4. PHILOLOGICAL STUDIES, PALEOGRAPHY, AND ETYMOLOGY

“Philology” generally refers to the critical study of texts and the languages they use, and of methodologies for the historical interpretation of texts. Philology may be seen as a foundational form of study; its goals are to resolve issues concerning the nature of textual evidence, so that literary, philosophical, and historical theories based on texts are less likely to be undermined by misdating or misreading of the textual data. Philological studies include such subfields as etymology (principally focused on the development of characters and their semantic basis), paleography (the study of “old” [Greek: paleo-] writing) and epigraphy (the study of inscriptions [Gr. epigraphē]), and historical phonology and linguistics. More broadly, philology engages issues of textual history (the lineage of editions behind extant texts), and the mastery of various historical, institutional, and cultural fields that relate to the essential formative environments within which texts were generated.

In this section we will distinguish three general approaches to philological studies, more narrowly understood as pertaining to matters of understanding the medium of written language in the traditional period, with an emphasis on recovering early meanings through research on the origins and principles of Chinese writing, and word etymology (concerning the origins of words and the characters that represent them). To keep this section manageable, it will generally consider only scholarship on language and writing prior to the Ch’ in era writing reforms. The approaches in each section are distinguished here for convenience, but in the practice of philological work there is no barrier between them.

I. Etymological research on principles of character forms and meanings
II. Paleography and ancient epigraphy: the database for the study of early writing
III. Historical phonology
IV. Online resources

In the 2000 revision of his field manual for the study of Chinese history, Endymion Wilkinson devotes considerable space to the issues raised in this section; he provides excellent discussion and includes references to a broader list of books and tools than is introduced here (though, naturally, new materials have been published since, and Wilkinson does not consider online resources). Where appropriate, references to discussions in Wilkinson 2000 included below, and it is worth noting at the outset that Wilkinson also includes in his manual a section on possible precursors of writing that archaeological work has uncovered, principally in the form of signs on pottery (pp. 373-388).

I. Traditional Approaches to Chinese Etymology

This section considers primarily traditional approaches to issues etymology, which focused on determining how character forms conveyed the “original” meaning of words. For most of Chinese history, these graph-centered approaches were dominant in issues of philology. In current scholarship, the relation of graph to meaning is generally treated as subsidiary to issues concerning the underlying spoken word, which is explored chiefly through phonetic analysis.
Interest in etymology, particularly the study of Chinese characters and their original meanings, while characteristic of Chinese scholarly traditions from the Han on, increased profoundly during the Ch’ing period. This was due largely to the growth of a “philological school” of textual criticism (k’ao-cheng hsueh-p’ai 考證學派), which began with some of the greatest thinkers of the early Ch’ing, Huang Tsu-nghsi 黃宗羲 (1610-1695) and Ku Yen-wu 顧炎武 (1613-1682). Members of this tradition were generally dissatisfied with the practical limitations of Sung Neo-Confucianism. Sung writings – particularly Chu Hsi’s commentaries on the Confucian classics – had been granted the status of state orthodoxy since the Yuan period and were the standard for official examinations. K’ao-cheng interpreters used philological methods to uncover “original” meanings in classical texts, meanings unfamiliar to Chu Hsi School commentators, whose interpretations were speculative rather than philological, and who tended to overlook issues of evolving word meaning. (On the history of this movement, see Benjamin Elman’s fine study, *From Philosophy to Philology*).

Originally, philological studies relied almost exclusively on variant usages found within the Classics themselves in their various editions, and on word definitions provided by early “dictionaries” of various types, such as the classic *Erh-ya* 爾雅 and the Later Han *Shuo-wen chieh-tzu* 書文解字. These materials were supplemented by very basic phonological techniques, based primarily on rhyme categories systematized in T’ang word books. By the late Ch’ing, k’ao-cheng writers had broadened their studies to include epigraphic materials. These consisted largely of stone inscriptions, such as a group of stone drums from the state of Qin inscribed with characters during the Spring and Autumn era, which became important objects of study as early as the T’ang – the poet Han Yü wrote a famous ode to them. Versions of the classics carved on stone by Imperial decree during the Han and later were also important sources for the early study of character forms, as were inscriptions on Chou bronze vessels, which had become a focus of study during the Sung. These materials were circulated in the form of ink rubbings, many of which were collected by antiquarians and published as books.

During the traditional period, however, the analysis of inscriptive materials was impeded by a relative lack of data, which prevented the development of theories of character forms and word etymologies that could supersede the enormously systematic and influential work of Hsu Shen 許慎 (c. 58- c. 147), the Han Dynasty compiler of China’s first true dictionary: the *Shuo-wen chieh-tzu* 書文解字. Philological scholar tended to rest on the bedrock of Hsu’s specific dictionary definitions (which remain important), and on his general analysis of the structural and semantic bases of Chinese characters, a system known as the *liu-shu* 六書.

*The “Six Principles of Character Construction”* (*liu-shu* 六書)

The phrase “*liu-shu*” denotes six principles that seem to underlie the construction of Chinese characters. The term itself may predate the Han: its first appearance is in the *Chou-li* 周禮, which some scholars date as Chou (though its final compilation was almost certainly later). The *Chou-li* instance, however, is of uncertain meaning; it is not definite that the phrase refers to character construction at all. It may refer to writing styles. The explication of the *Chou-li* phrase as a reference to types of construction principles begins, so far as we know, with Liu Hsin 劉歆 (46
BCE - 23 CE), who referred to it in his bibliographical treatise *Ch`i-lueh* (七略), now lost, but partially preserved in the *Han shu* 漢書. Liu’s use of the term was picked up by Hsu Shen when he compiled the *Shuo-wen chieh-tzu* about 100 CE. Hsu named and explained the *liu-shu* as follows:

1. 指事者：視而可識，察而見意：上、下是也．
2. 象形者：畫成其物，隨體詰詘：日、月是也．
3. 形聲者：以事為名，取譬相成：江、河是也．
4. 會意者：比類合誼，以見指撝：武、信是也．
5. 轉注者：建類一首，同意相受：考、老是也．
6. 假借也：本無其字，依聲託事：令、長是也．

Today, these six are generally treated in the following order: 2, 4, 1, 3, 5, 6.

There are two ways of approaching the *liu-shu*: one may debate the original meaning of the six terms as intended by Hsu and his contemporaries, or one may attempt to use the categories as a still viable means of understanding how to explore for early meanings of characters, especially as they are used in early texts. For the purposes of this course, only the second can be relevant. The explication here uses the categories only as a convenience, and adopts the theory of their original meanings which has been suggested by Lung Yü-ch’un 龍宇純, a philologist at Academia Sinica in Taiwan. (Lung’s theory is actually quite controversial in its explanation of #5, “chuan-chu.” There, he adopts a theory popularly known as the *yu-p’ang shuo* 右旁說. It may well be correct that Lung has not properly interpreted Hsu Shen’s intent, but if so, his interpretation represents a substantial advance over Hsu Shen’s in terms of usefulness.) Recent analyses of Hsu Shen’s use of these terms suggest that his purpose in enumerating them was not to expound an etymological theory, but to provide a cogent was of “thinking about” characters, principally as a learning device.

The *Liu-shu*

1. *Hsiang-hsing* 象形. Simplest of the categories. A *hsiang-hsing* character is one which is derived from a sketch of the whole or a part of the thing which the character denotes.

   Examples: *yue* 月 ; *niu* 牛 ; *yang* 羊

   Note that Xu’s choice of *ri* 日: ¯ is problematic because the dot in the center of the circle seems to represent an idea rather than a visual phenomenon.
2. *Hui-yi 會意*. The key criterion here is that the character be wholly composed of units representing independent characters used in their original sense.

Examples: *wu 武*: \( \text{foot & axe: Battle-march, war} \)

*hsin 信*: \( \text{man & word: one keeping to his word} \)

3. *Chih-shih*. The basic meaning is a sketch in which the meaning is not identical with the subject of the sketch but is suggested by it. There are several types:

a) Single element.

