

Writing systems

Pictures and written symbols

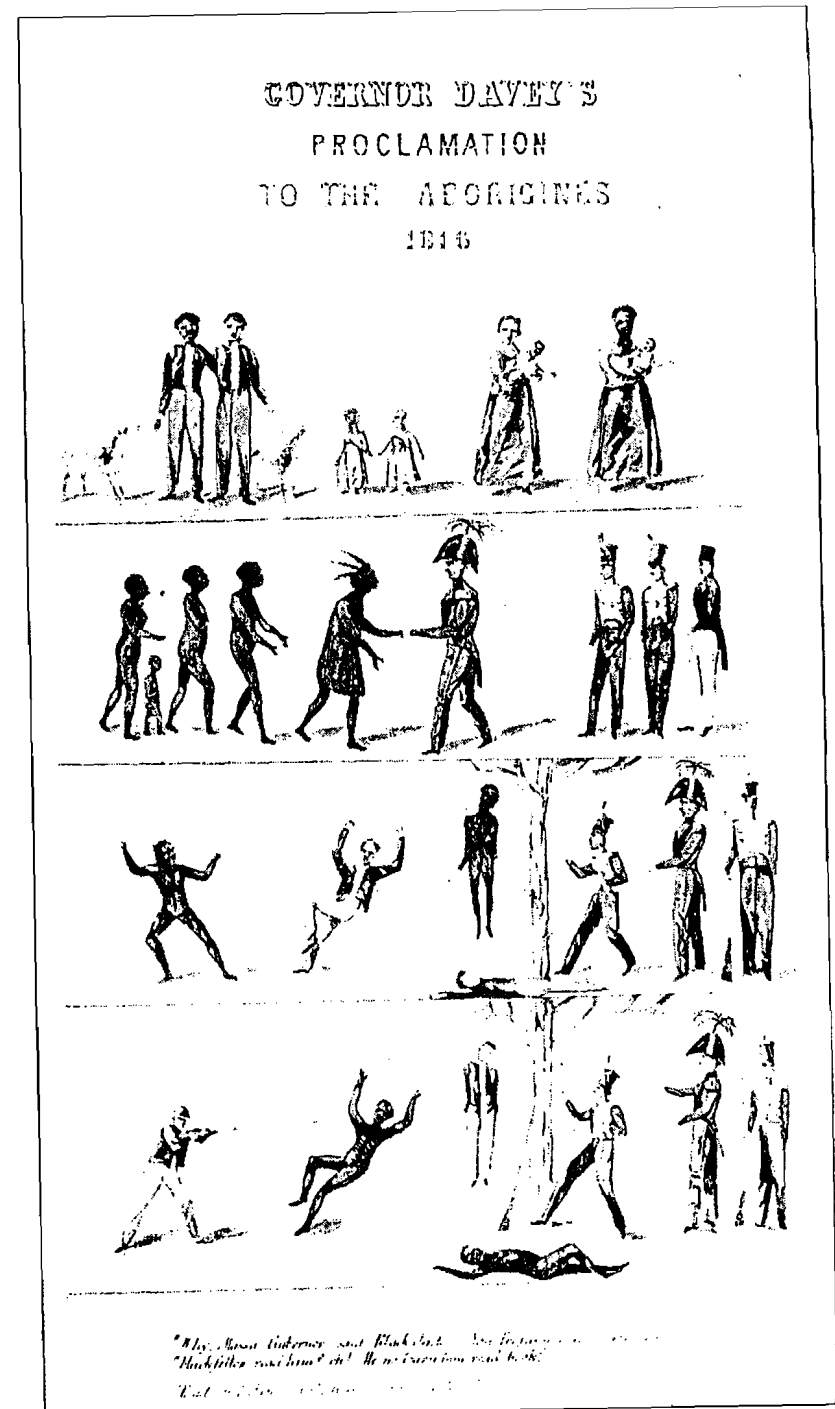
The key to language, then, is grammar, the level of 'words-in-structure'; since that is where the meanings are organised, processed, and packaged in a form that can be turned into an expression of some kind. (That is looking at it from the point of view of the speaker, the one who is doing the encoding. It is equally valid, of course, to look at it from the other end, from the point of view of the listener, the one who is doing the decoding. In his case, the grammar takes in the expressions, unpacks them, and sorts them out into the different semantic components.)

For about 99.5 per cent of the history of the human race, the only medium of expression for language has been sound, the sound produced by the organs of speech (from the larynx to the lips and nostrils) in modifying the stream of air that comes from the diaphragm. The 'sender' of the message has been a speaker, and the 'receiver' has been a listener.

Not that this was the only form of human communication. From at least a thousand generations back, our ancestors have been able to draw, and have made pictures on rock faces and the walls of caves (as well as, no doubt, on much less durable material that has not survived). Whatever the specific significance and social value of such artefacts at any particular place and time—whether adornment, or boasting of one's exploits, or marking a sacred site—they are bearers of meaning. Our ancestors long ago learnt to recognise and exploit the semiotic potential of the visual medium also.

