

## Chapter 10

# Writing

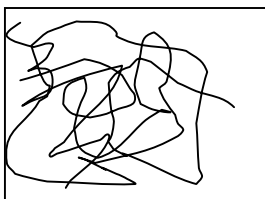
All writing systems in current use have, directly or indirectly, their roots in Asia or North Africa. These continents have played an exceptionally important role in the development of writing.

The connection between the use of writing and the development of a higher civilization is controversial. Some scholars see such development as a consequence of writing, while others see writing as a consequence (or an integral part) of such development. Like writing, all world civilizations have, directly or indirectly, their roots in Asia and North Africa.

However, Asia and Africa also have an immense number of languages without a writing system, of cultures based on orality rather than literacy. These continents therefore provide fertile ground for the study of the contrast between oral and literary cultures.

## 10.1 What is writing?

Writing includes, at least in its typical form, **visual marks** on a surface. There is, of course, also the Braille alphabet and other writing systems for the visually impaired that are based on tactile rather than visual marks, but these will not be discussed in the present chapter. Even inscriptions and carvings may have both visual and tactile elements, but they are usually read with the eye rather than with the sense of touch.



Visual marks, however, do not always constitute writing. A little child who has got hold of some crayons and a sheet of paper and covers part of the paper with colour is not producing writing. The visual marks have to be **representational**. In order to constitute writing, they must stand for something else.

Even representational visual marks, however, do not necessarily constitute writing. Otherwise all representational visual art would count as writing, from Leonardo's *Mona Lisa* to Picasso's *Guernica*. In addition to being representational, writing implies that the connection between the visual marks and what they represent is **conventional**. By convention, the Hebrew letter <א> is used to represent the sound /r/, the West-African Vai letter <I> stands for the syllable /na/, and the Chinese character 人 stands for the word which has the pronunciation (in standard Mandarin) *rén* and the meaning 'person'.



We all know, however, visual marks that are representational by convention, but still do not constitute writing. They may be simple symbols like traffic signs or the visual image of a man and/or a woman (and/or a wheelchair) to indicate the presence of a toilet. Or they may be



complex notational systems like musical notation. What writing represents is not just any kind of object or idea. It represents **language**. Writing, in this sense, is always linguistic. Although partly autonomous, it is never entirely divorced from the spoken language.

Writing may always be read aloud, and even when we read silently, we often do so by way of subvocalization. In fact, writing always represents not just language in a general sense, but one specific language. A language may borrow the writing system of another language, but seldom or never without adjustments. In contrast to writing, the image of a man and/or a woman and/or a wheelchair is not meant to represent any specific word for 'toilet' or 'restroom', just the presence in the immediate vicinity of a toilet. Likewise, although the notes and other symbols of musical notation may have names in a given language, their function is not to represent these names, but to represent the musical tones directly.



By excluding more or less closely related concepts, we have arrived at the following definition of writing:

**Writing** is a conventional system of representing language by means of visual marks.

This definition does not solve all problems, but it gives a good indication of the object of our study.

## 10.2 Writing goes via speech

One of the things this definition does is to make clear that the idea of **ideographic** writing is a misconception. Ideographic writing is conceived as a system of writing where the signs represent "ideas" directly without going via any particular language, so that people who speak different languages can still communicate in writing. The notion of ideographic writing has often been boosted by the fact that many Chinese characters may be understood by people who speak widely different languages, and



who therefore pronounce these characters differently. For instance, the Chinese character on the left means 'person' in both Chinese and Japanese, but is pronounced *rén* in Chinese and *hito* in Japanese. It has often been concluded that Chinese writing is ideographic, referring directly to the concept of 'person' rather than to the specific words *rén* and *hito*. But this is not and could not be true. Writing

always represents a specific language. In Chinese and Japanese, many characters with the same written form and the same general meaning may have widely differing pronunciations, but that does not mean that writing has bypassed speech, only that one writing system has been transferred and adjusted to another language. In Chinese, the written form of the character in question still represents the word pronounced *rén* and meaning 'person', while in Japanese, the same written form represents the word pronounced *hito* and meaning 'person'. In some respects, a writing system in which each sign refers to a word may be preferable to one in which each sign refers to a sound, and the effect may be closer to the desired effect of ideographic writing.

Mathematical notation is probably the closest we get to universal ideographic writing. All languages contain numerals, and most languages possess linguistic means for the expression of mathematical calculations, as in the following examples from English, Norwegian and Chinese:

One plus one equals two.  
En pluss en er lik to.  
一加一等于二。

The use of mathematical notation cuts across these language differences:

$$1 + 1 = 2$$

Mathematical notation is intended to relate directly to the ideas behind such calculations, not to the linguistic expression of them in any single language. In this sense, it is ideographic. But precisely for this reason, it is not, in our definition, writing, just a notational system with a very limited scope. The same is true of the notational systems of logic.

## 10.3 Orality and literacy

One of the founding fathers of modern linguistics, Ferdinand de Saussure (1857-1913) said the following about the relationship between language and writing:

Language and writing are two distinct systems of signs; the second exists for the sole purpose of representing the first.

By language he meant spoken language. Spoken language is primary, and written language secondary. Spoken language may exist without a corresponding written language, while the reverse is not true. In fact, the vast majority of languages in the world have never or only recently been reduced to writing.

But does writing exist for the *sole* purpose of representing speech? Of course not. Just like speech, writing exists primarily for the purpose of representing concepts and ideas. In order to do so, however, the written language always goes at least partly via the spoken language, while the spoken language does not need to go via the written language.

Writing is by no means fully subordinate to speech. In many respects, writing leads a life of its own and develops independently of speech.

Writing also influences speech. The most obvious example of this is spelling pronunciation, as when English pronounces the word *theatre* with a /θ/ instead of the original /t/, because <th> is pronounced /θ/ in other English words. More significantly, complex syntactic structures that are more common in writing than in speech, tend to increase in the speech of people who often write and in cultures that are dominated by writing. If we accept that language influences our way of thinking (see 2.5), this may have wide-reaching consequences.

