Writing Systems:
Their Properties and Implications for Reading

Brett Kessler and Rebecca Treiman

doi:10.1093/oxfordhb/9780199324576.013.1

Draft of a chapter to appear in:
The Oxford Handbook of Reading, ed. by Alexander Pollatsek and Rebecca Treiman.
ISBN 9780199324576.

Abstract

An understanding of the nature of writing is an important foundation for studies of how people read and how they learn to read. This chapter discusses the characteristics of modern writing systems with a view toward providing that foundation. We consider both the appearance of writing systems and how they function. All writing represents the words of a language according to a set of rules. However, important properties of a language often go unrepresented in writing. Change and variation in the spoken language result in complex links to speech. Redundancies in language and writing mean that readers can often get by without taking in all of the visual information. These redundancies also mean that readers must often supplement the visual information that they do take in with knowledge about the language and about the world.

Keywords: writing systems, script, alphabet, syllabary, logography, semasiography, glottography, underrepresentation, conservatism, graphotactics

The goal of this chapter is to examine the characteristics of writing systems that are in use today and to consider the implications of these characteristics for how people read. As we will see, a broad understanding of writing systems and how they work can place some important constraints on our conceptualization of the nature of the reading process. It can also constrain our theories about how children learn to read and about how they should be taught to do so.

Figure 1 shows examples of some of the writing systems that are used by modern societies. Each sentence expresses the sentiment 'I can eat glass' (Mollick, 1996). The visual similarity of Hebrew and Yiddish reflects the fact that those languages use the same *script*, or set of signs. Their similarity in appearance, or *outer form*, disguises the fact that the two languages are very different from each other and that their writing systems work differently in some respects. For example, in Hebrew, as we discuss later, many vowels are left unwritten, whereas that is not generally the case in Yiddish. Conversely, differences in outer form can disguise important similarities in how writing systems function, or their *inner structure* (see Gelb, 1952, for discussion of outer form and inner structure). As we will see, all systems of writing, even those written with different scripts, such as Classical Mongolian, Japanese, Hebrew, and Hindi, share some important properties, both properties related to their outer form and properties related to their inner structure. We discuss these commonalities in the first section of the chapter, focusing on those of most potential relevance for reading. That section is followed by an examination of some of the major differences across writing systems of the world. The final section of the chapter lays out several implications of the characteristics of writing systems for how people read and for the learning and teaching of reading.

Shared Properties Across Writing Systems

In this section, we discuss some of the more important properties that are shared by modern writing systems. In order to keep the discussion uncluttered, we will occasionally skip over counterexamples from special-purpose writing systems, such as the tactile systems used by blind people, or from writing systems that are no longer in use. Our focus is on properties that are common across current writing systems that are used for purposes of general literacy.

Writing Is Visual

Writing normally takes the form of visible marks on the surface of a relatively permanent object. The fact that writing is designed to be taken in by the eye leads to some characteristics that are essentially universal in modern writing. We discuss these properties of writing's outer form in what follows.

Visual processing, whether of writing, spoons, or coins, requires a certain amount of time and attention. Perceivers must determine the number of visual objects that are present, the location of the objects relative to one another, and the identity of the objects. They must often act before they have had time to perform a full analysis of the incoming information. For example, readers may not take in all of the visual information that is available to them, perhaps not fixating on a word at all or not processing all of the letters in a word on which they do fixate. Readers may make decisions about which words are present before they have fully processed the information they have taken in. Some universal tendencies of writing systems have developed to make reading possible despite people's limitations in vision and attention, to allow them to do a reasonable job even when going quickly.

One property of modern writing systems that aids visual processing is that there is a reasonable degree of contrast among the basic elements of a script. The distinction between the closed curve shape of ‹O› and the open curve shape of ‹C› is easy enough to see, but one would not expect a writing system to include several different C-like shapes that differ only by minuscule differences in the size of the opening. Nor would one expect to find a writing system in which a version of ‹C› with a narrow line and a version of ‹C› with a broader line constituted two different letters. To use a term traditional among historians of writing (Evans, 1894), the symbols of modern writing systems are *linear*. They require no shading, no fill, no color other than that needed to distinguish the writing from the background, and no distinctions between lighter and darker lines or wider and narrower lines.

Another visual property of writing that helps people to distinguish its elements is that most scripts contain few or no elements that differ only in their left–right orientation. English has the mirror-image pairs ‹p› and ‹q›, and Korean has ㅏ and ㅓ, but most writing systems do not include any pairs of this sort. This is probably because people find it hard to learn to assign objects that differ only in orientation to different categories (Kolinsky et al., 2011).

Yet another visual property of writing that helps both beginning and skilled readers is the visual redundancy within the elements of a script. In many cases, this redundancy can allow an element to be successfully identified even if some of its visual characteristics are overlooked. For example, a reader of English who fails to detect the crossbar on ‹A› can still identify the letter because there is no other letter ‹Λ› in the Latin alphabet. According to one estimate, in fact, the identity of the elements of modern writing systems can be determined when, on average, half of the strokes are removed (Changizi & Shimojo, 2005).

Within a writing system, the elements show a certain *stylistic consistency*. For example, Chinese characters, such as 圆 ‘round’ and 球 ‘ball’, are squarish and angular; they do not include full circles or semicircles as do a number of the letters of the Latin alphabet. The similarities among the elements of a script reflect the importance of the aesthetic qualities of writing, above and beyond the message that it conveys. The set of forms in (1a) is more pleasing to the eye than the set of forms in (1b) for example, (a) showing a set of Hebrew letters, which are stylistically similar to each other, and (b) showing Hebrew letters interspersed with Arabic letters. As we discuss later in the chapter, stylistic consistency can benefit readers, allowing them to develop a familiarity with the visual patterns that are shared across the elements of their writing system.

(1) a. לאכול
b. כסאפל

The aesthetic nature of writing means that writers may choose certain fonts or visual styles because of their beauty or the qualities they express, and readers must deal with these variations. For example, **Aberdeen Court** may appear on a street sign not because this font is easy for modern readers to read, but to convey adherence to tradition. The aesthetic benefits come with a cost: Readers must learn to deal with different forms of symbols, placing ‹𝔄› in the same category as ‹A›. For a child, it may be by no means obvious that these forms belong together.

Writing Represents Language

Communicating by visual marks on surfaces can work in any number of ways. For example, people can send a message by drawing or painting ad hoc images. This technique could, in principle, permit communication among people who do not share a language. However, communicating by drawing pictures is difficult and error-prone. Communication can be improved tremendously by using *symbols*: images that have conventionally assigned specific meanings for a specific group of users. Figure 2 shows some symbols that were developed in the 1970s to clearly and unambiguously direct travelers in the United States to essential services (AIGA, 1989). It is equally easy for speakers of any language to learn that these symbols direct travelers to departing flights, arrivals, and coffee shops, respectively. But the disadvantage of having a separate symbol for everything of interest is that life has an unbounded number of things of interest. Symbol sets such as those in Figure 2 almost never grow to more than a few hundred. This is because the effort in learning large numbers of symbols does not pay off in expressive power.

