

VISIBLE LANGUAGE

*The research journal concerned with all that is involved in our
being literate*

Volume XV Number 2 ISN 0022-2224

Special double issue

Visual Cues in Word Recognition and Reading

Guest editor: Keith Rayner, Ph. D.

Published in two parts

● **PART ONE / Volume XV Number 2**

- 125 **Visual Cues in Word Recognition and Reading: Introduction**
Keith Rayner
- 129 **Visual Factors and Eye Movements in Reading**
Robert E. Morrison and Albrecht-Werner Inhoff
- 147 **Visual Components of the Reading Process**
Ralph Norman Haber and Lyn R. Haber
- 183 **Toward a Computational Theory of Early Visual Processing in Reading**
Michael Brady
- 216 **The Authors**

PART TWO / Volume XV Number 3

- 219 **Children's Word Recognition in Prose Context**
Susan F. Ehrlich
- 245 **Disorders of Reading and Their Implications for Models of
Normal Reading**
Max Coltheart
- 287 **Word Processing in Reading: A Commentary on the Papers**
Arnold D. Well and Alexander Pollatsek
- 309 **The Authors**
- 311 **General Information**

Visible Language, Volume XV, Number 2, Spring 1981. Published quarterly by
Visible Language, Box 1972 CMA, Cleveland, OH 44106. ©1981 by Visible Language.
Second-class postage paid at Cleveland, Ohio, and at additional mailing offices

Merald E. Wrolstad, Ph. D., Editor and Publisher
P. O. Box 1972 CMA, Cleveland, OH 44106

Advisory Board

Colin Banks, Banks and Miles, London
Fernand Baudin, Bonlez par Grez-Doiceau, Belgium
George Bauer, University of Southern California
Pieter Brattinga, Form Mediation International, Amsterdam
Murray Eden, National Institutes of Health, Bethesda
I. J. Gelb, Oriental Institute, University of Chicago
Kenneth S. Goodman, University of Arizona
James Hartley, University of Keele, England
Albert Kapr, Hochschule für Grafik und Buchkunst, Leipzig
Paul Kolers, University of Toronto
Alexander Lawson, Rochester Institute for Technology
C. L. Lehman, Tigard School District, Oregon
Aaron Marcus, University of California, Berkeley
Dominic Massaro, University of Wisconsin, Madison
Alexander Nesbitt, Newport, Rhode Island
Thomas Ockerse, Rhode Island School of Design
G. W. Ovink, Amsterdam
P. David Pearson, University of Minnesota
Sharon H. Poggenpohl, University of Kansas, Lawrence
Marvin A. Powell, Jr., Northern Illinois University
Wayne Shebilske, University of Virginia
Mary Ellen Solt, Indiana University
Jack W. Stauffacher, The Greenwood Press, San Francisco
Robert St. Clair, University of Louisville
William C. Stokoe, Jr., Gallaudet College, Washington
Michael Twyman, University of Reading
Richard Venezky, University of Delaware
Dirk Wendt, Christian-Albrechts-Universität, Kiel
Michael Wood, University of Aberdeen
Bror Zachrisson, Grafiska Institutet (Emeritus), Stockholm
Hermann Zapf, Darmstadt, Germany

Composition by Creative Composition, Inc., Ashland, Ohio.
Printed by McNaughton & Gunn, Inc., Ann Arbor, Michigan,
on papers from French Paper Company, Niles, Michigan.

Crème Blanc 65-pound cover

Tex-O-Set natural white 70-pound text

General Information

Visible Language is concerned with research and ideas that help define the unique role and properties of written language. It is a basic premise of the Journal that writing/reading form a distinct system of language expression which must be defined and developed on its own terms. Published quarterly since 1967, *Visible Language* has no formal organizational affiliation. All communications should be addressed to

Visible Language
Box 1972 CMA
Cleveland, OH 44106 USA
Telephone 216/421-7340

Subscription Rates

	<i>One Year</i>	<i>Two Years</i>	<i>Three Years</i>
Individual subscription	\$15.00	\$28.00	\$39.00
Institutional subscription	25.00	47.00	66.00

Foreign subscribers: add \$1.00 for postage to each year's subscription.

All orders must be prepaid. To be honored free of charge, claims for missing issues must be made immediately upon the receipt of the next published issue.

Back Copies and Reprints

A folder listing the contents of all past Journal issues is available on request. Individual reprints are not available. A limited quantity of all back numbers is available at a per issue cost of \$3.00 to individuals and \$5.00 to institutions.

Manuscripts

Manuscripts, inquiries about research articles, and other contributions to the Journal should be addressed to the Editor. A guide for the organization, preparation, and submission of manuscripts is available. Three copies of a manuscript (including the original) should be accompanied by an abstract typed on a separate sheet of paper.

A Note from the Guest Editor

In this special issue of Visible Language I have attempted to bring together some integrative statements concerning visual factors in the reading process. It is my hope that the issue will be of interest to graphic designers, educators, computer scientists, and psychologists.

An effort such as this requires extra efforts on the part of many people, and I would like to thank the authors for agreeing to prepare their contributions. I would particularly like to thank my colleagues, Arnold Well and Sandy Pollatsek. They read each of the original versions of the papers in the volume and provided detailed comments to the authors on ways to revise the papers. They then undertook the rather difficult task of providing a commentary on the papers in an attempt to tie together all of the different themes and ideas running through the contributions. I believe that they have been able to provide a valuable analysis of the points made by the authors. In addition to Well, Pollatsek, and myself, the papers were all read by at least one other reviewer. I would like to thank Paul Kolers (University of Toronto), George McConkie (University of Illinois), Rod Barron (University of Guelph), and Dick Olson (University of Colorado) for so willingly agreeing to serve as reviewers for the original versions of the manuscripts. Finally, I would like to thank Merald Wrolstad for providing this opportunity for us to prepare this volume and for taking responsibility for the graphic design of the issue.

Keith Rayner

Visual Cues in Word Recognition and Reading: Introduction

Keith Rayner

Reading is a fascinating process to investigators in a wide range of fields: from experimental psychology, to education, to computer science and artificial intelligence, to linguistics, to graphic design. The ease with which most of us read tends to obscure the fact that it is a very complicated skill. However, even casual observation of illiterate adults as they attempt to acquire the skill makes it very clear that mastering reading is not at all a trivial process. Learning to read is also a time consuming (and often difficult) activity for young children. Despite widely different pedagogical strategies and techniques and often in spite of the variability that exists in the characteristics of beginning reading instruction, most children in literate societies seem to have pretty well mastered the skill by about the fourth or fifth year of schooling. Many children in our society show some retardation in reading ability. However, in many cases this merely represents a statistical and measurement artifact. That is, given the assumptions of standardized testing procedures, a certain percentage of children will be classified as reading one or two years below grade level. In many cases the poor reading will also be correlated with low intelligence scores. There will remain in each school system a certain percentage of children who have no obvious emotional or physical impairment and have normal intelligence scores, yet who have difficulty learning to read. Some of these children will eventually improve their reading performance. Some of them will not and will always have difficulty reading; such readers are often referred to as being dyslexic.

The study of reading as a process has had a very checkered and cyclic history. Within the field of experimental psychology, for example, there was considerable activity and research on the basic processes involved in reading between 1880 and 1920. The list of psychologists interested in the reading process around this time reads like a *Who's Who in Psychology*: Cattell, Pillsbury, Dodge, Dearborn, Huey, and Woodworth were among the most

Visible Language, XV 2, pp. 125-128.

Author's address: Department of Psychology, University of Massachusetts, Amherst, MA 01003.
0022-2224/81/0040-125\$02.00/0 ©1981 Visible Language, Box 1972, Cleveland, OH 44106.

prominent psychologists of their time. Most of the work they carried out is summarized very nicely by Huey (1908) and Woodworth (1938). Unfortunately, the prevailing doctrines and theories in experimental psychology underwent drastic and radical change in the early 1900's. This was due mainly to the emergence of the behaviorist position in American psychology. Since the basic tenet of the behaviorist revolution was that only observable behavior could be studied, basic research on reading virtually ceased because the most interesting questions concerning reading are related to what's going on inside the head and the answers must be inferred rather than directly observed. Thus, it's not at all surprising that when Huey's book was republished in 1968, it was still vitally relevant. In the last ten years or so, researchers in experimental psychology have again become very interested in a number of questions concerning basic processes involved in reading and learning to read.

The rise of the behaviorist position forced many researchers who were interested in reading to move their laboratories from psychology departments to schools of education. Since education is inherently more interested in applied issues, much of the research on reading that continued after 1920 was very applied in focus. However, it was during this period that Buswell and Tinker carried out their work on the relationship between eye movements and reading. Kolars (1976) has reviewed some of the major findings of Buswell's work and in this volume Morrison and Inhoff review many of Tinker's findings with regard to the effects of various print variables on eye movements and reading. Chall (1967) has summarized much of the work within the field of education between the early part of the twentieth century and 1967 concerning the various types of reading programs and their effectiveness.

One of the things that is very obvious in considering visual factors involved in reading is that there has often not been very much communication between researchers interested in reading and typographers and graphic designers who set text to print. Crouwel (1979) makes the point that it was not until the 1920s that the idea was accepted that a text could be made more comprehensible and more functional by deliberate emphasis on form. But Crouwel also clearly points out that changes were not the result of research findings, and he generally makes a plea for more communication between research workers interested in print and text, and graphic designers responsible for the layout of the text.

Of the various disciplines mentioned here, computer science and artificial intelligence represents the newest entry into the general field of reading. Artificial intelligence has become a field to be reckoned with in regard to many human endeavors that require cognitive or perceptual processing. There has been considerable interest in computer analysis of scenes and of handwriting.

Occasionally, there have also been forays into computer recognition of words (Hanson, Riseman, and Fisher, 1976). Brady's paper in the present volume represents a current approach to more directly interface the fields of computer science and psychology with respect to reading.

In short, the study of reading is now attracting workers from a number of disciplines. In the present volume we have put together a series of papers that we believe will be of interest to psychologists, educators, computer scientists, and graphic designers. Inasmuch as the topic of the volume concerns visual factors in word recognition and reading, there is not a contribution by a linguist. Likewise, there is not a contribution by a graphic designer, but certainly the first two papers in the volume are directly relevant to members of that community. In the first paper Morrison and Inhoff have described the visual factors that are important in terms of the eye movements that occur during reading. The first half of their paper focuses on how the alignment and printing of the text influence eye movements and readability. In the second half they look at more recent work on perceptual factors influencing eye movements. Haber and Haber present an account of how certain visual cues in text, particularly word shape, may influence the reading process and more generally they discuss the interaction of printing conventions and the knowledge the readers have stored in their heads. Brady's paper represents an attempt to integrate psychology and artificial intelligence in understanding reading. Ehrlich's paper specifically addresses the issue of learning to read and how certain cues are utilized to recognize words. Coltheart then presents a detailed overview regarding reading disability and a general model of where the process of reading may break down. Finally, Well and Pollatsek provide a critique of the different papers presented in the volume and they also provide a general integration concerning how visual factors influence reading and word recognition.

REFERENCES

- Chall, J. S. *Learning to read: The great debate*. New York: McGraw-Hill, 1967.
- Crouwel, W. Typography: A technique for making a text legible. In P. A. Kolers, M. E. Wrolstad, and H. Bouma (Eds.). *Processing of Visible Language 1*. New York: Plenum Press, 1979.
- Hanson, A.R., Riseman, E.M., & Fisher, E. Context in word recognition. *Pattern Recognition*, 1976, 8, 34-45.
- Huey, E.B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908. Republished by MIT Press, Cambridge, Mass., 1968.
- Kolers, P.A. Buswell's discoveries. In R.A. Monty and J.W. Senders (Eds.). *Eye movements and psychological processes*. Hillsdale, N.J.: Erlbaum, 1976.
- Woodworth, R.S. *Experimental psychology*. New York: Holt, 1938.

Visual Factors and Eye Movements in Reading

Robert E. Morrison
Albrecht-Werner Inhoff

The effects on oculomotor behavior of variations of physical attributes of text, both typographical and psychophysical parameters, and similar effects from physical (visual as opposed to semantic) word cues processed in the reader's parafoveal vision are discussed. It is concluded that the reliability of such effects underscores the primary nature of visual discrimination in reading. Furthermore, sets of data generated from typographically different stimuli suffer a serious confounding.

Psychologists have long been fascinated by the movements of the eyes in reading and have examined readers' oculomotor behavior in an attempt to elucidate the processes underlying fluent reading. The class of oculomotor acts observed during reading are well known. The eyes execute a series of jerky (saccadic) movements followed by pauses during which the eyes are held relatively stable (fixations). Since stimulation obtained during a fixation suppresses the activity due to stimulation obtained during the saccadic movement (Matin, Clymer, & Matin, 1972) it is only during fixations that the reader is able to abstract information from text. Fixations average around 200-250 msec and consume 90-95% of the total reading time (Tinker, 1965). Saccades that are executed in a forward direction (i.e., left-to-right for English readers) serve the function of bringing new information into foveal vision.¹ On the average, each forward saccade covers about 8-9 character spaces (Rayner, 1978). Saccades which are executed to a text location that has already been read are referred to as regressions and account for 10-20% of the saccades (Rayner, 1978). Right-to-left saccades that intervene between the final fixation of a line and the initial fixation of a new line are return sweeps. Frequently, the return sweep under-shoots the beginning of a new line and the reader performs a compensatory regression to the beginning of the line after a relatively short initial fixation

Visible Language, XV 2, pp. 129-146.

Morrison's address: Department of Psychology, University of Massachusetts, Amherst, MA 01003.
0022-2224/81/0040-129\$02.00/0 © 1981 Visible Language, Box 1972, Cleveland, OH 44106.

(Rayner, 1978). The return sweep can also be inaccurate in a vertical direction, causing fixation of the wrong line and a disruption of reading (Tinker & Paterson, 1955). Forward saccades, regressions, and return sweeps are voluntary eye movements and are ballistic movements whose trajectories cannot be influenced once the movement has been initiated.

Although the average values mentioned above are repeatedly reported, the oculomotor behavior in reading shows a considerable amount of variability, both within and between readers. Fixations for individual readers may range from 100 to more than 500 msec, saccade size may range from less than 1 character space (microsaccades) to about 20 character spaces, and the percent of regressions can vary from 0 to 40% (Huey, 1908; Rayner, 1978; Woodworth, 1938). Distributions of fixation duration and saccade length for college-aged readers are given in Figure 1.

There is accumulating evidence that the variability in the different eye movement measures is not due to random fluctuations of a "noisy" oculomotor system. Comprehensive reviews of eye movement data in reading (Huey, 1908; Rayner, 1978; Tinker, 1965; Woodworth, 1938) repeatedly stress systematic changes in oculomotor behavior occurring as a result of text (e.g., typographical factors, text difficulty) and/or reader characteristics (e.g., reading ability). There is evidence that the effects are not the result of a general tuning of the oculomotor system to global text characteristics or motivation levels. Rayner and McConkie (1976) reported no correlation between successive fixation durations or between the size of successive saccades as one would expect if the reader's eye behavior reflected a general tuning. Also, there was no correlation between fixation duration and subsequent saccade size, a result also reported by Walker (1938).

This suggests that the eye movement patterns are sensitive to the moment-to-moment processing activities of a reader (Rayner, 1977; Rayner & McConkie, 1976). Since visual inspection of the text is an obligatory part of reading, the perceptual salience of text should also affect eye behavior on a moment-to-moment basis. Indeed it has been well demonstrated that static characteristics of the text and reading situation (character size, contrast, illumination) and physical characteristics of what is available in parafoveal vision (word length, beginning letters) can alter the eye movement pattern of a reader (Rayner, 1979; Tinker, 1963). In the following pages we will proceed to analyze the effects of visual factors on eye movements in reading.

Effects of typography

Between 1927 and 1965 Miles Tinker and his colleagues conducted comprehensive investigations of the physical characteristics of the reading situation which affect reading speed (Tinker, 1963; Tinker, 1965). Tinker explored the

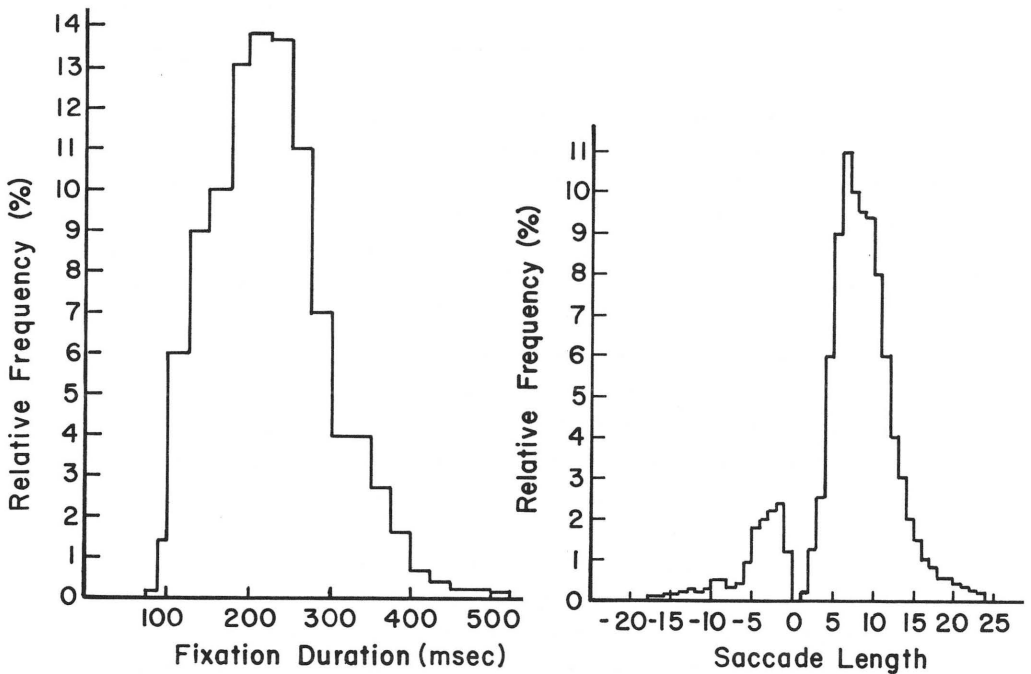


Figure 1. Frequency distributions of fixation duration (left) and saccade length for eight college-age readers. Return sweeps of the eye have been excluded from the distribution. Short fixations following the return sweep which are followed by corrective saccades have also been excluded. Reprinted with permission from Rayner and Inhoff (in press). Copyright The Chemical Rubber Co., CRC Press, Inc.

effects of typeface, type size, line width, leading (i.e., the amount of blank space between successive lines of type set text), illumination, contrast, color, and many other variables (see Tinker, 1963, for a bibliography of this work).² By using a check on reading comprehension to maintain a constant level of high comprehension from subjects, Tinker was able to use speed-of-reading as an operational definition of legibility of text. Legibility values were reported as percent changes in reading speed from that achieved with some standard in any given experiment. Comparisons of the magnitude of these values are valid only between conditions within the same experiment, not across experiments. In addition, readers' eye movements were photographed to determine what oculomotor changes were occurring to cause text to be read faster or slower. As Tinker's data is based on experiments involving over 30,000 subjects, it remains some of the most extensive research on the subject of legibility and how eye behavior is modified by different text parameters. His findings will be summarized in the remainder of this section.