The main differences between writing and speech are consequences of the simple fact that the former is visible, while the latter is audible:

1. Writing is permanent, while speech is evanescent.

2. Writing is independent of the discourse situation, while speech is not. The physical existence of a written text does not depend on the presence of the writer or the reader, while the spoken text (at least before the tape recorder) needs the presence of both the speaker and the hearer.
3. A writer can usually take his time preparing his text, with no reader present, while a speaker often has to produce it in the heat of the moment, with the hearer present.
4. A reader can usually take his time interpreting a written text, with no writer present, while a hearer usually has to interpret it there and then, with the hearer present.
5. Writing tends to be monologic and speech dialogic, since writing often leaves less room for the writer to further explain his intentions or even to discover whether he has been understood, and less room for the reader to inquire about the intentions of the writer.
6. Since a written text may always be reread, while a spoken text vanishes as soon as it has been uttered, writing reduces the need for memorization, while spoken language often develops a wide variety of techniques and formulas for memorization.
7. Due to the longer time available to both writer and reader, writing allows for more grammatical complexity as well as abstract conceptualization and argumentation than speech.
8. Through intonation, voice quality and similar means, speech allows for more nuanced expression of emotional tone than writing, which has fewer expressive tools of this type.
9. Being disconnected from the discourse situation, writing allows for more distance and objectivity than speech.
10. Writing allows the writer a greater freedom to pursue an argument independently of his social environment, possibly stimulating originality, while speech tends to be more traditional and conventional.
11. Writing allows a higher degree of fossilization of norms and patterns (for instance through written laws), while speech is often more flexible.
12. Writing reduces the need for redundancy and repetition as compared to speech.

These and other differences are not absolute. Different written texts contain different amounts of orality features, and vice versa. For instance, modern e-mail language often has many features usually associated with orality, and this may lead to problems when the writer of an e-mail forgets that the lack of an immediate discourse situation and the relative lack of means to express emotional tone may produce unintended interpretations in the reader. To some extent, the lack of means to express emotional tone has been alleviated by the introduction of conventional signs like :-)) and :-(.

Compare the following two translations of the opening lines from Genesis in the Bible:

1 In the beginning  
2 God created heaven, and earth.  
3 And the earth was void and empty,  
4 and darkness was upon the face of the deep;  
5 and the spirit of God moved over the waters.

1 In the beginning,  
2 when God created the heavens and the earth,  
3 the earth was a formless wasteland,  
4 and darkness covered the abyss,  
5 while a mighty wind swept over the waters.

The first translation is based on the Douay translation from 1610, when English language culture had still retained many features of orality (corresponding to similar features in the Hebrew original), while the second translation is from *The New American Bible*, published in 1970. Apart from some differences in the choice of words, the main difference is the degree of grammatical complexity. In the Douay translation, the four clauses comprising lines 2-5 are simply juxtaposed additively by means of grammatical coordination. In *The New American Bible*, both line 2 and line 5 have been made into a subordinate clause expressing time. In the Douay translation three clauses are introduced by the coordinative conjunction *and*, while *The New American Bible* only begins one clause with *and*, while two clauses are introduced by the subordinative conjunctions *when* and *while*. Subordination is grammatically more complex than coordination.

Writing is undoubtedly an extremely important cultural innovation, with a social impact at least as strong as later technological innovations like book-printing and, more recently, digitalization. Literate cultures tend to produce different literatures, different religious and philosophical ideas, and different forms of organization and administration from oral cultures. Would the pyramids of Egypt or the Great Wall of China have been possible without writing? What about the great Greek tragedies? Or the ancient philosophies of Greece and China?

Like any technology, however, writing has also been met with scepticism. Plato, although a writer himself, sees writing as a mechanical, inhuman way of treating knowledge, which destroys the ability to memorization, and which goes against his ideal of dialogic communication. In India and Persia, holy texts were (and to some extent still are) committed to memory long after the introduction of writing, because writing was considered inferior to recitation by heart.

As children of a highly literate culture, we sometimes underestimate the possibilities for literary, philosophical and organizational creation in oral cultures. For us it is difficult to fathom how the *Iliad* and the *Odyssey* could be produced and for a long time transmitted by people who could not read or write. It is almost incredible how the vedas of India and the holy books of Persia were recited in their original languages for a thousand years or more before they were committed to writing. And it is extremely impressive how 500 years ago the illiterate Incas of the Andes were not only masters of advanced architecture and engineering, but were also able to build an empire covering an area the size of France, Italy, Switzerland, Belgium and the Netherlands combined.

## 10.4 Elements of writing

A **writing system** is a set of rules relating a given set of written **signs** to the linguistic units represented (e.g. to sounds). The set of physical written signs is called a **script**.

Writing systems may differ from one another in the script used, in the underlying set of rules, or in both. For instance, the Korean writing system has borrowed many of its underlying principles, such as the combination of syllable representation and sound representation, from Indian models, but has no similarity in script. On the other hand, the writing system devised around 1820 for the Cherokee language makes heavy use of Latin (and partly Greek) script, but the underlying principles of representation are completely different, so that <A>, <B>, <C> and <D> represent the syllables /go/, /yī/, /tli/ and /a/, respectively. Finally, the Greek and Chinese writing systems have neither script nor underlying principles in common.

Sometimes, one and the same language may be written with two (or more) alternative writing systems. For instance, Hindi (written with the Indian devanāgarī system) and Urdu (written with a variant of the Arabic alphabet) are in many ways the same language. This is also the case with Serbian (written with the Cyrillic alphabet) and Croatian (written with the Roman alphabet). This is sometimes referred to as **digraphia**.

At other times, one and the same language makes simultaneous use of several scripts and underlying principles. Japanese, for instance, is written with a combination of Chinese characters and the two Japanese scripts hiragana and katakana, so that Chinese characters are used for many full words (nouns, verbs, adjectives etc.), hiragana is used for most functional words and affixes, and katakana for loanwords. The underlying principles behind the hiragana and katakana scripts are the same, both containing the same number of signs representing the same syllables. The underlying principles behind the Chinese characters are different, since these characters represent meaningful elements like words or morphemes. In addition to these, the Latin alphabet is now often used to represent Western loanwords, introducing a fourth script with yet another set of underlying principles. Japanese, therefore, is a good example of a **mixed writing system**.

The signs of a writing system may be divided into **basic signs** and **diacritics**, the latter being small extra signs added to the basic signs. For instance, French has four different diacritics that combine with vowels, <é>, <è>, <î> and <â>, and one that combines with a consonant: <ç>. With the exception of <â>, where the diacritic marks the historical presence of an <s>, all of them mark aspects of the pronunciation of the sounds to which they are attached. In many writing systems, suprasegmental features like stress, accent, tone, length and nasality are indicated by means of diacritics. In Semitic and Indian writing systems, vowels are often indicated by means of diacritics.

Some scripts contain **ligatures**, combinations of two individual signs to form a single sign. In Indian writing systems, ligatures are widely used to represent consonant clusters. In English, the ligature <æ> is sometimes used in words like *encyclopedia*.