Example: *hua 化*: \( \text{an upside-down [dead] person, denoting “change.”} \)

b) Multiple elements with all elements independent character forms.

Examples: *chi 及*: \( \text{hand [catching] man: to reach} \)

*chih 陟*: \( \text{one foot above another: to ascend} \)

c) Same as b), but with some elements used in sense unrelated to meaning as independent characters.

Example: *kuan 關*: \( \text{hands closing bolted doors: to close} \) where the element 一 is not *yi* (one) but *shuan 閂* (bolt).

d) Multiple elements where some are not independent characters.

Example: *ts'uan 災*: (to cook: hands placing meat on a hearth and wood in a fire; only the hearth is not an independent character)

Note: Type a) is easily confused with *hsiang-hsing*, type b) with *huiyi*. These confusions can have interesting effects. One of the major original theories of Western sinology rested on a type-a) confusion. Herrlee Creel’s important theory of the origins of the word “*t'ien*” 天 (Heaven) rested on a gloss of the character *ta 大*, which Creel argued obviously had an original meaning of “big man” since it was a clear and simple sketch of a big man 當.

4. *Hsing-sheng*. Always includes both a semantic (meaning) and a phonetic element. There are two types:

a) A simple *hsiang-hsing* with an added phonetic marker.

Examples: *chih 齒*: \( \text{a mouth with teeth showing; 止 is phonetic} \)

*hsing 星*: \( \text{stars; 生 is phonetic} \)

*chi 鷄*: \( \text{a chicken; 禽 is phonetic} \)
b) A *hsiang-hsing* or *chih-shih* semantic marker with added phonetic:

With *hsiang-hsing*: *ch'i* 旗: (a pennant; 其 is phonetic); *chiang* 江:  
With *chih-shih*: *nai* 氛 (modern: “neon”).

5. *Chuan-chu* Often confused with *hsing-sheng*, these are characters which have evolved from types 1, 2, and 3, through the need to distinguish among different meanings which have accrued to the original character. Here, the phonetic element represents the original character, and so is really primarily a semantic unit. What is usually designated the radical (a semanteme) is an added semantic unit.

Examples:  
Root character: *li* 豐 (ritual vessel)  
*Chuan-chu*: *li* 礼 (ritual); *li* 醴 (ritual wine)

6. *Chia-chieh*. There are two types of “loan” characters, only the first of which properly belongs to the *liu-shu* as a character construction method. The second describes the way in which already constructed characters function as used in texts.

a) Basic *chia-chieh-tzu* are characters whose denotative meaning is fundamentally unsuitable for representation by ideograph. Primarily, these are function words. Words such as these were assigned characters by early writers by “borrowing” the graph of a homophonous word without regard for semantic associations. In this way, for example, the subordinating particle *chih* came to be written with the graph 之 ( ), whose original sense (foot & ground line) was “to go” (a sense it still retains as a secondary meaning). Similarly, the word “*ch'i*” (a pronominal particle) borrowed the graph of the then homophonous word *chi* 其: (a gleaning basket), which is now written with a bamboo radical to distinguish it from the particle.

b) In early texts it was very frequent for writers to use graphs to represent words which, from today’s perspective, were not the primary graphs for the words in question. (This was particularly true for words whose graphs were evolving in the *chuan-chu* style, as described earlier.) Variant graphs (generally primary graphs for homophonous words) are frequently referred to as “loans” (*chia-chieh*) for the primary graph when they appear in texts, and if the appearance as a variant is frequent, the primary and variant graphs are said to have had a “loan relationship,” or to be “cognates” (the latter correctly applies only to true *chuan-chu* relationships). These types of “loan characters” are properly called *t'ung-chia-tzu* 通假字.

*Traditional Reference Works in Etymology*

*Shuo-wen chieh-tzu* 說文解字, by Hsu Shen 許慎 (30-124) [O.C. PL 1281 .H8]
Historically, it is doubtful that any dictionary in the world has had greater influence than the *Shuo-wen chieh-tzu* had in China. Completed by Hsu Shen about A.D. 100, it was, for over fifteen hundred years, an almost unchallenged authority on character etymology, and even modern Chinese encyclopedic dictionaries give its definitions pride of place. Actually, the *Shuo-wen* is filled with false etymologies and doubtful definitions, but it remains an enormously impressive accomplishment.

In the *Shuo-wen*, Hsu Shen invented the radical classification system for characters. He identified 540 graphemic/semantic roots and arranged his dictionary according to a loose concept of their natural order, beginning with the radical *yi* 一, and ending with *hai* 亥, the last of the “earthly branches.” Hsu also was the first we know of to analyze radicals systematically into semantic and phonetic elements: characters in the *Shuo-wen* are generally explained by the formula:

从 X , Y 声

that is, “taking X as the semantic element (or radical) and Y as the phonetic element.” This is the basic form for almost all entries (though the simplest pictographic or ideographic characters include no distinct phonetic).

Despite the *Shuo-wen*’s impressive qualities, it is frequently in error or unclear. These problems generated a long series of commentary works, many far longer than the *Shuo-wen* itself. The most famous of these is the *Shuo-wen chieh-tzu chu* 説文解字注, by the Ch’ing philologist Tuan Yü-ts’ai 段玉裁 (1735-1815).


For many years, the incomparable encyclopaedist Ting Fu-pao (1874-1952) labored to bring together in a single publication all the commentaries on the *Shuo-wen*, organized in a single, dictionary format. The result is the massive (over 16,000 tightly packed pages) *Shuo-wen chieh-tzu ku-lin*. The dictionary includes the text of about 350 commentaries on the *Shuo-wen*, with the relevant discussions included after each character in cut-and-paste fashion. The 540-radical order of the original dictionary is preserved (a K’ang-hsi radical table is provided).

The *Shuo-wen chieh-tzu ku-lin* is the most comprehensive single source of etymological scholarship available. If you are engaged in a serious quest for the origins of a character, it is almost imperative to consult the *ku-lin*. However, it is a formidably difficult book to find your way around in. On pages following, there are a series of annotated photocopies of sample pages. These can be useful to consult in finding your way around Ting Fu-pao’s magnum opus.

Using 丁福保《說文解字詁林》

**Step 1**
For the traditional boxed edition, go to volume 66 康熙 radical index (ex.: for 帝, go to 中)
For volume-bound reprint edition, the radical index is at the end of volume 1

Text is consecutively paginated: p. 19 is in vol. 5 (note that there is a variant on p. 6945); chūan 卷 numbers are not used here.

Step 2

Go to the Shuo-wen text and commentaries: the main body of the book
Commentaries often incorporate the *Shuo-wen* text.
Step 3

Locate bibliography and abbreviations in volume 1

When citing commentary, convention is to use the *SWCTKL* pagination, but to cite the original author of the commentary and, if appropriate, the commentary title.
Since Ting Fu-pao’s time, the various branches of specialized learning that contribute to the study of etymology have developed in dramatically different directions, influenced both by the developments in epigraphy and paleography discussed in section II below, and also by the influence of Western linguistic theory and historical linguistics, which accounts in large part for the growing impact of historical phonology, the subject of section III, on our understanding of word etymology (now sharply distinguished from analysis of character structures). Contemporary studies of etymology prioritize reconstruction of phonetic structure over graphemic structure, as discussed below. Moreover, they make extensive use of phonetic reconstructions for both Old Chinese and, increasingly, linguistic families such as Kam-Tai and Austroasiatic that were spoken by peoples on the periphery of territories dominated by Chinese speakers.

