Dating the Origin of Chinese Writing: Evidence from Oracle Bone Inscriptions

Kuang Yu Chen, Ph.D.
Rutgers University

ABSTRACT
The convergence of a visual sign containing semantic value with an audio symbol signifies the creation of writing. Writing was independently created at least four times in human history, hence, Sumerian, Egyptian, Chinese and Olmec/Mayan are considered the four original or pristine writings. The four pristine writings are all logographic and logosyllabic, containing both visual and phonetic elements. Among them, Chinese remains viable and logographic, the other three died out hundreds or thousands year ago. The dating of the origin of both Sumerian and Egyptian writing to about 3200 BCE is supported by strong archaeological evidence. The Mayan writing clearly was derived from the earlier Olmec writing, which could be dated with archaeological evidence to about 600 BCE or earlier. Although Chinese is the only surviving pristine writing, its origin has not been definitively dated. The earliest Chinese writing available today is OBI of late Shang (ca. 1300-1046 BCE), which was already a fully mature writing system. Clearly, the origin of Chinese would have to be much earlier than the time of late Shang. In the absence of definitive archaeological evidence, we propose two alternative approaches for dating the origin of Chinese writing, one based on the structural continuity of Chinese script, from OBI to modern Chinese, and the other based on the pre-Dynastic personal names identified in OBI. The role of taowen (pottery signs) in dating the origin of Chinese is also discussed.

1. DEFINITION OF WRITING SYSTEM
In contrast to spoken language, which is temporal and physiological, written language, expressed with scripts, is atemporal and mechanical, and requires material medium and space. With writing, a temporal speech can be transfixed into atemporal script on a medium that can be read and transmitted to future. It is generally agreed that writing arose ex nihilo only four times in human history: in Egypt, in Mesoptamia, in China, and in Mesoamerica. To determine whether a graph can be considered as a script, we will need to have a clear definition of writing. The following list a set of rules for clearly defining writing.

1) A graph (G) is defined as an instance of writing if it has a conventionally associated pronunciation (P), whether or not that pronunciation is accompanied by an associated meaning (S), thus G:{+P, ±S}.^2

2) The earliest evidence for writing, wherever it was invented ex nihilo, tends to show graphs that have an associated pronunciation and a correlated meaning, viz., G:{+P, +S}.

3) The {+P} feature alone defines them as writing; the {+S} feature is explained partly by the empirical fact that this is what the archaeological record of the earliest known evidence for writing shows and partly by the inference that if writing arose from a functional use of pre-writing graphs, those graphs must have had some meaning to start with. The questions is,

---

^1 We have adopted the symbols and argument put forward in Boltz (2003).

^2 Although G:{+P, -S} (i.e. G(+P)) can be considered as a script, we will not consider it in this paper.
when and how did that fusion of (+S) and (+P) happen.

For our purpose, we will define a writing as: \( G(\text{script}) = G: (+P,+S) \). The genesis of writing can be expressed by the transition from graph to script: \( G: (+S) \rightarrow G: (+P,+S) \). Here \( G \) represents a graphic sign and \( G: (+S) \) denotes a graph that contains ‘sense’ and information content. \( G: (+P,\pm S) \) represents a script, i.e. a graph that contains ‘sense’ and ‘phonetic value’. Thus, marks, signs, icons, graphs, glyphs, and scripts can all be considered to be \( G: (+S) \). But, it is only when a phonetic element is incorporated into the graph \( G: (+S) \), we have a writing, defined as \( G: (+P,+S) \).

2. DATING OF THE PRISTINE WRITINGS

The four pristine writing systems are Sumerian cuneiform of Tigris and the Euphrates Rivers, Egyptian hieroglyph of Nile River, Olmec/Mayan of Mesoamerica, and Chinese OBI of Yellow River. Whether or not Harappan of Indus River Valley represents writing is still controversial. Table 1 summarizes the key features of the four pristine writing systems.

<table>
<thead>
<tr>
<th>Writing</th>
<th>Time</th>
<th># of signs</th>
<th># of syllabus</th>
<th>Lifespan (yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sumerian</td>
<td>3200-1800 BCE</td>
<td>600</td>
<td>150</td>
<td>1500</td>
</tr>
<tr>
<td>Egyptian</td>
<td>3200-500 BCE</td>
<td>700</td>
<td>100</td>
<td>2700</td>
</tr>
<tr>
<td>Chinese*</td>
<td>2100 BCE – Now</td>
<td>50000</td>
<td>62</td>
<td>&gt;4000</td>
</tr>
<tr>
<td>Mayan**</td>
<td>600 BCE -1500 CE</td>
<td>500</td>
<td>50</td>
<td>2100</td>
</tr>
</tbody>
</table>

*The origin of Chinese writing was estimated to be no later than 2100 BCE (see below).
**The origin of Mayan was based on recent archeological evidence on Olmec civilization.

