BUDAPEST MONOGRAPHS IN EAST ASIAN STUDIES

1

Imre Galambos

ORTHOGRAPHY OF EARLY CHINESE WRITING: EVIDENCE FROM NEWLY EXCAVATED MANUSCRIPTS

IMRE GALAMBOS

ORTHOGRAPHY OF EARLY CHINESE WRITING: EVIDENCE FROM NEWLY EXCAVATED MANUSCRIPTS

BUDAPEST MONOGRAPHS IN EAST ASIAN STUDIES

SERIES EDITOR: IMRE HAMAR

IMRE GALAMBOS

ORTHOGRAPHY OF EARLY CHINESE WRITING: EVIDENCE FROM NEWLY EXCAVATED MANUSCRIPTS

DEPARTMENT OF EAST ASIAN STUDIES, EÖTVÖS LORÁND UNIVERSITY
BUDAPEST 2006

The present volume was published
with the support of the Chiang Ching-kuo Foundation.

ISBN 963 463 811 2
ISSN 1787-7482

CONTENTS

ACKNOWLEDGEMENTS

This book grew out of my Ph.D. dissertation. First of all, I wish to thank my advisers, Professors Jeffrey K. Riegel, Stephen H. West, and David N. Keightley. I thank them here not only for their guidance while writing this dissertation but also for their patience and support during the seven years that I spent at the University of California, Berkeley as their student. I especially would like to thank Professor David N. Keightley, who rigorously dissected my drafts over and over, correcting the multitude of mistakes and inconsistencies in both my thinking and writing.

Among the people who helped me revise the dissertation and shape it into a book, I am also most grateful to Susan Whitfield from the British Library for proofreading the manuscript and inspiring me to continue working on it. I also thank Haida Liang from the National Gallery, London for calling my attention to numerous problems hitherto unnoticed. Special thanks to Erzsébet Tóth for typesetting and preparing the manuscript for publication.

The main bulk of this book was written at the University of California at Santa Barbara while I was teaching there as a lecturer. I wish to thank the entire Department of East Asian Languages, the faculty and staff of which had been kind and supportive, rushing to offer help whenever I needed. I especially would like to thank in this respect Professor Sharon Yu.

I also want to thank Cecilia Chu from the Chinese Language Program at UC Berkeley, without whom I probably would not have been able to finish this work. She gave me support and encouragement when I needed it most and I am grateful to her for this. I also wish to thank the staff and faculty of the East Asian Library at UC Berkeley for their kind support. Among the staff of the Department of East Asian Languages and Cultures at UC Berkeley, I would like to express my gratitude to the Student Affairs Officer, Karen Ralls Mayers, who has been helping me with administrative affairs for the past few years.

Among the other people who helped me with their input and suggestions, I am thankful to Dr. Mathias Richter from Hamburg University, Tim Mansfield, Dr. Péter Vámos from the Hungarian Academy of Sciences;

Professor Michael Lewis from Michigan State University; Dr. Mie Lewis; Professor Eugenio Menegon from Boston University; Professor Christoph Harbsmeier; Gábor Kósa from ELTE University, Budapest; Vika and Tom Rendall; and last but not least, Márta Vajda and my father, Imre Galambos.

I wish to thank the Chiang Ching-kuo Foundation, especially Professor Wang Ch'iu-kui, for their generous grant that made the publication of this book possible. I am also grateful to Professor Imre Hamar from ELTE University, Budapest for his continuous encouragement and support for publication.

INTRODUCTION

The archaeological discoveries of the last few decades have provided an unprecedented amount of Warring States 戰國 (453–221 BC) texts in the form of manuscripts and inscriptions on various objects. From the point of view of palaeography, an intriguing challenge is how to fit all this new material into the early history of Chinese writing. Since these new texts predate the Qin 秦 dynasty (221–206 BC), they are able to provide the modern researcher with undigested data regarding the nature of writing in Warring States China. With the sudden increase of original documents, it has become clear that we need to revise our views regarding the nature of early writing, as well as the process and effect of the Qin unification.

The new material, as I shall argue, refutes the traditional linear model of the evolution of writing in China. According to this model, characters developed along a single line from the Shang 商 oracle-bone inscriptions to Zhou 周 bronze inscriptions, all the way to the Qin small seal 秦小篆 and Han clerical 漢隸 scripts. My contention is that this view is not only an oversimplification but in many cases incorrect. This model mirrors the ideologically motivated unilateral genealogy of traditional historiography which traced the mandate of Heaven from mythical emperors to the ruling house.

Both modern and traditional scholars have observed significant graphical variations between character forms in Warring States China. However, they attributed most of the differences to either temporal (i.e. evolution of character structure in time) or spatial (i.e. graphical differences between the writing systems of various regions where the various scripts showed local characteristics) factors.

It has not been commonly recognized, however, that variations also occurred within a corpus from the same general era and location. Longer Warring States manuscripts reveal that even a single document, undoubtedly written by the same person, could contain structurally distinct forms of the same character. Now the newly discovered documents provide a sufficient amount of material to document this phenomenon and to prove conclusively that the differences were not solely due to spatial or temporal factors. I will show that orthographic variability was an integral part of

the Warring States writing system, that character structure showed variation even within the same community and that, as seen from the concurrent presence of variant forms within the same "local script," the differences were due not to external factors but to the flexibility in the writing habits of Warring States scribes and the tolerance of their readers.

In reality, the differentiation of variant forms based on their structure is a modern distinction which was not significant for those people who wrote or read these character forms. They probably would not have even registered most of the differences between variants because they were primarily concerned with what meaning and sound a character could represent, rather than its graphical structure.

The high level of structural variability of character forms in excavated documents raises the question how fixed the structure of characters was and what constituted two variant forms of the same character versus two different characters. Another important question is whether these variants presupposed an initial form which they were the variants of or they were simply different visual representations of an "abstract character." The dominant view in both traditional and modern scholarship regarding variant characters has been that in each case there was a correct, or standard, form (*zhengzi* 正字), while all other forms were derivatives. In the case of Warring States character forms, scholars also commonly employed the category of "scribal errors" and "typos," in which they, once again, rated the character forms according to the *zhengzi* standard.

This assumption originated from examining the individual character shapes one by one and matching them against the vast corpus of historical forms which stretched from the Shang to modern times. The result unavoidably was that the variant form in question "lost out" and was regarded as an inferior form. It is only when we compare the character forms of the same script to each other, that we can see that the variants together constituted an aggregate character form. This aggregate character form is an abstract entity where each constituent form is equally "correct," only more or less frequently present. One could imagine the totality of individual character forms visually as a "cloud" of potential forms, as opposed to one discrete form. This representation is similar to the idea of an electron cloud where the actual location of the electron is represented only by the multitude of its potential locations. The character form became concrete once a scribe wrote it down, but prior to that moment any particular variant was only a possibility.

An often disregarded issue in the study of variant character forms is that beside the occurrence of a form it would also be useful to document the number of its occurrences. Most of the dictionaries and character compen-

dia have been deficient in this respect because they only recorded forms that were different from each other, but failed to include those that were identical. As a consequence, both common and rare forms came to be listed next to each other without any indication as to whether a certain form occurred once or hundreds of times. While the occurrence of concrete forms might seem random today, my analysis shows that the ancient scribes used them with a statistically consistent frequency. Despite the large number of potential forms, the scribes usually used a "preferred" form more often than any other form. The frequency of other forms declined along a curve, with the least common ones being used only once in the entire corpus.

I intend to show that when one looks at the character forms as a group, one can detect specific patterns behind their variability. The most important and obvious practice was the retention of the phonetic element. The scribes could abbreviate or leave out almost any other part of the character, could introduce new components, yet they retained the phonetic component in virtually every instance. This realization reinforces the priority of spoken language (sound) over writing (visual form), a connection easily forgotten when it comes to Chinese writing.

It is possible to establish that starting from the late 3rd century BC, there was a growing tendency to regularize the form, sound, and meaning of characters. The appearance of Han dictionaries and primers aiming to define any of these three aspects shows the increased efforts directed at standardizing Chinese writing as the regional centers of competing states were replaced by the centralized bureaucracy of Han hegemony. Chinese historiography attributed the standardization of writing to the reforms of Li Si 李斯 (280?–208 BC) but archaeological data reveals that the changes were, to perhaps an event greater extent, a result of a long historical evolution induced by political and administrative consolidation. The absence of any clear-cut border between the form and structure of characters immediately before and after the Qin reforms suggests that there was little distinction between the standards allegedly enforced by the government and the actual customs practiced by the people. Later in the Han, the establishment of the examination system and the erection of the stone classics were the means by which the government began requiring those who wanted to participate in the official administration to learn the official way of writing.

With respect to methodology, while I rely on previous, Chinese and Western, scholarship in working with manuscript sources, I set up six guidelines of my own that I have found particularly useful for the study of ancient Chinese character forms.

Firstly, I examine pre-Qin writing on its own terms, studying early manuscripts instead of relying on transmitted literature. The reliability and au-

thenticity of transmitted texts has been an issue that has occupied the minds of Chinese scholars for two millennia, and the new discoveries often confirm their suspicions. But textual faithfulness is only part of the problem. All transmitted documents written before the Qin dynasty had been transcribed into clerical script and had lost much of their inherent graphical information. When the transcribers rendered the characters into clerical script, they inevitably incorporated their own decisions and views into the text, thus transforming the original material. The fierce controversy in the Han between the *jinwen* 今文 and *guwen* 古文 schools over the authenticity of the classics shows that the process of transcription was problematic and the reading of ancient graph forms was by no means unambiguous.

Second, I use the original character forms, not transcriptions. During the last few decades, archaeological excavations in China have yielded a large amount of new palaeographic material. There are so many finds that in many cases the excavators do not have time to organize and publish the materials. These new texts are now forcing us to reevaluate our understanding of early Chinese writing. Before the discoveries, there simply was not enough reliable data to construct a clear picture of Warring States writing. Scholars could only work with transmitted sources and rely on tiny bits of information from the few available, sometimes erroneous, palaeographic sources. Needless to say, without a larger amount of reliable data, even the most brilliant studies will still remain on the level of hypotheses and are bound to contain errors. Since the 1980s, we have had enough data to establish the basic principles behind the evolution and use of the early Chinese script. The only obstacle to hinder our understanding of the true nature of the script was individual convictions rooted in traditional views. Therefore, my intention here is to stand back and let the data "speak" for itself, with minimum intervention. Because today one learns to write Chinese, even classical Chinese, using modern characters, one's understanding of the script is conditioned by the graphical information inherent in the modern script. This causes one to see early scripts in the light of modern writing, constantly referencing old forms to their modern "equivalents".

On the level of individual character forms I also make an effort to work directly with the original photographs and tracings instead of using transcriptions, thus relying on the original graphic forms as a point of reference. When analyzing ancient character forms, linking them to their modern equivalents presupposes an unequivocal relationship between the structure of the character in the old and new scripts. This is clearly not al-

ways the case. There are cases when even the transcription of easily identifiable components is ambiguous, depending on which components one chooses to represent in the transcription as a single unit.

Third, I believe that one must study ancient character forms on the basis of a statistically significant amount of data, because individual character forms by themselves are only examples of possible configurations. Finding solitary instances of certain orthographic variants is not sufficient for making general assumptions about the way scribes wrote that character, it only shows that a scribe *could also* have written it that way. Therefore, one cannot get an accurate and comprehensive image of the orthography of a character in Warring States China without considering multiple examples of that character.

Fourth, I establish the category of a "local script" which comprises the script shared by, and peculiar to, a living community. Obviously, one can find a great many variant character forms in the history of writing, but most of these examples would be centuries apart or written in places hundreds of miles away from each other. Even during the two and half centuries of the Warring States period, on the territory occupied by the seven major states, there were many distinct sub-scripts which differed from each other in many ways. These differences have already been addressed by former researchers. For my purposes I regard only instances of variant forms that are documented within the same local script as cases of inconstancy. I consider a local script in early China the sum total of the characteristics of a script shared by the people in the same locality and time for whom the script was unambiguously intelligible. For example, I consider that the Houma covenant texts 侯馬盟書 to have belonged to a single local script because, although written by a number of different scribes, they came from the same location and time period. I consider it more useful to compare character variants of an entire corpus than those of a single text because the former allows us to gather information on the nature of the script, including the variation tolerated by its readers, while the latter reflects only the writing habits of one scribe.

Fifth, I do not assume the existence of a correct or perfect form. Instead, choosing the path of descriptive linguistics, I regard each and all forms simply as different ways of writing a word, none of which is less or more correct than an other. This approach allows me to refrain from trying to pin down the perfect form, be it the modern form of the character or an occasional variant for which modern scholars can establish an etymological sequence. As a consequence, when referring to concrete orthographic forms in excavated manuscripts, I prefer to use the concept of "character form"

instead of "character" because the latter presupposes the existence of an abstract or ideal character that can be separated from its concrete physical appearance.

Sixth, whenever possible, I identify a character form based on the word it represented in context. Due to the lack of an orthographically perfect form in Warring States writing, it is not always possible to determine which two characters were identical and which different based on their constituents. Two character forms could be structurally different and stand for the same word, or be structurally identical and stand for different words. However, since writing is a means of recording language, each character form, regardless of its actual structure, stands for a specific word, and it is the context that permits us to identify that word.

The above six points represent the methodological framework of this study. In reality, all of these points revolve around the principle of looking at uninterpreted character forms in a local script without assuming the existence of a correct form.

The time frame of this study stretches from 497 BC to 221 BC, which loosely overlaps with the Warring States period. I am interested in the state of the script when there was no central authority controlling the entire country. I examine the state of Chinese writing prior to its consolidation and transformation into the clerical script stage, which automatically sets the upper time limit to around the Western Han 西漢 (206 BC–AD 23). Although I argue later that the standardization of writing happened over a long period of time and that some aspects of it lasted even beyond the Han, it is certain that both the Qin and the Han played a pivotal role in this process. As for the starting point in time, I chose the early 5th century BC marking the end of Spring and Autumn 春秋 and the beginning of Warring States periods. The reason for this choice is two-fold.

On the one hand, this was the time when a handful of states managed to increase their political influence and became regional centers. This process was important from the point of view the evolution of writing, because the strengthening of regional centers triggered the development of local scripts. These local scripts adapted to their new dialectal and cultural environment, as amply evidenced by the differences between the graph forms of the Six States.

On the other hand, the early 5th century BC is the earliest time from which a corpus of brush-written materials is available today. These documents from Houma 侯馬 and Wenxian 溫縣 were covenant texts written on jade tablets and buried together in pits as part of a collective covenant ceremony. Prior to the time of these covenant texts, most extant documents were either incised or cast on various objects, therefore they do not

reflect the handwritten script of everyday life.[1] In the course of this study, I mostly rely on the Houma covenant texts (497–490 BC) and the Guodian 郭店 manuscripts (ca. 300 BC) but also cite examples from other Warring States, Qin, and Han material.

With regard to the organization of this study, Chapter One, "Former Understandings," provides an example-based introduction to former scholarship on the subject of the structural variability of the Chinese script. Without trying to be comprehensive, my goal is to present a selected number of representative approaches from both traditional and modern scholars. I point out the merits of these works and incorporate these into my own examination. At the same time, I also try to learn from their shortcomings.

Chapter Two, "The Qin and Han Creation of the Standard," looks at the reasons behind the misconceptions regarding the Qin reforms and the discontinuity of Warring States scripts. I analyze the descriptions of Xu Shen 許慎 (AD 58?–147?) and Ban Gu 班固 (AD 32–92) on the evolution of writing and the Qin "unification of writing," which had been the major source of information on this subject for the past two millennia. In the course of this analysis, I show that while later interpreters simply misunderstood some of the described events, Xu Shen and Ban Gu were also striving to enforce and validate an orthographic standard which in reality did not exist at the end of the 1st century AD. The standardization of writing did not happen overnight in 221 BC but lasted several centuries. While the clerical script was already the dominant form of writing during the beginning of Western Han, the orthographic standardization of the script involved a longer process.

Chapter Three, "Identification of Character Forms," discusses the theoretical difficulties in identifying what constitutes a character. I show that although the common approach is that characters are graphically defined by their components, the structure of Warring States character forms is not constant and therefore cannot be used to define the character. In fact, the practice of identifying characters based on their components is the main reason behind the inability of scholars to discard the concepts of

[1] To use Qiu Xigui's (2000, p. 63) terminology, inscriptions on bronze and other objects represent the "formal script", while brush-written texts the "popular script". I believe that the current state of writing is always represented by the popular script, that is, the actual way in which people write during their daily routine. In formal scripts, not only the degree of archaicization and embellishment is high but also the inscriptions adhere to and imitate calligraphic traditions that in many cases are several centuries old.

"correct form" and "variant form." I demonstrate that, because of the variability of structure, it is more useful to define Warring States character forms based on the word they represented. Since writing is a graphical representation of language, the primary function of characters is to record words. It is the word that can serve as an unchanging entity behind all variant character forms.

Chapter Four, "The Standard," examines the concept of orthographic standard. First, I look at how this concept arose and developed in China, always being a desired ideal rather than reality. Then, for the sake of comparison, I look at the concept of standard with respect to the orthography of written English. I sample written English from late medieval manuscripts, late Renaissance spellings of the name of Shakespeare, and the language of modern Internet newsgroups. Each data set reveals a significant degree of orthographic variability, implying that a standard either did not exist or the members of the community did not follow it.

Chapter Five, "Variation," first shows that orthographic variability in Warring States writings was not always due to temporal or geographical factors but that it could occur within the same written dialect, sometimes within the same manuscript. Following this, I define my own methods of determining which character forms are variant forms. According to this definition, two forms are variants of each other when they stand for the same word and yet they are structurally different, i.e. they differ in at least one component. To illustrate this definition, I look at two versions of the same short text, both from the *Guodian Laozi* 郭店老子 manuscripts.

Chapter Six, "Patterns of Variability," analyzes every written form of three different words from the Houma covenant texts, probing into the reasons and patterns that lie behind their orthographic variability. By arranging the variant forms into a statistical chart, I show that while the ancient scribes wrote down a word with a specific form more often than other forms, they sometimes also consistently used other forms. Moreover, while the scribes wrote character forms with a considerable degree of orthographic flexibility, they almost always included the phonetic component.

Chapter Seven, "Conclusions," sums up the findings of this study and elaborates on their historical significance. Beside offering a more useful approach to both studying Warring States manuscripts and variant character forms in general, this study sheds new light on the development of the Chinese script, its transition into the clerical script stage, and the reality of the Qin reforms. The variability of Warring States character forms demonstrates that Chinese characters evolved not along a linear path that stretched from the oracle-bone inscriptions to the modern script but followed a complex process involving distinct cultures and languages. The

"fuzziness" of the line of evolution with respect to the spoken languages and dialects of ancient China raises questions regarding the national identity of the Chinese script. A related issue is how far can one go back in time and say with certainty that the various scripts were not only the predecessors of the Chinese script but were in fact Chinese.

CHAPTER ONE

FORMER UNDERSTANDINGS

This chapter looks at some of the works and theories related to various aspects of character inconstancy in early Chinese writing. One of the things I am interested in is how traditional and modern researchers have approached the phenomenon of structural variation, whether they described it as an ordinary feature of the script or condemned it as a deviation from a non-variable ideal. I call the first approach "descriptive" method, and the second "prescriptive."[2] Advocates of the descriptive method simply describe the script, including its variations as an outside observer without interpreting the data. Advocates of the prescriptive method represent a more active attitude towards the script; their primary focus is to establish or impose a standard which they regard as the proper state of the script.

Partly related to the above issue is the question whether the researcher sees variant character forms as a normal feature of writing or as mistakes. Even if there is no attempt to judge the way people wrote, as it is often the case with modern scholars looking at the pre-Qin script, there could still be an evaluation of the variant forms against a "correct" form. By doing this, the approach would be prescriptive with regard to the past because there is a presupposition of a standard way of writing a certain character.

I am also interested in how scholars specify the domain within which they compare variant forms. Although comparing character forms from different places and time periods with each other is useful to study the nature and evolution of Chinese writing in general, in order to examine productively the relationship between variant forms it is necessary to limit the scope of the corpus.

In observing how the literature treats the above concerns, I do not intend to provide a comprehensive historical review of all sources, but rather to discuss a handful of selected materials which have helped me to define

[2] In this study, I use the term "descriptive" in contrast with the term "prescriptive," as they are understood in the terms "descriptive linguistics" and "prescriptive linguistics."

my own approach to the problem.[3] I have arranged the literature in chronological order, starting with traditional views in imperial China, then proceeding to 20th century theories from both Chinese and overseas researchers.

1.1 Traditional views

As we can see from the official bibliographies in dynastic histories, starting from the Han dynasty, an increasing number of lexicographic works have been in use. The "Yiwenzhi" 藝文志 chapter of the *Hanshu* 漢書 lists 12 works under the category of linguistic works referred to as *xiaoxue* 小學, or "small learning."[4] The bibliography of the *Suishu* 隋書 lists about 110 works in the same category,[5] the *Xin Tangshu* 新唐書 over 130,[6] while the *Songshu* 宋書 over 200.[7] These *xiaoxue* works sought to enhance the understanding of the classics through studying their language. Research on the classics of the pre-Qin period, for example, launched the study of character forms, while the adaptation of Buddhist scriptures written in a foreign language propelled the study of phonetics. The *xiaoxue* works played a crucial role in enforcing the notion of a correct script, either with respect to the ancient (e.g. *Wujing wenzi* 五經文字, *Fugupian* 復古篇) or the modern (e.g. *Kuangmiu zhengsu* 匡謬正俗, *Sushu kanwu* 俗書刊誤) scripts.

The repeated carving of the Confucianist classics into stone represented the same conscious act of creating a standard sample script. The first version of stone classics was the Xiping Stone Classics 熹平石經

[3] For a more systematic historical review of the field of palaeography in general, see Gao Ming (1987), pp. 3–26. For an overview of Warring States palaeography in the 20th century, see Dong Lianchi (2000).

[4] *Hanshu*, pp. 1719–1720. The mere comparison of the number of works in the *xiaoxue* categories in different catalogues is useful only as an indicator of a general tendency because the catalogues sometimes differed in their categorization of works. For example, in the *Hanshu* (p. 1716) the *Erya* 爾雅 glossaries were under the *Xiaojing* 孝經, not the *xiaoxue* category, as they were listed in later bibliographies. We can get a glimpse of the degree of linguistic interest in imperial China by looking at the 12,067 known editions of 4,813 works in Yang (2002) which lists the linguistic works of China up to 1911.

[5] *Suishu*, pp. 942–946.

[6] *Xin Tangshu*, pp. 1447–1451.

[7] *Songshu*, pp. 5072–5079.

erected in the 4th year of the Xiping 熹平 reign period (AD 175); the second, the Wei Three-Script Stone Classics 三體石經 (AD 241); and the third, the Tang Five-Script Stone Classics 五體石經 (AD 623).[8] The creation of the stone classics in each case was an official project of the central administration directed at the enforcement of central authority.[9]

Below I discuss two medieval Chinese dictionaries, the *Ganlu zishu* 干祿字書 and the *Hanjian* 汗簡, which are two examples of how traditional Chinese scholars approached the problem of character variability. Because of the pivotal role of the Han dynasty in both the development of Chinese writing and the later image of Warring States writing, I do not include the *Shuowen jiezi* 說文解字 (hereafter: *Shuowen*) here but discuss it separately in detail in the next chapter.

1.1.1 *Ganlu Zishu*

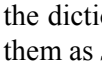
The majority of traditional books dealing with writing are concerned with standardization. After the Han, however, a smaller number of works also dealt with non-standard character forms. One of the earliest extant works on variant character forms is the *Ganlu zishu* 干祿字書 by Yan Yuansun 顏元孫 (d. 714) from the Tang dynasty. The term *ganlu* 干祿 in the title of the book referred to *Lunyu* 論語 2:18 where Confucius comments on Zi Zhang "learning about attaining an official salary" 學干祿. Thus the title of the book means a dictionary for scholars interested in official employment.

Yan Yuansun was the fourth generation grandson of Yan Shigu 顏師古 (581–645), the famous commentator of the *Hanshu*, who also wrote on nonstandard character forms in books such as the *Kuangmiu zhengsu* 匡謬正俗 and *Ziyang* 字樣.[10] In the *Ganlu zishu*, Yan Yuansun categorized

[8] Traditional sources habitually confuse the Han and Wei stone classics with each other. A common assumption, for example, was that the Han Xiping Classics were written in three different scripts, while in reality they were written in clerical script only. See Deng Dexiu (1979), pp. 5–13.

[9] Obviously, carving the text of classics into stone was also an act of creating a standard text, not only orthography. But the fact that the Wei and Tang stone classics were in different scripts shows that character forms were perhaps even more important than issues of textual criticism. Han, Wei, and even Tang stone inscriptions testify to the lack of orthographic consistency.

[10] In fact, according to the preface of the *Ganlu zishu*, Yan Yuansun largely based his work on Yan Shigu's now lost *Ziyang*. The final version of the *Ganlu zishu* was then rewritten by Yan Yuansun's nephew, the celebrated calligrapher Yan

the characters of his time into three major groups: *su* 俗 (popular), *tong* 通 (interchangeable), and *zheng* 正 (standard or correct). For example, the dictionary lists three forms for the character 毅 (), identifying them as *su*, *tong*, and *zheng*, respectively (上俗中通下正).[11]

Yan Yuansun did not regard non-standard characters as something to be avoided. Instead, he explained that each group had different uses: *su* characters were to be used in informal documents, such as registers and accounts, legal case files, and medical recipes; *tong* characters were suitable for submitted proposals, letters, and court verdicts; *zheng* characters could be used in writing essays at examinations or for carving them onto steles. Yan noted, for example, that if one used *tong* characters when those were appropriate then one could be certain of receiving no criticism. However, if one needed to prepare a composition in an educated language when applying for a position then it was more appropriate to use *zheng* forms.[12]

Yan's acceptance of non-standard characters for certain documents tells us that in the 8th century character forms falling outside the standard of the official examinations were in common use and that people were not trying to avoid them. The differentiation between the social uses of different character forms also tells us why later generations, including ours, saw a skewed picture of the writing habits of the Tang, even when looking at first-hand epigraphic evidence. Those documents that were important enough to be transmitted were written in *zheng* form, while more ephemeral documents were written using non-standard forms. The highly selective process of textual transmission has continuously weeded out documents written in non-standard forms. Consequently, the decisions of a relatively small circle of scholarly and political elites have shaped how we perceive the writing habits of the past today.[13]

Yan Yuansun's work demonstrates that the use of variant forms could be, and in his time was, an integral part of the writing system. Moreover,

---

Zhenqing 顏真卿 (708–784) around 770. This calligraphic copy became the ancestor of all extant versions, including the stele rubbings and the woodblock printed ones. For the history and relationship of various editions, see Shi Anchang (2001), pp. 1–9; for the rubbing of the stele recording the book in Yan Zhenqing's original hand, see Shi Anchang (1992).

[11] Yan Yuansun (1873), p. 22.

[12] *Ibid.*, pp. 3–4.

[13] The preoccupation of the elite with shaping their own image in the eyes of future generations was present throughout Chinese history. Starting with the references to sons and grandsons on bronze inscriptions, all the way to creation of newer and newer official histories, the Chinese expended great effort in trying to create a favorable view of themselves in the eyes of their descendants.

his account revealed that during the Tang, in terms of the general rate of literacy, only a very small circle of people used standard characters only on limited occasions, whereas the majority of the population used non-standard forms in their daily routine. Without trying to eliminate alternate character forms, people were accustomed to write with different degrees of sophistication in different spheres of society.

Although the *Ganlu zishu* studied only Tang-dynasty character forms, the reference to different levels of sophistication is significant for the current study because it shows that we cannot examine graphical inconstancy in Warring States writing without specifying the stratum of textual sophistication. The analysis of sacrificial inscriptions, for instance, might give the impression that there was little variation, whereas the study of bamboo slips with funerary inventories would lead to contrary conclusions.

1.1.2 *Hanjian*

During the Northern Song (960–1127) the practice of collecting antique artifacts and writings was at its height, which amplified the interest in ancient character forms. Among the studies dealing with ancient character forms, we find a work by Guo Zhongshu 郭忠恕 (d. 977) titled *Hanjian* 汗簡. The term *hanjian* literally means "to make the bamboo sweat", a procedure of preparing the bamboo slips for writing, more commonly called *shaqing* 殺青, or "killing the green."

Although the *Hanjian* had been important enough to be included in the "Yiwenzhi" chapter of the *Songshu*, after the Song the book became largely ignored and was nearly lost.[14] The few sources that mentioned it expressed their suspicion, or even conviction, that the character forms given by Guo were not authentic.[15]

The extant edition of the book lists almost 3,000 *guwen* 古文 character forms from 71 different *guwen* documents, including the *Guwen Shangshu* 古文尚書, *Guwen Laozi* 古文老子, and *Guwen Xiaojing* 古文

[14] During the Northern Song there was another work that followed the tradition of the *Hanjian*. This was Xia Song's 夏竦 *Guwen sishengyun* 古文四聲韻 which was basically an augmented version of the *Hanjian*, arranged phonologically. While the *Hanjian* followed the graphical arrangement of the *Shuowen*, the *Guwen sishengyun* organized the characters according to tones and rhymes.

[15] On the unfortunate fate of the *Hanjian*, as well as the arguments of its critics, see Huang Xiquan (1990), pp. 5–7.

孝經. Guo Zhongshu explained his reasons and methods behind the compilation of the *Hanjian* in the "Postface:"

> 汗簡者古之遺像，後代之宗師也…… 臣頃以小學蒞官，校勘正經石字。 繇是諮詢鴻碩假借字書， 時或採掇俄成卷軸。 乃以尚書為始， 石經說文次之，後人綴緝者殿末焉。遂依許氏各手部類不相間雜，易於檢討。
>
> The *Hanjian* is a collection of lost images (symbols) from the past, the great ancestral teacher of later generations. … Once I was in government employment in the field of *xiaoxue*, and I collated and edited the stone characters of the classics. Therefore, I have been learning from Hong Shuo's *Jiajie zishu* and the notes that I gathered here and there all of a sudden turned into a whole book. Its core primarily came from the *Shangshu*, then the Stone Classics version of the *Shuowen*. The amendments and revisions of later people were added last. After that, I arranged the characters based on Xu Shen's radicals so they would not be in disarray but could be easily located and consulted.

Under each *guwen* form, Guo included the equivalent modern character. He did not "transcribe the old form's structure into modern form" 不為隸古.[16] Instead, he recorded the characters in their original form, which today would be equivalent to listing Xeroxed or photographed copies. Generally speaking, he was trying to record character forms that differed from each other, and because of this there are not many structurally identical forms in the book. Guo also claimed to have omitted those character forms with respect to which he had the "slightest lack of knowledge." Figure 1.1 shows a sample page from the book, illustrating Guo's methodology of recording and annotating *guwen* forms.

The sample page on Figure 1.1 shows that a small-script character follows each *guwen* form in a smaller script. This small-script character is the modern equivalent of the *guwen* head entry. Certain entries also include the *fanqie* 反切 pronunciation. Characters such as 巫 (line 1, character #3), 會 (line 7, character #5), and 倉 (line 8, character #1) demonstrate that Guo translated the *guwen* forms directly into modern ones without trying to reproduce the original structure of those. Figure 1.2 shows the difference between these two approaches.

The examples demonstrate that Guo avoided *liguding* transcription and in each case picked a modern character form to explicate the character. The significance of this approach for the current study lies in treating the *guwen* forms in context and understanding them in the light of the word they represented. Consequently, the modern character forms that Guo

[16] I will discuss the method of transcribing character structure (i.e. *liding, ligu,* or *liguding*) and its limitations in the next chapter.

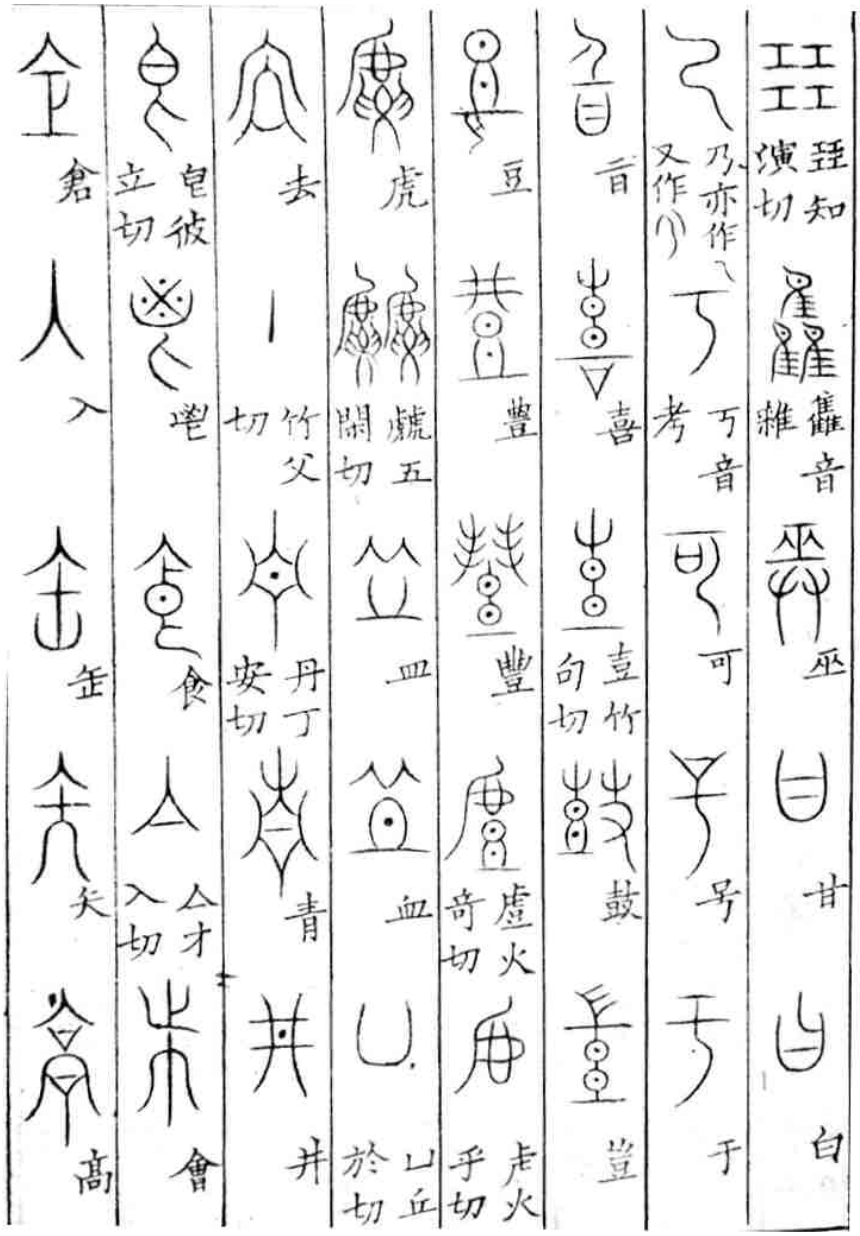


Figure 1.1. A sample page from an 1883 edition of the *Hanjian*.

added after each *guwen* form were not so much transcriptions as translations or glosses. In other words, Guo treated the various character forms as graphical representations of *words*.

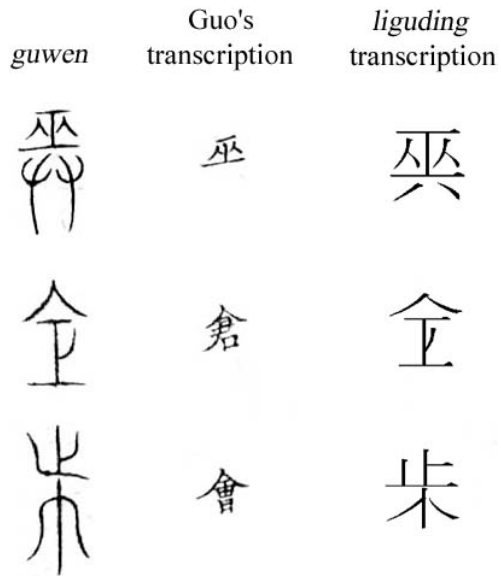


Figure 1.2. Comparison of Guo Zhongshu's transcription of *guwen* character forms with the *liguding* transcription of those. The first column shows the *guwen* entry in the *Hanjian*; the second, Guo's transcription of those entries; the third, my own *liguding* 隸古定 transcriptions which illustrate what the structure of the *guwen* forms would look if transcribed into modern forms component by component.

The treatment of *jiajie* 假借, or phonetic loan characters in the *Hanjian* reflects Guo's belief in the priority of the intended word versus individual character forms. Beside variant character forms, he also listed ancient forms of *jiajie* characters under the modern form for which they were used as a loan.[17] For example, Guo presented  as a *guwen* form of 日.[18] In reality, the graph is the *guwen* form of the character 晉 and has no graphical connection with any known form of 日. The *Shuowen* glossed 晉 as "to arrive," showing that there is no semantic link between 日 and 晉 either. But the pronunciation of 日 and 晉 was similar and sometimes one could be used as a loan character for the other.[19] There are many similar examples in the *Hanjian*, demonstrating that Guo understood every

[17] In fact, this was one of the reasons behind the mistrustful attitude of post-Song scholars towards the book, since it sometimes included character forms which were graphically unrelated to the main entry.

[18] Huang Xiquan (1990), p. 405.

[19] For concrete examples of such usage, as well as the phonetic connection between these two characters, see Huang Xiquan (1990), p. 404.

character in context, in the light of the word it represented in that particular environment.

On many occasions, Guo listed several *guwen* forms of the same character and, "in order to distinguish them from each other," he recorded the title of the text where each form came from. Variant forms of the same character in the *Hanjian* were usually from different texts. However, there were a few cases when Guo selected variant forms from the same document. For example, Guo claimed that he selected the forms of the characters 中, 契, and 好 (Fig. 1.3) from the *Guwen Shangshu*.

Figure 1.3. Variant *guwen* forms of the characters 中, 契, and 好 from the *Guwen Shangshu*.

Based on the list of *guwen* sources that Guo gave at the beginning of the *Hanjian*, we must assume that he used only one version of the *Guwen Shangshu* and that the above variant forms all had come from the same document. Accordingly, Guo sometimes worked with *guwen* documents that contained more than one variant of the same character.

The above examples demonstrate that the *Hanjian* is not only an important aid in deciphering obscure character forms in newly found manuscripts but also valuable because of its methodology. Guo treated writing on the level of words and not characters, revealing that for him it was the context, not the graphical form, that defined a character. The precedence of spoken words over written character forms is especially appropriate for the study of pre-Qin scripts because, as I will argue later in this study, Warring States writing had a relatively high level of graphical fluidity,

where a given character form was not an absolute entity but one of the several possible ways of recording a word.[20]

The *Ganlu zishu* and the *Hanjian* are but two examples of traditional works that touch upon the issue of variant character forms that date from periods of cultural revival, the Tang and Song, when *xiaoxue* studies reached a climax. The *Ganlu zishu* recorded contemporary character forms, whereas the *Hanjian* gathered forms from over a thousand years prior to its own time. Yet the two works shared their descriptive approach to written language.

Such a descriptive approach was in contrast with the majority of traditional works on characters and character structure which were concerned with teaching the correct way of writing characters. This is why variant forms, referred to under a number of different names, such as *suzi* 俗字, *biezi* 別字, *yitizi* 異體字, *tongzi* 通字, *qizi* 奇字, made sense to researchers only in contrast with the official *zhengzi* and never by themselves.

1.2 Modern views

The archaeological finds of the 20th century produced a large amount of material which had been unavailable to traditional researchers. While, as I have mentioned above, there had probably been more Zhou and Warring States manuscripts and inscriptional sources in Song China than we are accustomed to thinking, the discovery of the oracle-bone inscriptions at the turn of the century opened an entirely new direction in the study of Chinese writing. Being the direct ancestor of Warring States writing, yet predating it by eight hundred years or more, oracle-bone inscriptions carry valuable information both on the origins and nature of Chinese writing.