Because of these dramatic changes in the scholarly approach to Chinese etymological studies, the most recent comprehensive research tool for Chinese etymology bears little resemblance to the *Shuo-wen* and its descendants:


As easy to use as Ting Fu-pao’s compendium is difficult, the body of Schuessler’s dictionary (pp. 149-638) lists characters according to *pinyin* transcription of their modern Mandarin phonetic values (although the way that homonyms are grouped is a bit more complex, as described on pp. xv-xvi). The dictionary glosses and etymological discussions rely on reconstruction of Old Chinese pronunciation, represented in standard International Phonetic Alphabet (IPA) form [IPA charts are easily located online], Old Chinese being defined as the language of the late Shang to Han eras. Introductory sections provide discussions of the linguistic features of Old Chinese, with particular attention to relevant aspects of non-Chinese linguistic influences. An extensive scholarly bibliography is included, and referenced throughout the dictionary. The only barrier to using the dictionary is the dense set of symbols and abbreviated forms; all are, however, explained in prefatory material. An index of English language words is included after the dictionary.

**II. Paleography and ancient epigraphy**

Our understanding of early Chinese language and written texts has been immeasurably enhanced over the past century by enormous recoveries of ancient materials through scientific archaeology, consolidation of materials in museums, and judicious acquisitions of valuable materials from private dealers fencing grave robbed good. Through these activities, we now have original texts from the late Shang era (c. 1250-1045 BCE), chiefly inscribed oracle bones, but including some inscribed ritual bronze vessels; from the Western Chou era (1045-771 BCE), consisting of ritual bronze inscriptions; and from the Eastern Chou period (771-221 BCE), including bronze inscriptions, but more significantly large caches of ink-inscribed bamboo manuscripts.

Many of the issues discussed in this section are introduced in a useful volume that focuses
on this burst of emerging sources for paleographic research:


One of the particular values of this volume is that in addition to introducing the materials, discussion is designed to help sinologists become familiar with how the materials can be understood and used.

In terms of the implications that paleography has had on our understanding of character formation, a comparison of traditional *Shuo-wen chieh-tzu* commentary to the work of Ch’iu Hsi-kuei 裘錫圭, perhaps the foremost Chinese paleographer of recent decades, reveals a sea change in approaches.


A translation of Ch’iu Hsi-kuei’s, *Wen-tzu-hsueh kai-yao* 文字學概要 (Beijing: 1988; revised ed., Taipei: 1994). Ch’iu’s handbook quickly became the standard work for contemporary analysis of Chinese character forms. His analysis of the construction and early use of character forms goes well beyond the liu-shu model and includes a broad array of distinctions, of which the main categories are: semantographs (*piao-yi tzu* 表意字), phonograms (*hsing-sheng tzu* 形聲字), loangraphs (*chia-chieh tzu* 假借字), allographs (*yi-t’i tzu* 異體字), and homographs (*t’ung-hsing tzu* 同形字). Ch’iu’s many subcategories and detailed analyses introduced unprecedented nuance to paleographic studies.

Another way in which the study of the roots of Chinese writing has changed over recent decades has been growing awareness of the ways in which the origins of writing in Mesopotamia and Egypt, which have been analyzed through very well developed traditions of paleography that have the benefit of far greater access to phonetic information than is the case in China, may cast light on the Chinese process. An influential study of the origins of Chinese writing that is informed by these considerations is:


*Shang oracle texts* (*chia-ku-wen* 甲骨文)
Oracle bone inscriptions, as they are commonly known ("obi" for short), are the divination records of the late Shang court, inscribed on ox scapulae and turtle plastrons which had been employed as media for spirit communication. Using longstanding pyromantic practice (that is, fortunetelling through the use of fire), the Shang people, like their Neolithic ancestors and neighbors, learned the intention of spirits by applying hot pokers to prepared bones and shells, the sounds or shapes of the cracks that resulted serving as the form of spirit response. Based on the evidence we now have, it appears that sometime about 1250 BCE the Shang kings residing at the last Shang capital, located near the modern city of An-yang in northern Ho-nan, first ordered that divination records be inscribed on the durable materials that were employed in divination. The brief inscribed notations of the queries posed by the king to the spirit world about issues of sacrificial practice, state affairs, royal family welfare, and so forth provide a rich, though incomplete, repository of historical and social information about China’s first literate era. But in terms of philology, their most profound importance is that they are the earliest evidence of Chinese language and writing.

We know now that most Shang oracle records were buried in pits at the capital, which itself fell into ruins and was buried under accretions of soil, and there they remained for millennia, unknown to history. Periodically, erosion and farm ploughs would unearth these bones and shells, inscribed with characters so primitive that they were not recognized as writing at all; they were instead construed to be manifestations of spirit force, and the peasants who unearthed them learned that they could be marketed to dealers in medicine, who would grind up these spiritually efficacious bones to be included in medicinal prescriptions.

The first recognition of the significance of these inscribed bones and shells came about in 1899, when unground samples were delivered into the hands of a fever-stricken scholar, Wang Yi-jung, and his fellow paleographer Liu O, who realized that the signs were a form of early writing and that their source could be the repository of profound historical secrets. Wang soon died (not from his fever, but as a patriotic suicide), but Liu launched a search for the place where these inscribed bones could be found. Dealers in the bones were willing to provide Liu with the names of their sources, but they were careful to lie. After years of following false leads, Liu finally determined that the bones had been unearthed by peasants in fields and river banks near the city of Anyang. Over the first decades of the 20th century, scholars applying newly imported methods of archaeological research recovered thousands of texts from Hsiao-t’un, a village near Anyang, which they ultimately determined to be site of the “Wastes of Yin” (Yin-hsu 殷墟): the last capital of the Shang (or Yin) Dynasty.

Within a decade, the preliminary deciphering of the oracle bone inscriptions had set the parameters of the Shang lexicon and oracle text studies became a major philological sub-field, drawing leaders of both traditional scholarship, such as Wang Kuo-wei 王國維 and Lo Chen-yü 羅振玉, and of scholarship more influenced by Western methodologies, such as Kuo Mo-jo 郭沫若.

In terms of the field of oracle text studies itself, a wide range of research monographs have been published as scholars achieved increasing degrees of expertise in deciphering and interpreting the texts. In English, the standard work on the basic features of the field is:

No source in English has superseded Keightley’s meticulous study, but its review of research tools has been overtaken by subsequent scholarship. Wilkinson 2000 provides an excellent survey of these tools in a surprisingly detailed section on oracle bones (pp. 396-405).

Keightley also published several summary accounts of the nature and form of oracle texts, of which the most useful may be one published in the volume *New Sources of Early Chinese History*, referenced above. The most recent English introductory account of the state of the field at present can be found in:


For the present, the most extensive field overview is:


In 1955, the dean of Shang paleography, Tung Tso-pin 董作賓, published a classic field overview of the first fifty years of oracle bone studies, *Chia-ku-hsueh wu-shih nien 甲骨學五十年*. Tung’s guide was a small paperback book of a little over 100 pages. The growth of the field is reflected in this oversize volume of over 700 pages, the work of five co-authors. It constitutes an encyclopaedia of Shang studies at the turn of the present century, with chapters covering such issues as inscription periodization, decipherment, grammar, divination technique, social structure, religion, astronomy, economics, and so forth.

Oracle inscriptions generally consist of brief texts of anywhere from one character to a string of several dozen. Usually, a single bone or plastron included multiple inscriptions (thus there are many more “obi” than there are inscribed bones and shells). Almost all inscriptions follow a common template for presenting their reports of the divination, which includes the following elements:

- A preface, generally indicating according to the sexagenary dating system the day on which the divination was performed and the name of the diviner (frequently omitted)
- A “charge”: that is, a statement of the topic divined (the main content of most obi)
- A prognostication, made by the king, of the outcome indicated by the divination
A verification: a record composed at some later time indicating what events actually transpired

A postface, recording further information about the date or place of divination

Here is a sample inscription, including all elements (one missing character is supplied):

[癸]亥卜争貞 / 句亡禍 / 王□ 曰有□ / 旬壬申中師□ // 四月

Cracking on [guì]hai day, Zheng divined /  
Preface

about whether the coming week would have no disaster. /  
Charge

The King prognosticated saying, “There will be disaster.” /  
Prognostication

On the week’s renshen day a disaster occurred at the Zhong encampment. /  
Verification

Fourth Month.  
Postface

The elements of this template occur with varying frequency. Virtually all obi record the charge; the vast majority include prefaces. Certain types of divinations will routinely include a prognostication or postface, but overall these are relatively uncommon. Verifications are, perhaps, the least common element encountered.