2.1 Sumerian Writing
Sumerian writing emerged at the Uruk region of southern part of modern Iraq, the fertile land of Tigris and Euphrates. Triangular wedge-shaped reeds were used to write on moist clay to produce inscriptions in the shape of cuneiform, hence the namesake. The earliest collection of archaeological materials related to Sumerian writing consists of a large number of pictographic inscription tablets from the Uruk IV and III periods (ca. 3200-3000 BCE), termed “proto-cuneiform” texts, as the graphs have not yet assumed the distinctive wedge-shaped marks. Fig. 1 shows one example each from Uruk IV and III period. The Uruk IV tablet contains signs for numerals and objects, but we are not sure whether they represented scripts. In the Uruk III tablet, in addition to signs for numerals and objects, there was a composite sign at the bottom left, which also appeared in many other tablets. Scholars identified it as a personal name of an official in charge of storage facility of supplies for making beer. As a personal name, the composite sign must have pronunciation and thus qualified to be considered as a script \( G: (+P,+S) \). Thus, Sumerian writing should have been invented no later than Uruk III period, about 3200 BCE.

2.2 Egyptian Hieroglyphic Writing
The history of ancient Egypt, spanning 3000 years, started from Dynasty Zero (3200 to 3000 BCE) to 343 BCE when Egypt fell under Greek rule. Egyptian writing, already evident in 3000 BCE, lasted throughout the thirty-one dynasties, including the Greek and the Roman rule, and finally died out.

---

3 For a discussion on this issue, see Lawler (2004).
4 For a more detailed discussion on these two clay tablets, see Nissen et al. (1993).
5 The phonetic value of this composite sign was assigned as *ku-sim*. The name Kushim was probably the first personal name known in human history. See Nissen et al. (1993).
2.3 Olmec-Mayan Writing

Among the major Mesoamerican civilizations, Maya has the longest history and covers the largest area, including part of the modern Mexico, Guatemala, Belize, west part of Honduras, and part of El Salvador. Maya writing was widely used in Mesoamerica between 300 BCE and CE 1600. The Spanish colonization and Christianization of Mesoamerican led to the burning of thousands of Mayan codices and extinction of Mayan hieroglyphic and literary tradition. Fortunately, four Mayan codices survived the fire and found way into Europe. In addition, Mayan hieroglyphics can still be found on hundreds and thousands of objects, stela, stairs, and lintels in archeological ruins and several Mayan dialects are still spoken by millions of Mayan Indians. Yuri Knorosov (1922-1999), a Russian scholar, deciphered Mayan hieroglyphics in early 1950’s by showing convincingly that Mayan writing is logosyllabic. Although classical Mayan hieroglyphic writing was generally thought to start around CE 200, the Olmec-Maya writing can be traced back many centuries to pre-Classic Period. One of the earliest archaeological evidence was a greenstone seal excavated from La Venta in Tabasco, Mexico dated to about 650 BCE (Fig. 3A).\(^7\) The cylindrical seal contains the raised image of a bird with two signs emerging from its mouth. By comparing these signs to Maya hieroglyphs, Pohl and co-workers concluded that the seal image depicts that the bird symbolically pronounced, "King 3 Ajaw." In the Mayan language, the word ‘Ajaw’ could mean either ‘lord’ or a ‘day name’ in the 260-day calendar. It is highly likely that ‘3 Ajaw’ refers to the name of an Olmec ruler. This finding firmly established that writing in the La Venta polity has already existed at 650 BCE. Another archaeological evidence is a stone slab, named the ‘Cascajal block’, uncovered by road builders in 1999 while digging up an ancient mound at Cascajal, San Lorenzo (Fig. 3B). Its text consists of 62 signs, some of which are repeated up to four times. Some Mayan scholars suggested that they might represent Olmec glyphs, because the text “conforms to all expectations of writing”, because of its distinct

---

\(^6\) Baines, J. (2004)  
\(^7\) Pohl, M et al. (2002)
elements, patterns of sequencing, and consistent reading order. If their conclusion were right, it would push the origin of Olmec-Mayan writing to 900 BCE or earlier.

Most of the engraved signs on the surface of pottery, clay, wood or stone objects of neolithic time belong to the category of \(G:\{+S\}\), meaning they contained certain information content and served the function of non-verbal communication. As briefly described above, the archaeological evidence for dating Sumerian, Egyptian, and Olmec/Mayan writing suggested that the need for combining verbal communication with non-verbal signs might have started first with the symbolic pronunciation of personal name, be an official or a king like Kushim, Scorpion, and 3 Ajaw.