But this is not language; and the distinction is an important one. Painting a picture may be—perhaps always is—a form of communication, a symbolic act directed at other people. It may have a specific communicative purpose, such as recording past events or giving instructions on where and what to hunt. But this does not mean it is a form of language. If we use the word 'language' to refer to such activity, we are using it metaphorically, just as when we talk of music or mathematics as a kind of language. (There is no harm in this, obviously, provided we recognise the fact, and provided that we then have a clear way of indicating when we are talking about language in the primary sense of the term.)

Figure 2.1 Pictorial communication that is not writing



Let us use the term 'writing' in its exact sense, to mean a system of visual representation that is language. Such pictures, then, like those in Figure 2.1, may be a form of communication, but they are not a form of writing.

Children also learn to draw, usually some time after they have learnt to speak; and they then have to learn the distinction between drawing and writing. Here again there is a discontinuity: they have to 'leap' from one to the other, and the two are kept strictly apart. In the history of the human race, on the other hand, the line was not so clear. Drawing evolved gradually until it became writing.

From picture to character

Why do we say that a picture, even if it 'contains a message' or 'communicates something', is not writing?

First of all let us point out that the qualification 'even if it contains a message, or communicates something' really adds nothing at all, since any pictorial representation can be said to communicate something. Indeed, we are brought up to expect that it should do, as is shown by the common complaints of the picture-gazer: 'It doesn't mean anything to me'; compare also the language of art criticism, which makes frequent reference to what a picture 'conveys', its 'theme', 'symbolic significance', and so on.

But the question of whether something is writing or not can be answered in quite explicit terms. Writing is a part of language. More specifically, it is one kind of expression in language—an alternative to sound. We have said that a language consists of three strata: meaning, wording, and sound. We can now modify this, and say that a language consists of meaning, wording, and expression; and the expression may take the form either of sound or of writing.

One thing that follows from this is that writing can always be read aloud. If we look at a painting, or any other visual art form, we can describe it, make a commentary on it, 'say what it means'; but we cannot read it. We cannot decode it into wordings—because it is not an encoding of wordings in the first place. We could not list the elements of which it is made up, put them in a dictionary, and indicate how to pronounce them. They are not elements of a language.

The fact that we can make a clear distinction between what is writing and what is not does not mean there are no 'borderline cases'. There can always be instances that are mixed or indeterminate, however clearly defined the categories are in theory; and in the history of writing there must have been many, although none seems to have survived—which suggests that the transition from 'pre-writing' to writing may also have been fairly sudden.

But although we cannot document the process whereby writing first evolved, it is reasonably clear how it happened. Writing did not begin by somebody deciding to write language down instead of saying it aloud. It evolved from the coming together of two independent semiotic systems: language, on the one hand, and visual imagery on the other. Writing begins when pictures are interpreted as language.

Consider the shape shown in Figure 2.2. This is a picture incised on a bone, for purposes of divination, in China some time in the second millennium BC. It is a picture of a horse.

Figure 2.2 Earliest known form of Chinese character *horse*

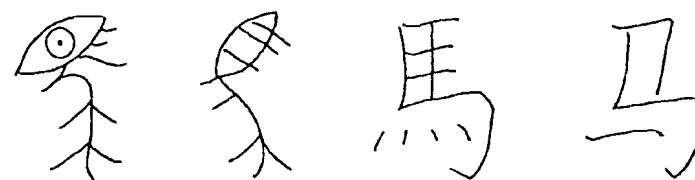


No doubt the Chinese had been making pictorial representations of horses for a long time before that; none of them has survived. But at some time in this process, an important change took place in the way such a representation was interpreted. At first, it represented a 'horse': that is, the animal itself—or, more accurately, a member of that class of animals recognised as a distinct category in the culture.

By the time this bone was cut, however, it no longer represented a 'horse'; it represented *horse*, a word of the language. (That is to say, it represented the Chinese word pronounced [mǎ], which means 'horse' in English). It could now be read aloud.

Let us express this change of function in linguistic terms. Functionally, the shape (Figure 2.2) is no longer a picture; it has become a CHARACTER. Since that time, the shape of this particular character has varied considerably, as shown in Figure 2.3. But its function has remained the same. Figure 2.3(d) is the modern Chinese character for the word *ma* (now pronounced [mǎ]).

Figure 2.3 Evolution of the *horse* character



(a) 1000-800 BC (b) c. 500 BC (c) 200 BC-1950 AD (d) modern 'simplified'

The same process took place with hundreds of other pictorial representations. From being pictures, representing classes of objects, they became characters, representing words. The shapes themselves did not have to change; what changed was the way they were understood.

