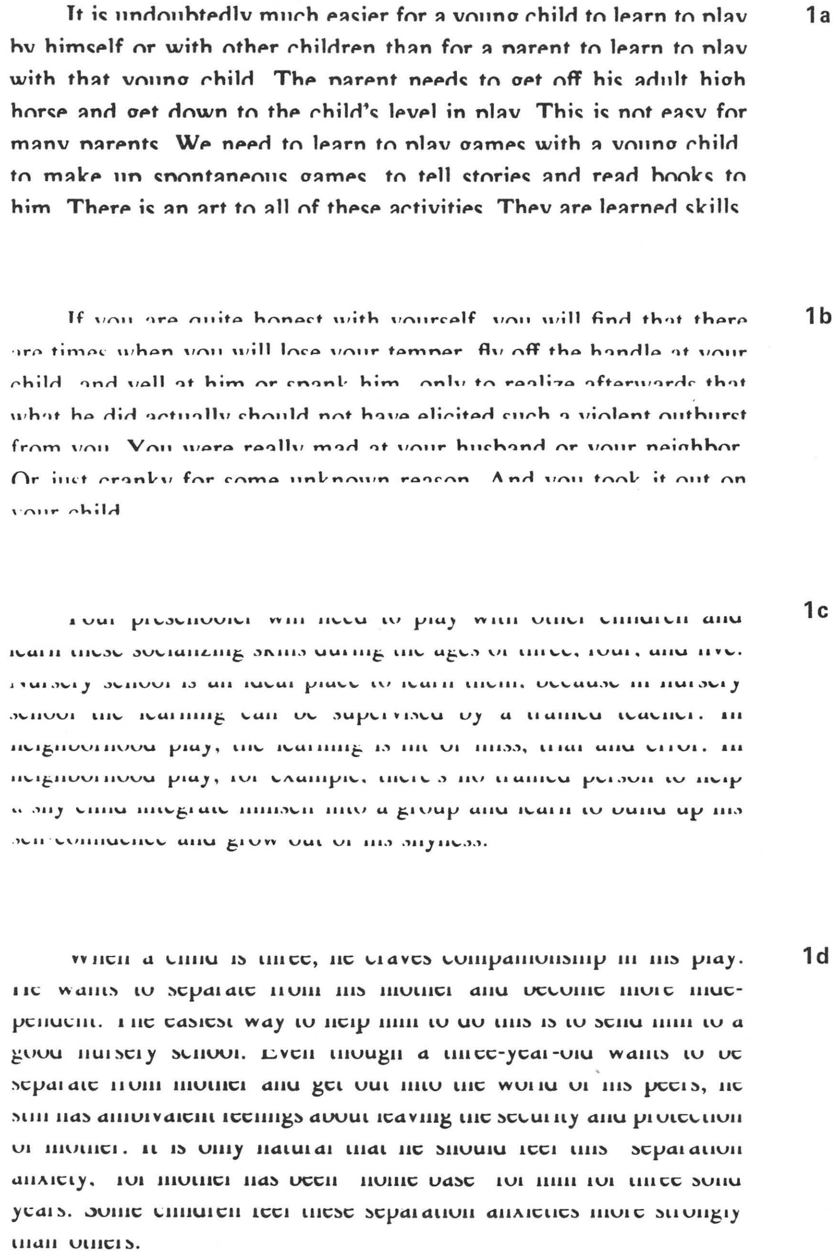
In most feature systems for English alphabetic characters, the ascender/descender/small letter difference is the basis of a distinctive feature, and intuitively it seems like an important one. Further, Huey (1908) suggested that the top portion of English letters are more informative than the bottom portion, in that deleting the tops interferes with reading more than does deleting the bottoms. Recently, Shimron and Navon (1980) tested this difference more carefully, and compared English to Hebrew letters (see Figure 6 for some examples). For English they found that reading speed decreased and oral reading errors increased more when the tops were removed as compared to the bottoms. Further, when the top 50% was removed (treating the middle as the horizontal middle of a lowercase o), performance was markedly worse than when only the top 25% was removed (75% reduction in speed, 200% increase in errors), whereas removing either the bottom 50%, or 25%, or nothing at all made little difference in speed or errors. They concluded that not only is there more information on the tops of English letters than on the bottoms, but that greater information is not confined to just the ascender portion of the letters, since all of the ascender was removed in both the 25% and the 50% deletion conditions.

They do not comment whether these results are consistent with a feature model of letter information. To say that there are more relevant features at the tops is circular, in the absence of a comparison of specific reading errors made in the top vs. bottom deletion conditions, with specific features being deleted. They did rule out a visual field or retinal sensitivity cause, since the opposite results were found for Hebrew letters – reading Hebrew is more affected by deletion of bottoms than tops of the letters.

Besides the importance of the vertical dimension, Kolers (1969) has shown that the right side of English letters is more informative than the left.

