Sumerian meanings and their pronunciations, the further burden of the native pronunciations of the Akkadian words equivalent to the Sumerian originals. An example is the Sumerian word *an*, which means

> 1. '*An*, god of the sky,' and in addition 'god' in general; (2. 'to be high'); 3. 'sky.' The Semites used this word 1. as ideographic sign for their god of the sky *Anu*, taken over from the Sumerians, and likewise for 'god' in general, Semitic *ilu;* and also for the concept 'sky,' Semitic *šamû;* 2. as phonetic (syllabic) sign for the syllable *an* (from the Sumerian) and the syllable *il* (from the Semitic) [Jensen 1969:94].[6]

The use by the Akkadians of one and the same symbol to represent the Sumerian syllable *an* and the Akkadian words *Anu, ilu, šamû,* and *il* is paralleled in other writing systems which adapted alien scripts to their own use. The cuneiform script adapted by the Hittites from Sumero-Akkadian exhibits a similar complex mixture (Moorhouse 1953:65–67). The same is true of the hybrid system developed by the Japanese on the basis of Chinese. The complexity of these hybrid systems will be illustrated in further detail in connection with the discussion of Japanese below.

For all its fiendish complexity, the Akkadian adaptation of the Sumerian system was also able to function as a medium of full communication, since it too had the requisite phonetic base at its disposal. Scholars differ in their assessment of how much use was actually made of the phonetic component over time and from one literary genre to another (Cohen 1958:90; Oppenheim 1977:238–239; Civil 1973:26–27; Wilhelm 1983:159; Kramer 1986:119–120). Their comments on this matter are generally limited to rather summary statements which fall considerably short of what is needed for a clear picture—within a particular writing system and between different scripts—of the relative use of the phonetic and nonphonetic elements that my "Duality Principle" notes are inherent in all full systems of writing.

Some samplings of cuneiform writing suggest that the phonetic component was much higher in the Akkadian system than in that of the Sumerians: 70–96 percent in the former as against 36–54 in the latter (Civil 1973:26). Also it appears that the high figure of 1,200–2,000 signs of all kinds in early Sumerian writing was reduced to some 500–700 in that of the Akkadians, of which 300–350 were used to represent syllabic sounds (Cohen 1958:90; Jensen 1969:90; Driver 1976:67).

These figures, and the preceding discussion, give some suggestion of the flexibility inherent in the cuneiform system. The symbols could be and were adapted to a great variety of languages. In one form or another they served as the main vehicle of written communication in the Near East throughout their long history, which began about 3100 B.C. and ended in the first century A.D.

Chinese

In contrast to Sumerian writing, whose history can be traced to its very beginning, the origins of Chinese writing are obscure and much debated. The earliest form known to us dates from the Shang dynasty (1200–1045 B.C.). By that time the script was already a highly developed system based on principles which have continued to characterize the system to the present.

A major point of contention is, How did the idea of writing come into the minds of the Chinese? There are two opposing views of the matter based on different explanations for the emergence of civilized societies and the existence of specific elements of culture, including writing. One approach stresses "independent invention," the other "stimulus diffusion."

Chinese scholars tend to espouse the first approach. In support of this view a number of writers have attempted to push Chinese writing back before Sumerian by claiming a connection between some symbols of Shang date and those inscribed on pottery several millennia earlier (Chang 1983; Cheung 1983). However, as noted in the previous criticism of scholars who see a connection between Easter Island symbols and those of the Indus Valley, it is methodologically unacceptable to advance as evidence miscellaneous instances of similarity among the necessarily limited number of scratchings that can be made using only two or three lines. There is involved here is a sort of chauvinistic scholarship that seeks to prove an independent invention for Chinese writing by methodologically suspect means.

Exponents of the "stimulus diffusion" approach sometimes use the same dubious methodology to claim that some of the symbols found in China have been copied from earlier ones found in the Near East. Most diffusionists advance a broader and more general argument based on the amount and quality of correspondence between civilizations. They argue that while some simple aspects of culture, such as stone knives and drawings of familiar things, may have been independently invent-

ed by various peoples, more complicated ones like writing must have had a single origin. In support of this thesis, diffusionists cite evidence of borrowing of some other specific cultural items to prove the borrowing of the idea of writing. Thus the prominent Near Eastern scholar Cyrus H. Gordon states that "China heard about casting bronze from the West; and what impelled China to invent her own system of writing was diffusion of the idea from the Near East" (Gordon 1971:16).

But here too the methodology is open to question. The borrowing of one item—if it really is a case of borrowing—does not necessarily prove the borrowing of another, though to be sure evidence of extensive borrowing is suggestive. Conversely, when instances of alleged influence are shown to be based on dubious scholarship, this saps confidence in the whole approach. Gordon's overall case for stimulus diffusion is hardly helped by his advancing the claim that "Chinese in pigtails" (actually found only in the Manchu period, from 1644 to 1911) "are portrayed unmistakably in the art in pre-Columbian Middle America" (Gordon 1971:171).

My own view of the matter is that the arguments for both approaches are seriously flawed and that at present there is simply not enough evidence to provide convincing proof for either claim about the origin of Shang writing. However, while keeping an open mind on the matter, I feel that the burden of proof rests with the diffusionists. I therefore incline toward the belief that the Chinese independently invented writing, not because there is any proof for this, but chiefly because I believe that human beings are sufficiently inventive to have come up with the same idea more than just once.

Moreover, it seems to me that while some of the principles underlying Chinese writing are in fact similar to those underlying Sumerian writing, in all probability the reason for this is not that one was influenced by the other. The distances in time and space, unlike the Sumerian–Egyptian and Phoenician–Greek situations (discussed in chapters 4 and 5, respectively), militate against such a hypothesis. A more reasonable explanation is that the two peoples independently thought up somewhat similar solutions to somewhat similar problems. It is underlying principles, not the superficial outward form of symbols, that should occupy most of our attention.

In approaching an analysis of the Chinese script we encounter quite a different problem from that of Sumerian. In the case of Sumerian, the language represented by the writing was completely unknown and

had to be reconstructed from scratch, in part with the help of information provided by peoples such as the Babylonians and Assyrians, who took over both the language and the writing system, applied the latter to their own language, and developed bilingual texts of various kinds. In the case of the Chinese, there is continuity, with gaps that can at least partially be filled in, between the earliest extant writing and that of the present day. And given the considerable amount of archaeological work going on in China and the exciting finds that have been made in recent years, there is hope of being able to fill in more gaps. Chief among these gaps is the huge void that must have contained an earlier stage of writing before the full-grown system emerged on the scene during the Shang dynasty.

Our knowledge of what is so far the earliest known Chinese writing is less than a hundred years old. Toward the end of the last century peasants working their fields in the district of Anyang, located in the northern part of Honan province a bit north of the Yellow River, turned up fragments of bone, some of which bore markings that Chinese scholars recognized as characters of an older form than any yet known. Owing to the turmoil attending the collapse of the imperial regime in 1911, it was several decades before scientific excavations could be conducted in the area. In the meantime large numbers of inscribed bone fragments found their way into the hands of scattered Chinese and foreign scholars. Scientific excavations began in 1928, were interrupted by the Sino-Japanese war of 1937–1945, and were resumed on a larger scale after the establishment of the People's Republic of China in 1949.

The result of all this activity has been the unearthing of over 100,000 inscriptions, many of them of a fragmentary nature, on pottery, stone, bronze, and, most important of all, bones and shells. The earliest date from about 1200 B.C. This was during the hitherto somewhat shadowy Shang dynasty that was overthrown in the middle of the eleventh century B.C. by conquerors who established the long-lived Zhou dynasty (1045–221 B.C.). The Zhou dynasty, in its middle course, produced the flourishing literate culture of which Confucius is an outstanding exemplar.