The earliest published inscriptions were collected in volumes of ink rubbings or hand copies, some privately compiled, but the largest being the products of major excavation projects operated by Academia Sinica in the 1920s. Keightley lists over sixty of these published collections, and the convention in early obi studies was to refer to individual inscriptions by using a combination of abbreviated collection title and fragment index number. Research on oracle inscriptions was dramatically facilitated by the first comprehensive index to the full range of this dispersed corpus:


Shima’s index was, in fact, much more than an index: it approaches a full concordance of the oracle inscriptions. It is organized as a sequence of Shang characters, with every occurrence of each graph in the oracle corpus provided under the entry for that character, with the complete inscriptive context being provided in every instance. However, for exceptionally common characters only representative occurrences in inscriptions are given. The inscription texts are hand copied in legible form, and each is capped by the standard reference system, a table of abbreviations to the 60 collections indexed being provided on p. 14 of the
prefatory section. Because oracle text characters do not conform with later orthographic norms, there was no way in which Shima could have used either stroke count or traditional radical categories for sorting obi graphs. He developed instead an index of 164 constituent forms, on analogy with the Shuo-wen chieh-tzu and K’ang-hsi tzü-tien models, but preserving Shang character structures, no attempt being made to reduce Shang forms to equivalents in later forms of standard orthography. Shima’s table of forms appears at the beginning of the book, and for each form two page numbers are provided: one to the page of the ten-page finding list for approximately 7000 characters, located in the preface, pp. 3-12, the other to the page in the full index where the simple character form is found. Only the first of these two numbers is generally useful, and, unfortunately, the numbers are systematically off by two pages. Names of Shang ancestors who appear in inscriptions are indexed at the end of the volume, and a useful lineage table appears on p. 556. Given the complexity of its task, the Sōrui index is remarkably simple to use, but it does include many errors and an abundance of cases where the representation of individual inscribed characters can be called into question.

Although Shima’s concordance was a major step forward in facilitating the development of Shang studies and philological research, it was soon made problematic by the publication of a nearly comprehensive collection that brought together in a single publication reproductions of the overwhelming majority of oracle bones and shells:


The Ho-chi, as it is generally abbreviated, provides consecutive index numbers for all 41,956 items in over 5000 folio pages. Inscriptions are arranged chronologically according to a five-period division devised by Tung Tso-pin; within each period, inscriptions are organized by topic categories, analyzing principal divination subjects according to such topics as slaves and commoners, patrician masters (these first two categories reflecting the Marxist legacy of PRC academics), military matters, tribute, agriculture, hunting, astronomy, spirit worship, and so forth. (Subsequent research has revised portions of the chronology.) A two-page general index at the close of the prefatory section in volume 1 should be consulted. The final volume is distinct because, unlike thefirst twelve, which reproduce rubbings, the final volume collects oracle texts that are preserved as facsimile tracings (mo-pen 坳本).
Shima Kunio’s chart of oracle text character forms
This new publication became the standard reference for oracle texts and citations of obi now generally include the index number of the corresponding bone or shell rubbing in Ho-chi, as it is generally abbreviated (for example, the inscription translated above would be cited as Ho-chi 5807). However, the Ho-chi was not indexed and Shima’s concordance provided no access to its renumbered fragments, making it less than ideally useful as a universal reference until a concordance was produced nearly a decade later:


The Lei-tsuan, as it is generally known, is closely based on Shima’s model and uses a slightly revised version of Shima’s table of forms to organize characters. The table appears on the first page of volume 1 and, like Shima’s index, indicates page references both for the finding list and for the body of the concordance. (Remarkably, the Lei-tsuan emulates the Sōrui in that there are discrepancies in the finding list page number.) The chief innovations of the Lei-tsuan are that inscriptions are given the index numbers of the Ho-chi, rather than the array of collections covered by Shima’s index, and the editors add to the handwritten replication of each inscription in Shang characters registers in which they provide their own transcriptions of every insessional entry into either traditional graphs in k’ai shu 楷書 form, or in the case of characters for which no traditional equivalent exists, a modern orthographic equivalent to the Shang form, based on analyses of the various component elements of the graph (a method termed li-ting 隸定). In cases where no form of equivalent could be determined, a tracing of the Shang character is supplied. In addition, the editors add a notation for each inscription of its chronological place within the 5-period scheme employed by the Ho-chi. (The transcriptions replicate those in a two-volume publication by the same editors which simply transcribes all inscriptions of the Ho-chi in numerical order, as well as those of the three other included sources, plus one additional collection.* The title is Yin-Chou chia-ku k’o-tz’u mu-shih tsung-chi 殷周甲骨刻辭摹釋總集 [Beijing: 1988] (O.C. PL 2456 .Y475 1988).) The hand-traced inscriptions of Ho-chi volume 13 are not included in the Lei-tsuan. However, it does include in its corpus some collections excluded by Ho-chi and the published inscriptions in one major find that postdated the editing of Ho-chi, located in the southern quadrant of the archaeologically rich village of Hsiao-t’un (Hsiao-t’un nan-ti 小屯南地). Thus, following a complete list of Ho-chi references for each character, examples appear that include one of three prefatory abbreviations:屯 stands for: Chung-kuo She-hui K’o-hsueh-yuan K’ao-ku-so 中國社會科學院考古所, ed., Hsiao-t’un nan-ti chia-ku 小屯南地甲骨, 2 vols (Shanghai: 1980, 1983); 英 stands for: Li Hsueh-ch’in 李學勤, Ch’i Wen-hsin 齊文心, Ai Lan 艾蘭 (Sarah Allan), eds., Ying-kuo so-ts’ang chia-ku chi 英國所藏甲

骨集; Oracle Bone Collections in Great Britain (Beijing: 1985); 懷 stands for: Hsu Chin-hsiung, Oracle Bones from the White and Other Collections (Toronto: 1979). At the close of the third volume there are radical, stroke count, and pinyin indexes, based on the transcription choices of the editors. Although many errors have been found in the Lei-tsuan and the transcription choices are in many cases not reflections of consensus, the concordance is currently the single most critical tool for the study of oracle inscriptions.

The Lei-tsuan does not include items omitted from the Ho-chi or subsequently discovered, apart from the three collections indicated in the description above. A compilation of all such items (over 13,000 in all, many very small), was published as:


The Ho-chi collection simplified the unreasonably complex system of reference citation for obi, but correlating inscriptions under the new and old systems was essential and required a table of correspondences. Only years after publication of the Ho-chi did a full set of tables emerge:


Different sets of tables allow for conversion from Ho-chi to traditional citation and vice versa.

In 1991, well after the publication of these major indexes, a major cache of inscribed turtle plastrons was discovered just south of Hsiao-t’un at a location known as Hua-yuan-chuang 毛園莊. The Hua-yuan-chuang finds, unique in the well preserved condition of the more than 1500 plastrons recovered, is particularly noteworthy because the divinations were all performed by or for a single Shang prince, rather than for the king.