The preceding discussion has tried to highlight the procedures used in the research literature to isolate visual features of alphabetic characters. In spite of a fair amount of empirical sophistication, it is still not possible, with respect to any typeface, to specify the visual features of its letters. All we have are ad hoc descriptions that are themselves untested for internal consistency or generality. However, there are a number of tests in the literature of particular ad hoc feature systems, tests that show the system predicts or explains some performance measure. As one example, Haber and Cole (1981) tested whether

Figure 6. Four examples of printing changes from Shimron and Navon (1980) in which the bottom 25%, bottom 50%, top 50%, or top 25% of each row of letters is covered over.



similarity in visual shape or similarity in pronunciation predicted the reaction time to respond whether an uppercase and lowercase letter, presented side by side, were different. The reaction times for all possible different pairs were correlated with the number of articulatory distinctive features on which the pair differed (taken from the articulatory distinctive feature matrix for English consonants of Halle, 1962). For the visual features, Haber and Cole made up the feature set shown in Table IX. While this set bears some relationship to those of Dunn-Rankin and of Bouma, it is not identical, in part because it includes features for both lowercase and uppercase forms of every letter. In three independent versions of the experiment (using different sets of letters and presentation procedures), they showed that it takes significantly and correspondingly longer to say two letters are different according to the number of visual features they share (shared articulatory features had little effect). Given the replicability of this result, the authors expressed some confidence in the usefulness of the feature set they generated. It is still an ad hoc set.

Besides the lack of any internal validity for any of the feature systems proposed, a more serious problem arises when one tries to apply the system as it might be used in normal reading. Every experiment reported on visual feature descriptions presented the letters in isolation, without any context. The letter tested never appears as part of a word, nor the word as part of a sentence. Thus, the identification of any letter is based only on the visual information available to the subject plus any response bias he may have, bias

Table IX. A visual feature system for letters used in the experiment by Haber and Cole (1980).

Balanced = equal "energy" on left and right side of the letter.

	B	C	D	F	G	H	L	N	P	Q	R	T	V	Z	b	c	d	f	g	h	l	n	p	q	r	t	v	z
Balanced	0	0	0	0	0	1	0	1	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1
Large	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Line features:																												
Vertical line	1	0	1	1	0	2	1	2	1	0	1	1	0	0	1	0	1	1	1	1	1	0	1	1	0	1	0	0
1/2 Vertical line	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
Vertical line below semicircle	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0
Diagonal	0	0	0	0	0	0	0	1	0	1	1	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	2	1
Horizontal	0	0	0	2	1	1	1	0	0	0	0	1	0	2	0	0	0	1	0	0	0	0	0	0	0	1	0	2
Curve features:																												
Open semicircle	0	1	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	0
Closed semicircle	2	0	1	0	0	0	0	0	1	1	1	0	0	0	1	0	1	0	1	0	0	0	1	1	0	0	0	0
Semicircle left of vertical	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0
Curved line – top	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
Curved line – bottom	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0

that would tend to yield excess responses of high frequency letters. Many of the experiments do show a response bias effect – highly frequent letters get reported more often than low frequency ones. Thus, if capital letter Q is presented, O is a typical error, but Q is rarely an erroneous response to O, because subjects are unlikely to think of Q as a possible alternative, and report C or D more often. But what about response biases introduced by the subject's knowledge of spelling rules?

The spelling rules of English create restrictions among letter sequences. Hence, having identified an initial t in a real word, the next letter is restricted to a, e, i, o, u, y, h, r, w – no other letters do follow initial t in English. If all that is perceived is an ascending stroke, since only one of these alternatives has an ascending extension, it must be h, and there is no further need to try to discriminate among the set of ascenders: b, d, f, h, k, l, t. If a descender is detected, it must be y, and no further feature tests are needed. Even more importantly, the initial sequence th accounts for over 65% of all tokens of words beginning with t, so once an initial t is encountered, the great likelihood is that the second letter is h: it is much less likely to be a vowel or r or w, and impossible that it be any of the other 17 letters. Pursuing the example one step further, the extreme bias toward h following t is even greater in the context of a real sentence, probably approaching 95% when the syntactic structure of the phrase specifies that a function word is required. On the other hand, tw is more likely if the context makes a numerical value, or a word with a numerical root more likely (two, twelve, twenty, twin, etc.).

What effects do these contextual constraints have on how readers use visual features to read letters and words? We simply do not know, but we suspect that for at least skilled readers, inferences about feature testing based upon tests with isolated letters may be all wrong.

Four quite different kinds of feature testing suggest themselves. First, readers may use visual features to identify letters, but may vary which features they use depending upon the likelihood of which letters they expect to find. For example, suppose a reader expects to find the letter e. He may test for roundness and internal horizontal line, but not test such features as ascending/descending, or presence/absence of hook at all. If this is the case, then the relative importance of the different features, as estimated by isolated letter experiments, may not reflect at all which features are used at any one time. Second, readers may use visual features, and may even use the same set in the same order each time, but that order is determined not by the perceptual salience of the features, but by the ease or speed with which the various feature tests yield a unique response. For example, while it is easy to tell whether a letter is z or x on the one hand, or q or g on the other, it would be silly to perform this test, since those four letters rarely occur; it rarely would add new information about any particular letter in question.