Tinker argued that more legible (more quickly read) text allows readers a wider perceptual span which requires fewer fixations and allows longer saccades. The perceptual span is defined as the amount of textual material which the reader effectively processes during an individual fixation (Paterson & Tinker, 1947). He divided the words or characters constituting a text by the number of fixations made to obtain an average measure of the perceptual span (in words per fixation or characters per fixation). Although recent evidence (McConkie & Rayner, 1975; Rayner, 1975) argues against the assumption underlying this computation, namely that the perceptual spans of successive fixations are contiguous and do not overlap; the assumption that a more legible text allows for a wider perceptual span and generally requires fewer fixations is not disputed.

Hence, although Tinker's words per fixation measure is not to be taken literally, we assume it bears a monotonic relationship to the mean width of the perceptual span. Furthermore, mean saccade length should also be strictly correlated with mean perceptual span size, and thus bear a monotonic relationship to the words per fixation measure. That is, when fewer fixations are made (more words per fixation) they must be more widely spaced, on the average (longer mean saccade length).³ Reviewing Tinker's data we shall attempt to infer the effects of legibility manipulations on mean saccade length.

Tinker asserted that fixation duration is a function of the ease of discrimination of the characters within the fovea. Reading speed then is primarily a function of: fixation frequency – the direct correlate of perceptual span and saccade length – as affected by parafoveal discriminability; and fixation duration, as affected by foveal discriminability. (Regression frequency and duration also influence reading speed.)

The authors could find no mention by Tinker of an attempt to rigidly control viewing distance nor of a need to in order to get reliable rankings of relative legibility. Presumably subjects selected the distance which most comfortably allowed them to perform the comprehension task Tinker used which involved crossing out an inappropriate word. For any subject this distance would be fixed with a head rest in Tinker's eye movement monitoring apparatus but was probably allowed to vary between subjects.

Although changes in viewing distance might be critically important if letters to be identified were at threshold size (for that distance), it doesn't seem that it would have an overwhelming effect in a reading task in determining the relative legibilities of type sizes that are all well above threshold size (when presented foveally). Parafoveal acuity limitations on the perceptual span would seem to be the major source of legibility differences. Of course there is an inverse relationship between viewing distance and retinal image size. Re-

ducing the distance increases retinal image size, but this means a constant unit of print, such as a word, extends into more eccentric retinal areas. Because of the linear decrease in acuity with increasing eccentricity (e.g., Anstis, 1974) small changes in viewing distance may have little effect on the perceptual span. (See Anstis for a demonstration of stimuli which "have the unusual property of being equally readable at any viewing distance" [1974, p. 590]). In any event, Tinker apparently found that relative legibilities of different typographies were reliable over moderate changes within the range of normal reading distances (i.e., within arm's length) and his conclusions generalize to such situations accordingly.

One of the most interesting facts discovered by Tinker is that factors which make alphabetic stimuli, both single letters and words, most readable in a number of threshold determination paradigms (for example distance, eccentricity in peripheral vision, or exposure duration thresholds) do not make them most legible in a normal reading situation. In fact, they may actually reduce legibility of text. For example, all-capital letters have a greater threshold distance than lowercase letters of the same nominal point size because they are actually larger. But they are less legible (reading speed decreased by 11.8%) in a reading situation (Paterson & Tinker, 1947; Tinker, 1963). ALTHOUGH FIXATION DURATIONS WERE SHORTER WHEN READING TEXT IN CAPITALS THAN WHEN READNG LOWERCASE (FROM A MEAN OF 230 MSEC TO 210 MSEC), 12.4% MORE FIXATIONS WERE NECESSARY ON THE CAPITALS AND TOTAL READING TIME WAS INCREASED. MEAN CHARACTERS PER FIXATION WENT FROM 8.1 FOR LOWERCASE TEXT TO 7.0 FOR CAPITALIZED TEXT. SINCE REGRESSION FREQUENCY DID NOT CHANGE, WE INTERPRET THIS AS A DECREASE IN MEAN FORWARD SACCADE LENGTH (IN TERMS OF CHARACTERS TRAVERSED). SURPRISINGLY, THE SACCADES WERE PHYSICALLY LARGER. SINCE CAPITALIZED TEXT OCCUPIES A 35% GREATER PAGE AREA, THE MEAN PICAS PER FIXATION INCREASED FROM 3.2 TO 4.3 (A 34% INCREASE IN PHYSICAL LENGTH), YET FEWER CHARACTERS OCCUPIED THIS DISTANCE. APPARENTLY THE PERCEPTUAL SPAN IS DECREASED FOR UPPERCASE TEXT EVEN THOUGH THE HORIZONTAL EXTENT OF THE SACCADES (IN TERMS OF DISTANCE OR VISUAL ANGLE) INCREASED IN THIS SITUATION. THE LARGER SIZE OF THE CAPITALS MAKES THEM PERCEPTIBLE TO A GREATER ECCENTRICITY IN PERIPHERAL VISION BUT NOT BY ENOUGH TO EQUAL THE INCREASED VISUAL ANGLE THAT A GIVEN CHARACTER STRING WILL SUBTEND. CAPITALIZED TEXT ALSO LACKS WORD SHAPE CUES WHICH COULD CONTRIBUTE TO THE DIMINISHED PERCEPTUAL SPAN.

A similar trend was seen with very large type (14-point versus a 10-point standard). Reading speed, hence legibility, was reduced by 6.4%. Though mean fixation duration was decreased for the larger type (from 230 to 210 msec), fixation frequency increased by 21% (regression frequency did not change significantly). Note that the size manipulation is not confounded with word shape – both sizes are in lowercase. Assuming the larger type would be perceptible to a greater eccentricity in peripheral vision, it is conceivable that saccades would cover more picas to traverse an equal number of characters. This was not so. Again, as with capitalized text, the mean picas per fixation were greater for the larger type (4.9 vs. 3.9), yet characters per fixation decreased (8.5 vs. 10.1). Readers did make physically somewhat longer saccades but these traversed fewer characters. There are two possible explanations for this, however. The study had used both 10-point and 14-point type in 19-pica line widths. The larger type had fewer characters in a line. Perhaps this constrained the perceptual span and not type size per se.

The study was replicated, this time keeping constant not the line width but the number of characters per line. Although the effect was diminished, the 14-point type was still less legible (by 3.3%). Apparently the 10-point type was inherently more legible than the larger (14-point) type – a rather unintuitive finding considering research on recognition thresholds for single letters. There seems to be a trade-off between size of type and decreasing parafoveal acuity which makes a moderate size most legible. At some point, the greater perceptibility of the larger type is offset by the further extension into areas of poorer parafoveal acuity by a given character string. It is also clear that any factors which affect legibility – type size, line width, etc. – interact with each other such that all variables not under experimental manipulation must be controlled in order to draw valid conclusions about independent variables.

Legibility is also reduced by very small type. Comparing 6-point type with a 10-point standard, legibility decreased by 5.8% when printed in constant 19-pica line widths and 6.0% when equated for number of characters per line. Eye movement records showed that when reading the smaller type total fixation frequency significantly increased even though the text occupied less area in this case. Regression frequency increased only slightly; thus the frequency of forward fixations (total fixation frequency minus regression frequency) increased. This implies by our previous reasoning that mean saccade size decreased, in this case both in terms of distance (from 3.5 to 2.7 picas) and characters per fixation (from 9.0 to 8.5). In addition, mean fixation duration was significantly increased for the 6-point type (240 msec) compared with the 10-point standard (230 msec). Thus, foveal processing took longer due to reduced perceptibility of the small characters, and the perceptual span was slightly reduced.

There has existed some uncertainty in the literature as to whether it is more appropriate to talk about saccade length in absolute terms of visual angle or in terms of character spaces (O'Regan, 1980; Rayner, 1978). Tinker's data, showing that a physically longer saccade can cover fewer characters of a less legible text arrangement (i.e., one with larger type) suggests that characters per saccade may be the more useful measure, for a couple of reasons. First, it is invariant over changes in viewing distance for saccades covering equivalent chunks of information; visual angle is not. Second, it is also invariant, for such equivalent-chunk saccades, over texts of different typographies, allowing comparisons between them. Consider the same text printed in two different type sizes. As noted above, more visual angle may be covered by saccades while reading the text of larger type. But without intervening calculations we don't know whether they cover more or fewer characters than the comparatively shorter eye movements made on the smaller sized type. It is not directly evident from a consideration of visual angle whether the perceptual span has increased or decreased with the large type. In addition, mean characters per saccade data will indicate the number of fixations made on a passage (of n characters) without having to adjust mean visual angle per saccade by the number of characters per degree of visual angle at a certain viewing distance. Of course, the number of characters per degree of visual angle in a study should always be reported anyway.

Tinker also determined the optimal line width, in picas, for various type sizes. Generally, a range of widths proves equally legible for a given type size. Larger type sizes are most legible in long lines, smaller types are most legible in short lines.

For 10-point type the optimal line widths ranged from 17 to 27 picas, and performance peaked at 19 picas, which results in a line of approximately 52 characters. Surprisingly, no constant ratio of optimal line width to type size was found. This means that the optimal line width for all type sizes did not result in an equal number of characters per line. An analysis of eye movements revealed why lines shorter and longer than the optimum for a given type size were less legible. For example, for 10-point type, a short line (9-picas) increased fixation frequency by 15.7% over a 19-pica standard.⁴ Tinker concluded that short lines, having fewer characters per line, prevent the full utilization of non-foveal vision. Mean fixation duration increased by 20 msec (to 240) as well. A common pattern observed in the reading of lines of optimal width is a decrease in mean fixation duration over the

last half of a line, presumably because the text has been glimpsed repeatedly in parafoveal vision during earlier fixations on the line (Dearborn, 1906; Tinker & Paterson, 1955). Text in short lines prevents this. Seemingly then, the longer the lines are, the better. However, a different factor decreases legibility for line widths beyond the optimum. Regression frequency dramatically increases (by 56.7% for 10-point type in a 43-pica line compared to a 19-pica standard). With very long lines Tinker found an increased incidence of compensatory regressions at the beginning of the line to correct undershootings and “inaccuracies” in executing the return sweep—one on practically every line! (Paterson & Tinker, 1940). Besides undershooting on lines that are too long, inaccuracies in the vertical plane can cause the wrong line to be fixated. Smaller sized type should be more susceptible to these errors, by demanding a more precise (vertically) return sweep, and restrict the upper limit on optimal line width more. In fact, this is what Tinker found—larger type has optimal width with longer lines than smaller type. In addition to more regressions, text in line widths longer than the optimum show slight yet significant increases in fixation duration (10 msec). The locus of this effect is also the trouble spot at the beginning of a line (Paterson & Tinker, 1940). Therefore, the only problem with long lines is an inability to make accurate return sweeps, and the corresponding disruption of reading.

Leading, or blank space between lines, has the general effect of expanding the acceptable range of line widths for any type size (Tinker, 1963). The increased blank area above and below words probably reduces lateral masking⁵ and widens the perceptual span. Although eye movements were not recorded for the experiments on leading, short lines are probably read with fewer fixations of shorter mean duration than when the text is set solid (without any leading). The improved legibility of very long lines with leading probably results from a less drastic increase in regression frequency. Presumably vertical inaccuracies on the return sweep are more tolerable when lines are spaced farther apart. Tinker found one or two points of leading to be optimal.

Since all the typographical features of a text interact in determining legibility, definitive experiments on the relative legibilities of type sizes necessitates each size being printed in its empirically determined optimal line width and leading. When this was done there were no significant differences between 9, 10, 11, and 12 point type. Smaller sizes, 6 and 8 point, were still less legible (Tinker, 1963). Unfortunately, sizes larger than 12 point were not used. Tinker (1965) recommends 10, 11 or 12 point type.

⁵10-point type, 43-pica measure, set solid and leaded gradually 1, 2, 3, and 4 points.

Effects of contrast and illumination

Brightness contrast between page and print can drastically affect speed-of-reading. Legibility is high if the reflectance of the page is 70% or greater and the reflectance of the characters is 10% or less. Less brightness contrast than this can severely reduce legibility. Tinker investigated the effect of colored prints and backgrounds and found that legibility differences for these stimuli were strictly a function of the brightness contrast between print and page. Color, per se, had no effect. Combinations with low brightness contrast, such as red print on green, were extremely illegible (e.g., there was a 39.5% drop in reading speed). Low contrast typographies increased fixation duration, regression frequency, and total fixation frequency.

Illumination is much less critical than might be expected, provided that legible text of adequate brightness contrast is being read. Tinker states, "Lighting specifications based upon visual acuity or visibility data overestimate the requirement for reading" (1963, p. 251). Illuminating engineers recommend 70 footcandles of illumination for reading (cited in McCormick, 1976, p. 323), but much less is sufficient, provided the reader is adapted to that level. Tinker found that subjects prefer illumination equaling that to which they are momentarily adapted. When adaptation was controlled, his subjects read 10-point type at maximum speed with only 3.1 footcandles of illumination. Actually, very strong illumination may be detrimental if it increases the effects of any glare already present. It is more important that the illumination be well diffused to avoid these effects (Tinker, 1963).

Tinker (1963) recommended illumination (20-30 footcandles) above the limits he empirically determined to be necessary, in order to provide a "margin of safety." Although there may be no decrement due to low illumination when typography is optimal, any non-optimal typographical feature, even one which alone does not cause a significant decrease in legibility, will have a combined effect in the presence of other conditions of marginal legibility, resulting in a significant drop in legibility. Furthermore, combinations of conditions which do individually reduce legibility will have a pooled effect, resulting in extreme drops of legibility – by increasing fixation frequency, fixation duration, and regression frequency. The combined decrement is less than the sum of the individual decrements.

There are other influences upon oculomotor behavior in reading such as text difficulty and changes in reading strategy (e.g., skimming vs. reading for detail). These are not physical properties of the reading task and have not been discussed. At issue were the well documented effects of purely visual phenomena on eye movements. To recapitulate, Tinker demonstrated that alterations of physical parameters of printed text cause significant differences

in legibility due to reliable changes in one or more of the usual oculomotor indices – fixation frequency, fixation duration, and/or regression frequency. Fixation duration is sensitive to the perceptibility of the characters in foveal vision, and also line width. It is a function of type size and line width. Fixation frequency, and inversely, saccade length, reflects the perceptual span a text allows or its perceptibility in extra-foveal vision. It varies as a function of type size, line width, and leading. Changes in regression frequency generally reflect the ease of making accurate return sweeps. It varies primarily as a function of line width, but also type size and leading. Poor contrast, poor illumination, or combinations of many non-optimal typographical features will inflate all three indices. The high reliability of these data emphasize the crucial and compulsory role of visual discrimination in reading.

Effects of parafoveal word characteristics

In addition to the factors discussed by Tinker, there has been recent evidence that visual word cues obtained from parafoveal vision influence readers' oculomotor behavior. The recent experiments have not dealt with typographical variations as the earlier work has. Factors such as type size and line width have been held constant while on-line eye movement monitoring techniques are employed to investigate the fixation-to-fixation processing of the textual stimulus in parafoveal vision and its effects on fixation duration and saccade length. In the following section, the effects of word length, word shape, and boundary (i.e., beginning and ending) letters will be discussed.