In time, of course, the visual shape does also tend to change. Once the visual symbol has become a character, and especially when this has happened in enough instances so that not just a few words here and there, but most of the words of the language, can be represented—in other words, once a writing system has evolved—then the shapes tend to become regularised and simplified, in ways that are strongly influenced by the materials that are used for writing on and with: incising on bone, casting in bronze, chiselling in clay, painting on silk, and so on. But the change of form is not a necessary consequence of the change of function. What creates writing is not the particular shapes that are

from J
charac

used; it is the way the shapes are interpreted.

This process, of reinterpreting representations of things as representations of words, took place—we presume independently—in three different parts of the world: (1) in south-west Asia and north-east Africa (Sumeria and Egypt), (2) in China, and (3) in central America (with the Mayans). We have no clear traces of the process itself; but then it would be difficult to recognise them if we had, since as already pointed out it does not necessarily involve changes in the forms themselves. All we can say for sure is that writing had effectively evolved by 4000 BC in Mesopotamia and the Nile delta, by 2000 BC in northern China, and by the turn of the era in what is now Mexico.

The Chinese system of writing

It is a little misleading to say, therefore, that writing evolved as 'language written down', since there were pictures long before there was writing, and one element in the origin of writing lies in the re-interpretation of pictures as characters. It is equally misleading, at the other extreme, to say that writing evolved 'independently of language', since it only becomes writing when the symbols are understood as linguistic symbols. A more accurate account would be to say that writing evolves when what are originally non-linguistic symbols get mapped on to the forms of the language.

However, not everything in language can be drawn a picture of. There are always forms ready at hand to serve as characters for *horse* and *mountain* and *tree*; but, if we were to start creating characters for English, we should find it difficult to produce a picture representing an *error*, or *to know*, or *dull*; to say nothing of words like *and*, *of*, *not*, and *the*. To be able to write some words but not others is already well worth doing; and it is likely that this was in fact the situation for quite some time, when writing was restricted to certain esoteric functions like divining. But as writing comes to be extended across a broader range of functions in the culture—recording achievements, marking property, keeping the calendar, making inventories, collecting taxes, conveying instructions, and so on—it inevitably evolves into a full WRITING SYSTEM: that is, a system in which all possible wordings in the language are able to be (more or less unambiguously) represented.

Let us see how this process took place in Chinese, as described with remarkable accuracy by a Chinese linguist of the first century AD named Xǔ Shèn. I have slightly modified his account, for clarity of exposition; but his theory was essentially correct. (See Figure 2.4 for the forms of the characters described.)

1. A picture is taken to serve as an indirect representation: for example, a picture of a tower for the word *high*; a man with arms and legs outstretched for the word *big*; a carpenter's square for the word *work*; a hand for the word *five*.
2. A new picture is created to give an iconic representation: for example, a dot above a line for the word *above*; one, two, three, and four parallel lines for (respectively) the words *one*, *two*, *three*, *four*; a cross (symbolising 'first unit' and 'first decade') for the word *ten*. There are not many of this type.

Figure 2.4 Development of Chinese characters

1. Pictorial							
early	modern			early	modern		
		rì	'sun'			fān	'sail'
		yuè	'moon'			zhuī	'dove'
		mù	'tree'			zhǐ	'foot'
		mén	'door'			shǒu	'hand'
		shuǐ	'water'			rén	'man'
		nǚ	'woman'			dà	'big'
2. Iconic							
early	modern			early	modern		
		shàng	'above'			sì	'four'
		xià	'below'			bā	'eight'
3. Semantic compound							
early	modern						
		xìn	'trust'	[man standing by speech]			
		píng	'submit'	[kneeling man under hand]			
		wǔ	'military'	[foot under dagger-axe]			

4. Phonetic transfer
early modern

来 来 lái 'wheat' → lái 'come'

豈 豈 kǎi 'drum' → qǐ 'how?'

凡 凡 fán 'sail' → fán 'all'

5. Phonetic/semantic compound
(same phonetic element)

early modern

semantic
element
(‘radical’)

phonetic
element

维 维 wéi 'tie, rope' = 纟 'silk thread' + 隹 zhuī

推 推 tuī 'push' = 扌 'hand' + 隹 zhuī

谁 谁 shuí 'who?' = 言 'speech' + 隹 zhuī

推 推 chuí 'hammer' = 木 'wood' + 隹 zhuī

(same semantic element)

洋 洋 yáng 'ocean' = 氵 'water' + 羊 yáng

江 江 jiāng 'river' = 氵 'water' + 工 gōng

汗 汗 hàn 'sweat' = 氵 'water' + 干 gān

3. Two pictures are combined to form a semantic compound: for example, kneeling man under hand for the word *yield*; standing man by the side of speech for the word *trust*; sun in the middle of tree for the word *east* (where the sun rises); foot under dagger-axe for the word *warfare* (marching under arms).