Third, it is possible that in the absence of any context, readers do test individual features one by one, but that as soon as the letters appear in context, then the tests involve groups or sets of highly correlated features. If there is no information about what the letter might be, then one feature may be as good as the next. But as soon as the letter choice is restricted, then several features could be tested together. While several feature tests must be made to separate h from all other letters, there is no reason why the test cannot be for a vertical line including a downward open bottom, rather than treating those as two separate features. Of course, if the test fails, the reader learns little from the negative answer, but if h is quite likely, he won't fail very often and will succeed more rapidly or more accurately with the combined test.

Fourth, when reading in context, especially where the syntax and semantics strongly suggest the upcoming word, readers may test for clusters of letters such as th, or even whole words, such as the. In this case, reader performance would not be predicted on the basis of a feature description of individual letters at all (see the last section of this paper).

It appears to us that while reading models must make some reference to letter processing, we know very little about what features comprise letters; how those features are extracted during letter identification; and whether those feature extraction processes are stable or variable as a function of local spelling contexts, or even larger contextual chunks. We are still a long way from definitive answers to any of these questions.

Nor are we any farther in considering alternatives to a feature principle in general. One possibility is that children learn their letters by a feature extraction and processing procedure, much as the more mechanical models of concept formation suggest for any set of arbitrarily defined concepts. But once learned, letters are perceived and identified in terms of their match to learned prototypes, in which the comparison process makes little reference to features. Rosch (1975) provides substantial evidence for such a concept performance model. Without arguing that her model applies to letter identification as well, it does seem likely that skilled reading performance may not depend upon the same kind of feature testing that was initially needed to learn the letters in the first place. The insights offered by Schneider and Shiffrin (1977) and Shiffrin and Schneider (1977) on the differences between controlled and automatic processing of perceptual tasks should be applied here.

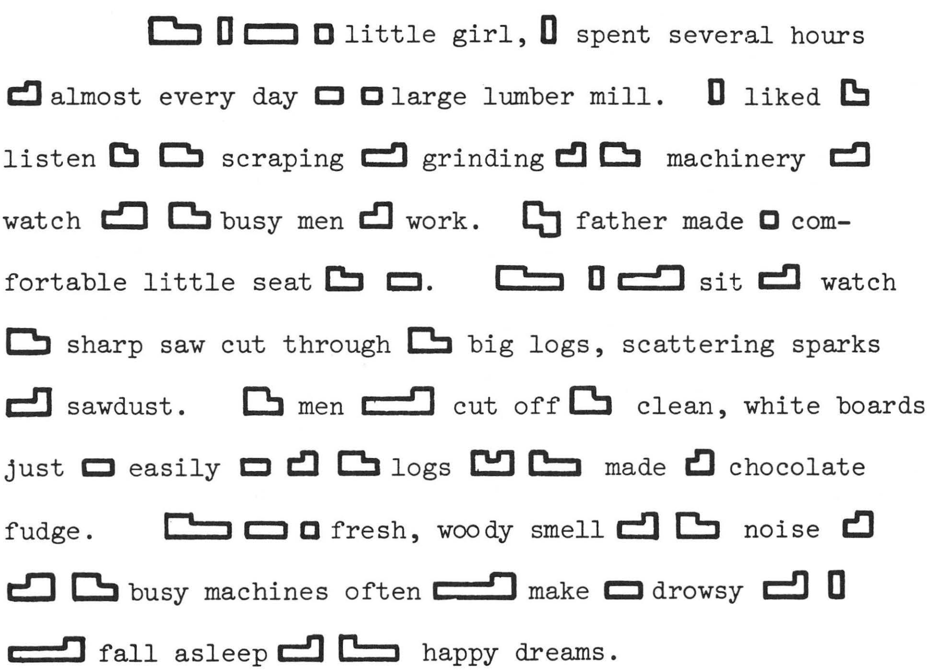
Visual shape of the configuration of entire words

While it seems obvious that readers are aided by the arrangements of the letters on the page, and that readers identify letters on some kind of basis of extraction of their features, it is less obvious that the shape alone of an entire word may be useful in specifying its meaning, without the identification of any of its letters. In this last section, we shall review the evidence about this third source of visual information in reading.

Figure 7 illustrates a brief passage of prose in which each of the very high frequency words, those in excess of 1000 tokens per million in printed text (Carroll, Davies, & Richman, 1971), has been replaced simply by an outline drawn around the extremities of each letter. While about half of the words in this passage have been so altered, it is not particularly difficult to read or understand. Informal evidence such as this shows that the shapes of the whole words, themselves, apart from the individual letter features, provide some visual information that readers could use when reading. However, we do not yet know much about the nature of that information.

Figure 8 shows several examples of word shape features. The first row illustrates shape based only on the overall outline. This is the system used in

Figure 7. A brief story in which every instance of every high frequency word (those whose tokens occur at least 1000 times per million words of text) has been replaced by its outline shape.



