141. Writing Systems

Bernard Comrie

1.   Defining the values

This map is concerned with the linguistic structure of different writing systems, in particular the unit of linguistic structure that is represented most directly in the writing system in question. Five basic systems can be represented, as given in 1–5 in the feature box; mixed systems are also possible, although only one, number 6 in the feature box, is included in the map.

@	1.	Alphabetic
@	2.	Consonantal
@	3.	Alphasyllabic
@	4.	Syllabic
@	5.	Logographic
@	6.	Mixed logographic-syllabic

In an **alphabetic** writing system, the basic unit of representation is the phoneme. Regular English spellings can be used to illustrate this type of writing system, e.g. *rat* /ræt/, *trim* /trɪm/, *spend* /spɛnd/. English can also be used to show that even in a writing system that is basically alphabetic, certain deviations from the strict principle of one-to-one correspondence of phoneme and grapheme may occur. Thus digraphs may represent a single phoneme, as with the initial two-letter sequence of *ship* /ʃɪp/, or more complex means may be used to represent certain phonemes, as when English uses the letter *e* after a consonant to indicate a different vowel quality of the vowel preceding that consonant, e.g. *fate* /feit/ versus *fat* /fæt/; these are regular phenomena in English spelling, even though they represent deviations from a strict alphabetic writing system. Primarily for historical reasons, English also has other,

less systematic deviations from the alphabetic principle, as in *rough* /rʌf/, *bough* /bau/. Some other languages with an alphabetic writing system, such as Finnish, have an almost complete correspondence between grapheme and phoneme. Similar deviations from the strict system are also found with the other types discussed below.

A **consonantal** writing system is a variant of an alphabetic writing system with the added feature that only consonants, not vowels, are represented. For instance, the Hebrew consonant sequence *p-h* in (1a) is a well-formed orthographic word; note that Hebrew is written from right to left and that the Latin transliteration follows this order. Indeed, depending on which vowels it is pronounced with, it can be several different words, in Israeli Hebrew pronunciation *po* 'here' or *pe* 'mouth'.

(1)   a.   פה   b.   פֹה   c.   פֶּה
              hp   hop   hep

In many consonantal scripts, it is possible to add diacritics to indicate vowels, as in the forms (1b–c), respectively *po* and *pe*.

It will be convenient next to take **syllabic** writing systems, or syllabaries. In an ideal version of such a writing system, there is a distinct grapheme for each syllable of the language in question. The Japanese hiragana syllabary comes close to this ideal, as in the examples in (2). Note in particular that there is no part of the graphemes that can be identified as denoting the consonant phonemes /h/ or /n/, nor is there any part that can be identified as denoting the vowel phonemes /i/ or /e/; the level of representation is directly that of the syllable.

(2)   ひ   へ   に   ね
      hi   he   ni   ne

An **alphasyllabic** writing system, or alphasyllabary, falls in between the alphabetic and syllabic types. In attested

alphasyllabic writing systems, the basic grapheme indicates a consonant, while a diacritic is added, or sometimes some other transformation applied, in order to indicate a combination of that consonant with a following vowel. The examples in (3) are from Thai, in which language the added diacritic can appear above, below, before, or after the consonant symbol, or some combination of these.

(3)　ม　　มิ　　มุ　　มา　　เม　　เมอ

　　　m　　mi　　mu　　maː　　meː　　məː

Note that, in contrast to the consonantal writing systems discussed above, in an alphasyllabic system the vowels must be indicated. (In practice, in alphasyllabic systems the consonant symbol without any diacritic usually has a so-called "inherent vowel"; in Thai, the leftmost example in (3) is often to be read /mo/.)

　　　In all the writing systems described so far, the linguistic unit that forms the basic representational unit of the writing system has been identifiable in phonetic terms. In a **logographic** writing system, the basic unit of representation is the morpheme, i.e. a grammatical-lexical unit. Chinese examples, with the Mandarin pronunciation indicated using the pinyin transcription, are given in (4).