In those listed up to this point there has been no connection made with the sound of the word. Two further strategies were adopted that involved taking account of sound, either (4) instead of or (5) as well as meaning.

what has
Chinese
evolved

4. A picture is 'borrowed' for a word of similar sound, by a process of phonetic transfer: for example, the character for wheat, Old Chinese [lǎg], for the word *come* (same pronunciation); the character for *war drum*, Old Chinese [k'ər] for the word *how?* [k'jər]; the character for *sail*, Old Chinese [biwǎm], for the word *all* (same pronunciation); the character for *flute*, Old Chinese [ŋjǎn], for the word *speech* (same pronunciation).
5. Two pictures are combined, one indicating the sound, the other indicating the meaning, to give a semantic-phonetic compound (a combination of the principles of 1 and 4 above): for example, the character for *dove*, Old Chinese [tjwər], used as phonetic element and combined with (a) the character for *silk thread* to represent the word *tie, rope* [djwər], (b) with the character for *hand* to represent the word *push* [t'wər], (c) with the character for *speech* to represent the word *who?* [djwər], (d) with the character for *tree, wood*, to represent the word *hammer* [d'jwər]. More than three quarters of all the characters used in the modern language are of this type.

The Chinese writing system, therefore, is a **CHARACTERY**; its symbols are characters. This means that they represent the wording of the language: the entities they stand for are words (or, more accurately, **MORPHEMES**, the smallest units of wording—to give an analogy from English, if the word *kindness* was written with a character there would be one character for the morpheme *kind* and another for the morpheme *ness*).

The technical term for a character, indicating its function in the language, is **LOGOGRAM**. Despite popular belief, characters are not ideograms, and Chinese writing is not ideographic. Characters stand for words, not for meanings. They are unambiguous when read aloud, and synonyms are not written alike; whereas if they were ideographic, synonyms would have to be written alike and there would be no unambiguous readings.

This kind of writing system is appropriate for the Chinese language. It is neither more nor less advanced than other writing systems, such as that of English; but the English writing system is different in a fundamental respect. In English, the written symbols represent the language not at the level of wording but at the level of sound. The next section describes how this system came about.

Chinese
'ideogr'
'logogr'

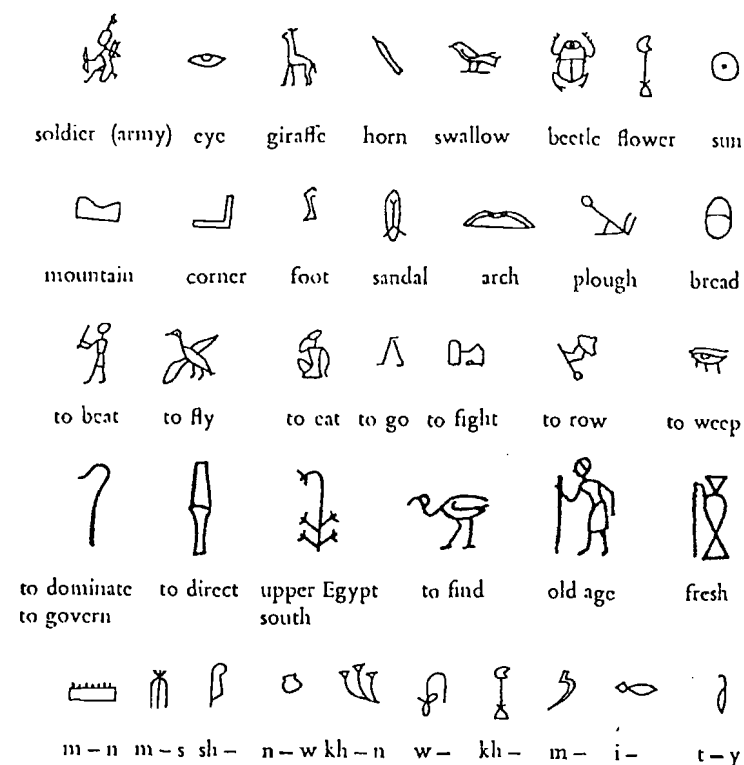
From ancient Egyptian to English

The first writing system developed in ancient Egypt was a character. Its characters are known as 'hieroglyphs', meaning 'sacred carving'.

Starting many centuries earlier, hieroglyphic writing had developed along the path that we have described above for Chinese (see Figure 2.5(a)). The principle of phonetic transfer was established by the third

millennium BC, and semantic indicators were added to form phonetic-semantic compounds (see Figure 2.5(b)).

Figure 2.5(a) Egyptian hieroglyphics, showing also development as syllabic symbols



2.5a. Hieroglyphic word signs. 1, symbols representing things shown
2, ideographs representing actions associated with things shown; 3, symbols representing abstract ideas; 4, hieroglyphic bi-consonantal signs.

At this point, however, the route taken was different from that in Chinese. In Old Chinese, all words were monosyllables; so there was never a case where one word was made up of the sounds of two or more others. In Old Egyptian, however, words could vary in the number of their syllables; so it frequently happened that one long word could be broken down into a number of syllables each of which was also a word, although quite unrelated to it; for example, the word *khesteb* 'turquoise' sounded like *khes* 'to stop' plus *teb* 'a pig'. This is the principle on which the game of charades is based, where we act out, for example, *can*, *knee*, *ball*, and then *cannibal*.

