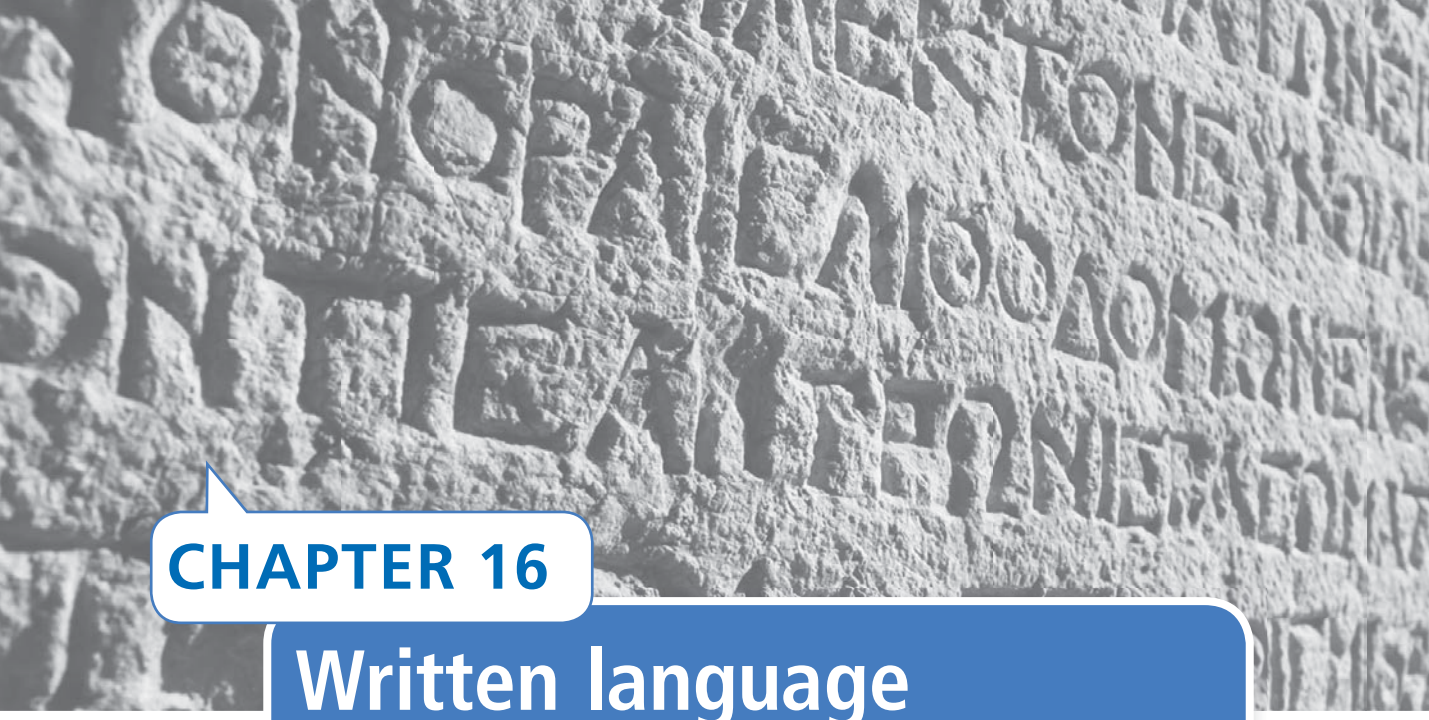


CHAPTER 16

Written language

The five-year-old child whose skills in symbolic representation are displayed here seems to be already familiar with some of the basic elements of writing. The sequence of letters goes from left to right, each letter is distinct from the next and generally well formed, and each word is separated from the next by larger spaces. The occasional spelling "mistake," a traditional problem in written English, cannot disguise the fact that this child has already learned how to write.