[20] Another important lesson we learn from the *Hanjian* is related not to its content but to its transmission. The archaeological discoveries of the second half of the twentieth century testified that many of the previously undocumented character forms in the *Hanjian* were not the product of the author's imagination but were actual examples of Warring States writing. Consequently, we know that the sources Guo worked from were available during the Northern Song. This is significant because it casts doubts on the notion that we are privileged in terms of the available archaeological material. The wealth of epigraphical data available to us today is not the only reason why we see early Chinese writing in different light as compared to traditional scholars. Song scholars already had much *guwen* material available to them, but they did not, or chose not to, utilize these sources the way we do.

Starting from the second half of the 20th century, there were also numerous finds of Warring States manuscripts. In terms of the amount of text, as well as its literary value, writings from the ancient state of Chu 楚 occupy an important place among the discovered documents. The Baoshan 包山 and Guodian 郭店 sites alone yielded tens of thousands of characters written in the distinct Chu script.[21] This sizeable corpus was further increased by the collection of Chu manuscripts that the Shanghai Museum acquired in Hong Kong in 1994.[22] The Houma and Wenxian covenant texts from the early 5th century BC are another sizable corpus of early Chinese texts.[23]

As the result of these discoveries, we now have a large amount of material for the study of pre-Qin writing. The application of Western historiographic and linguistic principles has also opened new horizons. As for the study of orthographic variability of ancient character forms, there have been very few dedicated studies on the subject, most researchers simply pointing out a few examples of variant forms. The following few scholars have discussed the phenomenon in more detail.

1.2.1 Noel Barnard

In 1958, Noel Barnard advanced his theory of character constancy in Chinese writing. He claimed to have examined an extremely large corpus of palaeographic data, including "20,000 oracle bone inscriptions," "some 10,000 bronze inscriptions," and "some hundreds of miscellaneous materials, e.g. Stone Drum text; pottery inscriptions, etc.," and as a result he was confident that the same principle of character constancy applies to Chinese writing in general.[24]

The main points of Barnard's theory were that 1) "inconstancy is not a characteristic of Chinese calligraphy, whether ancient or modern" and that 2) "it is quite permissible to regard inconstancy as a definite proof of forgery."[25] Barnard applied his own principle to prove that the *Maogong ding* 毛公鼎 and *Sanshipan* 散氏盤 vessels were modern forgeries because the inscriptions on these vessels showed cases of structural variation.

[21] *Baoshan* (1991) and *Guodian* (1998).
[22] *Shanghai Bowuguan* (2001).
[23] *Houma mengshu* (1976). The Wenxian covenant texts are in the process of being prepared for publication.
[24] Barnard (1959), pp. 28–29.
[25] *Ibid*, p. 29.

In his effort to define variation, Barnard distinguished between the structure and shape of a character, that is, structural and calligraphic attributes. He made a strong point that only structural differences could constitute inconstancy. He defined constancy as the fact that repeated characters within the same attested document had been "always written on the same structural principles – number of strokes, position and intrinsic features of stroke combinations."[26] At the same time, he did not consider differences in stroke number applicable when they were the result of simplification or ornamentation. He also discounted differences which he judged to be calligraphic mistakes or irregularities caused by the limitations of the physical space available for a character.

Barnard's principle of character constancy has been criticized by a number of people, perhaps most extensively, and passionately, Cheng Te-k'un (1971).[27] While Cheng pointed out many concrete problems in Barnard's research, a significant portion of his criticism was related to terminology. The two scholars interpreted concepts such as structure, stroke order and variation in a completely different way. Cheng, guided by his intuitions and training of a native Chinese scholar, saw even minor stroke-level changes as cases of inconstancy. Barnard, on the other hand, considered such changes insignificant and was looking for purely structural variations.

For my own purposes, I find Barnard's distinction between calligraphic and structural variations an important condition for defining variation. Even though I see character structure and its variation somewhat differently, I still commend Barnard's efforts for establishing clear and solid parameters for the study of character structure. At the same time, his application of his own criteria sometimes seems haphazard.

Generally speaking, the only difficulty I see in Barnard's approach is his lack of consistency in applying his own method. He allowed himself a great deal of flexibility defining the scope and meaning of his criteria in concrete situations. Since inconstancy depends on the definition of variation which is essentially a matter of convention, a study is valuable only if the researcher uses the same set of standards consistently across various segments of data. If Barnard, or anyone else, discerned a certain pattern in the epigraphy of Zhou bronzes then, regardless of the name he gave to this pattern, he would be able to apply this knowledge to examine the authenticity of other vessels.

[26] *Ibid.*, p. 37.
[27] Cheng (1971).

I have to add here that when I argue for the presence of character inconstancy or variability in early Chinese writing, I do not necessarily contradict Barnard's rules regarding the lack of inconstancy in Chinese calligraphy. Although at first glance our standpoints seem conflicting, we simply define the criteria for inconstancy differently. While for Barnard it is crucial to look at a single text, I examine a script shared by a community of people. I believe that Barnard's methodology is a valuable and useful approach, when applied consistently.

1.2.2 Cheng Te-k'un

Cheng Te-k'un discussed the issue of character constancy in detail in Cheng (1971), primarily to dispute Barnard's principle of character constancy.[28] Cheng's basic contention was that "inconstancy of character structures has always been a common feature of Chinese writing."[29]

To illustrate this, Cheng cited numerous examples from a multitude of sources: oracle records, attested bronze inscriptions, unattested bronze inscriptions, miscellaneous attested material, miscellaneous unattested material, Han bronze inscriptions, post-Han material, and modern writings. He moved freely from period to period, from Shang oracle-bones to the simplified characters of modern China.

Regarding the structure of Chinese characters, Cheng believed that it could not be separated from shape. He stated that "the shape of a Chinese character is determined not only by the number of strokes and the ways in which they are combined, but, more important, by the shape and movement of the strokes employed in the construction."[30] As a result, he managed to find a large number of variant forms, even though some of those differed only slightly from each other.

As a case against Barnard's theory of character inconstancy, I do not find Cheng's argument convincing. The primary reason for this is that Cheng did not pay attention to Barnard's definitions of terminology. The majority of the examples in Cheng's analysis are what Barnard defined as variations in shape, simplifications, omissions etc., therefore they are not useful to refute Barnard's theory. After discounting these examples and the unattested inscriptions from the analysis, only a handful of cases are left.

[28] This article was in fact a defense of Cheng's own views expressed earlier in Cheng (1965) and later criticized by Barnard (1965).
[29] Cheng (1971), p. 163.
[30] *Ibid*, p. 143.

In sum, more than anything, Cheng's article was a criticism of Barnard's terminology. However, for Zhou bronze inscriptions the traditional calligraphic terminology of imperial China is not necessarily applicable. Nevertheless, I agree with the general thesis that character inconstancy existed and was a common feature of Chinese writing.

1.2.3 William G. Boltz

In his book *The Origin and Early Development of the Chinese Writing System*, William Boltz described graphic variation as an apparent phenomenon in Han manuscripts. He defined it as the "variation in the way the same word is written on different occurrences."[31] He distinguished three different kinds of graphic variation: 1) "within a single text; 2) within a body of roughly contemporary texts; or 3) in the same (= corresponding) place in different versions of the same text." He himself discussed only the variations between characters in manuscripts and their transmitted versions.

Boltz also divided graphic variations into classifier and character variations. According to his definition, in classifier variation the semantic classifier changed while the phonophoric element remained the same; in character variation, the entire character changed, and the same word was written with a different character. He used the Mawangdui 馬王堆 silk manuscripts to demonstrate that both kinds of variation were common in Han writing. He uses the character 董 to illustrate that the character "may stay theoretically for any number of words" the Han pronunciation of which approximated that of 董.[32]

Observing the large degree of graphical variations in early Han manuscripts, Boltz put forward the hypothesis that had it not been for the forced systematization, Chinese writing might have evolved into an alphabetic or syllabary script. He noticed the large-scale tendency towards "desemanticization" and raised the possibility that a character often used for its phonetic value could have "become a genuine syllabograph or phonogram."[33] In this respect, while I agree with the existence of desemanticized character usage, I do not see desemantization as a trend. The reason for this is that the same phenomenon which Boltz observed in second

[31] Boltz (1994), p. 157.
[32] *Ibid*, pp. 162–163.
[33] *Ibid*, pp. 168–169. One of the examples Boltz uses is the graph 句, which in the Mawangdui corpus could equally represent the characters 後, 后, and 苟.

century BC manuscripts from Mawangdui and which he contrasted with the "equilibrium that must have existed for a thousand years, and that was characterized by a definitive use of determinatives," already existed for centuries. The Houma covenant texts I examine later in this study date to the beginning of 5th century BC and their script reveals that no such equilibrium existed. Consequently, we cannot talk about a noticeable trend towards phoneticization immediately before the Qin and Han reforms because the same semi-phonetic principle was already part of the script three hundred some years prior to the date of the Mawangdui manuscripts.

With respect to the subject of the current study, I find Boltz's discussion of graphic variation useful because it presented a logical system. Using a markedly non-Chinese linguistic approach, he always managed to look beyond the graphical shape of characters and seek out the words they represented in a particular context. His definition and analysis of graphic variation were an important incentive for this study.

1.2.4 Matsumaru Michio

Matsumaru Michio discussed the issue of character inconstancy in connection with the authentication of early Chinese bronze vessels based on their inscriptions. To prove the point that irregularities in character structure did not necessarily mean that a vessel was not authentic, he listed several cases of inconstancy (*itai* 異體) and character reversal (*hanten moji* 反轉文字) on scientifically excavated bronze vessels.[34]

He believed that the phenomenon of character inconstancy depended on the type of vessel. He divided the bronze vessels with inscriptions into two types. Matsumaru called the first type "vessels with self-initiated inscriptions" (*jiun meiki* 自運銘器), referring to primary bronzes with inscriptions that were written without a pre-existing model. He called the second type "vessels with copied inscriptions" (*rinsho meiki* 臨書銘器), referring to secondary bronzes whose inscriptions were copied from already existing inscriptions.[35]

Matsumaru claimed that cases of character inconstancy occurred only on copied inscriptions and never on initial ones. He argued that Barnard's

[34] Matsumaru (1980), pp. 66–67, 89–90. Character reversal refers to variations where the variant character forms are mirror images of each other flipped over a vertical axis. Matsumaru treats character reversal as a form of inconstancy and, as such, is only interested in reversed character pairs that occur in the same inscription.

[35] *Ibid.*, p. 90.

method of authenticating bronze vessels based on the existence or lack of character inconstancy was equivalent to seeing primary vessels as originals and copied ones as forgeries. Although technically speaking, noted Matsumaru, copied vessels could be understood as "Zhou forgeries," this was still a far cry from treating them as modern forgeries.

Matsumaru was of the opinion that the structural differences, as well as character reversals, were due to the fact that the artisans who made reproductions of initial vessels possessed a lower level of textual sophistication or were even illiterate. As an illustration to this theory, he showed that the same phenomenon could be commonly seen in writings of small children in modern Japan just learning to write.[36] He pointed out that character reversal occurred in children's writing either because the children lacked experience or because their nervous system was still immature. In the case of character reversal on ancient Chinese bronzes, the reasons were due to the problems in 1) memorizing the original form of a character and comparing it against the characters encountered in reading; 2) mastery of the distinctive features of a character necessary for character recognition. However, he cited Yibian 乙編 867 (Fig. 1.2) to illustrate that in oracle-bone inscriptions, character reversal was not uncommon on affirmative and negative charge pairs. Thus with regard to oracle-bone inscriptions, he suggested that people possessed a "certain amount of flexibility."

Although Matsumaru argued against Barnard's theory, saying that character inconstancy on bronze inscriptions did not disprove the authenticity of the vessels, both scholars shared the basic belief that scribes of early China wrote, or intended to do so, with orthographically consistent characters. Both scholars saw variant forms as errors and attributed their occurrence to some form of inadequacy on the part of the person who wrote the inscription.

Since Matsumaru only discussed this principle with reference to bronze vessels, I see his approach a valid standpoint. The differentiation between primary and copied vessels in analyzing inconstancy is also a useful distinction in understanding the reasons behind character inconstancy. However, I do not share his views that the variants were erroneous, or that the scribes who wrote inconsistently were necessarily illiterate.

The same holds true for his treatment of character reversal. The comparison of character reversal on bronze inscriptions with reversal mistakes

[36] *Ibid*, pp. 132–133.

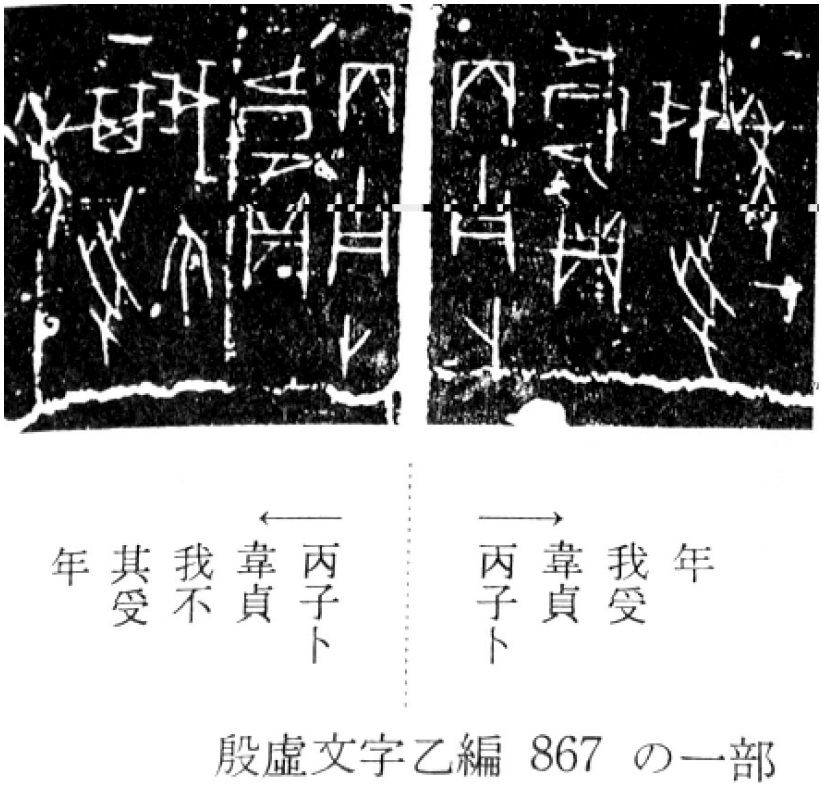


Figure 1.4. A case of reversed inscription in Yibian 867.
From Matsumaru (1980), p. 134.

in modern-day children's writing is helpful to probe into the possible reasons behind character inconstancy. But choosing children's writing as a comparison is obviously by itself an assumption about the degree of sophistication on the part of the scribes who committed such "mistakes." One would not cite, for example, variant spellings of Shakespeare's own signature and attribute their orthographic inconstancy to psychological immaturity or the lack of writing experience.

Despite this difference in opinions, I greatly benefited from Matsumaru's treatment of bronze inscriptions. His distinction between initial and copied inscriptions is an important requirement for the study of early Chinese writing. His other important contribution is that he pointed out that, with respect to character reversal on oracle-bone inscriptions, scribes and readers (literate or not) possessed a certain degree of flexibility or leniency. I shall argue later that this flexibility was perhaps the main reason behind character variability.

1.2.5 Qiu Xigui

Although in his book titled *Chinese Writing* Qiu Xigui talked about polygraphy in reference to either the current state of the script or throughout its history from antiquity up to present,[37] his discussion of terminology is extremely useful for the study of character inconstancy within written dialect. He pointed out the chaotic state of affairs with respect to technical terms and definitions used in both modern and traditional scholarship. Terms such as *tongyong* 通用 "interchangeable," *tongjia* 通假 "borrowing" have been commonly used in either their narrow or broad senses which sometimes overlapped with each other. While criticizing the confusing use of terminology, Qiu made an attempt to create better terms or to clarify the scope and meaning of already existing ones. For example, modern Chinese scholars have been using the term *yitizi* 異體字 in both a narrow and a broad sense. The narrow sense refers to the variant forms of the same character, the broad sense also includes the different written forms of the same word. In order to avoid the overlapping of terminology, some scholars coined the term *yitici* 異體詞 "variant word forms" to designate the different written forms of the same word. Qiu strongly disapproved of the use of the term *yitici* because the concept of "word forms" reflected a confusion regarding the relationship between spoken language and writing. Instead, he proposed to use the longer but more precise terms "different written forms of one and the same word" or "variant written forms of the same word."[38]

Another important issue related to character inconstancy is Qiu's distinction between "formal form" (*zhengti* 正體) and "popular form" (*suti* 俗體). He showed that these two forms have been in use simultaneously for different social purposes: "formal script refers to the standard script which was used for more serious occasions; what is referred to here as popular scripts pertains to those forms which were used daily for simplicity and convenience".[39] Qiu demonstrated that formal and popular forms could significantly differ from each other even within the same place and time.[40]

As I have mentioned it earlier in this chapter with respect to the *Ganlu zishu*, the awareness that dissimilar character sets coexisted in early China

[37] Qiu (2000), pp. 371–402.
[38] *Ibid*, pp. 394–395.
[39] *Ibid*, p. 63.
[40] Qiu's observation corresponds to Yan Yuansun's claim that *zheng*, *su*, and *tong* characters were used for different purposes. See 1.1.1 above.

is essential for the study of structural variability. It means that from the point of view of character structure, a corpus should be defined not only in geographic and temporal terms but also by the usage and social function of the script. The results of the analysis of only the popular or only the standard scripts are not automatically valid for the other type and, if taken to be representative of Chinese writing in general, create a distorted image.

As for Qiu's approach, I see its value for the current study in demonstrating the significance of defining terminology and eliminating the "gray areas" in the field. He showed that a set of clearly defined and understood terms is not only a tool to express related issues but also a mental framework to organize one's thoughts. Another important lesson from Qiu's work is that it is essential to be aware of the social function of texts we examine.

1.3 Conclusions

The majority of the issues in this chapter concerns terminology. The Barnard-Cheng debate is a good example of how important it is to work with clearly defined criteria organized in a logical way. As a counter-example, we have Boltz and Qiu both of whom are very careful and methodical in defining their terminology.

Generally speaking, all those who have discussed the issue have agreed that variation is a matter of structure and is unrelated to shape, i.e. to calligraphic style. The only exception to this consensus is Cheng who claimed that form and shape are inseparable. Another major issue in defining variation is whether the basis of variation is the word or the character. In other words, are we analyzing different graphical representations of the same word or just different forms of the same character? Qiu and Boltz were well aware of this distinction and clearly defined both approaches. Guo and Boltz analyzed the character forms in context, taking into account the words they were originally meant to denote. Others, such as Yan Yuansun, Barnard, Cheng, and Matsumaru were interested only in the variations of the character itself.[41]

[41] An important contribution with regard to the connection between words and characters is Professor Wang Li's scholarship on Chinese word cognates, especially his *Tongyuan zidian* 同源字典 (Wang, 1982). I do not discuss him in detail here because his work did not focused not so much on writing as on language itself.

As for the descriptive versus prescriptive methods of analysis, despite the predominantly prescriptive attitude of traditional works, both of the medieval dictionaries I discussed in this chapter, i.e. the *Ganlu zishu* and the *Hanjian*, treated the script from a descriptive point of view. Among modern scholars, it is Cheng, Boltz, and Qiu who have been most successful in employing a descriptive methodology.

With regard to limiting the corpus examined, we have seen a wide range of approaches. Yan Yuansun collected contemporary character forms from the mid-Tang period; Guo Zhongshu examined *guwen* character forms in Warring States documents; Barnard compared single texts from early China; Boltz compared characters in a manuscript with their equivalents in corresponding places of the same text's transmitted version. The other major issue bearing on the corpus was the coexistence of formal vs. popular spheres in writing. Qiu drew attention to this phenomenon and used it to explain certain problems in the evolution of the script. Among the traditional sources, the Yan Yuansun confirmed the same point by pointing out the distinct uses of the *tong*, *su*, and *zheng* scripts in Tang society.

As for the existence of character inconstancy, there have been conflicting opinions. The *Hanjian* occasionally listed structurally different character forms from the same document, without commenting on them. Barnard believed that variant character forms had never existed in the same document and their presence meant that the inscription was a modern forgery. Matsumaru's view was somewhat different, claiming that literate people in Zhou China did not write variant forms within the same document; it was the illiterate copyists who created variants when they made reproductions of pre-existing inscriptions. Cheng strongly argued in favor of the existence of character inconstancy, although his criteria was very different from that of Barnard. Boltz described graphic inconstancy as an apparent phenomenon in Han manuscripts.

CHAPTER TWO

THE QIN AND HAN CREATION OF THE STANDARD

Scholars of post-Han China have traditionally regarded the Qin writing reform as a turning point in the history of the Chinese script. They saw the reform as consisting of two main aspects. First, the Qin unified the writing system and eliminated the scripts of the other states; the standardized Qin small seal (*xiaozhuan* 小篆) form was the final product of these efforts. Secondly, the Qin administration adopted the clerical script (*lishu* 隸屬) for official use. With the fall of the Qin, the small seal script disappeared from general use and the clerical script became the dominant form of writing.

The earliest extant historical record of the reform occurs in the *Shiji* 史記. Only a hundred years after the reforms, Sima Qian 司馬遷 (145–186 BC) documented that, beside standardizing the legal code, axle widths, and weight and length measures, the First Emperor 秦始皇 (r. 246–210 BC) also unified the writing system.[42] In view of the overall significance of the reform, it is surprising to see that this is the only reference to it, as the *Shiji* gives many detailed explanations with respect to important historical events. Sima Qian referred, for example, to the infamous burning of books on numerous occasions.[43]

The study of the Qin reform is important because if we knew how it affected writing, we could retrace the changes made to the script and partially reconstruct the original state of Warring States writing. However, most of our knowledge of the Qin reforms derives from two sources. These two sources are the description after the *Xiaoxue* section of the *Hanshu* "Yiwenzhi" and the *Shuowen* "Postface." These two accounts resemble each other so closely that there is no doubt that they are two versions of the same story. In later times, traditional scholarship uniformly relied on these two sources to interpret the Qin unification and this has led to misconceptions. One such is that of accepting the characters in Warring States writing as having structural consistency.

[42] *Shiji* 6, p. 239.
[43] E.g. *Shiji* 6, p. 280; 28, p. 1371; 32, p. 1499; 121, pp. 3116, 3124, 3126; 130, p. 3319.

I believe that both traditional and modern scholars have had a skewed image of pre-Qin writing because of the ongoing orthographic standardization since the end of the Warring States period. The Han was a period of standardization of knowledge. Han scholars, as part of the contemporary philosophy to map the universe onto a systematic grid of correlations, depicted writing as a more consistent system than it really was. However, it is worth remembering that contemporary people saw this standardization not as the creation of a regulated system but as the depiction of the true nature of things. The ultimate role of an individual was to be in accord with the true nature of things. The Mawangdui manuscript sometimes referred to as *Huangdi Sijing* 黃帝四經 says: "Accord means life … disaccord means death" 順則生……逆則死.[44] Thus the Han image of the universe was, like any comprehensive ideological system, a prescriptive view that claimed to be descriptive. Later times looked back at the Han as the model for effective standardization, even if this image was not in accord with historical reality.

In this chapter, I try to look at the Han accounts with a fresh eye, without the bias of later interpreters. I employ epigraphical evidence to compare the actual state of writing during the Warring States and Qin-Han periods with the Han descriptions. This allows me to demonstrate that rather than describing the Qin reform as it really happened, Han scholars recorded an idealized version of the event. Similarly, their view of the writing habits of their own times differs from the epigraphical evidence.

Karlgren's account of the relationship between the Qin standardization of the script and the *Shuowen* account represents, and is itself an example of, former views. It is precisely this kind of views that I take issue with in this chapter.

> "The epoch-making work of Xu Shen is so much the more valuable as it was published only three centuries after Li Si and as therefore an unbroken tradition must have continued to the time of Xu about the interpretation of most of the characters."[45]

[44] Chen Guying (1995). The manuscript discusses in detail the concepts of accord and disaccord as a general principle of arranging one's conduct in harmony with the forces of the universe.

[45] Karlgren (1923), p. 3, n. 1.

2.1 The *Hanshu* and the *Shuowen* accounts

The "Postface" of the *Shuowen* is one of the most commonly cited descriptions of the Qin reform of writing. But almost the same description appears in Ban Gu's 班固 (AD 32–92) history following the *Xiaoxue* section of the "Yiwenzhi" chapter of the *Hanshu*. The two accounts, as I have already indicated, are so close to each other that one has to consider them two versions of the same text.

Ban Gu based the "Yiwenzhi" chapter on Liu Xin's 劉歆 (ca. 46 BC–AD 23) *Qilue* 七略 which, in turn, followed Liu Xiang's 劉向 (79–78 BC) *Bielu* 別錄. This information is known from Ban Gu's introduction to the "Yiwenzhi," which claims that the chapter recorded the essence (其要) of the *Qilue*.[46] This does not tell whether the account dealing with the history of writing came from the *Qilue* or the *Bielu*, and we can only treat the "Yiwenzhi" as a text that dates to the second half of the first century AD. If this is true, then Ban Gu's account predated the parallel passages in the "Postface" to the *Shuowen*, which was written around AD 100.

Comparing the two accounts reveals that *Shuowen* version is longer and more detailed, containing almost the complete text of the *Hanshu* version. The only additional information in the *Hanshu* version are a few bibliographical sentences not related to the history of writing. Still, there are also some discrepancies. For example, only Xu Shen mentioned that Li Si discarded the scripts which were not in accordance with the Qin writing. The same is true for the discontinuation of the *guwen* script. As a general rule, the *Shuowen* appears to be more subjective with respect to the role of the Qin than the *Hanshu*, which mentioned neither the termination of the scripts of the six states and the discontinuation of the *guwen* script, nor the burning of books.

A detailed study of the relationship between these two accounts would go beyond the scope of this study. For my purposes, it is sufficient to establish that these two descriptions come from a common source. Regardless of which one of them was borrowing from the other or whether both were based on a third source, we can ascertain that the account reflected a first century AD opinion. Because the *Shuowen* account incorporates most of the "Yiwenzhi" account, I use the former to discuss the Han views of the Qin reforms. I only cite the "Yiwenzhi" when it differs from the *Shuowen*, and only when that difference is relevant to the issue at hand.[47]

[46] Ban Gu described the genealogy of his bibliography at the beginning of the chapter. See *Hanshu*, p. 1701.

[47] An aligned parallel text and translation of both accounts is found in Appendix I.

2.2 The *Shuowen* small seal script

In extant editions of the *Shuowen*, a seal form of the character heads the entry, whereas the gloss and explanation follow in the modern script. In the "Postface" of the dictionary, Xu Shen also mentioned that he used the seal script for the head entries:

> 今敘篆文，合以古籀。
>
> Now I arranged the [dictionary] based on the seal forms, and supplemented them with *guwen* and *zhouwen* forms.

Although Xu did not state clearly whether "seal forms" referred to the small seal *or* great seal (*dazhuan* 大篆) forms, there are two major reasons to believe that he meant small seal forms. The first reason is that out of the three instances when he used the term *dazhuan*, twice he used it as the title of a book composed by the Grand Historian Zhou 太史籀.

> 及宣王太史籀，著《大篆》十五篇，與古文或異。
>
> In the time of King Xuan, Zhou, the Grand Historian composed the *Dazhuan* in fifteen chapters, [the character forms of which] were somewhat different from the *guwen* script.
>
> 秦始皇帝初兼天下， 丞相李斯乃奏同之， 罷其不與秦文合者。 斯作《倉頡篇》。 中車府令趙高作《爰歷篇》。 大史令胡毋敬作 《博學篇》。皆取史籀《大篆》，或頗省改，所謂小篆也。
>
> When for the first time [in history] the First Emperor of Qin united the subcelestial world, Li Si, his Grand Councilor presented a proposal to unify them (i.e. the scripts or character forms) and discard what was not in accord with the Qin script. Li Si wrote the *Cangjie pian*, Zhao Gao, the Keeper of Carriages, wrote the *Yuanli pian*, and Humu Jing, the Grand Historian wrote the *Boxue pian*. All of them borrowed [characters] from Historian Zhou's *Dazhuan*, although they sometimes altered and abbreviated those. This was what they called the small seal script.

In the body of the *Shuowen*, when he listed examples of alternate ancient character forms, Xu Shen only used the term *zhouwen* 籀文, never *dazhuan*, to refer to the character forms of Historian Zhou. This implies that Xu Shen used the term *zhouwen* to indicate the "great seal script". The physical appearance of *zhouwen* forms led some scholars, e.g. Wang Guowei 王國為, to believe that the *zhouwen* were in fact Zhou bronze inscription forms.

The third time Xu mentioned the *dazhuan* script, he used the term to specify one of the eight calligraphic styles in use after the fall of the Qin.

自爾秦書有八體： 一曰大篆， 二曰小篆， 三曰刻符， 四曰蟲書， 五曰摹印， 六曰署書， 七曰殳書， 八曰隸書。 漢興有草書。

From this time on, the Qin script had eight styles: the first one was the *da-zhuan* (i.e. great seal script); the second the small seal script; the third the tally carving script; the fourth the insect script; the fifth the seal stamping script; the sixth the title script; the seventh the spear inscription script; the eighth the clerical script. From the rise of the Han there was the draft script.

Although there is no indication to what the great seal script was, a few lines later Xu Shen explained the small seal script:

時有六書： 一曰古文， 孔子壁中書也。 二曰奇字， 即古文而異也。 三曰篆書， 即小篆。 四曰左書， 即秦隸書。 秦始皇帝使下杜人程邈所作也。 五曰繆篆， 所以摹印也。 六曰鳥蟲書， 所以書幡信也。

At this time, there were six scripts. The first one was the *guwen* script, which was the script on the documents found in the wall of the house of Confucius; the second the odd characters, i.e. the unusual *guwen* forms; the third the seal script, i.e. the small seal script; the fourth the attendant script, i.e. the Qin clerical script created by Cheng Mao from Xiadu under the commission of the First Emperor of Qin; the fifth the pseudo-seal script used for seal carving; the sixth the bird and insect script used to write on banners.

The above passage confirms that when Xu used the term “seal script” without a modifier, he was referring to the small seal script. Since in the above passage Xu used “seal script” as a technical term, we can exclude the possibility that he might have used it with reference to the great seal script.

In view of the above, we can ascertain that the seal forms Xu Shen used as head entries in the *Shuowen* were in fact small seal forms. He also stated that the small seal script was created by Li Si in the course of reforming the script. This could be, and often has been, taken to mean that the *Shuowen* recorded the character forms created by Li Si. It is important to establish, however, that Xu Shen never made such a claim. During the approximately three hundred years that separated Xu Shen from Li Si’s reforms, it would have been quite natural if the small seal script had evolved further and undergone some changes.

The official Qin small seal script can be seen on the numerous examples of the First Emperor’s edict on the standardization of the measures of capacity, weight, and length. The Qin administration had the edict inscribed on scales and weights, as well as separate bronze edict plates (*zhaoban* 詔版) which were attached to those scales and weights that passed the official inspection (Fig. 2.1).

The inscription on the edict plates usually was the same text:

廿六年， 皇帝盡併兼天下諸侯， 黔首大安， 立號為皇帝。 乃詔丞相狀、 綰： 法度量則， 不壹歉疑者， 皆明壹之。

Figure 2.1. A rubbing of a commemorative inscription from a Qin edict plate. After Wang Hui (1990), Part II., p. 102.

> In the twenty-sixth year [of his reign] (= 221 BC), the Emperor completely unified the feudal lords of the empire,[48] brought great peace to the people, and assumed the title of the emperor. Therefore, he ordered his Grand Councilors Zhuang and Wan to standardize the measures and normalize those which were suspect to being irregular.[49]

There have been many such inscriptions found on measure tools. The inscription was always the same, regardless of the object on which it had been inscribed. Often the Second Emperor of Qin 二世 had his own edict inscribed after the edict of the First Emperor,[50] suggesting that some of the measure tools had been used continuously during the reign of both emperors.

If we compare the *Shuowen* small seal script with the small seal forms on the Qin edict plates, we can see that in many cases the *Shuowen* small

[48] There has been some debate whether the sentence breaks before or after 諸侯. The two possibilities are: a) 併兼天下諸侯，黔首大安 – unified the feudal lords of the empire, brought great peace to the people; b) 併兼天下，諸侯黔首大安 – unified the empire, brought great peace to the feudal lords and the people.
I adopted the first punctuation in my translation, following Wang Hui's (1990, p. 109) argument.

[49] The meaning of the character 則 in 法度量則 is also debated. The two main opinions are whether it is a noun or adverb in this context. Wang Hui (1990, p. 109) argues, based on the gloss of 則 in the *Shuowen*, according to which the meaning of the character is to draw an object after a model, that 則 is a touchstone used to test weights and scales.

[50] The Second Emperor did not invent the custom of carving additional edicts onto existing ones. There is at least one extant example when the First Emperor appended his edict onto a Warring States weight. See Wu Hung (1979), p. 36.

seal forms differed from the actual Qin ones. For example, the *Shuowen* gave four forms for the character 則 (Fig. 2.2). Form #1 is the small seal form heading the entry for the character in the dictionary. Forms #2 and #3 are the *guwen* forms and form #4 is the *zhouwen* form. The *Shuowen* small seal form (#1) consisted of the components 貝 + 刀, just like the modern form of the character. In contrast with this, the edict plate forms could be written in a variety of different ways (Fig. 2.3).

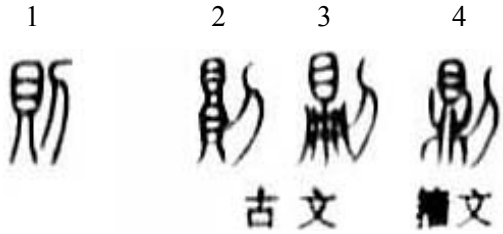


Figure 2.2. Four character forms given in the *Shuowen* for the character 則.

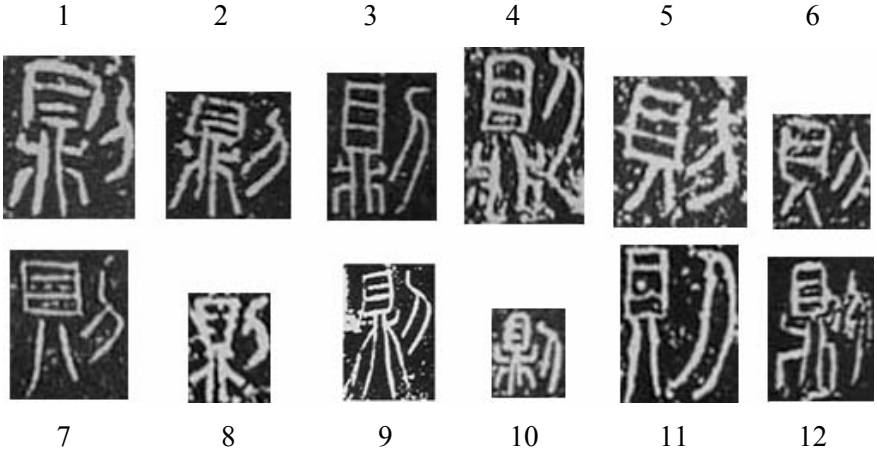


Figure 2.3. Forms of the character 則 on Qin edict plates. From Wang Hui (1990), Part II., pp. 98, 102, 108, 111, 131–133, 145, 147, 148, 156.

Out of the twelve forms, only four are structurally identical to the *Shuowen* small seal form: #5, #6, #7, and #11. The majority of the Qin character forms consists of not 貝 + 刀 but 鼎 + 刀. Xu Shen presented this structure as the *zhouwen* form in the *Shuowen*, indicating that this form, contrary to the evidence seen here, had been eliminated by Li Si during the reform of the script.

The observation that character forms on Qin bronze inscriptions were not identical to the small seal script of the *Shuowen* is not entirely new. Wu Hung 巫鴻 has performed an analysis of the Qin weights and in the course of his study, he touched upon the discrepancies of character forms.[51] Figure 2.4 displays a few examples of character forms from Qin bronze weights compared with each other and the *Shuowen* small seal form.

Figure 2.4. Variant character forms from inscriptions on Qin weights in Wu Hung (1979), pp. 36–37. The asterix sign (*) under some of the forms indicates that the form occurred on more than five different inscriptions. The numbers under the character forms indicate the object number in Wu Hung's catalogue.

Wu Hung called the first character form in each line, which he had most likely taken from the *Shuowen*, the "Qin standard small seal script" 秦代標準小篆. From the above evidence he deduced that the inscriptions were written in a less formal script rather than the standard Qin small seal script. He thus felt justified in rectifying earlier views on the nature of the script in which the Qin edicts were written. One such was Qiu Xigui's opinion that the Qin officials wrote the official decrees and laws with small seal script, never with the clerical script used for penal administration.[52] However, Wu regarded the *Shuowen* forms as the true image of the Qin small seal script and viewed the archaeological material in the light of this assumption.

[51] Wu Hung (1979).
[52] Qiu (1974).

As a slight modification to Wu's conclusion, I believe that the above evidence demonstrates that the small seal script displayed in the *Shuowen* did not match the Qin small seal script. Even within the short text of the First Emperor's edict there are cases of discrepancy when compared with the *Shuowen* small seal forms. This tells us that the *Shuowen* small seal script was not the original Qin seal script. Consequently, we have to distinguish between the Qin small script used at the end of the second century BC and a Han small seal script used around AD 100.

2.3 Structural variability within the Qin small seal script

Xu Shen organized the characters in the *Shuowen* under 9,353 head entries. While this arrangement was necessary for a convenient use of the dictionary, it also created the impression that the small seal script used for the entries was orthographically uniform. Although the *Shuowen* also listed 1,163 variant forms (重文), these appeared under, and could only be accessed through, the main entries. The same was true for the few cases of alternate forms (或體). Thus the organization of the *Shuowen* implied that the character forms of the small seal script were standardized.

However, we have already seen that sometimes the character forms on the Qin bronze inscriptions were different not only from the *Shuowen* forms but also from each other. In the case of the character 則, there were at least two major forms, both of which were commonly used. One of the twelve examples (#8) in Table 2.2 above can be understood as a transitional form between the two dominant forms.[53] In this transitional form, the left side of the character, written as 貝 in one form and 鼎 in the other, was a simplified version of 鼎 which differed from 貝 only in the couple of extra strokes at each side of the bottom part of the character.

The examples given by Wu Hung (Table 2.3) also testify to the lack of uniformity in Qin writing. Wang Hui cited several cases of inconsistent character forms on Qin measures as well.[54] One such case was the character 歉 which was written with different radicals on different inscriptions.

[53] The word transitional here, as well as later in my discussion, does not imply a progressional move between two final points. I only use the term as a temporary definition for specifying the position of a hitherto unknown form with respect to two already known forms. Depending on one's point of view, any form could be thought of as transitional.

[54] Wang Hui (1990), p. 108.

Based on the large number of discrepancies in attested bronze inscriptions, Wang came to the conclusion that "the Qin policy on the standardization of the script had not been put into practice rigorously."[55]

I have to agree with Wang Hui's conclusion. The differences between the character forms on the edict plates prove that the seal script in common use during Li Si's time was not thoroughly consistent. Despite the reforms, most of the population still wrote characters with variable structures. It is also possible that the Qin had a higher degree of tolerance towards variation and the forms on the edict plates still fell within the accepted scope of variation. All in all, it seems that Xu Shen, and his later interpreters, seriously overstated the effect of the writing reform.

2.4 Structural variability within the Han small seal script

Above we have seen that the Qin small seal script of the First Emperor was not identical to the small seal script used by Xu Shen in his dictionary. Beside the transmitted text of the *Shuowen*, the state of the Han small seal script can also be seen on the Yuanan stele 袁安碑 now held in the Henan Provincial Museum. The stele dates to AD 117, thus it is contemporary with the *Shuowen*. The 139 small seal characters of the inscription were executed with care and precision, closely resembling the uniform calligraphy of forms found in the *Shuowen*.

Even within the short text of the inscriptions, there are 25 characters that occur at least twice, some as many as 9 or 10 times, making the inscription an ideal resource for examining the variations in character structure. Comparing the repeated characters reveals that almost all of them are structurally identical. The only exception is the character 四, occurring four times but written in two distinct ways (Fig. 2.5).