One improvement to a set of symbols is to add a *grammar*, a set of rules for expressing relations between symbols. These rules allow users to express concepts that vastly outnumber the individual symbols. The grammar of musical notation, for example, allows musicians to represent an indefinite number of musical pieces with great accuracy using a fairly small set of musical symbols. An even more familiar system is mathematical notation. The positional system, where every digit weights a different power of ten, lets people precisely indicate any natural number using no more than ten symbols, the digits. The two symbols ‹98›, for example, represent the number obtained by adding nine tens to eight because part of the grammar says that the second position from the right is the tens place; ‹89› represents a quite different number. Communication systems such as these, which include a grammar for combining symbols, are referred to by a special term in many languages, such as *writing* in English. People *draw* symbols of departing airplanes, but they *write* sheet music and mathematical formulas.

Yet even powerful systems such as those for music and math are limited to particular domains. Many philosophers have labored hard to develop grammars and symbol sets that could visually represent ideas regardless of their domain—numbers, dining preferences, political humor—but none have caught on. Such systems are difficult to learn, and they are still incapable of representing with any precision everything that people are accustomed to talk about. Writing in the sense that we discuss it in this chapter is an ingenious solution to the problem of generality. Instead of trying to directly represent the concepts that people can talk about, it represents the words of language itself. A writing system that can represent words accurately can represent anything people can talk about. It is truly a general-purpose system of communication.

Systems such as musical and mathematical notation have been characterized as *semasiographic*, literally, ‘writing ideas’ (Gelb, 1952, p. 11). General-purpose writing, in contrast, has been called *glottographic*, literally, ‘language writing’ (Pulgram, 1976, p. 4). The distinction between the two types of systems can be appreciated by comparing the semasiographic ‹98› with glottographic forms like the English ‹ninety-eight› or the French ‹quatre-vingt-dix-huit›. The former follows the grammar of mathematical notation and works the same way around the world. In particular, it can be read off in innumerable different ways, depending on the language one speaks. The English and French spellings, in contrast, represent words in those languages, with specific pronunciations that would be meaningless to people who have not learned those languages. Although writing systems have many differences in their outer form, as we saw in Figure 1, all full-scale systems in regular use are glottographic.

Semasiography is sometimes mixed in, as when we write about "101 Dalmatians," but glottography is at the core.

If writing represented ideas directly, we might expect it to look like some sort of semantic network with nodes spread out all over the page, lines connecting the nodes, and so forth. But writing represents the words of a language, and the most natural way to do so is by arranging along a line a sequence of symbols representing units of language. This is what all modern writing systems do. We call the symbols that are laid out along a line the *characters*. Thus, ता is a single character in Hindi (representing the syllable [tɑ]) even though, as we discuss later, it includes one part (त) that stands for [t] and another part (the vertical bar on the right) that stands for [ɑ].

Stretches of speech can be indefinitely long, and writing media are decidedly finite. At some periods in the past, it was conventional when reaching the edge of a page to begin the next line directly underneath the end of that line. Writing would then proceed in the opposite direction, usually with all the letters reversed, so as to preserve to a reasonable extent the sequentiality between the two lines. In all modern writing systems the convention is now to write in straight lines until one runs out of space, then go back and begin another straight line in the same direction. This use of parallel straight lines is probably the best compromise between the desire to write words in their natural sequence and the desire to always present word symbols in the same orientation.

Although all general-purpose writing systems use symbols to represent language, language is complicated and open-ended. This leaves designers of writing systems many alternatives for mapping symbols onto language. As Martinet (1960) put it, language has a *double articulation*. Martinet's first articulation deals with words; to give it a more memorable name, we will call it the lexical level. At the lexical level, we can think of a sentence as being a sequence of words. Words, in turn, are sometimes composed of smaller meaningful units called morphemes. Thus *Kim's girlfriend uses a prepaid card* breaks down into the sequence shown in (2a); we use square brackets to enclose each word, some of which contain comma-separated morphemes. Martinet's second articulation, the phonetic level, deals with sounds. At this level, a sentence can be thought of as a sequence of syllables, which are in turn composed of phonemes; in (2b) we use brackets to enclose each syllable. Phonemes, in turn can be described in terms of their *distinctive features*, the differences that distinguish them from each other in a given language. For example, /k/ is a voiceless velar stop, distinct from the voiced velar stop, /g/. Each of Martinet's levels has its own grammar—the rules for combining words to make a clause have nothing in common with the rules for combining phonemes into syllables.

(2) a. [Kim, 's] [girl, friend] [use, es] [a] [pre, pay, ed] [card]
b. [k ɪ m z] [g ɚ l] [f r ɛ n d] [j u] [z ɪ] [z ə] [p ɹ i] [p e d] [k ɑ ɹ d]

The dual articulation of language pulls designers of writing systems in two directions. They could invent symbols at the lexical level, making a distinctive symbol for each word or morpheme. Such symbols are called *logograms*. Another choice is to focus on the phonetic level, making a distinctive symbol for each phoneme or each syllable. Such symbols are called *phonograms*. But even the most phonetic of modern writing systems do something special to stabilize word symbols and make them salient.

One way phonographic writing systems highlight the lexical level is through *lexical constancy*. In modern writing systems, there is typically only one conventional representation of

a word: A horse is always a ‹horse› in English. A related property is *lexical distinctiveness*: ‹horse› represents the word *horse* and no other word. When lexical constancy and distinctiveness are present, it is easy for readers to directly map from the written form to the corresponding word. These properties attract particular attention in writing systems in which a given phoneme is not always spelled with the same letter. No basic phonetic spelling rule of English precludes spelling *horse* as ‹horce› (as in ‹force›) as well as ‹horse›, but lexical constancy has encouraged English spellers to stick with a single spelling. Inconsistent mappings also permit lexical distinctiveness among homophones: *horse* is always distinguished from *hoarse*, even though the two are pronounced alike. There is some question as to whether distinguishing homophones is a primary principle of writing systems—there are, after all, exceptions like *sow* (either /so/ ‘plant seeds’ or /saʊ/ ‘female swine’)—or whether it is a side effect of the conservatism of writing that we discuss later in this chapter. Spellings often stay the same even when pronunciations change, and the different vowel letters in ‹horse› and ‹hoarse› reflect the fact that the vowels used to be pronounced differently. But regardless of the origin of lexical distinctiveness, the reader is presented with an additional cue to identifying which homophone the writer intended.

Another way most modern phonographic writing systems highlight the lexical level is through *lexical demarcation*: explicit, visual means of showing where one word ends and the next one begins. Nowadays, almost all languages separate words physically from each other, with spacing or other marks of punctuation, such as the colon-like symbol traditionally used in Amharic. Scripts sometimes demarcate words by other visual methods, as when Hindi writes most words with a solid bar at the top which connects with that of the adjacent elements in the same word, but, in modern times, not with the neighboring words (Figure 1, example 5). The bar reinforces the impression that the word is a unit. Similarly, Mongolian script (Figure 1, example 1) as well as Arabic joins letters within a word but not between words, a tactic used in Europe in many cursive handwriting styles but rarely in print. More subtly, Japanese does not space between words, but most content words start with a kanji character and end with a hiragana character, two types of character that look very different from each other. Thus, the transition between a simple, curvy hiragana character and a more complex, angular kanji character (as between を and 食 in Figure 1, example 2) usually is a visual marker of a lexical boundary. Not all writing systems have all three properties of lexical constancy, lexical distinctiveness, and lexical demarcation, but most have at least one.