Effects of word length

Dodge (1907), and more recently, Dunn-Rankin (1978) used an afterimage technique to determine spontaneous fixation locations for single words. Dunn-Rankin had observers look at a small light source until the light appeared to shimmer and a grey afterimage appeared whenever observers moved their eyes. The afterimage was a bright spot which was superimposed on the perceived material at the focal point. Hence, observers could fixate single words and, at the same time, see a light spot which indicated their focal point. Dunn-Rankin found that observers tend to fixate around the center of short and average length words. In addition, there was a slight tendency for some observers to adjust the focal point to the left on longer words. Rayner (1979) monitored fixation locations in a reading situation and found the same fixational pattern. Readers tended to fixate around the central character of short and medium length words, and again, there was a trend to the left for longer words (see Figure 2). If a word received more than one fixation, there was a tendency for an initial fixation near the beginning and a second fixation

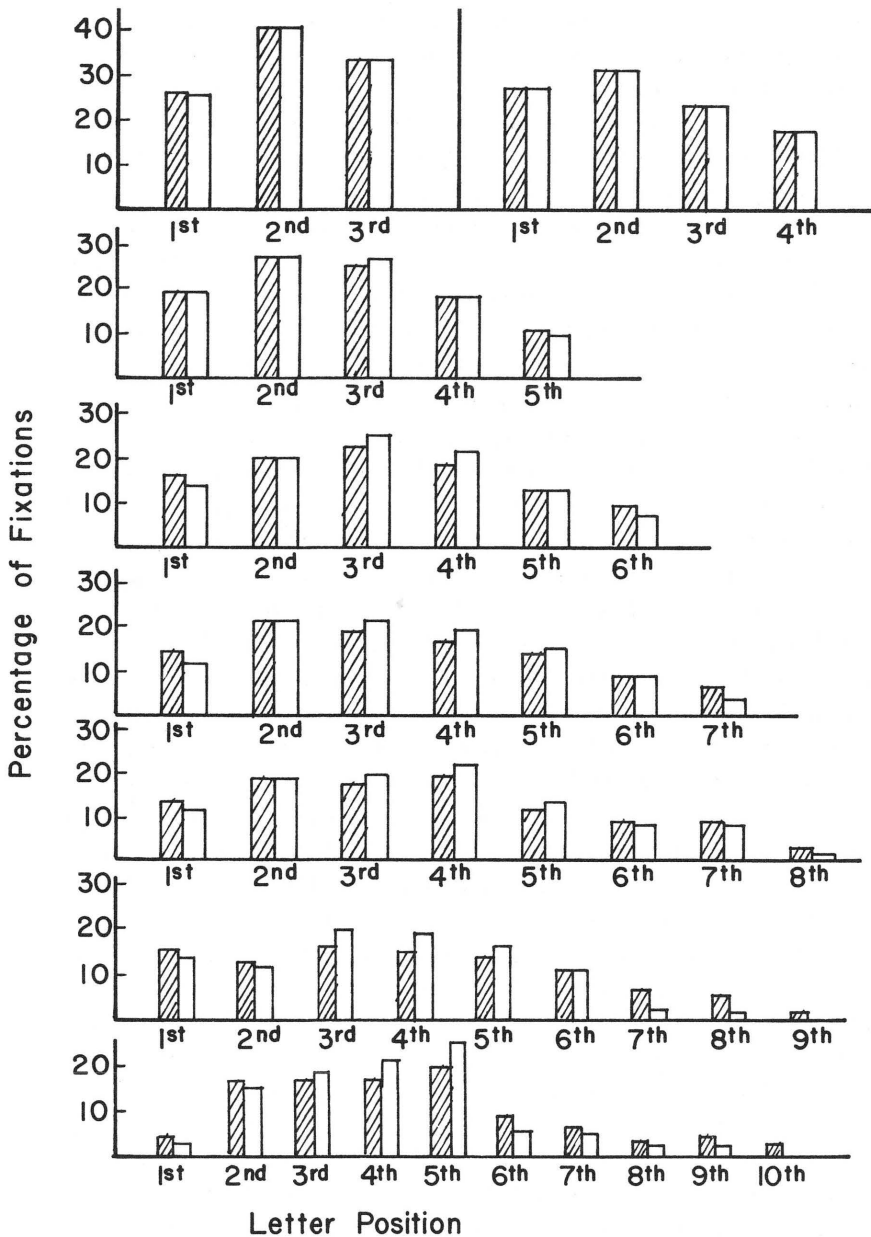


Figure 2. Histograms of percentage of time different letter positions of words of varying lengths were fixated in a forward fixation. Hatched bars indicate instances in which there was more than one forward fixation on a word; the open bars indicate the percentage of time that saccades landed on a word when the word was fixated only once. Reprinted with permission from Rayner (1979). Copyright 1979 by Pion, Ltd.

near the end of the word. Hence, word length was an effective cue in determining the reader's selection of a particular fixation position (i.e., saccade length). It should be pointed out that word length information must be sampled on fixation n if it is to determine the fixation location of fixation $n + 1$. That is, word length has to be identified with parafoveal vision.

Rayner (1979) also calculated the conditional probabilities with which particular words were fixated in reading. The probability was relatively low that one of two short words appearing in succession received a fixation. On the other hand, as either of the two words increased in length, so also did the probability that one of the words received a fixation. Rayner (1979) and also O'Regan (1980) concluded that saccade length was primarily a function of the length of the word to the right of fixation. McConkie and Rayner (1975) investigated the size of the perceptual span readers had for information about the length of words (in order to guide saccades). This, of course, reduces to the question, "What is the perceptual span for blank spaces between words?" They had subjects read computer-generated text (4 characters per degree of visual angle) displayed on a cathode-ray tube (CRT). An eye-tracking system monitored the reader's eye position and fed this information to the computer, which in turn could modify the text display around the readers fixation point in various ways dictated by the experimental manipulations. For example: The computer initially displayed a passage of mutilated text.⁶ However, as soon as the reader fixated a line an area of normal text was symmetrically displayed around the reader's fixation point. Both the size of this "window" and the type of information outside the window were manipulated (see Figure 3). If the window was very large, larger than the perceptual span, say, reading was not disrupted nor were subjects aware of the mutilated text pattern outside of the window. If the window was small enough to constrain the perceptual span though, reading was disrupted – mean fixation frequency and duration increased and mean saccade length decreased – but by different amounts depending on what characteristics of the original text pattern were preserved in the mutilated text outside of the window (these are apparently still perceived). Selectively eliminating visual characteristics of the original text from the mutilated text, such as word length and word shape, allowed the investigators to determine separate perceptual spans for those specific kinds of information. In some conditions ([XS], [CS], or [NCS], Figure 3) readers were provided with word length information outside the window. This was removed in other conditions by inserting letters into the blank spaces between the nonwords (conditions [XF], [CF], and [NCF]). It was found that saccade length was influenced by the presence or absence of word length information outside the window. When word length cues were removed,

Text Graphology means personality diagnosis from hand writing. This is a
 XS XXXXXXXXXXXX XXXXX XXXXonality diagnosis XXXX XXXX XXXXXXXX. XXXX XX a
 XF XXXXXXXXXXXXXXXXXXXXXXXXonality diagnosisXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
 CS Cnojkaiazp wsorc jsnconality diagnosis tnav kori mnlflra. Ykle le o
 CF CnojkaiapawSORcajsnconality diagnosisatnawakoriAmnlflrqaaaYklealeao
 NCS Hbfxwysyvo tifdl xiblONality diagnosis abyt wfdn hbemedv. Awel el f
 NCF HbfxwysyvoatifdlaxiblONality diagnosisaabytawfdnahbemedvaaaAwelaelaf

Figure 3. An example of a line of text and the various peripheral text patterns derived from it. On each line a window of size 17 is shown, assuming the reader is fixating the letter *d* in *diagnosis*. Reprinted with permission from McConkie and Rayner (1975). Copyright 1975 by The Psychonomic Society, Inc.

mean saccade length decreased. This difference between spaced and filled text was found up to window sizes of 25 and 31 character spaces (symmetrically around the fixation point). Hence, it was concluded that word length cues are perceived out to about 12–15 character spaces into the parafovea during a single fixation. Since a reader's saccade length averages around 8-9 character spaces, word length is acquired far enough from the fixation point to influence the selection of the next fixation location. Rayner and Bertera (1979) and Rayner, Inhoff, Morrison, Slowiaczek, and Bertera (1980), using a similar technique to that used by McConkie and Rayner (1975), have recently confirmed that word length cues are obtained about 15 characters to the right of fixation.

Rayner and Pollatsek (1981) did the definitive study showing that saccadic guidance based on parafoveally acquired word length information operates on a fixation-to-fixation basis as has been sketched. In the previous studies, window size had been varied between sentences or groups of sentences: The reader encountered the same window size throughout a sentence. It could be argued that saccade length was adjusted on a delayed basis, once window size had been determined following a couple of normal length saccades. To test this, Rayner and Pollatsek (1981) also varied window size within sentences, randomly from one fixation to the next. This would make a delayed strategy useless. The data showed the usual pattern – saccade size was a function of window size, smaller windows causing shorter saccades – with no difference between the random and blocked window size conditions. Hence, the visual information available during each fixation affects the oculomotor behavior immediately, or on a moment-to-moment basis.

Effects of word shape and specific letters

McConkie and Rayner (1975) also manipulated the availability of word shape information outside of the window by constructing mutilated text wherein each letter was either visually confusable or nonconfusable with the original text. It was found that graphical similarity of the mutilated text with the original text (conditions [CS] and [CF], Figure 3) resulted in shorter fixation durations than the conditions in which the word shape cues were not provided (conditions [NCS] and [NCF]). This difference was observed up to a window size of 21 character spaces. It was concluded that word shape information is acquired as far as 10 letter spaces from the point of fixation and influences fixation duration. Similar results were reported by Rayner (1975). Word shape was also reported to be an effective parafoveal cue in a naming task. Rayner, McConkie, and Ehrlich (1978) instructed readers to move their eyes to a parafoveal location which was occupied by various mutilations of a word. During the course of the eye movement, the initial, mutilated stimulus was replaced with the "base" word. The results seemed to indicate that prior availability of word shape information is an effective cue in parafoveal vision; i.e., response time to name the base word decreased if readers were initially provided with word-shape preserving mutilations. However, this was shown to be false in later studies.

Rayner, McConkie, and Zola (1980) found that facilitation due to word shape was a by-product of specific letter information. In the Rayner, McConkie, and Ehrlich (1978) experiment facilitating word-shape conditions also had some of the same letters as the base word. Shape preserving mutilations which shared no common letters with the base word had produced little facilitation. Rayner, McConkie, and Zola (1980) used stimuli printed in alternating case and reversed the case of each letter to derive the mutilated alternatives. For example, the stimulus "cHaRt" might appear in parafoveal vision, followed by the base word "ChArT" after the eye movement. This condition, devoid of normal word shape information and totally reversing the shape of the stimulus during the saccade, produced the most facilitation. Apparently some letters are identified or encoded into an abstract code from parafoveal vision and this facilitates processing of the base word in the fovea. McConkie and Zola (1979) showed the same effect during fluent reading. It appears that there is no effect of word shape information per se, but only of the identification of specific letters in the word. The letters most likely to be identified in parafoveal vision are presumably the beginning and ending letters of a word which suffer less from lateral masking because they are adjacent to a blank space.

That boundary letters are particularly prominent cues in word perception has been repeatedly demonstrated. For example, Oleron and Danset (1963) presented single words in a tachistoscope and blurred either the initial, middle, or final part of the words. Four levels of blur were used; at each level, the order of accuracy was the same: initial letters were reported best, final letters next, and medial letters poorest. Reaction time measures mimicked that order, with initial letters having shortest latencies.

Boundary letter information is particularly salient in parafoveal word presentations, although the most eccentric letter was found to be more perceptible than the inward-most (closest to the fovea) letter (Bouma, 1973). In addition, boundary letters are better identified to the right of fixation than to the left (Bouma, 1973). Boundary letters have also been hypothesized to control readers' eye movement behavior (e.g., Hochberg, 1970).

To explore the effect of boundary letters upon readers' eye behavior, Rayner (1975) controlled the availability of such cues. More specifically, it was possible to compare conditions in which a parafoveal nonword stimulus did or did not contain the boundary letters of a base word subsequently perceived in the fovea. The results showed that initial and final letters affected fixation durations on the base in a reading situation. That is, parafoveal word mutilations which preserved boundary letter information reduced subsequent fixation durations on the base word. Similarly, Rayner et al. (1978) showed that parafoveally available boundary letter information reduced response time for a subsequently fixated base word in a naming task. Unlike the data reported by Bouma (1973), these results showed that the initial letter in a word to the right of fixation, not the last (most eccentric) letter, had the strongest effect. It was concluded that specific letters, especially the initial letters of words, can be identified in parafoveal vision and allow shorter fixation durations when the words are fixated. This may account for the tendency of readers to have shorter mean fixation durations towards the end of a line than the beginning of a line (Dearborn, 1906; Tinker, 1965).

To summarize, recent research has shown that word length and the identity of specific letters (especially initial letters of words) in parafoveal vision affect oculomotor behavior. Word length primarily determines saccade length, while specific letter identification affects fixation duration.

Conclusion

The research surveyed provides evidence that the typographical features of text are determinants of the perceptual span in reading and thus affect eye movement measures such as fixation frequency and saccade length. Regression frequency and fixation duration are also affected by these purely

visual factors. Although comprehension requirements (e.g., skimming, reading for detail) and text difficulty may also have effects on eye movement measures within readers, and despite wide differences in eye behavior between individuals, the oculomotor effects of the physical characteristics of text are undeniable and should not be ignored.

Paradigms which use oculomotor measures (e.g., fixation frequency) as indices of text processing difficulty must control the effects of typography to avoid confounds. Absolute comparisons of eye movement data between studies utilizing different typographies will generally be invalid. Control groups yielding baseline data under typographical conditions identical to experimental groups are necessary to allow relative comparisons between experiments utilizing different typographies.

Oculomotor consequences of typography and visual word cues identified in parafoveal vision underscore the fundamental nature of visual discrimination in reading, including "normal" fluent adult reading. Reading depends upon the encoding of information physically represented in print. The speed of this process can be limited by characteristics of the print which impede the visual processing, particularly the efficiency of parafoveal information extraction. The availability of parafoveal word cues, such as word length and initial letters, may be crucial for fluent reading.

1. The fovea is a small, indented region of the human retina which extends about 2° of visual angle. The fovea contains the region of highest acuity although there is considerable drop-off even within the fovea (Riggs, 1965). Acuity continues to decrease beyond the fovea, both for the parafoveal region, which extends about 10° around fixation, and the peripheral region, which covers the remainder of the retinal system. Many experimenters have found that a number of visual acuity thresholds increase linearly with retinal eccentricity up to 30° (Anstis, 1974).
2. A point is a typographer's unit of vertical distance approximately equal to $\frac{1}{2}$ of an inch. Nine-point type (these lines) means that the vertical extent allowed a line of characters including the ascending and descending parts is approximately $\frac{9}{16}$ of an inch. A pica is a unit of horizontal distance approximately equal to $\frac{1}{6}$ of an inch.
3. This is necessarily true only if regressive fixations are removed from the fixation frequency before words per fixation is calculated. Although Tinker did not remove regressions (Tinker, 1955, p. 176), he does report regressive fixation frequency separately. Often the regressive fixation frequency shows no significant difference with legibility manipulations in Tinker's data. Significant changes in words per fixation in these instances are interpreted as real differences in mean forward saccade length.
4. Although Tinker's fixation frequency data include regression frequency, the increase must be solely due to an increase in the number of forward fixations, as regression frequency actually decreased or did not change on very short lines.

5. Lateral masking refers to the decreased perceptibility of stimuli when surrounded by adjacent contours. The effect is especially noticeable when letters to be identified in non-foveal vision are presented in isolation or amidst laterally adjacent letters. Performance decreases drastically in the latter condition (Mackworth, 1965).
6. In mutilated text each character is replaced with a different character, resulting in a passage of nonwords.

REFERENCES

- Anstis, S. M. A chart demonstrating variations in acuity with retinal position. *Vision Research*, 1974, 14, 589–592.
- Bouma, H. Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, 1973, 13, 767–782.
- Dearborn, W. F. *The psychology of reading*. New York: Columbia University Contribution to Philosophy and Psychology, Volume 14, No. 1, 1906.
- Dodge, R. An experimental study of visual fixation. *Psychological Review Monograph Supplements*, 1907, 8, No. 4.
- Dunn-Rankin, P. The visual characteristics of words. *Scientific American*, 1978, 238, 122–130.
- Hochberg, J. Components of literacy: Speculations and exploratory research. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.
- Huey, E. B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908.
- Mackworth, N. Visual noise causes tunnel vision. *Psychonomic Science*, 1965, 3, 67–68.
- Matin, E., Clymer, A., & Matin, L. Metacontrast and saccadic suppression. *Science*, 1972, 178, 179–182.
- McConkie, G. W., & Rayner, K. The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 1975, 17, 578–586.
- McConkie, G. W., & Zola, D. Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, 1979, 25, 221–224.
- McCormick, E. J. *Human factors in engineering and design*. New York: McGraw-Hill, 1976.
- Oleron, P., & Danset, A. Donnees sur l'apprehension des mots. *Psychologie Francaise*, 1963, 8, 28–35.
- O'Regan, K. The control of saccade size and fixation duration in reading: The limits of linguistic control. *Perception & Psychophysics*, 1980, 28, 112–117.
- Paterson, D. G. & Tinker, M. A. Influence of line width on eye movements. *Journal of Experimental Psychology*, 1940, 27, 572–577.
- Paterson, D. G. & Tinker, M. A. The effect of typography upon the perceptual span in reading. *American Journal of Psychology*, 1947, 60, 388–396.
- Rayner, K. The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 1975, 7, 65–81.
- Rayner, K. Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, 1977, 5, 443–448.
- Rayner, K. Eye movements in reading and information processing. *Psychological Bulletin*, 1978, 85, 618–66.