3. DATING OF CHINESE WRITING
The earliest archeological evidence of Chinese writing is the Shang oracle bone inscriptions (OBI), divination records inscribed on hundreds and thousands of turtle shell and animal bones. Almost all of them were discovered in Xiaotun (小屯), the last capital of Shang Dynasty (ca. 1600-1046 BCE). Si Maqian (司馬遷), the grand historian, wrote in the Shiji (史記, The Record of History): “King Ban Geng (般庚) moved the capital (to Xiaotun), from that time on, for the next 273 years until the demise of King Zhou (紂), the Shang capital did not change again.” Most oracle bones and shells belonged to the royal house of late Shang Dynasty (1300-1046 BCE) and dealt with various aspects of royal activities, from weather, harvest, hunting, war, to ancestral worship. The large corpus of OBI, over 150,000 pieces, provided a rich source for studying Shang history. The OBI writing contains more than 4000 scripts and reveals sophisticated grammar and lexical structures, not much different from ancient classical Chinese. As OBI is already fully mature, the origin of Chinese writing must have occurred before 1300 BCE. This brings our attention to taowen 陶文, signs or marks inscribed on pottery utensils, stone equipment or even shells and bones. Some of the neolithic taowen have been suggested to be writing, but the paucity of samples making the argument not convincing. In the absence of direct archeological evidence for Chinese writing prior to 1300 BCE, we began to explore other approaches in order to seek for indirect evidence for the earlier Chinese writing.

3.1 Fossils scripts for personal names in OBI
Among the archeological evidence for Sumerian, Egyptian and Olmec-Mayan writing, the personal name identified on pottery and other objects stood out prominently. For Chinese writing, no such evidence has been identified yet, despite the discovery of so many neolithic culture ruins all over China. On the other hand, many pre-Shang figures were mentioned in the transmitted texts, considered legendary because of absence of direct archeological proof of their existence. The discovery of Shang OBI provided archaeological evidence for the existence of all Shang kings and many high officials recorded in the Shiji and other ancient texts. Thus, if names of some pre-Shang or legendary figures could be identified in OBI, these names could be viewed as fossil scripts buried in the Shang OBI corpus, just like the name Narmer and Scorpion inscribed on pottery vase. Below, we listed some pre-Shang people whose names could be identified in OBI.

(1) The time of Cheng Tang 成唐 (ca. 1650 BCE)

---

8 Martinez et al. (2006)
9 For whether Cascajal Block can be considered as writing see Zhang He’s and Marin Mora’s papers in this volume.
10 A brief discussion on OBI corpus see Song Zhenhao in this volume.
11 With regard to the use of taowen to determine the origin of Chinese writing see Wang Hui in this volume.
The king list of the 31 Shang kings, from the founding king Cheng Tang 成唐 to the last king Zhou 紂, discovered in OBI showed almost identical sequence, in terms of enthronement order to that recorded in the *Shiji* and the *Zushujinian* 竹書紀年. All the 31 kings were historical figures and, as such, the name script of the founding king, Cheng Tang 成唐, must have existed at his time, about 1650 BCE. In addition to the temple names of kings and queens in early Shang Dynasty recorded in OBI, one famous OBI bone there listed the lineage of the Er 児氏 family for 13 generations. The bone was dated to about 1300 BCE. The name of the first member of the Er family listed was Cui 吹, his time would be 13 generations before 1300 BCE, i.e. about 1690 BCE. As such, the script for his personal name, Chui 吹, must have existed at that time.

(2) The time of Shang Jia 上甲 (ca. 1800 BCE)
According to OBI, the five-ritual cycle was performed annually during the late Shang Dynasty. The ritual cycle started with Shang Jia 上甲, known as the first pre-Dynastic king. Altogether there were six pre-Dynastic kings. The personal names of Shang Jia’s father and uncle, Wang Hai 王亥 and Wang Heng 王恒, respectively, were also found in OBI. Moreover, Wang Hai’s name was often preceded with the title Gaozu 高祖, the High Ancestor, signifying Wang Hai’s high status along the blood lineage. Thus, Shang Jia 上甲 and other pre-Dynastic kings and Wang Hai, Wang Heng should all be viewed as historic figures. As such, their name scripts, recorded in ancient text like the *Zushujinian* and now identified in OBI corpus, must have existed at the time of Shang Jia, around and before 1800 BCE.

(3) The time of Di Ku 帝嚳 (ca, 2100 BCE)
Wang Guowei identified the OBI script 帝嚳 as the name of Di Ku and read as kui 吹 or nao 藹.

According to transmitted texts, Di Ku 帝嚳 was the father of Qi 契, the founder of Shang clan. Thus, in the *Annals of Yin* of the *Shiji*, it recorded:

Qi of Yin’s mother was Jian Di, who was one of the daughters of Yousong and the secondary wife of Emperor Ku (Di Ku). She was going with her two sisters to bathe, when she saw a dark bird drop its egg. Jian Di picked it up, and swallowed it, and thus being with child gave birth to Qi.

