The Origins of Chinese Characters

Revised Edition

王宏源　著
Hongyuan Wang

英汉双语

汉字字源入门

—修订本—

漢字

漢字

汉字字源入门

—修订本—

The Origins of Chinese Characters

Revised Edition

王宏源 著

Hongyuan Wang

汉字字源入门（修订本）

PREFACE

Unicode is the foundation of mathematical platform for human beings, and in math we trust that it is the basic for the Earth. Meanwhile, Unicode is a new gift from the West to the East, not only due to the huge quantity of the ideographic characters in East Asia, but also for the convenience it offers for editing or compilation, and survival of the Chinese dictionaries themselves.

Having acknowledged the profound significance of this universal standard, one idea about or opportunity for its application appears in front of us. If we include the Unicode for a character entry in a dictionary, by means of the code we can input the character in our computer immediately. For example, MS Word has a simple method: Internal Code Input Method. When we type in the Unicode *5B9D* for 宝, then press ALT+X simultaneously, the character 宝 shows up on our screen, and vice versa. Moreover, we can also try to hunt the *5B9D* for 宝 or *5BF6* for its traditional form 寶 in BableMap (Unicode Character Map for Windows), a web search engine, the Wiktionary, or even the unicode.org.

The Origins of Chinese Characters is a dictionary-like book of mine first published 25 years ago. In this revised edition, the primary job is to introduce Unicode as the reference for each character entry, allowing readers to make use of these electronic characters. The other revisions in this Unicode edition are: (1) Corrections to some errors in the original edition; (2) Additions and replacement for some character entries; (3) Additions of many glyph shapes; (4) Additions of page numbers for reference characters in the text, to facilitate quick search; (5) Additions of some Chinese and English contents for most entries. The modifications are under the original page format, and all the revisions are moderate. In the mean time, I removed the bibliography since this book is used for learning and teaching Chinese as a foreign language, instead of for studying the ancient Chinese writing itself.

Chinese writing system is the centralized embodiment of primitive thinking mode, mythology, aesthetic ideology and creation rules of ancient Chinese, however, the source and development of Chinese characters have been sometimes illusory or incomprehensible. For example, as a common character and abstract concept, 毒 dú *poison*, is confused sometimes with its two components: 生 shēng *to live*, and 母 mǔ *mother*. Is it originally from the legend of tasting the herbs or from the understanding that harmful and poisonous plants grow wildly? Meanwhile, some readers will find that the component 女 nǚ sometimes has a negative meaning, such as in 姦=奸 jiān *wicked*, 嫌 xián *disapproving* and 奴 nú *slave*. However, fortunately we have the word 好 hǎo *nice, good, fine*, combining 女子 nǚzǐ, that is *woman* or *female* literally.

I wrote this bilingual book, *The Origins of Chinese Characters,* around 1990 when I was a student in the Department of Physics, Tsinghua University. It was reprinted many times and translated into French in 1995 and German in 1996. After leaving the university, I was engaged in international trade, and then studied abroad, earning a degree in electrical and computer engineering from the University of Waterloo, and then ran a database company. However, I have always had a great interest in the ideographic characters, and editing *The Unicode Han Dictionary* is a hobby and took years of hard work. This revised Unicode edition is as another episode of my regular time-consuming work on the dictionary. I have really enjoyed the experience, and hope I can update and enlarge this book within a decade.

Harrison Hongyuan WANG

October. 20th 2018

序 言

　　《汉字字源入门》成书于1990年前后，当时作者系清华大学物理系的学生，在与来华学习汉语的外国学生交流讲授汉字起源的过程中，整理完成了这部英汉双语对外汉语教学辅助读物。该书是作者30年前以拆分汉字构件的方式，利用古文字的知识讲解汉字的一次尝试。书中虽然收字有限，解文释字也多有不足之处，但是由于该书是较早出现的一部采用汉字起源作为对外汉语教学的类字典读物，书中讲解的641个汉字字头具有构件属性，组字能力强，图书出版25年来多次重印至今，发行量大，销售面广，其英文部分并被译成法文和德文，具有一定的影响力。