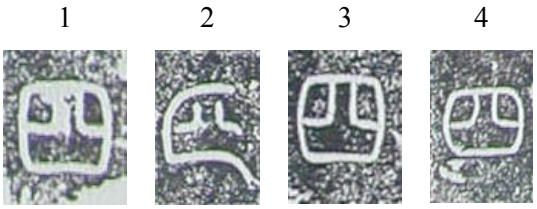


Figure 2.5. The instances of the character 四 on the Yuanan stele.

[55] *Ibid.*

Form #2 in Figure 2.5 resembles the character 匹 rather than 四. The context, however, makes it clear that this form stands for the word "four": "the fourth month of the fifth year" 五年四月.[56] The archaic pronunciations of 匹 and 四 were distant enough to exclude the possibility of a *jia-jie* usage. Since the *Shuowen* glossed 匹 as a length measure equivalent to four *zhang* 丈, there was a loose semantic connection between the two characters. Still, the most obvious link between the two variant forms is graphical similarity.

Despite this single graphic variation, we can say that the rest of the repeated characters on the Yuanan stele, show a high degree of consistency. The calligraphic style of the characters is close to the *Shuowen* small seal forms, but there are some differences in the structure of characters (Fig. 2.6).

Figure 2.6. Comparison of the Yuanan and *Shuowen* small seal forms.

The Yuanan and *Shuowen* forms of character #1 (徒) differ greatly in overall balance and arrangement, even though the basic components (彳, 止, and 土) are the same in both cases. The *Shuowen* form of character #2 (樸) has an extra vertical stroke, whereas the Yuanan form is identical to

[56] The fact that the variant form appears carefully incised into stone in a regular calligraphic hand shows that it cannot be written off as a "slip of the hand".

the modern form.[57] The difference between the two forms of character #3 (謁) lies in the right bottom part where, in the *Shuowen* form, one element encircles the other. The middle part of the Yuanan form of character #4 (賓) resembles 夕 (as in 多 or 名), whereas the *Shuowen* form has 正 at the corresponding place.

Despite these differences, both the *Shuowen* and the Yuanan stele, taken by themselves, presented a uniform image of the small seal script. Only when they are compared with each other do we realize that the uniformity was local. The conclusions are twofold. First, the differences between the Yuanan and *Shuowen* forms show that even in AD 100 the Han small seal script was not uniform. Second, the *Shuowen* idealized the small seal script and presented its own version of that ideal.

2.5 The reality of the Qin reforms

Xu Shen described the birth of the small seal script as an effort to create a uniform national standard. He ascribed this task to three people:

> Li Si wrote the *Cangjie pian*, Zhao Gao, the Keeper of Carriages, wrote the *Yuanli pian*, and Humu Jing, the Grand Historian wrote the *Boxue pian*. All of them borrowed [characters] from Historian Zhou's *Dazhuan*, although they sometimes altered and abbreviated those. This was what they called the small seal script.

However, the Qin edict plate inscriptions demonstrate that the Qin small seal script at the time of the reforms was still characterized by a large degree of inconsistency. These official documents recorded the imperial edict and thus had to be written in the official script of the state, which was the Qin small seal script. This implies that, contrary to Xu's claim, the script did not change overnight as the result of three people's work.

Epigraphical evidence shows that the changes were the result of a gradual historical process that began before the establishment of the Qin dynasty and lasted far into the Han, possibly even longer. Inscriptions made prior to the Qin unification testify that the small seal script was already in

[57] The *Ganlu zishu* records the same two forms of 樸 existing during the Tang, explaining that the one with the extended vertical stroke (*Shuowen* form) was the popular form, whereas the one without the stroke (Yuananbei) was the formal form. Yan Yuansun (1873), p. 163.

general use before the reforms. Zhu Dexi 朱德熙 noted that the inscription on the Xinqi hufu 新鄪虎符 from 16 years prior to Li Si's reform was already written in a script that had no noticeable difference with the small seal script.[58] Therefore, as far as the creation and the origin of the small seal script goes, Xu Shen's account does not correspond to the archaeological evidence.

The Mawangdui silk manuscripts provide an even more convincing example of the gradual processes shaping the evolution of the script. Although Tomb #3 dated to 168 BC, the documents found inside were written in different times. The editors of the *Mawangdui boshu yishu* 馬王堆帛書藝術 volume divided the documents of the silk manuscripts into three groups based on their calligraphy. These also represented temporal differences.[59] They categorized the first group as *zhuanli* 篆隸 (seal/clerical) manuscripts, referring to documents written in a script where the characteristics of the small seal script were stronger than those of the clerical script. The writing style of the documents written in the *zhuanli* script resembled the calligraphy of the Chu manuscripts found in Warring States tombs, whereas the structure of characters generally followed the Qin small seal script. The visual similarity with the Chu script is understandable because the manuscripts came from a tomb found on the former territory of the ancient state of Chu.[60] The second group of manuscripts was characterized as *guli* 古隸 (ancient clerical) manuscripts, referring to documents written in a type of clerical script that still retained elements of the seal script. The third group was called *Hanli* 漢隸 (Han clerical) manuscripts, indicating documents written in the Han clerical script which was completely distinct from the seal script.[61]

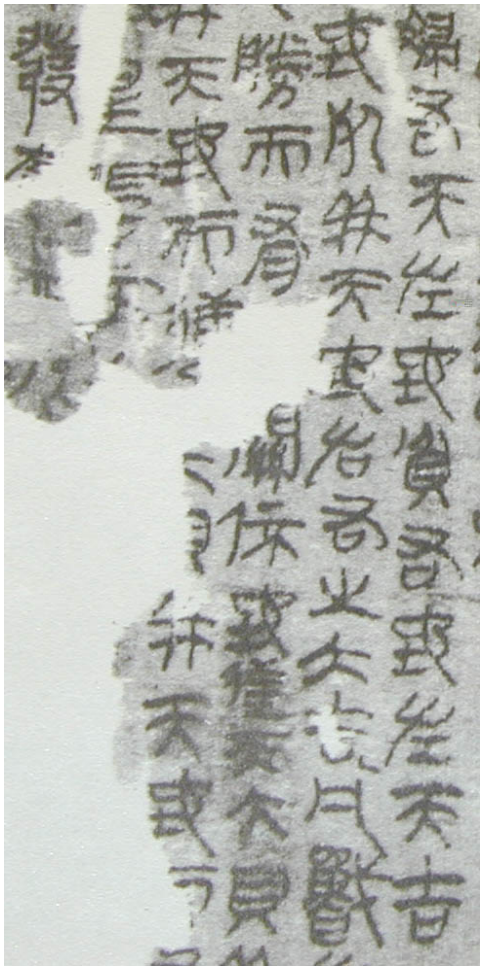
Beside demonstrating the lack of a clear-cut borderline between seal and clerical scripts, the Mawangdui manuscripts also reveal that character structure was not standardized during the Western Han. For example, the organizers of the Mawangdui corpus classified the *Yinyang wuxing jiapian*

[58] Zhu Dexi (1995), p. 78

[59] Demonstrating the transitional nature of these scripts, scholars have been using different terms to describe the scripts of the Mawangdui manuscripts. Some have distinguished, for example, a purely seal script from a purely clerical script version of the *Yinyang wuxing* 陰陽五行 manuscripts. See *Mawangdui* (1996).

[60] Boltz (1994) noted that the Mawangdui manuscripts, which were "fairly representative of the script that was in conventional use within a decade or two of the Qin reforms," were untouched by the reforms, but he attributed this to the fact that Chu was in the south.

[61] *Mawangdui* (1996), pp. 2–3.



凡道　婦右天左地貧右地左天吉怀地逆天娠
怀天逆地死并天地左右之大吉凡戰左天右地勝怀
天逆地勝而有　闢怀地逆天大貝并天地而左右
之一敫十并天地而逆　大貝并天地而右之王戰

Figure 2.7. Section of the Mawangdui *Yinyang wuxing jiapian* manuscript with a transcription of its larger context. From *Mawangdui* (1996), p. 24. In the transcription, the first vertical line of the manuscript portion shown here starts with the character 婦 and ends with the character 吉, which are characters #3 and #13 in the transcription.

陰陽五行甲篇 manuscript (Fig. 2.7) as a *zhuanli* document, representing the oldest cluster of writings. The interesting thing here is the two distinct forms of the character 左. In the section above, this character occurs three times; twice in the first line, in the phrases 右天左地 ("protect Heaven and assist the Earth") and 右地左天 ("protect the Earth and assist Heaven"), once in the second line, in the phrase 並天地左右之大 ("combine the greatness of the protection and assistance of Heaven and

Earth").[62] In the first two instances, the character is written with the 工 component as  , in the third with the 口 component as  .

Both of these forms are well-known ways of writing the character 左. The difference between them is that, structurally, the first one is the clerical form, whereas the second small seal. Since the context reveals that there is no difference in the usage of the word, we have to conclude that these two forms were in concurrent use, which speaks against a strict, or strictly observed, orthographic standard at the time. It also shows that the transition between small seal script and clerical script was gradual.

The orthographic inconsistencies in the above documents indicate that Li Si's reform of writing was not universally adopted. This is also confirmed by the fact that the Second Emperor of Qin had to reissue edicts regarding the standardization of weight measures. Apparently, the original edict of his father had not fully achieved its purpose. Indeed, analyzing the physical weight of the Qin weights, Wu Hung observed that the very objects that had edict plates attached to them as a sign of having passed the official inspection showed variation in weight.[63] Such discrepancies can be observed on attested, well-preserved weights that had concrete values inscribed on them.

In light of the above examples, we can say that Li Si did not fully transform the nature of writing in China through his reforms. The main role of his reforms on the grand scale of the evolution of writing lay in establishing the Qin small seal script as the official script in the country. However, the changes leading to the regularization of the Qin small seal and clerical scripts happened gradually over the course of several centuries. Evidence shows that the Qin efforts directed at standardization were not always as immediately successful as depicted in traditional sources.

2.6 Inconsistency as a moral issue in the Han

The Qin bronze inscriptions attest to the lack of a uniform standard in Qin times. The Mawangdui manuscripts from the Han show an even greater degree of structural variability. This variability manifested itself in both the composition of individual characters and the use of *jiajie* loan

[62] On the manuscript section in Figure 2.7, 左 is characters #4 and #9 in the first line, and #6 in the second.

[63] Wu Hung (1979), pp. 38–40.

characters. Such freedom testifies to the flexibility on the part of the writer and the tolerance on the part of the reader.

The existence of inconsistency in the Han can be deducted from the *Shuowen* "Postface" and the *Hanshu* bibliography.

Shuowen:

書曰：「予欲觀古人之象。」言必遵修舊文而不穿鑿。孔子曰：「吾猶及史之闕文，今亡矣夫。」蓋非其不知而不問。人用己私，是非無正，巧說邪辭，使天下學者疑。

The *Shangshu* says, "I wish to observe the images of the ancients." This means that one must follow and study the old text and not go into strained explanations." Confucius said, "I still go back to the times when the historian left empty spaces in the text. Today there is no such thing anymore!" This is a condemnation of not asking when one does not know. When people write according to their own fancy, right and wrong have no constancy, clever explanations and wicked words create doubts among the scholars of the world.

Hanshu:

古制，書必同文，不知則闕，問諸故老，至於衰世，是非無正，人用其私。故孔子曰：「吾猶及史之闕文也，今亡矣夫！」蓋傷其浸不正。

In the old system, documents had to be written with identical characters. If you did not know something, then you left an empty space and asked the elders. Once the world went into decline, right and wrong had no constancy, and people wrote according to their fancy. This is why Confucius said, "I still go back to the times when the historian left empty spaces in the text. Today there is no such thing anymore!" He was upset because of the inundation [of writing] with irregularities.

Both the *Shuowen* and the *Hanshu* quote *Lunyu* 15:25 where Confucius supposedly condemned the practice of his time to write without regard to the integrity of the text.

子曰：「吾猶及史之闕文，有馬者，借人乘之，今亡矣夫！」

The Master said: "I still go back to the times when the historian left empty spaces in the text. Those who had a horse, would lend it to someone else to ride. Today there are no such things anymore!"

The translation that I gave here is based on the Han interpretation of the passage as seen in the *Shuowen* and the *Hanshu*. In reality, however, the meaning of this passage is problematic. Numerous commentators and translators have offered different interpretations.[64] Most of them followed

[64] To illustrate the difficulties translators have encountered trying to interpret the passage, below I cite a few translators, whose translations are based on Chinese scholarship.

Xu Shen's and Ban Gu's interpretation and differed only in minor points. Still, I believe that the truth lies elsewhere. In an article specifically devoted to this problem, Zhou Cezong 周策綜 has convincingly demonstrated that the phrase 史之闕文 in the *Lunyu* originally referred not to the "empty space left by a historian" but to the official's lack of refinement.[65] The character 文 should be understood in the same way as it appears, for example, in *Lunyu* 6:18:

子曰：「質勝文則野，文勝質則史。文質彬彬，然後君子。」

The Master said: If a man's innate quality outweighs his learned sophistication, he will be wild; if his learned sophistication outweighs his innate quality, he will be like a historian. When his learned sophistication and innate quality are in harmony, only then can he become a true gentleman.

Here 文 refers to education of which a wild barbarian has very little and a bureaucrat has too much. Thus for Confucius, the historian is a person whose education and manners surpass his personal qualities. Viewed from this perspective, we ought to translate the phrase as "an official lacking refinement." The character 史 occurs in the *Lunyu* only 3 times, twice in the passages cited above (6:18 and 15:25). In the third case (15:6) it refers to the Historian Yu 史魚 and should be rather understood as a personal name. But we have seen that in *Lunyu* 6:18 it means a bureaucrat, an official in whom education or refinement exceeds his innate qualities. It does not specifically refer to historians per se but is used in a slightly pejorative way to indicate learned individuals who are too much bound by cultural values.

"I have come across doubtful places in historical chronicles. The owner of a horse would lend it to another to ride. There is no such thing today!" Perelomov (1993), p. 416. A footnote comments, "Commentators unanimously point out that there is an omission in the text, since there is no logical connection between the two phrases."

"I have arrived as it were at the annalist's blank page. – Once he who had a horse would lend it to another to mount; now, alas! It is not so." Jennings (1985), p. 176. Footnote: "When the annalist was disgusted with current events, or in uncertainty about them, he would leave a blank to be filled up afterwards. So Confucius lamented the degeneracy of his times. The latter sentence in this paragraph is not quite clear."

"I am old enough to have seen scribes who lacked refinement. Those who had horses would permit others to drive them. Nowadays, there are, I suppose, no longer such cases." D. C. Lau (1983), p. 155. Footnote: "One's carriages and horses are not things one should lightly permit others to use. To do so shows, therefore a lack of refinement."

[65] Zhou (1968).

Having gone this far, we can attempt to reconsider the entire passage which has been misunderstood for so long. To this day I have not seen a commentary that would have explained conclusively the relationship of the sentence regarding the horse (有馬者，借人乘之) to the rest of the passage.[66] From a grammatical point of view, the best reading for the passage would be the following: "I still come across officials who lack etiquette; to have someone who has a horse lend it to others, alas, there is no such thing anymore!" In other words, Confucius thought that officials had no manners anymore because they would not lend a horse to another person, an act which, presumably, would have been according to the etiquette in certain cases. Such a critical attitude is well documented with respect to Confucius.[67]

The above line of thought illustrates that Confucius was not concerned with incorrect writing habits in his time. I am inclined to think that the Han "interpretation" of the *Lunyu* passage was not a misunderstanding but a conscious use of the text of the classics to demonstrate a point, even if the original meaning of the text was different. A similar "misquoting" of the classics occurs when Xu Shen originates the principle of *liushu* 六書 from the *Zhouli* 周禮, a connection which I discuss later in this chapter.[68]

[66] Zhou Cezong believed that the officials who lent their horses to others were "uncultured" because they did not abide by the proper manners. As for the entire meaning of the passage, Zhou (1968, p. 24) interpreted it as Confucius, having once served as an official himself, still remembered the lack of manners of some of the other officials around him.

[67] A well-known case of Confucius feeling that he was not treated with proper respect is recorded in *Shiji* 47:17 where he leaves Lu because of not being given from the sacrificial meat.

[68] The "recycling" of the text of the classics to promote one's own agenda was probably an accepted methodology in the Han. My impression is that such reverse interpretations, when the text of the classics served to explain someone's argument, instead of the other way around, originally did not occur as commentaries to the classics. However, commentators in later times did draw upon these sources and, when faced with difficult passages, cited them. This is how the concept of the "historian's empty space" could have found its way into the *Lunyu* commentaries. Another related example is the reinterpretation of the concept of *zhengming* 正名 (*Lunyu* 13:3) in the bibliographic section of the *Suishu* (32:27, p. 946).

> 孔子曰：「必也正名乎！」名謂書字。「名不正則言不順，言不順則事不成。」
>
> Confucius said: "The necessary thing is to rectify the names." Names refer to written characters. "If the names are not correct, language will not be flowing, if language is not flowing, the deeds will not succeed."

The myth of truthful scribes in the past reflected a Han point of view, according to which correct and consistent (*zheng* 正) writing was a moral obligation expressing an ethical standpoint. The historian leaving an empty space would have demonstrated moral superiority over those who in later times simply substituted difficult characters with alternate ones "according to their own fancy." Inconsistent (*buzheng* 不正) writing supposedly came into existence with the world going into decline, a phenomenon that Han scholars invariably associated with "moral degeneration." Both the *Shuowen* and the *Hanshu* attributed the use of variant characters to the lack of constancy between right and wrong (*shifei wuzheng* 是非無正), i.e. moral standards.[69]

In a manner not uncommon during the Han, Xu Shen and Ban Gu projected the ideal writing habits back to the times of antiquity. They implied that the existence of a standard (*zheng*) was the original and right (*shi* 是) way of writing, whereas the lack of it was caused by the decline of moral standards in the world. The task of the Han literary elite was to re-establish these standards, an act which at least partially symbolized the restoration of the original moral order in the world.

Since the ability, or choice, to write correctly was a measure of the scribe's moral integrity, the standard way of writing had to be enforced for those in government service. Xu Shen and Ban Gu recorded the rule that if an official's characters were not correct, he was to be punished immediately.

Shuowen:

尉律：學僮十七以上始試。 諷籀書九千字， 乃得為史。 又以八體試之。 郡移太史並課。 最者， 以為尚書史。 書或不正， 輒舉劾之。 今雖有尉律，不課，小學不修，莫達其說久矣。

According to the regulations of official, students start their examinations at the age of 17. Once they memorize the 9,000 characters of the [Historian] Zhou's script, they can become historians. Their test also includes the eight [calligraphic] styles. The prefects send [the students] to the Grand Historian

---

Although the passage directly quotes from the *Lunyu*, it modifies the usage of the word *ming* 名, explaining it as "written characters." In the *Lunyu*, the term unquestionably referred to spoken language and was used in the sense of "words" or "concepts."

[69] The word *zheng* 正, which usually means "correct" could also be interpreted with reference to characters as "standard." This usage is documented in the compound word *zhengzi* 正字 or *zhengti* 正體 which not only refer to the correct way of writing a character, but also to a constant way of writing it, implying an existence of a *standard*. In the phrase *shifei wuzheng* 是非無正, the word *zheng* 正 is once again used in the sense of constancy, standard.

who tests all of them together. The best of them will become Clerks of the Imperial Secretariat.[70] If anyone writes irregularly, he is to be punished immediately.
Today, although we have the regulations of the officials, they are not enforced; students do not exercise themselves in the art of grammatology; there has not been anyone who understood its teachings for a long time.

Hanshu:

漢興， 蕭何草律， 亦著其法， 曰：「太史試學童， 能諷書九千字以上， 乃得為史。 又以六體試之， 課最者以為尚書御史史書令史。 吏民上書， 字或不正， 輒舉劾。」

At the rise of the Han, Xiao He (d. 193 BC)[71] drafted the regulations. He also composed their methods [of applications]: "The Grand Historian tests the students. Those who can memorize over 9,000 characters can become historians. Their text also includes the six [calligraphic] styles. The best of them will become Censors of the Imperial Secretariat and Calligraphy Clerks. If an official submits a document with irregular characters, he is to be immediately punished."[72]

These descriptions imply that irregularity in writing was not tolerated in government service during the Han. Naturally, regular or standard (*zheng*) writing in a centralized bureaucracy meant not only that a person always had to write the same character the same way, but that he also had to write that character the standard way and that all government employees had to write it the same way.

The other thing we can infer is that, in reality, irregular writing was an existing phenomenon even among officials. Were all officials writing standard characters, surely there would not have been a rule against those who wrote non-standard (*buzheng*) characters. Epigraphical evidence from contemporary stone inscriptions confirms that there was significant variation between character forms commonly used by the Han and post-Han literati (Fig. 2.8).

The structure of the stone inscription forms varied considerably, even though they are from the time period when the standard form was already

[70] In my translation of Han official titles, I follow Hucker (1985).

[71] Xiao He is mentioned several times in the *Hanshu* as someone who had helped Liu Bang 劉邦 (247–195 BC) to establish the empire by writing many of the Han laws and regulations.

[72] The recently published bamboo slips from Tomb #247 at Zhangjiashan 張家山 confirm that such regulations existed as early as 186 BC. The excavated slips to a large degree match the *Hanshu* and the *Shuowen*. See Li Xueqin 李學勤, "Shi shuo Zhangjiashan jian *Shilu*" 試說張家山簡《史律》, *Wenwu* (2002) 4, pp. 69–72.

Figure 2.8. Variant character forms (*biezi* 別字) on stone inscriptions of the Han and Three Kingdoms periods. From Qin Gong (1995), pp. 602, 286, 773. The first character in each row is the modern dictionary form of the character, the rest are transcriptions of stone inscription forms. I have adopted only the transcriptions of the original stone inscription forms, which are sufficient here to demonstrate the variability in character structure.

imposed. The examples confirm Xu Shen's and Ban Gu's complaints about the lack of standard and consistency in writing their own time.[73] According to Xu Shen himself, "students did not exercise themselves in the art of *xiaoxue*; there had not been anyone who understood those teachings for a long time."

The contrast between ideal and reality demonstrates the strong normative undertone in Xu Shen's work. He described Li Si's reform as the reestablishment of the regular state of writing that was in practice before Confucius's time but had been eclipsed during the Warring States period. He claimed that Li Si and his associates, by compiling the *Cangjie*, *Yuanli*, and *Boxue* volumes, successfully created a standard that restored the correct state of the script.[74] Xu Shen's basic assumption is obviously that

[73] I am referring here to the passages (translated in 2.6 above) where Xu Shen and Ban Gu quote *Lunyu* 15:25 regarding the "empty space left by a historian."

[74] It is not exactly known why were there three books with the new standard script. Since all of them were in the same small seal script, there seemed to be a redundancy in creating three works. A possible explanation for the necessity of three books is that they contained characters of different degrees of difficulty and frequency of usage. The *Cangjie pian* was probably the first in order. This is implied by the fact that even after the fall of the Qin empire, the book had been used as an elementary textbook for teaching children how to write. The title points the same direction: since Cangjie was the inventor of writing, his name was an obvious choice for the title of a beginner's textbook. The title of the second book was *Yuanli*, where *yuan* meant "to lead on to" and *li* means "to pass through" or "successive". The two characters together could have designated an intermediary level textbook. The third book, compiled by the Grand Historian himself, was the

the script was originally standard. He believed that during the Warring States period, this standard had been lost, primarily because of the moral decline of the feudal lords who disobeyed the Zhou monarch. Then, with the unification of the country, the Qin restored the original world order and re-established the lost standard of the script. From this point of view, there is a potential parallel between the First Emperor and King Xuan, as well as between Li Si and Confucius or Zuo Qiuming 左丘明.

Xu Shen's description of the sources the Qin used for the creation of the small seal script supports the connection between the First Emperor and King Xuan. In Xu's view, Li Si and his associates worked from an existing set of characters, the characters of the Historian Zhou, and reorganized the entire script. Accordingly, the Qin small seal script that emerged was a direct descendant of the script in the court of King Xuan, almost 600 years earlier. This view echoes the belief of the Qin who saw themselves as the descendants of the Zhou tradition. Having once lived in the Western Zhou homeland, the Qin adopted the Zhou script, while the eastern states where most of the philosophic schools of the Warring States thrived wrote in the increasingly diverse regional scripts.[75]

Although there might have been other standardizations during the Shang and Zhou that we are not aware of, the Qin-Han implementation of a standard script probably began with Li Si's reforms. Li Si and his associates recorded this script in the *Cangjie*, *Yuanli*, and *Boxue* compendia. Xiao He's regulations mentioned in the *Shuowen* and the *Hanshu* represented the Han official attempt to eliminate irregularity from official documents. Continuing this tradition, Xu Shen's dictionary was part of the same effort to create a standard script.

In later centuries, the government also put forward measures directed at the standardization of the script. Part of these measures was the repeated erection of the stone classics, the first of which happened in the 4th year of the Xiping reign period (AD 175), less than a century after compilation of the *Shuowen*. The "Biography of Cai Yong" 蔡邕列傳 in the *Houhanshu* 後漢書 described how Cai Yong 蔡邕 (AD 132–192) came up with the idea of a standard version of the classics.

> 邕以經籍去聖久遠， 文字多謬， 俗儒穿鑿， 疑誤後學， 熹平四年， 乃與五官中郎將堂谿典、 光祿大夫楊賜、 諫議大夫馬日磾、 議郎張馴、

Boxue pian, or the *Book of Wide Learning*. Both the author and the title of the book suggest that it would have contained the difficult and uncommon characters of the small seal script.

[75] Bodde (1938), pp. 151–153. Bodde reiterated Wang Guowei's earlier view of the Zhou-Qin cultural lineage.

> 韓說、太史令單颺等，奏求正定六經文字。靈帝許之，邕乃自書丹於碑，使工鐫刻立於太學門外。於是後儒晚學，咸取正焉。及碑始立，其觀視及摹寫者，車乘日千餘兩，填塞街陌。[76]

> Cai Yong felt that the classics, due to the length of the time since the days of the sages, had suffered many errors in graphs, and that ignorant scholars had made incorrect interpolations, thus misleading scholars in times that followed. Therefore, in 175 CE Cai Yong … [with a group of other officials] memorialized a request to make a definitive and standard edition of the graphs of the Six Classics. Emperor Ling (AD 181–234) assented. Cai Yong then wrote the texts on the stone tablets, and had workers engrave them and set them up outside the gates of the Imperial Academy. Thereby, future scholars and those who wished to study later would all have access to the correct versions. When the tablets were done and erected, those who came to look at them and copy from them numbered in the thousands of carts daily, blocking the streets and alleys of the city.[77]

The description claims that the standardization of the text of the classics was triggered by character errors in the text and, therefore, was designed to correct those. The carving of an official version of the classics into stone represented a decision of creating and enforcing a standard and was not the distillation of scholarly opinions in textual criticism. By standardizing the canon, the government picked *one* textual variant which from there on served as the authentic text, even though it was only an approximation. But according to the *Houhanshu* account, the main reason for the creation of an official version of the texts was not philological but orthographic. Accordingly, the stone classics served as touchstones for a standard orthography of characters within the classics. This was the reason why later classics used different scripts: the Wei 魏 Three-Script Stone Classics 三體石經 (AD 241) was carved in clerical, small seal, and *guwen* scripts; the Tang 唐 Five-Script Stone Classics 五體石經 (AD 623) in small seal, clerical, regular 楷, running 行, and cursive 草 scripts.

I believe that Xu Shen's compilation of the *Shuowen* represented, as the carving of the stone classics, a process of choosing character forms which were to be treated as the correct forms thereafter. In this sense, the purpose of *Shuowen* was analogous to that of the text of the stone classics. Therefore, the dictionary did not describe the contemporary state of writing but provided an idealized standard. Xu Shen believed that regularity and constancy (*zheng*) was the original state of writing and by establishing a standard he was reinstating this original condition.

[76] *Houhanshu* 50, p. 1990.
[77] Translation from Connery (1998), p. 70.

The analysis of the above sources shows that Han literati saw consistent orthography as a moral issue related to a person's integrity. This was the reason why they, departing from the inconsistencies in their own age, projected an ideal state of writing into the past prior to the moral decline of the world.

2.7 The *liushu* 六書

Xu Shen used the concept of *liushu* 六書 to identify the rules of character formation. Due to the influence of the *Shuowen*, generations of later scholars have been using these six categories to understand and to explain the structure of Chinese characters. However, as a number of traditional philologists had already observed, the *liushu* was a problematic system. Not only were some of Xu Shen's examples of the *liushu* categories etymologically incorrect, but also entire categories were meaningless from a historical point of view. Besides, many characters, depending on the point of view, could be grouped into more than one category.[78]

The discovery of oracle-bone inscriptions at the turn of the twentieth century confirmed former suspicions regarding the incorrect etymology of some characters.[79] Ever since, scholars have been quick to point out where Xu Shen's explanation of the structure of some characters did not accord with archaeological evidence. His "faulty" explanations have been commonly attributed to the lack of material available to him.

I believe, however, that Xu Shen did not intend the *liushu* to explain the etymology of characters. Instead, his primary purpose was to provide a teaching tool for the study of the nine thousand characters students had to master in order to become a historian. Therefore, he intended the *liushu* to be used as a set of mnemonic principles for the acquisition of characters. In later dynasties, the purpose and origin of the *liushu* became

[78] For a summary of the problems with the individual categories of the *liushu*, as well as some of the traditional and modern approaches to these problems, see Qiu (2000), pp. 153–163. Particularly the *zhuanzhu* category has been problematic, the meaning of which Xu Shen did not define clearly. As a result, there have been a large variety of interpretations (Qiu lists nine theories) trying to identify the exact meaning of this category.

[79] On the significance of excavated oracle-bone inscriptions with respect to our understanding of the evolution of Chinese characters, see Wang Yuxin (1999).

obscured and scholars interpreted it as an attempt to explain the etymology of characters.

A similar kind of misunderstanding on the part of traditional scholars took place with respect to the 540 components (*bushou* 部首) of the *Shuowen*. Gao Ming 高明 correctly saw the main reason behind the failure of traditional studies seeking the structural origins of characters in that these studies "confused the *Shuowen* components with the etymology of characters."[80] Gao pointed out that although the 540 components often included etymological information, not all of them could be understood in this light.

First, let us look at Xu Shen's own description of the six categories:

> 《周禮》： 八歲入小學， 保氏教國子， 先以六書。 一曰指事。 指事者， 視而可識， 察而見意，「上、下」是也。 二曰象形。 象形者， 畫成其物， 隨體詰詘，「日、月」 是也。 三曰形聲。 形聲者， 以事為名， 取譬相成，「江、河」是也。 四曰會意。 會意者， 比類合誼， 以見指撝，「武、信」 是也。 五曰轉注。 轉注者， 建類一首， 同意相受，「考、老」 是也。 六曰假借。 假借者， 本無其事， 依聲託事，「令、長」是也。
>
> The *Zhouli* says, "[children] entered school at the age of eight. Teacher Bao taught the sons of the state, starting with the *liushu* [categories.]" The first is *zhishi*. *Zhishi* characters are the ones that can be understood by looking at them, the meaning of which can be seen through observation. The characters 上 and 下 are like this. The second is *xiangxing*. *Xiangxing* characters are the ones that depict objects by reproducing their physical shape. The characters 日 and 月 are like this. The third is *xingsheng*. *Xingsheng* characters are the ones that take a [semantic] subject to indicate the word and combine it with a [phonetic] semblance. The characters 江 and 河 are like this. The fourth is *huiyi*. *Huiyi* characters are the ones that conjoin categories to present the indicated meaning. The characters 武 and 信 are like this. The fifth is *zhuanzhu*. *Zhuanzhu* characters are the ones that establish categories based on a single origin and which borrow their analogous meanings from each other. The characters 考 and 老 are like this. The sixth is *jiajie*. *Jiajie* characters are the ones that are assigned a written form, which did not exist originally, based on their pronunciation. The characters 令 and 長 are like this.

Ban Gu gave a similar description without explaining the categories and providing the two-character examples.

> 古者八歲入小學， 故周官保氏掌養國子， 教之六書， 謂象形、 象事、 象意、 象聲、 轉注、 假借， 造字之本也。
>
> In ancient times, [children] entered school at the age of eight. Thus a Zhou official by the name of Bao was in charge of the education of the sons of state. He taught them the *liushu*, which were the *xiangxing*, *xiangshi*, *xiangyi*, *xiangsheng*, *zhuanzhu*, and *jiajie*. These were the foundation of character creation.

[80] Gao Ming (1980), p. 42.

A third Han source, Zheng Zhong's 鄭眾 (5 BC–AD 83) commentary to the *Zhouli* from the second half of the first century AD, identified the six categories as *xiangxing* 象形, *huiyi* 會意, *zhuanzhu* 轉注, *chushi* 處事, *jiajie* 假借, and *xiesheng* 諧聲.[81] Thus we have three sets of definitions by three different authors (Fig. 2.9).

Xu Shen		Ban Gu		Zheng Zhong	
Zhishi	指事	*Xiangshi*	象事	*Chushi*	處事
Xiangxing	象形	*Xiangxing*	象形	*Xiangxing*	象形
Xingsheng	形聲	*Xiangsheng*	象聲	*Xiesheng*	諧聲
Huiyi	會意	*Xiangyi*	象意	*Huiyi*	會意
Zhuanzhu	轉注	*Zhuanzhu*	轉注	*Zhuanzhu*	轉注
Jiajie	假借	*Jiajie*	假借	*Jiajie*	假借

Figure 2.9. Comparison of the *liushu* categories in the three Han sources. The shaded areas in the table show those terms where Ban Gu and Zheng Zhong differed from Xu Shen. I took Xu Shen's definitions as the basis for comparison out of convenience, without implying the superiority or temporal priority of his version.

With respect to the differences between the three versions, there is only one case (*huiyi*) when two sources (Xu and Zheng) stand together against the third (Ban). In the other two cases of difference (*zhishi* and *xingsheng*), all three sources disagree with each other. However, the second character of these terms is identical in all three sources.

Despite the differences between the three accounts, they unquestionably referred to the same set of concepts described and illustrated by Xu Shen. This probably represented the general understanding of the *liushu* during the first century AD. The accounts trace the term *liushu* back to the "Diguan" 地官 chapter of the *Zhouli* which recorded what the official called Protector 保氏 taught to the children.

> 保氏掌諫王惡， 而養國子以道。 乃教之六藝： 一曰五禮， 二曰六樂， 三曰五射， 四曰五馭， 五曰六書， 六曰九數。[82]
>
> The Protector was in charge of correcting the king's wrongdoings and of educating the sons of the state in the Way. Thus he taught them the Six Arts. The

[81] Wang Yuzhi (1983).
[82] *Zhouli*, "Diguan."

> first was called the Five *Li*, the second the Six *Yue*, the third the Five *She*, the fourth the five *Yu*, the fifth the Six *Shu*, the sixth the Nine *Shu*.

In the above translation, I have not given a complete English translation of the Six Arts because the translation depends on the interpretation. One can see that the only way to treat them as logically parallel units is to understand them as activities. This is especially clear from the third and fourth arts, i.e. archery and chariot riding. *Li*, the first art, so important in Ru thought, referred to the observance of rituals and proper behavior rather than to the rituals themselves. If we look at the arts as activities, we could translate the Six Arts in the following way:

1. five forms of ritual behavior
2. six forms or styles of playing music
3. five styles of archery
4. five styles of chariot riding
5. six styles of writing
6. nine forms or styles of calculating

In sum, the term *liushu* in the *Zhouli* here did not mean the six evolutionary principles of character formation but six styles of writing, perhaps six calligraphic styles. In fact, in a passage I quoted earlier in section 2.2 of this chapter, Xu Shen used the term *liushu* once more, this second time describing the six types of calligraphic styles in use at the beginning of the Eastern Han.

> At this time, there were six scripts. The first one was the *guwen* script, which was the script on the documents found in the wall of the house of Confucius; the second the odd characters, i.e. the unusual *guwen* forms; the third the seal script, i.e. the small seal script; the fourth the attendant script, i.e. the Qin clerical script created by Cheng Mao from Xiadu under the commission of the First Emperor of Qin; the fifth the pseudo-seal script used for seal carving; the sixth the bird and insect script used to write on banners.

In this context, the meaning of the *liushu* is in accord with the original meaning of term in the *Zhouli*. In contrast with the bifurcated meaning in the *Shuowen*, Ban Gu's parallel account used the term *liuti* 六體 for the six types of calligraphy. Xu also used the term *ti* 體 to refer to calligraphic styles, but instead of the six mentioned in the *Hanshu*, he talked about eight styles. Therefore, from the point of view of terminology, we see some overlap, perhaps even confusion, in Xu Shen's account. He used the term *liushu* to indicate 1) the six principles of character formations and 2) the six calligraphic styles in the Eastern Han. On the other hand, he used the term *bati* 八體 to indicate the eight calligraphic styles that had emerged during the Qin.

When Xu used the term *liushu* as the six principles of character formation, he quoted the *Zhouli* as the source for the term. The relevant *Zhouli* passage stated that the Protector taught the Six Arts, including the *liushu*, to children. Xu Shen and Ban Gu reiterated the same belief, further adding that children had began learning at the age of eight. Upon entering school, they had immediately engaged themselves in learning the *liushu*, simultaneously with learning the characters. However, children of the age of eight who had not yet learned to read and write would have had no use for the evolutionary principles of character formation. They would have, on the other hand, benefited from a method that could have helped them to memorize the structure of newly learned characters.

Viewing the *liushu* from this perspective, the "incorrect" folk etymologies in the *Shuowen* can be easily understood as mnemonic explanations. Under the character 大, for example, Xu wrote:

> 大：天大，地大，人亦大。故大象人形。
>
> Character 大: Heaven is great, the Earth is great, man is great. Therefore, the character 大 depicts a human figure.

From the point of view of the historical evolution of the character, Xu Shen's explanation is not useful at all. The logic of his explanation is also faulty. But the phrase "Heaven is great, the Earth is great, man is great" helps to remember the connection of man with the graphical appearance of the character 大.

Another example is the character 信 given by Xu Shen to illustrate the principle of the *huiyi* category which included characters that consisted of two or more semantic components. In the character 信, the two semantic components would be 人 and 言, indicating that trust comprised of a man standing by his word. Today, however, most scholars agree that 人 component here is not a semantic but a phonetic component.[83] During the Warring States period, in the state of Chu the character was written as  , a form which consisted of the components 言 and 千. The *Shuowen* also documented this form as a Chu local feature. Thus in Chu people used  (千) component instead of the  (人) component. According to the *Shuowen*, 千 phonetically derived from 人. The close connection between the two characters could also be seen in their graphical structure: the only difference between them was that (千) had an extra horizontal line. Because 人 and 千 had no discernible semantic connection, their interchangeability supports the theory that 人 historically acted as a phonetic

[83] Qiu (2000), p. 155.

component in the character 信. Therefore Xu's explanation of the character etymologically was not correct, because 信 developed as a phonetic compound (*xingsheng*) character. However, since by his time the phonetic connection between 信 and 人 was perhaps not apparent anymore, he chose the *huiyi* principle as a mnemonic formula to aid the acquisition of the character. Indeed, even today Chinese language teachers use the same mnemonics to teach the character 信 to students.

Karlgren must have had a similar feeling about the benefits of some of the *Shuowen* categories and explanations because in the "Introduction" to his dictionary he wrote:

> Even some scholastic interpretations … I sometimes reproduce when they seem sufficiently amusing and fanciful to aid the memory – in such cases of course I mark the explanations as scholastic, thus indicating that in my opinion they must be replaced in future by interpretations that are more historically true.[84]

Even though Karlgren assumed that these "scholastic explanations" were mistakes, he acknowledged that they were useful for mnemonic purposes. He described the folk etymologies as "amusing and fanciful" and "farfetched" which are all common characteristics of an efficient mnemonic system.

In fact, Xu never claimed that the *liushu* described the etymological origins of characters. With respect to the historical process of character formation, he provided the following explanation:

> 倉頡之初作書也， 蓋依類象形， 故謂之文。 其後形聲相益， 即謂之字。[85]
>
> When Cang Jie first invented writing, he created graphic forms (*xiangxing*) according to categories; therefore these were called *wen* (patterns). After that, forms and sounds (*xingsheng*) mutually augmented each other; these were called *zi*.