The poem of *Xuan Niao* 玄鳥 in the *Shijing* 詩經 also read:

Heaven commissioned the swallow,
To descend and give birth to [the father of our] Shang.
[His descendants] dwelt in the land of Yin, and became great

Although Di Ku and Qi were mentioned in the transmitted texts, the finding of their names in OBI as high ancestors receiving lavish sacrificial ceremonies demonstrated that indeed they occupied high status in Shang lineage and should be viewed as historical personages. According to the *Shiji*, Qi

---

12 For a review on the Er Family Lineage OBI see Chen (2006)
was seven generations ahead of Wang Hai, the father of Shang Jia. Thus, Di Ku’s time should be around 2100 BCE. As such, the scripts for their personal names must have existed at their time. The script鹿was used only as the name for Di Ku, and could be viewed as a ‘fossil’ script passed down from 2100 BCE and buried in Shang OBI.

### 3.2 Mathematical Model
The four pristine writings are all logographic, meaning that the script is composed of three elements: shape/form (形), sense/meaning (义), sound/phonetic (音). Among them, only Chinese survives into modern age and remains logographic. The longevity and continuity of Chinese writing, from OBI to modern Chinese, is striking. Thus, this structural continuity allows us to analyze the epigraphic structure of Chinese scripts over a time span of 3000 years. Such analysis could provide clues on whether any structural element exhibits a time-dependent change over this time span. If a change can be identified, its time-dependent nature should enable us to trace such change to a much earlier time point than Shang Dynasty and thus may allow us to get a glimpse of the development of Chinese scripts at the earlier stage.

Traditional Chinese lexicography divided logographic scripts into six categories according to the six principles/rules (六書) on how they were constructed. Although the term六書appeared in the late Zhou time (ca. 400 BCE), Xu Shen 許慎 (ca. CE 58-147), the great philologist and lexicographer of East Han Dynasty, was the first to use the six书 principles to analyze over 9000 Chinese scripts and published the Shuowen Jiezi 說文解字, the first comprehensive Chinese dictionary. The六书 are: pictograph (象形 ‘imitating form’), simple ideograph (指事 ‘indicating event’), semantograph/composite ideograph (會意 ‘joining meaning’), loangraph (假借 ‘borrowing/rebus’), phono-semantic/logosyllabic graph (形聲 ‘form and sound’), and polyvalent (轉注 ‘reciprocal meaning’).

Although all logographic Chinese scripts contain phonetic element, there exists a difference in that whether the phonetic element is ‘visible’ or ‘invisible’. How do we differentiate visible from invisible phonetic elements? For example, the script鹿鹿 is a pictograph for deer. From the pictograph itself, we will not be able to tell its phonetic value is鹿, and hence the phonetic value is invisible. However, for the composite graph鹿（windlass, hauler), since the component車contains the semantic value whereas the component鹿 serves as a phonetic element, therefore the phonetic element is visible. Xu Shen described the evolution of Chinese script in the Shuowen Jiezi 說文解字 this way: When Cang Jie first invented writing, he created graphic forms (pictographs, 象形) according to categories; therefore these were called文(patterns). After that, forms and sound (形聲) mutually augmented each other; these were called字. The so-called文depicts objects; and so-called字, indicates propagating of scripts like spawn. 載頒之初作書，蓋依類象形，故謂之文。其後形聲相益，即謂之字。文者，物象之本；字者，言孳乳而漫多也。Since both文and字 are all script, they have to have both semantic and phonetic value. The difference is that the phonetic value for a文-script is embedded and thus invisible, whereas the phonetic value for a字-script is visible. We can use the symbol G:{+P_A, +S} to represent文, i.e. the script with invisible phonetic elements (+P_A) and the symbol G:{+P_B, +S} for字, the script
containing visible phonetic elements (+Ps). As such, the genesis and evolution process of Chinese writing can be expressed in a two-stage process: G:{+S} → G:{+Pa, +S} → G:{+Pb, +S}, where G:{+S} represents signs (fu 符) that were used for mnemonic and communication purpose before the invention of writing. The script created in the first stage: G:{+S} → G:{+Pa, +S} was limited in its function and ability in recording spoken language. As civilization progressed, the need to expand the corpus of lexicons and scripts to meet more sophisticated social demand, the evolution of Chinese writing entered the second stage: G:{+Pa, +S} → G:{+Pb, +S}. The script created in the second stage was called zi 字; literally, it means "sons of wen-scripts (i.e. pictographs and ideograms)."

According to the liushu principles as we discussed above, pictograph (xiangshing), simple ideograph (zhishi), and composite ideograph should all be consider to be in the category of G:{+Pa, +S}, scripts with invisible phonetic element. In contrast, logosyllabic graphs (xingsheng) are in the category of G:{+Ps, +S}, scripts containing visible phonetic elements. As to the rebus script (jiajie) and polyvalent script (zhuangzhu), they are more related to the extended use of existing script for additional function. In other word, their phonetic value as a rebus script (or a polyvalent script) depends on the context. As such we will not include them into the category of G:{+Pb, +S}.