In some writing systems, what is functionally the same sign comes in different variants according to its position in a word, a syllable or a sentence. Many writing systems, for instance, have special signs for **initial** consonants or special signs for **final** consonants. In several Greek-derived alphabets, the distinction between **uppercase** and **lowercase** letters sometimes plays a similar role, uppercase letters introducing a new sentence or proper names (earlier used to introduce all nouns). In

Greek-derived alphabets, various **styles** (plain, *cursive*, **bold**, underlined etc.) of the same letters may be used for a variety of purposes, such as emphasis.

**Punctuation marks** (such as <.>, <,>, <:>, <;>, <!>, <?>, <"> etc.) are used in many writing systems, both to aid parsing into sentences and clauses, to mark the distinction between declarative, imperative and interrogative sentences, and to mark emotional tone. The use of an empty space to mark word boundaries is a similar device. Historically, both punctuation marks and word boundary marking has varied, and word boundaries are still not marked in Chinese and Japanese (though they are marked in modern Korean).

The **direction** of writing is usually more or less fixed, and the following three types are by far the most common:

1. From right to left, lines running from top to bottom.
2. From left to right, lines running from top to bottom.
3. From top to bottom, lines running from right to left.

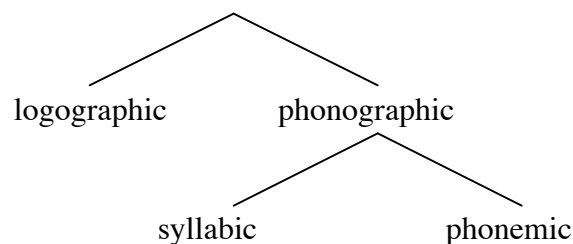
Type 1 includes most Semitic scripts, like Hebrew and Arabic. Type 2 includes Indian-derived scripts, like devanagari, Tamil, Thai and Tibetan, and Greek-derived scripts, like Latin and Cyrillic. Type 3 traditionally includes all Chinese-derived scripts, including Japanese and Korean, though these are now often written from left to right as in type 2. One rare, but interesting type is the so-called **boustrophedon**, where every other line runs from right to left and from left to right, as in some old Greek texts.

## 10.5 Types of writing

Writing systems may be divided in two according to what kind of linguistic unit their signs represent. In **logographic** writing, each sign represents a meaningful element like a word or a morpheme. In **phonographic** writing, each sign represents a phonetic or phonological element with no reference to meaning. Most modern writing systems are phonographic, Chinese writing being the most prominent example of logographic writing.

Phonographic writing systems may be **syllabic** or **phonemic**, according to whether each sign represents a syllable or a phoneme. Japanese hiragana and katakana are examples of syllabic systems, while the Latin alphabet and Korean hangŭl are examples of phonemic writing.

Very broadly speaking, therefore, there are three basic types of writing: logographic, syllabic and phonemic:



The difference between these three types is also reflected in how the linear ordering of written signs reflects the temporal ordering of speech elements. In logographic writing, the linear ordering of signs reflects the temporal ordering of words or

morphemes; in syllabic writing, the linear ordering of signs reflects the temporal ordering of syllables; and in phonemic writing the linear ordering of signs reflects the temporal ordering of phonemes.

This may not be surprising, but is useful to keep in mind when we want to evaluate whether a given writing system is, for instance, syllabic or phonemic. For there are, of course, no pure types. Many phonographic systems hover between syllabic and phonemic writing. In the English word *hope*, for instance, the phoneme sequence /ou/ is represented by a combination of the letters <o> and <e>, which, however, occur on each side of the <p>. Such breeches of linearity indicates that English writing is not purely phonemic. Another example is the placement of vowel signs in many Indian scripts, where it is not uncommon that a vowel sign is placed before the consonant sign even if the vowel actually follows the consonant, suggesting that Indian writing systems may be more syllabic than phonemic.

### 10.5.1 Logographic writing

Chinese writing is the only fullfledged logographic writing system in current use. It consists of a vast number of signs or characters, of which several thousand are in common use. Generally, each character represents a monosyllabic morpheme.

木

For instance, the character on the left represents a morpheme with the pronunciation *mù* and the meaning 'wood'. In modern Chinese, this morpheme is not commonly used as a free word, but often occurs in compounds like *hóng-mù* 'padauk' (literally 'red-wood') and *shù-mù* 'tree' (literally 'tree-wood'), and in derivations like *mù-tou* 'wood' (-*ou* being a derivational suffix). Some morphemes may constitute full words of their own, others may not.

咖

Not all characters, however, represent morphemes. This is especially true in loanwords like *kāfēi* 'coffee', which is written with the two characters to the left, none of which represents a morpheme of its own. In general, disyllabic morphemes are written with two characters. Note, however, that although *kāfēi* is clearly one morpheme, there is often a tendency to reinterpret such collocations to give each character (or, to be more precise, the speech element that each character represents) status as an independent morpheme. In Taiwan Mandarin, for instance, the first syllable of *kāfēi* is used to represent the whole word *kāfēi* ㄉㄧㄢˋ 'coffee-shop; café' in the compound *wǎng* ㄎㄚ̄ 'internet café'.

With very few exceptions, all Chinese characters represent a single syllable. Although *kāfēi* is only one morpheme, it would not have been possible to represent it with only a single character. On the other hand, a Chinese character always represents a syllable as it occurs in a given word or morpheme. The syllable *kā*, for instance, is written in at least three different ways, as indicated on the right. The first character is only used in the word *kāfēi* and, with a different pronunciation, in the word *gāli* 'curry'; the second character is used in the onomatopoeion *kāchā* (indicating the sound of something that breaks in two) and in some renderings of non-Chinese place names; and the third character, in

啡

咖  
喀  
刮



this pronunciation, is used as a verb meaning 'to slice with a knife'. Therefore, even when more than one character is needed to write a morpheme, each character still retains an indication of morphemic identity.

The term logographic seems to indicate that each sign represents a word (Greek *lógos*). This is clearly not the case in Chinese. Instead, each character represents a monosyllabic morpheme or a syllable in a polysyllabic morpheme. Some scholars have proposed, therefore, to call this writing system **morphosyllabic**.