The extension of our knowledge of Chinese civilization back from Zhou to Shang owes much to the inscriptions found on bones and shells. Inscriptions on other materials tell us little, since they are in general restricted to a few characters. Thus the second most important group, those on bronze vessels, consist in part of so-called "clan-

name" inscriptions. These are pictographs often encased in a sort of rectangular cartouche that is reminiscent of those found in Egyptian hieroglyphic inscriptions containing the names of royal personages. Some bronze vessels have slightly more extended inscriptions specifying who made the vessel and for whom it was made (Boltz 1986).

In contrast, the inscriptions on bones and shells are much more informative. This is due only in part to their somewhat greater length, for most of the texts are less than 15 characters in length, and very few exceed 50 characters (Mickel 1986:256). The nature of the inscriptions and how they came to be written provide most of the information.

The Shang people wrote on bones and shells for purposes of divination. Through their priests, who were scribes and had the power to communicate with their ancestors and their gods, they sought oracular advice on all sorts of matters, from the most serious affairs of state to lesser problems such as what to do about toothaches and how to interpret dreams. In this the Shang were like people today who run their lives by consulting astrologers, ouija boards, or other sources of spiritualistic "guidance." The Shang divination texts are known as "oracle bone inscriptions" (OBI).

The texts were written on the large shoulder bones of cattle and the shells of turtles. These were much prized. Some nondivination texts record the receipt of these valuable items. However, unlike the extensive accountancy of the Sumerian temple scribes, there are only a few Shang inscriptions of this sort.

The bones and shells, especially the latter, were prepared by being smoothed to a high polish and were stored until brought out for the divination ceremony. The diviner asked a question, such as "Shall an army of five thousand men be raised?" or "Will it rain?" Then heat was applied to the back of the bone or shell, causing cracks to appear on the face. The cracks were interpreted by the diviner, or perhaps even by the king himself. Finally a record of the whole matter, including the follow-up on the final outcome, was incised, or sometimes written with a brush, on the bone or shell.

It is of interest to note that at this early stage the characters were not yet standardized as to shape and size. It was only later that they came to occupy the uniform square space that has earned them the Chinese name of *fāngkuàizì* 'square block characters,' which one Western scholar has rendered as 'tetragraphs' (Mair 1988). Moroever, in the early inscriptions there was no fixed direction of writing. The sequence of characters varied from left to right, right to left, top to

bottom, and various mixtures of direction within the same inscription and even within the same sentence (Serruys 1974:16). When we add to this the fragmentary nature of many of the bones and shells, it goes without saying that the interpretation of the inscriptions is a highly complicated and controversial matter.

A full inscription consists of four parts: (1) a preface stating that on such-and-such a day a bone or shell was cracked and that So-and-so divined; (2) a "charge" asking a question; (3) the prognostication; and (4) the verification, telling what actually happened. Few inscriptions contain all four parts. The dating of the inscriptions is also generally limited to the specification of a day within a chronological system containing a cycle of sixty days. From this information, the naming of the diviner, and other clues, it is sometimes possible to achieve a more precise dating, as in the case of some solar-eclipse inscriptions during the reign of King Wu Ding (1200–1181 B.C.) (Keightley 1978:174).

A full oracle bone inscription dating from the same period is of particular interest because it reveals the existence, as early as the second millennium B.C., of concerns that are still very much in evidence in Chinese society. The inscription deals with the pregnancy of King Wu Ding's consort, Wife Hao. The preface states that on day 21 a shell was cracked and a certain Ge divined. The charge notes that Wife Hao was to give birth and asks: "Will it be a happy event?" The prognostication was undertaken by the king himself. Reading the cracks, he said that if the birth occurred on one of certain specified days, it would be a happy event; if it happened on any of some other specified days, it would be "hugely auspicious"; but if it occurred on another specified day, it would not be good. Three weeks and one day later came the verification telling what had actually happened: "The birthing was not a happy event. It was a girl" (Keightley 1978:41; Mickel 1986:255).

Shang divination practices of the sort just described are attested for a period of only about a century and a half, that is until the overthrow of the dynasty in 1045 B.C. The Zhou successors did not take over the Shang practice of divining by the use of bones and shells, but they did continue the practice of inscribing on bronze. The earliest identifiable Zhou text is on a bronze vessel with an inscription stating that it was commissioned only eight days after the Zhou victory over the Shang. It is typical of a large number of vessels cast during the early Zhou period that contained texts stating the date, discussing the meritorious deed which led to the casting of the vessel, noting the generosity of

the ruler as represented by the gift of metal, and expressing the desire that future generations take note and use the vessel. Some of the inscriptions are quite long. One that can be precisely dated as incised in the reign of a Zhou king who ruled from 946 to 935 B.C. contains 284 characters. The practice of incising inscriptions on bronze vessels continued for close to a thousand years, until the early Han period (206 B.C.–9 A.D.) (Mickel 1986:295).

Important as these inscriptions were, they were overshadowed by Zhou texts written with a brush on bamboo, wood, and silk, and carved on stone and other hard substances. These materials were the basis of the extensive literary remains that survive from before the invention of paper in the second century A.D. Following that there was a considerable increase in output, particularly after the invention first of block-printing about 600 A.D. and then, well before Gutenberg, of printing from movable type. This resulted in a veritable explosion of publications. The total Chinese output prior to the nineteenth century may have exceeded that of the rest of the world combined.

These developments occurred over a long period, more than three thousand years, and in a huge area populated by people who commanded various forms of speech that have conventionally been grouped together as "Chinese." This term is an umbrella designation for at least eight present-day varieties of what are usually called "dialects" but, since they are mutually unintelligible, might better be considered parallel to the various languages that make up the Romance group of languages. The main varieties are Mandarin (750 million speakers), Wu or Shanghainese (85 million speakers), and Cantonese (50 million speakers). Even these terms are imprecise, as emphasized in a study entitled *The Four Languages of "Mandarin"* which notes that "Mandarin" is itself an umbrella term for "Idealized Mandarin" (*Putonghua* 'Common Speech' or *Guoyu* 'National Language'), "Imperial Mandarin" (the largely uncodified language spoken by the scholar-official class in imperial China), "Geographical Mandarin" (the invention of twentieth-century linguists, created in an effort to delineate the language of a particular area sharing certain common phonological traits), and "Local Mandarin" (every locale, because of its unique linguistic composition, treated as an independent speech community) (Sanders 1987).

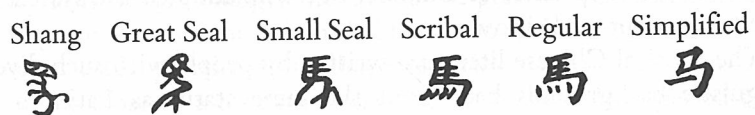In earlier times most literature was produced in a style loosely called "classical Chinese" that was written by people who spoke some form of language ancestral to the varieties of current "Chinese" noted above. Such literature was also produced by many people who were not even native speakers of Chinese, but who adopted the system for want of a script of their own.

The classical Chinese literature written by people with such diverse linguistic backgrounds had about the same status as Latin in the Romance situation. In modern times the primary emphasis in writing has been on a style based more or less on speech, chiefly Mandarin but to some extent also other varieties, notably Shanghainese and Cantonese. Just as these varieties of speech are mutually unintelligible, so are the varieties of written language based on them. It is a widespread myth that Chinese characters cut across boundaries of speech. This is no more true than the claim of universality for the Latin alphabet. It is necessary to learn the concrete application of the symbols to each specific form of speech. Since the version of Mandarin that is officially called Putonghua or Common Speech is the standard language in China, the writing system based on it is also the written standard. But it takes more effort for a Cantonese to learn to read and write in the national standard than it does for a Spaniard to learn to read and write French (DeFrancis 1984a).