Archeological excavations over the past century have yielded a rich corpus of Chinese scripts of pre-Qin times, including OBI (ca. 1300-1100 BCE), bronze inscriptions (ca. 1200-300 BCE), and bamboo scripts of the Warring States periods (ca. 400-200 BCE). With the pre-Qin scripts from various sources, we have in our hand Chinese scripts, from OBI to modern Chinese, across a time scale of 3000 years. This large corpus of scripts, created with the same liushu principles, could allow us to perform the liushu analysis to trace the change in the liushu categories over a 3000-year time span. We have previously assembled data from this kind of analysis and showed that there was a linear increase in the percentage of scripts containing phonetic elements, which we included scripts of xingsheng 形聲, jiajie 假借 and zhuangzhu 轉注 scripts. As discussed above, the phonetic value of jiajie 假借 and zhuangzhu 轉注 is context-dependent, and as such, they should not be considered in the category of G:{+Pb, +S}. Table 2 listed the liushu analysis of the script corpus at six different chronological time periods.

<table>
<thead>
<tr>
<th>date</th>
<th>~1300 BCE</th>
<th>~1000 BCE</th>
<th>~800 BCE</th>
<th>200 BCE</th>
<th>100 AD</th>
<th>1100 AD</th>
<th>bamboo</th>
<th>Shuowen</th>
</tr>
</thead>
<tbody>
<tr>
<td>liushu script %</td>
<td>Liu</td>
<td>Zhang</td>
<td>Zhu, Q</td>
<td>Zhu, M</td>
<td>Zhang</td>
<td>Zhang</td>
<td>Zhu, J</td>
<td>Zheng</td>
</tr>
<tr>
<td>xiangxing %</td>
<td>23</td>
<td>28</td>
<td>29</td>
<td>23.2</td>
<td>11</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>zhishi %</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2.2</td>
<td>1</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>huiyi %</td>
<td>32</td>
<td>46</td>
<td>47</td>
<td>35</td>
<td>25</td>
<td>12</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

14 For OBI analysis, we have included three independent studies here. More research has been done in this area; for example, Cao Jun 曹君: Yinshang jiaowen xingshengzi yanjiu 《殷商甲骨文形聲字研究》2010 廣州大學碩士論文; Yang Junhui 楊軍惠 Jiaowen xingshengzi yanjiu 《甲骨文形聲字研究》2009 浙江師範大學碩士論文. In all studies of the xingsheng analysis of OBI scripts, the value centered around 22%. 
Fig. 4 shows the plot of percentages of $G:\{+PB, +S\}$ against the chronological time. The linear increase of $G:\{+PB, +S\}$ is striking and allows us to apply extrapolation to estimate the time period that the transition step $G:\{+PA, +S\} \rightarrow G:\{+PB, +S\}$ started no later than 2100 BCE. It then goes without saying that the genesis step $G:\{+S\} \rightarrow G:\{+PA, +S\}$ of Chinese writing must have occurred before 2100 BCE, probably by several centuries. It is worth noting that this time point, estimated from a simple mathematical model coincides with the time of earliest ‘fossil script’ uncovered and identified from OBI.

### 3.3 Archaeological evidence

The simple mathematical model presented in Fig. 4 can predict the starting time for the second stage of evolution of Chinese writing, but not for the first stage $G:\{+S\} \rightarrow G:\{+PA, +S\}$. To determine when the first stage occurred, we have to rely on archaeological evidence. For the other three pristine writings, we have direct archaeological evidence to show that the beginning of Sumerian and Egyptian writing occurred somewhere around or before 3100 BCE. But for Chinese writing, despite the fact that we have unearthed hundreds and thousands bones and shells with the late Shang inscriptions; we do not have concrete direct archaeological evidence for the existence of writing before the late Shang.

Pottery marks, referring to etched marks or signs on the pottery utensils, stone equipment or even shells and bones, were found in Neolithic cultural sites almost all over of China. Scholars traditionally called pottery inscriptions as *taowen* 陶文, meaning pottery graph. Although some pottery marks, like those from the Dawenkou sites, have been suggested to be proto-writings, the direct precursor of Chinese writing, the general consensus is that there is no hard evidence indicating any of these pottery signs truly represents writing. A comparison of Chinese *taowen* with pottery signs from Mesopotamia, Egypt and Mesoamerica at the genesis stage may provide clues on whether any of the Chinese *taowen* could be considered as candidates for Chinese proto-writing. With this in mind we will examine the Dawenkou *taowen*, Liangzhu *taowen*, Dinggongcun 丁公村 *taowen* and Taosi 陶寺 *taowen* and assess their role in the genesis and evolution of Chinese writing.