To write a polysyllabic word of this kind, it would be a natural step to break it down into its component parts. But notice that these 'parts' are not morphemes; they are syllables. The word *snowball* consists of the morphemes *snow* + *ball*; if we write it with the character

Figure 2.5(b) Determinatives in hieroglyphic writing



1. Heaven, Sky, Ceiling, what is above. 2. Night sky with a star hanging like a lamp from it, darkness, night. 3. (above) Sky slipping down over its four supports, storm, hurricane; (below) rain or dew falling from the sky. 4. Sun, the sun-god Ra, day period, time in general. 5. Shine, rise (of a luminary), being of light. 6. Moon, month. 7. Star, morning star, hour, time for prayer, pray. 8. Flourish, blooming, year, time in general, last year of a King's reign. 9. Foreign country, desert. 10. Mountain. 11. Island. 12. City, town. 13. Nome, District. 14. Water, watery mass of the sky. 15. Skin, hide. 16. Worm. 17. Plant, vegetable, herb, dried up. 18. Field, garden. 19. Grain, corn. 20. Man, first person sing. 21. Woman, first and second person sing. 22. God or divine person. 23. Pray, worship, adore, entreat, praise. 24. High, lofty, exalt, make merry. 25. To see. 26. To weep, tear, grief. 27. Hair (of men and animals), bald, lack, want, lacuna (in manuscripts), colour, complexion. 28. Phallus, front, male, masculine, procreate. 29. Women, goddesses, cities. 30. Sweet, pleasant. 31. Incense. 32. Roll of papyrus, tie up, bind together, come to an end. 33. Roll of papyrus (ried round the middle), book, deed, document, register, group together, abstract ideas. 34. Oval round a royal name, known as *cartouche*. 35. Pair of tallies, count, tally, reckon, pass by, depart. 36. Bread, cake. 37. Sign of the plural. 38. Negation, no, not, nothing, lack, want, need. 39. Horn.

for *snow* plus the character for *ball*, we have not changed the function of the symbols. But *cannibal* does not consist of the morphemes *can* + *knee* + *ball*; these are quite different morphemes, which happen to be represented by its component syllables.

If we choose to write *cannibal* with the symbols for *can*, *knee*, and *ball*, we hardly need the semantic indicators. Taken by itself, the character for the word *can* 'container', if transferred to stand for the word *can* 'am able', might conceivably cause problems of understanding. But a string of characters for the words *can* 'container', *knee*, and *ball* would make no sense at all, unless each was reinterpreted as standing for the syllables /kæn/, /ni/, /bəl/ which taken together would make up the word *cannibal*.

But once this step has been taken, the whole nature of the writing system becomes transformed. As long as there is, in principle, a separate

the di
in an
toward
represent
sound

character for each word, as was the case in Chinese (where phonetic transfers *without* semantic elements were confined to instances where the original word was no longer in use—or else the semantic indicator was added instead to the original, as when the character for *cloth* was added to the original picture of a sail), the writing remains logographic—the symbols represent the language at the level of wording. As soon as the picture of a can comes to be used for something that is *not* itself a word (or morpheme), but simply an element in the sound of another word, it is no longer functioning as a character—it has become a syllabic symbol. It will then occur equally naturally in the representation of all words containing the syllable /kæn/, like *cannabis*, *pelican*, *incandescent*, and so on. The character has been replaced by a SYLLABARY.

Notice that, once again, there is no need for the form of the symbol to change; it can still be a picture of a billycan. Only its function has changed. It started as the representation of a class of objects recognised in the culture as belonging to a single category: a 'can'. It was then reinterpreted, to represent the word *can*, the name of this category in the lexicon of the English language. It has now been reinterpreted over again, so that it represents the *syllable* /kæn/, which is an element of English phonology. When this change has flowed through the whole writing system, the symbols no longer stand for words but for sounds. The script has become a phonological one.

Character, syllabary, alphabet

As a matter of fact, this change never did take place fully in the Egyptian writing system, which always retained some of the features of a character. But it did take place in languages whose speakers borrowed their writing system from the Egyptian, of which the one that is significant for our purposes is Phoenician.

Phoenician was a Semitic language, like modern Arabic and Hebrew. The Phoenicians took over a small number of Egyptian symbols and used them as syllabic signs. Thus the Phoenician word for 'water' was *mem* (cf. Hebrew *mayim*); the Phoenicians took the Egyptian character for *water* and used it to represent the syllable /ma/—keeping the word *mem* as the name of the symbol (as we have names for the letters of our alphabet: /ei/, /bi/, /si/, /di/, etc.).