The Hua-yuan-chuang publication includes its own concordance in volume 6 (though many of the character transcriptions have already been superseded by subsequent research). Research on this corpus is still in its early phases, with few monographs yet published in any language and only a single study, a yet unpublished PhD dissertation, available in English as of 2010. The content, form, and orthography of the inscriptions in this corpus, which dates from the earliest stage of Shang script development, are distinctive in significant ways, and promise to further refine our understanding of the formative development of Chinese characters. A searchable database for the Hua-yuan-chuang corpus can be found at: http://www.wenzi.cn/huadong/index.HTM.
In terms of the contribution of obi paleography to philological issues of language, the most important tools are a set of publications that explore the original development and meanings of Chinese characters, based on the Shang data. Most scholars believe that the obi corpus shows written language in so complex a form that there must have been a long period of script development preceding the late Shang phase; however, those earlier data are likely unrecoverable because writing was probably practiced in perishable media, such as ink on wood. Accordingly, the durable obi evidence will remain for us our best portrait of the dawn of writing in China, the ultimate origins remaining something that can be probed only in the speculative manner of Hsu Shen. Other scholars believe that the obi were themselves the origins of writing, an innovative mode of record keeping developed in the specialized environment of divination workshops, perhaps under the orders of the unusually powerful and religiously committed king Wu-ting, whose long reign began about the time of the earliest inscriptions.

The principal tools that update our understanding of character etymology in light of the obi finds are modeled on Ting Fu-pao’s cut-and-paste “ku-lin” format, listing in chronological order the comments of paleographic scholars on specific characters, correlating Shang character forms to later forms of Chinese writing wherever possible.

*Chia-ku wen-tzu chi-shih* 甲骨文集釋, compiled by Li Hsiao-ting 李孝定. 7 vols. (Nankang, Taiwan: Chung-yang Yen-chiu-yuan Li-shih Yü-yen Yen-chiu-so, 1965) [O.C. PL 2456 .L68]

Useful, though now somewhat dated, this dictionary collates scholarly discussion of *Shuo-wen* characters appearing in oracle bone inscriptions (often abbreviated in English as “obi”), and many other characters, some still unidentified, a total of 4766. It is modeled on Ting Fu-pao’s collated edition of the *Shuo-wen* commentaries, and employs the *Shuo-wen* arrangement of characters. The usefulness of the book is somewhat diminished by the fact that it is entirely handwritten (neatly, but in mild cursive) without punctuation. A stroke index appears in volume 1, pp. 143-181, with a stroke index for variant interpretations (pp. 183-204) and doubtful transcriptions (pp. 205-212) following. The bibliographic guide to abbreviations appears on pp. 213-232).


Much like Li Hsiao-ting’s compendium in concept, Yü Hsing-wu’s dictionary benefits from two decades of further research in oracle texts, incorporating research comments published through 1989. Like Li’s volumes, it is handwritten, but fully punctuated, more easily deciphered, and considerably easier to use. Characters are arranged according to the modification of Shima Kunio’s table of forms that is employed in the *Yin-hsu chia-ku k'o-tz'u lei-tsuang*. The chart and finding list appear at the outset of volume 1, but the finding list merely provides consecutive index numbers for the 3691 characters glossed, a true finding list appears at the end of
volume 4, along with lists by stroke and radical that provide character index numbers.

In addition to these large collections, the following one-volume index succinctly collects in tabular the modern Chinese character-equivalent interpretations for obi graphs of scholars.

*Kōkotsu moji jishaku sōran* 甲骨文字字釋綜覧, Matsumaru Michio 松丸道雄 and Takashima Ken’ichi 高嶋謙一 eds. (Tokyo: Tōbunken, 1993)

An unusual form of index, the *Sōran*, as it is known, arrays oracle text graphs in tabular arrangement and notes in columnar form the transcription equivalents provided in a very wide range of commentary works, from 1904 to 1988. Explanations and arguments are omitted, making this single volume work the simplest way to scan the varieties of interpretations available.

*Bronze inscriptions (jin-wen 金文)*

From the era of the last Shang kings at An-yang through the eight centuries of the Chou Dynasty, it was the frequent custom of patrician families to inscribe with some form of dedicatory text vessels of cast bronze that were principally used for ritual offerings of food and wine to ancestors. Inscriptions were typically made in the clay core of the vessel mold (thus generally appearing on the inner surface of the vessel); unlike the resistant media of bone and shell in which obi were etched, bronze inscriptions reflect the much freer calligraphic opportunities of a stylus scooping a soft substance, and provide a different view of the earliest forms Chinese writing.

Like oracle texts, bronze inscriptions have chiefly been used as sources of historical data about early Chinese society, politics, and religion. In this regard, the small number of late Shang inscribed bronzes provide only marginal additions to the obi records, and the bronzes of the Eastern Chou era (771-221 BCE) are also less important for their historical content. But study of the Western Chou period (1045-771), poorly documented in received texts, is centrally reliant on thousands of bronze inscriptions, many of which include full or partial dates and extend hundreds of characters in length.

The standard introduction to the field in English is:


Rubbings of bronze inscriptions have been collected since the Sung Dynasty, and paleographic expertise was developed enough by the dawn of the 20th century that Ting Fu-pao included commentary on bronze forms of characters in his assemblage of *Shuo-wen* scholarship, but it was only during the 20th century that the field of fully developed, spurred by the interest in paleography that discovery of the oracle texts had inspired.

Unlike obi, bronze inscriptions, of which known examples number well over 10,000, are
referred to individually by name, most commonly the name of the persons recorded within the
inscription as having commissioned the bronze vessel itself, plus a term representing the type of
vessel bearing the text (e.g., a cauldron [ting 鼎], basin [pan 盤], etc.). If multiple inscriptions
share these features, their titles may include distinguishing modifiers: for example, the titles of two
famous cauldron inscriptions of the tenth century BCE, commissioned by the same person, are
known as the Ta Yü ting 大盂鼎 and the Hsiao 小 Yü ting – the large and small cauldrons of Yü.

Bronze inscriptions represent a much more literary and diversified form of text than Shang
oracle inscriptions, but during the era of their greatest importance, the Western Chou, a large
percentage of them loosely conform to a template that became increasingly standardized by the
nineth century. Such inscriptions generally begin with a dating formula, which may include
narrative content denoting significant contextual events, a narrative involving official appointment
or recognition of the vessel commissioner’s merit by the king or other high patrician and an award
of gifts, and a final section dedicating the vessel to an ancestor, for use in making sacrificial
offerings at the ancestral shrine. Here is a brief tenth century example, known as the Shi Chü
fang-yi 師遽方彝 (Square vessel of Commander Chü) that illustrates these features:

In the first month, during the period of the waxing moon on the day ting-yu, the King was at
Chou conducting a ceremonial feast in the Chamber of K’ang. The valor of Commander Chü
was praised before his comrades. The King called out to the Steward Li to present Commander
Chü with a mien jade and four carved jade ornaments. Commander Chü bowed prostrate and
dared to raise in thanks the brilliant grace of the Son of Heaven. Wherefore was cast this
precious sacrificial vessel for my (i.e. Commander Chü's) patterned grandfather Ye Kung,
that I may pray for long life everlasting. May my descendants treasure it for a hundred
generations.

The ordering of bronze inscriptions published in collections of rubbings and hand
transcriptions has traditionally grouped inscriptions by the type of vessels on which they were
inscribed, and then according to the length of the inscription in terms of characters. This
arrangement, which reflects antiquarian approaches to cataloguing, continues to be standard, since
uncertainty and disagreement about the specific dates of individual inscriptions makes chronological arrangement impractical.

The standard collection for bronze inscription texts is:

Yin-Chou chin-wen chi-ch’eng 殷周金文集成, 18 vols. (Chung-hua shu-chü,
1984-95) [Oversize O.C. PL 2448 .Y54 1984]

Bronze inscriptions are generally cited by title and “Chi-ch’eng” number (unless, of course, their
date of recovery or original publication postdates the collection). For example, the inscription used
as an illustration above would cited as “Shih Chü fang-yi [Chi-ch’eng 9897].”