The realities of speech and writing in China that are obscured by the myth of the universality of Chinese characters and by the ambiguity of the umbrella term "Chinese" should be kept in mind in considering the evolution of Chinese writing. There has been continuity as well as change in this long period. Characters have been created by speakers of many different varieties of Chinese and often reflect the peculiarities of their speech. There have been historical changes in word order, and the characters have been shuffled around to adapt to these changes. Characters have died out as the words that they represented have become obsolete. New characters and new combinations of characters have been created to express new words that have entered into the vocabulary. In common with all writing systems with a long history, the pronunciation of the symbols has changed greatly over time, as is readily apparent in a recently published dictionary of Early Zhou Chinese (Schuessler 1987).

The forms or shapes of the characters have also undergone change. On the oracle bones many characters are clear pictographs. Their modern descendants have been so stylized, abbreviated, distorted, and otherwise modified that it is often difficult to see the relationship with the ancestral character, as can be illustrated by the following example

(from DeFrancis 1984a:83) showing the loss of iconicity in the evolution of the character for *mǎ* 'horse':

Shang   Great Seal   Small Seal   Scribal   Regular   Simplified

象   馬   馬   馬   馬   马

The styles illustrated are those of the Shang oracle bones, the Great Seal style of the Zhou dynasty, the Small Seal style of the short-lived Qin dynasty (221–206 B.C.), and the Scribal and Regular styles of the Han dynasty (206 B.C.–220 A.D.), the last being the most commonly used script until the official simplification of characters in the PRC in the 1950s. Figure 17 compares the OBI and modern forms of a number of characters.

As the examples presented in the illustration suggest, the task of deciphering Shang characters, of which some 1,000 out of a total of 4,500 have been identified to date, is facilitated by the fact that in many cases it is possible to trace the evolution of a character from its Shang to its contemporary form. There are many instances, however, of "descendantless" graphs. These constitute a problem in decipherment. Some descendantless characters appear to be completely unrelated to any modern symbols. Others are complex characters whose component elements have modern counterparts but are not now combined in the same way as in the Shang graphs.

This mention of the component elements in Chinese characters and how they are combined raises the crucial question of the principles behind the Chinese system of writing. To obtain a clear understanding of this system it is essential that we recognize what the principles are and—a much neglected aspect—how they were applied in different proportions over time.

It has been the traditional practice to classify Chinese characters into several groups. The first group, about which there is general agreement, consists of pictographs. As we have already seen, many of the earliest characters, as represented on the OBIs, are clearly pictographic in nature.

What might be called the "simple indicative principle" identifies a second group of characters which represent words not exactly pictorially but in some other representational manner. For example, the words for "one," "two," and "three" are respectively represented by one, two, and three horizontal lines:

OBI graph	modern character	modern reading	meaning
1.	象	xiàng	'elephant'
2.	求	qiú	'pelt' (mod. 裘)
3.	口	kǒu	'opening, orifice, mouth'
4.	目	mù	'eye'
5.	月	yuè	'moon, month'
6.	田	tián	'(cultivated) field'
7.	女	nǚ	'(kneeling) woman'
8.	其	qí	'(winnowing) basket' (mod. 箕)
9.	天	tiān	'overhead' > 'sky, heaven'
10.	羊	yáng	'sheep, ram'
11.	馬	mǎ	'horse'
12.	龜	quī	'turtle'
13.	魚	yú	'fish'
14.	鼎	dǐng	'tripod, cauldron'

Figure 17. Chinese Writing: A 1% Pictographic Script
The evolution of fourteen Chinese characters representative of the mere one percent that go back to pictographs, chiefly those found on oracle bone inscriptions (OBI) of around 1200–1045 B.C. All but two—12 and 14, the most complicated—have been used as phonetic symbols in the formation of multielement characters. Adapted with permission from William G. Boltz, "Early Chinese Writing." *World Archaeology* 17 (3) (1986):427.

	1	2	3
Shang	一	二	三
Modern	一	二	三

The words for "above" and "below" are represented by a horizontal line with another graphic element placed above or below it:

above   below

Shang   二   二

Modern   上   下

What might be called the "compound indicative principle" involves a somewhat more elaborate version of the preceding. It is frequently illustrated by the combination of the characters for "sun" and "moon" to represent the word *míng* 'bright,' as follows:

sun   moon   bright

日   月   明

The fourth principle is the familiar rebus device which we first encountered in Sumerian with the use of *gi* 'reed' borrowed to represent the homophonous word *gi* 'reimburse.' A Chinese example is the borrowing of a character representing *xiàng* 'elephant' to represent the homophonous word *xiàng* 'image.'

The final principle is one which combines a rebus-like symbol with another symbol giving, generally, a semantic clue to the meaning. One example is the addition (on the left) of the symbol for 'person' to that (on the right) for *xiàng* 'elephant' to produce an unambiguous character *xiàng* 'image':

亻 + 象 = 像

Another example is the previously cited character for *mā* 'mother' formed by combining the character for *mǎ* 'horse' (on the right) with another for 'female' (on the left):

女 + 馬 = 媽

I call this the SP principle, since it involves joining a semantic element S—such as the symbols for 'person' and 'female' in the previous examples—to a phonetic element P—such as the symbols for *xiàng* 'elephant' and *mǎ* 'horse.' I also designate as SP characters those, to be discussed below, that are formed by joining a phonetic element P to a semantic element S. We can think of both of these types as MS (meaning-plus-sound) characters.

The proportions in which the foregoing principles have been applied in the formation of Chinese characters have varied over time. This can be seen in the following table (adapted from DeFrancis 1984a:84) summarizing the structural classification of 977 Shang characters, 9,353 characters of a second-century dictionary by Xu Shen, and 48,641 characters of the great imperial Kang Xi dictionary of the eighteenth century:

Structural Classification of Characters

Principle	Shang Dynasty	2nd century	18th century
Pictographic	227 (23%)	364 (4%)	
Simple indicative	20 (2%)	125 (1%)	± 1,500 (3%)
Compound indicative	396 (41%)	1,167 (13%)	
Semantic-phonetic	334 (34%)	7,697 (82%)	47,141 (97%)
Total	977	9,353	48,641

The traditional view of Chinese characters summarized above has been challenged in some important respects by Peter A. Boodberg, a leading student of early Chinese writing whose views are receiving more and more acceptance since they were propounded several decades ago (Boodberg 1937, 1940, 1957). His criticisms center on two main points. The first is the failure of many people, sinologists included, to realize the importance of SP characters because of neglecting or minimizing the phonetic contribution of the P element and exaggerating the semantic contribution of the the S element. The second extends Boodberg's criticism of underestimating the significance of the phonetic element by dealing especially with those characters traditionally classified under the "compound indicative principle."

The majority of these characters, Boodberg contends, are in reality SP characters. Indeed, he goes so far as to claim that apart from "a few exceptional cases" there is simply no such thing as a class of characters constructed on semantic principles (1937:345–347). This view has recently been reiterated by another scholar, William G. Boltz, who has also done significant work on early Chinese writing. He asserts:

"Characters were not invented by just putting together two or more elements based on their semantic values alone. At least of one of the components must have had a phonetic function" (Boltz 1986:428).

Boodberg's "few exceptional cases" include chiefly single characters of clearly pictographic origin. There are at most only a few hundred of these, and the number has not increased for some two millennia. These simple characters of pictographic origin, examples of which appear in figure 17, comprise only about one percent of the total number of Chinese characters. The remaining 99 percent, examples of which are presented in figure 18, are compound characters whose main component is a phonetic element.

As an example of the need to rethink characters allegedly based on semantic principles, Boodberg cites the case of the previously mentioned character for *míng* 'bright.' He rejects the traditional approach which begins with a disembodied concept supposedly represented by a character formed by combining the symbols for "sun" and "moon." Instead he starts by assuming definite spoken words related to the meaning "bright." This leads him to note the existence of an earlier form of the character for *míng* 'bright' (Morohashi 1955–1960, 5:14, 366) that I present below in juxtaposition with the later version:

朙　明

In both cases the element 'moon' on the right-hand side of the characters is a semantic determinative. The element on the left-hand side of the first character is originally a picture of a window, with a pronunciation related to *míng*. In short, the present character representing *míng* 'bright' is simply a later variant with what is usually taken as a semantic "sun"—which has caused us to overlook an earlier version with a phonetic *míng* element that more closely relates the character to a spoken word (Boodberg 1937:344–345; 1940:270–274).