- Rayner, K. Eye guidance in reading: Fixation locations within words. *Perception*, 1979, 8, 21–30.
- Rayner, K. & Bertera, J. Reading without a fovea. *Science*, 1979, 206, 468–469.
- Rayner, K., & Inhoff, A. W. Control of eye movements during reading. In B. L. Zuber (Ed.), *Models of oculomotor behavior and control*. Boca Raton, FL: CRC Press, 1981, in press.
- Rayner, K., Inhoff, A. W., Morrison, R., Slowiaczek, M., & Bertera, J. Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 1981, 7, 167–179.
- Rayner, K., & McConkie, G. W. What guides a reader's eye movements? *Vision Research*, 1976, 16, 829–837.
- Rayner, K., McConkie, G. W., & Ehrlich, S. Eye movements and integrating information over fixations. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, 4, 529–544.
- Rayner, K., McConkie, G. W., & Zola, D. Integrating information across eye movements. *Cognitive Psychology*, 1980, 12, 206–226.
- Rayner, K., & Pollatsek, A. Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology*, 1981, in press.
- Riggs, L. A. Visual acuity. In C. H. Graham (Ed.), *Vision and visual perception*. New York: Wiley, 1965.
- Tinker, M. A. & Paterson, D. G. The effect of typographical variations upon eye movement in reading. *Journal of Educational Research*, 1955, 49, 171–183.
- Tinker, M. A. *Legibility of print*. Ames, Iowa: Iowa State University Press, 1963.
- Tinker, M. A. *Bases for effective reading*. Minneapolis: University of Minnesota Press, 1965.
- Walker, R. Y. A qualitative study of eye movements of good readers. *American Journal of Psychology*, 1938, 51, 472–481.
- Woodworth, R. S. *Experimental psychology*. New York: Holt, 1938.

Acknowledgement. Preparation of this paper was partially supported by Grant HD 12727 from the National Institute of Child Health and Human Development.

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.
0022-2224/81/0040-147\$02.00/0 © 1981 Visible Language, Box 1972, Cleveland, OH 44106.

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).

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

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

abcdefghijklmnopqrstu
vwxyz
 abcdefghijklmnopqrstu
 vwxyz
abcdefghijklmnopqrstu
vwxyz
 abcdefghijklmnopqrstu
 vwxyz
abcdefghijklmnopqrstu
vwxyz
abcdefghijklmnopqrstu
vwxyz
abcdefghijklmnopqrstu
vwxyz
abcdefghijklmnopqrstu
vwxyz
 abcdefghijklmnopqrstu
 vwxyz
abcdefghijklmnopqrstu
vwxyz
abcdefghijklmnopqrstu
vwxyz
 abcdefghijklmnopqrstu
 vwxyz
abcdefghijklmnopqrstu
vwxyz
abcdefghijklmnopqrstu
vwxyz
 abcdefghijklmnopqrstu
 vwxyz
abcdefghijklmnopqrstu
vwxyz
abcdefghijklmnopqrstu
vwxyz
 abcdefghijklmnopqrstu
 vwxyz
abcdefghijklmnopqrstu
vwxyz
 abcdefghijklmnopqrstu
 vwxyz

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 ensconced 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	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.

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.

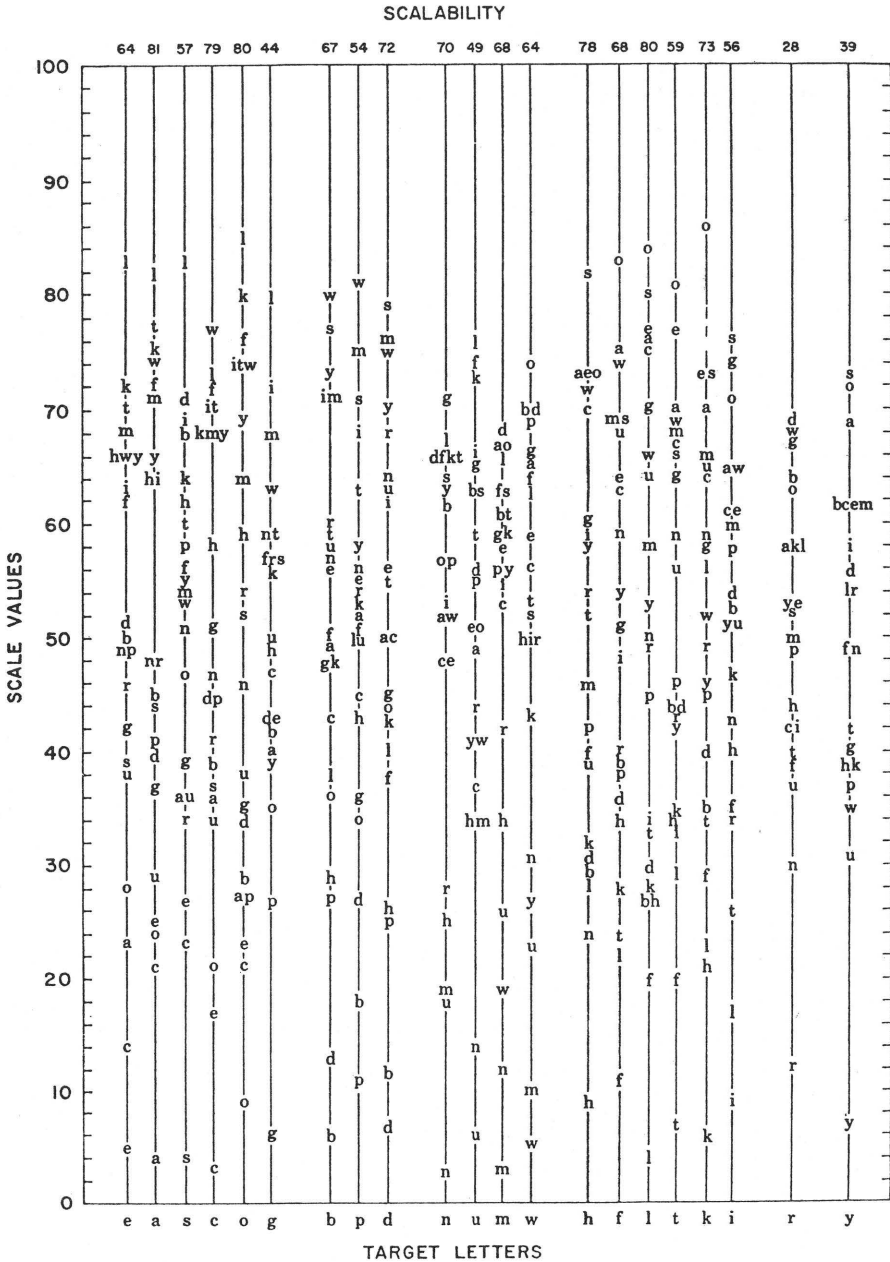
This kind of variation has been studied at the level of syntax and morphology in spoken language in contexts, where, as with letter features, the language permits the application of more than one rule. For example, in English, when forming the past tense of a verb, there are a number of possible rules 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.

		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 ₁	a	37	5	3	2	3	7	1	3	1	3	1		8	2				1	1							22	a	
	b	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	
S ₂	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								3				5	o	
	c	2		3		19	11	43	4	1	3	2	2	1									2	1		1	5	c	
S ₃	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		
S ₄	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	2			25	70													1	w	
A ₁	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											6	b
A ₂	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	5	7		y	
	q																		2				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.

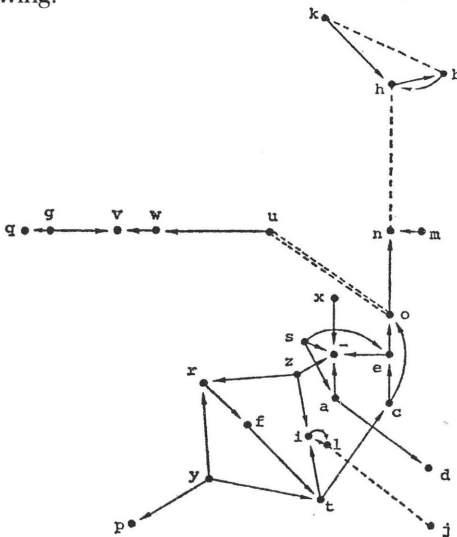


Table VIII. Upper estimates of the cue values of certain letter properties (upper cue value *C*).

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 <i>C</i>
(1) $H < 2.0$ mm or $H/W < 1.16$	Short letter	All <i>S</i> letters	All <i>S</i> letters	0.46	0.38	0.84
(2) $H > 2.5$ mm or $H/W > 1.16$	High letter	All <i>A, D</i> letters	All <i>A, D</i> 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 $H/W < 1.16$	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 a 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) $H/W > 1.22$	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.

It is undoubtedly much easier for a young child to learn to play 1a
by himself or with other children than for a parent to learn to play
with that young child. The parent needs to get off his adult high
horse and get down to the child's level in play. This is not easy for
many parents. We need to learn to play games with a young child
to make up spontaneous games, to tell stories and read books to
him. There is an art to all of these activities. They are learned skills.

If you are quite honest with yourself, you will find that there 1b
are times when you will lose your temper, fly off the handle at your
child, and yell at him or scold him, only to realize afterwards that
what he did actually should not have elicited such a violent outburst
from you. You were really mad at your husband or your neighbor.
Or just cranky for some unknown reason. And you took it out on
your child.

Four preschoolers will need to play with other children and 1c
learn these socializing skills during the ages of three, four, and five.
Nursery school is an ideal place to learn them, because in nursery
school the learning can be supervised by a trained teacher. In
neighbourhood play, the learning is hit or miss, trial and error. In
neighbourhood play, for example, there's no trained person to help
a shy child integrate himself into a group and learn to build up his
self-confidence and grow out of his shyness.

When a child is three, he craves companionship in his play. 1d
He wants to separate from his mother and become more inde-
pendent. The easiest way to help him to do this is to send him to a
good nursery school. Even though a three-year-old wants to be
separate from mother and get out into the world of his peers, he
still has ambivalent feelings about leaving the security and protection
of mother. It is only natural that he should feel this separation
anxiety, for mother has been home base for him for three some
years. Some children feel these separation anxieties more strongly
than others.

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	1	0	1	1	0	2	1	0	0	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	0	1	0	2
Curve features:																													
Open semicircle	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	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	1	0	0	0	0	0	1	1	0	1	0	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	1	0	1	0	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	1	0	0	1	0	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.

┌┐ ┌┐ ┌┐ ┌┐ little girl, ┌┐ spent several hours
└┘ almost every day ┌┐ ┌┐ large lumber mill. ┌┐ liked ┌┐
listen ┌┐ ┌┐ scraping └┘ grinding └┘ ┌┐ machinery └┘
watch └┘ ┌┐ busy men └┘ work. ┌┐ father made ┌┐ com-
fortable little seat ┌┐ ┌┐. ┌┐ ┌┐ ┌┐ sit └┘ watch
┌┐ sharp saw cut through ┌┐ big logs, scattering sparks
└┘ sawdust. ┌┐ men ┌┐ cut off ┌┐ clean, white boards
just ┌┐ easily ┌┐ ┌┐ logs ┌┐ ┌┐ made └┘ chocolate
fudge. ┌┐ ┌┐ ┌┐ fresh, woody smell └┘ ┌┐ noise └┘
└┘ ┌┐ busy machines often ┌┐ make ┌┐ drowsy └┘ ┌┐
┌┐ fall asleep └┘ ┌┐ happy dreams.


	<u>the</u>	<u>was</u>	<u>bed</u>	<u>may</u>	<u>what</u>
Outline					
Same Size	XXx	xxx	XxX	xx _X	xXxX
Density					
None	THE	WAS	BED	MAY	WHAT

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 acceptability 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-

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

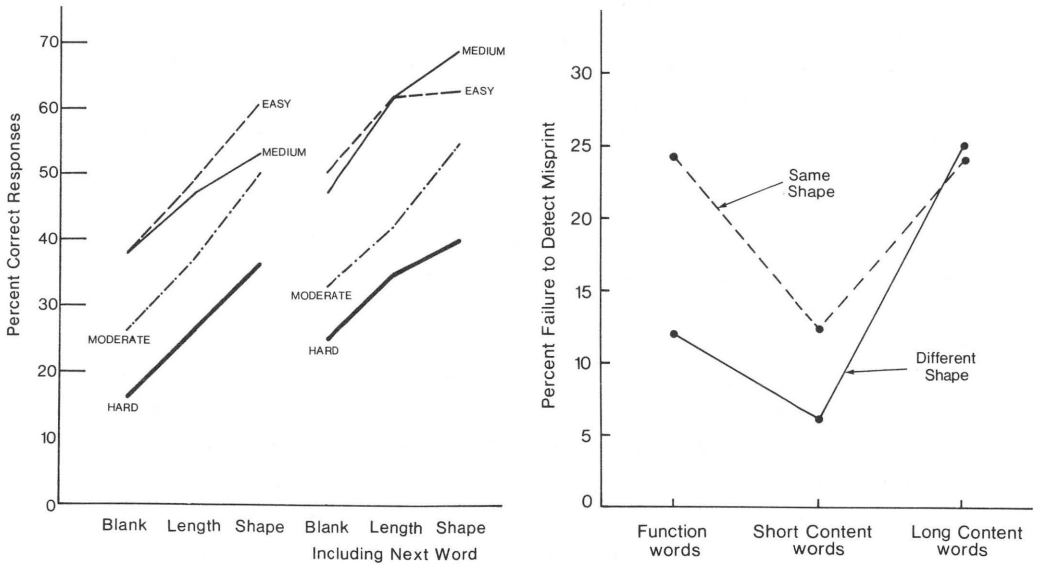


Figure 10. Accuracy of guesses of the word in the blank, when that word was only preceded by a full context of the story (blank), or in addition had its length specified, or in addition had its outline shape specified. Right hand panel is for conditions in which the word after the blank also was present. 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 encompass-

ing 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., & Hagedberg, 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. discriminability 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.

Toward a Computational Theory of Early Visual Processing in Reading

Michael Brady

This paper is the first of a series aimed at developing a theory of early visual processing in reading. We suggest that there has been a close parallel in the development of theories of reading and theories of vision in artificial intelligence. We propose to exploit and extend recent results in computer vision to develop an improved model of early processing in reading. This paper considers the problem of isolating words in text based on the information which Marr and Hildreth's (1980) theory of visual edge detection asserts is available in the parafovea. We show in particular that the findings of Fisher (1975) on reading transformed texts can be accounted for without postulating the need for complex interactions between early processing and downflowing information as he suggests. The paper concludes with a brief discussion of the problem of integrating information over successive saccades, and relates the earlier analysis to the empirical findings of Rayner.

1 Introduction

This paper presents computational and psychophysical evidence in support of a theory of one of the earliest stages of visual processing in reading, namely the isolation of words in text. As such it is the first step in the development of a computational theory of reading whose general direction is presented in the next section. A skeletal outline of the paper follows.

The goal of reading may be supposed to be the efficient extraction of meaning from text. Realising this goal involves integrating "upward flowing" information uncovered by early visual processing with "downward flowing" cognitive interpretations. In this paper, we present an approach toward understanding the visual aspects of reading which we believe may contribute greatly to an understanding of the overall reading process.

Existing theories of reading have relied on a primitive model of early visual processing. We suggest that as a result they have typically accorded too much emphasis to the role of "downward flowing" cognitive information, in effect

Visible Language, XV 2, pp. 183-214.

Author's address: MIT Artificial Intelligence Laboratory, Cambridge, MA 02138.

0022-2224/81/0040-183\$02.00/0 ©1981 Visible Language, Box 1972, Cleveland, OH 44106.

suggesting that its deployment is necessary for almost every aspect of reading. Indeed, over the past two decades there has been a close parallel between the development of theories of reading and theories of visual perception in artificial intelligence (AI). In particular, we note that a number of reading theorists have recently been attracted to complex processing models developed in AI. A major attraction of such models is that they seem to provide a mechanism supporting flexible behavior by which information available as a result of early visual processing could combine with downflowing information about the specific image domain to produce an interpretation or percept. Still more recently, AI has witnessed a fascination with "relaxation" style processing (see Section 2.1). This is not only claimed to support the interaction between low level and downflowing information, but to do so by local parallel interaction. A number of reading theorists have proposed similar mechanisms. For the most part, these theories have had limited success in explaining the empirical psychophysical data on reading. We argue that this is, in part, because they depend upon a primitive model of early visual processing. It is also partly because of an emphasis on the mechanism of integrating information from various sources, without addressing the issues of what purpose the information serves, what is the information which is passed, and how it is represented (see Marr, 1980; Marr and Nishihara, 1978).

Over the past few years there has been considerable progress in understanding early visual processing. The achievements of Horn, Marr, Poggio, Ullman, and others in developing a computational theory of natural visual perception has little or no counterpart in theories of reading. For example, Frisby (1979, page 108) and Allport (1980, page 235) equate early processing with feature extraction as developed in optical character recognition systems (Duda and Hart, 1973). A fuller account of the relevant empirical findings is given in Cohen (1978, page 65), but her analysis falls considerably short of being a precise and coherent theory. The computational theory of natural vision suggests that much richer information can be made available by early visual processing in reading, without the aid of downward flowing "higher level" knowledge of the domain being viewed. Reading has always attracted a great deal of attention from perceptual psychologists, in part because of the light it might shed on our understanding of human perception of the natural world. We claim that, temporarily at least, the boot is on the other foot, and that the recent developments in our understanding of real world perception can be gainfully applied to increase our understanding of reading. Specifically, Marr, Hildreth, Ullman, Poggio, and Richter have developed a computational theory

of edge detection whose findings closely match physiological data on ganglion cell responses. For the purposes of the work reported here, the important point about the Marr-Hildreth theory is that it makes available to reading theorists, perhaps for the first time, a detailed description of the information delivered by the earliest stage of visual processing of an arbitrary text sample viewed at an arbitrary eccentricity.