Similarly, the Phoenician word for 'snake' was *nun*; so they borrowed the Egyptian *snake* character and used it to represent the syllable /na/—calling the symbol itself *nun*. They borrowed about thirty symbols in all, and listed them a fixed order: first came the Egyptian 'ox' character, Phoenician word *ʔaleph* (from which we get our word *elephant*), beginning with a glottal stop, and hence used for the syllable /ʔa/; and second the Egyptian character for 'house', Phoenician *beth* (cf. Hebrew *beyth*), used for the syllable /ba/.

This kind of script was well suited to the Phoenician language, in which, as in modern Arabic, the root of a word is a sequence of (usually three) consonants; the vowels in between will vary (along with affixes before and after) to signal grammatical categories of person,

tense, number, and so on. So, for example, the consonant sequence /k-t-b/ means 'write, book', and yields a large number of words such as *katab* 'he wrote', *niktib* 'we write', *kitab* 'book', *kateb* 'clerk', *maktub* 'written', and so on. In a similar way the words *Islam*, *Muslim*, and *salaam* all come from the same root /s-l-m/ meaning 'peace'. In a language of this type, it is natural to have a writing system in which the symbol stands for a consonant plus any following vowel. The reader can be left to supply the appropriate vowel from the context; or alternatively, the vowel can be indicated by some additional diacritic, with perhaps the convention that if it is not marked then it is to be read as /-a/. There have been various forms of Semitic script, but all have been based on this kind of syllabic principle.

Next in line were the Greeks, who took over the Phoenician symbols and used them to write Greek. Greek, however, is a very different kind of language, in which vowels are just as much a fixed part of the word root as consonants are; moreover, there can be whole clusters of consonants in a single syllable, as in the word /straŋks/ meaning 'throat'. A syllabary, therefore, would be quite inappropriate. So the Greeks used each symbol to stand just for the consonant, without any following vowels; and they then added separate symbols for the vowels, either using Phoenician symbols for which they had no other use (like aleph—there was no glottal stop in Greek, so they adopted this symbol for the vowel /a/) or making up new ones for themselves. The result was an ALPHABET (so called because the Greeks also borrowed the Phoenician names for the symbols they took over, and these were the first two in order).

An alphabet resembles a syllabary in that its symbols stand for sounds, not words; but they stand for smaller units of sound—not syllables, but PHONEMES. In principle, one letter represents one phoneme; and that was more or less the case with the ancient Greek alphabet. This was then adapted to various dialects of Greek; and one of the dialect scripts was in turn borrowed by the Romans, who again adapted it slightly, left out some letters they did not need, and used it to write Latin. This Latin alphabet is essentially what we use for English today.

Table 2.1 summarises the various kinds of writing system.

Table 2.1 Kinds of writing system

Level of language represented:	lexico-grammatical (wording)	phonological (sound)	
Linguistic unit represented:	word/morpheme	syllable	phoneme
Type of symbol:	character ('logogram')	syllabic sign	letter
Type of script:	character	syllabary	alphabet

As usual, the categories themselves are clearly defined; but any given instance may be mixed or intermediate. Thus the Semitic scripts are not, in fact, pure syllabaries; they are in a sense intermediate between a syllabary and an alphabet. A stricter case of a syllabary would be the Japanese *kana* script, adapted from Chinese characters. And our

Greek
ment

sum
differ
writi



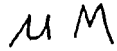


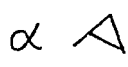


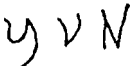
English writing system is certainly not purely phonemic.

There is a great deal of variation among different languages; but one thing is common to all: all languages are highly complex. When they are written down, the writing system has to be open-ended and flexible to accommodate the richness and complexity. When people try to design scripts, they usually make the mistake of making them too pure, and hence too rigid. When scripts evolve (which nearly always starts with borrowing—as we have seen, writing was developed independently in, at most, four contexts in human history, and even among these there may have been some transmission), they gradually adapt themselves to the needs of the particular language—which means they become somewhat messy and indeterminate. A writing system needs to be reformed now and again, because languages are always changing, whereas scripts, once codified, tend to stay as they are until someone takes positive action. But attempts to create ideal writing systems are bound to fail, because it is impossible to define what an ideal script should achieve—and if one could define such an ideal, it would certainly be impossible to attain.

A note on 'ideograms'

The symbols of all natural writing systems began as pictures. This is as true of the letters of our alphabet as it is of the characters of Chinese. Every time you write the word *man*, you are drawing three pictures—water, an ox head, and a snake (see Figure 2.6).

Figure 2.6 Evolution of letters m, a, n

Egyptian	Phoenician	Greek	Latin → English	
			M	m
water	/ma/	/m/	/m/	/m/
			A	a
ox	/ʔa/	/a/	/a/	/æ/
			N	n
snake	/na/	/n/	/n/	/n/

Functionally, on the other hand, no written symbol is ever a pictogram; in that sense, a 'pictographic script' is a contradiction in terms. If a symbol is part of a writing system, it must represent some element of a language; in that case it is not functioning as a picture.