Boodberg also cites the case of the following two characters:

目　見

In modern transcription, the first character is *mù* 'eye,' the second *jiàn* 'see.' But the first was also used to write the related word 'see.' Hence it represented two related meanings and had two quite different pronunciations, which have become modern *mù* 'eye' and *jiàn* 'see.' In order to distinguish the two meanings, the 'eye' character was supplemented with a phonetic determinative, the bottom part of the second

character, whose earlier pronunciation *ńźiĕn* provided a better phonetic clue than the modern pronunciation *rén* (Boodberg 1937:343).

In presenting these and other cases Boodberg stresses an aspect of Chinese writing that we have already encountered in Sumerian. That is the fact that many sounds are represented by more than one symbol (recall the 23 for Sumerian *du*) and that the same symbol may represent several different words (recall the different words represented by the pronunciations *gub, gin, túm* for the same symbol with the basic meaning 'leg'). Chinese words are also often written with different characters, and the same character may be read in several different ways. It is a major challenge to modern scholarship to unravel the interconnections that have grown up among Chinese characters in the several millennia that they have been handled and mishandled by millions of scholars with widely different backgrounds in the many varieties of spoken and written Chinese.

Because of his emphasis on relating writing to speech, Boodberg presents a clearer analysis of the evolution of Chinese writing than that suggested by the conventional listing of its underlying principles. It is summarized in even simpler terms by Boltz as a three-stage development: (1) a pictographic stage, which in its pure form could write only the limited part of the language that was clearly picturable; (2) a multivalent stage, which included the use of the rebus principle whereby the same symbol might stand for unrelated homophonous words, and the use of the same character to represent words semantically related but with different pronunciations; and (3) a stage in which the ambiguity that grew up with the multivalent use of characters was resolved by resort to semantic and phonetic determinatives, as in the case of Near Eastern writing (Boltz 1986). The following examples illustrate these stages:

Stage	xiàng 'elephant'	A xiàng 'image'	mù 'eye'	B jiàn 'see'
Pictographic	豸/象	(none)	𦓐/目	(none)
Rebus	象	象	目	目
Determinative (A. Semantic) (B. Phonetic)	象	像	目	見

	75	158	255	391
Syllabary	皇 huáng	辟 pì	馬 mǎ	堯 yáo
Derivatives	喤 huáng	僻 pì	瑪 mǎ	嶤 yáo
	徨 huáng	譬 pì	碼 mǎ	蟯 yáo
	惶 huáng	闢 pì	螞 mǎ	僥 jiāo
	湟 huáng	壁 pì	鎷 mǎ	澆 jiāo
	煌 huáng	避 pì	媽 mā	翹 qiáo
	鰉 huáng	壁 pì	榪 mà	磽 qiáo
	蝗 huáng	擘 pì	禡 mà	燒 qiāo
	鍠 huáng	劈 pì	罵 mà	嘵 xiāo
	隍 huáng	癖 pǐ	獁 mà	驍 xiāo
	遑 huáng	臂 pèi	嗎 ma	曉 xiāo
	篁 huáng	孼 pò		燒 shāo
	凰 huáng	蘗 niè		譊 náo
	堭 huáng			橈 náo
	艎 huáng			鐃 náo
				撓 náo
				髐 nào
				嬈 ráo
				蕘 ráo
				蟯 ráo
				饒 ráo
				繞 rào
				遶 rào

453	511	635	784	822
莫 mò	利 lì	門 mén	中 zhōng	黃 huáng
漠 mò	俐 lì	鍆 mén	鍾 zhōng	潢 huáng
膜 mò	唎 lì	捫 mēn	仲 zhōng	璜 huáng
寞 mò	浰 lì	悶 mèn	妕 zhōng	磺 huáng
摹 mò	痢 lì	們 men	忠 zhōng	蟥 huáng
鄚 mò	筣 lì	閩 mín	盅 zhōng	簧 huáng
嫫 mò	莉 lì	閱 wén	衷 zhōng	蟥 huáng
鏌 mò	鴷 lì	聞 wén	神 zhōng	黌 huáng
煤 mò	鷅 lì	問 wèn	蚛 zhòng	廣 huáng
瘼 mò	悧 lì	閃 shǎn	鴐 zhòng	廣 kuǎng
驀 mò	梨 lí	閑 xián	冲 zhòng	橫 héng
貘 mò	鬁 lì	閂 shuān	忡 chōng	
鬤 mò	蜊 lí	關 guān	沖 chōng	
摸 mō	犁 lí		宁 chōng	
謨 mó	猁 lì		笫 chōng	
模 mó	莉 lǐ		翀 chōng	
糢 mó	矖 lé		茽 chōng	
暮 mú	劙 lé		种 chóng	
墓 mù	猁 lè		岇 nōng	
慕 mù	詈 lián		馽 zhí	
募 mù	鋝 qián			
幕 mù				
蟆 má				

Figure 18. Chinese Writing: A 100% Syllabic Script
Examples of Chinese characters, which always represent syllables, showing their derivation (except for the 1% noted in Figure 17) from two elements—a primary phonetic element (i.e., one of the 895 syllabic elements in the "Soothill Syllabary"), and an added semantic element (i.e., one of the 214 elements traditionally called radicals or keys).

Note that in column A a semantic determinative, a variant of the symbol for 'person,' is added to the phonetic base *xiàng* to distinguish the meaning 'image.' In column B a phonetic determinative, another variant of the symbol for 'person,' is added to the semantic base *mù* 'eye' to distinguish the meaning 'see.' Both types are SP characters. Note also that the three stages should not be viewed as chronologically distinct. They define the stage of an individual character as determined by its function.

It is useful now to take a closer look at the preponderant category of characters of the determinative stage. Particularly illuminating is a comparison of Chinese and Sumerian in their approaches to an essentially similar problem of coping with the ambiguity inherent in writing systems in which one word might be written many different ways and one graph might be read many different ways. We can schematize the Sumerian and Chinese approaches as follows:

Sumerian   S   P?   P   S?

Chinese   SP?   PS?

The Sumerian examples have already been discussed in the preceding section. It was explained there that an ambiguous phonetic symbol (P?) is disambiguated by adding a semantic determinative (S), and an ambiguous semantic symbol (S?) is disambiguated by adding a phonetic determinative (P). Chinese does exactly the same thing, but it developed a variation that has had a profound influence on how the characters have been viewed. That variation is to weld the determinative, whether semantic or phonetic, with the ambiguous element to form a tightly knit symbol that is rigidly confined within its own square space of exactly the same size as that for every other character, regardless of simplicity or complexity.

With his usual perspicuity, Boodberg notes this important difference between Chinese and Sumerian (and Egyptian also) in the following passage:

Egyptian and cuneiform, where the use of semantic determinatives remained optional and the determinatives themselves detachable from the graphs they determined, moved on apace toward phonetization. In Chinese, the determinatives, semantic or phonetic, were welded securely to their graphs so as to form one single graphic body; diagram-

matic structure became thus the dominant type of character building. This may have been caused by a more pronounced homophony of the Chinese vocabulary, but it must have also been influenced by an aesthetic imperative in the Chinese which prompted them, apparently quite early in the development of the script, to enforce the principle of EQUIDIMENSIONALISM. . . . of the graphs [Boodberg 1957:115].

With respect to the two kinds of SP characters, namely those formed by adding an S to a P or a P to an S, which kind is more important? On this there has been considerable disagreement among specialists in Chinese. Earlier scholars, and a few still today, consider that the complex characters of this category were formed chiefly by adding a phonetic determinative to a semantic base. The popular names given to the semantic element reflect this view. It is frequently referred to as a "radical," sometimes as a "key," the latter being used especially in connection with its function as the unit (comparable to our *abc . . . xyz*) for filing characters in a dictionary. The semantic element is considerably less often called a "signific" or "determinative."