In the above description, Xu mentioned two historical processes of character formation: *xiangxing* and *xingsheng*. The former created the limited number of single-component *wen*, the latter the numerous compound *zi*. He used the same binary distinction in the title of his dictionary, which could be translated as "explaining the *wen* and dissecting the *zi*."[86] Xu's

[84] Karlgren (1923), p. 3.
[85] *Shuowen*, "Postface."
[86] The reference in the title of the *Shuowen* to the dissection (*jie* 解) of the *zi* also shows that those were compound characters. The *wen*, on the other hand, could only be explained (*shuo* 說) as single unit, without the possibility of dissecting them into smaller pieces.

account does not disagree with our modern understanding of the evolution of Chinese characters embodied in the "three-principle theory."[87] According to the original theory proposed by Chen Mengjia 陳夢家, the three principles consisted of the *xiangxing*, *jiajie*, and *xingsheng* categories. Qiu Xigui has amended the *xiangxing* category to *biaoyi* 表意 (semantographs) to include not only pictographs but also characters created from semantic symbols.[88]

Xu Shen's description of the evolution of characters differs from the three-principle theory only in not including the *jiajie* principle. However, an argument can be made that the *jiajie* principle is not a principle of *character formation*. Instead, it is a principle of the evolution of *character usage*. Because when a character is being used as a phonetic loan for another character, graphically it is still the original character. One could argue that since the form of the character has not changed, there has been no character-forming principle at work. Therefore, whether the *jiajie* principle should be counted as a principle of character formation, depends on the definition of the concept of character. If this concept comprises only the graphic form of the character, then the *jiajie* principle should not be included.

On the other hand, the development of *xingsheng* characters included the process of phonetic borrowing. Every *xingsheng* character started its existence as a *jiajie* character, and only eventually did it acquire an extra component that distinguished it from the "mother" character. Therefore, once again, the *jiajie* principle could be regarded redundant. The organization of the dictionary itself also reveals the same approach on Xu Shen's part. In the *Liushu lue* 六書略, Zheng Qiao 鄭樵 (1104–1162) of the Song dynasty had already noted that in the *Shuowen*, Xu effectively used only the *xiangxing* and *xingsheng* (called *xiesheng* by Zheng Qiao) principles.[89]

Therefore the organization of the body of the *Shuowen*, as well as Xu's description of the evolution of writing in the "Postface," was not very different from our modern understanding of the history of writing. The notion that Xu Shen used the *liushu* to explain the historical evolution of characters was solely a product of his later interpreters. Xu himself identified the *xiangxing* and *xingsheng* categories as the major principles behind character evolution.

[87] Qiu (2000), p. 167.
[88] *Ibid*.
[89] Zheng Qiao (1976).

Therefore the analysis of the original description of the *liushu* leads to the following conclusions: 1) the *liushu* was a mnemonic device in the *Shuowen*; 2) Xu did not intend to provide etymological information in his dictionary; 3) the *Shuowen* was designed as a teaching tool and as such represented a prescriptive system.[90]

Most of the modern criticism of the *Shuowen* concentrates on Xu Shen's application of the concept of *liushu* in describing the evolution of character structure. Scholars have usually accepted his other major system, the division of characters into the 540 *bushou*, as self-evident. Nevertheless, understanding the educational nature of the *liushu* helps us to realize that the *bushou* and *liushu* were but two sides of a single system. The 540 *bushou* dismantled the characters into elementary components (= deconstruction), whereas the *liushu* joined them back into characters (= reconstruction). The role of both the *bushou* and the *liushu* lay in establishing a modular system aiding the learning of characters, not in documenting character etymology.

The implications of these findings are that the character analyses in the *Shuowen* only reflected the idealized script of the Han literati and should be applied to the study of Warring States writing with great caution. The dictionary should not be treated as a systematic description of the history of characters, nor should the phonetic references be taken as representatives of the pre-Qin situation.

2.8 Conclusions

In this chapter, I followed two main directions. First, I re-examined the two conventional Han sources, namely in the *Shuowen* "Postface" and the *Hanshu* "Yiwenzhi," with respect to the changes in the script commonly attributed to Li Si and the First Emperor. Secondly, I contrasted these descriptions with the epigraphical evidence. Below is the summary of my findings.

[90] I see the prescriptive nature of the *Shuowen* not in telling people how to create new characters in the future but how to write the existing characters correctly. I do not think that the *Shuowen* was created as an attempt to prescribe the rules of character construction. On the contrary, I believe that the dictionary was trying to eliminate the custom of creating new characters by freezing the current state of the script into a consistent standard.

The Han dynasty, when the *Shuowen* and the *Hanshu* were written, was a period of standardization. People attempted to discern the innate order of the universe and map onto it all existing knowledge. The result was a worldview which for the Han people represented a faithful description of the universe. Xu Shen's description of the Qin reforms reflected the same Han belief in a state of original perfection. Xu did not believe that the contemporary way of writing, filled with inconsistencies, represented the true nature of the script. Instead, he claimed that the authentic state of the script lay in the past, prior to the moral decline of the world. Hence the ability, or will, to write "correctly" or "in a standard way" (*zheng*) reflected the moral standpoint of the scribe.

The changes that the Han sources attributed to Li Si did not happen overnight but involved a gradual process that began long before the Qin unification and lasted centuries after. This is evident from the following facts:

1. The Qin small seal script was different from the *Shuowen* small seal script. (Qin bronze inscriptions vs. *Shuowen*)
2. The Qin small seal script in itself was not consistent. (Qin bronze inscriptions)
3. The Han small seal script was not uniform either. (Yuanan stele vs. *Shuowen*)
4. Decades after the Qin reforms, the script still showed evidence of inconsistency. (Mawangdui manuscripts, Yuanan stele)
5. The Qin reforms were not comprehensive and not always successful. (Qin bronze inscriptions, repeated edicts by the Second Emperor)
6. Nonstandard writing existed in the time of Xu Shen. (punishment of officials who wrote nonstandard characters)

As part of the general Han approach to reality, the *Shuowen* was a prescriptive work that claimed to be descriptive by trying to restore the "original" state of writing. The main arguments in favor of this view are:

1. Xu Shen intended his work as a teaching tool to help students prepare for the examinations. As such, just as the stone classics did, the *Shuowen* created a standard against which people could be measured.

– The *liushu* were mnemonics to aid memorization, not principles of character evolution.

– Children used the *liushu* immediately upon entering elementary school in order to facilitate the acquisition of characters.

– The characters of the *Shuowen* were primarily designed to teach students the 9,000 characters needed to become officials in government service.

2. The *Shuowen* was the result of Xu Shen's intentional effort to systematize writing and create a standard for the script.

In light of the above arguments, it is evident that the standardization of the script was a long process. It had began before the Qin unification and continued long after. Bodde saw the *Shuowen* as the culmination of the process of standardization that began with Li Si.[91] While I agree with that the *Shuowen* was a continuation of the same process of standardization that included Li Si, I also think that the standardization did not end, or succeed, with the *Shuowen* but went on long after the Han as an ongoing effort.

[91] Bodde (1938).

CHAPTER THREE

IDENTIFICATION OF CHARACTER FORMS

In order to explore usefully the variability of character forms, one must clarify the terminology. One needs to define the parameters of variation, specify what varies and in what way. The foremost task in answering these questions is to understand what constitutes a character, to pinpoint the essential features or characteristics that define a certain character and distinguish it from all other characters in the script. Once one is clear about the object of variation, one can proceed to define variation.

In this chapter, I explore two aspects of characters: components, and words. Components pertain to the internal structure of character forms, words to the meaning and pronunciation conveyed by those. I intend to demonstrate that, due to the variability of character forms in the Warring States period, graphical attributes are not sufficient for defining the identity of a character form unambiguously. Instead, departing from the idea that writing is a graphical representation of language, one needs to compare character forms on the basis of the words those forms were meant to represent in context.[92] The word provides a solid reference point independent of the graphical attributes of a certain word.

3.1 Components

Xu Shen did not invent the idea of dissecting composite characters (*wen*) into components. In the *Shuowen*, he himself quoted two instances

[92] Context can be both textual and historical. Textual, or linguistic, context would be other character forms around the one in question, helping the researcher to understand the meaning from a linguistic point of view. Historical context would be the non-linguistic circumstantial evidence such as geographic location, time period, medium. Obviously, in the case of excavated texts the textual context is often missing. One cannot talk about textual context, for example, in the case of a seal that only has a single character carved on it. Nevertheless, the historical context of the seal might be sufficient to identify what word that character stood for, since other seals from the same area in the same time period were likely to carry similar inscriptions.

from the *Zuozhuan* 左傳 of people remarking on the structure of individual characters:

> 夫文止戈為武[93]
>
> In writing, 止 and 戈 make up the character 武.
>
> 故文反正為乏[94]
>
> Therefore, in writing the reversal of the character 正 constitutes the character 乏.

In the first case, the dissection was a means to prove the point that the purpose of warfare is peace. In this context, the sentence should be understood in the following way: 'As the example of the character 武 ("warfare") demonstrates, warfare consists of stopping weapons.' In the second example, the structure of the character 乏 is analyzed in a similar manner to prove a point. In context, the sentence should be understood as follows: 'As the example of the character 乏 ("demolition, downfall")[95] demonstrates, the downfall of a state is a result of going against what is right.'

The *Zuozhuan* examples[96] show that the literate elite of the Spring and Autumn period, long before the Qin and Han standardization of the script, understood characters as consisting of modules[97] and used this information in both teaching and rhetoric. At the end of the first century AD Xu Shen created a comprehensive system based on the modules of the script when he organized his dictionary into 540 groups (*bu* 部), choosing a character that acted as a basic building block in each character in that group. Xu called these basic building blocks "group heads" (*bushou* 部首). His organization has largely defined how later generations, including ours, understood the structure of characters.

[93] *Zuozhuan*, "Xuan 12."
[94] *Ibid*, "Xuan 15."
[95] I interpret the character 乏 as "demolition, downfall" following from Wang Zuoxin (2000, pp. 86–87) who demonstrated that in this particular phrase the character did not mean "shortage, deficiency" but stood for the word written after the Han as 泛or 芝.
[96] There is yet another instance in *Zuozhuan* "Shao 1," not quoted by Xu Shen, where the character 蠱 is dissected into components (皿蟲為蠱).
[97] I am indebted to Ledderose (2000, p. 10) for the application of the term "module" with reference to Chinese characters.

3.1.1 The derivation of components from characters

According to Xu Shen's interpretation, components initially evolved from characters. He described the development of such compound characters in the "Postface" to the *Shuowen*, saying that first were the *wen* which were graphical representations (*xiangxing*), followed by the *zi* which augmented each other based on their forms and sounds (*xingsheng*).[98] This implies that, at one point, every component had been a separate *wen*.

Because components originated as characters, they always have both semantic and phonetic values. This rule is valid even when the original meaning of a component is already obscure or when the component is used in a particular context for its phonetic value. The Chu form of the character 志 in the *Guodian Laozi* manuscript was written as . This character form consisted of the combination of the elements 之 and 心, written as and . In the character, 之 served as a phonetic component, but outside of this context it had its own semantic value. Indeed, components could only have a phonetic value because at one point they had represented a word, including its meaning and pronunciation. This original meaning and pronunciation usually does not completely vanish even when the character is utilized as a semantic *or* phonetic component in a new character.[99]

In the process of evolving from *wen* to modules, components went through a process of assimilation. As a result, some originally distinct units merged and became identical. An obvious example in the modern Chinese script is the assimilation of the components 肉, 月, and 舟 into the visually identical 月 component in characters such as 朔, 朗, 胡, 胎, and 服. Although graphically this particular component looks the same in each character, etymologically these components originated as three distinct characters which eventually merged into graphically identical forms.[100] Xu Shen's organization of the 540 *bushou* was certainly a significant step along this path. Still, it is worth pointing out that the system of the 540 *bushou* in the *Shuowen* does not mean that Xu Shen dissected the entire small seal character set into 540 modules. Rather, he selected 540 components which he thought were useful for the purpose of organizing the

[98] *Shuowen*, "Postface."

[99] Both Karlgren and Boodberg (1937, p. 335) have argued that even when a component is used for its semantic value, it also carries a "weak" phonetic value, and vice versa. Accordingly, the component 之 in the above character form would, at least partially, retain its semantic value and 心 its phonetic one.

[100] Zhao Ping'an (1993) discusses in detail the process of component assimilation and the general evolution of character structure from pre-Qin to the clerical scripts.

characters.[101] In the Ming dynasty, Mei Yingzuo 梅膺祚 (fl. 1570–1615) organized the modern *kai* 楷 characters in his *Zihui* 字彙 (1615) according to 214 radicals, a division which was later adopted in works like the *Zhengzitong* 正字通 (1672) and the *Kangxi zidian* 康熙字典 (1716). As for the actual number of modules in the Chinese script, Gao Ming puts the number at 140 plus.[102] He claimes that despite the differences that these components exhibited in time and space, only the way they were written changed, not the components themselves.

Simultaneous with the assimilation, there was also a process of divergence of components, in the course of which a single component evolved into two or more distinct shapes. All of the 5 Chu character forms in Figure 3.1 contain the 人 component, written as  .

In the characters 視[103] and 先, the component 人 evolved into 儿.[104] In the character 長, where the Chu form corresponded to the *Shuowen*

[101] In his study of the *bushou* system of the *Shuowen*, Serruys (1984, p. 657) has pointed out the existence of the principle of *glyph extraction*. He explained this principle the following way:

> It is called *glyph* extraction and not *graph* extraction, because it is a purely external, artificial process consisting in isolating a portion from a certain graph, regardless what role this extracted portion has been considered to play in the original graph. Often the extracted portion has been considered by many authors to have never existed as a separate and independent graph. The same objection has been made again for several other division heads, for which no well identified corresponding graph could be found in any of the existing bronze or bone inscription materials. Yet, there is no doubt that Xu Shen considered all of them to represent real words with actual specific pronunciations.

[102] Gao Ming (1980), pp. 48–49.

[103] The Chu form of 視 did not include the 示 component. Since the omission of a component was a common phenomenon in Warring States writing, we might want to transcribe 視 as "見 (視)," indicating that although graphically the Chu form was equivalent to the modern character 見, semantically and phonetically it represented the word that came to be written as 視 in the modern script. However, the character 見 was written as  , where the bottom consisted of a kneeling figure instead of a standing one. Yet Cui Renyi (1998, p. 101) lists both of these character forms under the 見 heading. The *Shuowen* small seal form of 見 consisted of 儿 and 目, just like the Chu form of 視. Obviously, there was a strong connection between the characters 見 and 視 both semantically and graphically, as it is evidenced by the *Shuowen*, where Xu Shen glossed 見 as 視也].

Based on the way the two characters were written in the *Guodian Laozi*, Qiu Xigui (1996) suggested to reconsider our reading of the graphs  and  , both of which have been habitually transcribed as 見.

[104] Since the Chu forms are not necessarily the direct ancestors of the modern forms, at least some of these cases of component divergence might have actually been cases

Figure 3.1. Character forms from the Guodian corpus, all sharing the component 

guwen form, the component in question changed into the lower part of 辰; in 皆 into 比. Only in the case of the character 化 did the component carry its original structure over to the modern form. Keeping in mind Boodberg's theory that "a similarity of graphs reflects phonetic or semantic relationship between the words represented by the graphs,"[105] it is possible that the common component in the above character forms reflected a similarity in pronunciation at some point during the Warring States period, even though it might not be apparent from the reconstruction of their archaic pronunciations.

3.1.2 Character structure across styles; clericization

Both traditional and modern approaches to character structure reflect the view that the structure of a character remained the same regardless of the calligraphic style used to write the character.[106] One might call this belief the myth of an *ideal structure* which had greatly hindered the ability to discern the true nature of the early Chinese scripts.

The consistency of structure across different styles is self-evident in cases of characters written in calligraphic styles where stroke number is the same. For example, the stroke number and the mutual arrangement of the components in the following two character pairs are almost identical.

---

of component assimilation, where several component forms in a pre-existing script merged into the single 人 component. On the other hand, with respect to the wide range of character forms in use during the Warring States period, as this study tries to demonstrate, we cannot say which of the Warring State forms were the primary ones.

[105] Boodberg (1937), p. 371.

[106] A representative view is held by Gao Ming (1980, p. 49) who claimed that "although modern and ancient character forms differ from each other significantly, their components are basically the same."

clerical style 隸體 

kai (regular) style 楷體 

The character 豐 in both styles is written with the same number of strokes and the same components. In the character 風 there is arguably some difference in the shape and stroke number of the component 虫. Still, the components are clearly identifiable and their arrangement remains the same. However, if we look at the same two characters in *xing* (running-hand) style, the differences between the character forms become much more pronounced.

xing (running-hand) style 行體 

The character 豐 now consists of eleven strokes instead of eighteen in the clerical and *kai* styles above. Its top part has been simplified to the degree that it is has become identical to the character 曲. In fact, the whole character now looks similar to the character 豊, which was used to write the word *li* "ritual." (In modern script the character 豊 has acquired an additional 示 component and is now written as 禮.) Early forms of 豐 and 豊 seen on oracle-bone and bronze inscriptions show that the two characters were very similar, sometimes even indistinguishable.[107]

The character 風 in the *xing* style, especially the component 虫, is also simplified. We can define the stroke as a continuous swipe of the brush that does not leave the writing surface. According to this definition, the component 虫 only consists of two strokes in the *xing* style, whereas in the *kai* style it consists of six.

The differences in the structure of the characters 豐 and 風 written in different styles show that a character cannot always be defined by its strokes or components. In other words, there is no ideal, purely graphical, structure independent of usage.

The concept of an ideal structure is especially faulty when it comes to the comparison of *guwen* and clerical character forms. As a way of both

[107] Lin Yun (1983, pp. 181–186) has shown that, despite the views of some other scholars, the characters 豊 and 豐 were originally separate characters and became graphically confusingly similar during the Warring States period.

preserving the structure of *guwen* character forms and bypassing some of the difficulties of transcription, emerged the custom of *liding* 隸定, *ligu* 隸古 or *liguding* 隸古定, which could be translated into English as "clericization." The practice of *liding* emerged in the Han as part of the process of transcribing the classics using the modern script, because the structure of characters in the modern script often differed from the *guwen* structure of the same characters.

The method entailed the transcription of *guwen* characters into clerical, that is modern, script while maintaining their original *guwen* structure.[108] The technique of *liding* was just the opposite of the interpretive transcription of *guwen* characters into their modern equivalents, regardless of their original structure in the *guwen* script. Using this interpretive transcription method, the transcriber translated *guwen* characters which structurally differed from their clerical counterparts based on their meaning; using the *liding* method, he reassembled them by identifying the components in the *guwen* form, converting those components into modern forms, and then reconstructing the character from components.[109]

A well-known historical example of applying the *liding* method was Kong Anguo 孔安國 (ca. 156–ca.74 BC) transcribing the *guwen* copy of the *Shangshu* 尚書 into clerical form. Kong Anguo's "Preface" to the *Shangshu* 尚書序 used the word *liding* to describe his method of transcription:

> 至魯共王好治宮室， 壞孔子舊宅， 以廣其居， 於壁中得先人所藏古文虞夏商周之書， 及傳論語孝經， 皆科斗文字。 王又升孔子堂， 聞金石絲竹之音， 乃不壞宅， 悉以書還孔氏。 科斗書廢已久， 時人無能知者， 以所聞伏生之書， 考論文義， 定其可知者， 為隸古定； 更以竹簡寫之，增多伏生二十五篇。[110]

[108] I follow the Chinese tradition in using the term *guwen* in this context as an umbrella term that includes all ancient (Zhou and Warring States) forms of writing preceding the clerical script. Clerical script, on the other hand, refers to the modern form of writing, regardless of the actual calligraphic style. The reason why clerical script refers to modern writing in general is that transcriptions began in the Han dynasty, when the clerical script was the modern script. The division of Chinese scripts into ancient (*guwen*) and modern (*clerical*) reflects the traditional view that the development of the clerical script and the disappearance of the seal script formed a dividing line in the history of Chinese writing.

[109] Even in the *liding* method, the transcriber matched the components against their clerical counterparts, so he employed interpretive transcription on the level of components. Therefore, *liding* could be understood as component-level clericization, while interpretive transcription, as character-level clericization.

[110] *Shangshu* "Shangshu xu."

> When King Gong of Lu was renovating his palace, in order to expand his residence, he tore down the old home of Confucius. In the wall he found *guwen* documents of the times of Yu, the Xia, the Shang and the Zhou, as well as copies of the *Lunyu* and *Xiaojing*. These books were hidden by people of former times and were all written in tadpole script. The king also went up into the hall of Confucius but when he heard the sounds of metal and stone, zither and flute, he did not tear down the house but returned all the books to the Kong family.
> The tadpole script had been discontinued a long time ago and modern people were not able to recognize the characters anymore. Using the books of Fusheng, I scrutinized the meaning of the text, established those [characters] that could be recognized and transcribed them using the *liguding* method. Moreover, I wrote them out on bamboo strips, augmenting [the books of] Fusheng by twenty-five chapters.

Sima Qian also recorded Kong Anguo's transcription of the *Shangshu* in the *Shiji*:

> 孔氏有古文尚書，而安國以今文讀之。[111]
>
> The Kong family was in possession of a *guwen* copy of the *Shangshu*. [Kong] Anguo interpreted it using the modern script.

Although the original transcription was lost, a version of the Kong Anguo *Shangshu* transcribed with the *liding* method, called *Liguding Shangshu* 隸古定尚書 ("Clericized *Shangshu*"), was found at Dunhuang 敦煌. The text dates to the second year of the Tianbao 天寶 reign (AD 743).

Figure 3.2. *Liding* characters from the *Liguding Shangshu* from *Dunhuang Baocang* 敦煌寶藏, v. 140, pp. 29–37.

Examples of *liding* character forms from the *Liguding Shangshu* manuscript (Fig. 3.2) illustrate that although on the level of components the *liding* character forms in the top row are written in the modern script, on

[111] *Shiji* 121:61, p. 3125.

the level of entire characters they do not belong to modern script. The structure of these *liding* forms closely resembles the structure of some pre-Qin forms (Fig. 3.3).

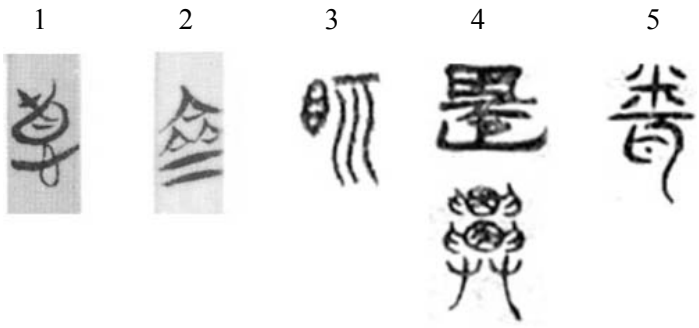


Figure 3.3. Examples of pre-Qin character forms resembling the *liding* character forms in the *Liguding Shangshu*. Character forms #1 and #2 are from the *Guodian Laozi*. Character form #3 is the *guwen* form of the character 視 in the *Shuowen*. Character forms #4A and #4B are from the Juyan 居延 corpus and the Tri-Script Stone Classics 三體石經, respectively. Each of these two forms partially resembles the *liding* form of 遷: the Juyan form on the top includes the 令 component, while the Stone Classic form includes the top of the 興 component. Character form #5 is also from the Tri-Script Stone Classics. These examples of pre-Qin character forms represent some of the possible *guwen* forms that the Han scholars could have had available to them when producing the *liding* forms of the *Liguding Shangshu* manuscript.

Even today, when transcribing epigraphic character forms into modern script, researchers often use the *liding* method because this way they can preserve the structure of the original *guwen* forms while providing a convenient visual format accessible to those who are not experts in reading pre-Qin scripts.

A major problem with the *liding* technique is that *guwen* components cannot always be unequivocally matched with clerical forms. Sometimes a manuscript form can be transcribed in several equally valid ways. In the *Guodian Laozi* manuscript, for example, the character 仁 was written as  , i.e. a combination of 身 and 心. The structure of this form could be transcribed as  . But 身 in the Chu script consisted of the components  ( , which generally represented 以 in context) and  (人). Thus 身, the top element in the Guodian form, could also be transcribed as a combination of  and 人, in which case the *liding* transcription of would consist of 人, , and 心, resulting in the form  .

The difference between the two transcriptions is that in the first case (), I divided the Chu character form into two components, whereas in the second case (), I divided it into three components. Obviously, a person only familiar with the modern script would not be able to see that the two *liding* forms (and) are transcriptions of the same Chu character form. No matter how I choose to transcribe the original *guwen* form, I will end up obscuring some of the inherent relationships between the components in the character, as well as the connection of the character with other characters that share those components.

In the case of the Chu form of 仁, transcribing the character as a combination of 身 and 心 conceals the connection of 仁 with 人. In the *Shuowen*, beside the small seal form (), Xu Shen also listed two *guwen* forms for 仁, one of which was written as  , and commented that this form consisted of 千 and 心 . At first glance there seems to be no connection between the 千 on the top of the *Shuowen* form and the 身 component in the *Guodian Laozi* form. However, in the Baoshan corpus we see that 身 was also written as  , which was a combination of  and 千. The upper half of (仁 in the *Guodian Laozi*) and (身 in the Baoshan corpus) only differed in a horizontal line. The reason for this is that 千 in the Chu script derived from 人, by the addition of an extra horizontal stroke.[112] Therefore, the original Chu forms reveal that the characters 人 and 千 were closely related and that, at least in the state of Chu, 仁 could be written either as 心 + 人 + or 心 + 千 +, in both cases amounting to 心 + 身.

The *Shuowen guwen* form, on the other hand, consisted of 心 + 千, which differed from the Chu forms only in leaving out the component. The *liding* transcriptions of the 心 + 千 (*Shuowen guwen*) and the 心 + 身 (*Guodian Laozi*) structures of 仁 suggest that the top parts of the two forms (千 and 身) acted as the phonetic component. The relationship between the Chu forms of 人 and 千 (人 + 十), and 身 (千 + or 人 +) reveals that, aside from their phonetic connection, the two forms were also related graphically, as well as semantically.[113] Figure 3.4 shows

[112] This corresponds to the *Shuowen* explanation, which claimed that 千 derived from 人 and 十.

[113] I see the semantic connection in the link between 人 and 身. Also, the other *guwen* form in the *Shuowen* consists of 尸 and 二, in which 尸 can also be linked to 人 and 身. In the second *guwen* form of the *Shuowen*, the  (尸) component is also graphically related to (人), the two sometimes being indistinguishable from each other in the Chu script. Therefore, 人, 身, and 尸 were sometimes interchangeable with each other. Another two characters that can be linked to this group are 骨 and its component 肉. The character 體, for example, is written on

the structural growth of the Chu form of the character 仁, which can be also represented as follows:

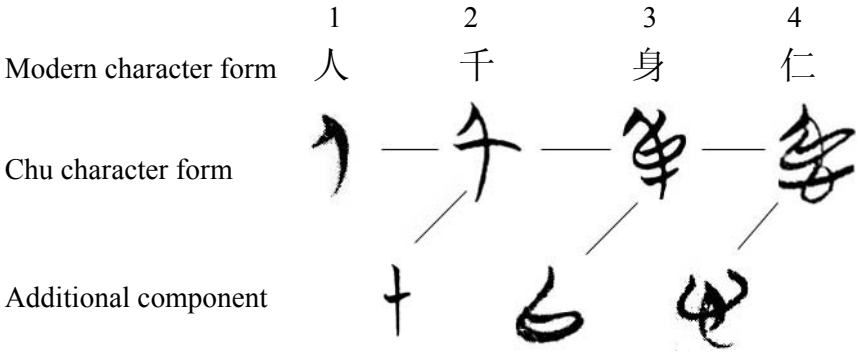


Figure 3.4. Structural evolution of the character 仁 in the Chu script.
The composition gradually increases as new additional components are added to it.

1. 人
2. 人 + 十 = 千
3. 千 +  = 身
4. 身 + 心 = 仁

The character 人 in stage #1 represents the basic component in all other character forms in the above chart. In stage #2, 人 receives an additional 十 component to form the character 千. The vertical stroke of 十 is assimilated into the vertical stroke of 人, thus graphically speaking the addition of the component 十 is equivalent to the addition of a single horizontal stroke. The phonetic value of the character 千 is represented by the component 人. In stage #3, the character 千 receives an additional component, thus forming the character 身. The pronunciation of 身 is represented by the component 千. In stage #4, the character 身 receives an additional 心 component, thus forming the character 仁. The pronunciation of the Chu form of the character 仁 is represented by the component 身.

The reconstruction of the structure of the Chu form of the character 仁 demonstrates how the graphical form of a pre-Qin character could include information no longer present in the modern script. If I were to transcribe  as either  or  , I would lose the structural information inherent in

---

the *Zhongshan wang fangding* 中山王方鼎 inscription with 身 instead of 骨 (as it is in the character's modern form). In the Shuihudi 睡虎地 and Mawangdui corpora, the component 骨 in 體 is written as 肉.

the original character and obscure the connection between the components and the entire character.

The Chu forms of the characters 信 (written as ) and 年 (written as ), for example, shared the same phonetic component. Both of these character forms derived from 千. In the state of Chu, the character 信 was a composite of 言 + 千 and the character 年 was a composite of 禾 + 千. The common 千 component played a phonetic role, and thus both characters, along with 身 and 仁 could be classified as members of the 千 *xingsheng* family.[114] However, while the connection was relatively obvious from the Chu forms of the characters, their modern forms no longer carried this information.

Another example of the transcription obscuring the relationship between characters is that of the characters 難 and 堇 (Fig. 3.5). The Guodian forms of the characters 堇, 菫, and 難 shared the same  component, which is equivalent to the modern character 黃. In the *Shuowen*, Xu Shen stated that in 難 the phonetic component was 堇, while in 堇 it was 黃. Thus all three characters, 黃, 堇, 難, as well as some other characters that shared the same component (e.g. 勤 and 僅), were phonetically related to each other, a relationship that is not apparent from neither their *liding* nor their modern forms.

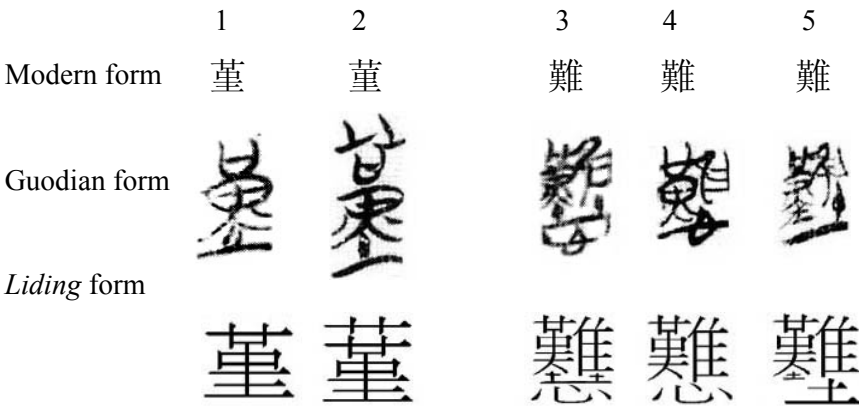


Figure 3.5. Variant forms of the characters 堇, 菫, and 難 from the *Guodian Laozi.*

[114] I would tentatively also add characters such as 聖 and 成, written in Chu as  and  , to the list of characters derived from 千. The *Shuowen* explained 壬, written as  in Chu, which was probably the phonetic component in both 聖 and 成 as a combination of 人 and 土. In the Chu script, however, 壬 was a combination of 千 and 一, perhaps a merger of 千 and 土.

The above examples illustrate that the *liding* method of clericizing *guwen* character forms into modern script, even if preserving their original structure, often conceals some of the information inherent in the original guwen form. The transcription could sever the connection of the character with other characters, as well as create new, false connections.

3.2 Words

A basic prerequisite of any writing system is the presence of a convention, that is, an agreement between members of a community regarding the relationship between a word and its graphical representation. This convention represents the rules according to which the writer encodes and the reader reliably decodes information. My objective here is to illustrate that in the case of Warring States scripts the convention did not impose a strict structural consistency of the graphical forms but defined the scope within which graphical forms could vary.

3.2.1 The identification of a character

I have already referred to the fact that although it is the components and their arrangement that constitute the structure of a character, in the case of structurally dissimilar variants, without assuming a standard form, the components are insufficient to identify the character forms. This problem is insignificant while we are dealing with the relatively standardized forms found in transmitted texts. But in both medieval and early Chinese manuscripts, including seal, pottery, and bronze inscriptions, character forms are characterized by a high degree of structural flexibility, making it difficult to identify them unambiguously. In some cases this flexibility can even prevent us from recognizing characters that are otherwise familiar to us.

With the absence of a standard form that could serve as an abstract form of a particular character, the solution to the problem of identification lies in determining the relationship of a character form with the word it stands for. Because written characters are graphical representations of words of a language, they are meaningful only in reference to the word they have been chosen to represent. Therefore, the word always determines the identity of the character, regardless of its graphical appearance.

This approach is essential for the study of pre-Qin character forms because of the relatively loose correlation between words and their graphical representations.

Naturally, orthographic flexibility also had its limits beyond which the meaning of a character became compromised. A vivid example of such a case is described in the *Lüshi chunqiu* 呂氏春秋 (22:6):

> 子夏之晉，過衛，有讀史記者曰：『晉師，三豕涉河。』子夏曰：『非也，是己亥也。夫「己」與相近，「豕」與「亥」相似。』至於晉而問之，則曰『晉師己亥涉河』也。
>
> When Zixia was going to Jin, he passed through Wey, where he read some historical records, one of which said, "The three pigs of the Jin army crossed the Yellow River." Zixia remarked, "That is wrong. The text should say 'the cyclical date *ji* and *san*, 'three,' are close and the characters for the cyclical date *hai* and *shi*, 'pig,' resemble each other." When he arrived in Jin, he asked about it, and the inscription did say: "The Jin army crossed the Yellow River on the cyclical day *jihai*."[115]

3.2.2 Relationship between a word and its graphical representation

In order to avoid the confusion arising from the variability of character structure, I arrived at the solution that the identity of a character is based on the word it represents in writing. Consequently, I decided to view the character forms of the pre-Qin script as graphical representations of *words* and not as variant forms of *characters*. Below I examine the relationship between word and character in more detail.

When a character form is associated with a single word there is little confusion. In the modern script the character 天 stands for the word *tian* "heaven"[116] and this graph always represents the same word. However, such an idealized view of a unique connection between a word and its graphical representation cannot be maintained with respect to the script of the Warring States period. It was a common phenomenon in early manuscripts that two or more distinct words could be written in a structurally identical way. For example, the Chu forms of the characters 夫, 天, and 而 were sometimes indistinguishable from each other. Although there was a strong graphical similarity in the these character forms, possibly reflecting

[115] Slightly modified translation from Knoblock (2000), p. 584.

[116] From here on, I mark words by providing an italicized *pinyin* pronunciation and English meaning in quotation marks. This convention is intended to signify that I am referring the word and not the character.

inherent phonetic or semantic links between them, the words they stood for were certainly distinct.

The Houma covenant texts often used the same  form to represent the words *fu* "abdomen" (written in the modern script as 腹) and *fu* "to restore" (written in the modern script as 復). Without context one would not be able to tell what word a given example of this form was to represent. At the same time, the structure of the form (彳 + 止 + 复) is different from the modern forms of the characters 腹 and 復. The context, however, makes it clear that in certain cases it referred to the word *fu* "abdomen" and in other cases to the word *fu* "to restore:"

#1. 敢不剖其腹心以事其宗….[117]

Should he dare not to split open his abdomen and heart in serving his ancestral temple …

#2. 而敢有志復趙尼及其子孫…

Should he dare to have the intention to restore Zhao Ni and his descendants…

In sentence #1 the form appeared in the phrase 腹心 ("abdomen and heart") in the sense of "body and mind." In sentence #2, it was used in the sense of restoring to power a clan. Therefore, despite the similarity in the actual structure of the character forms, the context disambiguates the meaning of both sentences.

The above example serves to demonstrate that in Warring States scripts sometimes the same graphical form was used to write distinct words. Another common phenomenon was that the same word was written in distinct graphical forms. In these cases the pronunciation and meaning (which defined the word) of two character forms were identical but their visual appearance was different. Sometimes the difference was small, involving one or more components (structural variation), sometimes large, involving the entire character (phonetic borrowing).

An example of structural variation is the forms which represented the word *cong* "to follow, adhere to" in the Houma covenant texts (Fig. 3.6). Both forms #1 and #2 appear in the same sentence on different covenant tablets:

敢不剖其腹心以事其宗，而敢不盡從嘉之盟、定宮平寺之命，…[118]

[117] *Houma mengshu*, p. 35. I will look at the Houma covenant text forms of these two characters, as well as their context, in detail in Chapter Six.
[118] *Houma mengshu*, p. 35.

Should he dare not to split open his abdomen and heart in serving his ancestral temple, should he dare not to fully adhere to Your covenant and the mandate granted at Dinggong and Pingsi temples…

	1	2
Houma covenant text form		
Word	*cong* "to follow, adhere to"	*cong* "to follow, adhere to"
Modern form used for the word	從	從

Figure 3.6. Two variant forms of the character 從 from the Houma covenant texts.

The identical context leaves no doubt that forms #1 and #2 stood for the same word, despite their dissimilar structure. The word *cong* "to follow, adhere to" in the modern script is represented by the character 從, thus it matches the structure of form #1. Form #2, however, is different in that it lacks the 彳 component found on the left side of form #1.

In the case of phonetic borrowing, the difference between two character forms involved the entire character. The English term "phonetic borrowing" is the translation of two traditional Chinese linguistic terms which are still in use today: *jiajie* 假借[119] and *tongjia* 通假. As Qiu Xigui has pointed out, the distinction between the two terms is often blurred but, generally speaking, *jiajie* refers to cases when a word which originally had no graphic form is written with a character used for a homophonous or nearly homophonous word, whereas *tongjia* refers to cases when a word with an existing graphic form is written with a character used for a homophonous or nearly homophonous word.[120] Since in modern works

[119] The term *jiajie* was first used by Xu Shen and Ban Gu as one of the six categories of character composition (*liushu* 六書). Xu described phonetic loan characters in the "Postface" to the *Shuowen* as characters produced by assigning an existing character, which already stood for a word, to another word with similar pronunciation but without written form.

[120] Qiu (2000, p. 167). For a more systematic analysis of the definition of phonetic borrowing, including the relationship between *jiajie* and *tongjia* categories, see Li Cunzhi (1995), pp. 19–40.

the terms are often interchangeable, or *tongjia* is incorporated within the *jiajie* category, I will use the English term "phonetic borrowing" for both phenomena.

The difference between a substitution of a component and the entire character was not always obvious. It is commonly assumed, for example, that the character 右 in certain contexts acted as a phonetic loan for the character 佑. The difference between the characters, however, is only apparent retrospectively because in the modern script the word *you* "to protect" is invariably written as 佑, whereas the character 右 always represents the word *you* "right side." The two words are unquestionably cognates and the similarity between their graphical appearances signifies this connection. But in the Warring States period the usual way of writing the word *you* "to protect" was 右. Therefore, in the scripts of the Warring States period, 右 was not a phonetic loan for the character 佑 but the customary way of writing both the words *you* "right side" and *you* "to protect."

In the case of early manuscripts, phonetic borrowing refers to a substitution of one character for another based on similarity in their contemporary pronunciation. In the *Guodian Laozi*, for example, a number of character forms were used to represent the word *shou* "to guard, hold" written in the modern script with the character 守 (Fig. 3.7).