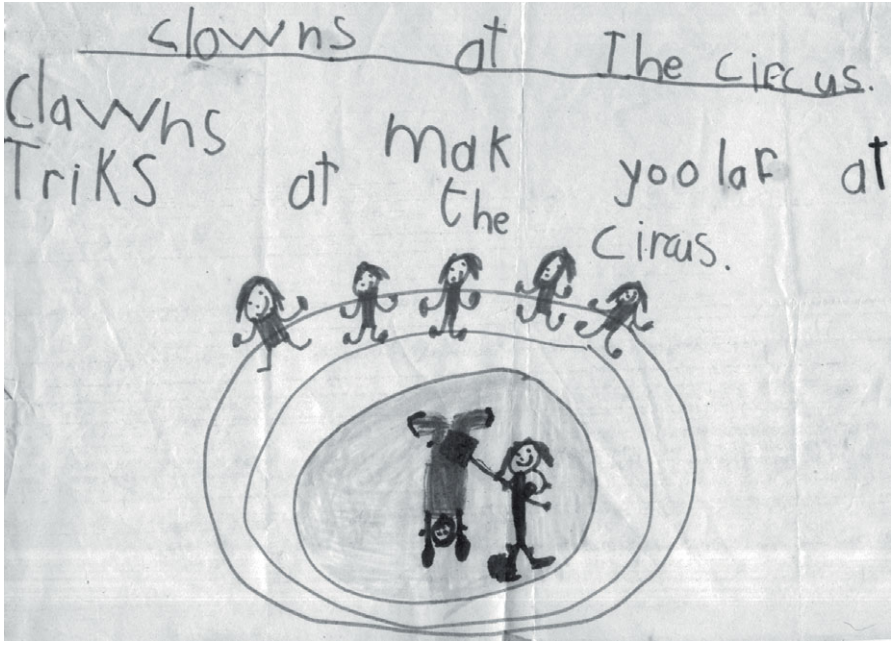


Figure 16.1

Writing

We can define **writing** as the symbolic representation of language through the use of graphic signs. Unlike speech, it is a system that is not simply acquired, but has to be learned through sustained conscious effort. Not all languages have a written form and, even among people whose language has a well-established writing system, there are large numbers of individuals who cannot use the system.

In terms of human development, writing is a relatively recent phenomenon. We may be able to trace human attempts to represent information visually back to cave drawings made at least 20,000 years ago, or to clay tokens from about 10,000 years ago, which appear to have been an early attempt at bookkeeping, but these artifacts are best described as ancient precursors of writing. The earliest writing for which we have clear evidence is known as "cuneiform," marked on clay tablets about 5,000 years ago. About 3,000 years ago, inscriptions were being used in an ancient script that has a more obvious connection to writing systems in use today. We know that we must have lost a lot of earlier inscriptions on perishable material, but working from the surviving inscriptions, we can trace the development of one writing tradition, lasting a few thousand years, with which humans have sought to create a more permanent record of what was going on.

Pictograms

Cave drawings may serve to record some event (e.g. Humans 3, Buffaloes 1), but they are not usually thought of as any type of specifically linguistic message. They are usually treated as part of a tradition of pictorial art. When some of the "pictures" came to represent particular images in a consistent way, we can begin to describe the product as a form of picture-writing, or **pictograms**. Modern pictograms, as in Figure 16.2, are language-independent and can be understood with the same conventional meaning in a lot of different places where a number of different languages are spoken.

Ideograms

Whereas the images presented in Figure 16.2 continue to reflect the physical forms of objects, we typically interpret them in a way that goes beyond simply recognizing

Figure 16.2

those objects. That is, the picture of a cup (with saucer) doesn't just let us know that there is a cup in this location, but also something to put in that cup and maybe something to go with it (neither of which is included in the picture). We actually interpret the images not as objects, but as symbols of the objects, with meanings associated with the symbol that may not be tied to the object.

At some point in the early development of pictorial representation, a symbol such as  came into use for referring to the sun. An essential part of this use of a representative symbol is that everyone should use a similar form to convey a roughly similar meaning. In time, this picture might develop into a more fixed symbolic form, such as  , and come to be used for "heat" and "daytime," as well as for "sun." Note that as the symbol extends from "sun" to "heat," it is moving from something visible to something conceptual (and no longer a simple picture). This type of symbol is then considered to be part of a system of idea-writing, or **ideograms**. The distinction between pictograms and ideograms is essentially a difference in the relationship between the symbol and the entity it represents. The more "picture-like" forms are pictograms and the more abstract derived forms are ideograms. A key property of both pictograms and ideograms is that they do not represent words or sounds in a particular language.

It is generally thought that there were pictographic or ideographic origins for a large number of symbols that turn up in later writing systems. For example, in Egyptian hieroglyphics, the symbol  was used to refer to a house and derived from the diagram-like representation of the floor plan of a house. In Chinese writing, the character  was used for a river, and had its origins in the pictorial representation of a stream flowing between two banks. However, it is important to note that neither the Egyptian nor the Chinese written symbols are actually "pictures" of a house or a river. They are more abstract. When we create symbols in a writing system, there is always an abstraction away from the physical world.