(i) Dawenkou *taowen*: Fig. 5 shows pottery marks from Dawenkou. In addition to Dawenkou, similar pottery marks were also discovered in Lingyanghe 陵陽河, Qianzhai 前寨, Dazhuijucun 大朱家村, and Hangtou 杭頭, all in modern Shandong Province, and in Yuchisi 尉遲寺 and Meng Cheng 蒙城 of Anhui Province, some 500 km from Dawenkou. So far, a total of about twenty individual signs have been discovered, dated to ca. 2800-2600 BCE. The signs were usually incised on the surface of
large-mouth pottery urns called *zun* 墟, mostly on the upper portion of the bodies, only a few at the bottom. Each urn bore a single sign, occasionally two signs, one up and one at the bottom, or one on each side of the upper body. Several signs still retain vermillion pigment, indicating that the urns probably served not just as wine vessels but more for some ritual functions, similar to the ritual bronze vessels of later Shang period. Some scholars speculated that the Dawenkou *taowen* could represent clan emblem inscription, like that on bronze vessels. Since we have no proof that any of the Dawenkou *taowen* contains phonetic value (like being used as a personal name), we still cannot consider them as scripts.\(^\text{15}\)

(ii) Liangzhu *taowen*: The neolithic Liangzhu 良渚 culture (ca. 2800-1900 BCE) covered an area of modern Jiansu and Zhejiang Province. All Liangzhu sites produced pottery marks. Several examples are listed in Fig. 6. Fig. 6A shows a black pottery from Chenghu 澄湖, Wu County 吳縣. It bears four signs that resembled OBI scripts, but the meaning is not decipherable.\(^\text{16}\) Fig. 6B shows a pottery shard with 9 signs, discovered in Longqiuizhuang. The style of the Longqiuizhuang signs is very different from that of the Chenghu pottery.\(^\text{17}\) Unfortunately, despite large number of pottery marks, samples with script-like signs are very few, making it difficult to draw any conclusion with regard to their function and meaning. Liangzhu culture is unique in its production and use of jade objects. At least 11 Liangzhu *bi* jade discs and four jade *cong* bear incised emblem-like motif associated with birds (Fig. 7A). They are all composite sign with a bird standing atop an inverted trapezoid (altar?) with additional sign inside the trapezoid. We do not know whether these bird-motif signs have anything to do with proto-writing. But, interestingly, when we compare these Liangzhu bird-motif signs with the Egyptian pharoah’s Horus serekh signs (Fig. 7B), the similarity is striking.\(^\text{18}\) Recall the close connection of the bird-motif with the origin of Egyptian (Fig. 2A and Fig. 7B) and Olmec hieroglyphics (Fig. 4A) and the Shang legend about their founding king Qi,\(^\text{19}\) it is tempting to speculate that the signs inside the trapezoid might have function of writing.

(iii) Dinggongcun *taowen*: The famous Dinggongcun *taowen* (Fig. 8) was unearthed in 1991-1992 from a pit (dated 2200-2100 BCE) at Dinggongcun, Zouping county 鄒平縣, Shandong.\(^\text{20}\) The inscription was not immediately recognized by the archaeologist, but by a worker while cleaning these shards in the laboratory. The eleven cursive signs were arranged vertically into five columns. It is not possible to establish any shape/form connection between these signs and OBI scripts. So far this is only sample bearing this type of cursive signs. Feng Shi suggested that these eleven signs represented the ancient divination record of Yi 羲 people.\(^\text{21}\) But without any knowledge of the ancient Yi language and without a critical amount of samples, we can only say that Dinggongcun *taowen* probably represents traces of some early proto-writing, which died out long before Shang Dynasty.

\(^\text{15}\) Scholarly work on Dawenkou *taowen* abound. For a recent review of the earlier work see Kang (2013).
\(^\text{18}\) Such comparison was first reported in Chen, K.Y. (2008a)
\(^\text{19}\) This legend about Qi’s birth after his mother Jian Di swallowed the egg of dark bird was described in the *Annals of Yin* of the *Shiji* cited above.
\(^\text{20}\) For the discovery and research of Dinggongcun *taowen* see Luan, F. (1994).
(iv) Taosi taowen: The Taosi 陶寺 cultural site at Xiangfen 襄汾, Shanxi 山西, dated to about 2500-1900 BCE, covers an area of 30,000 sq meter with about 1300 tombs; only 1% contains wooden coffins containing burial objects including painted pottery basins with coiled dragon designs, wooden drums, jade, and stone musical instrument qing 磬. Fig. 9 shows a pottery shard containing two signs painted by brush in vermilion color. Indeed these two signs do look like Chinese script. Particularly the first sign was almost identical to the script wen in OBI. Ge Yinhui interpreted these two signs as wen yao 文堯, the name of a legendary king, who abdicated his throne to Da Yu 文, the founding king of Xia Dynasty. Feng Shi read these two signs as 文邑, the name of capital of Xia Dynasty. Unfortunately, this shard is the only taowen sample available from Taosi site, making it difficult to draw definite conclusion about its relation to the genesis of Chinese writing.22