	1	2	3	4
Guodian form				
Word	*shoux* "to guard, hold"	*shou* "to guard, hold"	*shou* "to guard, hold"	*shou* "to guard, hold"
Modern form used for the word	守	守	守	守

Figure 3.7. Variant forms of the character 守 from the *Guodian Laozi.*

Forms #1, #2, and #3 are structurally identical to the modern form of the character 獸 which is used to write the word *shou* "animal." Form #4, however, is structurally identical to the modern form of the character 守. The word *shou* "animal" also occurs once in the manuscript and is written

as  , i.e. it matches the structure of the modern character 獸. This confirms that the form 獸 was also used in Chu in the sense we use it today. On the other hand, form #4 is structurally equivalent to the modern character 守, revealing that, at least in certain cases, it was used to represent the word *shou* "to guard, hold." Therefore, I might conclude, based on the context, that the first three forms, which primarily stood for the word *shou* "animal," were used as a phonetic loan for the character 守.[121]

However, it is important to realize that the concept of phonetic borrowing assumes a pre-existing association between a character and a word. In the above case, based on my knowledge of the modern script, I might have presumed that 守 was a more appropriate graphical rendition for the word *shou* "to guard, hold" than 獸, despite the fact that the word *shou* "to guard, hold" was written in the manuscript three times as 獸 and only once as 守. Based on the situation in the *Guodian Laozi* alone, I should have regarded 獸 as the graphic representation of the word *shou* "to guard, hold" and treat 守 as an occasional variant. Obviously, I tend to favor 守 over 獸 because of my knowledge of, and familiarity with, the modern script. In other words, I know "how the story ends." When I look at the connection between written character forms and the words they represent, I am inclined to perceive them from the point of view of the post-Qin modern script.

One could argue against seeing 獸 as the primary graphic representation of the word *shou* "to guard, hold," saying that the structure of 獸 represented the phonetic and semantic aspects of the word *shou* "animal." Once again, this claim would only mirror the modern understanding of the structure of the character. Xu Shen defined 獸 as a person on garrison duty (守備者) and 守 as an official (守官也).[122] This shows that the two characters, as well as the words they represented, were closely connected

[121] Since the number of characters on which I base these conclusions is rather limited, the conclusions may not hold true for the entire written corpus to which the *Guodian Laozi* belonged. Still, they reflect the writing habits of the scribe who wrote the manuscript, and we can assume that his "orthography" fell within the accepted boundaries of his community. Accordingly, anyone who was part of the community the members of which communicated using this script, would have been able to read the manuscript without difficulty.

[122] Most commentators and interpreters of the *Shuowen* pointed out that the component 犬 in the character 獸 was there because of the connection of the character with hunting. While this is a plausible observation, the component 犬 could easily be understood within the framework of an officer on garrison duty. This is supported by the *Shuowen*'s explanation of the double 犬 component in the character 獄: "The two dogs are used for guarding" (二犬所以守也).

with each other. The character 獸 certainly had the sense of protecting and holding, and was an entirely appropriate way to write the word *shou* "to guard, hold." In fact, on the basis of the *Shuowen* definitions alone it seems that 獸 was perhaps a better choice to write the word *shou* "to guard, hold" than 守.[123] The usage of the character forms in the *Guodian Laozi* confirms this hypothesis.

As seen above, the form 獸 occurs four times in the manuscript. Three times it stands for the word *shou* "to guard, hold" and once for the word *shou* "animal." The form 守 on the other hand, occurs only once and stands for the word *shou* "to guard, hold." Therefore, the *Guodian Laozi* manuscript and the *Shuowen* agree that the form 獸 was not only an appropriate way, but also the usual way, of writing the word *shou* "to guard, hold." Accordingly, one cannot really consider the usage of the form 獸 here a phonetic loan, since it was the customary graphical representation of the word in Chu. Indeed, one could equally well argue that, for the writer of the *Guodian Laozi* manuscript, 守 was a loan for 獸. In other words, 獸 and 守 were both valid ways of writing the same word.

To introduce another example, the word *zhi* "to know" occurs twenty times in the *Guodian Laozi*.[124] In all of these twenty instances, the word is written with the two different forms (Fig. 3.8).

	1	2
Guodian form		
Word	*zhi* "to know"	
Modern form used for the word	知	

Figure 3.8. Variant forms of the character 知 from the *Guodian Laozi.*

The difference between the two forms is that #1 lacks the 口 component under the 大 component on the top left side of the character. Otherwise,

[123] At the same time, the meaning of wild animals was not foreign to the character 守, as evidenced in the character 狩 which represented the word *shou* "to hunt."

[124] This number does not include an additional two cases which are indicated by a *chongwen* 重文 repetition marker.

both forms consist of the components 大, 于, and 日. Both of these forms include the component 日, which is why modern scholars transcribe this form as 智 acting as a phonetic loan for 知. However, this perception disregards two significant facts. Firstly, the two above forms are structurally not identical to 智; second, and more importantly, together they represented the only graphical representations of the word *zhi* "to know" in the *Guodian Laozi*.

To cite yet another example, in the Houma covenant texts the characters 腹 and 復, beside the two forms mentioned above, were written in a variety of different ways. Looking at a certain form separately, one could make a case for phonetic borrowing. The form 復, for example, is one of the many variant forms used to record the word *fu* "abdomen," which is written in the modern script as 腹. Since this form is structurally identical to the modern form of the character 復, one could say that in this context 復 was a phonetic loan for the character 腹. Yet the structure of other variant forms used to write the word *fu* "to restore" (written in the modern script with the character 復) in the Houma corpus almost never matched the structure of the modern character 復. Instead, in most cases the word *fu* "to restore" (written in the modern script as 復) was written with an additional 止 component at the bottom of the character (Fig. 3.9).[125]

	1	2	3	4	5
Houma form					
Word	*fu* "to restore"				
Modern form used for the word	復				

Figure 3.9. Variant forms of the character 復 from the *Guodian Laozi*.

All of these five characters forms share the basic structure of 彳 + 止 + 复, which does not exist in the modern script. Since the character 復 in the Houma corpus was not identical to its modern form, the fact that the form 復 matches the structure of the modern character 復 is merely a coincidence. Thus from the point of view of the scribe who wrote the text, the form 復 was not a phonetic loan for another character that otherwise

[125] This structure (彳 + 止 + 复) does not exist in the modern script.

customarily represented the word *fu* "abdomen." Instead, it was only one of the many possibilities the scribe chose to record the word. The variability of the graphic representation of the word *fu* "abdomen" in the Houma corpus demonstrates that the form 復 was only one of the many possible ways of writing the same word (Fig. 3.10).

	1	2	3	4	5	6	7	8
Houma form								
Word	*fu* "abdomen"							
Modern form used for the word	腹							

Figure 3.10. Variant forms of the character 腹 from the *Guodian Laozi*.

A common part in all of these forms is the component 复 which, according to Xu Shen, represented the phonetic in the character 腹. Other than this unchanging core, the rest of the structure seems completely flexible. The eight above forms can be dissected the following way:

#1 = 复 + 彳 + 月[126]
#2 = 复 + 彳 + 止
#3 = 复 + 彳 + 口
#4 = 复 + 月 + 止
#5 = 复 + 月
#6 = 复 + 彳 + 月 + 夂[127]
#7 = 复 + 月 + 口
#8 = 复 + 月 + 止[128]

All of the above eight character forms share the component 复 which represented the phonetic value of the character forms. The fluidity of other components shows that the form 復 (used by the Jin 晉 scribe to write the

[126] Even in these early character forms, the component 肉 looks identical to 月. This is why in the transcription I use 月 instead of 肉.
[127] The component 夂 that is on the top of the character is identical to the same component at the bottom of 复, only the position of it is different.
[128] This form is identical to #4 but the component 复 is missing the 日 from the middle.

word *fu* “abdomen” but incidentally matching the structure of the modern character 復) was not necessarily closer to the contemporary way of writing the word *fu* “to restore” than any of the above eight variants. All of these forms were simply different graphic “spellings” to write the word *fu* “abdomen,” used with various degrees of frequency. Therefore, the form  used for the word *fu* “abdomen” should not be regarded a phonetic loan character.

Going back to the Chu character form  , which is structurally equivalent to the modern character 獸, to classify it as a phonetic loan for the character 守 in the *Guodian Laozi* manuscript reflects a modern assumption about how certain words were, or rather should have been, written in Warring States Chu. This assumption obscures the nature of the relationship between word and character. Warring States manuscripts and inscriptions reveal that the principle of phonetic borrowing is meaningful only from the perspective of a standard, be it the modern script or the small seal script of the *Shuowen*.

Having realized this, one can see that the connection between variant forms categorized from a modern perspective as phonetic loan characters was often not solely phonetic. Although in the text of the *Guodian Laozi*, the character forms and  (transcribed into the modern script as 獸 and 守) appear to be phonetic loans of each other, they also stand for cognate words (*shou* “animal; originally: person on garrison duty” and *shou* “to guard, hold; originally: official”). Therefore, in trying to decipher unknown character forms, one should be careful with matching characters exclusively on the basis of phonetic similarity and take into consideration their potential semantic and etymological relationship.[129]

The variability of the graphical representation of words also explains why nearly a quarter of the Chu character forms does not appear in the *Shuowen*.[130] Most of the Chu character forms that cannot be found in the *Shuowen* were not “descendantless graphs,” as Barnard called those pre-

[129] Karlgren has already advanced the theory of cognate word families that share a common phonetic component. This phonetic component is graphically identical in the characters that represent cognate words. I am here going a step further in saying that two characters, which from a modern point of view appear to be loans for each other in a text, often represent cognate words, even if they do not share a graphically identical component.

[130] This number is based on the collection of Chu character forms in *So kanpaku moji hen* 楚簡帛文字編. (Katsu 1992, p. 4.) The *Chuxi jianbo wenzi bian* 楚系簡帛文字編 (Teng Rensheng 1995) has over 800 character forms which do not appear in any dictionary, including the traditional collections of archaic character forms.

Qin character forms that did not match the structure of any modern character form,[131] but unrecognized variant forms of known characters. These variant forms stood for words which are included in the *Shuowen* under other character forms.

The above discussion of phonetic borrowing is an example of the application and benefits of an approach that bases the analysis of Warring States character forms on the word those forms were meant to represent. At the heart of this approach lies the argument that writing is a graphical representation of language and that characters are graphical representations of words. Moreover, from the point of view of the contemporary users of a script, a written text was linguistically unambiguous and the audience could correctly read the words recorded by a scribe.

[131] Barnard (1979).

CHAPTER FOUR

THE STANDARD

The concept of variation presupposes the existence of at least two entities which are contrasted with each other. A character form can only be a variant when compared to another one that differs from it. Starting from the Han dynasty, scholars have viewed character variants as deviations from a standard form, or *zhengzi*, of a given character. In this chapter, I trace the origins of the concept of the *zhengzi*. To illustrate that orthographic variability also existed in other cultures using non-Chinese scripts, I touch upon the spelling of written English in three different time periods: late medieval, late Renaissance, and modern.

Before I begin discussing the concept of a standard, I would like to emphasize that I do not claim that a standard did not exist at all in the pre-Qin scripts. This would be equivalent to claiming that the script had no rules whatsoever. On the contrary, because writing is by definition a system based on an association of the elements of speech with graphical signs, convention lies at the core of any writing system. In the case of the pre-Qin scripts, even those cases of variation which from our point of view appear to be random coincidences might fall within the constraints of the rules and principles that the Warring States scribes observed.

These rules and principles, however, only partially overlapped with those that govern the use of the modern script. The consistent association of a word with a written form was not as strict as it is in the modern script. Not only individual components but the entire character form could vary, as long as this variation met certain requirements. The flexibility of the graphic form reflected a higher orthographic tolerance on the part of the writer and reader than we are accustomed to in the modern script. Writers and readers of Warring States China were used to these kinds of variations and were not surprised by them. Indeed, they probably did not even notice the differences between those forms that the modern researcher sees today as variants. Within a community that shared a script, the written character form was not an end to itself and a word could be associated with a graphic form in more than one way.

4.1 The concept of standard in the Chinese script

In his foreword to the *Yitizi zidian* 異體字字典 published by the Ministry of Education of Taiwan in 2001, Li Xian 李鎈, a co-editor of the dictionary, defines a character variant 異體字 the following way:

> 所謂「異體字」， 是指在一個正字標準下， 文獻上與此正字同音義而形體有異的字。
>
> The so-called "variant character" is a character which, in comparison with a standard form, appears in textual sources with the same pronunciation and meaning as the standard character but differs from that in its graphical form.[132]

Li Xian's definition is based on the concept of a standard form that acts as a touchstone against which all other character forms are measured. Thus the graphical appearance of a character could vary and that it is the pronunciation and meaning that would link it with its *zhengzi* equivalent. Li Xian assumes that, beside having the same pronunciation and meaning as the standard form, the variant is the *same character* as the standard form. Li, as evidenced by the arrangement of variant forms in the *Yitizi zidian*, does not consider the so-called *jiajie* characters to be variant forms because, despite representing the same sound and meaning, they are in fact different characters.

The structural variability of character forms in excavated texts implies that in Warring States times the concept of a standard form did not exist. The modern researcher cannot measure Warring States character forms against a correct form because there is no way of telling which of the many forms was the correct one. A standard is a result of a prescriptive effort of an authority, i.e. reform or standardization, aiming to eliminate a situation in which standard did not exist.[133] According to Wang Li, in the course of Chinese history, those character forms which were selected by the government of a given period as correct forms were the standard forms and all other forms were *suzi*, or popular forms.[134] Thus Wang was of the

[132] http://140.111.1.40/bian/shiu.htm#shiu3

[133] This is true of any standard and its enforcement. As for variant character forms, Zhang Yongquan 張涌泉 (1998, p. 3.) pointed out that the standard (*zheng* 正), and popular (*su* 俗), character forms are meaningful only in contrast with each other. Without a popular form, there is no standard form and vice versa. If all character forms were identical, there would be no reason for establishing a standard.

[134] Wang Li (1984, p. 498) also explains that certain character forms could be regarded by one administration as *zhengzi* and by others in other time periods as *suzi*.

opinion that the standard was always defined by the prescriptive authority of the government.

The concept of a standard form goes back to the Han dynasty. In the *Shiji*, Sima Qian mentioned an incident that had happened to Shi Jian 石建, the eldest son of the Lord of Ten Thousand Bushels 萬石君.[135]

> 建為朗中令， 書奏事， 事下， 建讀之， 曰：「誤書！『馬』者與尾當五， 今乃四，不足一，上譴死矣！」甚惶恐。其為謹慎，雖他皆如是。
>
> When Jian held the post of Chamberlain for Attendants, he wrote a proposal. Once the proposal was approved, Jian read it and exclaimed: "An error in writing! The character 馬 should have five [strokes] in the [horse's] tail. Here I only have four, missing one. I will be punished and put to death!" He was terrified. In other matters he was just as careful and alert.[136]

This incident implies that as early as the 2nd century BC, officials were expected to write with a high degree of orthographic uniformity and even the omission of a single dot or line in a character constituted a major mistake. It is possible, however, because the *Shiji* was completed in 91 BC, the story may reflect the habits of Sima Qian's rather than Shi Jian's time.

A few decades after the compilation of the *Shiji*, Liu Xiang 劉向 (79–8 BC), the bibliographer of the Western Han dynasty, wrote in the "Appendix" of the *Zhanguoce* 戰國策附錄.[137]

> 所校中戰國策書， 中書餘卷， 錯亂相糅舛 ……. 本字多誤， 脫為半字，以趙為肖，以齊為立. 如此字者多。
>
> Among the Warring States historical documents stored in the imperial archives, the documents I have edited had many volumes that were erroneous and disorganized. … The original characters [in these documents] had many mistakes, omitting half characters, such as writing 肖 instead of 趙 or 立 instead of 齊. Characters like this were in abundance.[138]

In this excerpt Liu Xiang stated that he had to work with documents in which the character forms were diverse, often extremely simplified. Recently excavated texts confirm that the characters 趙 and 齊 were sometimes indeed written as 肖 and 立 during the Warring States pe-

[135] The Lord of the Ten Thousand Bushels was a name given by Emperor Jing 景帝 (188–141 BC) to Shi Fen 石奮, who himself and his four sons (five people in total) were all officials of Two Thousand Bushels 二千石. See *Hanshu*, "Shi Fen" 石奮, p. 2194.

[136] *Shiji*, "Wanshi Zhang shu liezhuan" 萬石張叔列傳, p. 2766.

[137] Liu Xiang compiled the *Zhanguoce* between 26 and 8 BC, while he was heading the commission that examined the documents in the imperial archives. (Tsien 1993, p. 5.)

[138] *Zhanguoce*, "Fulu" 附錄, p. 1195.

riod.[139] Liu Xiang made note of this phenomenon because for him, as well as for the scholars of his time who were the implied audience of his "Appendix," this constituted a problem.

By the time of the Eastern Han, the concept of standard had already been formulated. In the "Yiwenzhi" chapter of the *Hanshu*, Ban Gu used the word *zhengzi* in the following context:

> 漢(書)[興][140]， 閭里書師合《蒼頡》、《爰歷》、《博學》三篇， 斷六十字以為一章， 凡五十五章， 并為《蒼頡篇》。 武帝時司馬相如作《凡將篇》， 無復字。 元帝時黃門令史游作 《急就篇》， 成帝時李長作《元尚篇》， 皆《蒼頡》中正字也。《凡將》 則頗有出矣。 至元始中， 徵天下通小學者以百， 各令記字於庭中。 揚雄取其有用者以作《訓纂篇》， 順續《蒼頡》，又易《蒼頡》中重復之字，凡八十九章。臣復續揚雄作十章，凡一百二章，無復字， 六藝群書所載略備矣。《蒼頡》多古字， 俗師失其讀，宣帝時徵齊人能正讀者群，張敞從受之，傳至外孫之子杜林，為作訓故，并列焉。
>
> With the rise of the Han, the teachers of the villages combined the *Cangjie*, the *Yuanli*, and the *Boxue* compendia and divided them into sixty-character chapters. This consolidated *Cangjie pian* totaled fifty-five chapters. At the time of Emperor Wu (r. 140–87 BC), Sima Xiangru (ca. 179–117 BC) created the *Fanjiang pian* which had no duplicate characters in it. At the time of Emperor Yuan (r. 48–33 BC), Shi You, the Director of Eunuch Attendants, wrote the *Jijiu pian*; at the time of Emperor Cheng (r. 32–7 BC), Li Chang, the Chamberlain for the Palace Buildings wrote the *Yuanshang pian*. All [of these books] used the *regular forms* of the *Cangjie.* The *Fanjiang* was especially remarkable. During the Yuanshi era (86–80 BC), hundreds of *xiaoxue* experts were summoned and ordered to record the characters [they knew] at the court. Yang Xiong selected those that were useful and compiled the *Xuanzuan pian*, [intended as] a continuation of the *Cangjie*. He also amended the duplicate characters in the *Cangjie*.[141] [These two works amounted to] a total of eighty-nine chapters. I (i.e. Ban Gu) further continued Yang Xiong's work in thirteen chapters, [augmenting it to] a total of one hundred and two chapters. This work has no duplicate characters and roughly contains all the characters used in the books of the Six Arts.
>
> The *Cangjie* had many ancient characters and ordinary teachers were not able to read them. At the time of Emperor Xuan (r. 73–49 BC), [the court] summoned people from the state of Qi who could read [the characters] correctly.

[139] For examples of such cases, see Lin Suqing (1986), pp. 74–75.

[140] The commentators of the *Hanshu* unanimously agree that the character 書 (in parentheses) in this case is an error and should interpreted as 興 (in square brackets).

[141] Primers in traditional China had been constructed to have as few repeated characters in them as possible. One of the most famous such primers was the *Qianziwen* 千字文 by Zhou Xingsi 周興嗣 during the reign of Emperor Wu of Liang 梁武帝 (535–543), which consisted of 994 characters with almost no duplicates. I believe that Ban Gu referred to the "repeated characters" in the *Cangjie pian* in a similar sense.

> Zhang Chang learned from them and passed [the knowledge] down to his great-grandson Du Lin who interpreted and arranged the characters.[142]

Ban Gu used the word *zheng* "correct, standard" twice in the above passage. First, he described how scholars at he beginning of the Han used the "correct characters" from the *Cangjie pian* in their own works. A few lines later, he described how the court of Emperor Xuan summoned some Qi scholars who were still able to read the characters of the *Cangjie pian* "correctly." The reason for this was, according to Ban Gu, that the *Cangjie pian*, which was the consolidated version of the book compiled by village teachers at the beginning of the Han, contained many ancient character forms. The Qi scholars were summoned to read these ancient character forms. The difficulties in reading could have resulted from the fact that this version of the *Cangjie pian* was compiled from the three original works by village teachers who were less educated than court scholars. In the course of the approximately 150 years that separated the reign of Emperor Xuan from these rural teachers, the original three works compiled during Li Si's reforms (i.e. Li Si's *Cangjie pian*, Zhao Gao's *Yuanli pian*, and Humu Jing's *Boxue pian*) were most likely lost and there was nobody to correct the errors and decipher the ancient character forms that were potentially already corrupted. Because the knowledge of these character forms was already lost at that time, the term "reading [the characters] correctly" (*zhengdu* 正讀) probably referred to translating the ancient forms into modern standard forms.

The characters of the *Cangjie pian* which the Qi scholars were still able to read correctly were the same ones which Sima Xiangru, Shi You, and Li Chang used in their own works. Sima Xiangru compiled the *Fanjiang* sometime between 140 BC (i.e. the beginning of Emperor Wu's reign) and 117 BC (i.e. his own death); thus the *Fanjiang* predates the summoning of Qi scholars by about 50–100 years. Shi You and Li Chang, on the other hand, compiled their works after the Qi scholars had rendered the ancient forms of the *Cangjie pian* into standard forms. In this case the readings of the Qi scholars would have also functioned as the standard reading of obscure characters.

I have already discussed in Chapter Two the moral undertone of the term *zheng* in the *Shuowen* "Postface" and the *Hanshu* "Yiwenzhi" chapters. Both Xu Shen and Ban Gu claimed that scholars in government employment had to write with *zheng* characters and if someone did not (或不正), he was reported and punished, as Shi Jian had feared. To give

[142] *Hanshu*, "Yiwenzhi."

weight to their argument, Xu Shen and Ban Gu cited a quote from *Lunyu* 15:25 (吾猶及史之闕文) which they interpreted, departing from the original meaning of the passage, as Confucius lamenting over the inundation of official documents with non-*zheng* characters. The opposite of writing correctly in this context, apart from writing "incorrectly" 不正, was the notion of writing "according to one's fancy" 用己私. Writing according to one's fancy referred to disregarding a convention, that is, deviating from an official standard.

Ban Gu's claim about the written standard implies that in his time this standard was enforced only in high-level government employment. The implementation of rules prohibiting incorrect forms shows that by the end of the first century AD, a standard script had still not been effectively established even in government service. However, the scholars of this time possessed a firm belief in the necessity of a standard script.

From a perhaps extreme point of view, one could say that Chinese writing has never been a completely standardized system. The stone classics, the woodblock prints of the Daoist and Buddhist canons, the dictionaries and lexicons, the textbooks and primers were all involved in producing, spreading, and enforcing a standardized script. Starting from the Northern Qi 北齊 (550–577), there was even an official post called *zheng-zi* 正字, which Charles Hucker rendered into English as Proofreader.[143] The *Tongdian* 通典 recorded the responsibilities of the Proofreader the following way:

> 祕書正字， 後漢桓帝初， 置祕書監， 掌圖書古今文字考合同異。 其後監令掌圖籍之紀， 監述作之事， 不復專文字之任矣。 今之正字盖令監之遺職， 校書之通制， 歷代無聞。 齊集書省有正書， 北齊祕書省有正字。隋置四人，大唐因之，掌刊正文字，其官資輕重，與校書郎同。
>
> As for the Palace Library Proofreader, at the beginning of the reign of Emperor Huan (AD 147–167) of the Latter Han, [the government] established the Director of the Palace Library who was in charge of studying ancient and modern character forms, collating their similarities and differences. Later on, the Director became in charge of archiving books and documents. He supervised the process of recording and no longer specialized in the script. Today's Proofreader probably derives from the Director's post. There has never been a consistent system throughout the various dynasties for the Editing Clerks. In Qi, the Department of Scholarly Counselors had Proofreaders.[144] In Northern

[143] Hucker (1985), p. 125.

[144] Hucker (*Ibid.*) rendered both 正字 and 正書 into English as Proofreader. Although this eliminates the distinction between these two titles that is present in the Chinese text, the English rendition is adequate because, as the *Tongdian* explained, these two titles were alternate designations that referred to titles with similar responsibilities.

> Qi, the Department of Palace Library had Proofreaders. The Sui appointed four scholars and the Tang followed [this practice]. [These four scholars] were in charge of editing and correcting characters [in documents] and their salary and rank equaled those of the Editors.

The above excerpt claims that starting in the Eastern Han, the government had appointed clerks whose main responsibility was to correct character forms in official documents. The effort the government had put into enforcing the standard script testifies to the continuous existence of non-standard character forms. The existence of the office of Proofreader shows that, on the one hand, variant character forms were not tolerated in official documents, on the other hand, the presence of such forms was an actual problem in government service.

In sum, the variability of character structure was a historical reality in both pre-Han and Han times. Starting from around the 1st century BC, the government began enforcing the written standard for people employed in its administration. Nevertheless, character forms outside of this standard were still commonly used in everyday life. Although the examination of the variability of character structure in the post-Han period falls beyond the scope of this study, the abundant presence of popular character forms in Dunhuang manuscripts and medieval stone inscriptions bears witness to the contrast between the orthography of standard forms used in official business and non-standard forms used in everyday life.

4.2 The concept of standard in written English

A comparison of character variation in pre-Qin China with spelling variations in late medieval England provides a useful perspective on the concept of standard. Today, in the age of printed and electronic texts we sometimes take for granted that the spelling system of the English language is highly standardized. However, Middle English, as recorded in manuscript material, was characterized by a high diversity of written form. Angus McIntosh et al. have compiled the *Atlas of Late Mediaeval English* (hereafter: *Atlas*) which records, based on hand-written texts, spelling differences in different parts of England in the period of 1325–1450.[145] The *Atlas* displays the spelling variants of a number of commonly used

[145] The *Atlas* includes a small number of texts beyond this general period. McIntosh (1986), p. 3.

words, such as "which," "each," "through," on the map of England, showing that differences between spoken dialects could manifest themselves in the spelling habits of these areas too. Nevertheless, spelling variants did not always reflect differences in the spoken dialects but, in many cases, were the result of alternate ways of rendering the same sound into written form. Therefore, McIntosh saw the *Atlas* primarily as an atlas of written English dialects.[146]

The *Atlas* shows the variability of orthographic forms in late medieval English. The word "which," for example, which has only five letters in its modern form, could be spelled in hundreds of different ways (Fig. 4.1).

The examined manuscript material altogether yielded a total of 411 variant spellings of the word "which" in the corpus, most of which occurred only once. Only 17 forms of the word occurred more then a couple dozen times in the entire corpus:

the-which	*whyche*
the-whiche	*wich*
weche	*wiche*
wheche	*wych*
which	*wyche*
whiche	*ye-whilk*
whilk	*þe-which*
whilke	*þe-wyche*
whych	

Out of these 17 variants, the variants "which," "whiche," and "þe-wyche" occurred in greatest number. This shows that despite the great variety of spelling forms, a small number of forms was dominant. In most cases these dominant forms were the ones that eventually evolved into the modern forms of the words. On the other hand, those forms that occurred only once in the corpus suggest an almost unlimited potential for variation; thus the number of known variants keeps growing as new manuscripts are being discovered.

McIntosh explained the orthographic diversity in written Middle English in terms of the decline of the use of Old English in government affairs:

> In the years following the Norman Conquest, the written standard of late Old English was largely displaced as the language of government by Latin and Anglo-Norman; its role as the national literary language was also at an end. In the sense that it does not reflect a *national* standard, therefore, almost any Middle English written before *ca*. 1430 is 'dialectal' as a matter of definition.

[146] *Ibid.*, p. 6.

huch	*qwylke*	*the-wilk*	*wylke*	*ye-whech*	*yo-whuche*	*þ^{e}-whiche*
huche	*thee-wiche*	*the-wilke*	*wysche*	*y^{e}-whech*	*yo-whyche*	*þe-whilk*
huich	*the-wilk*	*the-wuche*	*wζuche*	*ye-wheche*	*yo-whylk*	*þ^{e}-whilk*
chuiche	*thee-wilke*	*the-wycche*	*y^{e}-qhych*	*ye-whice*	*yo-whylke*	*þe-whilke*
husche	*thee-wyche*	*the-wych*	*y^{e}-qhylk*	*ye-which*	*y^{o}-wilk*	*þ^{e}-whilke*
huysch	*thee-wylke*	*The-wyche*	*ye-quech*	*y^{e}-which*	*yo-wilke*	*þe-whoche*
huysche	*the-huche*	*the-wyche*	*ye-queche*	*ye-whiche*	*y^{o}-wylk*	*þe-whuch*
hwich	*the-qhylk*	*the-wyilke*	*y^{e}-qhyche*	*y^{e}-whiche*	*y-whilke*	*þe-whuche*
hwiche	*the-quech*	*the-wylke*	*ye-quhilk*	*ye-whik*	*þee-whylke*	*þe-whulke*
hwych	*the-queche*	*tho-qhwych*	*ye-quhilkes*	*ye-whike*	*þe-hucche*	*þe-whych*
hwyche	*the-quelk*	*w^{c}*	*ye-quhylk*	*ye-whilc*	*þe-huch*	*þ^{e}-whych*
qhiche	*the-quhilk*	*wch*	*ye-quich*	*ye-whilk*	*þe-huche*	*þe-whyche*
qhilke	*the-quhilkis*	*w^{ch}*	*ye-quiche*	*y^{e}-whilk*	*þe-hwyche*	*þ^{e}-whyche*
qhyche	*the-quhylk*	*wche*	*ye-quiche*	*ye-whilke*	*þeo-wheche*	*þe-whylk*
q^{i}che	*the-quich*	*wech*	*y^{e}-quiche*	*y^{e}-whilke*	*þeo-whuche*	*þe-whylke*
q^{j}che	*the-quiche*	*weche*	*ye-quilk*	*ye-whoche*	*þeo-whyche*	*þ^{e}-whylke*
quech	*the-quilk*	*wh^{c}*	*ye-quilke*	*ye-wholke*	*þe-qheche*	*þe-wich*
queche	*the-quych*	*whch*	*ye-quych*	*ye-whych*	*þe-quech*	*þe-wiche*
qelk	*the-quyche*	*whech*	*y^{e}-quych*	*y^{e}-whych*	*þ^{e}-quech*	*þe-wilk*
quhilk	*the-quylk*	*wheche*	*ye-quyche*	*ye-whyche*	*þe-queche*	*þe-wilke*
quhilkis	*the-qwech*	*whi^{c}*	*y^{e}-quyche*	*y^{e}-whyche*	*þ^{e}-queche*	*þe-woche*
quich	*The-qweche*	*whicche*	*ye-quylk*	*ye-whylk*	*þ^{e}-quich*	*þe-wuche*
quiche	*the-qweche*	*whice*	*y^{e}-quylk*	*y^{e}-whylk*	*þe-quiche*	*þe-wych*
quiche	*the-qwhich*	*which*	*ye-quylke*	*ye-whylke*	*þe-quilc*	*þ^{e}-wych*
quilc	*the-qwhiche*	*whiche*	*ye-qvhilke*	*y^{e}-whylke*	*þe-quilk*	*þe-wyche*
quilk	*the-qwhilk*	*whichee*	*y^{e}-qwech*	*ye-wich*	*þe-quyche*	*þ^{e}-wyche*
quilke	*the-qwhilke*	*whik*	*ye-qweche*	*y^{e}-wich*	*þe-qwech*	*þe-wyhche*
quych	*the-qwhilkes*	*whilc*	*y^{e}-qweche*	*ye-wiche*	*þ^{e}-qwech*	*þe-wylk*
quyche	*the-qwhyche*	*whilk*	*y^{e}-qwelk*	*y^{e}-wiche*	*þe-qweche*	*þe-wylke*
quyilk	*the-qwhylk*	*whilke*	*y^{e}-qweylke*	*ye-wihche*	*þ^{e}-qweche*	*þ^{e}-wylke*
quylk	*the-qwiche*	*whilk-þat*	*ye-qwhilk*	*ye-wilk*	*þe-qwhech*	*þo-whiche*
quylke	*the-qwilk*	*whoch*	*ye-qwhilke*	*ye-wilke*	*þ^{e}-qwhech*	*þo-whik*
qweche	*the-qwilke*	*whoche*	*ye-qwhilkes*	*y^{e}-wilke*	*þe-qwheche*	*þo-whilk*
qwech	*the-qwych*	*whochee*	*ye-qwhilkis*	*ye-woche*	*þe-qwhich*	*þo-whilke*
qweche	*the-qwyche*	*whuc*	*ye-qwhilkys*	*ye-wych*	*þe-qwhiche*	*þ^{o}-whilke*
qwech-p^{t}	*the-qwylk*	*whuce*	*ye-qwhylk*	*y^{e}-wych*	*þe-qwhilk*	*þo-whoche*
qwelk	*the-qwylke*	*whuch*	*y^{e}-qwhylk*	*ye-wyche*	*þ^{e}-qwhilk*	*þo-whylke*
qwhiic	*the-weche*	*whuche*	*y^{e}-qwhylke*	*y^{e}-wyche*	*þe-qwhilke*	*þo-wilk*
qwhich	*the-welke*	*whyce*	*ye-qwich*	*ye-wylk*	*þe-qwhyche*	*ζe-qwhylk*
qwhiche	*the-whech*	*whych*	*y^{e}-qwich*	*y^{e}-wylk*	*þ^{e}-qwich*	*ζwch*
qwhik	*the-wheche*	*whyche*	*ye-qwiche*	*ye-wylke*	*þe-qwiche*	*ζwche*
qwhilk	*the-whicch*	*whylc*	*y^{e}-qwiche*	*y^{e}-wylke*	*þ^{e}-qwiche*	*ζwich*
qwhilke	*the-whice*	*whylk*	*ye-qwilk*	*yhe-whilk*	*þe-qwilk*	*ζwiche*
qwhilkes	*the-which*	*whylke*	*y^{e}-qwilk*	*y^{o}-quech*	*þ^{e}-qwilk*	*ζwuch*
qwhylk	*the-whiche*	*wich*	*ye-qwilke*	*y^{o}-queche*	*þe-qwilke*	*ζwuche*
qwhylke	*th-whiche*	*wiche*	*y^{e}-qwilke*	*y^{o}-quelk*	*þe-qwych*	*ζwyche*
qwice	*the-whike*	*wihch*	*ye-qwilkis*	*yo-quich*	*þ^{e}-qwych*	*þe-wache*
qwich	*the-whilk*	*wilc*	*ye-qwych*	*y^{o}-quyche*	*þe-qwyche*	*þ^{e}-w^{c}*
qwiche	*the-whilke*	*wilche*	*y^{e}-qwych*	*yo-qwilk*	*þ^{e}-qwyche*	*þ^{e}-w^{ch}*
qwiche-p^{t}	*the-whuch*	*wilk*	*ye-qwyche*	*yo-qwych*	*þe-vche*	*þe-wech*
qwich-p^{t}	*the-whuche*	*wilke*	*y^{e}-qwyke*	*y^{o}-qwych*	*þe-vich*	*þ^{e}-wech*
qwilk	*the-whycche*	*woch*	*ye-qwylk*	*yo-qwyche*	*þe-viche*	*þe-weche*
qwilke	*the-whych*	*woche*	*y^{e}-qwylk*	*y^{o}-qwyche*	*þe-vuche*	*þ^{e}-weche*
qwycche	*the-whyche*	*wuch*	*ye-qwylke*	*yo-qwylke*	*þe-wheche*	*þe-whche*
qwych	*the-whychy*	*wuche*	*y^{e}-qwylke*	*y^{o}-wheche*	*þ^{e}-wheche*	*þe-whech*
qwyche	*the-whylk*	*wych*	*ye-weche*	*yo-whiche*	*þe-which*	*þ^{e}-whech*
qwyilk	*the-whylke*	*wyche*	*y^{e}-weche*	*yo-whilk*	*þ^{e}-which*	
qwyk	*the-wich*	*wyhche*	*ye-welk*	*yo-whilke*	*þe-whiche*	
qwylk	*the-wiche*	*wylk*	*y^{e}-welke*	*yo-whoche*	*þ^{e}-whiche*	

Figure 4.1. The 411 different spellings of the word “which” in McIntosh (1986).

> In the course of the fifteenth century, however, and not at the same rate in all places, regional diversity gives way increasingly to Chancery Standard, the official language of the London administrators and the direct ancestor of modern Standard English. By the end of the same century, moreover, the establishment of printing was instrumental in the redevelopment of a national literary standard. The dialects of the spoken language did not die out, but those of the written language did; and although there are some late survivals, they are no sufficient basis for a dialect atlas.[147]

McIntosh saw the orthographic variability of written English as a temporary phenomenon that was preceded by the "written standard of late Old English" and succeeded by the "Chancery Standard." The major cause of diversity was that the language was not used in the official sphere. This notion closely parallels the Chinese situation where the orthographic diversity of the pre-Qin scripts was gradually replaced by the standard of the clerical script. However, while the appearance of this standard was the result of the employment of the script in a standardized bureaucratic system, variant forms have continued to be used in the non-official sphere up until modern times.

The co-existence of an official standard with the orthographic diversity of popular use might have been true for written English too. It is commonly known, for example, that Shakespeare (1564–1616) wrote his own name, and was referred to by others, in a variety of different spellings. There has been a traditional consensus that the differences in these spellings reflected differences in usage, which, at least according to some scholars, was also an indication that the names referred to different people. David Kathman argues against this consensus:

> – "Shakespeare" was by far the most common spelling of the name in both literary and non-literary contexts, and there is no significant difference in spelling patterns when we take into account such factors as handwritten vs. printed and Stratford vs. London spellings;
>
> – there is no evidence that the variant spellings reflected a consistent pronunciation difference, but there is considerable evidence that they were seen as more or less interchangeable;
>
> – there is no evidence whatsoever that hyphenation in Elizabethan times was ever thought to indicate a pseudonym, and other proper names of real people were also sometimes hyphenated.[148]

In support of his argument, Kathman lists the different spelling variants of the name of Shakespeare in two tables (Figs. 4.2 and 4.3).

[147] *Ibid.*, p. 3.
[148] Kathman (http://shakespeareauthorship.com/name1.html).

Table 1. Non-literary references (1564–1616):

	Total	Outside London	In London
Shakespeare	71	8	63
Shakespere	27	25	2
Shakespear	16	16	0
Shakspeare	13	9	4
Shackspeare	12	11	1
Shakspere	8	7	1
Shackespeare	7	7	0
Shackspere	6	5	1
Shackespere	5	5	0
Shaxspere	3	3	0
Shexpere	2	2	0
Shakspe~	2	0	2
Shaxpere	1	1	0
Shagspere	1	1	0
Shaksper	1	1	0
Shaxpeare	1	1	0
Shaxper	1	1	0
Shake-speare	1	0	1
Shakespe	1	0	1
Shakp	1	0	1
with first ‘e’	128 (71%)	61 (59%)	67 (87%)
w/o first ‘e’	52 (29%)	42 (41%)	10 (13%)

Figure 4.2. Kathman’s first table with the non-literary references of Shakespeare’s name.

Table 2. Literary references (1593–1616):

	Total	Printed	Hand-written
Shakespeare	115	108	7
Shake-speare	21	21	0
Shakspeare	9	5	4
Shaxberd	4	0	4
Shakespere	3	0	3
Shak-speare	2	2	0
Shakspear	2	0	2
Shakspere	1	0	1
Shaksper	1	0	1
Schaksp.	1	0	1
Shakespear	1	1	0
Shakespheare	1	1	0
with first ‘e’	141 (88%)	131 (95%)	10 (43%)
w/o first ‘e’	20 (12%)	7 (5%)	13 (57%)

Figure 4.3. Kathman’s first table with the literary references of Shakespeare’s name.

The two tables show that although the modern form of the name was the dominant form even in Shakespeare's time, a "considerable variation" in spelling still existed. According to Table 2, the name was more consistent in print than in hand-written sources. Printing, through its influence on large numbers of people, undoubtedly played a decisive role in the standardization of the script both in Europe and China. Beside the contrast between printed and hand-written forms, there was also a noticeable difference in the uniformity of spelling in and outside London, the spelling in the city being more consistent. The orthographic consistency of urban areas probably resulted from the higher concentration of government offices and employees using the standard script, as well as from the more common access to printed media.