　　这次《汉字字源入门》的修订出版，主要的目的是实现了"字宝"化改造，即在原书字头下加注Unicode编码。同时，"旧瓶"中装了些"新酒"，在保持旧稿版式、页码不变的基础上，全书内容有如下更动：（1）修改原版中的一些错误；（2）添加个别字头；（3）添加一些古文字字形；（4）对行文中的参考字，如出字头，则标注其所在的页码，便于读者快速翻检查验；（5）适量添加一些中英文内容，并在字头下引入一些相关联的衍生疑难字、游戏叠加字等，以增加趣味性。

　　半年来，作者利用25年前的旧稿修订《汉字字源入门》，感觉当年临摹字体、"剪刀加浆糊"出书的干劲没有变，面对汉字快乐的心态没有变，唯一改变的就是自己的年龄。当年拿到《汉字字源入门》样书时，年轻的作者在扉页上抄写下王国维《晓步》诗中的一句，在25年后的今天，情景和心境可能更贴切：

　　四时可爱唯春日，一事能狂便少年。

<div style="text-align: right">

Harrison Hongyuan WANG

王宏源

2018年10月20日于瀚堂

</div>

关于"字宝"

　　2000 年国际统一编码标准（Unicode）发布以来，对全球信息化的发展产生了重要影响。作者在十多年来依据 Unicode 东亚表意字符标准（Unicode Han）编撰字典的实践过程中，逐步形成了一个概念，即在纸质字典的字头附注 Unicode 编码数值。由于一个字符的编码完全等同于该字符，读者可根据编码，在电子编辑器上快速将编码转换为字符。作者将采用这种方法的出版物定名为"字宝"（Unicode Dictionary），"字宝"是字典在数字化时代的升级，是任意字符从纸质转化到电子设备中的桥梁，也是纸质字典在数字时代可以延续功用的有力武器。

　　大家在使用传统中文字典时，查询到任何字，不存在无法利用的问题，只要有笔和纸，就可以将其写出来。在数字化时代，读者在字典中查询到的字，如果属于生僻字，很多时候将其录入到电子设备中是困难的。"字宝"则是根据 Unicode 编码编撰的字典，读者将纸质字典上的 Unicode 编码输入到编辑器中，即可立即显示出字符，可以说利用"字宝"，任何字符都可以快速录入。作者在一年前出版的类字典著作《殷周金文字宝》一书中，首次明确提出此概念，并采用了"字宝"的形式，对 Unicode 发布的 2300 多个相当生僻的殷周金文隶定字，附注 Unicode 编码数值，使得读者在需要录入特定字符时，可立即获取之。

　　随着《殷周金文字宝》一书的出版，"字宝"这一概念引起了一些读者的兴趣。采用"字宝"这种方式，不仅对殷周金文隶定字这类生僻字符的"录入"有效，同样也可以将其引入到常见汉字中。作者进而萌生了将本人早期撰写的对外汉语教学辅助教材《汉字字源入门》一书，进行"字宝"化改造的想法。就对外汉语教学而言，常见汉字的录入对非汉语母语的学生来说也是难题，若采用"字宝"的方式，则可利用字符编码将汉字便捷地"录入"到电子文档中，《汉字字源入门（修订版）》中每个字头附带 Unicode 编码数值，就是作者这种构想的尝试结果。利用"字宝"，只需要开启 MS Word，输入字符编码，随即同时点击键盘中【ALT】+【X】两键即可。反之亦然，即光标在一个汉字字符后面，同时点击【ALT】+【X】，立即出现字符的 Unicode 编码。当然，读者亦可以下载安装最新版本的 Unicode 字符映射表，例如 BabelMap（Unicode Character Map for Windows），或者利用网络直接检索编码数值，以获取相应的汉字。

INTRODUCTION

This book is intended as a guide to the origins and histories of Chinese characters. Although it could be used as a Chinese etymological dictionary, this book is an attempt to find a new way to teach the practical ideography of Chinese to those whose native scripts are alphabet-based.