Because writing represents language, its outer form is constrained both by the properties of speech in general and the properties of the specific language that it represents. One property of speech is that it is not highly repetitive. Repetition rarely goes beyond doubling. At any level, whether going by individual phonemes, syllables, or words, it is rare to see three instances of the same unit in a row. Because writing represents speech, this degree of repetition is uncommon in writing as well. We might accept that ‹Lorem ipsum dolor sit amet› means something in some language, but we would be very surprised to learn that there was a language in which ‹Looooooremmmm iiiiipppsoooo› was a normal way of writing anything. In contrast, in mathematical notation ‹1,000,000› is a well-formed and unsurprising number.

Another general property of language that is reflected in the outer form of writing is its redundancy. Not all of the sounds in a word or all of the words in a message need to be processed in order to make sense of the message. Because writing represents language, it has some of the same redundancy. For example, standard French indicates negation by placing *ne* before the verb and *pas* after it. A reader who fails to perceive the word *ne* in *Je ne sais pas* ‘I don’t know’ can still determine that the sentence is negative because of the *pas*. Or, because

there is a word *mariposa* ‘butterfly’ in Spanish but not a word *maruposa*, a reader can identify the word without resolving whether the second vowel letter is ‹i› or ‹u›.

Writing also reflects many of the properties of the specific language that it represents. For example, it would be rare to see ‹the› at the end of an English sentence or を at the beginning of a Japanese sentence; this Japanese symbol in reality appears only immediately after the direct object of a verb. Likewise, readers of English would notice something peculiar about a word beginning with ‹ng›, ending with ‹pv›, or totally lacking a vowel letter. Such words are odd because the sounds they would represent violate the *phonotactics* of English, namely, the constraints on the ordering and position of phonemes in words. English words cannot begin with [ŋ], end with [pv], or lack a vowel sound. There should be no reason to come up with the spellings that represent these sounds, unless one is designing an experiment to tap people’s processing of illegal sequences. Although the restriction against words like ‹ngim›, ‹lupv›, and ‹scvnkls› is motivated by the phonotactics of English, nothing would stop the beginning reader from learning about such restrictions as *graphotactic* irregularities—namely, those that violate the normal patterns by which letters are assembled.

At the same time, other graphotactic violations have little or nothing to do with the phonotactics of a language. For example, ‹hevvi mettl› could arguably be used as an improved spelling of *heavy metal* [ˈhɛvi ˈmɛtl̩] in English were it not for the fact that English spelling generally avoids ‹vv›, word-final ‹i›, and word-final ‹l› after a consonant. As this example shows, writing takes on a life of its own. Although readers normally see through the visual forms to the language they represent, writing is a visual system with its own patterns.

Writing Does Not Represent All Aspects of Language

Writing represents a language, but it does not represent all aspects of the language. With rare exceptions, writing systems only represent distinctions that result in *lexical contrast*. For example, Spanish has the sound [d] (as in the English *<u>d</u>en*) and the sound [ð] (as in the English *<u>th</u>en*), but the choice between them is determined entirely by factors such as whether they come after a vowel. There can never be a situation where replacing one sound with the other could result in a new word. Thus, [d] and [ð] are *allophones* of the same phoneme in Spanish. Writing systems rarely have different symbols for different allophones of the same phoneme, and indeed Spanish [deðo] ‘finger’ is spelled ‹dedo›. The failure to distinguish allophones never seems to even be noticed by native speakers of a language, and it probably benefits them by reducing the number of symbols they must learn and distinguish when reading.

More surprisingly, it is not at all uncommon for writing systems to fail to represent distinctions that are lexically contrastive. Sometimes this *underrepresentation* is due to a change in a language’s phonology that makes allophones become phonemes. For example, English spelling fails to differentiate the phoneme /θ/ (as in *<u>th</u>umb*) from /ð/ (as in *<u>th</u>en*), in part because the two used to be allophones. In modern English they have become separate phonemes, but apparently there has not been sufficient confusion to prod people into inventing a new spelling for one or both of the sounds. In Hebrew, the letter פ once stood for the phoneme that had the allophones [p] and [f]. In modern times, these are now two separate phonemes, /p/ and /f/, but both are still written with the same letter. The fact that the same spelling can represent two different phonemes makes reading somewhat more complicated, but knowledge about words (e.g., /θɛn/ isn’t an English word but /ðɛn/ is) and the use of discourse context help readers. Generally speaking, people are conservative and fairly tolerant of inconsistencies in a writing

system, resisting spelling reforms that would break with tradition and require them to learn new spelling rules.

A related historical factor has to do with what writing system another writing system was borrowed from. People who devise a writing system by adapting that of another culture are often reluctant to change it greatly, even if ambiguities and inconsistencies arise. A great many modern writing systems use the Latin script, which, in its original and most widely disseminated form, has no basic way of representing more than five different vowel qualities. There is no direct way of representing vowel length, stress, tone, or virtually any sound that wasn't found in Latin. If the Latin script had had ways of representing these things, possibly many more languages today would as well. Instead, languages that use the Latin script, such as English, often let phonemic distinctions go unexpressed—‹wind› is either "moving air" with a short vowel or "turn coils" with a long vowel—or rely on often indirect ways of distinguishing them, such as consonant doubling—‹planing› /e/ versus ‹planning› /æ/—or silent ‹e›, as in ‹bite› /aɪ/, as opposed to ‹bit› /ɪ/. *Digraphs* (two-letter sequences that represent a single phoneme) are also widely used to supplement the alphabet, often with values that are hard for readers to deduce from their composition. For example, without considering its history, it is not obvious why ‹au› in English *taut* spells a vowel that has little to do with the common values of either ‹a› or ‹u›. Other languages that use the Latin script, such as Czech, have added *diacritics* to certain letters to express phonemic distinctions. Adding a small mark to a letter, such as the diacritic on ‹č›, comes across as a smaller change than creating a wholly new letter.

Another example of the effect of borrowing is afforded by the many other writing systems that use scripts derived from ancient Aramaic, which had no mechanism for writing short vowels. The most widespread of these scripts are Arabic and Hebrew. To this day, Hebrew and most of the languages that use the Arabic script, including Persian and Urdu, do not represent some of their vowels in the texts that are typically read by adults. Figure 3 illustrates how Hebrew writing omits representation of some but not all of its vowels. Hebrew and Arabic do have methods of representing all vowels, but these are used only in special situations, including texts for beginning readers. It is also quite likely that Hindi and related scripts of South Asia descended from Aramaic or one of its precursors. They mostly write vowels as diacritics and they leave out one vowel entirely. This suggests that vowel symbols may be late additions to the script. In Hindi, for example, the vowel /ə/ is not represented: The word सकता /səktɑ/ 'can' in Figure 1, example 5, has overt representations for /s/ स, /k/ क, and /t/ त, and a diacritic for /ɑ/, the vertical bar to the right. However, it does not have anything overt for /ə/. While it is tempting to speculate on why some languages do not write many of their vowels, history has furnished us not with a natural experiment but with little more than a single anecdote: There is a single lineage of scripts that omit vowels. Languages that inherited scripts that descend from Aramaic may simply be retaining its ancestral properties because cultures tend to be fairly conservative about reforming their writing systems.