Two additional published versions of the Chi-ch’eng corpus include both copies of the
original rubbing or hand copy and transcription into modern character-form equivalents are:
Yin-Chou chin-wen chi-ch’eng shih-wen 殷周金文集成释文, 6 vols. (Hong Kong: Chinese University of Hong Kong, 2001) [O.C. PL 2448 .Y56 2001]

Each page includes a rubbing and transcription, as well as information concerning the total number of characters, period to which the inscription may be dated, place of origin (if it is an excavated piece), and current location.


Only the rubbing and transcription are given on the primary page of each inscription, but an index to each volume provides the information given in the previous item, along with publication histories.

In addition, a concordance of the Chi-ch’eng collection has been compiled (a broader concordance for bronzes has also been published, but IU does not hold it):


This concordance includes three basic components: (1) a transcription of the texts of all inscriptions included in the Yin-Chou chin-wen chi-ch’eng (pp. 1-180); (2) a finding list of characters included in the concordance section (a radical index to the finding list appears on pp. 181-84, with the list itself on pp. 185-224; note: a stroke index appears on pp. 1545-75); (3) the concordance, providing the context of every usage of each character (pp. 225-1478). Use the concordance by locating concordance page range of the target character in (2), finding the instances of the character’s usage in the Chi-ch’eng inscriptions in (3), and referring as needed to the full inscription text for each instance in (1).

By far the most extensive commentarial analyses of bronze inscriptions known at the time the Chi-ch’eng was compiled was done by the Japanese sinologist Shirakawa Shizuka. Although decades have passed, this work is still routinely cited:


This is actually a set of 56 issues of the bulletin of a museum in Kobe, Japan. The last two volumes are word indexes to the entire bronze corpus known in Shirakawa’s time.

While there have been few finds of oracle texts in the years since the major Ho-chi collection was compiled, inscribed bronzes are constantly being recovered archaeologically, many bearing lengthy and important inscriptions. Two publications have collected these more recent finds:


In addition, the following newly published item includes transcriptions, but no copies of rubbings:


This collection is designed on the model of the *Yin-Chou chia-ku k’o-tz’u mu-shih tsung-chi* (noted above). On the upper register of each page are clear hand copies of the original character forms for every inscription, and in a separate, lower register, a full transcription into modern character equivalents. It includes all the inscriptions in the *Chi-ch’eng* and *Chin-ch’u Yin-Chou chin-wen chi-lu*, plus more recently recovered inscriptions.

A number of annotated anthologies of Chou bronze inscriptions have been published that provide helpful access to this early form of text. The range of the following item is far broader than others, and the scholarly annotation is exceptionally strong:


A total of 925 inscriptions are transcribed. The first two volumes include rubbings, volumes 3-4 include reduced images of the rubbings, transcriptions, annotations, and detailed information concerning each vessel. Vols. 1 and 3 cover Shang and Western Zhou inscriptions; the other volumes cover Eastern Zhou. The earlier range are ordered chronologically, the later by state.

Just as specialists in oracle texts produced dictionaries in the form of collected character commentaries, modeled on Ting Fu-pao’s *Shuo-wen chieh-tzu ku-lin*, a prominent paleographer in Taiwan compiled and for many years updated such a dictionary for bronze texts.

Li-shih Yü-yen Yen-chiu-so, 1982 [O.C. PL 2448 .C44 Suppl.]

Not only a philological dictionary, but also a concordance of Chou bronze texts, the Chin-wen ku-lin series is more ambitious than Li Hsiao-ting’s earlier oracle text compendium. In particular, the fulu and pu volumes include translations into Chinese of the commentary of major Japanese scholars. The text for all volumes are hand written, but clearly and with punctuation. Indexes are found in ku-lin v. 1, pp. 161-210 (stroke index) and v. 16, pp. 3-196 (by radical); fu-lu v. 4, pp. 2685-2709 (stroke); pu v. 8, pp. 11-222 (radical). Unfortunately, supplements have not kept up with the proliferation of newly excavated bronze inscriptions, and the concordance function of the work is slowly eroding.

Bamboo texts (chu-shu 竹書)

The third major medium of new materials for paleography consists of ink-written manuscripts produced prior to the script reforms of the Ch’in Dynasty (221-208 BCE). Until the early 1990s, a growing stream of such texts had begun to emerge through archaeological recovery, but although their value was recognized by paleographers, their impact on the field was modest. Then in 1993, archaeologists working in Hubei Province at a site near the village of Kuo-tien 郭店 in Jingmen 荊門 Prefecture unearthed a significant hoard of bamboo manuscripts in tombs associated with the Warring States era state of Ch’u, datable to c. 300 BCE. These texts were subsequently published as:

*Kuo-tien Ch’u-mu chu-chien 郭店楚墓竹簡, Ching-men-shih Po-wu-kuan 荊門事博物館, ed. (Beijing: Wen-wu ch’u-pan-she, 1998)*

The Kuo-tien finds include a total of eighteen distinct texts, including three separate partial versions of the Tao te ching 道德經, a version of the Liji 禮記 chapter Tzu-yi 緇衣, and fourteen previously unknown texts. The volume provides a section of photographic plates that reproduces black-and-white photographs of every recovered bamboo strip, followed by transcription of the characters into modern character equivalents and commentarial glosses. The editorial staff wisely invited paleographer Ch’iu Hsi-kuei 裘錫圭 to add comments to these annotations.

The Kuo-tien texts represented a sudden acceleration of a trend. The Jingmen region had previously yielded a rich set of bamboo text findings at the site of Pao-shan 包山, but the nature of the texts, mostly medical and mantic, were too specialized to attract broad interest. With the recovery of the earliest known version of the globally popular Tao te ching, the study of Warring States era bamboo texts became perhaps the fastest growing subfield of sinology, both within China and in the West. Therefore, it may be unsurprising that the next major find was not the product of archaeological excavations, but an evasion of them: a massive number of bamboo texts robbed from graves, almost certainly in Hubei, and made available for private purchase by dealers in Hong Kong, well aware of their suddenly substantial value. Ultimately, almost all of these strips were purchased on behalf of the Shanghai Museum, which began publishing lavishly produced editions of them in 2001:
Of course, bamboo was not the only medium for ink writing in early China: wood tablets were used early on and silk manuscripts are known from the pre-Ch’in era as well, and became common during the Han. Examples of both types have been recovered, but are now much rarer than bamboo texts. (Writing is found in other media as well, for example, on coins and tokens of trade.) For an account of the most valuable cache of early wood strips, mostly employed for written covenants, see:


For an introduction to the most famous pre-Ch’in silk text, see:


For a full account of all archaeologically recovered texts on bamboo, wood, and silk, see:


The basic structure of this handbook includes a survey the history and structure of the field of early text studies (pp. 3-376), a chronological inventory of major finds (pp. 379-479), and an annual bibliography (pp. 483-702).

Dictionaries of archaic characters

There exist a number of dictionaries for oracle bone, bronze, and archaic characters. While these may be useful when learning to read oracle texts and initial reading guides, they are not generally employed for scholarly purposes, as they may convey a false sense of consensus in their glosses. Nevertheless, some of the editors of these dictionaries are leading scholars and their dictionaries can make deciphering texts much easier.

This is essentially a list of character forms, with references. The next two items are also character-form lists rather than dictionaries.


Arranged according to the Shuo-wen chieh-tzu radical system, there is a stroke index after the introductory material. About 1000 characters are included, with proper names and doubtful graphs excluded. The dictionary provides exemplars for each character under the chieh-tzu 解字 section, but its main value lies in the shih-wen 釋文 sections, where glosses are illustrated with text examples (unfortunately, indexed to the older array of collection sources, rather than to the Ho-chi).


This dictionary appears particularly useful in guiding readers through the evolution of character forms.


Schuessler’s dictionary gives definitions for characters and provides sample phrases and sentences of characters used in oracle text and bronze inscriptions, along with both reconstructed archaic readings and modern Mandarin equivalents. The dictionary is difficult to use because it is arranged according to pinyin transcription and includes no index. Thus one must know how Schuessler wishes to read an archaic character in modern Mandarin before the character can be located.