Finally, we review some empirical findings about the earliest stages of visual processing in reading, and we settle upon the isolation of words as the first goal of the reader's perceptual processing. We note that eye movement studies show that a great deal of processing is carried out on text prior to foveation. It follows that it is reasonable to conjecture that word isolation is effected on the basis of information available in the parafovea. As part of an investigation of this conjecture, we suggest that Fisher's (1975) results on transformed text provide some insight into the isolation of words based on information available in the parafovea, and so we analyze his results carefully. We argue that they can be explained on the basis of Marr and Hildreth's (1980) theory of edge detection without postulating the need for "higher order visual processing" as was claimed by Fisher. The explanation leads to a number of empirical predictions, which are confirmed by essentially replicating Fisher's experimental technique. The concluding section sketches a theory of word isolation in the parafovea, and notes that the decision to activate the reading process in the first place is also not very mysterious.

2 Background to the study

2.1 Past approaches to theories of reading

From the earliest days of experimental psychology there has been a constant stream of research findings about reading (see, for example, Huey, 1908, and Henderson, 1977). All of the major schools of perception have considered reading to some extent and have attempted to exploit various mathematical and computational insights to develop their theories. We are particularly concerned with the growth of interest over the past two decades, during which time a number of theories have developed, the majority being expressed in terms of information processing.

Relative to the behaviorists' reliance on a simple mechanism, which bore many of the characteristics of the pattern recognition systems developed in the 1950's and 1960's, and the extreme wordiness of the gestalt and new look theorists, information processing accounts of reading are refreshingly precise. They consist of individuated stages, at which some particular

functionally defined "process" is carried out (say, to extract features or to consult a lexicon), together with interconnecting arrows, which represent the flow of information through the system under consideration. An important property of such models is that they describe the way in which a perceptual or cognitive process being studied unfolds over time. The particular class of individuated stage processes, and the topology of interconnecting arrows, are carefully chosen to account for relevant empirical findings. While the power of such formalisms is clearly sufficient to account for any given set of descriptions, in the absence of a wholly precise mathematical or computational account of reading, any particular model is inevitably vague in places. The extent to which it does or does not adequately explain the available empirical data (and the precision of the predictions which can be made from it) are limited. For example, Gough (1972) presents a flow diagram of "one second of reading" which embodies the theory that phonological recoding is obligatory. Marcel and Patterson (1979) present an alternative in which it is not. For further examples, see Estes (1977), Cohen (1978), and McClelland and Rumelhart (1980).

The box and arrow diagrams which feature in most information processing accounts of perception are highly reminiscent of the system flowcharts which used to be prepared by programmers in the early stages of developing a program. Flowcharts have fallen into disrepute in computer science as it has been realized that they provide an impoverished representation of such a key issue as the structure of a program. They are also wholly inadequate as a representation of process interaction and parallelism being essentially restricted to the description of a single sequential process. Of course, they are merely the simplest first approximation to a model of processing, though one should be aware of the computer science experience that they unacceptably straitjacket thinking.

Several authors have argued that it is not possible to develop a theory of an ability such as reading, in which the flow of information is wholly unidirectional, that is, a flow that proceeds from the processes which embody relatively general knowledge, and which make contact with the intensity levels of the image to the processes embodying knowledge about the specific objects and situations depicted in the image (see, for example, Allport, 1979; Frisby, 1979; Cohen, 1978; Rumelhart, 1977). It is supposed that "downward flow" of knowledge about such objects and situations is also necessary to account for the remarkable abilities and flexibility of human perception.

The invocation of "downward flow" as an explanation for reading abilities has an interesting (perhaps not co-incidental) parallel with the history of

computational theories of natural visual perception in the field of artificial intelligence. The period 1963 to the early 1970's in the development of AI was most notable for extensive experimentation with edge detecting or region finding operators, designed ad hoc in accordance with the needs of some particular project. Authors time and again noted that the results of applying their operators to digitized images were essentially unpredictable; many concluded that it was simply not possible to develop a theory of early visual processing capable of generating predictably rich and useful descriptions that could then be used as the basis for computing the visible surfaces and objects in a scene. It was supposed therefore that, just as in the case of reading (although the AI workers involved would not have known of the parallel), "downward flow" of knowledge about the objects and situations imaged in the scene was essential to explain the remarkable abilities of human visual perception. The interaction between upward flowing information generated by relatively unknowledgeable early processing modules and downward flowing information was essentially dynamically determined and could not be completely defined in advance. It was conjectured by Minsky and Papert (1972) that among the tools developed in computer science, the best way to achieve this dynamically determined behavior was through process interactions, which, it was noted, need not be restricted to the simple patterns of (serial) activity provided in a language like Fortran or Algol. These were the considerations which lay behind the development of a rash of complex "heterarchical" programs to understand natural language, perceive utterances from a speech signal, and see in various narrowly defined domains. Programs such as Hearsay 2 (Lesser and Erman, 1977), Margie (Schank et al., 1973), Barrow and Tenenbaum's (1976) Interpretation Guided Semantics, and the author's own program for "reading" Fortran code (Brady, 1979; Brady and Wielinga, 1978) are typical of the genre.

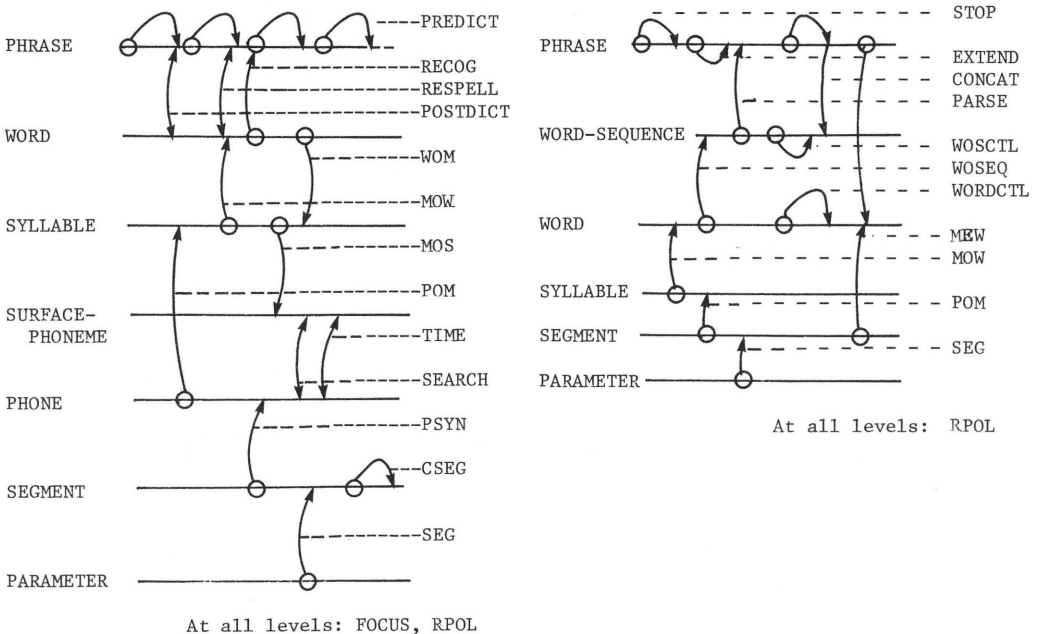
The development of complex "heterarchical" programs such as Margie and Hearsay 2 is paralleled by the adoption of those computational models of processing by reading theorists eager to explain the use of downward and upward flow as determinants of a percept. Examples are Cohen's (1978) discussion of Speechlis (Nash-Webber, 1975), and Allport's (1979) detailed explanation of the operation of Margie.

In fact, a number of difficulties emerged in the dynamic processing account of perception as soon as vague theoretical notions like "process interaction" needed to be made precise (see Brady, 1979). There are two basic difficulties, one technical, the other more empirical in nature though reflecting a theoretical shortcoming. Technically, the potency of process

interactions, and the stock of ideas about how to control and analyze them, remain very limited indeed. Secondly, and most notably, the presumed power of heterarchy never materialized. It repeatedly became evident that a small increase in the early processing capabilities of programs could have a far greater impact on the performance of a program as a whole than a vastly greater amount of "higher level reasoning."

Consider in particular the case of Hearsay 2 (Lesser and Erman, 1977). One of the main innovations of Hearsay 2 was the introduction of a centralized data structure called the "blackboard," on which the findings of a number of "knowledge sources" (which performed such tasks as isolating phonemes, syllables, words, or larger syntactic units) were presented. At any stage of the processing of a speech signal corresponding to an utterance, the contents of the blackboard represented the state of the system's interpretation. The addition of a piece of information by one knowledge source could enable the activity of several others. At any given stage, there were typically many runnable processes (up to two hundred), each of which was assigned a numerical priority value indicating its apparent importance. This design is illustrated in Figure 1a, which shows the Hearsay 2 system as of January

Figure 1. The structure of the blackboard state descriptor for the Hearsay 2 speech understanding system. 1a: the system as of January 1976. 1b: the second version as of September 1976 (from Lesser and Erman, 1977).



1976. The authors note that “this implementation had poor performance; e.g., 10% of sentences correct in 85 million instructions per second of speech on a 250 word vocabulary” (Lesser and Erman, 1977, page 790). A second design, shown in Figure 1b, was aimed at “making the lower levels of processing more sequential and bottom up” (page 795). The authors reported that “this configuration performs substantially better; e.g., 85% correct in 60 million instructions per second of speech on a 1000 word vocabulary.”

Some AI researchers (see, for example, Davis and Rosenfeld, 1978, 1981; Barrow and Tenenbaum, 1978; Rosenfeld, Hummel, and Zucker, 1976; Waltz, 1978; Zucker, 1978) concluded that the main drawback of the heterarchical process organisations discussed above was that they were essentially serial. They argue that much of their complexity arises because one is forced to choose a particular sequential order in which to carry out a number of processes. Since this order is inevitably often inappropriate (being unpredictable), one is then required to incorporate sufficient mechanism to facilitate recovery. Instead, such authors suggest the use of globally constrained local parallel processes, which some authors, notably Rosenfeld et al. (1976), have likened to relaxation processes for solving systems of equations in numerical computing.

In essence, the idea underlying “relaxation processing” is as follows. A large number of simple individual processors are postulated, say, one for each image position (there are after all several million cones in the human retina, and billions of neurons in visual cortex.) For present purposes, it may be supposed that the processors are laid out on a flat plane. Each processor is connected to just a small number of the other processors that are physically near it. The processors all operate in strict synchrony. At each step, each processor changes its “state” depending upon its previous state and the states of the processors to which it is connected. The system proceeds in this manner until no processor changes state, at which time the system is said to have “converged.” This model of computation has attracted some attention because of its inherent parallelism coupled with its (limited) possibilities for context sensitivity.

Note that in common with the heterarchy approach, the structure of the mechanism is developed and fixed in advance of the analysis of the particular perceptual problem being studied. The only issues which the theorist is left to settle in most accounts are parameter settings, such as the size of neighborhoods, thresholds, and the like (see Davis and Rosenfeld, 1981). We argued above that a major drawback with heterarchical accounts of perception was the difficulty in analysing and controlling them. It is important to realise that analogous problems arise with relaxation processes. It is usually extremely

hard to guarantee that such a process settles down to a steady state ("converges"). As an example, consider the difficulty that Marr, Palm, and Poggio (1978) had in analysing the behavior of the Marr and Poggio (1976a, 1976b) cooperative algorithm for computing stereo disparity. If this is difficult for a single level of relaxation processing, it is considerably more so for the hierarchical or multi-stage processes which have been advanced (though usually not implemented and tested) in the literature (e.g., McClelland and Rumelhart, 1980; Davis and Rosenfeld, 1978; Zucker, 1978). Few, if any, results are known regarding the convergence (including speed of convergence) of such relaxation processes (see Ullman, 1979; Zucker, Leclerc, and Mohammed, 1979). Without such results, the uncritical proposal of complex locally parallel processes is of questionable significance.

2.2 The computational approach to vision

Against this background of ad hoc experimentation and the construction of uncontrollable complex processing models in artificial intelligence, the computational theory of natural visual perception developed by Horn, Marr, Ullman, Poggio, Binford, and others is quite remarkable. A fuller account of the current state of computer vision can be found elsewhere (Marr, 1980; Brady, 1981a, 1981b; Horn, 1978; Marr and Poggio, 1979; Marr and Hildreth, 1980; Grimson, 1980). For the purposes of this article, it is sufficient to note that there now are mathematically precise theories and highly parallel, robust computer implementations of a variety of (human) visual processes. These include edge detection, stereopsis, shape from shading, shape from texture, early motion detection, and surface interpolation. In each case these theories concern processes which occur at an early stage of perception, and they embody knowledge about the world which is of considerable generality; for example, that the world mostly consists of smooth surfaces. In short, the computational theory of vision is a compelling argument in support of the power of early visual processing. More significantly perhaps, it promotes a research methodology which defers consideration of knowledge rich, domain specific, downward flow of information until the considerable scope of early processing is more clearly understood. It also makes little sense to develop an understanding of the role of downward flow until we have a better appreciation of what information early processing can and does provide.

The computational theory of visual perception referred to above is also interesting for the research methodology which has developed from it. The first step is to isolate a perceptual ability for which there is empirical evidence for considerable competence on the basis of early processing. For example, Horn (1974) has studied the determination of lightness and the

computation of and shape from shading (1978) from an image. Marr and his colleagues have considered edge detection (Marr and Hildreth, 1979), stereopsis (Marr and Poggio, 1979; Grimson, 1980), and motion computation (Ullman, 1978; Marr and Ullman, 1979; Ullman and Richter, 1980). The particular problem is then studied in three parts. First, we consider what information must be extracted from the scene, in order for the system to exhibit this competence, and what constraints on the world the system needs to assume in order to extract this information. Second, one attempts to devise a representation which makes explicit the information required to explain the competence. Only then is it reasonable to devise algorithms to discover the appropriate representation instance for a scene. Finally, one can conduct experiments to discover the extent to which the algorithm explains human performance. Notice that in contrast to this methodology, the heterarchical and relaxation processes outlined above start with an algorithm (or commitment to a particular restricted kind of processing) and only then examine competence, devise representations, and analyze the basis of the competence.

2.3 Edge detection in the human visual system

As an example of the results of the computational approach to early visual processing, we take a brief look at Marr and Hildreth's (1980) theory of edge detection. The reason for this choice is quite simple. The theory addresses the very first stage of analysis of the visual input, and this is the stage which is most relevant to the study of parafoveal processing in reading which is presented in the balance of the paper.

A feature of Marr's (1976) original development of the "primal sketch" representation was its direct reference to neurophysiology and psychophysics, a commitment Marr was to continue to stress in later work. Marr's algorithm for computing the primal sketch from an image had a number of interesting features. First, being inspired by neurophysiology, Marr applied the findings of Hubel, Wiesel, Barlow, and others, which seem to suggest that an early stage in the processing of visual information consists of convolving the image with edge and bar masks. As we observed above, such masks signal an approximation to the first and second (directional) derivatives of the brightness function. Marr based his algorithm on an analysis of the response of bar and edge masks to ideal instances of the scene events which give rise to intensity changes. The algorithm itself consisted of convolving an image with a number of edge and bar masks and then "parsing" the results by matching the actual responses to those predicted for ideal scene events. It was noted that bar masks seemed to give more reliable information than edge masks,

an observation whose explanation awaited the later development of $\nabla^2 G$ operators which have a similar cross section. The algorithm convolved the image with masks of different panel widths. Although the later justification for this would be in terms of separate processing channels, the original explanation was based on the need for noise reduction, although this idea was never formulated precisely. In any case, the outputs of the individual channels was combined, not only to reduce the effects of noise, but to compute measures such as the "fuzziness" of an edge.

Marr and Hildreth (1980, page 189) point out that "a major difficulty with natural images is that changes can and do occur over a wide range of scales, so it follows that one should seek a way of dealing with the changes occurring at different scales." One way to do this, which has been proposed several times in the image processing literature, is to pass the image through a number of band limited filters. Of course, the difficult issues concern the choice of filters (bar mask, Fourier, Gaussian), the number of them, and the exact band pass characteristics of each.

Intensity changes are localised in space, a fact which derives from their physical causes (see Horn, 1977; Marr, 1976; Marr and Hildreth, 1980, page 189). Marr and Hildreth argue that they are also localised in the frequency domain, since the world is mostly composed of visible surfaces of roughly uniform texture, the idea being that a texture is essentially periodic. Marr and Hildreth (page 191) note that "unfortunately, these two localization requirements, the one in the spatial and the other in the frequency domain, are conflicting." They show that bar masks localise changes in the image, but can generate echo effects on textures. Conversely, a Fourier filter localises changes in the frequency domain but produces unwanted echoes around edges. They point out that only the Gaussian optimises localisation in both domains simultaneously, and so it is chosen as the band limiting filter in the theory. Recently, Binford (1981) has questioned the argument that texture corresponds to limited variance in the frequency domain, and by implication questions the optimality of the Gaussian.

In order to locate edges, one can either find places where the first derivative of the intensity function reaches a maximum, or equivalently where the second derivative is zero. To locate edges at arbitrary orientations with equal facility, we require a differential operator which is not directional. The Laplacian is the only first or second order differential operator with this property. Thus the Marr and Hildreth theory asserts that following Gaussian smoothing, the image is convolved with a Laplacian and zero crossings noted. In fact, by the convolution theorem,

$$\nabla^2(G*Image) = (\nabla^2 G)*Image,$$

where G is a Gaussian operator, and $*$ denotes convolution. Marr and Hildreth (page 193) point out that the $\nabla^2 G$ operator is bandpass and closely resembles the difference of Gaussian (DOG) operators proposed by Wilson and Giese (1977; see also Wilson and Bergen, 1979). Indeed they show that $\nabla^2 G$ is the limit of a DOG, and that the DOG closely approximates it. Actually, doubly differentiating an image causes severe numerical noise, and so the difference of Gaussians can actually be more accurate in practice, as well as being more efficient. The two-dimensional cross section of the $\nabla^2 G$ operator is a smoothed version of a bar mask cross section, and may explain the heuristic observation mentioned earlier. Wilson and Bergen's work suggests that there should be four DOG channels at each retinal eccentricity, and that their characteristic sizes should scale linearly with eccentricity, being smallest in the fovea and doubling in size by about 4° . Recently Marr, Hildreth, and Poggio (1979) have noted evidence for a fifth, smaller channel in the fovea, and Stevens (1981b) has shown that the fifth, finest resolution channel plays the most important role in determining the information we compute foveally.