Noel Barnard, who has done some of the most important research on this aspect of Chinese, is firmly of the opinion that the phonetic element is the real core of compound characters. For the most part semantic elements were added to phonetic elements, not the other way around (Barnard 1978). This is the prevailing view among most specialists today. I hold strongly to this opinion also.

Leaving aside the matter of priority, it should be noted that the result of adding an S to a P is essentially the same as that of adding a P to an S. That is to say, PS = SP. The order in which the two elements merged is now largely of only historical interest. And the location within a character (e.g., left side versus right side) is also of secondary importance.

If almost all characters are of the SP variety, and if most of these were formed by adding a semantic to a phonetic, then we need to take a closer look at just how the combination was effected. Part of the task is easy. It has been the tradition, as illustrated in the Kang Xi Dictionary, to identify exactly 214 key semantic elements. Until the PRC simplification of the 1950s, all characters were analyzed, sometimes quite arbitrarily, as having one of these 214 keys, and they were listed in dictionaries under the appropriate key. However, there is a good deal of artificiality and arbitrariness in all this, as is indicated by the fact that the first Chinese dictionary, of the second century B.C., listed characters under 540 keys, while the most recent PRC dictionaries

have variously classified them under 186, 191, 225, 226, and 250 keys.

What of the phonetic elements? The Chinese have in general paid much less attention to this aspect, though some philologists have compiled rhyming dictionaries based on the sounds of the characters. Some scholars, including some foreign pathbreakers like Bernhard Karlgren, have made good use of the phonetic elements in reconstructing the pronunciation of earlier stages of Chinese. A few have also attempted to use the phonetic elements in teaching. Two well known examples of pedagogical use are works by Wieger (1965) and Soothill (1942) that classify characters under 850–900 phonetics based on Mandarin pronunciations.

An extremely useful, if somewhat flawed, study was published in 1814 by the missionary-scholar Joshua Marshman, who analyzed the characters in the eighteenth-century Kang Xi dictionary. He excluded from consideration more than a third of the characters on various grounds, such as their being mere stylistic variants or lacking explanations. This left him with about 25,000 characters that can more or less be viewed as the total unabridged lexicon of Chinese over the past two millennia. Removing from each character what he called its "Element," that is, the semantic element or key under which the character was classified, he arrived at a figure of 3,867 residual components, which he called "Primitives." He concluded that all characters, apart from the few hundred consisting only of a single component, are formed by combining one of the 214 elements with one of the 3,867 primitives. He referred to the combinations thus formed as "Derivatives."

By his use of the term *primitives* we can conclude that Marshman correctly assigned the primary role to this category of components that enter into the composition of Chinese characters. At the same time, however, he was so firmly convinced that the primitives "convey a general idea" that he failed to appreciate the significance of the fact (which he himself pointed out) that, for example, 11 of the 16 derivatives (actually there are more) formed with a primitive *li* had exactly the same pronunciation, and all but one had the same initial. Despite his myopia regarding the precise function of the primitives, which was of course chiefly phonetic, his work remains valuable, for it shows that Chinese characters are not all idiosyncratic entities like, as is frequently alleged, our numerals 1, 2, 3.

All Chinese characters, or at least all the characters one is likely to encounter in reading a text written within the past two millennia or so, and excluding a few of direct pictographic origin, are actually combinations of some 200 semantics and 4,000 phonetics. These numbers are large, but they are not open-ended, and above all they are finite enough to make the Chinese system manageable. It works because the phonetic elements are syllabograms that comprise a sort of syllabary. It is, to be sure, an outsized, haphazard, inefficient, and only partially reliable syllabary. Nevertheless it works, as is apparent from the examples given in figure 18.

Perhaps it will help to visualize the structure of Chinese characters if we imagine a huge "Semantic-plus-Phonetic Matrix" composed by listing the 214 semantics on the left and the 3,867 phonetics across the top. Of course not all semantics combine with all phonetics, so that of the over 800,000 cells contained in our matrix, only some 25,000 would be occupied by the derivatives that Marshman selected for study from the Kang Xi dictionary. We can also imagine a smaller matrix based on Soothill's classification of 4,300 of the more frequently used characters (approximately the number needed for full literacy) under 895 phonetics, combined of course with the usual 214 semantics. We extract from the imagined overarching matrix a few examples (from DeFrancis 1984a:106) of cells filled by derivatives that are actually formed by combining one of the 3,867 phonetics with one of the 214 semantics. The numbering system follows that of what might be called the "Soothill Syllabary" of 895 phonetics that is contained within the "Marshman Syllabary" of 3,867 phonetics.

Semantic-plus-Phonetic Matrix

semantic	Phonetic 264 敖 (*áo*)	Phonetic 282 參 (*cān*)	Phonetic 391 堯 (*yáo*)	Phonetic 597 甫 (*fŭ*)
9 亻 'person'	傲 (*ào:* 'proud')	傪 (*cān:* 'good')	僥 (*jiǎo:* 'lucky')	俌 (*fŭ:* 'help')
64 扌 'hand'	摮 (*áo:* 'shake')	摻 (*shán:* 'seize')	撓 (*náo:* 'scratch')	捕 (*bŭ:* 'catch')
75 木 'wood'	檖 (*áo:* 'barge')	槮 (*shēn:* 'beam')	橈 (*náo:* 'oar')	椨 (*fú:* 'trellis')
85 氵 'water'	滶 (*áo:* 'stream')	渗 (*shèn:* 'leak')	澆 (*jiāo:* 'sprinkle')	浦 (*pŭ:* 'creek')

As can be seen by reading across the rows, in many but not all cases the semantic element on the left provides a sort of thesaurus-like clue

to the meaning of the items on the right. All those to the right of no. 85, for example, have something to do with water. The phonetics noted at the top of the chart appear to give some clues to the pronunciation of the characters of which they form a part. Although the clues vary in the degree to which they suggest the pronunciation of the full characters, overall they are far more specific than the semantic clues.

Some phonetics are more productive of derivatives than others—from as few as two or three to as many as almost two dozen. In figure 18 the *lì* phonetic (no. 511) and the *zhōng* phonetic (no. 784) both occur as a component in 20 characters. *Yáo* (no. 391) and *mò* (no. 453) each has 22 derivatives. The phonetics are more likely to be evident in less frequently used characters, as attested by the fact that they enter into an average of 6.5 characters in the Kang Xi list of 25,000 characters but only about 5.0 in the Soothill selection of 4,300. The lower ratio in the latter shows the effect of attrition in more frequently used characters, where the original structure of the graphs has often become so distorted as not to be readily recognizable.

Some semantics also occur more frequently than others in compound characters. The "vegetation" semantic occurs in hundreds of characters. The "step forward" semantic occurs in 17 derivatives, all but two or three of which are quite rare.

The illustration in the matrix of the way semantics and phonetics combine to form new characters in Chinese can be used to expand on the important difference between Sumerian and Chinese mentioned earlier. If Chinese combined the two elements along the same lines as Sumerian, the two characters for *ào* 'proud' and *ào* 'stream' might appear as on the left below instead of as they actually do on the right:

亻敖		傲	*ào* 'proud'
	vs.		
氵敖		滶	*ào* 'stream'

The detachability of the semantic elements for 'person' and 'water' in the characters to the left would incline us to view these characters rather differently from those on the right. We would surely pay as much if not more attention to the phonetic elements like *ào* than to the determinatives and would view them all as separate entities. In counting symbols we would therefore most likely say that the virtually unabridged historical lexicon based on the Kang Xi dictionary has

$3867 + 214 = 4081$ different symbols instead of the astronomical 25,000 that we see in the more closely-knit derived characters. Similarly, the abridged modern selection presented by Soothill might be said to comprise $895 + 214 = 1109$ different symbols instead of the 4,300 obtained by combining semantics and phonetics.