III. Historical Phonology

The exceptional interest of the Chinese writing system and the continuing relation between graphic form and semantic significance long overshadowed the importance of phonetics in studies of word etymology. Chinese writing is often mistaken for a direct transcription of ideas, with characters functioning as “ideograms” (idea-graphs), basically an elaboration of pictographs. However, true writing is a transcription of spoken language, words composed of sounds. While the liu-shu model may suggest that etymological analysis is principally a matter of analyzing how characters are
constructed to reflect meanings, contemporary scholarship places much greater emphasis on the way in which characters, from the earliest extant examples, were used to record pronounced words.

Chinese writing represents one of only four or five truly independent inventions of writing, Mesopotamian cuneiform and Egyptian hieroglyphics being the only others that persisted and evolved (Mayan cuneiform and, perhaps, the undeciphered Indus Valley civilization writing are the others). Cuneiform and hieroglyphics, like Chinese, began by using semi-pictographic forms to denote spoken words, but early in their development simplified versions of these forms came to be used to represent sound independent of meaning, ultimately evolving into alphabets. Only Chinese writing did not take this further step. We can understand this as a reflection of the unique circumstances of early Chinese polities, such as the late Shang and the Chou, which stretched over vast territories occupied by speakers of an evolving language. Over time, dialect divergence among speakers of Old Chinese combined with an expansionist need to communicate in writing with members of non-Chinese speaking peoples within and at the borders, would have made the cost of fully phoneticizing writing very high. Indeed, the persistent recurrence of the expansive Chinese state over time was greatly facilitated by the disjunction between writing and phonetics. Although in his Shuo-wen chieh-tzu, Hsu Shen was attentive to the ways in which characters reflected phonetic values, ultimately, the gap between spoken practice and writing obscured the importance of phonetics to understanding the etymology of words and their relation to characters.

It does not take much reflection to recognize the reason why phonetics received much less attention in etymological studies than did character forms: the characters were preserved, but the sounds of ancient Chinese were lost with the passage of time. Although Hsu Shen and others recognized that at the time that individual characters were fixed to refer to particular spoken words, the sounds of those words was often indicated through phonetic components of characters, even by the late Han the actual pronunciations of words that shared a single phonetic component had varied enough that the phonetic value that had at one time could be shared was unrecoverable. (We can easily see this phenomenon in modern Mandarin, where words whose characters show them to have been at one time homonyms now diverge broadly in pronunciation: for example, chih 至 and tie 堆; yi 義 and wo 我; shih 是 and t’i 提; etc. Of course, such changes occur in Western alphabetic languages as well, but in those contexts, spelling tends to provide far more detailed information concerning archaic pronunciation, as does alphabetic notation in hosts of related languages.

In the Chinese case, the best clues to archaic pronunciation are provided by a genre of texts known as “rhyme books” (yun shu 韻書), a term denoting word books initially compiled to provide aid in finding poetic rhymes. The earliest of these that is in any form extant is the Ch’ieh yun 切韻, complied by Lu Fa-yan 陸法言 in 601 CE, during the Sui Dynasty. Its most important descendant was probably the Kuang yun 廣韻, edited four centuries later.

The Qie yun provides historical phonologists a portrait in print of a type of standard dialect from its era. Although the specific pronunciation values represented may have been more an idealized type of court speech than a transcription of living language values, it probably reflects a notion of "correct" speech that was widely shared within a literati class that, despite profound differences in spoken dialects, wished to communicate as a group through rhymed poetry.
The *Qie yun* uses a type of phonological notation known as *fan-ch’ieh* 反切, which became a standard manner of recording word pronunciations until the twentieth century. In this system, each monosyllable is understood to have two components: an initial and a final. (In today’s parlance, the "final" would be analyzed as a composite of two elements, either of which could be absent: a "medial" vowel [-i-, -u-, -ü-] plus a final consonant or vowel.) Phonetic notation for a target character would take the form: X, Y Z 切, meaning that the pronunciation of character X consists of the initial of character Y plus the final of character Z. (Many texts use *fan* 反 rather than *qie* 切 to denote the meaning “cut and combine.”) For example:

直除力切

This example, taken from the *Kuang yun*, illustrates change in pronunciation from the era of the medieval rhyme books to modern Mandarin, since we would not analyze the sound of 直 (zhì) as a combination of chu plus li. But in the pronunciation of the era, known as “Middle Chinese,” 直 would be transcribed as djok, 除 would be transcribed djwo (in qu-sheng, or falling tone) and 力 as ljok, (that is, df + ak = djok).

The rhyme books may be the most important tool historical linguists use when reconstructing the pronunciation of the earliest period of Chinese texts, but they are not the only tools. Early poetry, particularly the rhymes of the *Shih ching* 詩經 anthology, which dates from approximately 1000-600 BCE, provides important evidence, even though the period of the poems may vary by several centuries and reflect both original rhymes in different dialects and also a regularizing editorial hand at a later date. Moreover, it is an axiom of the field that in cases where characters include a phonetic signifier (which most phonologists will assert means the vast majority of characters), identical signifiers convey homophony or near homophony of the underlying words.

One of the major contributions of Western philology to the study of Chinese character etymology has been to focus on the significance of phonology for etymological understanding. Led by the work of the Swedish sinologist Bernhard Karlgren and the theories of Peter Boodberg, who taught for many years at Berkeley, Western philologists have emphasized that true etymology must ultimately focus on the spoken word as prior to the written graph. By recognizing that Chinese characters were initially less attempts to map ideas into written form than to map spoken sounds with associated meanings, Western scholarship has helped clarify the semantic contributions of many character elements classified by Hsu Shen and others as simply phonetic markers.

The first etymological handbook to reflect these sorts of ideas was composed by Karlgren and originally published in the *Bulletin of the Museum of Far Eastern Antiquities* (Stockholm) 12 (1940) as *Grammata Serica* (it was later reprinted as an independent volume). Subsequently, Karlgren revised this as *Grammata Serica Recensa* (published in *BMFEA* 29 [1957] and issued by the museum in 1972 as a monograph).

*Grammata Serica Recensa* (GSR) is significant for many fields of sinology. In it, Karlgren groups characters by phonetic-graphemic “families” – that is, he groups together characters that share graphemic elements reflecting the phonetic value of the words they represent. (These are
commonly known as *hsieh-sheng* 諧聲 series; *GSR* lists 1235 such series.) The phonetic value of each character is given in terms of Mandarin (using Karlgren’s own transcription system), and also in terms of archaic and middle (T’ang-Sung) Chinese, using Karlgren’s reconstructions. Karlgren then provides simple definitions for characters, allowing users of *GSR* to track shared semantic features among members of character families and among members of different but approximately homophonous families. It is this feature of *GSR* that makes it so useful to etymological studies.

The information *GSR* provides allowed philologists to search much more easily for semantic linkages among characters that share phonetic indicators, and for phonetic linkages among characters that are close in meaning. The result is a powerful tool in the search for root words and meanings that may link characters, particularly those that exhibit *t’ung-chia* 通假 type loan relationships.

*GSR* is not very large, but it is organized in a unique manner and not particularly easy to use. Annotated copies of key elements of the book appear on the following two pages to help users find their way around *GSR*.

Use of *GSR* was somewhat facilitated by the publication of a version that organized the dictionary alphabetically:


Despite its alphabetic organization, Ulving’s book is not simple to use. It serves as an index to the phonological information about each of the 8398 characters provided in *GSR*: a finding list in *pinyin* at the outset gives both the Karlgren word-family code for the character and the index number of the phonological information provided in the main portion of the book, which is alphabeticized by Karlgren’s Old Chinese transcription. The word families are listed on pp. 16-22, but only the first character of the series, as listed by Karlgren, is provided; no overview of the series elements is provided. A very useful table of the phonetic principles of Karlgren’s Old Chinese transcription appears on pp. 12-14, with a helpful table showing how *GSR*’s 1235 word family groups were sequenced by Karlgren according to the 26 rhyme categories he attributed to Old Chinese.
Using Grammata Serica Recensa

Example: Locating information on the word 帝.