One of the novel aspects of the implementation of the theory concerns the sizes of the $\nabla^2 G$ operators. Edge finding operators used in computer vision are typically at most seven pixels square; the smallest operator used in the implementation of the Marr-Hildreth theory at MIT is 35 pixels square. Not only are the resulting operators much closer approximations to the Gaussian (or any other filter for that matter), but the signal to noise characteristics of the smoothed images is vastly improved. One practical consequence of these seems to be that one can approximate differential operators (to compute the orientation of visible edges) by simple difference operators. Conventional edge finding operators confound filtering and differentiation, and have poor and essentially unpredictable filter characteristics. The first implemented version of the Marr-Hildreth theory took on the order of three hours to compute the zero crossings in the coarse channel of an image 512 pixels square. A prototype hardware implementation reduced this to thirty minutes. Nishihara and Larson (1981) report a TTL implementation which computes and displays the zero crossings in any channel of an image 128 pixels square in under 0.25 seconds.

We can reproduce the information that the Marr-Hildreth theory proposes is computed at any eccentricity ϵ by the channel which subtends m minutes of arc in the fovea. The figures used in this paper, and the arguments derived from them, rely upon examining such information, and so it should be understood how the parameters of the programs are set. If we digitise a text image, say at a resolution of μ microns (where typically $\mu = 50$), we can compute the size of mask to use in a computer program which precisely models the

information available in the any channel at eccentricity ϵ . Suppose that an average character from a text sample has width d microns. Its digitised image is

$$\frac{d}{\mu} = P$$

pixels wide. Suppose that the text is viewed from a distance of D microns, so that it subtends an angle of $\frac{d}{D}$ radians or

$$\frac{d \cdot 180 \cdot 60}{D \cdot \pi} = A$$

minutes of arc. The channel which subtends m minutes of arc in the fovea subtends roughly $2m$ minutes at 4° . Hence at eccentricity ϵ the channel subtends $\frac{m\epsilon}{2} = e$ minutes. Finally, we can choose the panel width of a digital mask corresponding to the m minute DOG at eccentricity ϵ by equating

$$\frac{w}{P} = \frac{e\sqrt{2}}{A},$$

where the $\sqrt{2}$ factor is required to account for the rotational invariance of the mask. Using this relationship, one can for example choose m to be the coarsest channel, which subtends 21 minutes of arc, ϵ to be 3° , and so determine the value of w . Examples of the result of applying this process can be found in Figure 6.

3 The isolation of words in text

3.1 Introduction

It is usual to equate early processing in reading with the extraction of character features, such as line endings, T-junctions, holes, and concavities. We are presently more concerned with an even earlier processing stage, namely the point at which the visual system first makes contact with (the gray level intensities forming the image of) a portion of text. Let us suppose for the moment that the "reading process" is already active. The work of Rayner (1975a, 1975b, 1977, 1978a, 1978b, 1979; Rayner and McConkie, 1976; Rayner, McConkie, and Ehrlich, 1978; McConkie and Rayner, 1975) and others (e.g., McConkie, 1979; O'Reagan, 1979; Levy-Schoen and O'Reagan, 1979) on eye movements demonstrates clearly that text is substantially processed before it is foveated. The extent to which eye movement control is either (1) autonomous, being entirely determined by information computed by early processing from the gray level array; or (2) is capable of being explicitly controlled by downward flowing task specific information, say, by knowledge of the syntax

Itn owb eca mee vid ent tha tth eci tym ust bea ban don eda
ton ceT her ewa sad iff ere nce ofo pin ion inr esp ect tot
heh our ofd epa rtu reT ayt ime itw asa rgu edb yso mew oul

Figure 2. Text into which spaces have been randomly introduced after elision.

and semantics of the text in question, is controversial. This is, of course, the invariance of the issue raised in Section 2.1 about system organization.

The goal of reading may be supposed to be the efficient extraction of meaning from text. Given the nature of written language, particularly English, a presumably necessary primitive subgoal is the isolation of words. In normal text, words are clearly separated by spaces which are substantially wider than the spaces between individual letters. It would seem that the “program” controlling eye movements could be trivial given a reasonable theory of the separation of words from inter-word spaces such as that provided by the Marr-Hildreth theory outlined in the previous section. Evidence in support of the contention that the control program is quite simple is easy to find. Firstly, it is well known that inter-word spaces, even when they are of varying width, are seldom foveated (Levy-Schoen, 1979; Rayner and McConkie, 1976, report that they are fixated about 10% of the time). Conversely, if spaces corresponding to word boundaries are randomly introduced into previously elided text (as shown in Figure 2), reading becomes exceptionally difficult. In this situation the inconsistent information provided by a simple space finding algorithm and its utilisation by the processes which analyze the text, produce a complex pattern of foveations and a significant increase in the duration of any individual foveation. Intermediate behavior results when inter-letter spaces are made nearly equal to those between words.

However, as is equally well known, spaces are not unique in avoiding foveation. In particular, function words such as “and” and “the” are often not foveated. This partly explains the difficulty difficulty we have in proof reading “Paris in the the spring” relative to this sentence as a whole. This raises the ever present question: how “intelligent” does the eye movement controller need to be? Is the word “the” omitted on the basis of information available in the parafovea, where individual letter recognition is poor (Bouma, 1971), or alternatively does it rely on knowledge about linguistic context?

3.2 Fisher’s results on reading transformed text

In fact, the trivial word isolation process sketched above does not work in every circumstance in which people can read quite easily. This was demonstrated in an elegant experiment performed by Fisher (1975). Building upon the earlier work of Smith (1969) and Hochberg (1970), Fisher used the trans-

1

The government of Henry the Seventh, of his son, and of his grandchildren was, on the whole, more arbitrary than that of the Plantagenets. Personal character may

4

The+government+of+Henry+the+Seventh,+of+his+son,+and+++of+his+grandchildren+was,+on+the+whole,+more+arbitrary++than+that+of+the+Plantagenets.++Personal+character+may++

7

ThegovernmentofHenrytheSeventh,ofhisson,andofhis grandchildrenwas,onthewhole,morearbitrarythanthat ofthePlantagenets.Personalcharactermayinsomedegree

2

THE GOVERNMENT OF HENRY THE SEVENTH, OF HIS SON, AND OF HIS GRANDCHILDREN WAS, ON THE WHOLE, MORE ARBITRARY THAN THAT OF THE PLANTAGENETS. PERSONAL CHARACTER MAY

5

THE@GOVERNMENT@OF@HENRY@THE@SEVENTH,@OF@HIS@SON,@AND@@@OF@HIS@GRANDCHILDREN@WAS,@ON@THE@WHOLE,@MORE@ARBITRARY@THAN@THAT@OF@THE@PLANTAGENETS.@@PERSONAL@CHARACTER@MAY@

8

THEGOVERNMENTOFHENRYTHESEVENTH,OFHISSON,ANDOFHIS GRANDCHILDRENWAS,ONTHEWHOLE,MOREARBITRARYTHANTHAT OFTHEPLANTAGENTS.PERSONALCHARACTERMAYINSOMEDEGREE

3

ThE GoVeRnMeNt oF HeNrY ThE SeVeNtH, oF HiS SoN, aNd Of HiS gRaNdChILDrEn wAs, On tHe wHoLe, MoRe aRbItRaRy ThAn tHaT Of tHe pLaNtAgEnEtS. PeRsOnAl cHaRaCtEr mAy

6

ThE@GoVeRnMeNt@oF@HeNrY@ThE@SeVeNtH,@oF@HiS@SoN,@aNd@@@Of@hIs@gRaNdChILDrEn@wAs,@On@tHe@wHoLe,@MoRe@aRbItRaRy@ThAn@tHaT@oF@tHe@pLaNtAgEnEtS.@@PeRsOnAl@cHaRaCtEr@mAy@

9

ThEgOvErNmEnToFhEnRyThEsEvEnTh,oFhIsSoN,AnDoFhIs GrAnDcHiLdReNwAs,oNtHeWhOLE,MoReArBiTrArYtHaNtHaT oFtHePlAnTaGeNeTs.pErSoNaLcHaRaCtErMaYiNsOmEdEgReE

Figure 3. The nine type and boundary variations used by Fisher.

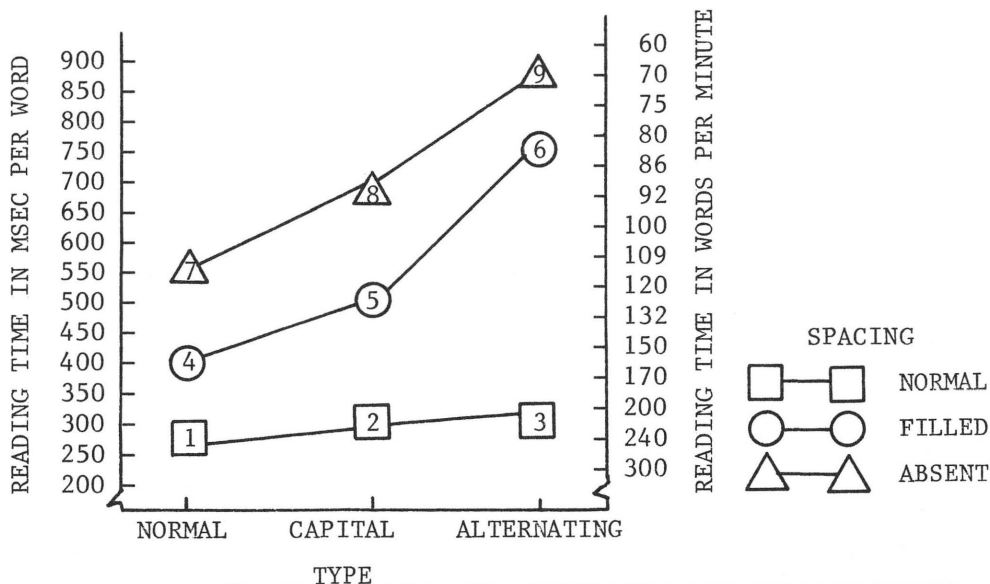


Figure 4. Fisher's results (from Fisher, 1975, page 189).

formed texts illustrated in Figure 3 to investigate the effect of manipulations of word shape and word boundary on reading. Word shape was manipulated via three *type variations*: normal, all uppercase, and alternating uppercase and lowercase letters. These are illustrated in samples one to three of Figure 3. *Word boundaries* were also manipulated in three ways: normal spacing, replacing an inter-word space by the filler character "+" or "@", and elision to remove inter-word spaces. These manipulations are illustrated for the uppercase type variation in samples two, five, and eight. In all, there are nine possible type and word boundary combinations.

Fisher recorded the length of time taken by subjects to read nine paragraphs of approximately equal length and complexity, whose texts had been randomly manipulated in the ways described above. As a safeguard against skim reading without understanding, a subject was required to answer a number of questions (typically four) about the passage just read, and was required to get a certain number correct for the data point to be recorded. The results are presented in Figure 4.

Fisher noted that the "interdependence of cues causes a reduction in reading speed to nearly one third of the speed of the separate cue manipulations", and he suggested that this "interdependence of word shape and word boundary cues tends to implicate higher order visual processing that might be required simply for word identification" (page 190).

3.3 The role of early visual processing in the isolation of words in text

In the Introduction we commented on the difficulty of devising and controlling processes which embody an interaction between upward flowing and downward flowing information, and argued for a model where early visual processing plays a bigger role. Since word isolation is clearly one of the first steps in reading, we start by examining Fisher's results more closely, in the hope of discovering an explanation of his findings without resorting to higher level cues. Firstly, the reading time per word in sample seven is significantly lower than that in sample eight. This might be explained on the grounds of the latter's lesser shape variability. However, sample nine has *greater* variability in shape than sample eight, and yet the time to read eight is significantly lower than that for nine. Similarly, there is greater variability in the shape of sample three than sample two, and yet the time to read three is significantly greater. Clearly, one possible explanation is that in the absence of spaces, capital letters can be used to signal word boundaries. According to this explanation, samples three and nine provide information (random capitals) about word boundaries inconsistent with that discovered by the processes which analyze the text. (Compare Figure 2 and its discussion in the text.) It would then follow that the eye guidance system could make the distinction between uppercase and lowercase characters and makes use of that information in isolating words.

This leads to our first empirical prediction: if the paragraphs used by Fisher are transformed by first capitalizing the initial letter of each word and then eliding, so as to appear as in Figure 5a, the resulting text should be significantly easier to read than the elided text sample shown in Figure 5b (compare sample seven in Figure 4). To test this prediction, and to maximise

Figure 5. Typical data for experiment one. 5a: text which has been elided after capitalizing the initial letter of each word. 5b: elided normal text like that in sample seven of Figure 4.

5a ItNowBecameEvidentThatTheCityMustBeAnandonedAtOnceThere
WasADifferenceOfOpinionInRespectToTheHourOfDepartureThe
DaytimeItWasArguedBySomeWouldBePreferableSinceItWouldEnable

5b ItnowbecameevidentthatthecitymustbeabandonedatonceThere
wasadifferenceofopinioninrespecttothehourofdepartureThe
daytimeitwasarguedbysomewouldbepreferablesinceitwouldenable

comparability with the results discussed above, we replicated Fisher's experiments with a number of additional font variations. The experimental details are given in the Appendix. The font variations used in the experiments are shown in Figure 13. Subjects were required to read texts which had been transformed in various ways similar to those shown in Figure 4. The average reading time per transformed word was compared for significance between two variations. According to this metric, the phrase "significantly easier to read" means that the reading per word was significantly shorter.

It turns out that the capitalized elided text shown in Figure 5a is indeed significantly easier ($p < 0.01$) to read than the elided normal text shown in Figure 5b. This supports the hypothesis that we are capable of distinguishing between uppercase and lowercase characters on the basis of information available in the parafovea. Significantly, however, it leaves open the precise details of the way in which that distinction is made.

Some evidence bearing upon this distinction can be gleaned from the results for samples five, six, eight, and nine in Figure 4. Whereas sample five is significantly easier to read than sample eight, there is insignificant difference between the ease of reading samples six and nine. This is a puzzle. The advantage of sample five over sample eight suggests that we are capable of dynamically modifying our eye movement control system to exploit the delimiter "@," and this contention is supported by the significant advantage of sample four over sample seven. However, if we are capable of distinguishing uppercase characters and the character @ in the parafovea in a way which is entirely robust and reliable, we could expect to find a similar significant advantage for sample six over sample nine; but we do not. One possible resolution of this puzzle would be to show that it is often difficult to distinguish @ and uppercase characters when they are viewed in the parafovea. If that were so, the use of @ as a filler would give some advantage in sample five relative to sample eight, but the advantage would be offset by the inconsistent information provided by fillers and text in sample six.

To investigate this question precisely, we need a detailed representation of the information which is actually available in the parafovea. Fortunately, such a representation is now available, having recently been developed by Marr and Hildreth (1980), and it was sketched in the previous section. Figure 6 shows the result of applying the digitisation process described in that section to sample five of Fisher's data (Figure 4) at an eccentricity of 4° . Figure 6b explicitly marks the convolved @ characters. It can be seen quite clearly that while some of them are relatively easy to distinguish on the basis of shape, others (for example, those marked in Figure 6c) are not.

This evidence does indeed seem to show that it is often difficult to

distinguish @ and uppercase characters when they are viewed in the parafovea. We suggest that this resolves the puzzle of Fisher's results discussed above without the need to postulate any downward flow of high level information. It further suggests that while uppercase and lowercase characters can be clearly and reliably distinguished (in most fonts), the model of "uppercase character" used by the early visual system in guiding eye movements is actually quite crude. Tentatively we may suppose that the model of an uppercase character amounts to an assertion that they are relatively large compared to those in lowercase and have relatively lower curvature. This simple model normally serves the reader well, since written text consists mostly of uppercase and lowercase characters. However, being a simple model, it is easily confused, and is particularly unreliable at making the distinction between uppercase characters and @.

A number of predictions follow from this analysis. Firstly, it suggests that a font in which the distinction between uppercase and lowercase is difficult to make on the basis of size and shape would be quite hard to read. Figure 7 shows such a font. Indeed, as we point out in the Conclusion, the analysis here can be viewed as a first step towards making font design less subjective than it has been in the past (see, for example, Spencer, 1968). Secondly, the analysis suggests that on the basis of the information available in the parafovea, it would be difficult for the visual system to distinguish the capitalized elided text shown in Figure 8a and the text filled with @ shown in Figure 8b. This translates into a prediction that there should be insignificant difference in the ease, that is to say speed per word, of reading the samples in Figure 8. Experiment 2 confirms this prediction; the relative advantage of one sample over the other failing to reach significance at the 10% level.