### 10.5.2 Syllabic writing

One of the standard examples of syllabic writing is Japanese kana, which actually represents two different scripts, hiragana and katakana, with the same underlying principles for interpretation. In both, each sign corresponds to a syllable. The following is an overview of hiragana (blue) and katakana (green), with the initial consonant indicated on the left and the vowel indicated at the top:

	A		I		U		E		O	
	あ	ア	い	イ	う	ウ	え	エ	お	オ
K	か	カ	き	キ	く	ク	け	ケ	こ	コ
G	が	ガ	ぎ	ギ	ぐ	グ	げ	ゲ	ご	ゴ
S	さ	サ	し	シ	す	ス	せ	セ	そ	ソ
Z	ざ	ザ	じ	ジ	ず	ズ	ぜ	ゼ	ぞ	ゾ
T	た	タ	ち	チ	つ	ツ	て	テ	と	ト
D	だ	ダ	ぢ	ヂ	づ	ヅ	で	デ	ど	ド
N	な	ナ	に	ニ	ぬ	ヌ	ね	ネ	の	ノ
H	は	ハ	ひ	ヒ	ふ	フ	へ	ヘ	ほ	ホ
B	ば	バ	び	ビ	ぶ	ブ	べ	ベ	ぼ	ボ
P	ぱ	パ	ぴ	ピ	ぷ	プ	ぺ	ペ	ぽ	ポ
M	ま	マ	み	ミ	む	ム	め	メ	も	モ
Y	や	ヤ			ゆ	ユ			よ	ヨ
R	ら	ラ	り	リ	る	ル	れ	レ	ろ	ロ
W	わ	ワ							を	ヲ
N	ん	ン								

One reason why this works so well for Japanese is that the language has an unusually small number of syllables. Most syllables consist of a vowel or a diphthong which may or may not be preceded by a single initial consonant. The only possible syllable-final consonant is /n/. With a small number of possible syllables, a syllabic writing system works quite neatly. A similar system for English would have been completely unmanageable, due to the large number of possible syllables in English, probably more than 10000.

Note, furthermore, that each sign does not necessarily represent an entire syllable. Instead it represents a **mora**, a unit corresponding to the time it takes to pronounce an initial consonant followed by a single short vowel. For instance, the hiragana sign **か** represents the mora /ka/. If the vowel is long, however, the syllable takes longer to pronounce and counts as two moras, which must be written with two kana signs, as when **かあ** (corresponding to *ka + a*) represents the syllable /ka/. The same applies if the vowel is followed by an /n/, as in **かん** /kan/. In fact, therefore, kana is not strictly speaking syllabic, but **mora-based**.

Like most phonographic writing systems, kana has its irregular features. For instance, the topic particle pronounced /wa/ is written as though it were pronounced /ha/: **は**. In such cases, it is impossible to know how to write a given sound sequence unless one knows which morpheme or word is involved. In other words, this is a logographic element within a basically phonographic system. In Japanese, such elements are relatively few.

### 10.5.3 Phonemic writing

The Latin alphabet, as well as other alphabets derived from Greek, is designed as a phonemic writing system. English, however, although it makes use of the Latin alphabet, is an extremely bad example of a phonemic writing system. We have already mentioned how it breaks the most elementary demands for linearity in *hope*, *rote*, *scope*, *rope* and other words where the discontinuous combination of <o> and <e> represents the diphthong /ou/. In English, furthermore, one and the same sound or sound sequence may often be written in a number of different ways, and one and the same letter or letter combination may often be pronounced in a number of different ways. The author Mark Twain proposed in jest that the English word pronounced /fɪʃ/ ought to be spelled <ghoti> rather than <fish>, with <gh> pronounced /f/ as in *rough*, <o> pronounced /ɪ/ as in *women*, and <ti> pronounced /ʃ/ as in *nation*. It is most often impossible to know the spelling of a word on the basis of its pronunciation alone, and it is often impossible to know the pronunciation of a word on the basis of its spelling. To a very large extent, the written form of each word has to be learned separately. The logographic element is much higher than what we saw in Japanese.

Korean hangul writing is a much better example of a phonographic writing system. In its modern form, it has few irregular features. It may seem odd that the same sign, **ㅇ**, is used both for the lack of a consonantal initial and for the presence of the consonantal ending /ŋ/, but this is a regular feature throughout the system. It also deviates from actual pronunciation to keep the graphic shape of a morpheme constant, in the same way that English writes its most common plural marker <s> whether it is pronounced /s/, /z/ or /ɪz/ (**morphophonemic** spelling). Apart from that, it is usually possible to know the written form of a Korean word on the basis of its pronunciation,

and it is as good as always possible to know the pronunciation of a Korean word on the basis of its written form.

The consonants in Hangeul:

ㄱ	ㅋ	ㄴ	ㄷ	ㅌ	ㄹ	ㅁ	ㅂ	ㅃ	ㅇ
k	kk	n	t	tt	r	m	p	pp	silent consonant
ㄸ	ㅆ	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
s	ss	ng	ch	jj	ch'	k'	t'	p'	h

The vowels and diphthongs in Hangeul:

ㅏ	ㅑ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅠ	ㅡ
a	æ	ya	yæ	eo	e	yeo	ye	o
ㅘ	ㅙ	ㅚ	ㅝ	ㅘ	ㅙ	ㅚ	ㅜ	ㅠ
wa	wæ	we	yo	u	wo	we	wi	yu
ㅞ	ㅟ	ㅣ						
eu	eui	i						

Each sign in the Korean writing system represents a phoneme. However, these signs are arranged in a way that also gives a high prominence to the syllable. For instance, the syllable consisting of the three phonemes /p/, /o/ and /m/ is not just written as a succession of the signs ㅍ ㅜ ㅁ, but as a cluster representing the entire syllable in one square of a given size: ㅍㅜㅁ. Although each sign represents a phoneme, the syllable is also a primary unit, so hangeul has elements of a syllabic writing system.

In addition, a comparison of the various signs in the chart above reveals that hangeul not only represents syllables and individual sound segments, but even some of the phonetic properties of these individual sound segments. For instance, most of the letters that are pronounced with the tip or the blade of the tongue up against the alveolar ridge include the same right-bending hook: ㄷ ㅌ ㄴ ㄹ ㅈ. It has been argued, therefore, that hangeul is a **featural** writing system, where the phonetic features of each sound are represented in writing.

The hangeul writing system, therefore, although basically phonemic, also has elements of syllabic and even a featural writing system.

### 10.5.4 Consonantal writing

Phonemic writing systems more often represent consonants than vowels. This applies especially to Semitic writing systems, reflecting the fact that word roots in Semitic languages tend to consist of three consonants, with the vowels expressing grammatical distinctions that are largely deducible from the context.