Certain features of pronunciation appear to be especially susceptible to underrepresentation in writing. All languages have *intonation*, using variations in pitch and timing over the range of a sentence to indicate attitude and emotion, to focus attention on part of a statement, and so forth. Intonation can be an important cue in understanding speech. But no writing system, to our knowledge, represents intonation in all but the most rudimentary and ad hoc way. The modern system of punctuation that has been adapted by most scripts is better than nothing—it can occasionally help the reader discern whether a panda who "eats shoots and

leaves" is a herbivore or carries a gun ("eats, shoots and leaves"; Truss, 2004)—but no language employs it to mark all aspects of intonation. Furthermore, much of the small set of symbols that is used for punctuation marks features other than intonation. For example, ‹?› marks a sentence as interrogative, regardless of its intonation. This widespread inattention to intonation may be due in part to the fact that writing systems focus on words. Representing words in different ways when they appear in different intonational patterns would undermine lexical constancy.

Writing systems are better at representing pitch differences when they help identify words, but they are not much better. Lexically contrastive pitch is called *tone*, and it is a feature of the majority of the world's languages. For example, in Hausa, one of the dominant languages of Africa, /baːǰba˥/ 'father' and /baː˥baǰ/ 'mother' differ only in the tones. 'Father' has a low tone in the first syllable (indicated by the ǰ after the first vowel) and a high tone in the second (indicated by the ˥ after the second vowel). 'Mother' has the opposite pattern. Despite the fact that tone is very important in Hausa and there are established conventions for writing tone with diacritics, tone marks are almost never written in Hausa. The situation with Hausa is not unique: Kutsch Lojenga (2011) reported that speakers of tonal languages throughout Africa have greatly resisted writing tone.

Part of the resistance to writing tones is due to the perception that writing should look like the system on which it is modeled. In Africa this is mostly either English or French, which do not write tones. But tones have several other issues that appear to contribute to the reluctance to write them. Specifically, tones tend to be low in *segmentability*, *stability*, and *proprioception*. Segmentability refers to the inherent difficulty of explicitly factoring out tone as a phonetic feature that is different from the more concrete vowels and consonants over which it is superimposed. Even if a speaker understands that /á/ is a combination of the vowel /a/ and a high tone, the requirement to write separate symbols for the vowel and the tone impresses her as inconvenient clutter. Stability refers to the fact that tones tend to be pronounced differently in different phonetic and grammatical contexts. For example, in many African languages a high tone is pronounced with a very high pitch at the beginning of a sentence but with a much lower pitch toward the end of a sentence. Such things can make it difficult to associate a word with a specific tonal pattern. Proprioception refers to the fact that people can scarcely feel any differences in what their speech organs are doing when they make different tones, depriving them of a useful cue in reading and writing. It is not clear which of these issues is the most important, but underrepresentation of tone must be due to factors beyond adherence to models of writing that use the Latin alphabet. Many independently developed writing systems for tonal languages such as Cherokee, Vai, Japanese, and Chinese do not represent tones at all.

Similar factors appear to contribute to a tendency to underrepresent several other features in writing systems. Lexical stress, such as the distinction between /ɪnˈsɛns/ 'enrage' and /ˈɪnsɛns/ 'substance burned for its fragrance' in English, is not explicitly represented in the spelling: both of these examples are ‹incense›. Contrasting phoneme lengths often go underrepresented as well. For example, in Fijian, words may be distinguished by whether a vowel is pronounced longer or shorter: /mama/ means 'a ring', but /mamaː/, with the second vowel lengthened, means 'chew it', and /maːmaː/, with both vowels lengthened, means 'lightweight'. Although the words differ phonemically, they all are written *mama*. Writers may find it difficult to identify length as a feature distinct from the vowel itself, especially because length is not stable: All vowels sound long when one tries to read or write a word by sounding it out slowly. Graphical considerations may also contribute to the underrepresentation of stress, length, and tone that we see across writing systems. In part for the historical reasons mentioned earlier, these features of speech are

usually represented by small marks that clutter and complicate the written text (Kutsch Lojenga, 2011). Writers may omit the marks for this reason. That makes writing less visually complex than it would otherwise be, but it can also lead to ambiguity that readers must deal with.

Writing Changes, but Lags Behind Language Change

We have already mentioned that writing may change over time, as when diacritics are added to indicate certain previously unexpressed aspects of pronunciation. In this section, we consider how writing changes and why such changes usually occur more slowly in writing than in spoken language.

Table 1 illustrates some changes to the shapes of letters that took place during the evolution of Arabic. The ultimate ancestors of these letters were the early Phoenician letters shown in the second column. As the script developed, it become more and more cursive—that is, simplified and adapted for quick writing—until the Aramaic letters begin to resemble today's Hebrew script, with many of its letters clearly pointing to the left, leading to the next character. By the time the earliest Arabic texts appeared, it had become customary to join many of the letters together, as in the traditional cursive form of such scripts as the Latin of English and French. Further, the letter forms had become so simplified that several of them had become indistinguishable. For example, as Table 1 shows, the forms for /b/ and /t/ in early Arabic were identical. This degree of similarity proved to be too high, so a system of dots to distinguish letters that stood for the same sound was standardized in the seventh century A.D.

As the Arabic example shows, changes to the shapes of symbols are often driven by a need for economy when writing by hand. The symbols become simpler, possible to be produced in fewer strokes and with fewer lifts of the writing implement. Symbols that have a pictorial quality (the Phoenician symbol in the first row in Table 1 was originally meant to depict the head of an ox) tend to lose that quality over time. However, a writing system must maintain a sufficient degree of contrast between symbols so that the human visual system, working quickly, can tell them apart. When contrast breaks down, methods may develop to allow for differentiation, as with the system of dots that was introduced in Arabic.

Although writing changes over time, it does not usually change quickly. This conservatism reflects, in part, the characteristics of its users. People become attached to the status quo, valuing the familiar in the case of writing as in the case of other things. For writing, conservatism is fostered by the fact that revered documents from the past, such as a sacred scripture, could become harder to read if the writing system changed. Spoken language is less conservative than written language, probably because speech fades quickly without special technology. One would have to compare the present against one's memory of the past to notice a change in how people speak. In contrast, the permanence of writing means that one need not rely on memory to notice changes in how people write. These considerations mean that the pace of change is generally faster for spoken language than for written language.

One manifestation of the conservatism of writing is a tendency to retain the spellings of words after they have been borrowed from another language. For example, in English, French, and German, the sound /f/ is spelled *ph* in words and morphemes that were borrowed from Greek, such as *philosophy*, but as *f* in non-Greek morphemes. This etymological rule is very regular and would doubtless come naturally to children who learned to read and write in Greek before learning their native English, French, or German. But most children do not learn Greek first, of course, and so these etymological spellings come across to them as difficult.