Writing is a system of conveying ideas by means of conventional symbol that form visible marks. These symbols are traced, incised, drawn, or written on the surface of materials such as tortoise shells, bones, stone, metal, bamboo, papyrus, parchment, or paper. Writing gives permanence to human knowledge and enables communication over great distances.

Writing grows out of pictures. This is as it should be, since the most natural way of communicating visually is through pictures. Some time in the Upper Paleolithic period, perhaps about 20,000 BC, early man in southern France and northeastern Spain drew sketches of his prey — horse, buffalo, deer and other animals — on the wall of his cave and colored them with earth and vegetable dyes. Several factors — some aesthetic, some spiritual or magical — may have led to the creation of these primitive drawings. This may have been the beginning of art, but it was hardly the beginning of writing. Such pictures do not represent writing because they do not belong to a system of conventional signs and their significance can be understood only by the man who drew them or by his family and close friends who had heard of the event depicted. However, genuine writing, whether it retains a pictorial form or not, serves purely to communicate.

In the process of using pictures to identify and recall objects or beings, complete correspondence is gradually established between written signs and objects or beings. These simple pictures contain only those elements that are important for the communication of meaning and lack the embellishments that are included in an artistic representation. Since these objects and beings have names in spoken words, a correspondence is also established between the written signs and their vocal counterparts. When individual signs are used to express individual words and syllables, it can lead to the development of a system of word signs; that is, script or logography. In logography, one sign or a combination of signs

expresses one word or a combination of words. However, pure logography is not found in any known system of writing. It exists normally only in conjunction with syllabography or syllabic writing, as best represented in logo-syllabic writing.

Logo-syllabic writing, that is, writing with signs express that words and syllables, is found in the East, including continental Asia from the eastern shores of the Mediterranean Sea to the western shores of the Pacific Ocean. Egypt and the area of the Aegean Sea, at least in the pre-Hellenic period, are included within the orbit of Oriental civilizations.

This large area is home to seven original and fully developed logo-syllabic systems of writing: Sumerian in Mesopotamia, 3100 BC to AD 75; Proto-Elamite in Elam, 3000 to 2200 BC; Proto-Indic in the Indus Valley, around 2200 BC; Egyptian in Egypt, 3000 BC to AD 400; Cretan in Crete and Greece, 2000 to 1200 BC; Hittite in Anatolia and Syria, 1500 to 700 BC; and Chinese in China, 1300 BC to the present. Other logo-syllabic systems may exist, but at the present to the above list stands. The Proto-Armenian inscriptions discovered within the last few decades are too short and too little known to allow any safe conclusions. The mysterious Easter Island inscriptions are not writing even in the broadest sense of the word, as they are probably nothing but pictorial concoction for magical purposes. Finally, the systems of the Mayas and the Aztecs do not represent a full logo-syllabic writing system; even in their most advanced stages they never attained the level of phonographic development of the earliest stages of the Oriental systems.

Of the seven systems, three — Proto-Elamite, Proto-Indic and Cretan — are not yet deciphered. Consequently, modern understanding of the logo-syllabic systems is limited to the other four systems: Sumerian, Egyptian, Hittite and Chinese. Chinese writing is the only logo-syllabic writing system still in use today and it has developed a formidable degree of sophistication.

History does not throw much light on the beginning of Chinese writing. Chinese mythology attributes the invention of writing to *Ts'ang Chieh* 仓頡 Cāng Jié. It is said that he got his ideas from observing animals' footprints and birds' claw marks on the sand as well as other natural phenomena. We can divide the development of all Chinese characters into three stages, which overlap: ideography, phonetic borrowing, and picto-phonetic writing.

Ideography is the earliest stage, the forerunner of writing. It consists of a drawing or a combination of drawings to represent the thing or action shown. The drawings are simple and give a clear impression as a memory-aid device. Samples of these devices can be found all over the globe. The following Chinese ideographic characters are good examples: 木 mù

stands for a *tree*, 鱼 yú for a *fish*, 虫 chóng for a *snake*, and 射 shè, literally *to shoot*, is the combination of a hand and the bow and arrow. The number of ideographic characters in Chinese writing system is limited. However, the surviving ideographs have static and definite meanings with simple but distinctive strokes. They are basic and easy to understand and many of them, especially the pictographs of single objects, are used as *radicals* in the Chinese writing.