The writing system of modern Hebrew, so-called Hebrew Square, is an example of a pure **consonantal** system, with no vowel signs:

י	ט	ח	ז	ו	ה	ד	ג	ב	א
Yod	Tet	Chet	Zayin	Vav	He	Dalet	Gimel	Bet	Alef
(Y)	(T)	(Ch)	(Z)	(V)	(H)	(D)	(G)	(B/V)	(silent)
ע	ס	נ	נ	מ	מ	ל	ך	כ	
Ayin	Samech	Nun	Nun	Mem	Mem	Lamed	Khaf	Kaf	
(silent)	(S)	(N)	(N)	(M)	(M)	(L)	(Kh)	(K/Kh)	
ת	ש	ר	ק	ץ	צ	ף	פ		
Tav	Shin	Resh	Qof	Tsadeh	Tsadeh	Feh	Peh		
(T)	(Sh/S)	(R)	(Q)	(Ts)	(Ts)	(F)	(P/F)		

This may seem like a deficiency, but is hardly more so than the lack of means to represent suprasegmental features like stress, tone, length and nasalization in many writing systems. As long as the lack of such means does not impair reading, it is no problem. And just as suprasegmental features may be represented by diacritics when they are felt to be needed, the same is true of vowel segments in many Semitic writing systems. In modern Arabic, for instance, there are signs for long vowels, which are always written, and diacritic marks for short vowels, which are, however, used less often.

### 10.5.5 Semi-syllabic writing

Most Indian writing systems, such as devanāgarī in the chart below, combine features of syllabic and phonemic writing.

Devanāgarī signs include both vowel signs and consonant signs.

The vowel signs found in the chart below, however, are only used when the vowel is not preceded by a consonant. Whenever a vowel is preceded by a consonant, it is instead represented by a diacritic mark attached to this consonant. For instance, the syllable /ke/ is written with the consonant sign क for /k/ and a diacritic mark <sup>̄</sup> for /e/: के. The vowel sign ए for /e/ is not used, since the /e/ is preceded by a consonant.

Consonant signs without a vowel diacritic are interpreted as containing an inherent /a/. For instance, if क is used without a diacritic, it is usually not just /k/, but /ka/. Hence there is no diacritic mark for /a/. Sometimes, however, consonant signs may also represent the consonant without an inherent /a/. The "muting" of the inherent /a/ may be marked by the "stop" sign <sub>॥</sub>, but often is not marked at all, as in Hindi

*mālik* मालिक 'employer'. Consonant clusters are typically represented by ligatures.

For instance, the consonants क /k(a)/ and व /ʃ(a)/ combine into = /kʃ(a)/.

अ	आ	इ	ई	उ	ऊ
a	ā	i	ī	u	ū
ऋ	ॠ	ऌ	ॡ		
ṛ	ṝ	ḷ	ḹ		
ए	ऐ	ओ	औ	अं	अः
e	ai	o	au	aṃ	aḥ
क	ख	ग	घ	ङ	
ka	kha	ga	gha	ṅa	
च	छ	ज	झ	ञ	
ca	cha	ja	jha	ña	
ट	ठ	ड	ढ	ण	
ṭa	ṭha	ḍa	ḍha	ṇa	
त	थ	द	ध	न	
ta	tha	da	dha	na	
प	फ	ब	भ	म	
pa	pha	ba	bha	ma	
य	र	ल	व		
ya	ra	la	va		
श	ष	स	ह		
śa	ṣa	sa	ha		

By themselves, all devanāgarī signs represent full syllables. The consonant signs, however, are used for individual phonemes in all cases except where these are followed by /a/. This also implies that all vowel phonemes except /a/ are marked, albeit often with diacritic marks rather than full signs. Thus, devanāgarī has features both of a syllabic and a phonemic writing system.

An important question is whether the linear ordering of signs in such a writing system reflects the temporal ordering of phonemes or only of syllables. The answer is that the linear ordering of signs, if diacritic marks are included, does not always reflect the temporal ordering of phonemes. For instance, the syllable /ki/ is written कि, with the diacritic  $\bar{i}$  actually preceding the consonant sign क /k/. This is an argument for treating devanāgarī as a syllabic rather than a phonemic writing system.

## 10.6 The history of writing

Writing may have been invented at least three times in history:

1. **Sumerian** writing in Mesopotamia around 3200 BC
2. **Chinese** writing along the Yellow River before 1200 BC
3. **Mesoamerican** writing in Guatemala and Southern Mexico a little after 500 BC

Of these three, only Chinese writing is still in use today, although both Sumerian and Mesoamerican writing were also used for several thousand years.

A fourth type, the **Egyptian** writing system, emerged just shortly after Sumerian writing appeared in Mesopotamia. While it may have been inspired by Sumerian writing, the signs of the two writing systems do not resemble each other.

A fifth type, the North Semitic **Phoenician** writing system, became more influential than any other. Phoenician writing is the ancestor of most modern writing systems, including Greek, Latin, Cyrillic, Arabic, Hebrew and probably also the large number of Brahmi-derived scripts in South and South-East Asia. Phoenician writing may ultimately descend from Egyptian writing, but this is uncertain.

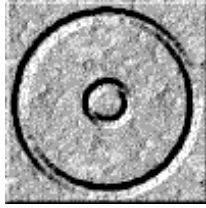
Many ancient writing systems are as yet undeciphered. This is the case with Proto-Elamite (2900 BC in Persia), the Indus script (2500 BC in the Indus valley) and Linear A (1800 BC on Crete). In most cases, we have no certain way of relating such writing systems to other known forms of writing.

### 10.6.1 From drawing to writing

The origin and development of all early writing systems follow remarkably similar patterns.

First of all, drawing is the origin of writing all over the world. Most signs of early writing systems are **pictographic** in origin. The fact that the Egyptian hieroglyph for 'sun' and the earliest known Chinese character for 'sun' are similar does not mean that the two writing systems are related, only that both originate in attempts at drawing the sun.





*The Egyptian hieroglyph for 'sun'*

The decisive stage in the development comes when such drawings are stylized and become representations not of the sun itself, but of the word meaning 'sun':

pictorial sign → object

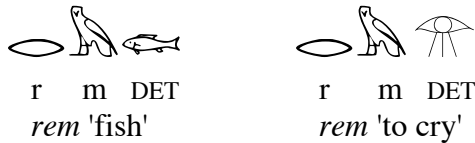
pictorial sign → word → object

At this point, drawing becomes writing. In other words, pictograms are transformed into logograms. The earliest writing systems are overwhelmingly **logographic**.

Second, with the transformation from drawings representing objects to written signs representing words, the emphasis shifts from imitation of the object depicted to differentiation from other written signs. It no longer matters that the Chinese character 日 'sun' is square instead of round, but it is very important that it is distinguished from 月 'moon'. Each sign is increasingly **standardized**, and so are the stroke types, stroke number and stroke order of the elements of which the sign consists. Furthermore, the **linear ordering** of signs is also fixed, usually to represent the ordering of words in a sentence.