The same computational argument can be turned around, in which case it leads to the prediction that using a "visually striking" character as a filler would produce text that is significantly easier to read than when @ is used. Indeed, insofar as this can be shown empirically, it essentially enables us to frame a precise definition of "visually striking." In Experiment 3 we compare the effect of using "\ " and "@ " as fillers. The choice of \ was quite deliberate. Figure 9 shows a sample of text which has been digitised and convolved according to the Marr-Hildreth theory at a number of eccentricities in the manner sketched earlier. Figure 9b shows the information available way out at 9° (corresponding to about 36 letter spaces), and Figure 9c shows the instances which every one of a group of five subjects chose when they were instructed to simulate an unintelligent program to extract \ from Figure 9b. Figure 9d illustrates the information available at 7°, and shows that the subjects correctly isolated each and every instance of \. Finally, Figure 9c shows

GOVERNMENT OF HENRY THE SEVENTH, OF HIS
HIS GRANDCHILDREN WAS, ON THE WHOLE, MORE
NOT MISTAKABLE THAN IN THE PRESENT 'COURT OF COMMONS'

6a

GOVERNMENT OF HENRY THE SEVENTH, OF HIS
HIS GRANDCHILDREN WAS, ON THE WHOLE, MORE
NOT MISTAKABLE THAN IN THE PRESENT 'COURT OF COMMONS'

6b

GOVERNMENT OF HENRY THE SEVENTH, OF HIS
HIS GRANDCHILDREN WAS, ON THE WHOLE, MORE
NOT MISTAKABLE THAN IN THE PRESENT 'COURT OF COMMONS'

6c

Figure 6. The result of convolving sample five of Fisher's data to show the information available at 4°. 6b: all instances of the character @. 6c: instances of the character @ which are difficult to distinguish on the basis of shape.

Figure 7. A font in which the distinction between uppercase and lowercase would be difficult to make. It is reproduced from Spencer (1968, page 16): "A new kind of type proposed in the 1880's by Andrew Tuer in which 'the tailed letters projecting above or below the line, have been docked' to provide maximum type size 'where economy of space is an object - as in the crowded columns of a newspaper'."

An inquiry which has just been held at Brighton once more illustrates the kind of leading strings in which local municipalities are kept. An inspector of the local Government Board has been holding a kind of public inquest on the proposal of the Brighton Corporation to borrow 55,000*l.* This enterprising public body in its desire to in-

8a ItNowBecameEvidentThatTheCityMustBeAbandonedAtOnceThere
WasADifferenceOfOpinionInRespectToTheHourOfDepartureThe
DaytimeItWasArguedBySomeWouldPreferableSinceItWouldEnableThem

8b It@now@became@evident@that@the@city@must@be@abandoned@at@
once@There@was@a@difference@of@opinion@in@respect@to@the@
hour@of@departure@The@daytime@it@was@argued@by@some@would@be@

Figure 8. 8a: text sample in which words have been elided following capitalizing each initial letter. 8b: text in which spaces have been filled by @ (compare Figure 4, sample 5).

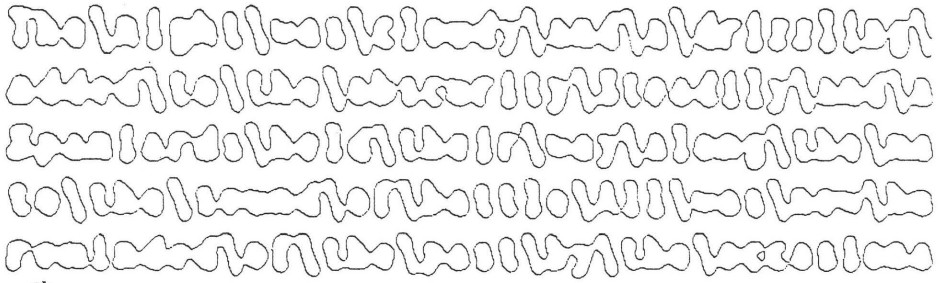
the information available at 4° . It is clear that the early visual system could more easily and reliably find instances of \ than @, and so we are led to predict that the Fisher like sample of text shown in Figure 9a would be significantly easier to read than the same thing with \ replaced by @. Experiment 3 confirms this prediction. Indeed, in Experiment 4, we compared the visually striking filler \ and normal spacing (sample 1 of Figure 4), and we find that the relative advantage of normal spacing fails to reach significance even at the 10% level.

The final Experiment 5 is a tribute to the versatility of the computing facilities available for this research. Consider the text sample in Figure 10a, in which the forward slash character is used as a delimiter. Since the downstrokes of ascender characters such as b and f slope slightly forwards but not nearly so much as the slope of /, we would expect a similar significant advantage for / over @. It turns out that this is the case. More interestingly, we were able to design a font in which the only change compared to that of characters in Figure 10a is that the forward slash character had precisely the same slope as the downstroke of an ascender (see Figure 10b). Figures 10c and 10d show the convolved images of the samples in Figures 10a and 10b, respectively. The analysis developed above leads us to predict that text samples of the form shown in Figure 10a will be significantly easier to read than those in the special font shown in Figure 10b, though we might expect that there will be a reduced advantage compared to that shown by / or \ over @. Experiment 5 confirms this prediction, the significance being only at the 5% level.

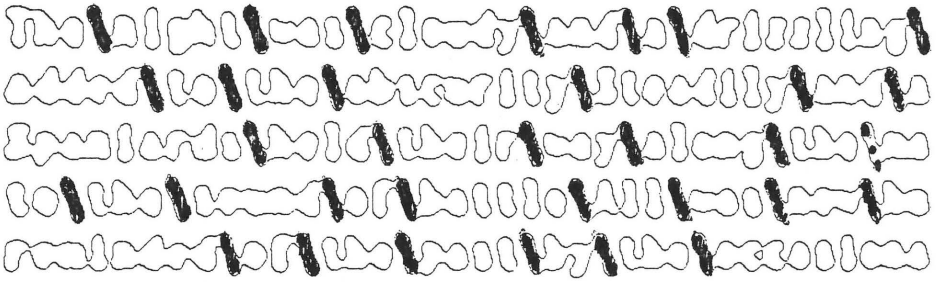
Figure 9. 9a: text sample in which \ is used as a filler between words. 9b: resulting of convolving the sample in (a) to show the information available at 9° . 9c: instances of \ found in (b) by a group of subjects simulating an unintelligent program. 9d: information available in the convolved image at 7° eccentricity. 9c: information available to early visual processing at 4° .

The\night\was\cloudy\and\a\drizzling\rain\which\fell\without
intermission\added\to\the\obscurity\Steadily\and\as\noiselessly\as
possible\the\Spaniards\held\their\way\along\the\main\street\which

9a



9b



9c

The\night\was\cloudy\and\a\drizzling\rain
added\to\the\obscurity\Steadily\and\as
Spaniards\held\their\way\along\the\main
to\the\sound\of\battle\All\was\now\hu
reindeed\of\the\past\by\the\occasional

9d

The\night\was\cloudy\and\a\drizzling\rain
added\to\the\obscurity\Steadily\and\as
Spaniards\held\their\way\along\the\main
to\the\sound\of\battle\All\was\now\hu
reindeed\of\the\past\by\the\occasional

9e

Figure 10. 10a: text sample filled with /. 10b: text sample in the special font in which the forward slash character has precisely the same slope as the ascender of d. 10c: convolved image of (a) at 4°. 10d: convolved image of (b) at 4°.

It/now/became/evident/that/the/city/must/be/abandoned/at/once/The/difference/of/opinion/in/respect/to/the/hour/of/departure/The/day/argued/by/some/would/be/preferable/since/it/would/enable/them/to

10a

It|now|became|evident|that|the|city|must|be|abandoned|at|once|The|difference|of|opinion|in|respect|to|the|hour|of|departure|The|day|argued|by|some|would|be|preferable|since|it|would|enable|them|to

10b

*It/now/became/evident/that/the/city/must
difference/of/opinion/in/respect/to/the/hour
argued/by/some./would/be/preferable./sin*

10c

*It|now|became|evident|that|the|city|must|be
difference|of|opinion|in|respect|to|the|hour
by|some.|would|be|preferable.|since|it|woul*

10d

4 Conclusion

This paper began by sketching the background against which this investigation of word isolation in the parafovea has been conducted. Our aim has been to show how published empirical data, especially that of Fisher (1975), could be accounted for using the rich theories of early visual processing of the natural world which have recently been developed in artificial intelligence. On the basis of a precise representation of the information available in the parafovea, we proposed an explanation of Fisher's results by postulating a crude, though mostly reliable, model of uppercase versus low-

erace characters. The same computational evidence led us to frame a number of predictions, each of which was then confirmed by psychophysical experimentation. As a side effect, we were required to consider how the idea of a character being “visually striking” might be made precise. This approach provides a method for the study of legibility to add to those listed by Spencer (1968, page 21).

As we pointed out in the Introduction, this study is merely the first step on the long haul towards understanding through computation the exquisite human skill of reading. The results reported here have encouraged us to proceed to consider the next step in the process of acquiring meaning from the sea of gray level intensities which form the image. We consider the next step to be the problem of integrating information over successive saccades. Rayner’s work (1975a, 1975b, 1977, 1978a, 1978b, 1979; Rayner and McConkie, 1976; Rayner, McConkie, and Ehrlich, 1978; McConkie and Rayner, 1975) provides a rich background of empirical data for our study, which is intended to exploit detailed computational models of natural vision in the manner of this paper. It is clear for example that the notion of word shape needs to be made more precise by defining an appropriate representation of the information available when a word is convolved at 2° . Rayner’s (1975, page 76) finding that the first and last letters of a word (his NS condition) cause a significant increase in foveation duration is entirely consistent with the approach pursued here. When two nearby lines are convolved, they produce a smeared blob. This occurs not only for strokes within a character, but for nearby strokes of two adjacent characters (Figure 11). Such inter-character smearing confounds any process whose goal is to elicit structure within a word, and in particular to discover the precise locations of its individual characters. The extremal characters are relatively unaffected by such intercharacter smearing, and hence the information gleaned at 4° will closely match that computed on a subsequent (foveal) saccade. A similar argument applies to ascenders and descenders, so long as they are relatively isolated. It is not inconceivable that we have learned that such shape information at the extremities of words and from isolated ascenders and descenders within a word are preserved over a typical 2° saccade, and have based our word representation scheme, which develops over several such saccades, and

Figure 11. The smearing of nearby lines by convolution. Left: “emi” – for strokes within a character. Right: “nWh” – between two characters.



the corresponding processes for eliciting substructure, upon it. Further study is needed to make the representation and matching process precise.

For the moment at least, we are left with a reasonably detailed model of eye movement control whose goal is the isolation of words in text on the basis of the information which is available in the parafovea.

1. We can reliably isolate spaces above a size which is yet to be determined, but is about one character space in normal text. We assume that such spaces delimit words, and mostly this inference serves us well. We are confused (and our reading is inhibited) when they do not.

If a space is located on either side of a blob which subtends a visual angle of roughly the same size as an individual saccade, we initiate an eye movement to the beginning of the as yet unprocessed blob. O'Regan's (1979) data gives us some evidence on which to develop the details of this process. In particular, the control may involve a crude representation of the sort discussed earlier for uppercase characters, in which case it would presumably be easy to confuse. Again, this requires detailed investigation.

2. If spaces are not available, but words are delimited by some filler character, we dynamically adjust our scanning strategy to locate instances of that filler. This requires that we first compute a description of the appearance of the filler in the parafovea, and secondly that we search for instances of the description in the convolved parafoveal image. This strategy is reliable to the extent that the filler is "visually striking," that is to say, its instances can reliably be extracted from the available information. The backwards and forwards slash characters are visually striking in this sense, the @ sign is less so. It is expected that the first foveation of text in which spaces are routinely filled in this way would be considerably longer than subsequent ones (there is some evidence that this is generally true; see Levy-Schoen, 1979, page 12). It may be conjectured that this can be explained on the basis of the considerations discussed in this paper.

In particular, our model leads to the following prediction. Consider a text sample which consists of a sequence of "segments," each of which can be several words long and is associated with a particular filler character. For example, a segment filled with \ might be followed by a segment filled with / and so on. We would expect that there would be a significant increase in the duration of foveations at the boundary between two segments as the parafoveal processing fails to discover an instance of its currently "loaded" filler, and has to locate and load the description of the filler for the next segment.

3. We distinguish between uppercase and lowercase characters on the basis of size and lower curvature only. Capital letters mark important linguistic events in English, such as proper names and the beginnings of sentences. As

before, we assume that this importance has been translated into a coarse description which often can be reliably computed in the parafovea. While it often serves us well in isolating uppercase characters and drawing our attention to the corresponding linguistic event, it is a coarse description and is easily confused.

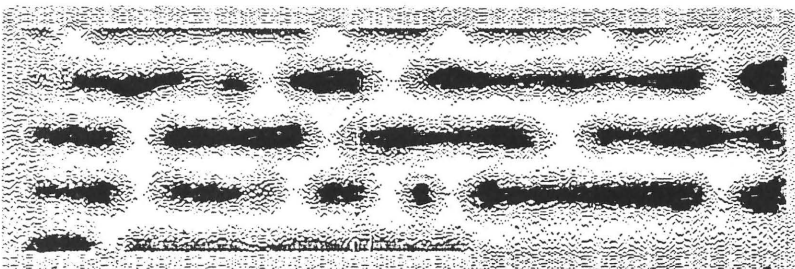
Other work, not reported in detail here, shows a slight though not statistically significant advantage over sample seven in Figure 4 for a word sequence in which words are alternately printed in a roman font and in italics. This effect is less than that which occurs when bold font is alternated with regular roman. This is consistent with the findings of legibility research. Various researchers, including Tinker (1955), have found that italics actually retard reading and that readers mostly do not like italics. Tinker (1955) found that 96% of his adult subjects were of the opinion that they could read lowercase roman more easily than italics.

This study assumes that the word isolation process is already activated at the time when the text is initially encountered, and it might be thought that high level knowledge would be required to effect this activation. Figure 12c shows a sample of text (Figure 12a) convolved with a mask size which corresponds to foveation at a distance of 5.83 metres. The regular texture of lines of blobs is quite clear, even though it is impossible to make any sense of the text. In short, the image looks like text even at a distance, as does the image in Figure 12g, although in this case it is in fact the convolution of the image shown in Figure 12d. Once again, the theory being advanced here is that we interpret a particular image as a piece of text on the basis of quite a crude representation, which, however, mostly serves us well.

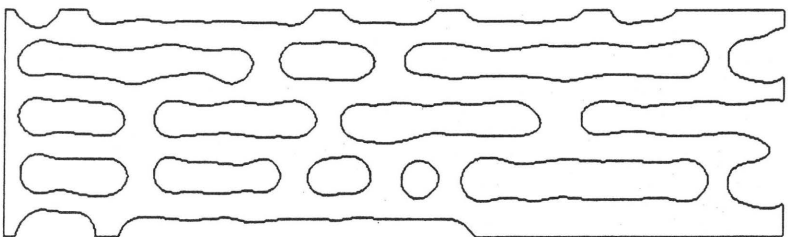
We conclude with one final remark on the notion that the ease with which a text can be read is directly related to the ease with which information can be reliably computed from its convolved image, and it concerns font design. A great deal of research on font design (see, for example, Spencer, 1968) is depressingly subjective. Recently however, Julesz (1980) and his colleagues have begun a study which is analogous to that pursued here. They apply their ideas about texture discrimination to define a set of so-called "textons" and then advocate the design of fonts based on the discriminability of textons. Our approach also relates the legibility of a font to the processes of natural perception, but we are currently more concerned with understanding the perceptual basis of the efficacy of using serifs and so forth than with the aesthetics of font design. There is nevertheless a good deal of similarity between our goals. Much more work is necessary to develop the ideas sketched in this section into a coherent and precise theory.

i scanned the text at roughly the resolution of the human system, viewing the text at a distance of 30 centimeters.

12a



12b



12c

i scanned the text at roughly the resolution of the human system, viewing the text at a distance of 30 centimeters.

12d

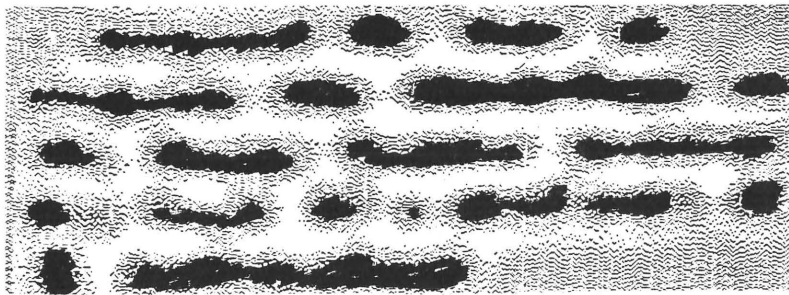
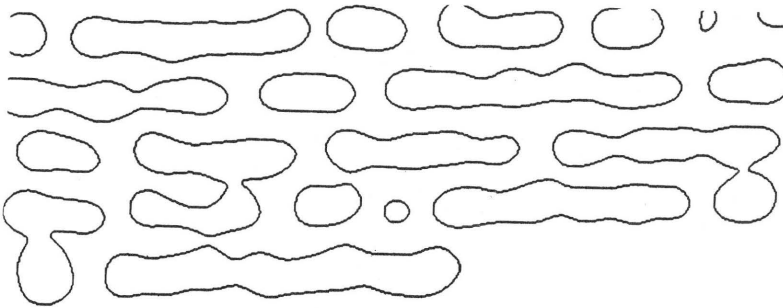
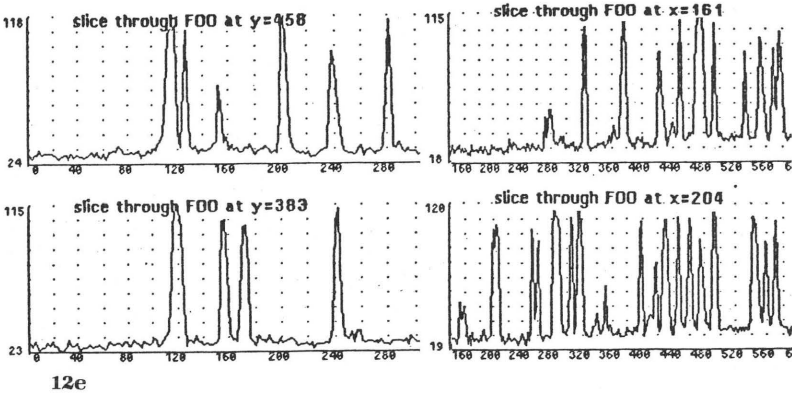


Figure 12. 12a: a sample of text displayed, after photoc scanning at a resolution of 100 microns, using a psuedo grey level system devised and constructed by Berthold Horn. 12b: the result of convolving the text in 12a with a mask whose central panel width is 36. This corresponds to foveating the text at a distance of 5.83 metres. 12c: zero crossings of the convolution shown in 12b. The pattern of blobs corresponding to words is evident. 12d: a set of random marks produced by filling in the regions which arise from tracing around the text sample given in 12a. 12e: a number of cross sections of the intensity profile shown in 12d in the x and y directions. 12f: the result of convolving the image shown in 12e in the same way as 12b. 12g: the zero crossings of the convolution in 12f. The result is quite similar to 12c.