In contrast with the orthographic diversity of late Middle English and the variability of Shakespeare's name in the late Renaissance England, modern written English is perceived as a highly standardized script. Spelling variations for the most part are understood as mistakes and a dictionary of English usage would not include them. There are relatively few exceptions from the rule that each word has one orthographic form. Such exceptions are the so-called variant spellings usually included in the dictionaries. Most modern dictionaries list variant spellings for words such as "gauge/gage."

Although such variant forms do occur in modern written English, they represent rare exceptions to the general rule of orthographic uniformity. The spelling of the vast majority of English words is constant and occurs in the same form in all major dictionaries. We turn to these dictionaries when in doubt because we, together with Xu Shen and Ban Gu, would rather leave an empty space and look up the correct word than to write the word according to our "own fancy." One's ability to spell "correctly" today is an important indication of one's level of education and plays a decisive role in one's success in society.

In order to obtain a descriptive view of modern written English, as opposed to how educators believe it should be written, one should examine how common users of English, who represent the majority of the users, write. Since even dictionaries of usage are compiled from books, articles and other printed material, they embody a different level of orthographic uniformity than the writings of ordinary people. Writers, editors, scholars and all those people whose works appear in print represent a tiny but highly specialized fraction of the total number of users of written English. By virtue of their professional involvement in the production of written texts in the official domain, this "professional elite" is significantly better at adhering to the written standard than the general population.

Internet newsgroups are a convenient source for analyzing the writing habits of the general population. Several word frequency lists have been compiled by modern linguists from various newsgroups, but the compilers have removed the spelling variants from these lists, creating an artificial sense of orthographic constancy. Another factor leading to increased uniformity is the source of the text. Wordlists are generally derived from soc.* newsgroups which represent the more verbose and grammatically sophisticated segment of all newsgroups.[149] But the analysis of random spelling variants of individual words on a comprehensive newsgroup search engine shows that English words on the Internet are misspelled surprisingly frequently. Consider, for example, the spelling variations of the words in Figure 4.4.[150]

The examples show that, although for an educator these variants are but spelling mistakes, in their total number they constitute a pattern. The word "occurrence" is spelled correctly only about half of the time, and its second most frequent variant, "occurance," from a statistical point of view, is a common written form of the word. Some words have been "misspelled" so often that their misspelled forms were accepted as an alternate spelling. In the same study, the word "minuscule," for example, 68% of the time is spelled in alternative ways. The form "miniscule" has become so common that now it is included in some standard dictionaries as a variant of "minuscule."[151]

In the above examples, the "misspellings" are highly phonetic in nature. Almost all variant forms are phonetically valid renditions of the word they represent. Based on its pronunciation, the word "privilege" could easily be spelled "privlidge" and the reason that its "correct" spelling is different lies not in its modern pronunciation but in the historical connection between its pronunciation and written form, as well as the choices

[149] Obviously there are also other factors that distort the image of orthographic habits in modern written English. One such factor is the use of spellcheckers. Another one is that those who have access to computers and know how to type probably have better literacy skills those who do not. On the other hand, one could argue that the use of spellcheckers today is part of the process of writing, therefore the resulting text represents the writing habits of the people. Similarly, those people who use computers and the Internet write much more, and thus statistically produce more texts than those who do not have access to computers.

[150] The following data are cited from Cornell Kimball's article called "A Study of Some of the Most Commonly Misspelled Words" (http://www.barnsdle.demon.co.uk/spell/error.html). The study is based on the material derived from Deja-News, now called Google Groups (http://groups.google.com).

[151] *Ibid.*

mischievous 1203	occurrence 5508	privilege 7035	definitely 90565
mischevious 255	occurance 2109	priviledge 1050	definately 17904
mischievious 182	occurence 1983	privelege 521	definitly 2434
mischeivous 51	occurrance 221	privledge 331	definatly 1609
mischeivious 30	ocurrence 43	priveledge 238	definitley 322
mischevous 15	ocurrance 8	privilage 212	definatley 247
mischivous 8	ocurence 3	privelage 134	defenitely 158
mischieveous 7		privelige 61	defenitly 48
mischiveous 7		privlege 40	defenatly 44
mischeiveous 2		privalege 38	
		privelidge 36	(plus a few each of:
		privlage 21	defenately
		privaledge 18	defanatly
		privalage 15	defanately)
		privellege 12	
		privlidge 12	
		privlige 11	
		(plus a few each of:	
		privillege	
		privalige	
		privalidge	
		privillige	
		privillage	
		privelledge)	

Figure 4.4. Frequency of some misspelled words on Internet newsgroups. From http://www.barnsdle.demon.co.uk/spell/error.html

made by lexicographers. Because of the phonetic nature of the above variants, none of these misspellings would prevent the reader from understanding which word is meant by that particular orthographic form. The forms are unambiguous and fully functional. The only reason why they are not used in print is that they violate the standard.

From a statistical point of view, these variants are an important part of modern written English. Some of the misspelled variants are used more often than other correctly spelled, but otherwise, rare words.[152] The nearly complete absence of these alternate spellings from printed works demon-

[152] For example, a search today (August 25, 2004) on Google Groups for the misspelled form "definately" produced 863,000 hits, whereas only 1,850 for the correctly spelled but relatively rare word "abeyant." The word "ducal" occurred 6,600 times, "cinchonism" 184 times, "podetia" (plural form of "podetium") not once. Thus from a statistical perspective, the misspelled form "definately" is much more part of modern written English than the words "abeyant," "ducal," "cinchomism," and "podetia," all of which appear in *The American Heritage Dictionary*.

strates, once again, the distinction between official and popular writing. In the case of modern English, printed texts function as the official way of writing. When children study to read and write in schools, they all aim to acquire the skill of being able to write in this correct way. Their advance in society to a significant degree depends on their ability to write and speak in the standard way. There is, however, the popular domain, comprising the texts written by ordinary people outside the boundaries of the school or workplace. These texts include hand-written or typed material people use in their daily life, from shopping lists to notes left on refrigerator doors, from personal letters to scribbled memos. A member of the "professional elite" would naturally use the same standard spelling he is accustomed to in his work when writing informal things, yet the majority of the population would not consider important how they spelled a word on their shopping list.

Such spelling inconsistencies in English demonstrate that, from a descriptive point of view, written English has been, and still is, characterized by orthographic variability. The appearance of an orthographic standard in both English and Chinese has created a duality between this ideal standard and the way people wrote in practice. While the standard could be enforced in official spheres, such as schools and publications, there was no authority when it came to the way people wrote in their private life.

In this chapter, I examined the concept of the standard. I have shown that a standard script is a result of a prescriptive measure enforced by an authority. When it comes to Warring States writing, today we have no records of an authority prescribing a standard. The earliest references to standardization are found in Han sources such as the *Shiji*, the *Hanshu*, and the *Shuowen*. However, the enforcement of an orthographic standard by the government did not mean the complete elimination of variant forms. While people conducted official business using the standard script, in their private life they did not observe the same level of orthographic uniformity.

Similar to Warring States writing, late medieval English showed a great degree of orthographic diversity. The same word could be written in a number of different ways, some forms being more frequent than others. In the case of the word "which," there were three most common spelling forms which included the modern form of the word. The spelling variants of the name of Shakespeare show that spelling variants were common in late the 16th and early 17th century too. At this time, orthographic inconsistency was a common phenomenon and did not indicate low literacy skills on the part of the writer. Finally, the examination of some frequently

misspelled words from Internet newsgroups shows that even today there is a significant disparity between the spelling practices of the official (printed media, schools, workplace, etc.) and private (non-printed media, personal correspondence, recreation, etc.) spheres.

When examining Warring States writing, my aim is to look at character forms without the assumption of a standard. Although it is probable that there had been reforms even before the Qin and Han standardization of writing,[153] the Warring States archaeological finds present an image of orthographic diversity. By looking at the character forms alone, it is not possible to demonstrate that any of the several forms was "the correct one." Therefore, I do not treat character variants as character forms differing from a correct or standard form but as alternate ways of writing the same word none of which was more correct than the other.

[153] A possible reform of writing would have been represented by the Historian Zhou's supposed creation of the *Dazhuan* compendium in the court of King Xuan of Zhou, an event referred to by Xu Shen in the *Shuowen* "Postface."

CHAPTER FIVE

VARIATION

In Chapter Three, I defined character forms in terms of the words they represented, regardless of their graphical appearance or resemblance to modern character structures. In Chapter Four, I showed that one cannot study Warring States character forms from the point of view of standard or correct forms, and that one should not presume normative differences between variant forms.

Up to this point, I have only given a basic definition of variation, according to which two character forms were different when they both represented the same word and yet they differed structurally. I stated in Chapter Three that the structure of a character form was determined by the components and their arrangement. Components, in turn, were modular building blocks that originated as individual characters with their own semantic and phonetic values. However, in order to compare variant forms with each other, it is necessary to define two more parameters. First, one needs to limit, and define, the scope of the material examined; second, to determine what constitutes a structural difference between two character forms.

Limiting the scope of the material means setting the boundaries of the community that used the same script. I shall call this script "local script." Any character form within a local script would have been clear and unambiguous for all literate members of the community. I consider documents originating from the same time and location, written on the same medium for similar social use, to be part of the same local script. In this study, the concept of local script pertains only to grammatological attributes and does not include the peculiarities of the language of a particular document. Thus even if a document was written elsewhere but then copied by a local scribe, it would then belong to the local script of the copier.

The first task in determining the structural differences between two character forms is to define which character forms one considers identical in a local script. The emphasis here is on the structure of a character form, as opposed to its identity in the context in which it was used. In reality, the differentiation of variant forms based on their structure is an artificial distinction which did not exist for those people who wrote or read

these character forms. They probably would not have even registered most of the differences between variants because they were primarily concerned with what meaning and sound a character could represent. Such differentiation in this study is done purely for the sake of demonstrating the dissimilarity between the character forms and observing the patterns of variation.

5.1 The local corpus

It has never been doubted that Chinese writing has variant forms. Out of the 50-60,000 characters featured in comprehensive Chinese dictionaries, such as Morohashi Tetsuji's 諸橋轍次 *Daikanwa jiten* 大漢和字典, the majority of character entries are variants of characters listed under other entries. Most of these variants have been gathered from various printed editions of transmitted texts and are considered to be part of the literary and linguistic heritage of China. Character variants from excavated Warring States texts have only recently started to appear in dictionaries.

Most modern scholars attribute character variants to temporal and regional peculiarities. On the one hand, this view is not incorrect. During the 3,000 years that separate the oracle-bone inscriptions from the simplified script of modern China, the Chinese script has undergone changes that produced numerous alternate character forms. On the other hand, newly excavated texts show that the pre-Qin scripts included a large number of variants, the differences of which cannot be attributed to temporal and spatial factors. Some modern scholars regard these variants as mistakes or typos. The large number of such "mistakes" that modern transcribers of early manuscripts have identified in the course of their work created the impression that scribes in early China lacked literacy skills. Beside disagreeing with the basic assumption of orthographic uniformity, I think that it is unreasonable to assume that we, who in many cases cannot even decipher the inscriptions correctly, have higher literacy skills with respect to Warring States writing than professional scribes who wrote these inscriptions on a daily basis. The quality of the literary output of the Warring States period also negates the low literacy skills of literate people in this period.

Xu Shen referred to differences in the scripts of the seven states in his "Postface" as the main factor that necessitated the Qin reform of writing. In the *Shuowen*, he listed over 170 examples of variant character forms in

the non-Qin states, most of them from the states of Chu and Qi.[154] Pre-Qin texts also occasionally referred to the linguistic divergence of the spoken dialects and written scripts of the various states. The following passage from *Mengzi* 孟子 3B:6 uses the analogy of language acquisition to discuss how one could learn to be a good ruler.

> 孟子謂戴不勝曰：『子欲子之王之善與。我明告子。有楚大夫於 此，欲其子之齊語也，則使齊人傅諸，使楚人傅諸。』曰：『使 齊人傅之。』曰：『一齊人傅之， 衆楚人咻之， 雖日撻而求， 其齊也不可得矣。 引而置之莊嶽之間，數年雖日撻而求，其楚亦不可得矣。』
>
> Mencius said to Dai Busheng, 'Do you wish your King to learn goodness? I shall speak to you plainly. Suppose a Counselor of Chu wished his son to learn the language of Qi. Would he have a man from Qi to tutor his son? Or would he have a man from Chu?'
> 'He would have a man from Qi to tutor his son.'
> 'With one man from Qi tutoring the boy and a host of Chu men chattering around him, even though you caned him every day to make him speak Qi, you would not succeed. Take him away to some district like Zhuang and Yue for a few years, then even if you caned him every day to make him speak Chu, you would not succeed…'[155]

The archaeological discoveries of the last few decades confirmed that the scripts in the various states could differ from each other not only in their calligraphic style but also in their character structure. Many modern scholars have compared the character forms of the various scripts and pointed out their dissimilarities. Figure 5.1 lists some of Zhu Dexi's 朱德熙 examples of the characters 馬, 安, and 乘 as they were written in the six non-Qin states during the Warring States period.

Zhu Dexi's examples show the diversity between the character forms of the six states. The Qi forms of the character 馬 ( ,  , and ) were so different from the Yan and Jin forms ( ,  , and ) that a person not familiar with Warring States writing would not be able to tell that these forms were related to each other. Other forms can be seen as intermediary forms between the otherwise graphically very distinct character forms. The lower part of Chu-1a (), for example, represents the horse's legs, while the top part of the character form, the horse's eye. Qi-1a (), Qi-1b (), and Qi-1c () include the legs but omit the eye,[156] whereas Jin-1a () and Jin-1b () include the eye but completely omit the legs. Thus

[154] Ma Zonghuo (1959), p. 1.
[155] Translated by D. C. Lau in *Mencius* (1970), pp. 111–112.
[156] The top part of the three Qi forms depicts the hair of the horse, which is also part of the Chu forms.

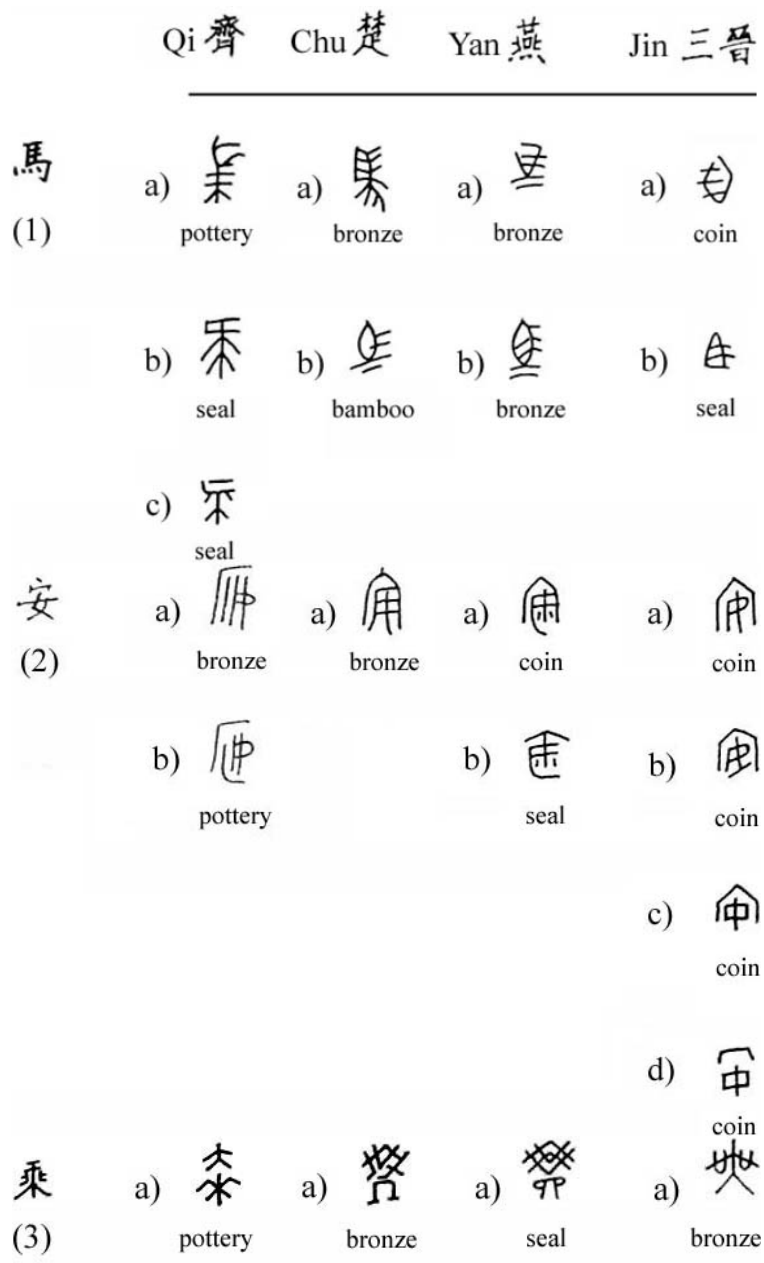


Figure 5.1. Examples of variant character forms used in the six states. Adopted from Zhu Dexi (1995), pp. 74–76. I added the type of media on which each character form was written. The Three Jin states in the table indicate the states of Han 韓, Zhao 趙, and Wei 魏.

one could see Chu-1a () as the most complete form and the Qi ( and ) and Jin ( and) forms as simplified forms that omit different parts of the complex form.

Zhu Dexi's examples illustrate that character forms in different scripts *could* significantly vary from each other. On the other hand, the examples convey a false impression that within the same script, character forms were more or less uniform. Because there are only 1–4 examples of each character per state, the selected forms imply that they epitomize the regular way of writing a given character in that state. In reality, however, Zhu's examples represent character forms that were used on different material. For example, the above forms of the character 馬 were written on the following types of media.[157]

Qi-1a		Pottery
Qi-1b		Seal
Qi-1c		Seal
Chu-1a		Bronze
Chu-1b		Bamboo
Yan-1a		Bronze
Yan-1b		Bronze
Jin-1a		Coin
Jin-1b		Seal

Pottery forms can considerably differ from seal or coin forms, not to speak of the brush-written bamboo strip forms. For example, the two Chu forms of the character 馬 were written on different media. Chu-1a () comes from the "Ejun Qi jie" 鄂君啟節 bronze inscription and Chu-1b () from the Yangtianhu 仰天湖 bamboo slip manuscripts. Clearly, these two forms are quite dissimilar, even though they are both from the state of Chu. Indeed, Qiu Xigui uses the "Ejun Qi jie" inscription and the Wangshan 望山 bamboo slips, the character forms of which are nearly identical with those of the Yangtianhu slips, to demonstrate the difference between popular and formal scripts (Fig. 5.2).

Qiu regards the "Ejun Qi jie," and bronze inscriptions in general, as examples of the formal script, and the Wangshan texts, along with other bamboo slip manuscripts, as examples of the popular script. He points out

[157] *Ibid.*, pp. 74–75.

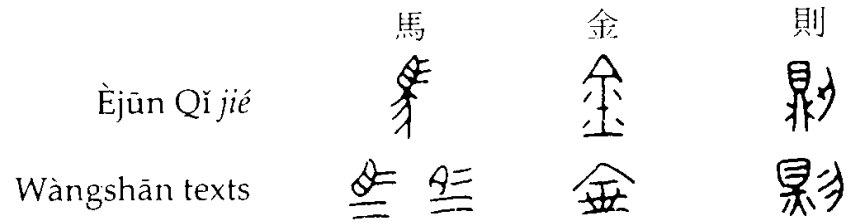


Figure 5.2. Demonstration of the differences between popular and formal scripts in Qiu (2000), p. 88.

that even though these two sources were close in date, their character forms exhibited considerable differences.

The lesson from the above examples is two-fold. First, character forms differed not only between the six states but even within the same state. The differences between the formal and popular forms confirm the existence of such disparity. Second, one cannot compare individual character forms written on different media in different states to show the variability of character structure between the states. Although Chu-1b () appears to be similar to Yan-1a () and Yan-1b (), the first was written on bamboo and thus represents the Chu popular script, whereas the second two forms were written on bronze and represent the Yan formal script. In the state of Chu, the character 馬 was written on bronze as seen on Chu-1a (), a form very different from the two Yan forms.

In reality, the relationship between individual character forms was even more complex. The Chu bamboo slip forms of the same character 馬 show that there were notable differences between character forms within the same script written on the same type of media.[158]

Although all of these forms come from Chu bamboo slips and therefore in Qiu Xigui's definition are popular forms, the disparity between two forms could be as large as that between the forms written on different media.

One might argue that these discrepancies still resulted from different forms having been used in different areas of the state or in different periods of the Warring States period. Twentieth-century examples illustrate

[158] From Teng Rensheng (1995), pp. 750–755. I have chosen only a few examples out of the larger number of forms which Teng considers variants.

how character forms can undergo significant changes within the course of only a few decades.[159] However, there are numerous examples in which Warring States character forms on single documents belonging to the same time period and geographical area still show variation. The *Guodian Laozi* manuscript has a total of six forms of the character 難, representing all of the occurrences of the word *nan* "difficult" in the manuscript (Fig. 5.3).[160]

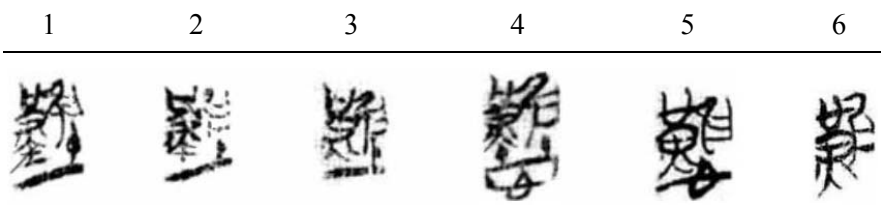


Figure 5.3. The six forms of the character 難 in the *Guodian Laozi*.

The main difference between these six forms is in the inclusion or exclusion of the 土 and 心 components. The six character forms can be dissected into components in the following way:

#1 = (難 + 土) + 土
#2 = (難 + 土) + 土
#3 = 難 + 土
#4 = (難 + 土土) + 心
#5 = 難 + 心
#6 = 難

Forms #1 and #2 are structurally identical. Form #3 is similar to the previous two but omits the small 土 component under the 堇. Forms #4 and #5 both feature an additional 心 component below the entire charac-

[159] I am primarily referring here to the simplification of Chinese characters in mainland China that began in the 1950s. There had been two major lists of simplified characters adopted by the PRC government, one in 1956, and the other in 1964. However, modifications have been carried out even until the 1980s. See Qiu (2000), pp. 404–407.

[160] The organizers of the Guodian bamboo slips identified three sets of slips with texts affiliated with the transmitted version of the *Laozi*. These were called manuscripts *Laozi A* (甲), *B* (乙), and *C* (丙). The organizers differentiated between these three sets of manuscripts solely on the basis of the length and shape of the slips. There was no obvious difference in handwriting between the three groups, containing a total of almost 1,700 legible characters.

ter, but while #4 has a double 土 component, the bottom parts of which are joined together, #5 completely omits the 土 component. Form #5 is identical in structure to the existing character 戇 defined in the *Shuowen* as "to revere" 敬也; it is also a variant attested in the Ming dynasty dictionary *Zixue sanzheng* 字學三正 (1601) as a *guwen* form of the character 然. Form #6 omits both the 土 and 心 components and thus matches the modern form of the character 難. Since all of these forms are from the same document, they most likely represented the hand of the same scribe.[161] This means that the random addition or omission of certain components was within the boundaries of the variation tolerated by the community in which the manuscript was written and read.

The variant forms of the character 難 within the *Guodian Laozi* are not an isolated case of structural variability. In fact, variability is common in those excavated pre-Qin documents that are long enough to have an array of duplicate characters. But structural variation can appear even on two sides of the same seal. Lin Suqing 林素清, for example, notes that both of the following imprints, taken from two sides of the same seal, carry the same Jiang Quji 江去疾 name (Fig. 5.4).[162]

Although the parallel forms of the characters 江 (and) and 疾 (and) also show some degree of orthographic dissimilarity, it is the two forms of the character 去, written as (#1) and (#2) that differ from each other the most. Form #1 consists of the components 大 + 口, although the component 口 lacks its top horizontal stroke. Form #2 consists of 彳 + 止 + 大 + 口.

Thus differences in structure sometimes can be seen even on inscriptions which consist of only a few characters. My contention is that in the case of an entire local script, such as the ones defined by the Guodian or

[161] Since there is a difference in the length of slips in the three sets of slips the organizers called *Laozi A, B*, and *C*, it is possible to argue that they do not represent one document. However, five out of the six above forms of the character 難 occur in *Laozi A*. Only form #5 comes from *Laozi C*, the other five forms are all from *Laozi A*. Moreover, these five forms are all in close vicinity of each other, from slips #12, #14, #15, and #16. Forms #1, #2, and #3 even occur within the same line of thought:

> 多易必多難。是以聖人猶難之，故終無難。
>
> Who has many easy things is bound to have many difficulties. Therefore the sage seems to regard things as difficult, thus in the end there is no difficulty for him.

[162] Lin Suqing (1976), p. 36.

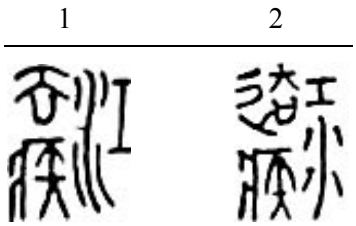


Figure 5.4. Two sides of the same seal with the identical inscription of the name Jiang Quji 江去疾. From Lin Suqing (1976), p. 36.

Baoshan corpora, which include all writings from the same region within a given time period, variability of character structure was *de facto* a part of the script.

Another important parameter of a local corpus in my definition is that the texts must be used in the same social domain, i.e. popular or formal.[163] I do not regard texts written on bronze vessels (formal script) part of the same local script as texts written on bamboo and wood tablets (popular script). Therefore, beside time and space, social use represents the third parameter of a local script.

The category of local script is more useful for the study of variation than that of a single manuscript because it is based on the notion of a community that is able to communicate via that script. The study of a single manuscript is essentially the study of one person's handwriting, which cannot represent "writing" at any given time and place in history, just as the utterances of one person do not represent the entire spoken dialect.[164] As speakers of the same spoken dialect understand all other speakers of the same dialect, literate people using the same local script were able to read texts produced by any other user of the same script.

Analyzing an entire local script is important because this way one has a chance to look at character forms in their totality. By gathering variants together, one can discern the graphical, semantic, or phonetic limits that

[163] Although Qiu Xigui's distinction of popular and formal script can be applied to most examples of Warring States writing, there are a small number of cases which appear to fall outside of these two categories. Such cases are represented by stylized and often highly simplified character forms that occur on pottery, coins, and certain types of seals. I am of the opinion that the script used on these types of media should be grouped into a separate category, perhaps called "stylized script."

[164] By using the word dialect I do not refer to a non-standard way of speaking but to "a variety of language that with other varieties constitutes a single language of which no single variety is standard" (*The American Heritage Dictionary*).

were involved in representing a particular word in writing. Because one person's handwriting did not represent the entire community, a single manuscript is not useful for such study. Similarly, a larger corpus of texts is also inadequate because it goes beyond the limits of a community that shared the same spoken language.

Tolerance plays a key issue role in reading variant forms. Strictly speaking, there are no two completely identical character forms. Even if the forms appear to be indistinguishable, there are inevitably minute differences between them. When the character forms are written by different people, the dissimilarities are larger, sometimes even visible. Therefore, tolerance plays an important role in the recognition of any handwritten script, including the modern scripts. People recognize the same character despite the fact that every time they see a new example of it, it differs from any other they have seen before. In the case of the written dialects of the Warring States period, the structural dissimilarities between variant character forms fell within the limits tolerated by the community. I have already pointed out that the structural variability within the script did not indicate the complete absence of rules. Instead, it meant that the community possessed a higher degree of tolerance towards such variability than we do today.

5.2 Degree and nature of dissimilarity

Variation was not deviation from a correct or absolute form but dissimilarity between two forms, none of which was more correct than the other. This distinction is important in order to avoid the common mistake of looking at variants in the light of modern character forms. Although the structure of modern character forms is for the most part based on the structure of dominant pre-Qin forms, this becomes evident only in retrospect. The linear connection between the structure of dominant pre-Qin and modern forms is not automatic and in many cases does not hold true. There are many cases when a modern character form derives from a regional variant, rather than the mainstream Qin form. A frequently cited example is that of the character 明, which is written with the component 日 in its modern form, whereas in the Qin script the component 日 was almost always written as 囧. This structure was also commonly used in the Han clerical form of the character. In the pre-Qin Chu script, on the other hand, the character was written with the components 目 or 日. On post-

Han manuscripts and stone inscriptions, 明 is written either as 日+月 or 目+月, whereas the 囧+月 form gradually disappeared. Thus in the case of the character 明 the clerical form, contrary to Xu Shen's claim on the nature of the Qin reforms of writing, evolved from a non-Qin form.

In order to see early character forms without a bias, one should avoid giving preference to any form over the other. A view that favors the antecedents of modern forms and regards them as the correct forms would inevitably conflict with the reality of multiple concurrent character forms in early China.

In Chapter Three, I identified a character form in terms of the word it represented. According to this definition, any two graphical forms that were accorded the same meaning and the same pronunciation are variants of each other. Even though the structure of two character forms corresponds to the structure of different characters in the modern script, if the two forms represented the same word in context, they were alternate graphical representations of the same word and, therefore, are variants of each other. Because the identification of character forms is not based on graphical appearance, the degree of similarity or dissimilarity between written forms of the same word becomes irrelevant.

However, in order to determine the rules that governed the variability of character forms from the point of view of the modern observer, it is necessary to determine the minimum difference between two character forms that would qualify them as distinct forms. In this study I consider two character forms to be different if they differ in at least one component. In Chapter Three, I followed Xu Shen in defining components as those character modules that are, or at one point had been, individual characters with a set of associated phonetic and semantic values. If the part that graphically distinguishes two forms is not an entire component, I consider the two forms to be structurally identical.

To illustrate this definition, I am going to look at a passage that occurs twice in the *Guodian Laozi* manuscripts, once at the beginning of *Laozi A* and once at the end of *Laozi C* (Fig. 5.5). The transcription of the two versions follows below, each in three lines, corresponding to the arrangement of the photographic forms in Figure 5.5. Although I use the transcription from the *Guodian Chumu* (1998) volume, I do not transcribe the original character structure of the Chu character forms but instead substitute the modern characters for the *words* represented by those character forms. For example, instead of transcribing the third character form as 谷 and putting the character 欲, which is its modern reading, into parentheses, I directly write the character 欲.

Figure 5.5. Two parallel passages from the *Guodian Laozi*. For the sake of convenience, I broke up the original vertical orientation of the text and align the two versions horizontally. The photographs of the first version are from *Guodian Chumu* (1998), p. 4, those of the second version are from p. 10. In my arrangement, the order of the text goes from left to right, the character forms in the first version appear above the corresponding character forms of the second version. Because the passage was too long to fit on one line, I had to break it into three sections. Nevertheless, they should be understood as a continuous text.

Version I.

聖人欲不欲，不貴難得之貨。
　教不教，復眾之所二過。是故聖人

能輔萬物之自然，而弗能為。

Version II.

[聖]人欲不欲，不貴難得之貨。
　學不學，復眾之所二過。是以
能輔萬物之自然，而弗敢為。

Version I.

The sage desires not to desire and does not value the goods that are hard to obtain. He teaches not to teach and restores that in which the masses have erred. Thus the sage could support the spontaneity of the myriad things but is not able to act upon this.

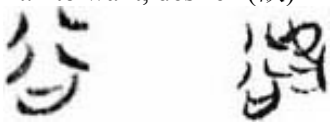
Version II.

[The sage] desires not to desire and does not value the goods that are hard to obtain. He learns not to learn and restores that in which the masses have erred. Thus he could support the spontaneity of the myriad things but dares not to act upon this.

Apart from lexical differences, such as alternation between the phrase 是故聖人 in Version I and the phrase 是以 in Version II, which are not related to character structure, there are 12 discernible dissimilarities between the character forms of the two versions (Fig. 5.6).[165] The differences between these 12 character pairs vary in degree. Below I examine each pair in detail in order to demonstrate my definition of variation on concrete examples.

Figure 5.6. The dissimilarities of corresponding character forms in the two parallel passages.

1. *Yu* “to want, desire” (欲)



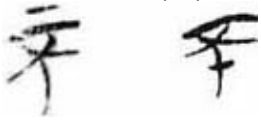
1 2

These two forms represent the word *yu* “to desire,” written in the modern script as 欲. Form #1 coincides in structure with the modern character

[165] A few of these character forms occur twice in the text but, since they are structurally identical to each other, I only list them once.

谷, form #2 with 欲. The *Shuowen* claims that the character 欲 was a *xingsheng* character which consisted of the semantic 欠 and phonetic 谷 components. Both of these components are complete characters by themselves with their associated sound and meaning: 欠 is pronounced *qian* and carries the meaning of "to lack; to yawn;" 谷 is pronounced *gu* and carries the meaning of "valley." The omission or inclusion of the 欠 component in the character forms thus constitutes, according to my definition, a case of structural variation. Therefore the above two forms are variants.

2. *Bu* "no, not" (不)



1 2

The two written forms of the word *bu* "no, not," written in the modern script as 不, differ from each other in the placement of a single horizontal stroke. This stroke appears at the top of the character in the form #1 () and at the lower part of the character in the second (). One could think of it as a displacement of the stroke. However, in the Guodian corpus these two strokes sometimes appear together in the same character (), in other cases they are both omitted (). There seems to be no significance to their presence or absence. The main question in determining whether forms #1 and #2 are variants is whether either of these strokes represents a component by itself. Xu Shen wrote that 不 was a pictograph and listed it as one of the 540 *bushou*.

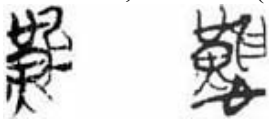
不，鳥飛上翔不下來也。从一，一猶天也。象形。

The character 不 represents a bird flying upwards and not coming down. It consists of 一. This 一 represents the sky. The character 不 is a pictograph.

In contrast with Duan Yucai's 段玉裁 edition of the *Shuowen*, quoted above, Xu Xuan's 徐鉉 edition has 从二猶天也 instead of 从一，一猶天也. The difference lies in how closely the two horizontal lines would have followed each other in the vertical text: if they had been close, they should be read as the character 二, if they had been further apart, they should be read as two consecutive occurrences of the character 一. In either case, the small seal form of the character given in the *Shuowen* () does not include either of the two "displaced" strokes seen in forms #1 and #2.

The "sky" in Xu Shen's explanation of the character most likely referred to the top horizontal stroke in 帀. In the Chu forms an addition or omission of either of those two strokes appears to have been no more than a calligraphic preference. Because of this, I conclude that the forms #1 and #2 do not differ structurally from each other.

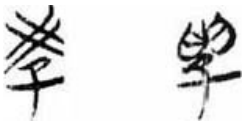
3. *Nan* "hard, difficult" (難)



1 2

These two forms represent the word *nan* "difficult, hard," written in the modern script as 難. The difference between the two forms is that form #2 has an additional 心 component at the bottom. Form #1, on the other hand, matches the character's modern structure which consists of 堇 and 隹. Although the component 隹 is slightly abbreviated in form #1, both forms include these two basic components. However, the presence of the component 心 in the form #2 represents a structural variation. Therefore, I regard the two forms as variants.

4. *Jiao* "to teach" / *xue* "to learn" (教/學)



1 2

According to the editors of the *Guodian Chumu* volume, these two forms represented the words *jiao* "to teach" and *xue* "to learn," written in the modern script as 教 and 學. If they were indeed two different words, then the variation is lexical and not orthographic. From the point of view of grammar, both words fit the context. The sage can "teach not to teach," just as he can "learn not to learn." However, the graphical similarity between the two character forms implies a connection. The difference between the two forms is that form #1 has the component 爻 on the top, whereas form #2 has the component 臼, with a vertical stroke in its center. The modern form of the character 學 is in fact a combination of these two forms, since its upper part consists of the component 爻 placed inside the component 臼.

Form #1 appears in the *Shuowen* as a separate entry, consisting of 子 and the phonetic 爻, with a small seal form written as 斆. Xu Shen glossed

 with the character 放 ("to release") which, in turn, he explained as 遂 ("to obey"). Duan Yucai's commentary added that the characters 放 and 仿 were interchangeable and that Xu's gloss of the character meant "to follow and rely on." Duan's commentary further explicated that both the characters 教 () and 學 were *huiyi* compound characters derived from the character .

It is clear that the modern characters 教 and 學, as well as the words they represent, are closely related and that there is no reason to favor the reading of form #1 as 教 over 學.[166] Theoretically speaking, form #1 could stand for either of these words. However, since forms #1 and #2 occur in the exact same context I favor the idea of reading them as two written forms of the same word. In light of the unquestionable reading of form #2, I believe that both forms stood for the word *xue* "to learn." This conclusion is further supported by the image of the sage presented in the *Laozi*. The sage's attitude towards the people is strictly pragmatic and consists of "emptying their minds and filling their bellies, weakening their wills and strengthening their bones" 虛其心，實其腹，弱其志，強其骨 (*Laozi* 3). He is never engaged in the education of the masses. In fact, he is against it. On the other hand, the sage does perfect himself, even if he does this in a negative direction by discarding the knowledge he has learned. The words "those who study increase daily, those who follow the Dao decrease daily" 為學日益，為道日損 (*Laozi* 48) express the same idea as "learning not to learn."

Thus forms #1 and #2 stood for the same word. At the same time, their top parts comprise different components. Therefore, I consider the two forms structural variants.

5. *Fu* "to restore" (復)

1 2

The two forms of the word *fu* "to restore," written in the modern script as 復, share the same basic structure: 彳 + 止 + 目 + 夂. The difference between the two forms is that form #1 leaves out the bottom stroke of 立, merging it with the top of the 目 component. Since the presence or absence of this stroke does not alter the identity of the components, I do not consider the two forms variants.

[166] Wang Li (1982, p. 300) also grouped the two characters together as cognates.

6. *Zhi suo* “that which … of” (之所二)

1 2

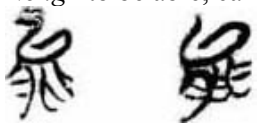
A special case involves the two ways of writing the characters 之所. In form #1, the two characters are written separately, whereas in form #2 they are joined into a single *hewen* 合文. As usual, the presence of a *hewen* is indicated with a repetition mark affixed at the bottom right corner of the character. However, form #1 in which the two characters are written separately also has a repetition mark. Neither of the two functions of the repetition mark, i.e. indicating the merge of two characters into a single *hewen* or the repetition of the entire character (*chongwen* 重文), seem to be appropriate here. Thus there is no explanation to the function of the repetition mark in this case. However, my main interest lies in whether these two ways of writing 之所 constitute a case of character variation. In form #2, the characters are not only closer to each other but also share a horizontal stroke. Nevertheless, I cannot consider the two forms variants because the characters 之所 stood for a string of words and not a single one.

7. *Guo* “to err” (過)

1 2

These two character forms represent the word *guo* “to err,” written in the modern script as 過. Form #2 consists of form #1 (彳 + 辶) and an additional 彳 component. Although in the modern script the components 彳 and 止 together make up the component 辶, in pre-Qin scripts they were separate entities, each of which was often used separately from the other. Because of the absence of the component 彳 in form #1, I consider the two forms variants of each other.

8. *Neng* “to be able, can” (能)



1 2

Form #1 of the word *neng* “to be able, can,” written in the modern script as 能, consists of the components 厶 + 肉 + 匕 + 匕. This form is

structurally identical to the *Shuowen* small seal form, written as  . Form #2 consists of the components 厶 + 肉 + 爪 + 爪. The two forms differ from each other in that the two 匕 () at the right bottom side of form #1 are written in form #2 as two 爪 (). Xu Shen defined the character 能 in the *Shuowen* as follows:

能，熊屬。 足似鹿 ， 从肉厶聲。 能獸堅中， 故稱賢能，而彊壯稱能傑也。

The character 能 means a kind of bear. Its legs are like those of a deer. The character consists of 肉 and the phonetic 厶. The animal *neng* is resolute [in his nature,] that is why they call him able and worthy (*xian-neng*). At the same time he is physically powerful, that is why they call him mighty and heroic (*neng-jie*).