What about the term 'ideogram'? I have avoided using that term because it is not at all clear what it means; it is used rather inconsistently in most discussions of language. But it is possible to make sense of it. If we refer back to the diagram in Figure 1.1, we can see that the level of representation of the writing systems so far described has been either lexico-grammatical or phonological. Nothing has been said of writing representing the semantic elements of a language.

As a general phenomenon, a semantic writing system would be an impossibility. The semantic systems of natural languages are so complex, with so many intersecting dimensions of meaning involved, that they could not be reduced to writing—for exactly the same reason, in fact, that they cannot be expressed in a protolanguage. The only possible writing systems are those whose symbols represent, as a general principle, either wording or sound.

In the seventeenth century there were various attempts, in England, Holland, and France, to create semantically based writing systems, attempts that were encouraged by a misunderstanding of the nature of Chinese characters. Scholars hoped in this way to produce a writing system that would be the same for all languages, and would serve to express the new scientific knowledge and new ways of reasoning. These schemes did not work; but a great deal was learnt about language in the process. One of the by-products, 150 years later, was *Roger's Thesaurus*.

But it is possible for a script to embody some use of semantic representations, as a minor theme; and perhaps the clearest example of this is Japanese. Until the fifth century AD, Japanese was not written down; then there were two large-scale invasions from China, as a result of which Japanese borrowed from Chinese both the writing system and a large amount of vocabulary. Japanese, however, is a very different kind of language from Chinese (to which it is also quite unrelated—Japanese is recognised to be an Altaic language, probably with an earlier substratum of Austronesian), and a character is entirely unsuited to it. What does suit it is a syllabary, and after a century or two the Japanese modified and simplified two sets of Chinese characters to create a syllabic writing system of their own.

By this time, however, there were many Chinese words in the language, which although they could be written in the syllabary (their phonology having become adapted to Japanese) were also entirely at home in character; so the Japanese retained the characters and used them side by side with their own syllabic signs. Various patterns grew up, the predominant one being Chinese characters ('kanji') for the lexical roots and Japanese syllabic signs ('kana') for the affixes and for grammatical words.

In this process, however, the characters were not confined to words borrowed from Chinese; they were also used to represent the roots of native Japanese words that were similar in meaning—the same character often being used for more than one Japanese word. Thus, for example, the character for the Chinese word *měi* 'beautiful' (Middle Chinese pronunciation [mjwi]) stands in Japanese for the following: (1) the word *mjwi* borrowed—twice, from different dialects—into Japanese, now pronounced either [bi] or [mi]; (2) the native Japanese word *utsukushii*, meaning 'beautiful'; and (3) the native Japanese word *yoi* meaning 'good'.

From the Chinese point of view, a character is tied absolutely to a particular word. From the Japanese point of view, however, that same character may stand for three or four different words, unrelated to each other in sound or form but related to each other in meaning. In other words, it tends to have for them a semantic as well as a lexico-

Can
system
stand
mean

some
Japane

grammatical significance: part logogram, part what we might call 'semogram'. It is in this sense that the Japanese themselves often refer to their writing as 'ideographic'. The term is inaccurate, since it is not 'ideas' but meanings that are being represented; but at least it makes an intelligible use of the term, so there is no reason why we should particularly reject it.

Some Japanese claim that it is because of their mixed script, partly phonological (the syllabary) and partly lexico-grammatical with a dash of semantic (the character), that they have little or no dyslexia in the population. It is impossible to prove this one way or the other; but it is an interesting idea—the virtues of a script that has something in it for everyone.

The English writing system

There is a tendency for mixed languages to get mixed scripts. Japanese is one example; English is another.

Like Japanese, English has been through a great deal of outside influence. After the English overran Britain, their language was strongly influenced by the native Celtic languages—hardly at all in vocabulary, but quite considerably in some aspects of its grammar. Next it was successively invaded by Norwegian, Danish, and Norman French; then in the Renaissance, it took over massive doses of Latin and Greek, not only lexical roots but also large numbers of affixes and the morphological processes that went with them.

Like every other European language, English inherited an alphabetic writing system; and after a few letters had been added (Latin had a very simple phonological system, so its alphabet is rather impoverished from the point of view of most other languages), it was excellently suited to the writing of Old English (Anglo-Saxon). The Norman French scribes destroyed some of its good qualities, by refusing to write the symbols they did not recognise; but what really perturbed it were two phenomena that took place in the language itself. One was the great internal upheaval that took place in Middle English (1100–1500), when the language changed extremely quickly and a dialectally mixed standard variety evolved; the other was the inflow of Graeco-Romance elements from 1450 onwards, already referred to above.