Third, all early writing systems make use of the **rebus principle**, by which a pictographic sign originally representing one word is used to represent another word with the same or similar sound value. For instance, the Chinese character 足 'foot', which originally was a drawing of a foot, was used to represent a homophonous word meaning 'sufficient' (both are pronounced *zú* in modern Chinese). In a famous Sumerian example, a pictographic sign for 'reed' pronounced *gi* is also used for the word *gi* 'to reimburse'. In Egyptian, the sign for 'swallow' pronounced *wr* is also used for the word (or rather, the consonantal root) *wr* 'big'. In Maya, the sign for 'fish' pronounced *xoc* is also used for the word *xoc* 'to count'.

Fourth, all early writing systems make use of **determinatives** in order to cope with the increasing ambiguity resulting from the use of the rebus principle. Determinatives give a rough indication of the meaning of the sign, thus distinguishing it from other usages of the same sign. In Chinese writing, for instance, the character 止 was originally used for *zhǐ* 'toe', but was then extended to represent *zhǐ* 'to stop' as well, and in order to distinguish the two usages, the former usage was replaced by a complex character using 足 'foot' as a determinative: 趾. In Egyptian hieroglyphs, where the use of the rebus principle had given rise to signs indicating individual consonants, both the word *rem* 'fish' and *rem* 'cry' could be written with the ◯ (for the consonant *r*) followed by 𓄀 (for the consonant *m*), but in order to distinguish between the two, the former would add the determinative 𓄁 for fish, while the latter would add the determinative 𓄂 for an eye shedding tears:



Sumerian and Mesoamerican writing also made use of determinatives to disambiguate polyvalent signs.

Fifth, early writing systems tend to be **nuclear** in the sense that each sign represents a word in any of its inflected forms, leaving inflectional affixes unrepresented. Only with extensive use of the rebus principle does it become common to represent inflectional affixes by indicating their sound value.

### 10.6.2 From words to phonemes

Some scholars have proposed that the development of writing is **unidirectional**. By extensive use of the rebus principle, logographic writing systems evolved into **phonographic** writing systems. The earliest phonographic writing systems were either **syllabic** (each sign representing a syllable) or **consonantal** (each sign representing one or more consonants). Only with the ancient Greeks did we get fully **phonemic** writing systems (each sign representing one phoneme).

Historically, it is true that all writing has its origin in logographic systems, and that syllabic and consonantal writing developed before purely phonemic writing. On the other hand, many written cultures have not taken the full step to phonemic writing. Chinese writing is still largely logographic, Japanese kana writing is still largely syllabic, and Arabic and Hebrew writing is still largely consonantal. Some scholars see this as an effect of **cultural conservatism**, a resistance towards progress. Others maintain that different writing systems are suitable for different languages. The Chinese writing system, which we have characterized as morphosyllabic, is clearly designed for a language in which morphemes tend to be monosyllabic. Japanese kana is functional in a language with a limited number of syllables, though it would not have worked in a language like English. The Arabic and Hebrew systems may be well suited to languages where words are built on consonantal roots.

Cultural conservatism does play an important role in writing. Even languages with a phonemic writing system are seldom spelt the way they are spoken. In this sense, Modern English spelling is much less purely phonemic than the spelling of Old English. In some cases, such as the *k* of *knee*, *knight* and *knot*, sounds have disappeared in speech, but the letters used to represent them are still used in writing. In Japanese, what is conceived of as a high level of culture is often expressed in the frequent use of Chinese characters (logographic) instead of kana (syllabic), and it is even common to replace kana with Chinese characters when ancient texts are reprinted. To some extent, cultural conservatism may be functional, since it stimulates stability and historical continuity.

Some of the great innovations in the history of writing have come when a writing system has been transferred from one language to another. Foreigners are less likely to be hampered by cultural conservatism. More importantly, innovations are often the results of **adaptation** of the original writing system to the new language. For instance, when the Greeks adopted the Phoenician alphabet, which was consonantal, they needed to create vowel symbols. When the Japanese began to write their own language with Chinese characters, they needed to find a way to represent inflectional suffixes.



### 10.6.3 The emergence of writing in Sumer

The earliest known form of writing developed in Sumer in Mesopotamia, between the rivers Tigris and Euphrates in present-day Iraq, around 3200 BC. It is called **cuneiform** from Latin *cuneus* 'wedge' because it was written by creating wedge-shaped forms in wet clay.

As in other ancient forms of writing, most signs are pictographic in origin. In addition, some signs derive from the impression into wet clay of clay tokens stemming from various geometrical forms used within a very ancient Mesopotamian accounting system.

Sumerian writing was first used in the administration of the temple economy. Scribes went through training to write words correctly, indicating that the standardization of pictorial signs had begun, but the earliest texts are only itemized lists containing at most a number and a word referring to a concrete object. Gradually, however, we get continuous texts that more unequivocally represent natural Sumerian, a language of unknown affiliation which later disappeared as a spoken language. In the mid-second millennium BC, the direction of writing becomes fixed from left to right in horizontal lines.

Sumerian writing made use of around one thousand different signs. By employing the rebus principle, one sign could refer to a number of different words (or parts of words) with similar pronunciation, such as the sign for *ti* 'arrow' being used for *ti(l)* 'life'. In addition, one sign could also be used for different words with related meanings, such as the sign for *ka* 'mouth' being used for *inim* 'word' and *dug* 'to speak'. In order to determine which meaning was intended, one could add either semantic determinatives or phonetic indicators. The problem was that one and the same sign in principle could have four different functions: 1) as a logogram, 2) as a rebus sign, 3) as a semantic determinative, and 4) as a phonetic indicator.

The system became even more complex when it was employed to write other languages, such as the Semitic language Akkadian, which gradually replaced Sumerian after around 2800 BC. Now Sumerian signs were used to write Akkadian words, but since Akkadian also borrowed many Sumerian words, the same sign often represented both a Sumerian and an Akkadian word.

Cuneiform writing was in use for more than 3000 years, the last known text dating from 74 AD. In addition to Sumerian and Akkadian, it was used to write the Semitic languages Babylonian and Assyrian, the Indo-European languages Hittite and Persian as well as Elamite and Hurrian, both of unknown affiliation. Some of these, especially Persian, make little use of logograms, and basically use cuneiform as a syllabic or sometimes even phonemic writing system.