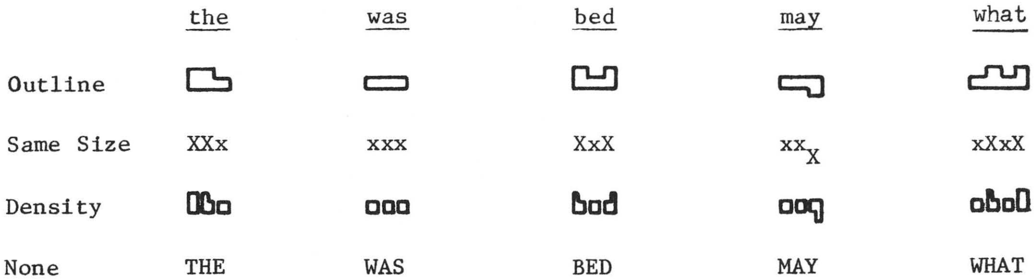
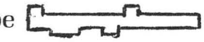


Figure 8. Four different ways in which the shape of words might be represented. For the words indicated along the top, shape is specified by an outline drawn around the word as a whole, by representing the height of each letter separately, and by the height, width and center of gravity of each letter separately. The last line removes all height information for word shape.

the passage in Figure 7. The second keeps each letter separate, so that shape is given by the letter sizes, not the outline envelope. The third example includes density within the letter as well as the letter sizes. Thus the adjacent letters db look quite different from bd when their density is taken into account, though they would be equivalent by either the outline or size criterion. All of the research has used the first two word shape procedures, with none exploring whether there is any usable visual information in letter density.

The last line of Figure 8 represents a system with little word shape information, such as occurs when text is presented in all capital type. Since it is easy to read text in all capitals, and we do so frequently, outline shape is not a necessary source of information. Both Smith (1969) and Fisher (1975) compared reading times for text printed in normal and in all capital type. They found a 5 to 10 percent difference in speed favoring the mixed case type, suggesting some usefulness for word shape. Since readers have substantial familiarity with the single alphabetic characters in their capital cases, changing to all capitals removes only the overall configuration information about the shape of the word. It is a more easily interpreted manipulation than that of case alternation, where not only is familiar shape information removed, it is replaced by drastically new and unfamiliar shape information. A 5 to 10 percent reduction in reading time attributable to word shape alone suggests the importance of this variable. That the reduction in reading time is not larger probably attests to our familiarity with reading all capital text, so that we have learned to depend upon other sources of information when word shape is missing.

Every word in the language could be described by its shape, but shape would not be expected to be equally useful as a source of information for all words. While only one word in English fits the shape  , few readers normally learn that shape. Nevertheless, given its great length and unusual shape, readers can readily reject nearly all words in English as possibilities from this shape alone, and if the story is about animals, especially jungle animals, the shape may be sufficient for a unique identification even if a hippopotamus has not appeared in the story previously. Therefore one use of word shape information for low frequency words might be the elimination of possible incorrect alternatives. Further, if the word has appeared in the story before, subsequent identification of such unusual words might be done entirely on the basis of overall shape.

Even a very general perception of the outline shape of a low frequency but expected word may enable the reader to identify it correctly. Here even the word shape may be picked up as approximate length and ascender/descender pattern, without attention to the specific pattern of the word. In this case the specific kinds of cues to which individual readers attend (length, where changes in overall letter shape occur in the word) may vary within and among readers, just as we suggested with letter feature extraction.

Quite a different case has been made for word shape information of highly frequent words. For some highly familiar and frequent words, the reader might be able to make unique identification of the word from its shape alone – for example, only one word fits the shape  (by) or the shape  (the). Since readers have enormous exposure to and practice on the very high frequency words of their language, they certainly can learn the shape of these words and therefore be able to respond to that shape directly.

Context is also a critical factor for high frequency words, but for a very different reason than described for low frequency words. As we show below, while a number of high frequency words often share the same shape, context so restricts the occurrence of many of these words that when word shape plus context is considered, the word is uniquely specified by its shape.

Failure to appreciate the importance of context in processing word shape has been a source of some theoretical controversy. For example, Groff (1974, 1975) notes that even for the very frequent words in English (about the first 200), i.e., those expected to be most aided by shape features because of their great familiarity, only about 25% of them have a unique outline shape (such as the or by) not shared with any other word among those top 200. How then, Groff asks, could shape be a useful cue to word identity, when it uniquely specifies only 25% of the words? And how can shape be useful for those instances where 12 of the top 200 words all have the same shape? When this argument is coupled with the evidence that we can easily read text devoid of word shape information – that is, in all capital type – many theorists have relegated word shape to the useless bin.