Another manifestation of writing's conservative nature is that the grammar of the language that is encoded in print is somewhat different from the grammar that people normally use when they speak. It is more old-fashioned. Finnish is justly famous for having an almost perfectly consistent phonemic writing system, but it also has a wide gap between colloquial and formal styles of speech. The colloquial register is what people acquire as children and use for practically all spoken communication; the latter is what appears in practically all printed material (Karlsson, 2004). For example, schoolchildren in Finland are taught *Olemme talossa* for 'we are in the house'. That can be read off as /olemːe talosːa/ using simple relationships between letters and phonemes. In daily life, however, people would hear and say /me olːaːn talos/. The gap between written and spoken language is even larger for speakers of some other languages. For example, the standard literary form of Arabic is very close to the classical Arabic language spoken over 1,000 years ago and very different from the colloquial forms that are spoken today. Such is the conservatism of writing that historical and cultural factors are the main determinants of why a particular language has the writing system that it does.

Summary

The writing systems that are used by people with normal perceptual abilities employ visual symbols to represent language. In its outer form, writing appears as strings of characters that are arranged along lines. In its inner structure, it concentrates on representing the words of a language. Although writing represents language, it does not represent all aspects of language. Some features tend to be left out. Too much underrepresentation can make reading difficult, but readers can get by with a surprising amount of it because they normally know the language they are reading. As we discuss later in the chapter, readers can often add missing information by using context: what they pick up elsewhere on the page or what they know about the language or about the world.

Differences Among Writing Systems

Writing Represents Words in Different Ways

All general-purpose writing systems use symbols to represent language. It is common to categorize writing systems by how their most fundamental symbols do so (e.g., Sampson, 1985). Writing systems that use symbols that stand for morphemes are called *logographies*. Phonographic systems in which the basic elements represent syllables are *syllabaries*, and phonographic systems in which the basic elements represent phonemes are *alphabets*. However, placing writing systems into mutually exclusive categories such as these invites unproductive quibbling about which level is the most fundamental in a particular writing system and tends to neglect the fact that readers are likely to pick up on multiple levels of representation. In practice, most writing systems are mixed. Some contain logograms and phonograms side by side. Some have symbols for one level of representation that are composed of symbols of another level of representation. In what follows, we consider the choices that are made by several writing systems.

Chinese writing, like all modern writing systems, consists of a linear sequence of characters. Chinese gives priority to the lexical level of Martinet (1960) that was described earlier. Specifically, each Chinese character corresponds to a morpheme, which, by definition, has both a meaning and a pronunciation. In Chinese, those pronunciations are monosyllabic. The

Chinese character 生 represents the specific morpheme that means ‘born’, for example. Because that morpheme has a pronunciation [ʂəŋ˥], then 生 itself represents the syllable [ʂəŋ˥]. Lexical constancy means that the morpheme [ʂəŋ˥] ‘born’ is always represented as 生, and a very strong drive toward lexical distinctiveness in Chinese means this morpheme is represented differently from the morpheme that is pronounced [ʂəŋ˥] and means ‘raise’ 升 and from the morpheme that is pronounced [ʂəŋ˥] and means ‘nephew’ 甥. Lexical demarcation is handled by packing together into one character all the visual elements that contribute to the representation of the morpheme. Characters in Chinese can be identified by the fact that they all occupy the same amount of square space in linear sequence. Thus, in the character 甥 [ʂəŋ˥] ‘nephew’, it is obvious that both the elements 生 [ʂəŋ˥] ‘born’ and 男 [nan˧˥] ‘male’, being squeezed into one square space, contribute to the representation of a single morpheme. Most modern Chinese words have more than one morpheme, and therefore they must be expressed by a sequence of characters. For example, the compound word [nan˧˥ʂəŋ˥] ‘male student’ is written 男生. Indeed, the structure of Chinese is such that all polysyllabic words must be spelled with multiple characters, even if their morphemic structure is not clear to present-day speakers. Chinese writing takes no special notice of words; it does not even put spaces between them. The fundamental level of representation in Chinese writing, at which all lexical demarcation is done, is at the level of the morpheme.

An interesting feature of Chinese writing is that a great number of characters, which by themselves represent a specific morpheme, can also occur as components of other characters, which stand for a different morpheme. In that usage they give hints as to what morpheme the character represents, typically by reference to other morphemes with related meaning or pronunciation. For example, the 男 in 甥 [ʂəŋ˥] ‘nephew’, representing a morpheme that means ‘male’, gives a hint as to the meaning of 甥. The 生, being pronounced [ʂəŋ˥], gives a very good hint as to the pronunciation of 甥. In this character, 生 is used phonetically. Some 81% of Chinese characters contain component symbols that give a hint to how the character is pronounced, although in only about half of those characters does the phonetic component match the pronunciation exactly, even if we ignore tone mismatches (Hsiao & Shillcock, 2006).

Chinese represents a situation where there are many symbols that can stand for syllables but none that stand for smaller units of sound. Symbols standing for syllables also exist in several other writing systems. In the Japanese sentence shown in Figure 4 (repeated here from Figure 1), all symbols except for the first are pronounced as one syllable. Many of the symbols were borrowed from Chinese. Others are members of Japanese syllabaries that can be used to represent syllables in a variety of different words. The katakana syllabary is used, among other purposes, to represent words borrowed from languages that do not use Chinese characters: in this case, [garasu] from English *glass*. The hiragana syllabary is used more generally than katakana to phonetically spell out morphemes for which there is no ready Chinese character. These morphemes include function words and inflectional endings. For example, ます [masu] is an ending expressing politeness. The word [taberaremasu] ‘can eat’ illustrates an intermediate use of hiragana. The character 食 was borrowed from Chinese to represent the root morpheme [tabe-] “eat.” Logically, one might expect that character to be followed here by hiragana that spell out only the inflectional endings, [raremasu]. Instead, it is followed by hiragana that spell out [beraremasu], which includes a spelling of the last part of the root. This apparent redundancy is useful because 食 can also represent other root morphemes in Japanese, such as [kuw-] ‘feed

on’. Adding a phonetic spelling for the last part of the root helps the reader determine which morpheme the writer intended to represent by the symbol 食. This strategy is used throughout Japanese, because the great majority of characters borrowed from Chinese can represent at least two different morphemes with different pronunciations.

In Japanese, the symbols that represent syllables do so atomically; they are not composed of smaller elements that represent the individual phonemes. For example, the symbol ら for /ra/ has nothing in common with the symbol れ for /re/, even though those syllables begin with the same consonant, nor with the symbol ま for /ma/, even though those syllables have the same vowel. This is true in other syllabaries as well. In Cherokee, Ꮻ represents [la], but no part of that symbol is reprised for other syllables that contain [l], such as Ꭼ [le], or for other syllables that contain [a], such as Ꮆ [ma].

In some other scripts where a character represents a syllable, the character is decomposable. In Korean, each syllable consists of elements that symbolize individual phonemes. For example, the symbol for [m] is ㅁ, which can be found at the top left of the character 명 [mjʌŋ] and at the bottom of the character 함 [ham]. When these two characters are sequenced along a line, one gets the word 명함 [mjʌŋham] ‘business card’. Thus if we focus on characters, Korean can be said to represent syllables, like a syllabary. If we focus on the smallest component symbols, however, Korean represents phonemes. In this latter sense, it is an alphabet.