*One of the earliest examples of continuous writing, from Sumer around 3200 BC (from the Martin Schøyen collection in Norway).*



#### 10.6.4 The invention of writing in Egypt

Egyptian writing emerges quite suddenly as a fully developed system almost as early as Sumerian writing, around 3000 BC. Its invention may well have been inspired by the existence of writing in Sumer, but the two writing systems are fundamentally different.'

The shape of Egyptian hieroglyphs is highly pictorial and unlike most other scripts remains so throughout the more than 3500 years during which they are in use. In fact, however, most of the signs used represent sounds rather than words. From the very beginning, the Egyptian writing system is a mixture of the following three types:

1. logographic, such as ☉ for *r*' 'sun' and ✱ for *sb*' 'star'.
2. consonantal, each sign representing either one, two or three consonants, such as  for *m*,  for *nb* and ✱ for *sb*'.
3. determinative, such as  for 'man' and  for 'death'.



Many signs may be used in more than one function, such as ✱ for 'star' or for the consonant combination *sb*'.

Note that there are no signs for vowels, only for consonants and consonant combinations, reflecting the consonantal form of word roots in Egyptian and other Semitic languages. This is probably the origin of consonantal writing, as in Phoenician and, in modern times, Arabic and Hebrew.

The consonant signs build on the rebus principle. Signs representing two or three consonants usually correspond to a full word root, such as 𐤀 representing the word *hr* 'face' or the two consonants *hr*, or 𐤃 representing the word *sb* 'star' or the three consonants *sb*'. Signs representing one consonant build on the **acrophonic** principle, i.e. they represent the initial consonant of the word on which the pictorial sign is based, such as 𐤁 representing the consonant *d* because the sign depicts an open hand, and the word for hand is pronounced *dert* with an initial *d*. The acrophonic principle is also used in the first purely consonantal writing systems, the ancestors of Phoenician writing.

While hieroglyphs are flexible with regard to linear ordering, the two cursive scripts derived from hieroglyphs, the hieratic and the demotic scripts, are arranged horizontally from right to left. This is yet another point of similarity with Phoenician and Phoenician-derived writing.

### 10.6.5 The birth of the alphabet in Phoenicia and Greece



Although the phonographic elements increased with time in both cuneiform and Egyptian writing, both continued to use a number of logograms. Writing systems based entirely on phonographic principles seem to have evolved in the middle of the second millennium BC. The most famous and influential example is the Phoenician alphabet and its Canaanite predecessors.

The Phoenician writing system consists of 22 consonant signs, usually listed in a specific order:

Sign	Name	Sound value
𐤀	'āleph	'
𐤁	bēth	b
𐤂	gīmel	g
𐤃	dāleth	d
𐤄	hē	h
𐤅	wāw	w
𐤆	zayin	z
𐤇	hēth	ḥ
𐤈	tēth	t
𐤉	yodh	j
𐤊	kaph	k
𐤋	lāmedh	l
𐤌	mēm	m
𐤍	nūn	n
𐤎	sāmekh	s
𐤏	'ayin	'
𐤐	pē	p
𐤑	sādhē	s
𐤒	qōph	q
𐤓	rēš	r
𐤔	šin	š
𐤕	tāw	t

*The Phoenician alphabet*

It should not be difficult to recognize similarities with the Latin alphabet, which is indirectly derived from the Phoenician one.

Where the Phoenicians got their alphabet from is subject to controversy, but many think its consonantal nature points to Egyptian roots. The signs are pictographic in origin and are, like the monoconsonantal Egyptian hieroglyphs, based on the acrophonic principle, the sign  probably being derived from an earlier  (Proto-Sinaitic), depicting a bull's head, the Phoenician letter name 'āleph originally meaning 'bull, cow'. As already mentioned, the Phoenicians wrote horizontally from right to left, like the Egyptians and unlike the users of cuneiform writing.

When the Greeks learnt writing from the Phoenicians, they changed the system in one fundamental way: they created vowel symbols. The Phoenician letter 'āleph, which had represented a glottal plosive, came to represent [a], hē came to represent [e], yodh came to represent [i], and 'ayin came to represent [o]. Later long vowels were distinguished from short vowels, some new consonant signs were introduced, and diacritic signs were used to express accent and other features of pronunciation. The result was an almost perfectly phonemic writing system, the first in history.

At the beginning, the Greeks wrote from right to left like the Phoenicians, but both horizontal directions were possible, sometimes even alternating on every other line of the same text. By the 5<sup>th</sup> century BC, writing from left to right had become the standard, as it is in all Greek-derived scripts, including the various Latin and Cyrillic alphabets.

To the Phoenicians, the letter names functioned as a reminder of the pictographic origin of the letters, although the pictorial nature of the signs was no longer readily recognizable. To the Greeks, letter names like *alpha*, *beta*, *gamma*, *delta* did not carry any other meaning, and the last traces of pictographic writing were lost.

It has been argued that the Greeks were not the first creators of vowel symbols. As early as the 14<sup>th</sup> century BC, the inventors of the Ugaritic writing system (which was cuneiform in shape but Phoenician in underlying principles) had split the original 'āleph sign in three, purportedly representing the vowels /a/, /i/ and /u/. This, however, is hardly correct, since the three signs in question did not represent the vowels as such, but the consonant /ʔ/ (glottal plosive) plus one of these vowels. Vowels following any other consonant were still left unmarked.

To this day, other writing systems derived from the Phoenician alphabet or its close relatives continue to privilege consonants over vowels. In Hebrew and Arabic, for instance, no separate vowel signs exist, although vowels may be indicated either by an extended use of consonantal signs or by the use of diacritics. In Arabic, the three long vowels /a:/, /i:/ and /u:/ are written with the consonant signs ا /ʔ/, ي /j/ and و /w/, while the short vowels are either left unmarked or marked with diacritics. In Hebrew, similar strategies are available, but modern Hebrew is usually written without any indication of vowels at all.

### 10.6.6 The creation of Brahmi writing in India

The Brahmi writing of India, which first appeared around 500 BC, shows many points of resemblance with the Phoenician alphabet, as can be seen from the comparative table on the next page. During its first phase, Brahmi was also written horizontally

Phoenician		Brāhmī	
ʾ			a
b			b
g			g
d			d
h			h
w			v
z			ṇ
ḥ			gh
t			th
y			y
k			k
l			l
m			m
n			n
s			ṣ
ʿ			e
p			p
ṣ			c
q			ch
r			r
š			ś
t			t

are used to represent consonant clusters.

Instead of creating a phonemic writing system with equal representation for vowels and consonants, as the Greeks did, the creators of Brahmi made a semi-syllabic writing system, in which both the individual phoneme and the syllable are given prominence.

from right to left, though from the 3<sup>rd</sup> century BC, this changed, and all later Brahmi-derived writing systems are written from left to right. It seems highly probable that Brahmi writing derives from Phoenician or a related Semitic writing system.