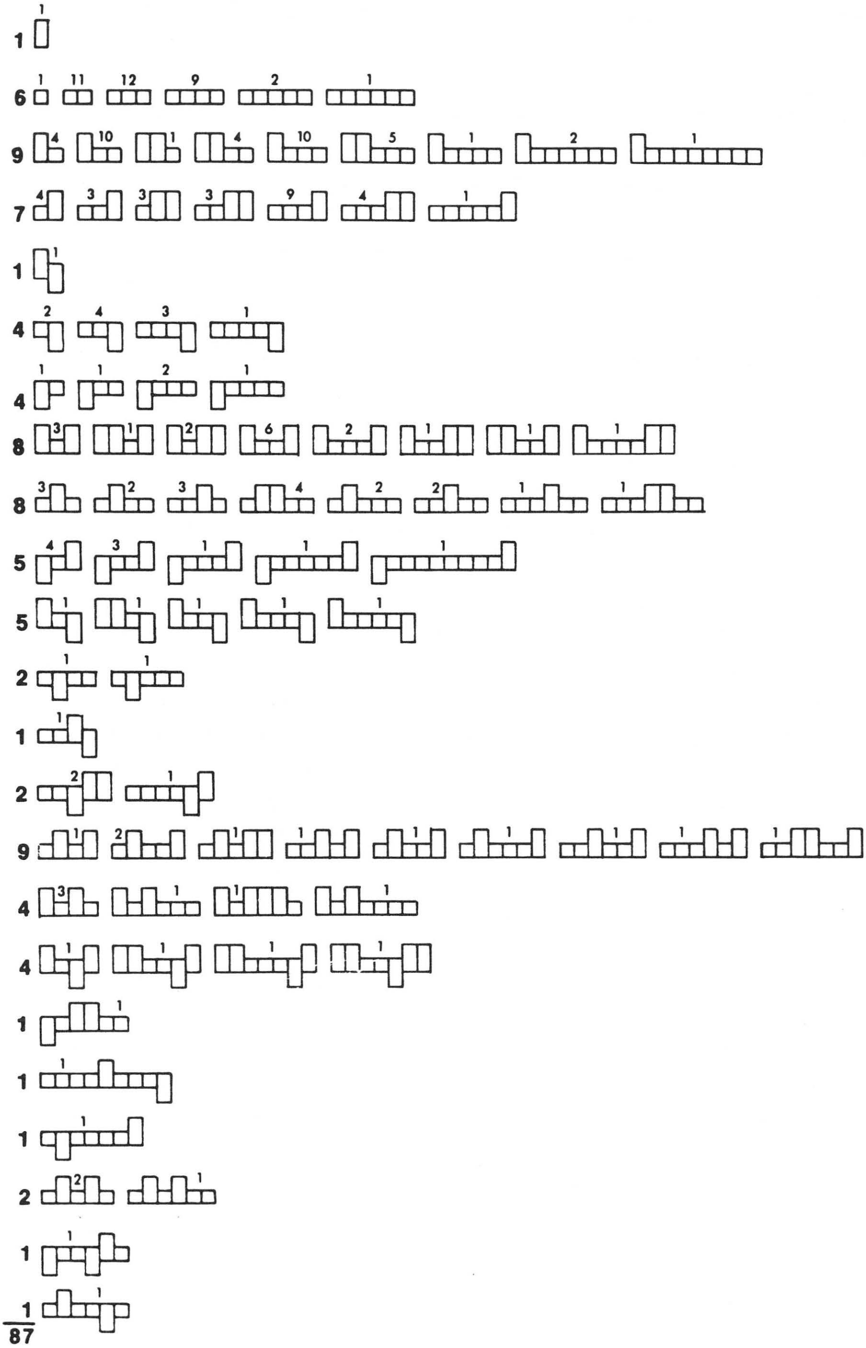
To demonstrate why Groff's argument is faulty, by showing the importance of word shape when considered in context, Haber and Haber (1981) took a typical story from a second grade basal reader and examined the shapes of each of its very high frequency words, those drawn from the 200 most frequent words in English (Kucera & Francis, 1967). In theory these words should have the greatest chance of being learned by their shape alone, because of their great frequency. Just these 200 words account for 66% of all the tokens found in printed English; thus two-thirds of all the words that appear in typical texts are on this list. Figure 9 illustrates the 87 different outline shapes that describe the lowercase printed versions of these words. They are arranged in the figure according to their patterns of ascender/descender and small letters. Thus, the second row has all the patterns of only small letters; and the third row all the patterns that begin with tall letters and end with small letters.

When we examined our target story, we found, as expected, that 25% of the story's frequent words had a unique shape, but the remaining 75% had a shape which was shared with other highly frequent words – counting all the frequent words, each one had a shape that could be one of 4.1 possible words from the top 200. To demonstrate that context helps reduce the number of alternatives below 4.1, we asked 18 adult judges to make context acceptibility judgements among all of the high frequency words with similar word shape in each position in the story. The judge looked at a card which contained on it all of the story up to its first high frequency word. In place of that word, the card listed all the words from the high frequency list that shared the same shape. This number ranged from 1 to 14, with a mean of 4.1. The judge circled all words from the list that were acceptable in the story up to that point. Then he looked at the next card, which had all of the story from the beginning on up to the next high frequency word, and so forth. The mean number of circled words was 1.4, and for over 80% of the words, only one alternative was circled by each judge (the same word in every case). While in isolation, only 25% of the high frequency words have a unique shape, in context over 80% of them are unique in shape.

Thus, the argument that because few high frequency words have unique shapes, word shape is a useless source of information, makes sense only when words are considered in isolation. Since reading always has a context, and context so extensively restricts word choice, shape certainly has the potential of providing valuable information about the identity of a word.

Several kinds of research evidence have appeared recently on the impor-

Figure 9. All 87 different outline shapes needed to represent the 200 most frequent words in English. They are arranged according to the different patterns of ascending, small and descending letters. The superscript above each shape indicates the number of different words that share that shape (from Haber & Haber, 1980).