So far, the archaeological findings, particularly pottery marks found in many neolithic sites, do not provide concrete evidence for dating the origin of Chinese writing. Nonetheless, the fact that some pottery marks exhibited more than two signs arranged together in strings suggests that these signs could be lexical. These string-signs were not decipherable and their connection to Chinese writing was unclear. If the Dinggongcun taowen and Longqiuzhuang taowen are not related to OBI, they may still represent some early non-Chinese proto-writings. In that case, it would give a tangible evidence that by the time of late Longshan Culture period (2300-2000 BCE), some forms of writings might have existed in China.

4. FUNNEL MODEL

Pottery marks (taowen) were discovered in almost all neolithic culture sites across China, covering an area much larger than that of Egypt, Mesopotamia, and Mesoamerica combined. Most of pottery marks existed as a single sign on the object. However, after about 2500 BCE, multiple signs (two or more in string) signs appeared in Longshan culture sites and Liangzhu culture sites. These cultural sites spanned a time period from 5000 BCE to early Shang. Almost with no exception, these cultural sites were near Yellow River, Yangtze River and Huai River, or near the tibutary rivers and lakes that connecting them. As discussed above, multiple signs like the Dinggongcun taowen, the Liangzhu taowen, and Longqiuzhuang taowen were sufficiently long to represent something lexical. As Dinggongcun taowen and Longqiuzhuang taowen did not resemble OBI in form, they might represent some non-Chinese early writings. In light of the large spatio-temporal distribution of pottery marks, it is possible that in ancient China during the third millennium BCE there co-existed several proto-writing systems, and the proto-OBI (i.e. proto-Chinese writing) was only one of them. The extensive waterways connecting various neolithic cultural sites would allow interactions among people with different languages and writing systems and as such, through competition, absorption, modification and elimination, eventually the proto-Chinese writing survived into pre-Shang and Shang Dynasty and became what was known as OBI. This scenario appears similar to the process of protein folding in cell biology, where immediately after a protein molecule was translated off ribosome, it existed as a nascent polypeptide with only linear secondary structure, but with capacity of assuming many different conformations. Throughout a trial-and-error process including mutual interaction, absorption, modification, and elimination, many meta-stable conformations were ‘funneled’ out, until eventually the polypeptide folded into a most stable (with the lowest free-energy state) and versatile tertiary structure as a mature protein. Scientists have used the energy funnel model to illustrate this protein stabilization process. We propose here that a similar funnel model can be used to illustrate the selection and eventual stabilization of proto-Chinese writing.23

---

22 With regard to the discussion of the Taosi taowen, see Feng, S. (2008); Ge, Y. (2007).
5. CONCLUSION

Human communication has relied on two separate and parallel tools: visual signs and audio symbols. Spoken language has been used for 100,000 years and pictographs/signs have been around for over 30,000 years. The gestation and genesis of writing began with the fusion of these two communication tools. Thus the first script was created by associating a fixed sound with a pictograph: \( G:{+S} \rightarrow G:{+PA}, +S \). This genesis/gestation step independently occurred four times in human history, in Egypt, Mesopotamia, Mesoamerica and China. The four original pristine writings, Egyptian hieroglyphic, Sumerian cuneiform, Olmec/Mayan hieroglyphic, and Chinese OBI, were all logographic. Sumerian cuneiform and Egyptian hieroglyphic were eventually replaced first with syllabic then alphabetic spelling systems. Mayan hieroglyphic died out because of colonization. Today, all the writing systems in the world, except Chinese, are either alphabetic or syllabic. The following diagram illustrates these two different evolutionary paths, one for logographic Chinese and the other for the rest of writing systems.

**Chinese path:**

\[
G:{+S} \rightarrow G:{+PA}, +S \rightarrow G:{+PB}, +S, \]

icon \hspace{1cm} logographic script \hspace{1cm} logosyllabic script (word)

**Other path:**

\[
G:{+S} \rightarrow G:{+PA}, +S \rightarrow G:{+PB1}, +S \rightarrow [G:{+PB}, +S],_n \rightarrow G:{+PB2}, +S \]

icon \hspace{1cm} logographic script \hspace{1cm} alphabet \hspace{1cm} word \hspace{1cm} syllable