Another point we have to consider is this: Just how useful are these semantic and phonetic elements? The former, it is clear, can at best suggest only a general semantic area. Thus we know that characters containing semantic no. 85 most likely have something to do with water, and those containing semantic no. 140 with vegetation. In fact the so-called semantic in many characters does not provide even this limited amount of information. They often offer no real semantic information at all and merely serve to differentiate one character from another, as do our spelling distinctions in *hair* and *hare*.

There is a wide range in the usefulness of the phonetic elements. We can distinguish four degrees of correspondence between a phonetic and the derivative of which it forms part:

1. In some cases the phonetic tells us with 100 percent accuracy the pronunciation, even as to tone, of the full character of which it forms part. So phonetic no. 74, *huáng* (see figure 18), indicates exactly the pronunciation of the 14 derivatives of which it forms part. An example is one in which it combines with the semantic key no. 142, 'insect,' to form the first part of the two-syllable word *huángchóng* 'locust':

key	phonetic	derivative
虫	皇	蝗
insect	*huáng*	*huáng*

2. Some phonetics indicate the pronunciation of the derivative character except possibly for tone. Phonetic no. 255, *mǎ*, is such a phonetic in 10 derivatives. One example is the character for the word *mā* 'mother,' in which, as already noted, it represents the pronunciation with complete accuracy except for tone:

key	phonetic	derivative
女	馬	媽
female	*mǎ*	*mā*

3. Some phonetics indicate only part of the sounds which comprise the syllable represented by the derivative. Usually it is the final, the major component of a syllable, that is represented. Thus phonetic no. 391, *yāo*, enters into 22 derivatives variously read as *yáo, yǎo, jiāo, jiāo, qiāo, qiáo, xiāo, xiāo, náo, nǎo, nào, ráo, rào, shāo*. One example is the character for *jiāo* 'to sprinkle' composed of this phonetic and the 'water' determinative as its key:

key	phonetic	derivative
氵	堯	澆
water	*yāo*	*jiāo*

4. Some so-called phonetics provide no useful phonetic clue. This is sometimes due to the mistaken analysis and classification of characters by dictionary makers, including Soothill, but perhaps even more, as Ramsey reminds us (personal communication, 5/26/88), to the extensive phonological changes that have taken place during the long stretch of time over which the series was built. Sound changes of various kinds have obscured some of the homophony or near-homophony that once existed.

An example of a useless phonetic appears in the character *xià* 'below.' Its actual etymology, as mentioned earlier, goes back to a simple indicative graph consisting of a dot or dash below a horizontal line. The modern character is mechanically analyzed by Soothill, whose popular and convenient work I largely follow despite some points of disagreement, as a derivative made up by combining key no. 1 'one,' under which it is customarily classified, and phonetic no. 119 *bǔ* 'to divine':

key	phonetic	derivative
一	卜	下
one	*bǔ*	*xià*

Clearly *bǔ* is completely useless as a phonetic for *xià*.

A somewhat different group of characters in this category consists of those which some specialists in Chinese, though not ordinary readers, might be able to identify as having useful phonetics. A case in point is the last character, *guān* 'gate,' which Soothill places under

phonetic no. 635 *mén* (see figure 18). Specialists like Karlgren (1940: 187) may be able to correct Soothill's misplacement of this character under the phonetic *mén* by noting that it is a derivative made up by combining elements which include key no. 169 'door' and a rare phonetic (the bottom part of the character) which had the early pronunciation *kwan*:

	key	phonetic	derivative
Soothill		門	
	門	*mén*	關
Karlgren		卝	
	door	*kwan*	*guān*

The potential utility of phonetics in this last group of characters is not reflected in figure 18, which is mainly limited to examples of the first three groups of characters.

There is ample evidence that while what we might call the "spelling" of derived characters indicated by the phonetic element is not a completely reliable guide to pronunciation, any more than is the case with English spelling, nevertheless it is by no means useless or unused. Readers of Chinese frequently guess at the pronunciation of unknown characters by referring to the phonetic component. Writers frequently make mistakes by writing wrong characters that have the same or similar sounds as the intended graphs.

It is pertinent, therefore, to look a bit more closely at the issue of phoneticity that was mentioned earlier in citing Y. R. Chao's estimate that Chinese writing is 25 percent phonetic as against 75 percent for English. Research done on this issue indicates that if one has memorized the pronunciation of the 895 phonetic elements singled out by Soothill, it is possible in 66 percent of the cases to guess the pronunciation of any given character one is likely to encounter in reading a modern text.

If we apply only those phonetics like the aforementioned *huáng*, which reflect pronunciation with complete accuracy, we have a 25 percent chance of guessing the pronunciation of the characters in a given text. It is probably not coincidental that this figure is identical with Chao's estimate. His definition of phoneticity may well have been based only on such cases, where a phonetic precisely matches the pro-

nunciation of a character of which it forms part, even including the tone.

But symbols which represent accurately the phonemes of a syllable other than the tones are also generally useful. This is attested by the fact that much has been published in a variety of Chinese scripts which do not indicate tones. One such is the Latinxua or Latinization scheme in use before World War II (DeFrancis 1950). Since the early 1950s newspapers, poetry, fiction, and works on linguistics, history, and politics have been published in a Cyrillic transcription of Dungan, a dialect of Northwestern Mandarin. This dialect is spoken by some 36,000 people in Soviet Central Asia descended from Muslim Chinese refugees who fled persecution at the hands of the Manchus in the nineteenth century (Isayev 1977:186–187; Rimsky-Korsakoff 1967:356, 410–413; Rimsky-Korsakoff Dyer 1987:235). If we add the 17 percent of phonetics of this type, represented by the phonetic *mǎ*, phoneticity increases to 42 percent.

Even phonetic elements of the *yāo* type are useful since they give hints about the pronunciation of part of a syllable, usually the final part, which is the most distinctive part of the syllable. Hence they generally permit a good guess at the pronunciation of a character in context. If we add the 24 percent of phonetics of this type, phoneticity increases to the figure of 66 percent mentioned above (DeFrancis 1984a:105–110).[7]

The 66 percent figure represents a conservative estimate of the phoneticity of Chinese characters. Scholars with specialized knowledge of Chinese historical phonology can often derive additional phonetic information from the previously mentioned fourth category, which I have dismissed as providing no useful phonetic clues. This rejection is based on my estimate of utility for the average, linguistically unsophisticated reader of Chinese texts. More knowledgeable readers such as the specialized scholars Bernhard Karlgren and William S.-Y. Wang are able to discern corespondence between phonetic elements and full characters that is not apparent to ordinary readers. They arrive at a somewhat different classification of characters, primarily those in my fourth category. Those of the *xià* 'below' type they consider as not belonging to the SP catergory at all. Some others, such as *guān* 'gate,' are reclassified into my third category on the basis of a more refined phonological analysis that is made possible by a more sophisticated understanding of the history of a character and the sounds attached to it. The net result of all this is that their estimate of

phoneticity (as defined by my first three categories) rises to as high as 90 percent (Karlgren 1923:4; Wang 1981:232).[8]

Apart from the multielement (SP) graphs which contain phonetics of the varying degrees of utility described, there are also single-element graphs which themselves comprise phonetics. (Some also function as keys.) Here are a few examples of characters we have already encountered:

> 馬 *mǎ* 'horse'
>
> 門 *mén* 'door'
>
> 目 *mù* 'eye'
>
> 人 *rén* 'person'
>
> 象 *xiàng* 'elephant'

The overall distribution of the various kinds of characters can be roughly summarized as follows:

