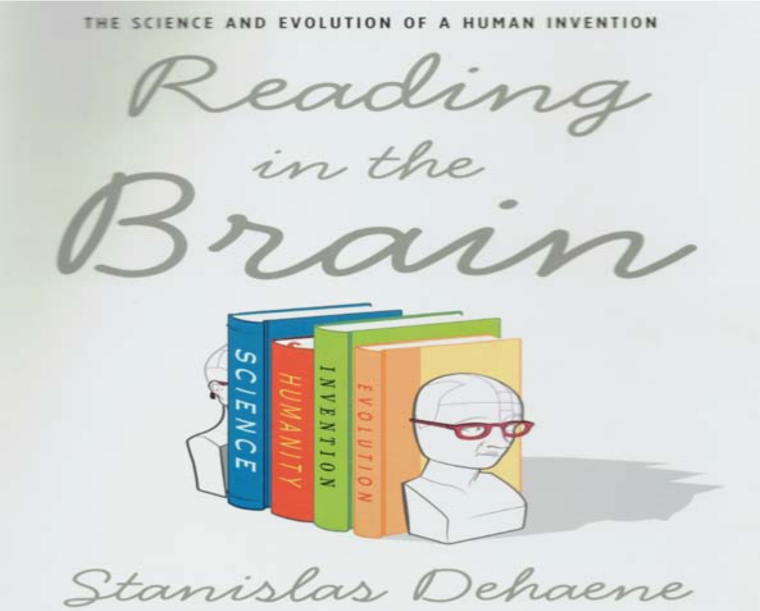




## The Search for Invariants

Reading poses a difficult perceptual problem. We must identify words regardless of how they appear, whether in print or handwritten, in upper- or lowercase, and regardless of their size. This is what psychologists call the *invariance problem*: we need to recognize which aspect of a word does not vary—the sequence of letters—in spite of the thousand and one possible shapes that the actual characters can take on.

If perceptual invariance is a problem, it is because words are not always in the same location, in the same font, or in the same size. If they were, just listing which of the cells on the retina are active and which are not would suffice to decode a word, much like a black-and-white computer image is defined by the list of its pixels. In fact, however, hundreds of different retinal images can stand for the same word, depending on the form in which it is



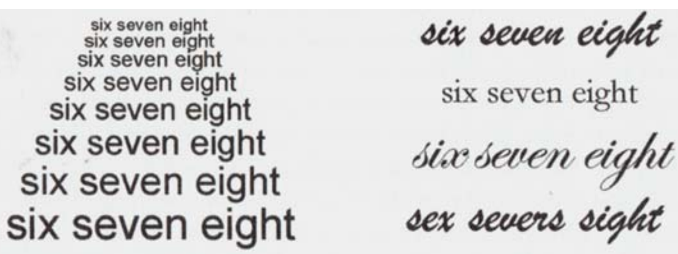


**Figure 1.2** Visual invariance is one of the prime features of the human reading system. Our word recognition device meets two seemingly contradictory requirements: it neglects irrelevant variations in character shape, even if they are huge, but amplifies relevant differences, even if they are tiny. Unbeknownst to us, our visual system automatically compensates for enormous variations in size or font. Yet it also attends to minuscule changes in shape. By turning an "s" into an "e," and therefore "sight" into "eight," a single mark drastically reorients the processing chain toward entirely distinct pronunciations and meanings.

written (figure 1.2). Thus one of the first steps in reading must be to correct for the immense variety of those surface forms.

Several cues suggest that our brain applies an efficient solution to this perceptual invariance problem. When we hold a newspaper at a reasonable distance, we can read both the headlines and the classified ads. Word size can vary fiftyfold without having much impact on our reading speed. This task is not very different from that of recognizing the same face or object from a distance of two feet or thirty yards—our visual system tolerates vast changes in scale.

A second form of invariance lets us disregard the location of words on the page. As our gaze scans a page, the center of our retina usually lands slightly left of the center of words. However, our targeting is far from perfect, and our eyes sometimes reach the first or last letter without this preventing us from recognizing the word. We can even read words presented on the periphery of our visual field, provided that letter size is increased to compensate for the loss of retinal resolution. Thus size constancy goes hand in hand with normalization for spatial location.

Finally, word recognition is also largely invariant for character shape. Now that word processing software is omnipresent, technology that was formerly reserved to a small elite of typographers has become broadly available. Everyone



knows that there are many sets of characters called "fonts" (a term left over from the time when each character had to be cast in lead at a type foundry before going to the press). Each font also has two kinds of characters called "cases," the UPPERCASE and the lowercase (originally the case was a flat box divided into many compartments where lead characters were sorted; the "upper case" was reserved for capital letters, and the "lower case" for the rest). Finally, one can choose the "weight" of a font (normal or **bold** characters), its inclination (*italics*, originally invented in Italy), whether it is underlined or not, and any **combination** of *these options*. These well-calibrated variations in fonts, however, are nothing compared to the enormous variety of writing styles. Manuscript handwriting obviously takes us to another level of variability and ambiguity.

In the face of all these variations, exactly how our visual system learns to categorize letter shapes remains somewhat mysterious. Part of this invariance problem can be solved using relatively simple means. The vowel "o," for instance, can easily be recognized, regardless of size, case, or font, thanks to its unique closed shape. Thus, building a visual o detector isn't particularly difficult. Other letters, however, pose specific problems. Consider the letter "r," for instance. Although it seems obvious that the shapes r, R, *r* and *r* all represent the same letter, careful examination shows that this association is entirely arbitrary—the shape e, for instance, might serve as well as the lowercase version of the letter "R." Only the accidents of history have left us this cultural oddity. As a result, when we learn to read, we must not only learn that letters map onto the sounds of language, but also that each letter can take on many unrelated shapes. As we shall see, our capacity to do this probably comes from the existence of abstract letter detectors, neurons that can recognize the identity of a letter in its various guises. Experiments show that very little training suffices to DeCoDe, At An EsSeNtIaLly NoRmAl SpEeD, EnTiRe SeNtEnCes WhOsE LeTtErS HaVe BeEn PrInTeD AlTeRnAtElY iN uPpErCaSe aNd In LoWeRcAsE.[15] In the McConkie and Rayner "evil genius" computer, this letter-case alternation can be changed in between every eye saccade, totally unbeknownst to the reader![16] In our daily reading experience, we never see words presented in alternating case, but our letter normalization processes are so efficient that they easily resist such transformation.

In passing, these experiments demonstrate that global word shape does



not play any role in reading. If we can immediately recognize the identity of "words," "WORDS," and "WoRdS," it is because our visual system pays no attention to the contours of words or to the pattern of ascending and descending letters: it is only interested in the letters they contain. Obviously, our capacity to recognize words does not depend on an analysis of their overall shape.