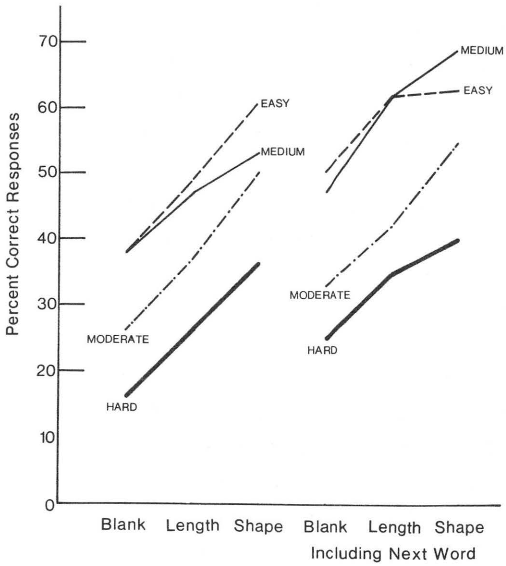


tance of word shape in reading, one type demonstrating that it provides useful information, another exploring the types of words benefiting most from word shape information, and a third concerning the relation between eye position and the locus in the text in which word shape is useful. Rayner (1978) has reviewed much of this literature up to 1977, and we describe several more recent studies below.

As an example of the first type of study, in which the usefulness of word shape is shown, Haber, Haber, and Furlin (1981) asked college subjects to read passages which appeared on a CRT screen in segments. Initially the first few lines appeared, ending in the middle of the second or third sentence. The subject had to guess what he thought the next word in that sentence would be. After his guess, some more text appeared, again ending in the middle of a sentence, and the subject had to guess the next word. This procedure provides a baseline estimate of how much information the subject can extract from context in the absence of any information about the next word. For some passages, the subject was also provided with the number of letters in each word to be guessed, specified by providing an X for each letter. For other passages, the subject was given the length of the word and its outline shape as well.

The results strongly show the usefulness of the outline shape information. Guesses were more accurate (both in exactly matching the correct word, and in matching its meaning and syntactic function) when the subjects knew the shape of the word as compared to prior context alone or only length, and this was true regardless of the difficulty level of the text (see Figure 10). Further, when the subjects did make an error, their erroneous responses more often matched the word in shape when shape information was provided, again indicating that readers use such information. Finally, the part of speech of the word to be guessed interacted with the shape manipulation. Long and relatively low frequency common nouns did not significantly increase in accuracy over the context-only condition when their shape was provided, whereas accurate guesses of the short highly frequent function words (e.g., determiners, prepositions, auxiliary elements of the verb phrase) increased greatly when shape was indicated. This study, therefore, provides several lines of evidence that readers are able to utilize word shape information in extracting meaning from text.

The second kind of recent study on word shape explores the relation between frequency, word length, and word class, and the usefulness of word shape. As an example of this kind of study, Haber and Schindler (1981) used a proofreading task to determine if word shape is an equally useful source of information for all types of words. College subjects read long passages for



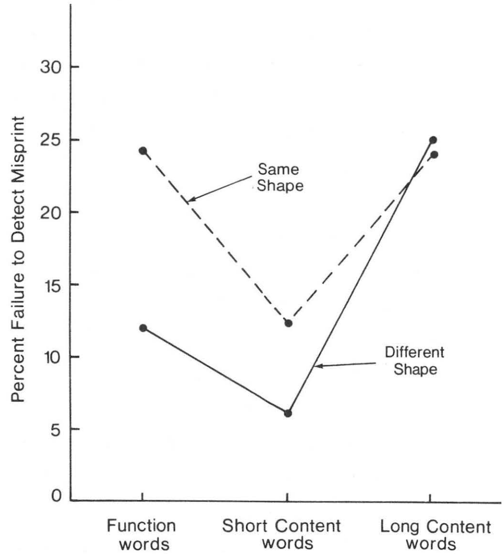


Figure 10. Accuracy of guesses of the word in the blank, when that word was only preceded by a full context of the story (blank), or in addition had its length specified, or in addition had its outline shape specified. Right hand panel is for conditions in which the word after the blank also was present. Data are for four difficulty levels (data from Haber, Haber, & Furlin, 1981).

Figure 11. Failures to detect misprints in proofreading tasks, as a function of the kind of word (function word, short content, or long content word) and whether the misprint preserved or altered the original shape of the word (data from Haber & Schindler, 1981).

comprehension (on which they were carefully tested) but were also asked to mark any misprints they noticed. Only 25 misprints per 1000 words of text were present, all of which were misspelled to become a nonword. Some were short function words, some short content words, both of very high frequency, and some longer and lower frequency content words. Half of the misspellings substituted an incorrect letter for a correct one of the same size, thereby preserving outline word shape, whereas in half the substituted letter was a different size, thereby changing word shape. Figure 11 presents the results, in which several findings are apparent. Word shape has a much larger effect on the function words, less on the equally short content words, but no effect at all on the longer and less frequent content words. Thus for the shorter, more familiar words, some of the information readers use is from the outline shape alone, and not exclusively from each of the letters, so that when a misprint still resembles the correct word, the misspelling is not noticed nearly as often. As in the Haber, Haber and Furlin study, word shape information is used most in the identification of the familiar function words, and plays little if any role in decoding less familiar longer content words.