**STEP 1**

Go to the radical index at the rear of the text. Note: *all* graphs in the 帝 word family (that is, that take the form 帝 as a phonetic element) are found by looking up 帝, which is located under the 巾 radical: e.g., look up 諦 under the 巾 radical (帝), not the 言 radical.

The number given in the chart is the word family number, not the page number.
STEP 2

Karlgren’s reconstructed
Chou pronunciation
(Old Chinese)

T’ang pronunciation
(Middle Chinese)

Modern Mandarin
(Karlgren’s transcription)

Word definitions are followed by source references
(Shi: 詩[經])

Following convention, asterisks indicate hypothetical reconstructions
Karlgren’s reconstructions of Old Chinese have been superseded many times over in the years since GSR was published, although the usefulness of GSR is so great that they are still sometimes employed by scholars who are not specialists in historical phonology. Subsequent reconstructions by others such as Li Fang-kuei, Chou Fa-kao, E.G. Pulleyblank, William Baxter, Laurent Sagart, and others have addressed defects in Karlgren’s specific reconstruction. However, no single system has ever been so elegant as to unify the field behind a single set of choices. That situation may be changing. Nearly two decades ago, William Baxter published a massive analysis of Old Chinese which entailed a new transcription system that was widely cited:


Baxter’s “handbook” (at over 900 pages, the hand in question needs to be hefty) set a new standard for the field, and has become the basis for a new “companion” volume to GSR:


Although Schuessler does not generally include glosses for Old Chinese words, confining information to phonological issues (word meanings are the topic of Schuessler’s etymological dictionary, noted above), this volume is far simpler to use than Ulving’s and brings the phonological analysis up to date by replacing Karlgren’s notation with a modification of Baxter’s. A pinyin index at the back of the book indicates for each character a hyphenated number representing first one of the 38 rhyme groups into which Schuessler divides Old Chinese words, and then the number of the word family within each group (after a slash, the number of the GSR word group is also provided).

Schuessler notes that despite the fact that he relies on Baxter’s system, that system no longer represent’s Baxter’s own views. These have been much influenced by a work subsequent to his Handbook:


Sagart’s analysis chiefly expands a theory first developed in 1930 by the great French sinologist Henri Maspero, which holds that Old Chinese words typically included a range of “affixes” (prefixed, infixed, and suffixed features) which served syntactic functions, much like inflected languages in the Indo-European family. Although the approaches of Baxter and Sagart are different in significant respects, the two have since teamed on a project to develop a new system for the reconstruction of Old Chinese, and their preliminary efforts have been made available to the field online through a series of “Beta” versions, the latest being the “Baxter-Sagart System, Version 0.99” (http://sitemaker.umich.edu/wbaxter/home).
Online Resources

The most important single online site for the study of recovered texts, whether oracle bones, bronzes, or bamboo strips, is the CHina ANcient Texts database, or CHANT, website, which is maintained by the Institute for Chinese Studies at the Chinese University of Hong Kong.

The CHANT site holds databases for archaeologically and otherwise recovered text materials, including oracle and bronze texts, bamboo texts, and also silk texts recovered from early Han Dynasty sites (such as the well known Ma-wang-tui texts, unearthed in 1973). In addition, it includes a repository of received texts that includes virtually all known texts from the pre-Ch‘in period, and a wide range of texts from the Han and Six Dynasties eras. All texts are fully searchable, and the received texts and transcriptions of bamboo and silk materials are extensively annotated.

This section will introduce only those portions of the site that are devoted to recovered materials, that is, obi, bronze inscriptions, and bamboo and silk texts.

CHANT site a subscription site. IU Libraries has an institutional subscription: link through IU Libraries: https://www-chant-org.ezproxy.lib.indiana.edu/member_login/login.aspx for access to CHANT without a personal subscription. This will lead you to the CHANT homepage:
The second screen provides two banks of buttons: recovered texts are on the left, received texts on the right.

Before accessing recovered texts, it will be necessary to download the CHANT fonts. This is accomplished through the *hsia-tsai* 下載 link in the upper register.

The *chia-ku-wen* button will take you to the oracle text area of the site. Enter the search functions through the appropriate link on the top register of the first screen and at the next screen a bank of buttons on the left allows you to select from a number of search options, such as by direct entry of a character string, by *Ho-chi* or other index inscription number, or by radical index of the character sought. Ultimately, your search will link to a screen that reproduces the relevant inscriptions in a standardized Unicode version of Shang script, with character equivalents according to the *Lei-tsun* reading. For example, if you look up the key character 甲 in *Ho-chi* 5807 (the example inscription provided earlier in this reading), you will see the following screen:

![Download link and H05807 screen](image-url)
In this way, by using either index numbers or key characters, any oracle inscription can be accessed. However, since no images of rubbings are provided, there is no way to check the accuracy of the site’s standardized replication of the text.

The *chin-wen* 金文 button will access the bronze inscription area of the site. From the initial screen, the easiest way to call up inscriptions is to use the *chin-ch’eng sou-hsun* 進階搜尋 function at the right hand tab under the upper register of links. That will lead to a search screen that allows search by *Chi-ch’eng* number, character string (either input directly or using radical or stroke-based index screens), or by vessel title. For example, to locate the text for the sample inscription translated earlier in this reading, we could input index number 9897, the vessel’s title (*Shih Chü fang-yi* 師遽方彝), or a distinctive character string from the text. This will lead us to a list of appropriate vessel titles – including, as is often the case, multiple existing versions of identical inscriptions. Clicking on one of these will bring up a screen that includes both the rubbing of the inscription and a transcription into modern character equivalents, arranged in parallel with the rubbing (passing the cursor over individual characters in the transcription will call up enlargements of each character in the rubbing).
There are two separate entry points for bamboo and silk texts (*chu-chien po-shu* 竹簡帛書). The first leads to a database of seven collections of recovered strips and silk texts dating from the Ch’in and Han eras.

One accesses each collection by clicking on the appropriate book icon. The subsequent screen provides a table of contents in the left hand register, broken down into links for text transcriptions, commentary, and a variety of other useful databases for the relevant text.

While the Ch’in and Han period manuscripts are valuable for many reasons – apart from providing early versions of texts such as the *Lun-yü, Tao te ching*, as well as previously unknown medical, religious, and philosophical texts on, they are particularly rich in records of governmental administration and law – for the types of philological issues that are the focus of this reading, they are somewhat less critical than texts that predate the Ch’in script reform. These, the texts recovered from the Kuo-tien site and those purchased and published by the Shanghai Museum (periodically updated to reflect additional publications), are in the second of the two “chien-po” 簡帛 databases.
The sample screen below, from the Ch‘u chu-shu link, illustrates the rich resources of the site. The left register lists the database texts, each opening out into a series of resource screens: transcription, bibliography, and comparisons with relevant received texts. The main register in this illustration is of a text transcription, with the cursor on note 1, which calls up extensive annotation in the yellow block (only partially captured in this illustration).
In addition to the CHANT databases, a great deal of research concerning paleography and related issues is now hosted online by units of Wu-han University. The earlier of the two is known as Jianbo (或 簡帛研究) at URLs http://jianbo.org or http://bamboosilk.org; it is supported by the Harvard- Yenching Institute and now maintained by the 武漢大學中國傳統文化研究中心. The second, now somewhat more developed, is known as BSM (或 Bamboo and Silk Manuscripts; 簡帛); the URL is http://www.bsm.org.cn/; it is maintained by the 武漢大學簡帛研究中心.