When the relationship between the symbol and the entity or idea becomes sufficiently abstract, we can be more confident that the symbol is probably being used to represent words in a language. In early Egyptian writing, the ideogram for water was  . Much later, the derived symbol  came to be used for the actual word meaning "water." When symbols are used to represent words in a language, they are described as examples of word-writing, or "logograms."

Logograms

An early example of logographic writing is the system used by the Sumerians, living in Mesopotamia, in the southern part of modern Iraq, around 5,000 years ago. Because of the particular shapes used in their symbols, these inscriptions are more generally described as **cuneiform** writing. The term cuneiform (from Latin *cuneus*, "wedge") means "wedge-shaped" and the inscriptions used by the Sumerians were produced by

pressing a wedge-shaped implement into soft clay tablets that created a permanent symbol when the clay hardened, resulting in forms such as: 

The form of this symbol really gives no clue to what type of entity is being referred to. The relationship between the written form and the object it represents has become arbitrary and we have a clear example of word-writing or a **logogram**. The cuneiform symbol above can be compared to a typical pictographic representation of the same fishy entity: 

We can also compare the ideogram for the sun, presented earlier as,  with the logogram used to refer to the same entity found in cuneiform writing:  .

Modern logograms in English are forms such as $, 8, &, where each symbol represents one word, as is also the case with @, now one of the most commonly used logograms (in email addresses). A more elaborate writing system that is based, to a certain extent, on the use of logograms can be found in China. Many Chinese written symbols, or **characters**, are used as representations of the meaning of words, or parts of words, and not of the sounds of spoken language. In some treatments, this type of writing is technically described as "morphographic," because the symbols have come to be used for morphemes rather than words. (See Chapter 6 for the distinction between words and morphemes.) One of the advantages of such a system is that two speakers of very different dialects of Chinese, who might have great difficulty understanding each other's spoken forms, can both read the same written text. Chinese writing, with the longest continuous history of use as a writing system (i.e. 3,000 years), clearly has many other advantages for its users.

One major disadvantage is that quite a large number of different written symbols are required within this type of writing system, although the official "list of modern Chinese characters for everyday use" is limited to 2,500 characters. (Other lists contain up to 50,000 characters.) Remembering large numbers of different composite word-symbols, however, does seem to present a substantial memory load, and the history of most other writing systems illustrates a development away from logographic writing. To accomplish this, some principled method is needed to go from symbols representing words (i.e. a logographic system) to a set of symbols that represent sounds (i.e. a phonographic system).

Phonographic writing

The development from pictographic representation to logographic writing, even among the Sumerians, did not take place without some symbols being used in a similar way if certain words had similar sounds (not meanings). Following Gelb (1963), we can observe this process taking place as the physical shape in the form of an arrow () was first used in representations of the word for "arrow" (*ti*), then later adopted for the more abstract concept "life" (*ti*), simply because both words sounded the same. Similarly, the symbol () for a physical object "reed" (*gi*) is

taken over for the abstract concept "reimbursement" (*gi*) on the basis of similar pronunciations. In this development, the symbols are adopted to represent the sounds of the words, also known as **phonographic writing**.

The rebus principle

This general pattern of using existing symbols to represent the sounds of words in a language is often described in terms of a process known as the **rebus principle**. In this process, the symbol for one entity is taken over as the symbol for the sound of the spoken word (or part of it) used to refer to that entity. That symbol then comes to be used whenever that sound occurs in any words.

We can create an example, working with the sound of the English word *eye*. We can imagine how the pictographic representation  could have developed into the logogram  . This logogram is pronounced as *eye* and, with the rebus principle at work, you could then refer to yourself as  ("I"), to one of your friends as  ("Crosseye"), combine the form with the logogram for "deaf" to produce "defy," with the logogram for "boat" to produce "bow-tie," and so on.