Many of the scripts of South Asia, including Hindi, are similar to Korean in that individual characters often represent syllables but may be composed of smaller symbols that represent phonemes. Such systems are sometimes called alphasyllabaries, syllabic alphabets, or *abugidas*. The Hindi example in Figure 5 reprises the ‘I can eat glass’ sentence from Figure 1, but inserts breaks between characters instead of between words. Even in this short extract, we can see that some of the characters are compositional. The characters for [kɑɲ] and [k] share the shape क, and syllables containing the vowel /ɑ/ have a vertical stroke at the right side of the character.

We talked earlier about the fact that underrepresentation is common across writing systems. It is particularly common in writing systems that represent syllables in a noncompositional way, like Japanese and Cherokee. Often, such systems are far from having different symbols for all of the syllables found in the language. In Cherokee, for example, distinctions of tone and vowel length are unrepresented. The phonemic distinction between plain and aspirated consonants is not noted for most consonants, and the glottal sounds [h] and [ʔ] are not noted at the end of syllables (Montgomery-Anderson, 2008).

Apart from neglecting tone, modern Japanese has eliminated such ambiguities in its syllabaries. But it has syllabic symbols—syllabograms—for only 44 syllables: one for each of five short vowels (e.g., あ for [a]), and 39 for combinations of a single short consonant and a short vowel (e.g., か for [ka]). Instead of using extra syllabograms for the hundreds of remaining syllables in the language, additional symbols are combined with the basic syllabograms to indicate such features as vowel length (かあ for [kaː]), consonant length (っか for [kːa]), syllable-final nasals (かん for [kan]), and consonant clusters in syllable onsets (きゃ for [kja]). From our perspective, the principle behind the design of the Japanese syllabaries is to minimize the number of syllabograms by providing them only for the simplest vowel and consonant–vowel syllables and using those syllabograms as the basis of more complex syllables. Many other writing systems agree with Japanese in handling things such as phoneme length by adding an

element to that for a short phoneme. Another common perspective is to say that Japanese syllabaries are *moraic* by design. A *mora* is the length of a rime that consists of nothing but a short vowel. Rimes that have a long vowel, or include a consonant, have two moras. Moras are important to a phonological analysis of Japanese; for example, haiku have lines of five, seven, and five moras. Because of the importance of moras in Japanese, it is often considered significant that most one-mora syllables, such as [ka], are spelled with one symbol (か), whereas most two-mora syllables are spelled with two symbols, such as かあ for [kaː], かん for [kan], and かっ for [ka] followed by a stop. It is sometimes argued that the symbols spell moras, not syllables. However, syllables with onset consonant clusters break this rule, in that they have more symbols than moras, as in きゃ for the one-mora syllable [kja]. In our view, it is more useful to connect symbol counts with syllable complexity than with mora counts, because the former accounts for the use of multiple symbols in onsets as well as in rimes.

In many writing systems, the individual characters that are arranged along the lines of print represent phonemes. To use a more familiar terminology, the characters are letters. Alphabets based on the Arabic, Cyrillic, Greek, Hebrew, and Latin scripts are familiar examples of systems that rely almost entirely on letters. The relationships between letters and phonemes are often not one-to-one, however. There are several historical sources of such mismatch. Sometimes when existing writing systems are adapted to new languages that have phonemes not provided for in the existing system, the lack is mitigated by assigning the sound to a sequence of two letters, a *digraph*. For example, the English sound [θ] did not exist in the Latin script, so the digraph ‹th› was used. Another reason for letter–phoneme mismatches is the conservatism of writing that was mentioned earlier. When the pronunciations of words change, spellings often do not change to keep up with them. In much of the English-speaking world the pronunciation of [ɹ] in a syllable coda has changed or been lost, resulting in many homophones such as *lord* and *laud*, but an ‹r› remains in the spelling. This conservatism in the spelling means there are more ways to spell several sounds such as [ɔ], which makes spelling harder for people who have this sound change. At the same time, the conservative spelling means that people who have already learned to spell ‹lord› do not have to change their ways, and existing books with such spellings still look fresh. Perhaps most importantly, reluctance to drop the ‹r› from the spellings avoids conflict with dialects in which the [ɹ] still exists in words like *lord*. Conservatism in writing means that spelling has remained more consistent throughout the world than pronunciation has.

Because sounds sometimes change in particular positions of words or when near other specific sounds, the pronunciation of a letter may depend on the position in which it occurs in a word or on the surrounding letters. For example, the vowel of English *fall* was originally the same as that of *fat*, but it moved further back in the mouth under the influence of the following /l/, which is pronounced here with the back of the tongue raised. The spellings did not change. As a result, *fall* now has the same vowel phoneme as *fault* in most dialects of English, but a different vowel spelling. A reader who considered the following letter when deciding on the pronunciation of the ‹a› would have a better chance of being correct than a reader who used a context-free rule that linked the letter ‹a› to the phoneme /æ/. This type of situation is very common in English (Kessler & Treiman, 2001). Purely graphotactic patterns can also make links between letters and phonemes more complex. For example, English spelling generally avoids word-final ‹c› or ‹k› after a single vowel letter, especially in one-syllable words. The spelling ‹ck› is typically used instead, as in *back*. This means that the reader must learn an additional digraph for no phonetic reason.

To some extent, alphabetic writing systems represent other levels of language besides the phoneme. Almost all such writing systems clearly demarcate individual words and spell words in a consistent manner. This lexical constancy helps readers to recognize lexical elements as well as individual phonemes. However, syllables tend not to be separately demarcated in alphabets, nor, in general, are sublexical morphemes such as prefixes and suffixes. But some systems do make an attempt at morphemic constancy, such that a morpheme has the same spelling in different words even if the pronunciation varies somewhat. For example, the first part of the English *cleanliness* /ˈklɛnlinɪs/ is spelled the same way as *clean* /ˈklin/, even though ‹ea› is a rare spelling for /i/. As another example, German spells *Tag* /ˈtɑːk/ ‘day’ with the same ‹g› that appears more phonetically in the plural *Tage* /ˈtɑːgə/. Although many cases of morphemic constancy may be due to the conservative retention of older spellings, they probably help to make morphemes easily identifiable.

Some alphabetic writing systems, often called *shallow*, have primarily one-to-one links between letters and phonemes. Finnish fits this description, at least if readers use the pronunciations of its formal register, as mentioned earlier. The term *deep* is often used to refer to writing systems such as English that have “orthographic inconsistencies and complexities, including multi-letter graphemes, context dependent rules, irregularities, and morphological effects” (Seymour, Aro, & Erskine, 2003, p. 146). However, these different sorts of patterns probably have different effects on readers. Readers may find it easier to use a rare pronunciation of a letter when that allows for morphemic constancy, for example, than when it does not. Indeed, when Chomsky and Halle (1968) introduced the term *deep* in reference to writing, they used it for cases such as *cleanliness* in which spellings reflect morphological considerations. These are cases in which people must go deeper than the surface phonology in order to make sense of the spelling. Chomsky and Halle would not consider multiletter graphemes such as ‹ck› or letter-to-sound rules that depend on context to fall into the same category as the deep spellings to which they refer.