As in Semitic writing, consonant signs form the core of the Brahmi writing system. In contrast to Semitic writing, however, vowels are also always given written representation. In Brahmi writing, vowels may be represented in one of the following three ways:

1. Word-initial vowels have their own signs.
2. Every consonant sign has an inherent *a*, so that, for instance, the sign for *k* is pronounced *ka* unless otherwise indicated.
3. Other non-initial vowels are indicated by adding diacritics to the consonant sign, so that, for instance, the *k(a)* sign with a diacritic for *u* is pronounced *ku*.

Two strategies are used to mark the absence of the inherent *a* in syllable-final consonants and clusters of two or three consonants:

1. The inherent *a* can be muted by adding a special diacritic.

*Correspondences between Phoenician and Brahmi signs. Note that Brahmi signs without corresponding Phoenician signs have not been included in the chart.*

2. Ligatures (signs formed by merging two or more signs into one) rather than sequences of individual consonant signs

Brahmi writing is the ancestor of a large number of widely different writing systems across South and South-East Asia, having spread as far as Mongolia in the North and the Philippines in the East. In addition to a large number of indigenous Indian languages, it is or has been used to write languages such as Tibetan, Burmese, Thai, Khmer, Javanese and Balinese.

### 10.6.7 The endurance of Chinese logographs

The first certain instances of written texts in China are the inscriptions on bone and tortoise shells used for divination purposes (so-called oracle bones) around 1200 BC. It is likely that the history of writing in China goes further back, but we do not have concrete evidence.

Chinese writing has several traits in common with other early forms of writing, such as Sumerian cuneiform and Egyptian hieroglyphs:

1. It is basically logographic, each sign representing a root morpheme.
2. It is pictographic in origin.
3. With time the pictorial nature of the signs often becomes unrecognizable.
4. The rebus principle is employed to extend the meaning of a sign to homophones.
5. Determinatives are added to distinguish between different meanings of a sign.

For instance, on oracle bones, the sign to the right was a pictorial representation of a foot (with three toes!), and the character was used to represent a root morpheme meaning 'foot'. Later, the shape of the character was stylized, and it was written 止, no longer betraying its pictorial origins. Furthermore, by making use of the rebus principle, this character was also used to represent a homophonous morpheme meaning 'stop'. Finally, the determinative 足 'foot' was added to the character when it was used for the 'foot' morpheme: 趾. (In Modern Chinese, the character means 'toe' rather than 'foot'.)



However, Chinese writing also differs from Sumerian cuneiform and Egyptian hieroglyphs in several respects:

1. While many Sumerian and Egyptian root morphemes are polysyllabic, the root morphemes represented by Chinese characters are always monosyllabic, and in the few cases where a root morpheme is polysyllabic, it is written with one character per syllable.
2. The rebus principle is less freely employed than in Sumerian and especially Egyptian. The character 止 cannot be used for just any homophone, only a few selected ones. Thus, the tight connection between sign and morpheme is maintained.
3. While Sumerian and Egyptian determinatives are used as individual signs in addition to the signs they modify, Chinese determinatives are used to create new complex characters, as when 足 and 止 are combined to create 趾.
4. While Sumerian and Egyptian were written horizontally, Chinese was usually written vertically (lines running from right to left) until influence from Western languages largely changed this in the course of the 20<sup>th</sup> century.

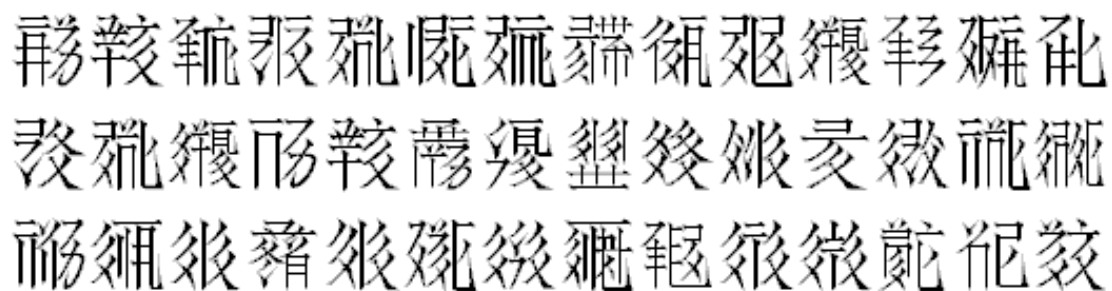
While both Sumerian and Egyptian writing disappeared almost two thousand years ago, Chinese writing is still in use and follows the same basic principles as more than 3000 years ago.

Chinese writing has spread to a number of neighbouring languages, most notably Japanese, Korean and (until the 20<sup>th</sup> century) Vietnamese. While the ancient ideal in all these countries was to write Classical Chinese rather than one's native language, the writing system was gradually adapted to represent the native language as well. This adaptation process has been most thorough in Japanese and has typically involved the following changes:

1. Because one Chinese character is often used to represent both a Chinese loan morpheme and an indigenous Japanese word, each character tends to have a number of different pronunciations. For instance, the Chinese character 人 'person' has two Chinese-derived pronunciations, *jin* and *nin*, as well as one indigenous pronunciation, *hito*. Each pronunciation has different usages and might well be seen as different morphemes.
2. Since Japanese root morphemes are often polysyllabic, the syllabic nature of the Chinese writing system is lost, as exemplified by the pronunciation of 人 as *hito*.
3. In some cases, new characters are created in order to represent indigenous concepts, usually by making new compound characters based on already existing ones, such as combining the determinative 亻 'person' with the character 動 'to move' to create the complex character 働 (pronounced *hataraku*) 'to work'.
4. Japanese also differs from Chinese in having a large number of verbal suffixes, such as *-ta* (indicating past tense). For this and other reasons, Japanese has developed its own phonographic writing systems hiragana and katakana, and the syllable *ta* is written with the hiragana sign た.

The result is a highly complex system, combining Chinese characters (with Chinese and Japanese readings) with Japanese hiragana and katakana and sometimes even Western loanwords written in the Latin alphabet.

Throughout history, other neighbouring languages have developed their own writing systems inspired by Chinese writing, such as the Khitan, Tangut and Jurchen writing systems developed between the 10<sup>th</sup> and the 12<sup>th</sup> centuries.



Examples of Tangut writing

*Examples of Tangut writing*

### 10.6.8 The rise and fall of Meso-American writing systems