The effect on the writing system was likewise twofold. Just when the spelling was becoming standardised, it had suddenly grown rather archaic; the language had changed, and the spelling continued to reflect its earlier phonological patterns. Secondly, the Latin and Greek borrowings brought with them new phonological patterns from outside that had somehow to be reconciled with the native ones—while at the same time the Latin (and Latinised Greek) spelling was retained largely unaltered. There were thus two partially distinct phonological systems, compatible but not homogenised, each represented by different spelling conventions neither of which was particularly appropriate.

The French, who had similar problems, tackled them by setting up an Academy, which would legislate about the language and its orthography; the result was a writing system that is consistent but mas-

sively archaic. The English, equally characteristically, let things take their course, and ended up with a writing system that looks incredibly muddled, but in which the superficial messiness hides a rather effective compromise between the old and the new, the native and the foreign. It is far from perfect; but it has many virtues—not the least of which is that it is quietly neutral among all the various native and non-native forms of English that are now spoken around the world. When it was confined to England, and other English-speaking areas of the British Isles, it had already proved its ability to represent the various local accents of standard English. (These are not the original dialects, which have now largely disappeared except in some rural areas; they are vastly different and have their own orthographies.) Now, it serves not only the 'first language' English of Australia, New Zealand, South Africa, the Caribbean, Canada, and the United States but also the 'second language' English of many other parts of the world—South Asia, many countries of Africa, Singapore, and the South Pacific.

Like the Japanese script, English writing ought to be impossible to learn; but—again like Japanese—it is not. Its mixed character is also what makes it accessible. In the first place, to the extent that it is a phonological script, it is not phonemic. It allows various other principles to override the phonemic one. For example, it writes *photograph* in *photograph*, *photography*, and *photographic* all alike, even though their phonemic structure is very different; and similarly with many hundreds of other related sets. It allows Anglo-Saxon and Graeco-Romance words to have different spelling conventions rather than forcing one to adapt—incongruously, as it would be—to the other. It embodies strange, minor, but very useful conventions of its own, like the two-, three-, and four-letter rule (grammatical words can have two letters, lexical words must have at least three, and proper names, at least four; cf. the well-known example *Mr Inne is in the inn*). But at the same time, it works by tendencies and not by rules—which is exactly how language works as well.

In the second place, it is not entirely phonological, but also partly logographic. There are many sets of words in English that are pronounced in an identical fashion, but are kept apart in the spelling. There is no necessity for this, of course; but it is useful for two reasons. One is the dialectal neutrality referred to above. For example, in my own dialect *paw*, *poor*, *pour*, and *pore* are all identical, whereas for many speakers of English there are two or even three different syllables among them; on the other hand, we distinguish *higher* and *hire*, which many English speakers pronounce alike. The spelling allows for all sorts of different groupings.

The other reason is that, although such homonymy causes no trouble in speech, written language is not spoken language written down. It has a life of its own, in which it is useful to be able to use words without the same environmental cushioning that is characteristic of speech. When we talk, there is always a context; it poses no great problem that many words are pronounced alike. In writing, however, where the whole object is to get away from dependence on the immediate environment while still remaining unambiguous, it is useful to be able to put up a notice saying *wait for pause after whole lessons* without

the evol
English
system

Some fo
English
limited
kind of
purely l
sound;
tendenc

at the same time saying *weight four paws after hole lessens*. With a purely phonological script, the written text makes the same demand on the context as the spoken one.

Up to this point, we have been exploring the origin of speech and the development and nature of writing. It is now time to turn to the exploration of written language.

Chapter 3

Written language

Codified and codable expressions

Up to this point we have been assuming that whatever is spoken can also be written—that writing is simply an alternative form of expression to speech. We now need to examine this assumption a little more closely.

In the broadest sense, the assumption can be allowed to stand. That is to say, a writing system is capable of **representing** all possible wordings in the language: (1) by providing ready-made ('codified') expressions, for the majority of elements, and (2) by providing the means of creating ('coding') expressions for elements that are not already codified—new borrowings and coinings, an individual writer's neologisms, mistakes (for example, children and foreigners), and the like. So English, for example, contains (1) recognised **spellings** for the great majority of its words, and (2) recognised **principles** of spelling that can be applied where the spellings do not yet exist.

At an earlier stage in the language, the early Modern English period when standard English was emerging and printing had just begun, there was much less codification; writers used a variety of different spellings. But if principle (1) had not yet been generally applied, principle (2) held good: the variation was within the limits of tacitly agreed practice, and there was no problem of intelligibility—texts could be read without difficulty (and the literacy rate among adults was for that time extremely high; it has been estimated that in the fifteenth century over half the population could read). We tend to take it for granted that spelling should be totally uniform; but there is no compelling reason why it should be, provided the principles are clear. We understand each other's spoken language throughout the English-speaking world, unhampered by the wide variation in dialect—because all dialects are underlain by what is, by and large, a single phonological system. The same principle will work for writing.

As a rule, however, writing systems tend to engender conformity once they come into general use; partly for convenience, and partly because the development of writing tends to be associated with norma-

English
provide
dardise
for old
(2) prin
spelling