The third approach to word shape is best exemplified by Rayner (see 1978). Several of his studies concerning the locus of word shape information have shown that word shape information is picked up farther from the center of the current fixation than is letter shape information. In one study Rayner (1975) monitored the reader's eye movements and changed a single word in a sentence when the eyes came within a predetermined distance to that word. He found greater disruption in the pattern of eye movements, indicating that the reader had already picked up some information from the about-to-be changed word, when the change altered word shape than when word shape was preserved, and this effect accrued farther to the right of the current eye position than did changes in spelling that did preserve word shape. Thus, the shape of a word is picked up farther away than is the identity of its individual letters.

Presumably the printing conventions, especially those of spacing between words (specifying word length) and those specifying phrase and sentence boundaries, are picked up even farther to the right in each fixation, but we know of no evidence yet on that.

This is by no means an exhaustive review of word shape research, though it is representative. More such research is badly needed, both to study the different kinds of shape information conveyed by words, and to determine when such information becomes important in the acquisition of the reading process. Virtually nothing has appeared on the different kinds of shape, and only a few studies on acquisition of shape information in learning to read (see Rayner & Hagelberg, 1975).

These studies give no indication of how important a source of information is word shape to the normal skilled reader. They do suggest that shape can potentially provide information and that readers do use shape information in processing text. Further research is needed to show which aspects of shape are used when, and how the use of word shape varies according to contextual constraints. Further, all of the above discussion and research has focused on whole-word shape. It is possible that there are part word shapes that are learned and used as visual cues, such as the initial th of so many function words.

We would no more wish to make a case for word shape as an important source of information in isolation than we would for letter shape in isolation. It is much more likely that readers use word shape – shape here encompassing both the entire shape of a word or of some of its parts – in combination with other kinds of information, both visual and non-visual, provided by context.

Look back at Figure 9. The words in each of the 23 lines contain a similar configuration, either all small letters, small followed by ascender(s), etc. If word shape is defined literally, then the shapes within each line differ in length: a single small letter differs from two small letters, for example. Perhaps, however, when a functor is highly expected, readers test looser hypotheses about word shape: is the target word composed of all small letters? Is an ascender followed (at some point, unspecified) by a descender? Such a more general definition of word shape might account for how readers verify low frequency outsized words like hippopotamus. Just as it seems unlikely that the reader who expects a hippo tests each letter, it is reasonable to expect a more general testing of overall rather than specific shape: here comes a long word, or I see an embedded ascender roughly where I expect.

Another possibility concerns the Shimron and Navone (1980) results (Figure 6). In using word shape, readers may attend more to ascenders than to descenders, and especially to the configurations where shapes change (as where a small letter follows an ascender), focusing just above the middle of letters. Thus, alterations in shape just above the center of small letters may be the most important aspect of word shape.

These variables: where vertically the reader attends to shape, whether he attends to configurations or where shapes change, and whether he perceives specific shapes or more general envelopes, all need to be explored – singly, in combination, and above all, in context. We expect their combination will prove to play an important role in the reading process, and that word shape, like print arrangement rules and letter shapes, is a major component.

In conclusion, we have reviewed work on three areas of visual information that are available to readers: information provided by the conventional arrangement of the print on the page; information provided by the features of each letter; and information provided by the shapes of entire words. While intuition as well as some empirical evidence suggest overwhelmingly that all three of these sources are important, the empirical work is either quite incomplete, or still insufficient to support much confidence in identifying the precise nature of this visual information or how it is extracted. This lack is particularly noticeable with respect to visual processing of single letters.

Since all theories or models of reading depend in part or in total on the specification of visual information extraction from the page as a whole, words as a whole, or individual letters, in order to make progress in testing and improving these theories, it is critically important to further our knowledge of visual processing.

REFERENCES

Bouma, H. Visual recognition of isolated lower case letters. *Vision Research,* 1971, *11,* 459-474.

Browning, R. My Last Duchess. In H. Spencer, W. E. Houghton & H. Barrows (Eds.), *British literature from Blake to the present day.* Boston: D. C. Heath, 1952.

Carroll, J. B., Davies, P., & Richman, B. *The American heritage word frequency book.* Boston: Houghton-Mifflin, 1971.

Coffin, S. Spatial frequency analysis of block letters does not predict experimental confusions. *Perception and Psychophysics,* 1978, *23,* 69-74.

Cole, R. A., & Jakimik, J. A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1980.

Cummings, E. E. Portrait. In R. P. Warren (Ed.), *Understanding poetry.* New York: Henry Holt, 1950.

Dunn-Rankin, P. The similarity of lowercase letters of the English alphabet. *Journal of Verbal Learning and Verbal Behavior.* 1968, *7,* 990-995.

Dunn-Rankin, P. The visual characteristics of words. *Scientific American,* 1978, *238,* 122-130.

Fisher, D. F. Reading and visual search. *Memory & Cognition,* 1975, *3,* 188-196.

Fowles, J. *The French Lieutenant's Woman,* New York: New American Library, 1969.