A similar process is taking place in contemporary English texting where the symbol "2" is used, not only as a number, but as the sound of other words or parts of words, in messages such as the one on the left below ("(I) need to speak to you tonight"). In this message, the letter "u" also illustrates the process of rebus writing, having become the symbol for the sound of the spoken word "you," as also illustrated in the other example, with "c" and "8" also representing sounds.

nd2spk2u2nite *cul8r*

Let's take another, non-English, example, in which an ideogram () becomes the logogram  , for the word pronounced *ba* (meaning "boat"). We can then produce a symbol for the word pronounced *baba* (meaning "father") which would be  . One symbol can thus be used in many different ways, with a range of meanings. What this process accomplishes is a sizeable reduction in the number of symbols needed in a writing system.

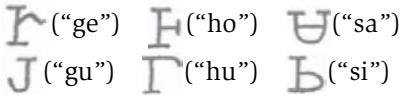
Syllabic writing

In the last example, the symbol that is used for the pronunciation of parts of a word represents a unit (*ba*) that consists of a consonant sound (*b*) and a vowel sound (*a*). This unit is one type of syllable, described in Chapter 4. When a writing system employs a set of symbols, each one representing the pronunciation of a syllable, it is described as **syllabic writing**.

There are no purely syllabic writing systems in use today, but modern Japanese can be written with a set of single symbols representing spoken syllables, called

"hiragana," and is consequently often described as having a (partially) syllabic writing system, or a **syllabary**.

In the early nineteenth century, a Cherokee named Sequoyah, living in North Carolina, invented a syllabic writing system that was widely used within the Cherokee community to create written messages from the spoken language. Papers with writing were described as "talking leaves." In the Cherokee examples below, we can see that the written symbol in each case does not correspond to a single consonant (C) or a single vowel (V), but to a syllable (CV).



Both the ancient Egyptian and the Sumerian writing systems evolved to the point where some of the earlier logographic symbols were being used to represent spoken syllables. However, it is not until the time of the Phoenicians, inhabiting what is modern Lebanon between 3,000 and 4,000 years ago, that we find the full use of a syllabic writing system. Many of the symbols that the Phoenicians used were taken from earlier Egyptian writing. The Egyptian form  (meaning "house") was adopted in a slightly reoriented form as  . After being used logographically for the word pronounced *beth* (still meaning "house"), the symbol came to represent other syllables beginning with a *b* sound. Similarly, the Egyptian form  (meaning "water") turns up as  and is used for syllables beginning with an *m* sound.

So, a word that might be pronounced as *muba* could be written as  , and the pronunciation *bima* could be written as  . Note that the direction of writing is from right to left, as it still is in the writing system of languages such as Arabic. By about 3,000 years ago, the Phoenicians had stopped using logograms and had a fully developed syllabic writing system.

Alphabetic writing

If you have a set of symbols being used to represent syllables beginning with, for example, a *b* sound or an *m* sound, then you are actually very close to a situation in which the symbols can be used to represent single sound types in a language. This is, in effect, the basis of alphabetic writing. In principle, an **alphabet** is a set of written symbols, each one representing a single type of sound or phoneme. The situation just described is what occurred in the development of the writing systems of Semitic languages such as Arabic and Hebrew. Words written in these languages, in everyday use, largely consist of symbols for the consonant sounds in the word, with the appropriate vowel sounds being supplied by the reader (or rdr).

This type of writing system is sometimes called a **consonantal alphabet**. The early version of Semitic alphabetic script, originating in the writing system of the Phoenicians, is the basic source of most other alphabets to be found in the world.

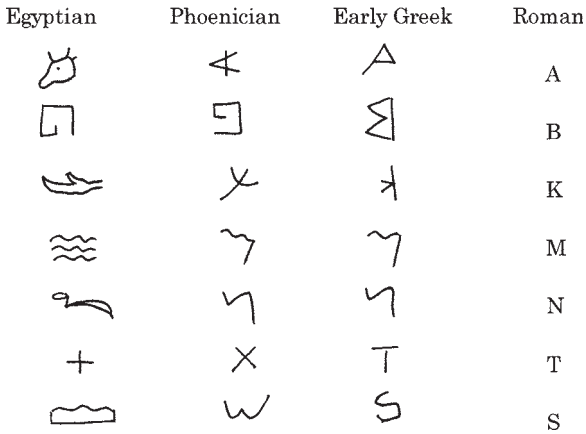


Figure 16.3

Modified versions can be traced to the East into Iranian, Indian and South-East Asian writing systems and to the West through Greek. The basic order of letter symbols in the first "A-B-C-D. . ." was created about three thousand years ago by the Phoenicians and continues to be used as our primary ordering device for lists in everything from dictionaries to telephone directories to grades for academic performance.