The individual characters that are arrayed along a line of print may represent morphemes, syllables, or phonemes, but there is no writing system in which they represent distinctive features such as velar place of articulation or aspiration. In a few writing systems, including Korean, the shapes of some letters bear a relationship to the features of the sounds that they represent. For example, ㅌ for the aspirated stop [t^h] and ㅋ for the aspirated stop [k^h] add a horizontal line to the letter that represents the corresponding unaspirated stop, ㄷ for [t] and ㄱ for [k]. However, these patterns are not always geometrically or phonetically consistent.

Differences in Outer Form

In discussing differences among writing systems, we have focused so far on differences in their inner structure: how they represent language. In this section we briefly consider differences among writing systems in their outer form.

As we discussed earlier, all scripts lay out their characters sequentially. They differ, though, in the *direction* in which the characters are laid out. Nowadays the great majority of languages are written in horizontal rows from left to right. The popularity of that direction invites explanation. One might speculate that left-to-right writing has won out because most people are right-handed, and ink is less likely to be smudged if the right hand proceeds to the right after writing a word. Vertical writing may be less common because the effective field of vision extends further horizontally than vertically, and because the eye may be more used to tracking

horizontal movement than vertical movement. Upward writing may be least common of all because objects fall more often than they rise. The prevalence of left-to-right writing systems may be misleading, however, because so many writing systems descend from just a couple of ancestors—Classical Greek and Latin (among them English and most other European systems) and Brahmi (most scripts of south Asia)—or were inspired by one of those descendant systems. For example, Chinese and Korean were traditionally written vertically but are predominantly written nowadays from left to right, in order to better integrate with European books and other writing technologies.

To a small extent, the outer forms of scripts vary in how *pictorial* the symbols are. Certain ancient scripts, such as those of Egyptian and Mayan inscriptions, contained a high proportion of symbols that pictured objects. In modern scripts, pictoriality is harder to come by. Korean consonant letters are supposed to be based on images of the vocal tract—ㅁ represents the closed lips, for example—but it is doubtful that people would recognize such pictures without being taught their meaning. Chinese characters are often characterized as pictorial, and indeed a few hundred of them started out as pictures three thousand years ago. Nowadays, however, people who have not been taught what those characters mean cannot usually guess what they represent (Xiao & Treiman, 2012). In general, even scripts that start out with largely pictorial symbols make them highly simplified from the very beginning, and the symbols typically become unidentifiable as they are changed to make writing faster and more compact.

A third way in which scripts vary in their outer form is in the *complexity* of their characters. Among the factors that most commonly contribute to character complexity are the number of contrasts the characters must represent and whether they are composed of multiple elements. In alphabetic writing systems that treat letters as characters, characters are fairly simple—a little less than three strokes on average (Changizi & Shimojo, 2005). Characters are more complex in writing systems such as Hindi and Korean, in which symbols that stand for phonemes combine into characters that stand for syllables. Chinese characters are quite complex. This is in part because Chinese needs a way to make several thousand visually distinct characters and in part because it achieves that by combining multiple components in each character.

As mentioned earlier, each script has a certain stylistic consistency: A character in a particular script tends to look more like other characters in that script than like other signs or images. The specific stylistic features that are shared differ from one script to another. For example, circles are absent from modern Chinese character forms but are present in some letters of the Latin alphabet. Such patterns let people make generalizations and simplifying assumptions when reading and writing. For example, a Chinese writer who imperfectly remembers that a character contains a certain closed shape does not need to consider whether that shape was a circle. A learner of the Hebrew script will apprehend that letters tend to open to the left but not to the right, so if the correct left–right orientation of כ is not immediately recalled per se, it can quickly be deduced. Of course, learning a rule may be injurious when encountering exceptions. When a letter of the Latin alphabet contains a vertical line to which additional segments are attached, the appendage is usually on the right side of the vertical line, as with ‹b› and ‹p›. Children who learn that regularity may have trouble writing the few exceptions, such as ‹d› and ‹q› (Treiman & Kessler, 2011).

Summary

Although all writing systems represent language, they do so in different ways. Systems differ in the emphasis that they place on the lexical level and the phonological level, although many systems include elements of both. Writing systems also differ in their outer form, including the visual properties of their characters and the kinds of lines along which the characters are arranged.

Implications for Reading and for Learning to Read

Having discussed the nature of writing systems, including their similarities and their differences, we focus in this section on the implications for reading. We briefly discuss some general implications for how the reading process must take place, both in skilled readers and in those who are learning to read. We do not have the space to review the empirical evidence for the predictions, but we point to a few studies that have addressed some of the issues.

Reading Involves the Recovery of Linguistic Form

Because writing represents language, successful reading involves recovering the linguistic form that is represented in print. Reading begins with the eyes, because writing is visual, but "successful skilled reading enables the language system to take over from the visual system with astonishing speed" (Perfetti, 2012, p. 299). Indeed, much research has documented the involvement of phonology in silent reading across a variety of writing systems (see Pollatsek, this volume). And it is not just phonology that is recovered from print, but other aspects of language as well. Models of the cognitive processes involved in skilled silent reading must specify how linguistic processing shapes and then takes over from visual processing, and our consideration of the nature of writing systems suggests that models that postulate a close relationship between visual and linguistic processing have more psychological plausibility than models that do not.

Given the nature of writing systems, children cannot get very far in learning to read until they have learned that elements of writing stand for specific elements of language. Thus, ‹cat› stands for *cat* and not *kitty*. Children must learn that reading a word is different from naming an object or a picture, where it could be appropriate to use either *cat* or *kitty* for a young feline. Learning to read is easiest, moreover, if children already know the language that is represented. By six years of age or so, when children in many societies begin learning to read, language development is quite advanced. Children know many words and many aspects of grammar. However, the language that is represented in writing may include certain lexical items and grammatical structures that are not common in colloquial speech and that may be unfamiliar to children. As mentioned earlier, this is especially true in certain cultures, including those that use Arabic. In any language, though, people learn some aspects of language primarily from reading.

Readers Need Not Always Take in All Aspects of Writing's Visual Form

As we have discussed, there is some redundancy in writing by virtue of the redundancy in language itself and by virtue of the redundancy in symbols of writing. This means that readers can sometimes get by with only partial uptake of visual information. Thus, a reader of English could recognize ‹Alligator› even if he missed the crossbar on the ‹A› and even if he did not take in the last few letters. This is because there is no other letter in the Latin alphabet that is the same

as ‹A› except for the crossbar and no other word that differs from ‹Alligator› in just the last few letters. As another example, a reader of English could recognize ‹campus› even if she did not resolve the relative order of ‹m› and ‹p›. Readers learn which aspects of writing's visual form are critical to take in and which are less critical, and this can depend on which language they are reading (see Frost, this volume). In any written language, though, it is rarely necessary to process each and every visual element. In order to write a word correctly, on the other hand, complete information about the identity and the order of the elements is required. This is a major reason why, across writing systems, spelling is more difficult than reading (see Treiman & Kessler, 2014, for a discussion of spelling).