Frase, L. T., & Schwartz, B. J. Typographical cues that facilitate comprehension. *Journal of Educational Psychology,* 1979, *71,* 197-206.

Gibson, E. J., Gibson, J. J., Pick, A. D., & Osser, H. A. A developmental study of the discrimination of letter-like forms. *Journal of Comparative and Physiological Psychology,* 1962, *55,* 897-906.

Groff, P. The topsy-turvy world of sight words. *Reading Teacher,* 1974, *27,* 572-578.

Groff, P. Research in brief: shapes as cues to word recognition. *Visible Language,* 1975, *9,* 67-71.

Haber, L. R. The muzzy theory. In R. Grossman, J. San, & T. Vance (Eds.). *Proceedings and papers from the eleventh annual meeting of the Chicago linguistic society.* Chicago: University of Chicago Press, 1975.

Haber, L. R., & Haber, R. N. Visual processes in reading. In F. J. Pirozzolo & M. C. Wittrock (Eds.), *Neuropsychological and cognitive processes in reading.* New York: Academic Press, 1981.

Haber, L. R., Haber, R. N., & Furlin, K. R. Word length and word shape as sources of information in reading. Under editorial review, 1981.

Haber, R. N., & Cole, R. A. Evidence for the direct visual access to letter identities. *Acta Psychologica,* 1981, *46,* in press.

Haber, R. N., & Haber, L. R. The shape of a word can specify its meaning. *Reading Research Quarterly,* 1981, *16,* in press.

Haber, R. N., & Schindler, R. Errors in proofreading: evidence of syntactic control of letter processing? *Journal of Experimental Psychology: Human Perception and Performance,* 1981, *7,* in press.

Halle, M. Phonology in a generative grammar. *Word,* 1962, *18,* 54-72.

Hartley, J. Spatial cues in text. *Visible Language,* 1980, *14,* 62-79.

Hartley, J. Introduction to special issue: Spatial arrangement of text. *Visible Language,* 1981, *15,* 4.

Hochberg, J., Levin, H., & Frail, C. Studies in oral reading: VII: How interword spaces affect reading. Mimeograph, Cornell University, 1966.

Huey, E. B. *The Psychology and Pedagogy of Reading.* New York: Macmillan, 1908. Reprinted Cambridge: MIT press, 1968.

Jakobson, R., Fant, C. G. M., & Halle, M. *Preliminaries to speech analysis: the distinctive features and their correlates.* Cambridge: MIT Press. Original edition 1951; sixth printing 1965.

Kolers, P. A. Clues to a letter's recognition: Implications for the design of characters. *Journal of Typographic Research* (now *Visible Language*), 1969, *3,* 145-168.

Kolers, P. A., & Perkins, D. N. Orientation of letters and errors in their recognition. *Perception & Psychophysics,* 1969, *5,* 265-269.

Kolers, P. A., & Perkins, D. N. Spatial and ordinal components of form perception and literacy. *Cognitive Psychology,* 1975, *7,* 228-267.

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

LeCarre, J. *Tinker, tailor, soldier, spy,* New York: Bantam Books, 1975.

Luce, R. D. *Individual choice behavior.* New York: Wiley, 1959.

Luce, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.) *Handbook of mathematical psychology,* Vol. I. New York: Wiley, 1963.

McNamara, W. G., Patterson, D. G., & Tinker, M. A. The influence of size of type on speed of reading in the elementary grades. *The Sight Saver Review,* 1953, *23,* 28-33.

Naus, M. J., & Shillman, R. J. Why a Y is not a V: A new look at the distinctive features of letters. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, *2,* 394-400.

Neisser, U. *Cognitive psychology.* New York: Appleton Century Crofts, 1967.

Poe, E. A. The sleeper. In R. P. Warren (Ed.) *Understanding poetry,* New York: Henry Holt, 1950.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. Asymmetries in the perceptual span for Israeli readers. *Brain and Language,* in press.

Rayner, K. The perceptual span and peripheral cues in reading. *Cognitive Psychology,* 1975, *7,* 65-87.

Rayner, K. Eyemovements in reading and information processing. *Psychological Bulletin,* 1978, *85,* 618-660.

Rayner, K., & Hagelberg, E. M. Word recognition cues for beginning and skilled readers. *Journal of Experimental Child Psychology,* 1975, *20,* 444-455.

Rosch, E. H. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General,* 1975, *104,* 192-233.

Schneider, W., & Shiffrin, R. M. Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review,* 1977, *84,* 1-66.

Shiffrin, R. M., & Schneider, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review,* 1977, *84,* 127-190.

Shimron, J., & Navon, D. The distribution of visual information in the vertical dimension of Roman and Hebrew letters. *Visible Language,* 1980, *14,* 5-12.

Smith, F. Familiarity of configuration vs. discriminatibility of features in the visual identification of words. *Psychonomic Science,* 1969, *14,* 261-262.

Tinker, M. A. *The legibility of print.* Minneapolis: University of Minnesota Press, 1954.

Townsend, J. T. Theoretical analysis of an alphabetic confusion matrix. *Perception and Psychophysics,* 1971, *9,* 40-50.