Ideography is a natural method, but a complete system of ideographs has probably never existed either in antiquity or in modern times. To create and memorize thousands of signs for newly acquired words and names is so impracticable that ideographic writing either can be used only as a very limited system, or it must be adapted in some new way in order to develop into a useful system.

Phonetic borrowing is the second stage. The number of function words and abstract nouns is very limited, but they are used frequently and difficult to draw or show. Therefore, signs for function words and abstract nouns were borrowed from the ideographs with similar pronunciation. Such signs should also be simple and distinct. A borrowed character should
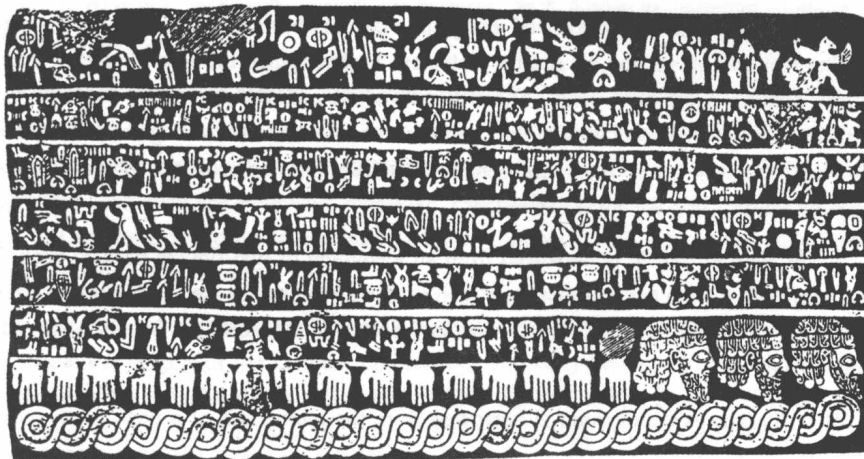
Copy of a hieroglyphic Hittite inscription from the site of Carchemish (*after Renfrew, p.* 48).
赫梯语象形文字（卡尔凯美什遗址）。

Word-Signs on Potsherds   陶器上的刻画符号

上：西安半坡　　下：临潼姜寨

仰韶文化（5000~3000 B.C.）

马家窑文化（3300~2500 B.C.）

have few strokes; and if the borrowed ideograph was not obsolete, it should have its original meaning restored by adding an auxiliary element to distinguish the new character from its original borrowed pictorial form. These borrowed ideographs are called Phonetic Loan Characters (PLC) in this book. The phonetic loan characters can be regarded as symbols. However, in many cases there are semantic connections between the original and the borrowed forms. Let's see some examples:

北 běi, literally *north*, is borrowed from an ideograph of two figures back to back. The *north* may derive from the fact that early man sat facing the sun to the south with his back to the north, thus 北 běi is a PLC. The word for *back* is 背 bèi, adding a *human body*-radical below.

自 zì, literally *self* and *from*, is borrowed from the pictograph of a nose. Both *self* and *from* are difficult to express using drawings. Here, *self* may derive from a man pointing at his nose to express *himself*, and the character was also adopted to indicate the abstract concept *from*. 自 zì thus is a PLC. The word 鼻 bí, *nose*, includes the pictorial element 自 zì and a phonetic element 畀 bì below.

Finally, 枼 yè, meaning thin, derives from a pictograph of a tree with leaves, while 葉= 叶 yè, meaning leaf, a *plant*-radical on the top. Note that since antiquity, many of the pronunciations have changed, so that the original *borrowed* and *returned* forms no longer sound alike sometimes.

The adoption of the borrowing method in the Chinese writing system was a watershed transition from memory-aid picture writing to practical logo-syllabic writing. As in the societies become more developed, more and more objects and beings were named more precisely. Therefore, picto-phonetic characters or pictophones emerged.