Kind of Character	Example	Percent
A. single-element characters	馬 *mǎ* 'horse'	1
B. multielement (SP) characters		
1. completely useful phonetic (represents all the phonemes of the derivative)	蝗 *huáng* 'locust'	25
2. generally useful phonetic (represents all the segmental phonemes, but perhaps not the tone)	媽 *mā* 'mother'	17
3. contextually useful phonetic (represents most of segmental phonemes)	澆 *jiāo* 'sprinkle'	24
4. useless phonetic (represents no significant phonemes)	下 *xià* 'below'	33
Total		100

The utility of the first three categories of phonetic elements becomes even more apparent if we look at the phonetics not merely in isolated graphs but also in characters as we normally encounter them, which is,

of course, in context. Even the minimal environment provided by two-character expressions illustrates this point, as in the following examples consisting of several pairs of phonetics and their derivatives:

phonetics		derivatives		meanings
分方	*fēnfāng*	芬芳	*fēnfāng*	fragrant
免厲	*miǎnlì*	勉勵	*miǎnlì*	encourage
山夭	*shānyāo*	訕笑	*shānxiào*	ridicule
士原	*shìyuán*	志愿	*zhìyuàn*	aspiration
亡生	*wángshēng*	忘性	*wàngxing*	forgetfulness

The wider contexts in which these words are normally encountered will enable readers to handle the disparity in pronunciation between phonetics and derivatives, just as readers of this book will no doubt, either consciously or unconsciously, carrect the preceding misspelled word.

Chinese spelling as represented by its phonetic elements is erratic, inefficient, and difficult to master. But the same has been said about English spelling. Chinese writing deserves these opprobrious labels even more than does English, but this should not obscure the fact that phoneticity, deficient though it has become, far surpasses iconicity, which actually approaches zero.

Yet this fact is indeed commonly overlooked by people who mistakenly call Chinese "pictographic" or "ideographic." These labels are popularly attached to Chinese characters by Western writers. The Chinese themselves are also almost universally convinced that theirs is a unique system that they call *biǎoyì* 'semantic' or 'ideographic' writing. The writing system does contain some symbols that might, very loosely, be so labeled, but a few, or even a few hundred, such symbols do not make a system of writing. In actual fact, there never has been, and never can be, a full system of writing based on the pictographic or ideographic principle.

What then of the frequent designation, especially in academic circles, of Chinese as a logographic or morphemic system of writing? Other writing systems, such as Sumerian, are also described by these terms, but Chinese is usually taken as the example par excellence of this category of scripts. I think this too is a serious error, and the error is compounded by sinologists because they have been unduly influenced by the previously mentioned difference between Chinese and Sumerian in the way they handle semantic and phonetic determinatives. In contrast to Sumerian writing, in which the determinatives are detachable from the graphs they determine, Chinese writing welds these elements so tightly together that the characters, surrounded as they are by white space in their little square cubicles, are usually viewed as unitary symbols, or at least as the basic unit in the writing system.

The error here will become clearer if we invoke the concepts of grapheme and frame (or lexeme) that were discussed in the section "The Forest of Family Trees" in chapter 2. The grapheme, as we recall, is the indispensable meaningless unit that corresponds to the smallest segment of speech represented in the writing system. The frame is the dispensable meaningful unit that corresponds to the smallest segment of writing conventionally receiving special status, such as being surrounded by white space and listed in dictionaries.

In English, the grapheme is a letter or combination of letters corresponding to one of the approximately forty smallest units of speech, the phonemes, that are represented in the writing system. The frame is a word, the smallest unit of writing that is conventionally surrounded by white space and listed in dictionaries.

It is my contention that in Chinese the syllabic element P, such as that in the overwhelmingly preponderant SP characters, must be viewed as the grapheme, the indispensable phonetic unit without which the system would not work. Whole characters are frames or lexemes, secondary units that in a reformed system of writing could be dispensed with entirely, along with the semantic element S. The Chinese frame is a derivative, as is true also of the English frame. Chinese writing, consisting as it does of derived characters, can be called logographic (or morphemic) only if English writing is also called logographic because it too consists of derived frames, in this case called words. But by this standard most, if not all, systems of writing must be called logographic, which then becomes a vacuous term utterly lacking in any power to differentiate systems of writing.

The Chinese system must be classified as a syllabic system of writing. More specifically, it belongs to the subcategory that I have labeled meaning-plus-sound syllabic systems or morphosyllabic systems.

I use the term *morphosyllabic* in two senses. The first applies to the Chinese characters taken as individual units. Individual characters are morphosyllabic in the sense that they represent at once a single syllable and a single morpheme (except for the 11 percent or so of meaningless

characters that represent sound only). In this usage the term is intended to replace the more widely used expressions *logographic, word-syllabic,* and *morphemic,* all of which are applied to individual characters taken as a unit. The second sense of the term refers to the structure of Chinese characters and is intended to draw attention to the fact that, in most cases, a character is composed of two elements, a phonetic grapheme which suggests the syllabic pronunciation of the full character, and a semantic element which hints at its meaning.

The aspect of the Chinese system of writing covered by this second sense of the term often receives little if any attention. This applies particularly to the phonetic grapheme. Its neglect leads to widespread errors in viewing characters as either (1) unitary symbols with no representation of sound, or (2) compound symbols made up by combining semantic elements with no representation of sound, or (3) compound symbols with phonetic elements of so little importance that they can be disregarded.

Whereas English graphemes represent phonemes, Chinese graphemes represent syllables, or better still, accepting Boodberg's felicitous label, "syllabic phonemes" (Boodberg 1937:331). There are in current Chinese some 1,277 syllabic phonemes counting tones, about 400 not counting tones. For purposes of comparison, let us say that, in round figures, there are 40 phonemes in English and 400 toneless and 1,300 tonal syllabic phonemes in Chinese.[9]

A sampling of the characters in a dictionary with about 4,800 entries indicates that 44 percent represent free words, 45 percent are bound morphemes, like *-er* in English *teacher,* and 11 percent are meaningless symbols that represent only sounds, like the *cor* and *al* of English *coral* (DeFrancis 1984a:184–187). On the basis of these figures Chinese characters are at best 44 percent logographic and 89 percent morphemic. But the Chinese *writing system* is 100 percent syllabic since all characters (except that for the suffix *r*) represent syllables, either as single-element graphs which themselves comprise phonetics or as multielement graphs which include phonetics of the varying degrees of utility noted earlier. It is a mistake to take Chinese frames, or lexemes, as the basis for defining a writing system, just as it would be a mistake to call English logographic because its frames are words. Yet the superficial approach of equating Chinese characters with the Chinese writing system is often adopted. People fail to look below the surface of the characters to what makes the characters work and allows new ones to be generated as needed.

It would be quite impossible to write Chinese exclusively with logographic or morphemic frames not further divisible into components that minimally include a phonetic grapheme. The number of words, in the order of hundred of thousands if not more than a million, is much too large.

The number of morphemes is harder to estimate. If we accept the conventional view that Chinese characters represent morphemes, which as noted above is approximately 89 percent true, then there are at least 25,000 morphemes in the Kang Xi dictionary. An incomplete study based on only 4,200 characters estimates the number of morphemes at 5,000 in modern Mandarin. The author notes that this figure includes only "frequently used morphemes" and should be increased to 7,000 or 8,000, which even so excludes polysyllabic morphemes, foreign loanwords, and personal and place names (Yin Bin-yong 1984 and personal communication, 1/8/87). Regardless of the precise figure, it is obviously very great, much too large for a purely morphemic script.

On the other hand, it would be a relatively simple matter to write Mandarin Chinese with a standardized syllabary of only 1,277 signs, which could be reduced to 398 if tones were separately indicated. Figure 19 presents such a standardized syllabary based largely on the Soothill Syllabary (DeFrancis 1984a). The pronunciation of the characters is indicated by the Pinyin transcription that was adopted in 1958 as the official way of transcribing the characters.