Appendix: Experimental details

The experiments were designed strictly in accordance with the method devised by Fisher (1975) to maximize comparability with his results.

Method. Twelve members of the Artificial Intelligence Laboratory who were naive with regard to the purpose of the experiment took part.

Materials. The nine paragraphs of the 1960 revised Nelson Denny Reading Text (Denny, 1960) were used, together with three paragraphs of similar length (about 200 words) and complexity. The Nelson Denny texts were used by Fisher because they "had a very high degree of standardization from high school through college aged adults" (Fisher 1975). A Times Roman 10-point font was used throughout the experiments. There were several variations to the basic font:

It now became evident that the city must be abandoned at once. There was
(i)

ItnowbecameevidentthatthecitymustbeabandonedatonceTherewasadifferenceof
(ii)

ItNowBecameEvidentThatTheCityMustBeAbandonedAtOnceThereWasADifferenceOf
(iii)

It@now@became@evident@that@the@city@must@be@abandoned@at@once@There@was
(iv)

It\now\became\evident\that\the\city\must\be\abandoned\at\once\There\was
(v)

It/now/became/evident/that/the/city/must/be/abandoned/at/once/There/was
(vi)

it|now|became|evident|that|the|city|must|be|abandoned|at|once|There|was
(vii)

i regular spacing between words ("normal").

ii all words elided together, that is, inter-word spacing removed.

iii words elided together after the initial letter of each word had been capitalised

iv inter-word spaces filled by "@".

v inter-word spaces filled by "\".

vi inter-word spaces filled by "/".

vii inter-words spaces filled by a special character of the same slope as the descenders in the font.

The experiments (1-5) described in Section 3 were designed to compare the relative ease of reading several pairs of the variations listed above. Specifically, the following hypotheses were tested:

- 1 ii vs iii: It was hypothesized that it would be significantly easier to read variation iii than variation ii.
- 2 iii vs iv: It was hypothesized that there would be insignificant difference between the ease of reading variations iii and iv.
- 3 iv vs v: It was hypothesized that it would be significantly easier to read variation v than variation iv. A similar hypotheses was that variation vi would show significant advantage over iv.
- 4 i vs v: It was hypothesized that there would be insignificant difference between the ease of reading variations i and v.
- 5 vi vs vii: It was hypothesized that it would be significantly easier to read variation vi than variation vii.

The variations i to vii were divided into two overlapping sets i, ii, iii, iv, v and ii, iii, iv, vi, vii. The subjects were divided into two groups of six and each group was associated with one of the two sets of variations. Each subject had an individually prepared booklet consisting of the twelve paragraphs. The booklets comprised two instances of paragraphs in three of the variations and three instances of two of the variations. The choices of variations and the order of presentation of the variations was counterbalanced over all subjects. "After each paragraph, a set of four multiple choice questions was presented which had to be answered. The questions were taken from the Nelson Denny Reading Test. A digital clock graduated in [steps of 0.1 second] provided a display of the time to read and was clearly visible to all subjects" (Fisher, 1975).

Procedure. Each subject was given a page of instructions containing the variations of text which would appear, the individually prepared booklet of twelve paragraphs, and a question and answer sheet. "When subjects finished reading, they were to look at the time . . . They were then to turn the page, answer the questions, and wait for instructions to go on to the next paragraph" (Fisher, 1975).

Results. As there was a substantial spread in the reading speed of the subjects, averaging the data points over all subjects for a particular text produces an unacceptably large standard deviation. As we are in fact most interested in the relative ease of reading two variations, the relevant hypothesis for comparing one text variation α against another β is the null hypothesis:

$$H_0: \mu \left[\frac{\alpha}{\beta} \right] = 1.$$

We can use the simple t statistic defined by

$$t = \frac{r - 1}{s\sqrt{\frac{1}{v}}}$$

where there are $v + 1$ subjects, r is the mean of the individual values of $\frac{t_\alpha}{t_\beta}$, where t_α is the time taken per word to read the paragraphs in variation α , and s is the standard deviation of that measure from r . The actual results were given in Section 3.

REFERENCES

- Allport, A. Word recognition in reading. In Kolers, P.A., Wrolstad, M.E., and Bouma, H. (Eds.), *Processing of visible language*, Vol. I. New York: Plenum Press, 1979.
- Barrow, H.G., and Tenenbaum, J.M. Experiments in interpretation guided semantics. Technical Note, 123, SRI International, 1976.
- Barrow, H.G., and Tenenbaum, J.M. Recovering intrinsic scene characteristics from images. In Hanson, A., and Riseman, E.M. (Eds.), *Computer vision systems*. New York: Academic Press, 1978.
- Bouma, H. Visual recognition of isolated lowercase letters, *Vision Research* 1971, 11, 459-474.
- Brady, Michael. The development of a computer vision system (in Italian: English version available from the author), *Ricerche Psicologiche*, 1979.
- Brady, Michael. The changing shape of computer vision, *Artificial Intelligence* (Special issue on computer vision), 1981, 17, 1-18.
- Brady, Michael. Artificial intelligence approaches to image understanding. In Kittler, J. (Ed.), *Pattern recognition and its applications*. Amsterdam: Reidel Press, 1981.
- Brady, Michael, and Wielinga, B.J. Reading the writing on the wall. In Hanson, A., and Riseman, E.M. (Eds.), *Computer vision systems*. New York: Academic Press, 1978.
- Binford, T.O. Inferring surfaces from images, *Artificial Intelligence* (Special issue on computer vision), 1981, 17.
- Cohen, Gillian. Speech perception and reading. In *Cognitive psychology* (part 2). Milton Keynes: Open University Press, 1978.
- Davis, Larry S., and Henderson, T.C. Hierarchical constraint processes for shape analysis. Computer Sciences Dept., University of Texas, Austin, TR-115, 1979.
- Davis, Larry S., and Rosenfeld, Azriel. Hierarchical relaxation for waveform parsing. In Hanson, A., and Riseman, E.M. (Eds.), *Computer vision systems*. New York: Academic Press, 1978.
- Davis, Larry S. and Rosenfeld, Azriel. Cooperating processes for low-level vision: a survey, *Artificial Intelligence* (special issue on computer vision), 1981, 17.
- Duda, R. and Hart, P. *Pattern classification and scene analysis*. New York: John Wiley, 1973.
- Estes, W.K. Interaction of perception and memory in reading. In Laberge, David, and Samuels, S. Jay (Eds.), *Basic processes in reading: perception and comprehension*. New York: John Wiley, 1977.

- Fisher, D.F. Reading and visual search, *Memory and Cognition*, 1975, 3, 188-196.
- Frisby, J.P. *Seeing*. Oxford: Oxford University Press, 1979.
- Gough, P.B. One second of reading. In Kavanagh, J.F., and Mattingly, I.G. (Eds.), *Language by ear and by eye*. Cambridge: MIT Press, 1972.
- Grimson, W.E.L. *From images to surfaces: a computational study of the human early visual system*. Cambridge: MIT Press, 1981.
- Hildreth, E.C. Implementation of a theory of edge detection (MS dissertation), also AI-TR 579, MIT, 1980.
- Henderson, L. Word recognition. In Sutherland, N.S. (Ed.), *Tutorial essays in experimental psychology*. Potomac: Erlbaum, 1977.
- Hochberg, J. Components of literacy: speculation and exploratory research. In Levin, H., and Williams, J.P. (Eds.), *Basic studies in reading*. New York: Basic Books, 1970.
- Horn, B.K.P. Determining lightness from an image, *Computer Graphics and Image Processing*, 1974, 3, 277-299.
- Horn, B.K.P. Understanding image intensities, *Artificial Intelligence*, 1977, 8, 201-231.
- Huey, E.B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908.
- Julesz, B. Spatial nonlinearities in the instantaneous perception of textures with identical power spectra, *Phil. Trans. R. Soc. Lond. B*, 1980, 290, 83-94.
- Lesser, Victor R., and Erman, Lee D. A retrospective view of the Hearsay-II architecture, *Proc. Int. Jt. Conf. Artificial Intelligence*, 1977, 2, 790-800.
- Levy-Schoen, A., and O'Regan, K. The control of eye movements in reading. In Kolars, P.A., Wrolstad, M.E., and Bouma, H. (Eds.), *Processing of visible language*, Vol. I. New York: Plenum Press, 1979.
- Marcel, T., and Patterson, Karalyn. Word recognition and production: reciprocity in clinical and normal studies. In Requin, J. (Ed.), *Attention and performance*. Potomac: Erlbaum, 1979.
- Marr, D. Early processing of visual information. *Phil. Trans. R. Soc. Lond. B*, 1976, 275, 483-524.
- Marr, D. *Vision*. San Francisco: Freeman, 1981.
- Marr, D., and Hildreth, E. Theory of edge detection, *Proc. R. Soc. Lond. B*, 1980, 207, 187-217.
- Marr, D., Hildreth, E., and Poggio, T. Evidence for a fifth, smaller channel in early human vision. MIT, AI memo 541, 1979.
- Marr, D., Palm, G., and Poggio, T. Analysis of a cooperative stereo algorithm, *Biol. Cybernetics*, 1978, 28, 223-229.
- Marr, D., and Nishihara, H.K. Representation and recognition of the spatial organisation of three dimensional structure, *Proc. R. Soc. London B*, 1978, 200, 269-294.
- Marr, D., and Poggio, T. Cooperative computation of stereo disparity, *Science*, 1976a, 194, 283-287.
- Marr, D., and Poggio, T. Cooperative computation of stereo disparity. MIT, AI Memo 364, 1976b.
- Marr, D., and Poggio, T. A theory of human stereo vision, *Proc. R. Soc. Lond. B*, 1979, 204, 301-328.
- Marr, D., and Ullman, S. directional selectivity and its use in early visual processing, *Proc. R. Soc. Lond. B*, 1981, 211, 151-180.

- McClelland, James L., and Rumelhart, David E. an interactive activation model of the effect of context in perception (part 1). Center for Human Information Processing, University of California, San Diego, 8002, 1980.
- McConkie, G. W. On the role and control of eye movements in reading. In Kolers, P. A., Wrolstad, M. E., and Bouma, H. (Eds.), *Processing of visible language*, Vol. I. New York: Plenum Press, 1979.
- McConkie, G. W., and Rayner, K. The span of the effective stimulus during a fixation in reading, *Perception and Psychophysics*, 1975, 17, 576-586.
- Minsky, M. and Papert, S. Artificial intelligence progress report, MIT, AI Memo 252, 1972.
- Nash-Webber, B. L. The role of semantics in automatic speech understanding. In Bobrow, D., and Collins, A. (Eds.), *Representing and understanding: studies in cognitive science*. New York: Academic Press, 1975.
- Nelson. *The Nelson Denny reading test*. Boston: Houghton-Mifflin Company, 1960.
- Nishihara, H. K., and Larson, N. G. Toward a real time implementation of the Marr-Poggio stereo matcher. In Baumann, Lee (Ed.), *Proceedings of the image understanding workshop*. McLean: Science Applications, 1981.
- O'Regan, Kevin. Moment to moment control of eye saccades as a function of textual parameters in reading. In Kolers, P. A., Wrolstad, M. E., and Bouma, H. (Eds.) *Processing of visible language*, Vol. I. New York: Plenum Press, 1979.
- Rayner, K. The perceptual span and peripheral cues in reading, *Cognitive Psychology*, 1975a, 7, 65-81.
- Rayner, K. Parafoveal identification during a fixation in reading, *Acta Psychologica*, 1975b, 39, 271-282.
- Rayner, K. Visual attention in reading: eye movements reflect cognitive processes, *Memory and Cognition*, 1977, 5, 443-448.
- Rayner, K. Eye movements in reading and information processing, *Psychological Bulletin*, 1978a, 85, 618-660.
- Rayner, K. Foveal and parafoveal cues in reading. In Requin, J. (Ed.), *Attention and performance*. Potomac: Erlbaum, 1978b.
- Rayner, K. Eye movements in reading: eye guidance and integration. In Kolers, P. A., Wrolstad, M. E., and Bouma, H. (Eds.), *Processing of visible language*, Vol. I. New York: Plenum Press, 1979.
- Rayner, K., and McConkie, G. W. What guides a reader's eye movements? *Vision Research*, 1976, 16, 829-837.
- Rayner, K., McConkie, G. W., and Ehrlich, S. Eye movements and integrating information across fixations, *Jour. Exp. Psychology: Human Perception and Performance*, 1978, 4, 529-544.
- Richter, J. and Ullman, S. A. model for the spatio-temporal organization of X and Y-type ganglion cells in the primate retina. MIT, AI Memo 573, 1980.
- Rosenfeld, Azriel, Hummel, R. A., and Zucker, S. W. Scene labelling by relaxation operations, *IEEE Trans. Systems, Man, Cybernetics*, 1976, SMC-6, 420-433.
- Rumelhart, David E. Toward an interactive model of reading. In Dornic, S. (Ed.), *Attention and performance*. Potomac: Erlbaum, 1977.
- Schank, R., Goldman, N., Rieger, C., and Riesbeck, C. MARGIE: memory analysis response generation and inference on English, *Proc. Int. Jt. Conf. Artificial Intelligence*, 1973, 3, 255-261.

- Smith, F. Familiarity of configuration vs. discriminability of features in the visual identification of words. *Psychonomic Science*, 1969, 14, 261-262.
- Spencer, H. *The Visible Word*. London: The Royal Academy, 1968.
- Stevens, K. A. On the visual detection of fine detail. MIT (in preparation), 1981.
- Ullman, S. *The interpretation of visual motion*. Cambridge: MIT Press, 1978.
- Ullman, S. Relaxation and constrained optimisation by local processes, *Computer Graphics and Image Processing*, 1979, 10, 115-125.
- Waltz, D. L. A parallel model for low level vision. In Hanson, A., and Riseman, E. M. (Eds.), *Computer vision systems*. New York: Academic Press, 1978.
- Wilson, H. R., and Giese, S. C. Threshold visibility of frequency gradient patterns, *Vision Research*, 1977, 17, 1177-1190.
- Wilson, H. R., and Bergen, J. R. A four mechanism model for spatial vision, *Vision Research*, 1979, 19, 19-32.
- Zucker, S. W., Leclerc, Y. G., and Mohammed, J. L. Continuous relaxation and local maxima selection – conditions for equivalence, *Proc. Int. Jt. Conf. Artificial Intelligence*, 1979, 6, 1014-1016.

This paper originally appeared as MIT Artificial Intelligence Memo AIM-593. Support for the Laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643. The author would like to thank the following people for valuable discussions at various stages of the research described here: Bob Berwick, Phil Gough, Eric Grimson, Ellen Hildreth, David Marr, Marilyn Matz, Keith Rayner, Kobi Richter, Shimon Ullman. The referees' comments were valuable. This research owes a great deal to the flexibility of the software and hardware facilities available at the Artificial Intelligence Laboratory. In particular, Bob Sjoberg provided the font generation system and was always ready to give help about its use. Keith Nishihara implemented most of the programs which were used in this research, including the display system and edge detection programs. Ellen Hildreth carried out some of the earlier experiments, including the production of Figure 12.

*This special double issue of Visible Language
on visual factors in the reading process
is being published in two parts,
in accordance with regulations
for the journal's mailing permit.
Contents of the second half are given
on the first page of this issue.*

The Authors

Keith Rayner is professor of psychology and director of the Center for Cognitive Science at the University of Massachusetts (Amherst, MA 01003). He received his doctorate from Cornell University and from 1973-1978 was assistant professor of education, psychology, and visual science at the University of Rochester. His primary research interests are in perceptual and language processes in reading.

Robert Morrison is a doctoral-candidate in psychology at the University of Massachusetts (Amherst, MA 01003). He is interested in eye movement control and reading.

Albrecht-Werner Inhoff is a doctoral candidate in psychology at the University of Massachusetts and at Christian-Albrechts-Universität in Kiel, West Germany. He is interested in reading, eye movements, and parafoveal word processing.

Ralph Norman Haber is professor of psychology at the University of Illinois at Chicago Circle (Chicago, IL 60680), where he has been since 1979. Previously he taught and was chairman of the Department of Psychology at the University of Rochester. Dr. Haber's main research interests are in visual perception, including how we perceive the layout of space in scenes and in pictures, and in the perceptual components of reading.

Lyn R. Haber is associate professor of psychology at the University of Illinois at Chicago Circle, where she has been since 1979. Her doctorate is in linguistics from the University of California at Berkeley in 1970. Her main research interests are in language acquisition and in language disabilities.

Michael Brady is a senior research scientist with the Artificial Intelligence Laboratory at Massachusetts Institute of Technology (Cambridge, MA 02138). His doctorate in mathematics is from the Australian National University in Canberra. His current research interests are in surface perception, reading, and trajectory control of robot arms. He has published *A Programming Approach to the Theory of Computer Science* (Wiley, 1977) and edited four other books.