A picto-phonetic character consists of at least two parts: one part refers to the meaning of the character and is usually called the radical; the other, the phonetic element, gives the sound. In most cases, the phonetic element also has meaning value as a *picto-phonetic* element. The following examples are similar to rebus writing: 蝶 dié and 鰈 dié, literally butterfly and flounder, with the *worm* and *fish* radicals respectively, and the picto-phonetic element 枼 yè: thin. This is the last stage of development. Picto-phonetic characters comprise about ninety to ninety-five percent of all Chinese characters. Today, the number of characters is almost fixed. When the Chinese needed to introduce a word from another language, they use a combination of characters to form a new Chinese word, e.g., 吉普 jípǔ for *jeep*, 浪漫 làngmàn for *romance*, 歇斯底里 xiēsīdǐlǐ for *hysteria*, 激光 jīguāng in mainland China or 镭射 léishè for *laser*.

Whether borrowed or created, a character generally begins its life in Chinese with one meaning and its ancient pronunciation. Yet no living language is stagnant, and in time words develop new pronunciations or meanings and lose old ones. However, the forms of the ideographic characters are somewhat quiescent, especially the pictographs which derive the letters of Chinese from common objects or beings. This book will show you a large sample of these fascinating characters.

A look at the origins of the characters making up the Chinese writing system involves a look also at the origins of the Chinese civilization. The early history of both China and the Chinese dates from the Neolithic period (about 5000 BC) to the *Han* Dynasty (206 BC to AD 220), including the *Shang* Dynasty (1523 to 1028 BC) and the *Zhou* Dynasty (1027 to 221 BC). By the *Han* Dynasty, the number and forms of common characters were fixed on the whole.

Etymology is not an exact science. Often we are unable to discover the origin of a

character, but more often there are multiple genetic stories for one character. Ingenious etymological theories are put forward frequently, some plausible and attractive, others wildly improbable. I have chosen the most likely explanations to be included in this book because the purpose of the book is not to introduce scholarly debates but to serve as a new way to learn Chinese characters. A picture is worth one thousand words. The origins and histories of Chinese characters should be not the missing piece in the Chinese puzzle, but a key to resolving the puzzle.

Some Chinese characters have been simplified as *Jianhuazi* 简化字, and these simplified forms are used in mainland China. But, the ancient form of Chinese characters must be referred to in such a book as this. And although the explanations for each entry are bilingual, they are not equivalent, the Chinese part being simpler with easy characters. For transcription, the Chinese phonetic, alphabet, or *pinyin* 拼音 as it is known, has been used with *four-tones* 四声 in this book.

INSCRIPTIONS

甲 **Shell and bone character** Inscriptions were carved on oracle bones with practical and angular strokes, *Shang* Dynasty.

金 **Bronze character** Inscriptions are found on bronze vessels of the *Shang* and *Zhou* Dynasties. Bronze characters derive from prehistoric picture writing, and their lines are smooth yet forceful.

古 **Ancient character** Inscriptions, which appear on the surface of bamboo, stone, pottery or ancient seals, were used mainly during the Warring States Period (475-221 BC).

篆 **Seal character** A kind of standard or decorative character which appeared in the *Qin* Dynasty (221-207 BC).

参 **For reference**

Hongyuan WANG

December 1991

目录　　**CONTENTS**

Chapter 1

MAN　人类

The first inhabitants on Chinese soil whose remains are known to us were the race to which Yuanmou Man belonged 1,700,000 years ago. However, Peking Man is more famous, and it was the focus of worldwide attention in 1927. Peking Man was later and more "human" than Java Man. Peking Man's bones were firstly discovered by Professor Pei Wenzhong at *Zhoukoudian* near Beijing. Unfortunately, all the *Zhoukoudian* hominid remains disappeared when being transferred from Peking to an American ship during the Japanese invasion of China prior to the Second World War. In Chinese characters, the form of a standing man means "great", perhaps because standing was a great feat in man's evolution.