(4)　人　　　　木　　　　湖　　　　壺

　　　rén　　　　mù　　　　hú　　　　hú

　　　'person'　'tree'　'lake'　'pot'

Since Chinese morphemes are nearly always a single syllable, one might wonder whether the Chinese writing system could not equally be described as syllabic. That this is not the case can be seen by comparing the written representations of the words for 'lake' and 'pot' in (4). They have identical phonetic representations, but are different morphemes with different

meanings, and therefore have different representations; in a syllabary they would receive the same representation.

Most writing systems currently in use in the world fall into one or another of these five types. The main exception is Japanese, which is a **mixed** system, incorporating both logographic and syllabic principles. Roughly speaking, lexical morphemes (e.g. nouns, verb roots) are represented logographically, while grammatical morphemes (suffixes, particles) are represented by means of the hiragana syllabary.

(5)　人　　が　　犬　　を　　　見た。
　　　hito　ga　　inu　o　　　mi-ta.
　　　man　NOM　dog　ACC　　see-PST
　　　'The man saw the dog.'

In (5), the nouns *hito* and *inu* and the verb root *mi-* are represented logographically, while the particles *ga* and *o* and the past tense suffix *-ta* are represented syllabically.

2.　Geographical distribution

Representing the geographical distribution of the world's writing systems on a single map necessarily involves some oversimplification, in particular where more than one writing system is used in a given area, or where a particular writing system is used only in a very small area. For more detailed information, reference may be made to Daniels and Bright (1996). Moreover, the discussion in this chapter relates exclusively to the linguistic classification set out in section 1, and disregards cultural links among writing systems. Thus, although most languages using Arabic script have a consonantal writing system, like that of Arabic itself, the versions of Arabic script used for writing Uyghur in northwestern China and for writing Kurdish represent all vowels, and are therefore considered alphabetic.

At present, the alphabetic writing system is clearly

dominant in the world as a whole, being used in almost all of Europe (including the Caucasus), northern Asia, Turkey and parts of Central Asia, insular Southeast Asia and the Pacific islands, sub-Saharan Africa, and the Americas, in addition to pockets of use elsewhere, such as Korea and Vietnam. It is therefore in general easier to specify where the other writing systems are used.

The consonantal writing system is used throughout the Arabic-speaking world, and also for a number of other languages traditionally associated with Islam, in particular Persian, Pashto, and Urdu and other languages of Pakistan; it is also used for Hebrew, and occasionally for Tuareg and some other Berber languages. The consonantal writing system, in particular in its Arabic script version, was earlier more widespread; it is still used occasionally for such languages as Malay, Swahili, and Hausa, although it is now obsolete for Turkish and other Turkic languages.

The alphasyllabic writing system is found in three distinct parts of the world. The first is the "Indosphere", comprising nearly all of India, Nepal, Tibet, Bhutan, Bangladesh, Sri Lanka, Myanmar, Thailand, Laos, and Cambodia. Related scripts are also in occasional use in insular Southeast Asia, from Sumatra through Java and Bali to South Sulawesi and parts of the Philippines. The second area is Ethiopia with an extension into Eritrea; in Ethiopia, the alphasyllabic Ethiopic script is used for most of the indigenous languages, including the national language Amharic. The third area is northern Canada, where an alphasyllabic script originally developed for the Algonquian languages Cree and Ojibwa is now also used for Athapaskan and Eskimo (Inuit) languages. The scripts of southern India and Ethiopia have numerous irregularities in the combination of consonant and vowel, thus moving them somewhat in the direction of a syllabary. The other alphasyllabic scripts of Asia and of northern Canada use almost exclusively regular combinations; those of northern Canada introduce the innovation of marking some vowels by

rotating the consonant symbol, as in example (6) from Inuktitut (an Eskimo language).

(6)　　Λ　　>　　<　　ᖠ　　ᑯ　　ᖄ
　　　　pi　　pu　　pa　　ki　　ku　　ka

Pure syllabic writing systems are rare in the modern world, even Japanese using its syllabaries only in conjunction with logograms. Two syllabaries developed in the first half of the nineteenth century are still in use, one for Cherokee, an Iroquoian language now spoken mainly in Oklahoma and North Carolina (USA), and one for Vai, a Mande language spoken in Liberia.