Thus Xu Shen saw the component 比 common to the characters 能 and 鹿 as the legs of those animals, even though elsewhere he defined 比 as an image of two people following each other and 匕 as an image of a reversed person 人. This shows that he took the double 匕 component to be in fact a simplified version of the double 爪 component. Forms #1 and #2 share the same components: 厶 + 肉 + 爪 + 爪, only in form #1 the two 爪 components are written with one less stroke each.[167] Therefore, I do not consider the two forms variants.

9. *Fu* “to support” (輔)


1


2

According to the editors of the *Guodian Laozi* volume, these two character forms represented the word *fu* “to support,” written in the modern script as 輔. Form #1 consists of 父 + 田 + 又, form #2 consists of 木 + 甫. At first glance the two forms have nothing in common. However, since the component 甫 () itself was a combination of 父 () and 用 (),[168]

[167] Other forms of the character from Guodian sometimes even further simplify this component, either omitting one of the 爪, as in the form  , or merely referring to both components by two lines, as in the form  .

[168] The structure of the character 甫 is also clearly visible in its *Shuowen* small seal form written as  . Xu Shen wrote that the character 甫 consists of the components 用 and 父, 父 also acting as the phonetic component.

the two character forms share the phonetic component 父 (). Although in the modern script both the 父 + 田 and the 父 + 用 combinations were assimilated into the character 甫, in the above two Chu forms they were clearly distinct. Because the two forms differ in at least two components (田+又 and 用), I regard them as variants.[169]

10. *Wan* "ten thousand" (萬)


1 2

The above two forms represent the word *wan* "ten thousand," written in the modern script as 萬. The structure of the two forms is identical, except for the additional 土 component at the bottom of form #2. Because of this discrepancy, I regard the two forms as variants of each other.

11. *Fu* "no, not" (弗)


1 2

According to the editors of the *Guodian Chumu*, these two character forms represented the word *fu* "do not," written in the modern script as 弗. The editors of the volume not only think that the forms represented the same word, they consider them structurally identical to the modern character 弗.[170] The two above forms, however, differ in structure. While form #2 does match the structure of the modern character 弗, form #1 could be dissected differently. Its top part resembles the character 女, whereas the bottom right is unidentifiable. The usual form of the character 女 in the manuscript is  but in the character 奴, written as  , the component 女 is identical to the top part of  . This difference between forms #1 and #2

[169] Although form #1 is the equivalent to the modern character 尃 and thus one could make a case for *jiajie* borrowing, I would argue that this was a case of graphical variation rather than substituting the original character with a homophonous one.

[170] Whenever the editors of the volume think that there is a structural dissimilarity between the Chu form of a character and its modern equivalent, they transcribe the structure of the Chu form and indicate in parentheses which modern character it equals. They transcribed the above two character forms, however, directly into the modern character 弗. Cui Renyi (1998, pp. 39, 45) follows them in this respect.

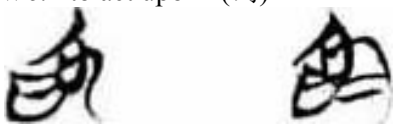
of the character 弗 was not an isolated incident in the *Guodian Laozi* manuscripts. Below are some examples of these forms in the *Guodian Laozi*:

Form #1 

Form #2 

While form #2 occurs elsewhere in the Guodian corpus, form #1 occurs exclusively in the *Laozi* manuscript. The examples demonstrate the graphical consistency within the two distinct groups. The two forms are also used consistently in the manuscript: whenever the word represented by the form occurs more than once in a sentence, the same form is used in each case. The analysis of the usage of these two forms, which falls outside the scope of this study, could reveal some differences in their grammatical function. However, if I follow the editors of the *Guodian Chumu* volume and accept that these two forms indeed stood for the same word, I would have to regard them as variant forms.

12. *Wei* "to act upon" (為)



1 2

These two forms of the word *wei* "to act upon," represented in the modern script by the character 為, differ from each other in two strokes. In other forms of the same character in the Guodian corpus, the two horizontal lines seen in form #2 are sometimes heavily accentuated:  . Xu Shen listed the character 為 under the 爪 *bushou* and interprets it as pictograph of a mother monkey.

為，母猴也。 其為禽好爪， 爪， 母猴象也；下腹為母猴形。 王育 曰：爪，象形也。

古文為，象形，兩母猴相對形。

The character 為 is a mother monkey, which is an animal that likes grabbing. The component 爪 (grab; claws) is an image of a mother monkey. The lower part of the character represents the shape of a mother monkey. Wang Yu claimed that 爪 was a pictograph.
The *guwen* form of the character 為 is a pictograph which represents the shape of two mother monkeys facing each other.

The *Shuowen* small seal form of the character shows that the two horizontal lines in the Chu forms of the character () are basically a simplification of the four legs of the mother monkey. The claws and body of the monkey, however, are represented even in the form omitting these two lines. Therefore, I do not regard forms #1 and #2 as variants.

Among the above 12 pairs of character forms, the dissimilarities between 4 pairs did not qualify as structural changes (Fig. 5.7). These forms are not variants because there is no component-level discrepancy between them. Also, the character forms in pair #6 stood for a string of two words and thus did not satisfy the requirement that the variants must stand for single words.

	2	6	8	12
Modern form	不	之所	能	爲
I				
II				

Figure 5.7. Pairs of character forms that do not qualify as structural variants.

On the other hand, 8 pairs of character forms differed from each other in at least one component (Fig. 5.8). None of the character forms in these pairs was more correct than its variant form. One cannot talk about *zhengzi*

or *zhuzi* 主字 because these concepts are meaningful only from the point of view of an existing standard. These variants were not variations from a standard form but variants of each other, and there is no reason to think that contemporary Chu people regarded any one form to be more accurate than the other.

	1	3	4	5	7	9	10	11
Modern form	欲	難	學	復	過	輔	萬	弗
I								
II								

Figure 5.8. Pairs of character forms that qualify as structural variants.

It is not always clear whether the difference between two forms is structural or not. One might argue that the horizontal strokes appearing and disappearing on the Chu forms of the character 不 were meaningful as individual components. Therefore, regardless of how clearly I strive to define variation, I cannot eliminate the subjective element. Besides, it is worth keeping in mind that although a modern researcher recognizes variation between structural components, the overall visual design of a character form could have sometimes been more important to the members of a community than its structural composition. If the principles of symmetry (composition) were important enough, they may have taken precedent over literal (semantic) issues.

The examination of the two above parallel passages from the *Guodian Laozi* manuscript serves to illustrate my definition of structural variation. I attempted to draw a line between component-level and other, less significant changes, such as differences in strokes or stroke order. The practical purpose of such demarcation was to create a working model for the comparative analysis of variant character forms of the Houma covenant texts in Chapter Six.

CHAPTER SIX

PATTERNS OF VARIABILITY

In this chapter I demonstrate that while on the level of a single character form structural variability occurs as a random phenomenon, when examining a multitude of forms together, the differences in character structure exhibit a clearly distinguishable pattern. By identifying this pattern, I intend to show that the literate members of the community that shared a script in the Warring States varied the structure of the character forms, consciously or unconsciously, within certain principles and did not permit, or tolerate, variations outside these principles.

For the purpose of detecting the patterns of variability, I chose the Houma covenant texts, which consist of hundreds of nearly identical copies of the same few basic oaths. Thus the same word often occurs in the same context hundreds of times, making the covenant texts an ideal corpus for examining the various ways ancient scribes wrote a particular character. For my analysis, I selected three words which appear in the corpus in identical grammatical context: *zhi* "will, intention," *fu* "to restore," and *fu* "abdomen." For the evaluation of the differences, or similarities, between character forms, I utilize the criteria I defined in the previous chapter.

6.1 Introduction to the corpus

The Houma covenant texts are a group of inscribed jade and stone tablets discovered in 1965–1966 near the city of Houma 侯馬, Shanxi province. The total number of excavated fragments exceeded 5,000 but not all of them were inscribed.[171] Most of the tablets were in the shape of gui 圭, that is, long rectangular plaques with a pointed end. The texts on

[171] Starting from the 1930s, covenant texts were also unearthed in Wenxian. In 1942, a few dozen jade tablets, some inscribed, were found in Qinyang 沁陽. The inscriptions on these tablets were similar to those found at Houma.

the tablets date to the end of the Spring and Autumn period around the beginning the 5th century BC, sometime between 497 and 470.[172] Because the texts were not incised but written with ink and a stylus or pen, they represent the way people actually wrote in the time of Confucius and his disciples. The texts were so-called "covenant texts" from the state of Jin 晉 that recorded the oaths of vassals swearing alliance and loyalty to a covenant lord. The organizers of the texts identified this lord as Zhao Yang 趙鞅, also known as Zhao Jianzi 趙簡子.[173]

In 1976, the Shanxi Cultural Relics Working Committee 山西文物工作委員會 published hand-tracings of over 600 tablets and 200 photographs of the original tablets in a volume called *Houma mengshu* 侯馬盟書. Figure 6.1 shows a representative example of a jade tablet, categorized by the editors of the volume as "lineage covenants" 宗盟類.

My own transcription of the text on the above tablet appears below. I do not reproduce the structure of the original character forms (i.e. do not use the *liding* technique) but directly use the modern character forms used today to write the words of the covenant text. For example, in the first line, I use the character 盟 instead of 明 which would be the modern-day structural equivalent of the character form on the jade tablets, because the original character form in this case recorded the word meng "covenant." When the modern equivalent of a character is not certain ( , 擅, 渙,  , etc.), I include the transcription supplied in the *Houma mengshu* volume.

> 朝，敢不剖其腹心以事其主，而敢不盡從嘉之盟、定宮平畤之命；而敢或改擅及渙卑不守二宮者，而敢有志復趙尼及其子孫、先之子孫、先直及其子孫、  之子孫、史醜及其子孫、于晉邦之地者，及群乎盟者，吾君其明殛視之，靡夷非是。
>
> Should Zhao[174] dare not to split open his abdomen and heart in serving his lord,[175] should he dare not to fully adhere to His covenant and the mandate of [the covenants concluded] at the Dinggong and Pingzhi temples; or should he

[172] Based on the comparison of the Houma covenant texts with the Wenxian covenant texts, Zhou Fengwu (1994, pp. 126–127) claimed to have determined the exact date of the covenant: November 12, 496 BC.

[173] *Houma mengshu*, pp. 65–68. The identity of the covenant lord is by no means certain and is still subject of an ongoing debate among scholars.

[174] The person pledging alliance in this document is Zhao, who talks of himself in the third person. I choose not to use the "I, Zhao, …" form in my translation because it creates problems later when the document refers to Zhao with the pronouns 其 and 之 which should be translated as "his" and "him" rather than "mine" and "me."

[175] Huang Shengzhang (1981, pp. 28–29) has convincingly argued that the character transcribed by the editors of the *Houma mengshu* as 宗 ("ancestor" or "ancestral temple") should be transcribed as 主.

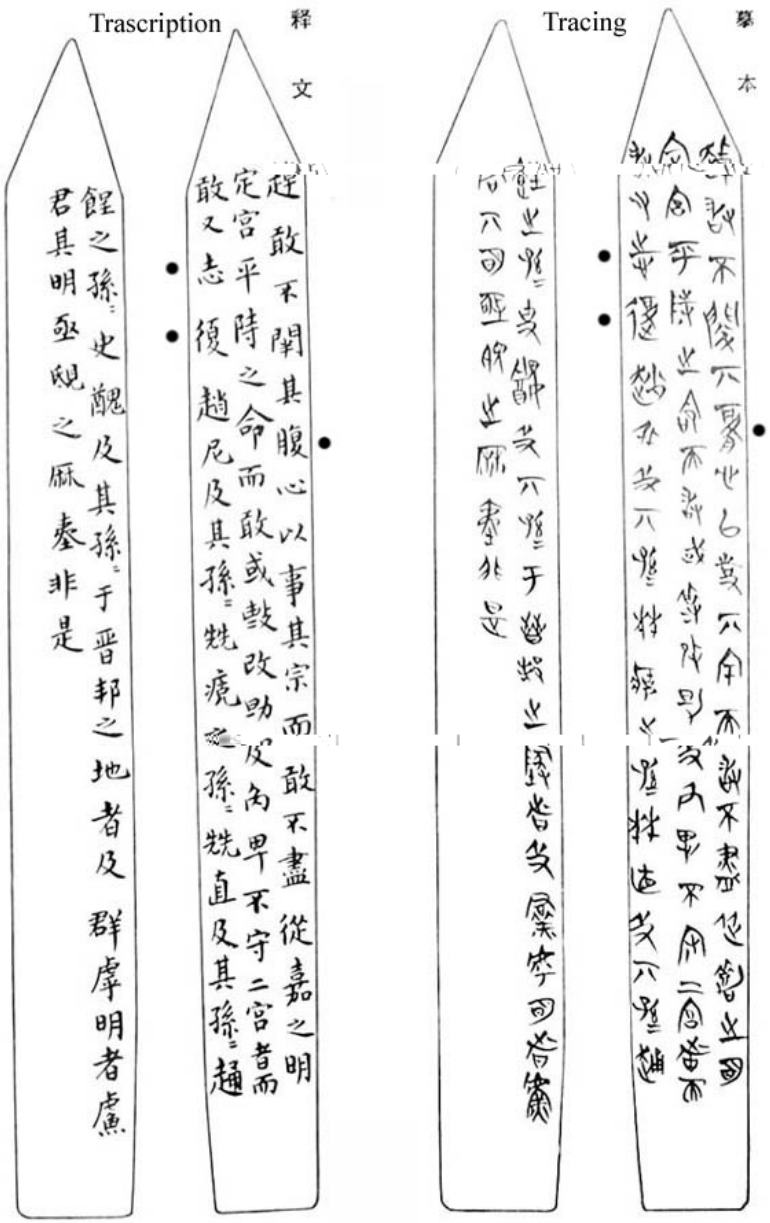


Figure 6.1. An example of the Houma covenant texts. The tracing of the tablet is on the right side, the transcription of the editors of the *Houma mengshu* is on the left. Adopted from *Houma mengshu*, p. 35. Although there are a few other texts in the Houma corpus, this oath includes all three words I examine in this chapter. I marked these words with black dots by their side on both the tracing and the transcription.

dare, in any respect, to initiate breaking of the faith or dispersion [of the alliance], interruption in the guardianship of the two temples;[176] or should he dare to have the intention to restore Zhao Ni[177] and his descendants, the descendants of Xian Ke, Xian Zhi and his descendants, the descendants of Yong Cheng, Shi Chou and his descendants on the territory of the state of Jin; or join in covenant with them; may the bright spirits of our [former] lords punish and scrutinize him[178] with no [degree of] destruction being inappropriate.[179]

[176] I am uncertain about the meaning of the phrase 而敢或𢦏改擅及渙卑不守二宮者. I am not entirely convinced by the interpretation of the scholars who worked with the text. For the sake of convenience, I follow Susan R. Weld's (1997, p. 142) translation.

[177] Zhou Fengwu (1994, pp. 113–117) claims that Zhao Ni was the same person known from historical sources as Zhao Ji 趙稷.

[178] Susan R. Weld (1997, p. 142) translates the phrase 吾君其明亟視之 as "may our former rulers, far-seeing, instantly detect me." I have departed from this translation in several points:

1. I take the character transcribed by the editors of the *Houma mengshu* volume as 亟 (ji "immediately") to stand for the word written in transmitted documents as 殛 (ji "to punish, kill"). There are examples in the *Zuozhuan* where the same word is used in an identical context: "should someone transgress this covenant, may the bright spirits punish him" 有渝此盟，明神殛之 ("Xi gong" 僖公 28). The *Shangshu* also used the same word as a verb: "Because the Xia had many offenses, the Will of Heaven punished them" 夏多罪，天 命殛之 ("Tang shi" 湯誓).
2. I believe that the character 明 refers to the spirit of former lords, i.e. ancestors. The justification for my interpretation is as follows:
   a. This meaning of the character is attested in the compound word *mingqi* "spirit vessel" 明器. The *Liji* 禮記 explains, "They call it 'spirit vessel' because they belong to the brightness of spirits" 其曰明器，神明之也 ("Tangong" 檀弓).
   b. Parallel sentences in the *Zuozhuan* (in addition to the one cited above) indicate that, in this context, the character 明 referred to the brightness of the spirits: "should someone transgress this covenant, … may the bright spirits and former lords correct and punish him" 有渝此盟……明神先君，是糾是殛 ("Xi gong" 28); "… may the many spirits of famous mountains and rivers, the bright spirits of former kings and dukes and of the ancestors of the seven clans and twelve states punish him" 名山名川羣神羣祀, 先王先公七姓十二國之祖，明神殛之 ("Xiang gong" 襄公 11; note: 祀 should be read as 祇, meaning "spirit of the earth").
   c. In other sources the compound word "bright spirits" 明神 appears as "brightness of the spirits" 神明. For example: "should someone transgress this covenant, may the brightness of the spirits punish him" 有渝此盟，神明是殛 (Fan Ye's 范曄 [398–445] commentary to the *Houhanshu* 後漢書, "Yuan Shao" 袁紹).
   d. In the *Sanguozhi* 三國志, the character 明 occurs in an identical context: "should someone transgress this covenant, … may the brightness of our

The above oath represents a basic template at the beginning of which the oath takers inserted their own names, thus producing a large number of nearly identical texts with different names. In this particular oath on the above image, the name of the oath taker was Zhao 朝.

6.2 Analysis

The three words whose written forms I examine in this chapter are *zhi* "will, intention" (written in the modern script as 志), *fu* "to restore" (written in the modern script as 復), and *fu* "abdomen" (written in the modern script as 腹). The word *zhi* "will, intention" is written with a relatively high degree of consistency, the structure of the other two words is less constant. My intention is not to choose the words with the most versatile

---

great spirits examine him." 有渝此盟……于爾大神，其明鑒之 ("Wushu," 吳書). This line also suggests that the character 視 in the Houma covenant texts, written as 鑒 in the above quote from the *Sanguozhi*, should be understood in the sense of "to examine, scrutinize" and not as "to see, detect."

[179] The last four characters of this oath, transcribed by the editors of the *Houma mengshu* as 麻夷非是, has been a subject of a number of scholarly studies. Most of the scholars agree that the characters 麻夷 mean "to destroy, annihilate." (麻 is equated with modern character 摩, whereas 夷 in itself has the meaning of "annihilation.") Only Qi Guiyan (1979, pp. 272) believed that 麻夷 was the name of Hebo 河伯, the God of the Yellow River. The character 是 is commonly interpreted as "clan," written in the modern script as 氏. Opinions varied, however, regarding the identity of the third character (非). Chen Mengjia (1966), erroneously, transcribed it as 我; Zhu Dexi and Qiu Xigui (1972, pp. 73–74) were of the opinion that 非 should be understood as the demonstrative pronoun 彼 ("that"). However, as I argue in Galambos (2005), the characters 非是 appear in the covenant texts over two hundred times the same way, with very little variation; not once was 彼 (or any form that could be linked to it) used as a substitute for the character 非. Therefore, on the basis of statistical frequency throughout the corpus, I am of the opinion that the last two charactes should be both transcribed and interpreted as 非是. The first character should be transcribed as 靡 in the sense of "not have; there is no," as suggested by the fact that a few oaths have the character 亡 (equivalent in usage in Warring States times to the modern character 無) in the same place. In view of the above considerations, the four characters in questions should be transcribed as 靡夷非是 and interpreted in context as a) "with no [degree of] destruction being inappropriate"; or b) "it would not be right if destruction did not [befall him]".

graphic forms but to show that the degree of variability itself was not consistent: some character forms varied more than others. The last two words, *fu* "to restore" and *fu* "abdomen," are particularly interesting because they are similar both graphically and phonetically, while standing for semantically distinct words.

6.2.1 *Zhi* "will, intention" (志)

The word *zhi* "will, intention," written in the modern script as 志, appears in the covenant texts in the phrase "should dare to have the intention of restoring…" 敢有志復… The character forms below all appear in this phrase. Based on their structure, the 232 legible character forms can be grouped into 5 categories (Table A).[180]

A1.	221		之+心
A2.	6		之+寺 (=寺)
A3.	4		之+心+寸
A4.	1		心
A5.	1		之+寸+阜

Table A. Frequency of variant character forms representing the word *zhi* "will, intention" in the Houma corpus

Form A1, which is also the most common form, matches the description of Xu Shen who wrote that the character 志 meant "intention" 意 and consisted of the component 心 and the phonetic component 之.[181] Form

[180] Appendix II contains the complete list of character forms.

[181] There is an apparent graphical similarity between the Warring States forms of the characters 之 () and 止 (). Because the Warring States form of the character 之 () resembles the modern form of the character 止, the two characters in pre-Qin manuscripts can be easily confused. Moreover, when the character 之 () appears as a component in a composite structure such as, it is usually equivalent to the components 土 (e.g. = 寺) or 士 (e.g. = 志).

A3 has an additional 寸 component at the lower part of the character. This extra component represents a hand and can also be transcribed into modern form as 又. Thus the lower part of the character has two components: 心 and 寸.

Form A2 also includes the component 寸 but omits the component 心. Thus the structure of this form can be derived from either form A3, by way of omitting the component 心, or form A1, by way of substituting 寸 for 心. Without form A3, forms A1 and A2 would create the impression that the components 寸 and 心 were interchangeable. However, form A3 reveals that these two components in the above character forms were complimentary, rather than interchangeable.

Form A4 consists of the component 心, without the top part of the character 志 that Xu Shen identified as the phonetic component. Since form A4 occurs only once in the Houma covenant texts, it is possible that the character form is damaged and its original form was similar to forms A1 or A3.[182] However, it is also possible that form A4 is intact and that it was the ancient scribe who omitted the top part of the character.[183]

As for the frequency of the above four character forms with respect to the total number of forms, Figure 6.2 shows the distribution of the 232 forms. According to the diagram, form A1 was by far the most common way (95.3%) of writing the word *zhi* "will, intention," whereas other forms were used seldom.

6.2.2 *Fu* "to restore" (復)

The word *fu* "to restore," written in the modern script as 復, appears in the covenant texts in the phrase "should he dare to have the intention of restoring Zhao Ni and his descendants…" 敢有志復趙尼及其子孫…[184]

[182] Evidently, this was the opinion of the editors of the *Houma mengshu* volume because this form is not even listed among the other distinct forms in the dictionary part of the volume. See *Houma mengshu* (1976), p. 310.

[183] It is also possible that the ancient scribe, when writing the character 心, intended to write not the word *zhi* "will, intention" but the word *xin* "heart" which also had the meaning of "will, intention." This usage is seen, for example, in the phrase "two men of the same will" 二人同心 (*Guanzi* 管子, "Xinshu shang" 心術上).

[184] The word *fu* "to restore" occurs in yet another sentence in some of the covenant texts: "or [dares to] once again let them enter the state of Jin" 或復入之于晉邦之地. However, in this case the word is used as an adverb and should be translated as "once again." Since in modern Chinese both meanings are written with the same character (復), the editors of the *Houma mengshu* volume (pp. 331–332) listed both the words *fu* "to restore" and *fu* "once again" under the same entry.

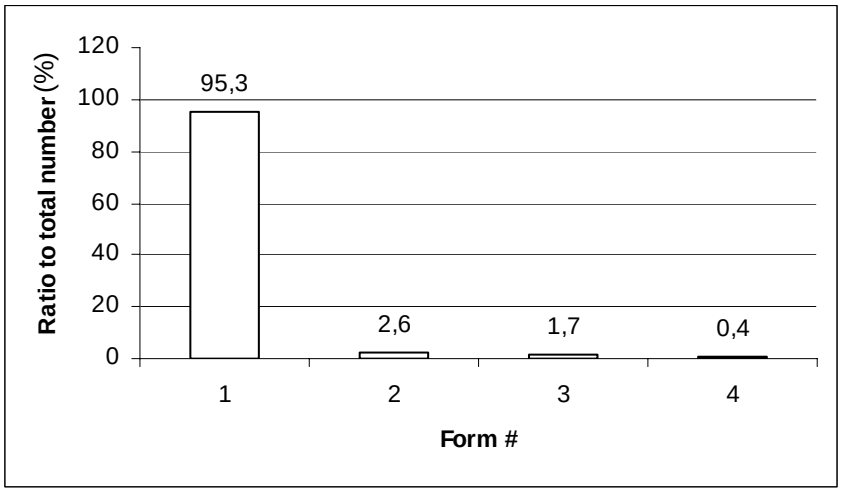


Figure 6.2. Frequency of the variant forms of the character 志 within the Houma corpus.

B1.	153		复+彳 +止	B6.	3	 复
B2.	19		复+肉+彳	B7.	2	复+肉
B3.	14		复+彳	B8.	1	复+止
B4.	5		复+彳 +口	B9.	1	复+口
B5.	5		复+彳 +止+口			

Table B. Frequency of variant character forms representing the word *fu* "to restore" in the Houma corpus

In order to provide identical context for all of the character forms I compare here, I only include character forms occurring in the context "should he dare to have the intention of restoring Zhao Ni and his descendants…"

The character forms above all appear in this sentence. Based on their structure, the 207 legible character forms can be grouped into 9 categories (Table B).[185]

Among the listed character forms, form B3 matches in structure the modern form of the character 復. However, as seen from the frequency of character forms, the most common way of writing the word *fu* "to restore" was form B1 which, in contrast with the modern form, includes a 止 component at the bottom.[186] The components 彳 and 止 are consistent with the meaning "to restore". The component 彳 often occurs in characters with a meaning associated with "going" or "moving", e.g. 行 ("to travel"), 往 ("to go"), 征 ("to go on an expedition"). The combination of the components 彳 and 止, as seen in forms B1 and B8, typically appears in the modern script as the component 辶/辵 which usually signifies movement and appears in characters like 巡 ("to go around"), 適 ("to approach"), 過 ("to go across"), 進 ("to advance"), 返 ("to return"), 運 ("to move, ship"), 道 ("path").[187] Therefore, the choice of the ancient scribe to write the forms B1, B5 and B8 with the components 彳 + 止 could reflect the notion that Zhao Ni, whose restoration the contracting parties of the Houma covenants were aiming to avoid, would have come back to Jin from another region.

Forms B4, B5, and B9 include the component 口 which is not part of the modern form (復). A possible explanation for the presence of this component is that the restoration of Zhao Ni's clan was viewed as a political act which involved the notion of "declaring." Another possibility is that 口 was added not as a separate component but as a component linked to the 止 (appearing in form B5), together forming the component 足. In forms B4 and B9, which do not include the component 止, the component 口 could simply be the abbreviation of 足. The component 足 could signify a "base" (as suggested by the component 止) or movement (as suggested by the component 辶). The presence of the component 口 in group C below is likely a result of an influence of group B.

Forms B2 and B7 include the component 肉 which, from the point of view of character structure, could suggest that for the ancient scribe the

[185] Appendix II contains the complete list of character forms.

[186] Once again, it is important to keep in mind that one can consider the 止 component additional only from the perspective of the modern form.

[187] In some modern characters the components 彳 and 止 remained separate. The meaning of these characters, however, is also often related to movement. E.g., 徙 ("to move"), 從 ("to follow"), 徒 ("to walk on foot"). In other characters, such as 復 and 後, the component 止, originally present in the pre-Qin forms, disappeared from the modern form.

word *fu* "to restore" also had some relationship with the human body. While form B7 occurs only once in the corpus, form B2 with 19 instances is the second most common form. Form B2 is identical in structure to the most common written form of the word *fu* "abdomen" (C1 below) that appears in the first half of the same sentence. This form is also identical to the modern form of the word *fu* "abdomen," written as 腹. Regarding the rather loose connection of the concept of restoration (which in this concept certainly does not involve resurrection) with the human body, I am of the opinion that the relatively high frequency of this form in the corpus could be caused by a sort of structural cross-influence between the written forms of the words *fu* "to restore" and *fu* "abdomen" within the same document.[188] It is quite possible that if the word *fu* "abdomen" had not been part of the same document, then the ancient scribes would have written form B2 with a significantly lower level of frequency.

As for the frequency of the above character forms in relation to the total number of forms, Figure 6.3 shows the distribution of the 207 forms. According to the chart, although form B1 was by far the most common (73.9%), forms B2 (9%) and B3 (7%) were also common ways of writing the word *fu* "to restore." Other forms occurred less commonly, some only once or twice.

6.2.3 *Fu* "abdomen" (腹)

The word *fu* "abdomen," written in the modern script as 腹, appears in the covenant texts in the phrase "should dare not to split open his abdomen and heart in serving his lord " 敢不剖其腹心以事其主… The character forms below all appear in this phrase. Based on their structure, the 225 legible character forms can be grouped into 22 categories (Table C).[189]

These 22 different forms show that the ancient scribes who recorded the covenant texts allowed themselves a great degree of orthographic flexibility in writing the word *fu* "abdomen." Form C1, the most common character form, is identical in structure to the modern form of the character 腹. Although there is some discrepancy in the location of the 肉 component, sometimes appearing on the left side, sometimes at the bottom (), the basic components are the same in each case: 复 + 肉.

[188] Other examples of such cross-influence within the same document are the word *fu* "abdomen" analyzed below, as well as the word *ming* "bright" sometimes occurring in a form identical to the modern character 盟.

[189] Appendix II contains the complete list of character forms.

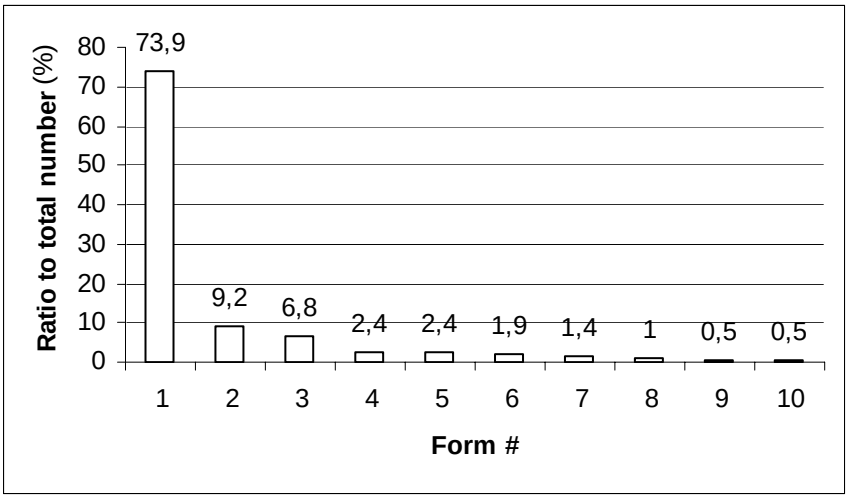


Figure 6.3. Frequency of the variant forms of the character 復 in the Houma corpus.

C1.	74	复+肉	C12.	2	复+夂
C2.	46	复+肉+彳	C13.	2	复+肉+冖
C3.	27	复+止+彳	C14.	3	复+肉+彳+冖
C4.	14	复+肉+夂	C15.	2	复+肉+彳+夂
C5.	13	复+肉+止	C16.	2	复+彳+夂
C6.	9	复+彳	C17.	2	复+彳+止+夂
C7.	8	复+肉+口	C18.	1	复+冖
C8.	5	复+止	C19.	1	复+彳+心+冖
C9.	4	复+彳+口	C20.	1	复+口
C10.	4	复+彳+口+止	C21.	1	复+彳+心
C11.	4	复	C22.	1	复+心

Table C. Frequency of variant character forms representing the word *fu* "abdomen" in the Houma corpus

The second most common form in the corpus (C2) differs from form C1 in the presence of an additional 彳 component. The same 彳 component is also present in forms C2, C3, C6, C9, C10, C14, C15, C16, C17, C19, and C21, being used in a total of 100 forms, that is, nearly half of all 225 forms. Therefore, it is obvious that for the ancient scribes the component 彳 was an important orthographic element for writing the word *fu* "abdomen," even though they omitted it in the dominant form (form C1). Form C2 also coincides with form C2 for the word *fu* "to restore,"[190] showing that the scribes sometimes wrote the words *fu* "abdomen" and *fu* "to restore" in a graphically identical way. Since the component 彳 does not seem to serve an apparent semantic purpose in writing the word *fu* "abdomen," I am of the opinion that the reason why the ancient scribes added this component was because they were influenced by the various forms of the word *fu* "to restore" within the same texts.

Form C4 includes the component 攵, which also occurs in forms C12, C15, C16, and C17. This form is identical to the lower part of the component 复 but appears at the top of the character form. The component ⟙ in forms C13, C14, C18, and C19 is a possible variant of 攵.

Forms C7, C9, C10, and C20 include the component 口 which, as seen above in case of the word *fu* "to restore,"[191] might have been an abbreviation of the component 足 used interchangeably with 止.

Forms C19, C21, and C22 include the component 心, which could be tentatively explained the following way. First, the component 心 generally indicated that the meaning of the character had an emotional or mental aspect to it. There is also an obvious physiological connection between the abdomen and the heart, therefore the component 心 could represent the body. This view is supported by the absence of the component 肉 (used for a similar purpose) in all three forms in which the component 心 occurs. Another explanation would be that the character immediately following 腹 (心 in the compound word *fuxin* "body and mind" 腹心) influenced the ancient scribe in writing the structure of 腹.

As for the frequency of the above character forms in relation to the total number of forms, Figure 6.4 shows the distribution of the 225 forms. According to the chart, the number of structurally distinct spellings of the word *fu* "abdomen" (22) significantly exceeded those of the words *zhi* "will, intention" (4) and *fu* "to restore" (10) analyzed earlier in this chapter.[192] At the same time, the frequency of the dominant form (form C1,

[190] See section 6.2.2 above.
[191] See section 6.2.2 above.
[192] See sections 6.2.1 and 6.2.2 above.

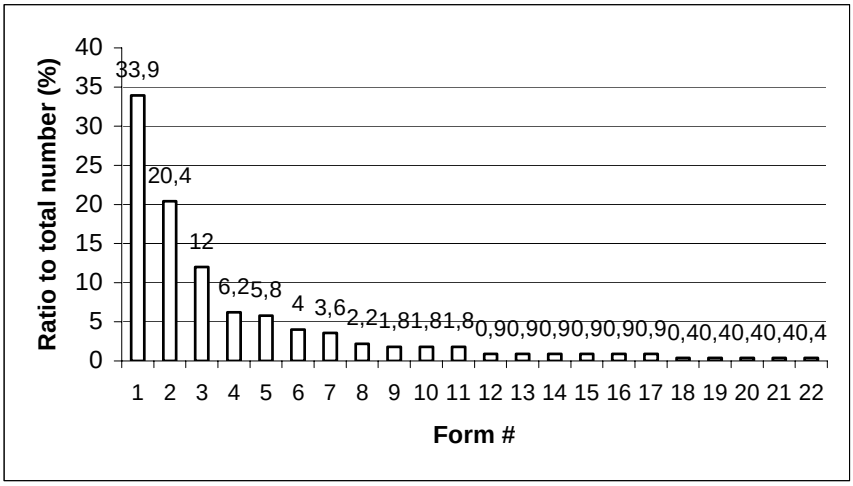


Figure 6.4. Frequency of the variant forms of the character 腹 in the Houma corpus.

33.9%) is much lower than in the case of the other two words (95.3% and 73.9%). Although form C1 was the most common form, the frequency of forms C2 and C3 together reached approximately that of form C1. Other forms occurred less frequently, some only once.

A closer look reveals, however, that the high number of distinct forms, 22 in total, is the result of the combination of the same eight basic components: 复, 肉, 彳, 止, 攵, 口, ⟌, 心. Figure 6.5 displays the frequency of these eight basic components within the 225 character forms used for the word *fu* "abdomen."

The first basic observation one can make is that the component 复 occurs in all 225 character forms. Since the component 复 remains unchanging, while other components are added or omitted with varying degree of frequency, one could regard 复 as an orthographic root. Because this root is phonetic in nature, sometimes it can serve as a root for other characters as well. For example, the component 复 also represents the orthographic root of the character forms used in the covenant texts for writing the word *fu* "to restore."

The second most common component in the above graph is 肉, present in 161 character forms, that is, in 71.6% of the cases. Our understanding, corroborated by Xu Shen, is that 肉 plays a semantic role in the character 腹, but it now appears that this semantic component could sometimes be omitted. However, whenever the ancient scribes omitted the semantic component 肉, they nearly always added another one. The word *fu* "abdomen" occurs in the covenant texts as a bare phonetic component (复) only

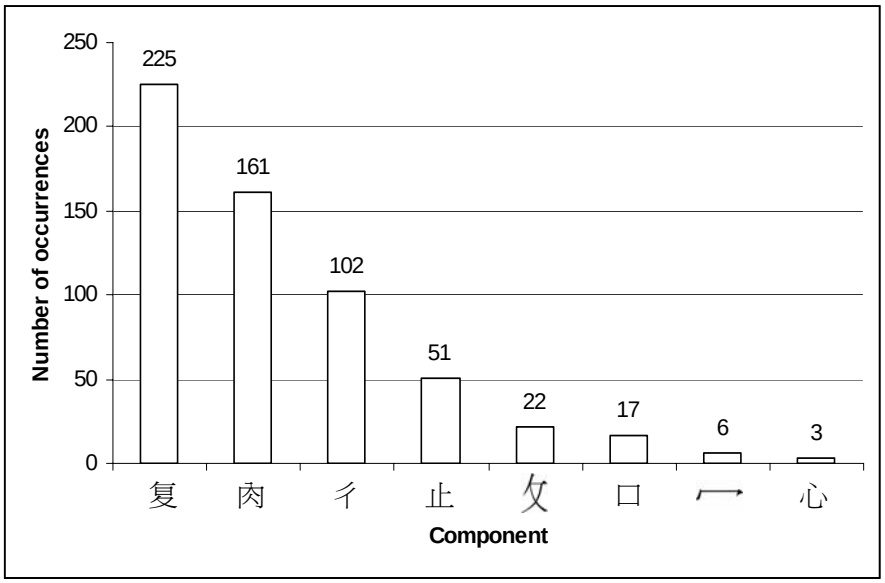


Figure 6.5. The frequency of components in variant forms of the character 腹 in the Houma corpus.

4 times (form C11), amounting to only 1.8% of all forms. In 98.8% of the forms, the ancient scribes augmented the phonetic 复 with at least one additional, presumably semantic, component.

6.3 Orthographic patterns

The above analysis of the three words from the Houma covenant texts permits the following observations:

Inconsistency of variability

The Jin scribes of the early 5th century BC wrote different words with a different degree of variability. They wrote some words with a fairly consistent structure, whereas others with a considerable amount of discrepancy. They wrote the word *zhi* "will, intention," for example, with a 95.3% consistency. The scribes wrote the most common form for the word *fu* "abdomen," on the other hand, with only 34% frequency. The reason why some words were written consistently, while others with numerous struc-

tural variants, would be a subject of another study. It is certainly not the case that character forms that appear in the modern script as multi-component compounds varied more than those consisting of only a single component. For example, the three examples analyzed in this chapter, all written in the modern script as two-element compounds, showed different degrees of variability.

Dominant form

The analysis of variant forms in this chapter reveals that there was generally a dominant character form, which represented the most common way of writing a word. The existence of a dominant form shows that the scribes of the Warring States period adhered to certain conventions regarding structure, but with a larger degree of tolerance than we are accustomed to. An interesting question is how the dominant form relates to the structure of the same character in the modern script. While it is reasonable to assume that the Qin-Han standardizers of the script, in their attempt to create a unified script, preferred to establish the dominant form as the standard and discard all other variants, this was not always the case. Among the three examples in this chapter, for instance, the structure of the dominant form of only two (*zhi* "will, intention" and *fu* "abdomen") coincides with their modern form.

Phonetic consistency

Among the variant forms, the components that varied were usually semantic in nature. The phonetic component remained the same in nearly every case, forming a phonetic root. In the case of the word *zhi* "will, intention," the phonetic root was the component 之, present in 99.6% of character forms. In the case of the words *fu* "to restore" and *fu* "abdomen," the phonetic component 复 was present in every single character form, that is, in 100% of the time. The retention of the phonetic root implies that in Warring States time, the primary aspect of a character form, that tied it most intimately to the word it stood for, was its sound value.