But instead of a relatively small number of syllabic graphemes, Chinese has, according to the Marshman study cited earlier, something like 4,000 such basic signs. It is partially coincidental, but not completely unrelated, that the figure approximates the picture of maximum syllabic complexity attributed to a sixth-century dictionary which divides all the sounds of the language into 3,877 groups (Kennedy 1964:113–114). Actually this figure is suspect, and it is unlikely that Chinese ever had this many syllables. Boodberg makes the startling suggestion that the number of different syllables in the still earlier phase of Chinese, which some scholars consider to have been phonologically the most complex, was more limited than in modern Mandarin (Boodberg 1937:360).

Regardless of the precise number of syllables in Chinese in the various periods of its evolution, it is clear that there have always been many symbols for the same sound. Chinese writing never underwent the reduction in number of symbols that characterized the evolution of

Figure 19. Chinese Writing: A Simple Syllabary and a Simpler Alphabet

A standardized syllabary of 398 signs which with four additional marks for tones would enable Chinese to be written with full accuracy and relative ease in a way comparable to the Japanese use of kana symbols. The Pinyin transcription of the 398 basic syllabic signs combines an initial consonant (shown in the left-hand column) with a final (shown in the top row) made up of one or more vowels and an occasional ending in n or ng. Reprinted with permission from John DeFrancis, *The Chinese Language: Fact and Fantasy.* (Honolulu: University of Hawaii Press, 1984.) p. 27.

the cuneiform scripts. Indeed, the Chinese seem to have almost a penchant for avoiding simplification and standardization. This is seen also in the failure to make efficient use of a syllable-telescoping technique that has some similarity with that devised by the Sumerians.

The Chinese variation of this technique, which they call *fǎnqiè* 'reverse-cutting,' indicates the syllabic pronunciation of an unknown character C through the intermediary of two presumably known characters A and B by cutting off the final part of the syllable from A and removing the initial in B. This is as if in English we indicated the spelling of *cat* by telescoping *cup* and *rat* as follows: c(up r)at.

With a stock of only about 40 A's and about 200 B's this could have have been made into a fairly simple standardized system capable of expressing all the syllables. But the Chinese never standardized the system, and indeed selected characters, some of them quite obscure, at random, as if we spelled *cat* indiscriminately as *cup-rat, cow-fat, coal-hat, cap-mat,* and so on. This failure resulted in the haphazard use of about 500 A's and 1,200 B's. And sometimes the "reverse-cutting" was circular, with C being explained by reverse-cutting A and B, and A by reverse-cutting C and B (Kennedy 1953:8, 146–147).

But the shortcomings of the Chinese "reverse-cutting" device, which not surprisingly confused Gelb (Gelb 1963:87–88; DeFrancis 1950:40–47), are not particularly relevant, since little use was made of it. Unlike the Sumerian technique, which played an important role in that writing system, the Chinese variation did not form part of the writing system itself but was confined to lexicographic use. Modern dictionaries have now abandoned this inefficient way of indicating the pronunciation of characters in favor of newer techniques closer to the alphabetic principle.

Traditional Chinese writing never attained even the limited degree of simplification that marked the evolution of cuneiform writing. Throughout its history the actual sound-to-symbol relationship in Chinese has approximated on the syllabic level the much-maligned situation in English on the phonemic level. In contrast to the one-to-one relationship, where there is close correspondence between sound and symbol, both writing systems are characterized by a highly complex many-to-many relationship. Thus English spells the same sound *o* in at least ten different ways: *so, sow, sew, oh, owe, dough, doe, beau, soak, soul.* It uses the same letter *o* to represent at least 8 different sounds in *so, to, on, honey, horse, woman, borough* (DeFrancis 1984a:112). The situation is the same, on the syllabic level, in Chinese. Here some syllables are represented by many different symbols, which may be either whole

characters or the phonetic components in more complex characters of the SP type. And some symbols have several different pronunciations.

The poor fit between sound and symbol in both English and Chinese should not obscure the key fact that both are based on phonetic principles, with the 40 phonemes of English being represented by various alphabetic spellings, and the syllables of Mandarin Chinese being represented by various syllabic spellings. The number of different spellings for the 40 English phonemes has been variously estimated at 600 (Zachrisson 1931:4), 1,120–1,768 (Nyikos 1988; see 298 below), and 2,000 (Alisjahbana 1965:530; Daniels 1985:34). The ratio between syllabic spellings and syllabic phonemes in Chinese is much smaller, but the greater complexity of its graphic symbols makes for a system the cumbersomeness of which considerably surpasses that of English and perhaps of all other systems ever created.

It hasn't needed to be so. As noted earlier, it would be possible to write Mandarin Chinese quite simply and accurately with only 1,300 different signs. It would be possible to manage with only 400 symbols if tones are separately indicated, or not indicated at all. But writing with a simple phonetic script, whether syllabic or alphabetic, would be impossible without the adoption of a further feature that has characterized Chinese written in an alphabetic script. This is a literary style that is more closely based on actual speech.

Over the past hundred years there has been a long-running debate regarding the Chinese literary style and the Chinese character system of writing. Proponents of reform are urging a more colloquial style of writing and the extended use of the simple romanization system called Pinyin, not as a replacement for the characters, but as part of a policy of digraphia, that is, the use of two more or less equal systems of writing, each to be used in the areas for which it is best suited, such as Pinyin for computers, characters for historical research (DeFrancis 1984a, 1984b). The promulgation in July 1988 of rules for Pinyin orthography, that is, rules for such things as punctuation and use of blank space, hyphens, and closed juncture between syllables, is expected by Chinese reformers to help create digraphic literates who would extend the use of Pinyin from a mere tool for annotating characters to an auxiliary system for writing the language.

However, reformers seeking to speed China's modernization by modernizing the writing system through a policy of digraphia have to contend not only with the natural attachment of Chinese to their familiar script but also with chauvinistic and mindless claims for its superiority. For years the official *People's Daily* has promoted a cabal of conservative dabblers in the area of writing, headed by a wealthy returned expatriate, as part of a campaign attacking the reformers and extolling the traditional characters. The intellectual level of the campaign is indicated by an item, carried in *China Daily* (11/15/1984) under the headline "Characters 'easier than ABC to read'," which retailed the preposterous claim of an establishment psycholinguist that "children aged 2 to 4 can easily learn 3,000 characters." China's writing reformers and forward-looking educators, in their uphill battle against such drivel, are beleaguered in an atmosphere of intimidation and quackery reminiscent of the intellectual climate in the Stalinist period that earned Soviet linguistics and genetics, long dominated respectively by N. Y. Marr and T. D. Lysenko, the contempt of scholars throughout the world. In contrast to countries like Turkey (Heyd 1954; Bazin 1983), North Korea (Blank 1981), and Viet Nam (DeFrancis 1977), where writing systems and writing styles were reformed in a matter of a few years or a few decades, it appears that controversy over basic problems of writing is likely to drag on indefinitely in China.

Mayan

The origin of Mayan writing is the subject of the same disagreement between original inventionists and diffusionists as Chinese. The diffusionist Cyrus H. Gordon and the Mayanist David H. Kelley have seen a connection between the "Phoenician-Hebrew alphabet" and Mesoamerican lists of signs for days of the month, and they consider that "the origin and the transmission of both the alphabet and the lists are the achievement of intercontinental mariners" (Gordon 1970:193; also Gordon 1971 and Moran and Kelley 1969). It appears, however, that most Mayanists strongly reject the notion of contact with the Old World (Lyle Campbell, personal communication, 5/13/87). The generally accepted view is that Mayan writing is an indigenous creation—the third in the history of syllabic writing—and that in the development of their system the Maya were influenced by antecedents in the form of pictographic or hieroglyphic symbols that were executed by various neighboring peoples. However, the precise nature of these symbols and their relationship to those developed by the Maya has yet to be fully established (Justeson et al. 1985).

Despite this general agreement regarding the origin of Mayan writing, there has been sharp disagreement among Mayanists on other matters that have occupied most of their attention. Chief among these