The early Greeks took the alphabetizing process a stage further by using separate symbols to represent the vowel sounds as distinct entities, and so created a remodeled system with vowel symbols. This change resulted in the Phoenician consonant "alep" becoming a symbol for a vowel sound as *A* ("alpha") to go with existing symbols for consonant sounds such as *B* ("beta"), giving us single-sound writing or an "alphabet." In fact, for some writers, it is the Greeks who should be given credit for taking the inherently syllabic system from the Phoenicians and creating a writing system with the single-symbol to single-sound correspondence fully realized.

From the Greeks, this revised alphabet passed to the rest of Western Europe through the Romans. As a result, we talk about the Roman alphabet as the writing system used for English. Another line of development took the same basic Greek writing system into Eastern Europe where Slavic languages were spoken. The modified version, called the Cyrillic alphabet (after St. Cyril, a ninth century Christian missionary), is the basis of the writing system used in Russia today.

The actual form of a number of letters in modern European alphabets can be traced from their origins in Egyptian hieroglyphics. The examples in Figure 16.3 are based on Davies (1987).

Written English

If indeed the origins of the alphabetic writing system were based on a correspondence between a single symbol and a single sound type, then one might reasonably ask why

there is such a frequent mismatch between the forms of written English ("you know") and the sounds of spoken English ("yu no" or /ju noʊ/). Other languages (Italian, Spanish) have writing systems that hold much more closely to the one-sound-one-symbol principle of alphabetic writing. English is not always so consistent. As we noted in Chapter 3, there is a lot of variation in how each sound of contemporary spoken English is represented in writing. The vowel sound represented by /i/ is written in various ways, as shown in the first two columns on the left below, and the consonant sound represented by /ʃ/ has various spellings, as in the other two columns.

i (*critique*)	ee (*queen*)	s (*sugar*)	ch (*champagne*)
ie (*belief*)	eo (*people*)	ss (*tissue*)	ce (*ocean*)
ei (*receipt*)	ey (*key*)	ssi (*mission*)	ci (*delicious*)
ea (*meat*)	e (*scene*)	sh (*Danish*)	ti (*nation*)

English orthography

As we have just seen, English **orthography** (or spelling) is subject to a lot of variation. Notice how often a single phoneme in the two lists is actually represented by more than one letter. Part of the reason for this is that the English language is full of words borrowed, often with their spelling, from other languages, as in *ph* for /f/ in the Greek borrowings *alphabet* and *orthography*, where two letters are used for a single sound. A combination of two letters consistently used for a single sound, as in *ph* /f/ and *sh* /ʃ/ is called a **digraph**.

The English writing system is alphabetic in a very loose sense. Some reasons for this irregular correspondence between sound and symbolic representation may be found in a number of historical influences on the form of written English. The spelling of written English was largely fixed in the form that was used when printing was introduced into fifteenth-century England. At that time, there were a number of conventions regarding the written representation of words that had been derived from forms used in writing other languages, notably Latin and French. For example, *qu* replaced older English *cw* in words like *queen*. Moreover, many of the early printers were native Flemish speakers and could not make consistently accurate decisions about English pronunciations, hence the change from Old English *gast* to *ghost*, with *h* derived from the Flemish version (*gheest*).

Perhaps more important is the fact that, since the fifteenth century, the pronunciation of spoken English has undergone substantial changes. For example, although we no longer pronounce the initial *k* sound or the internal *gh* sound, we still include letters indicating the older pronunciation in our contemporary spelling of the word *knight*. These are sometimes called "silent letters." They also violate the

one-sound-one-symbol principle of pure alphabetic writing, but not with as much effect as the silent final *-e* of so many English words. Not only do we have to learn that this letter is not pronounced, we also have to know the patterns of influence it has on the preceding vowel, as in the different pronunciations of *a* in the pair *hat/hate* and *o* in *not/note*.

If we then add in the fact that a large number of older written English words were actually "recreated" by sixteenth-century spelling reformers to bring their written forms more into line with what were supposed, sometimes erroneously, to be their Latin origins (e.g. *dette* became *debt*, *doute* became *doubt*, *iland* became *island*), then the sources of the mismatch between written and spoken forms begin to become clear. In Chapter 17, we will look more closely at other aspects of the historical development of English and other ways in which languages change.