In this diagram, the transition from icon \( G:{+S} \) to logographic script \( G:{+PA}, +S \) signifies the genesis of writing. Direct archaeological evidence shows this transition occurred in Sumer and Egypt around 3200 BCE. For the Chinese writing, so far no concrete evidence is available for dating this transition. Instead, we show here that based on the continuity of structural feature of Chinese scripts, we can use a simple mathematical model to estimate the beginning time of the the transition step from logographic to logosyllabic scripts. As shown in Fig. 4, this step began at about 2100 BCE. The names of some pre-Dynastic Shang legendary heroes have been identified in OBI. Among them, Nao (Di Ku) was the earliest one, dated to ca. 2100 BCE. The name script for Nao can be viewed as a fossil script buried in the OBI corpus. Thus, we argue that this script, hence writing, must have existed at the time of Nao, around 2100 BCE. It may not be a coincidence that the estimate from both mathematical model and name scripts pointed to 2100 BCE as the time when the transition step \( G:{+PA}, +S \rightarrow G:{+PB}, +S \) began.

In the absence of direct archaeological evidence, when, where, and how the first transition step occurred in China remains murky. We have proposed to use the funnel model to illustrate how the proto-Chinese writing could have come about through competing with other contemporary proto-writings (like Dinggongcun *taowen*), but we have no idea how many decades, or centuries that process took. We can only hope that future arcaheological findings will enable us to answer all these questions eventually.

6. REFERENCES

Baines, John * The earliest Egyptian writing: development, context, purpose* in “First
Chen, Kuang Yu 陳光宇 (2008a) On the origin of Chinese writing and comparison with other original writings, Wen Bo, 《試論漢字起源定點與世界古文字溯源比較》《文博》4, 26-34. (in Chinese)
Chen, Kuang Yu 陳光宇《從甲骨文推測漢字起源與發展模式》，《首届中国文字发展论坛暨纪念甲骨文发现110周年学术研讨会论文集》，2009年《中国文字博物馆第二期》第9-18頁。
Chen, Kuang Yu 陳光宇《兒氏家譜刻辭綜述及其確為真品的證據》《甲骨文與殷商史-新六輯》第 267-297 頁，2016年。
Coe, M.D. Breaking the Maya Code, Thames and Hudson, Inc. 1992
Feng Shi 馮時《山東丁公龍山時代文字解讀》，《考古》1994年第1期，第37-54頁。
Feng, Shi 馮時《試論文邑，見《文邑考》《考古學報》第3期，第273-290頁，2008年。
Ge, Yinghui 葛英會《破譯帝堯名號》《古代文明研究通訊》第32期，2007年
Kang, Hanyu 康翰予《文明史視野中的大汶口文化陶文研究——兼與良渚文化刻符比較》（2013年煙台大學碩士論文）
Lawler, Andrew The Indus Script—write or Wrong? Science 306, 2026-2029, 2004
Li, Xiaoding 李孝定 The Origin and Development of Chinese Writing in History of Chinese Writing《中國文字的原始與演變》《漢字史話》聯經出版社，1977
Luan, Fengshi 欧豐實《丁公龍山城址和龍山文字的發現及其意義》《文史哲》1994年第3期，第85-89頁。
Rao Zongyi 饒宗頤《讀高郵龍虬莊陶片的刻畫圖文》，《東南文化》1996年第4期
Wang, Guowei 王國維 Yin buci zhong suojian xianwang xiangong kao, Guantang Jilin vol. 9《殷卜辭中所見先公先王考》，《觀堂集林》卷 9，中華書局，1959


Zhang Su 張溯 On the Incised Liangzhu Pottery Marks from Chenghu, Jiangsu, Dongnan Wenhua 《論江蘇澄湖遺址出土的良渚刻符》《東南文化》2015 年第 5 期，頁 68-74。
Figure 1 Sumerian clay tablets from Uruk sites

Figure 2. A) U-j vase incised with serekh sign of Narmer. B) A parade scene on Narmer pallete. C) a U-j vase incised with a scorpion sign. D) An incised scene on a macehead.
Figure 3. A) The San Andres ceramic roller stamp. B) The Cascajal Block.

Figure 4. Plot of percentages of G:{+P, +S} in Chinese scripts at different chronological time.

\[ y = 0.0374x + 73.931 \]

\[ R^2 = 0.9649 \]
Figure 5. (Left) The Dawenkou vase with sign d incised at the center. (Right) from a-h, different single signs incised on Dawenkou vases.

Figure 6. A) Incised marks on a black pottery jar from Chenghu. B) Pottery marks on a shard from Longqiu Zhuang.
Figure 7. A) The bird-altar motif inscribed on Liangzhu jade *bi* discs. B) The horus-serekh motif inscribed on Egyptian U-j pottery urns.

Figure 8. Dinggongcun *taowen*. Inscribed pottery marks on a pottery shard from Dinggongcun site.
Figure 9. A) The Taosi jar with red-ink marks. B) The full view of the two marks on the Taosi jar.