The Chinese script, used for all varieties of Chinese, is the only widespread purely logographic script in use today, although adaptations of Chinese script have been used for a number of neighboring languages, and one is still current for Yi, spoken in southern China and belonging to the Tibeto-Burman family.

The mixed logographic-syllabic system used by Japanese has already been described in section 1. Until recently, a similarly mixed logographic-alphabetic writing system was used by Korean. At present, only the alphabetic system is used in North Korea, and it is also dominant in South Korea.

3.   The ancient world

The alphabetic writing system first arose with the Greeks, at a date that remains controversial but was probably in the eighth or ninth centuries BCE. Alphasyllabic writing systems are even younger. In the ancient world, then, we are dealing primarily with consonantal, syllabic, and logographic writing systems, including combinations thereof. The precise definition of the "ancient world" is of course open to interpretation, and in any event covers a period measured in millennia, but for practical reasons I have here taken the end point in the Old World to be the invention of the Greek alphabet; because of its completely separate history,

Mesoamerican writing, though developing much later, is also included.

The oldest attested writing system is that of Sumerian, which arose in the late fourth millennium BCE; it is a mixed logographic–syllabic system. This Sumerian system was adopted and adapted directly or indirectly by a number of other languages of the region, most notably the Semitic language Akkadian. The Ancient Egyptian (hieroglyphic) script vies with Sumerian in terms of ancientness, and is a mixed logographic–consonantal script. In contrast to the general pattern among consonantal scripts, Egyptian script had a number of signs for sequences of two or three consonants (irrespective of intervening vowels), somewhat paralleling the use of the letter *x* to represent the two–phoneme consonant sequence /ks/ in Roman script.

The first purely consonantal writing systems arose in the Middle Bronze Age (first half of the second millennium BCE) among speakers of West Semitic languages, with an Egyptian model being the most probable main precursor. Among the earliest known purely syllabic scripts probably is Linear B, used for writing Mycenean Greek from the middle of the second millennium BCE.

Chinese script is logographic from its earliest clear attestations, around 1200 BCE, and has remained so ever since.

In Mesoamerica writing is attested from the second half of the first millennium BCE, with the Mayan script being by far the most widely attested and also the only one to have been substantially deciphered. The writing system is mixed logographic–syllabic.

A number of features going beyond the above classification into the main types of writing systems characterize Sumerian, Egyptian, and Mayan as early writing systems. One is the rebus principle, whereby the symbol for a particular concept can also be used for homonyms; this principle was used especially in the creation of signs for abstract concepts and in the creation of phonetic (syllabic, consonantal) signs, as when the Sumerian sign

for 'body', read *su*, is also used for the homonymous 'to replace'. The same principle is also found in the earliest Chinese writing, and is also used currently in transliterating foreign proper names into Chinese; it continues in modified form to play a role in that one character can be used as part of another character with the same sound, as in the examples in (7).

(7)   正   证
          zhèng   zhèng
          'straight' 'to prove'

Another principle is the use of so-called determinatives, namely signs that have no phonetic realization but serve to indicate the semantic field to which the word in question belongs. Finally, phonetic complements can be added to clarify part of the phonetic reading of a logographic sign or (in Egyptian) a two- or three-consonant sign. In the Akkadian example in (8a), a determinative 'woman' (the left-most sign) is written before the proper name 'Shibtu' (written with the three syllabic signs *ši-ib-tu*). A parallel is again found in Chinese: the left-hand side of the second character in (7) is the character for 'speech', thus indicating the semantic field of the overall character. Akkadian example (8b) consists of a logogram (the first sign) meaning 'god' and pronounced *ilum*, followed by the syllabic sign for *lum*; since the same logogram can also be read *šamû* 'sky', the phonetic complement here helps to ensure the correct interpretation.

(8)