Reading Often Requires Information That is Not on the Page

All writing systems represent language, but no writing system represents all aspects of the language. For example, we have seen that intonation and stress are often not represented. Underrepresentation means that in order to understand a text, readers sometimes must supplement what is on the page with other things that they know about the language or the world. For example, readers of English might decode the spelling ‹give› as [gaɪv], but because they do not recognize that as a word of their language, they will quickly rule it out. They can use their knowledge about the world to suggest that ‹sewer› in *The dirty water drained into the sewer* stands for [ˈsuɚ] 'a conduit for carrying off waste' as opposed to [ˈsoɚ] 'a person who sews'. Sometimes, the information that is needed to disambiguate a word comes only after the word has been read. Consider the sentence *Since the old sewer was awful, the tailor's shop got a bad reputation* (Folk & Morris, 1995). When the reader gets to the part about the tailor's shop, it becomes clear that *sewer* is a person who sews. But before this point it is possible that *sewer* could be a conduit for carrying off waste. Readers have difficulty with sentences such as this, sometimes moving their eyes back to *sewer* when they reach the point where the meaning is disambiguated. Readers' knowledge about things that are not on the page can often reduce or even eliminate the effects of any remaining ambiguities.

Learning to Read Is Challenging

Writing is not a part of the human genetic endowment, as spoken language is; it is an "optional accessory that must be painstakingly bolted on" (Pinker, 1997, p. ix). The ease with which educated modern people read can be misleading: That ease is achieved only through years of practice. Even when a child already knows the language that is represented, learning to read accurately and fluently is hard. Some of the challenges, such as the need to learn that a written word represents its meaning differently than a picture does, are the same for learners of different writing systems. Others are different, such as learning the direction the characters are written in or what specific shapes correspond to what specific sounds or morphemes.

Learning to read and write typically requires a degree of explicit instruction that is not required in order to learn to speak and listen. That instruction should be based on a solid understanding of how the writing system works, but that is not always true. For example, phonics instruction for English does not typically give adequate consideration to the fact that the pronunciation of a letter or digraph can be influenced by the context in which it occurs. It does not typically consider how morphology and graphotactics can help in choosing among alternative pronunciations. Children could benefit from learning about such patterns as that ‹oo› is typically pronounced as /ʊ/ when it occurs before ‹k›, as in *book* and *look*, but as /u/ in other contexts.

They could benefit from learning that the final ‹e› of words like *give* and *have* is motivated by graphotactic considerations (English words do not normally end with ‹v›), and the final ‹e› does not influence the pronunciation of the vowel as it does for *guide* and *hate*.

Knowledge of Written Language Is Used for More Than Reading and Writing

Writing developed as a tool to allow people to freeze language, which fades quickly without special technology. By placing language in a relatively permanent form, writing permits communication among people who are distant in time and space. Once learned, however, writing becomes a tool for other purposes. For example, people use writing to help remember things. Because writing represents language, people use their knowledge about the written forms of words in thinking about the words' spoken forms. Indeed, a number of studies support the idea that literate people's ideas about speech can be influenced, sometimes accurately, sometimes inaccurately, by their knowledge of writing (see Kolinsky, this volume). For example, people often judge that the word *lagoon* /ləˈgun/ contains the syllable /læg/ because it contains the letters ‹lag› (Taft & Hambly, 1985). Sometimes, awareness of the spelling even leads people to change their pronunciation to more closely reflect the written form of a word. An example of a *spelling pronunciation* that has become widespread in English is *often*, which is now often pronounced with a [t] that had been silent for 500 years.

Conclusions

Theories about how people read and about how they learn to read should be based on a good understanding of writing systems and how they work. If our knowledge about writing is too narrow, perhaps limited to our own writing system and others like it, our theories may be too narrow. The teaching of reading may also suffer. In this chapter, we have presented some basic information about writing systems. We have seen that although writing's outer form can vary a good deal from one writing system to another, there are a surprising number of similarities across writing systems in both outer form and inner structure. The properties of writing mean that skilled reading involves the recovery of linguistic forms. Readers do not always have to take in all aspects of the print in order to recover the linguistic form; however, they often need to use information outside the print itself.

References

AIGA. (1989). *Tools and resources*. http://www.aiga.org/symbol-signs

Changizi, M. A., & Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B, 272*, 267–275.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper and Row.

Evans, A. J. (1894). Primitive pictographs and a prae-Phoenician script from Crete and the Peloponnese. *Journal of Hellenic Studies, 14*, 270–372.

Folk, J. R., & Morris, R. K. (1995). Multiple lexical codes in reading: Evidence from eye movements, naming time, and oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1412–1429.

Gelb, I. J. (1952). *A study of writing: The foundations of grammatology*. Chicago, IL: University of Chicago Press.

Hsiao, J. H., and Shillcock, R. (2006). Analysis of a Chinese phonetic compound database: Implications for orthographic processing. *Journal of Psycholinguistic Research, 35*, 405–426.

Karlsson, F. (2004). *Finnische Grammatik* [Finnish grammar]. 4th ed. Transl. of Suomen peruskielioppi. Hamburg, Germany: Buske.

Kessler, B., & Treiman, R. (2001). Relations between sounds and letters in English monosyllables. *Journal of Memory and Language, 44*, 592–617.

Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-Cabral, L., & Morais, J. (2011). Enantiomorphy through the looking glass: Literacy effects on mirror-image discrimination. *Journal of Experimental Psychology: General, 140*, 210–238.

Kutsch Lojenga, C. (2011, January). *Orthography and tone: Tone system typology and its implications for orthography development.* Paper presented at the meeting of the Linguistic Society of America, Pittsburgh, PA.

Martinet, A. (1960). *Éléments de linguistique générale*. Paris, France: Colin.

Mollick, E. (1996). The "I Can Eat Glass" project. http://web.archive.org/web/20040201212958/http://hcs.harvard.edu/~igp/glass.html

Montgomery-Anderson, B. (2008). A Reference Grammar of Oklahoma Cherokee (Doctoral dissertation, University of Kansas). Retrieved from http://kuscholarworks.ku.edu

Perfetti, C. (2012). Thru but not wisht: Language, writing, and universal reading theory. *Behavioral and Brain Sciences, 35*, 299–300.

Pinker, S. (1997). Foreword. In D. McGuinness, *Why our children can't read and what we can do about it* (pp. ix–x). New York, NY: Free Press.

Pulgram, E. (1976). Typologies of writing systems. In E. Pulgram & W. Haas (Eds.), *Writing without letters* (pp. 1–28). Manchester, England: Manchester University Press.

Sampson, G. (1985). *Writing systems: A linguistic introduction*. Stanford, CA: Stanford University Press.

Seymour P. H. K., Aro M., & Erskine J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174.

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language, 24*, 320–335.

Treiman, R., & Kessler, B. (2011). Similarities among the shapes of writing and their effects on learning. *Written Language and Literacy, 14*, 39–57.

Treiman, R., & Kessler, B. (2014). *How children learn to write words*. New York, NY: Oxford University Press.

Truss, L. (2004). *Eats, shoots & leaves: The zero tolerance approach to punctuation*. New York, NY: Penguin.

Xiao, W., & Treiman, R. (2012). Iconicity of simple Chinese characters. *Behavior Research Methods, 44*, 954–960.