中国大地上埋藏有十分丰富的古人类化石和旧石器时代遗物,至今已发现的早、中、晚各个时期的地点共 200 多处,包括直立人、早期智人、晚期智人各个阶段的人类化石。这当中据认为最早的为距今 180 万年的西侯度文化(山西省)和距今 170 万年的元谋人*(云南省)。1927 年裴文中教授在北京周口店发现的北京人则是最早发现,最具有影响的中国直立人。不幸的是,大量极其珍贵的北京人遗物,包括五个头盖骨和其它骨和牙齿标本在太平洋战争爆发前,全部在几个美国人手里弄得下落不明。
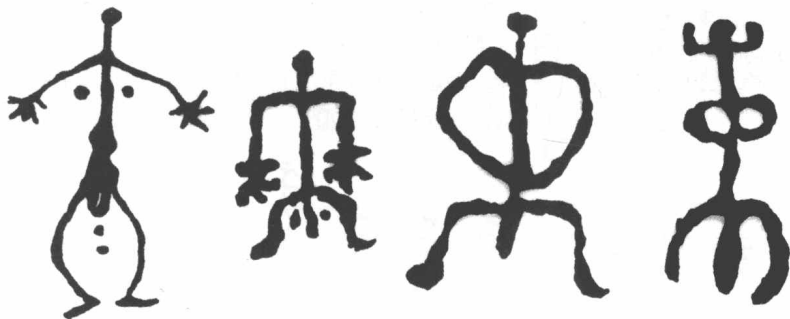
Figures in rock art

岩画中所见人形

* 系根据古地磁学测定的数值。1983 年有人对此提出不同见解,认为元谋人化石年代不超过 73 万年,可能距今 50 至 60 万年。北京人距今 70 至 20 万年。

1.1 **Man's Body   身体**

rén

人 甲 〈figures〉 金 〈figures〉 古 〈figures〉 篆 〈figure〉

4EBA — man, person.

认=認rèn；從 5F9E=从 4ECE，
丛=叢cóng，苁，枞cōng；
众=眾zhòng.

A man, a figure in profile. The original form of 人 rén may reveal evolution from anthropoid apes to man. ☞ 亻 4EBB as radical.

人的侧视形。仌，古文虞。夵，古文盗。

tǐng

壬 甲 〈figures〉 篆 〈figure〉 參 〈figures〉

2123C — good; erect (archaic).

聽 807D.tīng: listen；廷 5EF7，
庭，莛，蜓，霆tíng；挺，
艇，铤，梃tǐng.

A man standing on land. 挺tǐng: to erect; with a hand — radical which indicates a verb. 壬 58EC.rén is very close to 壬, and also a useful element in 妊，衽，任rèn，荏rěn. ☞ rén.1, tǔ.2.

人挺立于土地上。

wù

兀 甲 〈figures〉 金 〈figure〉 古 〈figure〉 篆 〈figure〉

5140 — proud; upright; bald.

阢wù；堯 582F=尧 5C27，嶢yáo；
浇jiāo；侥jiǎo；挠，蛲náo；硗，
跷qiāo；翘qiáo；娆，桡，饶ráo；
绕rào；烧shāo；骁xiāo；晓xiǎo.

A derivative of 人 rén.1 by adding a stroke on the man's head. 尧 yáo, a legendary monarch in ancient China around 2300 B.C., is also a useful phonetic.

人上一横。《说文》兀，高而上平也。从一在人上。尧，高也。从垚在兀上，高远也。机，木无枝。檽机，篆作檽柮。杌陧，不安貌，篆作黜隉、阢隍。

yuán

元 甲 〈figures〉 金 〈figures〉 篆 〈figure〉

5143 — basic, first, primary.

阮ruán；玩，顽，完5B8C，
烷wán；莞，皖，脘wǎn；沅，
园=園，鼋yuán；远=遠yuǎn；
垸，院yuàn.

Two strokes drawn on a man's head; a composite of the characters for man and above. ☞ 上shàng.195, 人rén.1.

元字上从人头形，所以元即首（头）。元首二字重文迭义。"原来"一词原作"元来"，明灭元后，汉人忌讳"元"来，改为"原"来。