Semantic relevancy

Despite the seemingly haphazard variability of semantic components, there usually was a semantic connection between the varying component and the usage of a character form. In turn, this meant that components that were not related to the meaning of the word the character form stood for

were not part of its orthography. The frequency of character forms with semantically "irrelevant" components was low.[193] An important question with regard to the variability of semantic components would be how this variability influenced, or triggered, the development of *xingsheng* characters during this period.

The analysis presented in this chapter revealed the above patterns as the basic rules of structural variability. The analysis of the entire Houma corpus, including every character form would provide a more solid foundation for describing the orthographic patterns of Warring States scripts. Such an endeavor, however, is beyond the scope of this study.

[193] An exception to this pattern were cases of "graphical cross-influence," when the structure of a character form, by its mere presence, influenced the structure of another character form that was either adjacent or structurally similar to the first one.

CHAPTER SEVEN

CONCLUSIONS

The Chinese script was one of the most important inventions of Chinese civilization and one of the key elements in terms of which the Chinese people still define their national identity. Xu Shen wrote in the "Postface" of the *Shuowen*:

> 蓋文字者，經藝之本，王政之始。前人所以垂後，後人所以識古。
>
> For writing is the foundation of the classics and the arts, the beginning of royal government. It is the means by which people of the past reach posterity, by which people of the future know the past.

This quote sums up the significance of writing, showing that the study of Chinese writing is not just a means of learning about the past but is also important from the point of view of the national identity of the Chinese people.

In the previous chapters, I examined the variability of character forms in the Warring States period, the reasons why the existence of this phenomenon has not been commonly recognized, and the orthographic conventions of Warring States scribes. I present my major findings, their implications, and some further speculations below.

7.1 Findings

In the first part of this study, I demonstrated that the traditional understanding of the nature and evolution of the pre-Qin script was primarily based on the accounts of Xu Shen and Ban Gu, as well as, to a lesser degree, on sporadic comments of Sima Qian. However, these accounts reflected the not unbiased contemporary effort to establish a unified image of the empire and its history, in order to provide an ideological validation for the Han dynasty.

One of the main products of this politico-ideological program was the portrayal of an ideal stage prior to the Warring States period, allegedly referred to by no smaller a personality than Confucius himself, when scribes

would have rather left a blank space than written a non-standard character. Xu Shen and Ban Gu claimed that this ideal stage of writing had been lost during the Warring States period and that the seven states had begun to write in local scripts. According to these two sources, when Li Si and the First Emperor established the modified Qin script as the official script of the empire, they in fact were re-establishing the ideal state of writing that existed before the turmoil of the Warring States period. Such a description was part of the larger Han concept of history which saw the role of the Han as restoring order 治 by wiping out disorder 亂, that is, re-implementing the central power of the Zhou that had allegedly preceded the chaos of the Warring States period. Therefore, Xu Shen and Ban Gu did not portray the enforcement of the orthographic standard as the creation of a new, formerly unknown, concept but as the reinstatement of an initial, ethically superior, state of writing. Xu Shen and Ban Gu, as well as their later interpreters, accordingly regarded orthographic variability as a deviation from the correct state of writing.

It is worth mentioning here that Xu Shen and Ban Gu were not the only ones responsible for the skewed image of the Qin unification of writing. Scholars of the Han and later dynasties played just as important a role in this respect when they chose the texts that were to be transmitted. Historiography has routinely regarded the Qin burning of books as a heinous act against the literary output of the contemporary China, ignoring the fact that the quiet but continuous process of selecting works for copying by the scholars of dynastic China was perhaps an even more influential factor in textual transmission, which in turn determined later views of history.

Xu Shen and Ban Gu attributed the pivotal event of re-establishing the uncorrupt way of writing to the Qin reforms which immediately followed the unification of the empire in 221 BC. The archaeological evidence, however, reveals that the character forms from shortly after this date still exhibited a significant degree of variability. Moreover, orthographic variability was present not only on silk and bamboo manuscripts, commonly categorized as examples of popular writing, but also on the bronze and pottery edict plates distributed by the government. Although these edict plates were official documents, the structure of character forms on them was not always uniform, showing that the reforms were either not comprehensive or not successful. In either case, Chinese writing did not become unified during the reign of the First Emperor. The archaeological material suggests that the standardization of the Chinese script was a gradual, and not always unidirectional, process lasting several centuries.

Returning to the writing practices of the Warring States period, I found that the scribes did not always write in an orthographically consistent

way. Indeed, variability was an integral part of Warring States writing. Distinguishing the writing habits of a single scribe from those of an entire community, I demonstrated that the writers and readers, whether humans or spirits, as a group tolerated a considerable degree of structural fluidity. Although on the level of a single manuscript, the product of one scribe's hands, the degree of variability was somewhat lower, variant "spellings" of the same character still existed. However, in my examination of the problem, I chose to concentrate on an entire community, not individual scribes – just as a single person's speech cannot adequately represent a whole language or dialect, a single scribe's hand cannot represent the entire local script.

On the level of the community, represented by an entire corpus of excavated material (e.g. Guodian corpus, Houma corpus), the variability of character forms exhibited specific patterns. One such pattern was the presence of a dominant form. This dominant form often coincided structurally with the subsequent clerical form of the same character, suggesting that the Qin-Han standardizers of the script habitually adopted the most common Warring States form to eventually develop the concept of a "correct" form.

Scribes used less common forms with decreasing frequency. Because some forms were used only once or twice in a large corpus, we can conclude that the scribes may have "invented" some orthographic configurations on the spot, never to use them again. Such a spontaneous way of writing could easily have obscured the meaning of the character, had it not been for the phonetic core that the scribes left unaltered in nearly every case.

Therefore the variability of character forms generally meant the variability of the semantic components. The phonetic component usually did not change. To be exact, it was the sound value of the phonetic component that remained the same, since the scribes sometimes substituted the phonetic component for another, homophonous or nearly homophonous, component. From the perspective of the modern script, but only from this perspective, these instances were cases of phonetic borrowing.

My opinion is that orthographic variability in pre-Qin scripts was the prerequisite for the evolution of the script. Dominant forms evolved gradually, simply because the scribes favored some newer "spelling" over an older one, eventually transforming the new form into the dominant one. There was no need to change the script by way of reforms because the writing habits were flexible enough to let the script evolve in the direction that facilitated communication. The creation of an orthographic standard also meant the "freezing" of the then current script, a measure that prevented further evolution.

7.2 Implications

The recognition of orthographic variability in Warring States manuscripts has concrete consequences for the work of archaeologists and historians. First, one cannot use isolated examples of orthographic forms by themselves for dating texts and objects with certainty. While researchers commonly associate certain orthographic forms with a particular time period, it must be understood that a single example of a character form could easily be one of the less common, perhaps even atypical, orthographic variants. Unless one looks at a larger pool of samples that tie the dominant form to a specific period, one cannot gain a reliable picture of orthographic conventions of a particular time period and region.

Similarly, one cannot use isolated examples of orthographic forms by themselves to determine the geographic origin of manuscripts and artifacts. One form could be the usual form in one region but it could still occur, even if less frequently, in other regions.

Furthermore, the occurrence of dissimilar orthographic forms within the same document does not necessarily have any bearing on the authenticity of a manuscript. While some scribes might have written with a greater degree of orthographic consistency than others, the fact that a scribe used variant forms within the same document would simply have been part of his writing habit. Neither variant forms nor scribal errors necessarily bear on the authenticity of a document.

Beside the above practical points, the lack of a unified orthographic system in Warring States writing inevitably raises the question of how writing prior to the Warring States period had evolved. Were there "ideal" forms in the early period, when writing was less widespread and was there an early writing system that was "unified"? In other words, was the model one of a unified writing system fragmenting into regional and even individual idiosyncrasies during the Warring States period, but then gradually coming back to a standardized orthographic system? Or was it one of local option from the very beginning with the eventual standardization representing not a return to some original ideal, as Xu Shen and Ban Gu implied, but an evolution consonant with the general move towards the political and cultural unification of China?

Since the Chinese script originated from a single cultural entity, the Shang and later the Zhou, it might be argued that there had been a sort of initial unity which eventually disappeared, and the scribes of the various regions of the Warring States period gradually broke free from the dominance of the Zhou culture. However, this model would be true only for

the stylistic characteristics of the script, not its orthography. The orthography of Shang and Western Zhou inscriptions was similarly inconsistent. Although it might be possible to ascribe the various "spellings" of certain words to different diviner or scribe groups, from the perspective of a written dialect, the Shang and Zhou scripts were orthographically not standardized.

Accordingly, I am of the opinion that the Qin-Han standardization represented not a return to an original state of orthographic uniformity but the initial, even if not immediately successful, standardization of earlier writing practices. Just as the Qin did not "re-unify" China (which did not even exist before the fusion of the seven major states into a single military and administrative entity), they did not re-standardize writing either. Instead, for the first time in history, they united the various states on the territory of later China and began the standardization of writing. However, since this study does not address the issue whether the Shang and Western Zhou periods possessed an orthographically uniform writing system, this model is still a hypothesis awaiting verification.

One cannot fail to notice, however, that this view of a once existing, then lost and later restored, state of orthographic integrity parallels the traditional view of history according to which the Xia, Shang, and Zhou dynasties followed each other in implementing the Heavenly Mandate in the world. Once the rulers of one dynasty lost their Virtue, Heaven conferred the Mandate on a new dynasty that overthrew the corrupt system. The victory of the new dynasty was seen not just as an expansion of dominance but as the victory of good over evil, the reinstatement of order in the subcelestial world. Bagley describes the sharp contrast between this worldview and the picture emerging from the archaeological evidence:

> The civilized world on the eve of the Zhou conquest was large, diverse, and intricately interconnected. In the light of archaeology, therefore, the most striking feature of traditional history is the absence from it of any such world. Transmitted texts present us instead with an ancient China in which the only civilized powers were Zhou and Shang, and with an ancient history in which the principal event was the transfer of rule from one to the other. Ever since the Eastern Zhou period the Zhou conquest has been viewed as an event of towering significance, not because of anything tangible connected with it, such as a building project or a reform of script or a standardization of weight and measures, but because it provided a model for the morally correct transfer of power and for the maintenance of power through dynastic virtue. In that model a unified political order coextensive with civilization was ruled by the Shang until their rule grew oppressive, whereupon the Shang were replaced by the Zhou. This is a distinctly schematic account of the past, one that left us quite unprepared for archaeology's discovery of a wider civilized world, and

> if we are to understand its emphases and omissions, we must begin by reminding ourselves that the tradition in which the Zhou figure so centrally is a Zhou creation.[194]

The traditional view of history thus is gradually changing as new archaeological evidence comes to light. We began to realize that the Shang at Anyang were neither the central nor the dominant civilization on the territory of modern China and that there were other thriving civilizations not mentioned in historical sources. Yet the traditional view of the development of the Chinese script has generally remained.

On the one hand, there is the distinctly phonological approach based on the principles of Western linguistics which attempts to reconstruct the original sounds of the language. The problem is that we cannot be sure which language we are trying reconstruct. In many cases, the analyses are performed in an abstract linguistic space based on the modern and small seal script structure of characters, with occasional, but random, references to bronze and oracle-bone forms.

In contrast with this approach, there is the graphical approach, based on the traditional discipline of palaeography, which tries to determine the origins of characters based on their modular structure. Scholars belonging to this group still cannot liberate themselves from the influence of the principles of the *liushu* which, as I have already demonstrated, never meant to describe the evolution of characters. The fundamental problem with the palaeographic approach is its effort to match archaic character forms with modern equivalents, simply treating the old forms as variant ways of writing post-Qin characters. These scholars assume that the changes in character forms were usually due either to "technical changes" effected as a means of facilitating the writing process, or "corruption."

7.3 Further speculations

The following section is a loose string of hypothetical speculations which I include only as future directions for research.

An intriguing topic with respect to the early development of Chinese writing is whether the Zhou and Warring States scripts were direct descendants of the Shang script. I am under the impression that this is the general view today. Yet the Shang and Zhou could have been completely

[194] Bagley (1999), p. 230.

distinct peoples with their own culture and, in all likelihood, language. Although there are arguments as to whether the Zhou homeland was in the Wei River Valley or farther east, without conclusive evidence we cannot automatically assume that the two cultures were linguistically related to each other. Archaeological evidence shows that the Zhou began using the Shang script even before the conquest.

From the end of the Spring and Autumn period we see the development of regional scripts, a process in the course of which linguistically distinct peoples used and modified the Zhou script. Many of these peoples, such as the Qin originating from Northwest China and the Yue of Southeast China, were ethnically, culturally, and historically diverse groups that most likely spoke different languages. Yet all these peoples used the same Zhou script, even if some of them eventually modified it. When the Qin, for example, began using the Zhou writing, they produced bronze inscriptions with no discernible differences from the Zhou inscriptions.

During the millennium that separated the earliest examples of oracle-bone inscriptions from the end of the Warring States period, the script had not changed significantly. Indeed, there are greater differences between the small seal and the clerical script than between the Qin small seal and Shang bronze inscription scripts. The process of evolution between the Shang and the Qin mainly involved calligraphic modifications and the development of *xingsheng* characters.

Both traditional and modern scholars have habitually attributed the development of *xingsheng* characters to societal changes, including the appearance of new tools and concepts, as well as the necessity to describe reality more precisely. The problem with such explanations is obvious. Although the changes of society might introduce some new concepts which require new words and perhaps new characters, I see no justification for the large-scale increase of *xingsheng* characters. While the proportion of *xingsheng* characters to the total number of characters in oracle-bone inscriptions was small, by the time of Xu Shen, this ratio has grown to an overwhelming majority.[195]

As for the theories citing the need for a more precise description of reality, they suggest that during the late Shang the script lacked clarity and precision which is obviously a projection of the simplistic concept of "historical progress" moving from primitive to more advanced societies. My contention is that during the late Shang, the Shang script was just as precise and unambiguous to the Shang people as the Zhou script to the Zhou people. Writing does not describe reality, words do. The function of

[195] Qiu (2000), p. 52.

writing is to record language. To assume that at any given moment in a people's history their language was not adequate to describe their world reveals a basic misunderstanding about the nature and function of language. Surely, every language undergoes continuous changes as the outside world changes. The users of the language integrate new words for new concepts and abandon obsolete ones. These changes are a natural part of linguistic evolution and happen on a daily basis. Whether some changes become widespread enough to remain permanently in the language is yet another matter. However, all these changes happen in the realm of the spoken language. There is no need to modify existing characters to record existing words better.[196]

Modification of the script becomes necessary when the script begins to lose its phonetic connection with spoken language. Since Chinese writing, even in its purportedly pictographic stage, was to a significant degree phonetic, the increasing gap between a word's pronunciation and its graphic representation probably created a problem. As long as a character was used in its primary graphical sense, the change in pronunciation did not pose a serious problem, but when the character became used to record an originally homophonous, or nearly homophonous, word, the disappearance of the phonetic similarity between the two words impeded comprehension.

I see the reason for the increase of *xingsheng* characters in the increasing differences between the spoken languages of the peoples that adopted the Shang and Zhou scripts. In order to achieve this, they had to adjust the script to fit their own language with its inherent phonetic correlations. The *xingsheng* characters could have arose from the practice of using signifiers to re-establish the bond between the script and pronunciation.

The above points are only speculations and hypotheses arisen in the process of working on this study. Just as this entire study, they argue against a linear line of evolution in the history of Chinese writing and bring into question the national identity of the "Chinese" script. The demonstration and evaluation of these hypotheses will be a subject of future research.

[196] Naturally, there was a change in the function of writing from oracle-bone inscriptions to Zhou bronze inscriptions, accompanied by the change of the audience. Still, I think that the linguistic bearing of audience falls far behind that of the adoption of the script by a new people.

APPENDIX I.

COMPARISON OF THE *HANSHU* AND *SHUOWEN* ACCOUNTS

This section aligns the text of the *Shuowen* "Postface" with the explanation to the *Xiaoxue* section of the "Yiwenzhi" chapter of the *Hanshu*. The order of the text follows that of the *Shuowen*, without assuming the priority of either version. Because the two versions differ in arrangement, to be able to align the two texts, I had to break up the original structure of the *Hanshu* account.

For the sake of clarity and convenience, I divided the texts into smaller sections. Parallel texts of the two versions appear in the table in the same row. Whenever the *Shuowen* has no parallel section in the *Hanshu*, the right column remains empty, and vice versa. The few *Hanshu* sentences that are not in the *Shuowen* are omitted.

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
01	古者庖羲氏之王天下也，仰則觀象於天，俯則觀法於地，視鳥獸之文與地之宜，近取諸身，遠取諸物；	
	In ancient times when Pao Xi ruled the country, he looked up to observe the signs of Heaven, he looked down to observe the patterns on Earth. He saw how the [visual] designs of the birds and beasts corresponded with those on the ground. Near, he borrowed [designs] from himself; afar, he borrowed from other objects:	
02	於是始作《易》八卦，以垂憲象。及神農氏，結繩為治，而統其事。庶業其繁，飾偽萌生。黃帝史官倉頡，見鳥獸蹄迒之跡，知分理之可相別異也，初造書契。百	《易》曰：「上古結繩以治，後世聖人易之以書契，百官以治，萬民以察，蓋取諸夬．」「夬，揚於王庭」，言其宣揚於王者朝廷，其用最大也．

	From the *Shuowen* “Postface” 說文解字《敘》	From the *Hanshu* “Yiwenzhi” 漢書《藝文志》
	工以乂，萬品以察，蓋取諸夬。「夬，揚於王庭」，言文者，宣教明化於王者朝庭，「君子所以施祿及下，居德則忌」也。	
	this is how he began composing the eight trigrams of the *Yijing*, with the aim of passing on these patterns and signs. In the time of Shen Nong, people organized their affairs by using knots on threads as a means of creating order. As the many professions flourished, embellishment and fabrication arose too. Cang Jie, the historian of the Yellow Emperor, saw the foot and paw prints of the birds and beasts on the ground and understood that these patterns could be distinguished from each other. Thus for the first time [in history], he created writing. [As a result], the hundred craftsmen could be regulated and the ten thousand categories [of things] could be inspected. [Cang Jie] probably borrowed from the *guai* [hexagram]. [The *Yijing* says:] “*Guai*: to display in the royal court.” This means that writing displays the teachings and elucidates the transformations at the royal court. “This is the means by which the gentleman bestows benefits to his subordinates. If one abides in virtue, he will be cautious.”	The *Yijing* says: “In ancient times, people used knots on threads as a means of creating order. Later, the sage changed this by creating writing. [As a result], the hundred officers could be regulated and the ten thousand people could be inspected. [The sage] probably borrowed from the *guai* [hexagram].” “*Guai*: to display in the royal court.” This means that [writing] is displayed in the royal court. [Therefore], it has the greatest use of all.
03	倉頡之初作書也，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，言孳乳而寖多也。著於竹帛謂之書。書者，如也。以迄五帝三王之世，改易殊體，封於泰山者七十有二代，靡有同焉。	

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	When Cang Jie first invented writing, he created graphic forms (*xiangxing*) according to categories; therefore these were called *wen* (patterns). After that, forms and sounds (*xing-sheng*) mutually augmented each other; these were called *zi*. The *wen* were the origin of the depiction of objects; the *zi* increased in number as the vocabulary grew. When [characters] were written on bamboo and silk, they were called books. A book is a portrayal. By the time of the Five Sovereigns and Three Kings, the forms of [writing] have undergone great changes. Of [the writings left behind by] the 72 generations of rulers who performed sacrifices at the Mount Tai, not two are alike.	
	《周禮》：八歲入小學，保氏教國子，先以六書。一曰指事。指事者，視而可識，察而見意，「上、下」是也。二曰象形。象形者，畫成其物，隨體詰詘，「日、月」是也。三曰形聲。形聲者，以事為名，取譬相成，「江、河」是也。四曰會意。會意者，比類合誼，以見指撝，「武、信」是也。五曰轉注。轉注者，建類一首，同意相受，「考、老」是也。六曰假借。假借者，本無其事，依聲託事，「令、長」是也。	古者八歲入小學，故周官保氏掌養國子，教之六書，謂象形、象事、象意、象聲、轉注、假借，造字之本也.
	The *Zhouli* says, "[children] began studying grammatology at the age of eight. Teacher Bao taught the sons of the state, starting with the *liushu* [categories.]" The first is *zhishi*. *Zhi-shi* characters are the ones that can be understood by looking at them, the meaning of which can be seen through	In ancient times, [children] began studying grammatology at the age of eight. Thus Teacher Bao, an official of the Zhou house, was in charge of educating the sons of the state, teaching them the *liushu* [categories:] the *xiangxing*, the *xiangshi*, the *xiangyi*, the *xiang-*

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	observation. The characters 上 and 下 are like this. The second is *xiangxing*. *Xiangxing* characters are the ones that depict objects by reproducing their physical shape. The characters 日 and 月 are like this. The third is *xingsheng*. *Xingsheng* characters are the ones that take a [semantic] subject to indicate the word and combine it with a [phonetic] semblance. The characters 江 and 河 are like this. The fourth is *huiyi*. *Huiyi* characters are the ones that conjoin categories to present the indicated meaning. The characters 武 and 信 are like this. The fifth is *zhuanzhu*. *Zhuanzhu* characters are the ones that establish categories based on a single origin and which borrow their analogous meanings from each other. The characters 考 and 老 are like this. The sixth is *jiajie*. *Jiajie* characters are the ones that are assigned a written form, which did not exist originally, based on their pronunciation. The characters 令 and 長 are like this.	*sheng*, the *zhuanzhu,* and the *jiajie*. These are the foundation of generating characters.
04	及宣王太史籀，著《大篆》十五篇，與古文或異。至孔子書六經，左丘明述春秋傳，皆以古文，厥意可得而說也。	《史籀篇》者，周時史官教學童書也，與孔氏壁中古文異體.
	In the time of King Xuan (827/25–782 BC),[197] Zhou, the Grand Historian composed the *Dazhuan* in fifteen chapters, [the character forms of which] were somewhat different from the *guwen* script. When Confucius wrote the Six Classics and Zuo Qiuming recorded his commentary on the *Chunqiu*, they both used the	The *Shizhou pian* is a book written in Zhou times by the Historian for teaching students. The character forms in it are different from the *guwen* characters in Confucius's Books from the Wall.

[197] *Cambridge History of Ancient China*, p. 25.

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	guwen script, [thus] their meaning can be understood and explained.	
05	其後諸侯力政，不統於王。惡禮樂之害己，而皆去其典籍。分為七國，田疇異畝，車涂異軌，律令異法，衣冠異制，言語異聲，文字異形。	
	After that, the feudal lords grew in power and did not obey the [Zhou] king. Disliking the harm done to themselves through the rules of etiquette and through the music, they did away with the statutes and records of these.[198] The divided and formed the seven states which had cultivated fields of different sizes, carriage axles of different width, regulations and ordinances of different legal codes, clothes and caps of different systems, languages that sounded different, and written characters that looked different.	
06	秦始皇帝初兼天下，丞相李斯乃奏同之，罷其不與秦文合者。斯作《倉頡篇》。中車府令趙高作《爰歷篇》。大史令胡毋敬作《博學篇》。皆取史籀《大篆》，或頗省改，所謂小篆也。	《蒼頡》七章者，秦丞相李斯所作也；《爰歷》六章者，車府令趙高所作也；《博學》七章者，太史令胡母敬所作也. 文字多取《史籀篇》，而篆體復頗異，所謂秦篆者也.
	When for the first time [in history] the First Emperor of Qin united the subcelestial world, Li Si, his Grand Councillor presented a proposal to unify them (i.e. the scripts or character forms) and discard what was not	The seven chapters of the *Cangjie* were composed by Li Si, the Grand Councillor of the Qin. The six chapters of the *Yuanli* were composed by Zhao Gao, the Keeper of Carriages. The seven chapters of the *Boxue*

[198] "Disliking the harm…" Here I rely on Bodde's interpretation. Bodde (1938, pp. 149, 162–163) links this phrase to *Mengzi* Vb, 2., where Mencius says, "Their particulars cannot be learned, for the feudal lords, disliking their injuriousness to themselves, have all done away with the records of them."

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	in accord with the Qin script. Li Si wrote the *Cangjie pian*, Zhao Gao, the Keeper of Carriages, wrote the *Yuanli pian*, and Humu Jing, the Grand Historian wrote the *Boxue pian*. All of them borrowed [characters] from Historian Zhou's *Dazhuan*, although they sometimes altered and abbreviated those. This was what they call the small seal script.	were composed by Humu Jing, the Grand Historian. The characters [in these compendia] are all borrowed from the *Shizhou pian*, although the seal character forms were also somewhat different. This was what they call the Qin seal script.
07	是時，秦燒滅經書，滌除舊典。大發吏卒，興戍役。官獄職務繁，初有隸書，以趣約易，而古文由此絕矣。	是時始造隸書矣，起於官獄多事，苟趨省易，施之於徒隸也.
	At this time, the Qin burned the classics and destroyed the old records. They made great levies of troops and raised armies, and the work of the official clerks grew complex. The clerical script first appeared out of the need [to write in a] simple and easy way. From this time on, the *guwen* script was discontinued.	At this time, for the first time the clerical script was created. [The clerical script] arose because the official clerks' work increased and they, out of carelessness and haste, began abbreviating and simplifying [the characters] when applying them for convicts.[199]
08	自爾秦書有八體：一曰大篆，二曰小篆，三曰刻符，四曰蟲書，五曰摹印，六曰署書，七曰殳書，八曰隸書。漢興有草書。	
	From this time on, the Qin script had eight styles: the first one was the great seal script; the second the small seal script; the third the tally carving script; the fourth the insect script; the fifth the seal stamping script; the sixth the title script; the seventh the spear inscription script;	

[199] The word *tuli* 徒隸 means labor convicts, therefore according to Ban Gu the character 隸 used in the term *lishu* ("clerical script") should be translated as "punitive, penal" and not as "clerical." For the sake of consistency, however, I stick to the translation "clerical script."

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	the eighth the clerical script. From the rise of the Han there was the draft script.	
09	尉律：學僮十七以上始試。諷籀書九千字，乃得為史。又以八體試之。郡移太史並課。最者，以為尚書史。書或不正，輒舉劾之。	漢興，蕭何草律，亦著其法，曰：「太史試學童，能諷書九千字以上，乃得為史．又以六體試之，課最者以為尚書御史史書令史．吏民上書，字或不正，輒舉劾．」
	According to the regulations of official, students start their examinations at the age of 17. Once they memorize the 9,000 characters of the [Historian] Zhou's script, they can become historians. Their test also includes the eight [calligraphic] styles. The prefects send the [examinees] to the Grand Historian for collective examinations. The best of them become Clerks of the Imperial Secretariat. If anyone writes irregularly, he is to be recommended for punishment immediately.	At the rise of the Han, Xiao He (d. 193 BC) drafted the regulations. He also composed their methods [of applications]: "The Grand Historian tests the students. Those who can memorize over 9,000 characters can become historians. Their text also includes the six [calligraphic] styles. The best of them will become Censors of the Imperial Secretariat and Calligraphy Clerks. If an official submits a document with irregular characters, he is to be punished immediately."
10	今雖有尉律，不課，小學不修，莫達其說久矣。	
	Today, although we have the regulations of the officials, they are not enforced; students do not exercise themselves in the art of grammatology; there has not been anyone who understood its teachings for a long time.	
11	孝宣皇帝時，召通《倉頡》讀者，張敞從受之。涼州刺史杜業，沛人爰禮，講學大夫秦近，亦能言之。	《蒼頡》多古字，俗師失其讀，宣帝時徵齊人能正讀者，張敞從受之，傳至外孫之子杜林，為作訓故，并列焉．
	At the time of Emperor Xiaoxuan (r. 73–49 BC), [the court] summoned the interpreters of the *Cangjie* and	The *Cangjie pian* had many ancient characters and ordinary teachers were not able to read them. At the

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	Zhang Chang learned [to interpret the book] from them. Du Ye, the Regional Inspector of Liangzhou, Yuan Li, a man from Pei, and Qin Jin, the Grand Lecturer could also explicate it.	time of Emperor Xuan (r. 73–49 BC), [the court] summoned people from the state of Qi who could read [the characters] correctly. Zhang Chang learned from them and passed [the knowledge] down to his great-grandsom Du Lin for interpretation and comparison.
12	孝平皇帝時，徵禮等百餘人，令說文字未央廷中，以禮為小學元士。黃門侍郎揚雄，采以作《訓纂篇》。	至元始中，徵天下通小學者以百數，各令記字於庭中．揚雄取其有用者以作《訓纂篇》，順續《蒼頡》，又易《蒼頡》中重復之字，凡八十九章.
	At the time of Emperor Xiaoping, Li and over a hundred other scholars were summoned and ordered to explain the graphs [of the *Cangjie*] at the court of the Weiyang Palace. Li was made the Senior Serviceman of Grammatology. Yang Xiong, the Gentleman Attendant at the Palace Gate, collected [the interpretations] and compiled the *Xunzuan pian*.	During the Yuanshi era (86–80 BC), hundreds of experts on grammatology were summoned and ordered to record the characters [they knew] at the court. Yang Xiong selected those that were useful and compiled the *Xunzuan pian*, [intended as] a continuation of the *Cangjie*. He also amended the duplicate characters in the *Cangjie*. [These two works amounted to] a total of eighty-nine chapters.
13	凡《倉頡》以下十四篇，凡五千三百四十字，群書所載，略存之矣。	漢書[200]，閭里書師合《蒼頡》、《爰歷》、《博學》三篇，斷六十字以為一章，凡五十五章，并為《蒼頡篇》.
	From this time on, there were fourteen sections in the *Cangjie*, a total of 5,340 characters. Nearly all of the graphs recorded in various versions [of the *Cangjie*] were now preserved.	With the rise of the Han, the teachers of the villages combined the *Cangjie*, the *Yuanli*, and the *Boxue* and divided them into sixty-character chapters. This consolidated *Cangjie pian* totaled fifty-five chapters.

[200] According to the commentators, the original character 書 here is a mistake which stands for the character 興.

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
14	及亡新居攝，使大司空甄豐等校文書之部。自以為應制作，頗改定古文。	
	When the now deposed Xin dynasty usurped the throne, Zhen Feng, the Grand Minister of Works and others were ordered to revise the various categories of writings. They believed that it was neccessary to systematize and organize, [therefore] they amended and regularized the *guwen* script somewhat.	
15	時有六書：一曰古文，孔子壁中書也。二曰奇字，即古文而異也。三曰篆書，即小篆。四曰左書，即秦隸書。秦始皇帝使下杜人程邈所作也。五曰繆篆，所以摹印也。六曰鳥蟲書，所以書幡信也。	六體者，古文、奇字、篆書、隸書、繆篆、蟲書，皆所以通知古今文字，摹印章，書幡信也。[201] 與孔氏壁中古文異體[202]
	At this time, there were six scripts. The first one was the *guwen* script, which was the script on the documents found in the wall of the house of Confucius; the second the odd characters, i.e. the unusual *guwen* forms; the third the seal script, i.e. the small seal script; the third the attendant script, i.e. the Qin clerical script created by Cheng Mao from Xiadu under the commission of the First Emperor of Qin; the fifth the pseudo-seal script used for seal carving; the sixth the bird and insect script used to write on banners.	The six styles were the *guwen* script, the odd characters, the seal script, the clerical script, the pseudo-seal script, and the bird and insect script. By means of these one was able to fully master the ancient and modern characters, to make seal carvings and write on banners.

[201] This section is a duplicate of H08. Although it appears only once in the original text, the structure of the parallel text of the *Shuowen* necessiated its repetition in the table.

[202] This phrase is a duplicate from H04.

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
16	壁中書者，魯恭王壞孔子宅，而得《禮記》、《尚書》、《春秋》、《論語》、《孝經》。又北平侯張蒼獻《春秋左氏傳》。郡國亦往往於山川得鼎彝，其銘即前代之古文，皆自相似。雖叵復見遠流，其詳可得略說也。	
	The Books from the Wall were the documents that King Gong of Lu found when he was knocking down the mansion of Confucius. These were the *Shiji*, the *Shangshu*, the *Chunqiu*, the *Lunyu*, and the *Xiaojing*. Other than this, Marquis Zhang Cang of Beiping presented the *Chunqiu Zuozhuan*. In different provinces and states [the people] also often found various types of vessels in the mountains and rivers; the inscriptions on these were the ancient script of former ages. All [the characters on these inscriptions] were similar to each other. Although we cannot see anymore their original state, we can still interpret their features.	
17	而世人大共非訾，以為好奇者也，故詭更正文，鄉壁虛造不可知之書，變亂常行，以耀於世。諸生競逐說字，解經誼，稱秦之隸書為倉頡時書，云：「父子相傳，何得改易！」乃猥曰：「馬頭人為長，人持十為斗，虫者，屈中也。」 廷尉說律，至以字斷法：「苛人受錢，苛之字止句也。」若此者甚眾，皆不合孔氏古文，謬於《史籀》。俗儒鄙夫，翫其所習，蔽所希聞。不見通學，未嘗睹字例之條。怪舊埶而善野言，	

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	以其所知為秘妙，究洞聖人之微恉。又見《倉頡篇》中「幼子承詔」，因曰：「古帝之所作也，其辭有神僊之術焉。」其迷誤不論，豈不悖哉！	
	However, ordinary people all disagree [with the result of the work of these men], thinking that they are fond of oddities and purposefully falsify the correct characters. [People believe that these scholars] deceitfully created some unrecognizable characters in front of a village wall, bringing disarray onto normal customs, all for the sake of shining in the eyes of the world. Many scholars compete in explaining the characters and elucidating the meaning of the classics, calling the Qin clerical script the script of Cang Jie's time. They say, "[these characters] have been handed down from fathers to sons, how could they change!" Then they ignorantly say, "a man 人 with a horse's 馬 head makes up the character for 'leader' 長. A man 人 holding a cross 十 makes up the character for 'peck' 斗. The character for 'worm' 虫 consists of the crooked character 'center' 中. Even when the Chamberlains for Law Enforcement interpret the legal code, they make decisions based on the shape of characters: "[The code says to] condemn a man for accepting bribes. The [clerical form of the] character for 'condemning' 苛 consists of stopping 止 him and ceising 句 [his assets.]" Explanations like this are in abundance and none of them conforms with the *guwen* characters from Confucius's home or the *Shizhou*. These mediocre scholars and boorish	

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	people are mislead by what they are used to and blinded by what they have scarcely heard of. They have not encountered comprehensive scholarship and have never seen the classification of character categories. They rebuke old principles and admire wild beliefs, thinking that what they know is esoteric and mysterious and that they thoroughly comprehend the subtle meaning of the sage. Moreover, when they see the words "the youngest son received the mandate" in the *Cangjie pian*, they say: "This is written by an ancient emperor. His words incorporate the art of immortality." These people do not understand their own confusion. How absurd!	
18	書曰：「予欲觀古人之象。」言必遵修舊文而不穿鑿。	
	The *Shangshu* says, "I wish to observe the images of the ancients." This means that one must revere the old text and not go into strained explanations."	
19	孔子曰：「吾猶及史之闕文，今亡矣夫。」蓋非其不知而不問。人用己私，是非無正，巧說邪辭，使天下學者疑。	古制，書必同文，不知則闕，問諸故老，至於衰世，是非無正，人用其私．故孔子曰：「吾猶及史之闕文也，今亡矣夫！」蓋傷其[illegible]不正．
	Confucius said, "I still go back to the times when the historian left empty spaces in the text. Today there is no such thing anymore!" This is a condemnation of not asking when one does not know. When people write according to their own fancy, right and wrong have no constancy, clever teachings and wicked words create doubts among the scholars of the world.	In the old system, documents had to be written with identical characters. If you did not know something, then you left an empty space and asked the elders. Once the world went into decline, right and wrong had no constancy, and people wrote according to their fancy. This is why Confucius said, "I still go back to the times when the historian left

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
		empty spaces in the text. Today there is no such thing anymore!" He was upset because of the inundation [of writing] with irregularities.
20	蓋文字者，經藝之本，王政之始。前人所以垂後，後人所以識古。故曰：「本立而道生。」知天下之至賾而不可亂也。今敘篆文，合以古籀；博采通人，至於小大；信而有證，稽譔其說。 將以理群類，解謬誤，曉學者，達神恉。分別部居，不相雜廁也。萬物咸睹，靡不兼載。厥誼不昭，爰明以喻。其稱《易》孟氏、《書》孔氏、《詩》毛氏、《禮》周官、《春秋》左氏、《論語》、孝經，皆古文也。其於所不知，蓋闕如也。	
	For writing is the foundation of the classics and the arts, the beginning of royal government. It is the means by which people of the past reach posterity, by which people of the future know the past. This is why they say: "When the foundation is erected, the Way is born." If you know the most profound things in the world, you cannot get confused anymore. Now I arranged the [dictionary] based on the seal forms, and supplemented these with their *guwen* and *zhouwen* forms. I widely employed the opinions of experts, both minor or great. As long as they were credible and were based on sufficient proof, I incorporated their views. My purpose was to organize the various categories [of characters], to unravel mistakes, to enlighten scholars and make them understand the divine meaning	

	From the *Shuowen* "Postface" 說文解字《敘》	From the *Hanshu* "Yiwenzhi" 漢書《藝文志》
	[of characters. I organized [the characters] into groups, so they will not be confused with each other. All ten thousand things are here, I did not omit anything. If a certain principle is not obvious, I explain it with an example. I used the Meng[203] version of the *Yijing*, the Kong[204] version of the *Shangshu*, the Mao version of the *Shijing*, the Zhouguan version of the *Li*, the Zuo version of the *Chunqiu*, as well as the *guwen* versions of the *Lunyu* and the *Xiaojing*. If I did not know something, "I did not offer any opinion."[205]	

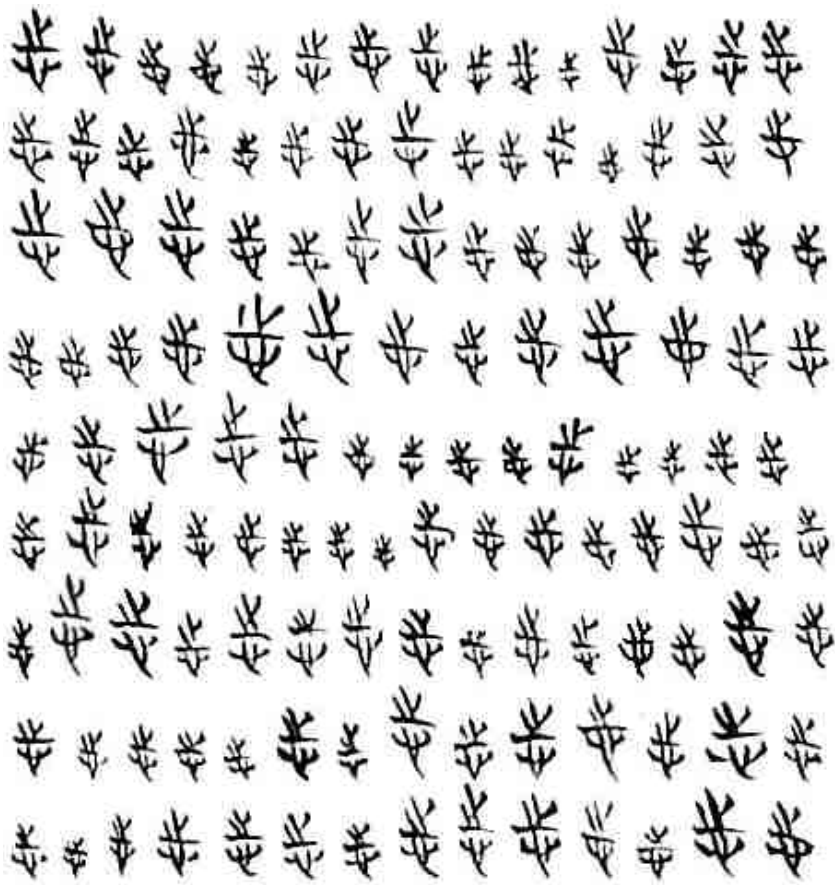
[203] Meng Xi's 孟喜 dates are unknown. He was the founder of the *jinwen* 今文 school of the *Yijing* 易經 during the Western Han.
[204] Kong Anguo (ca. 156 BC–ca. 74 BC).
[205] A quote from *Lunyu* 13:3: "Where a gentleman is ignorant, one would expect him not to offer any opinion." Lau (1979), p. 118.

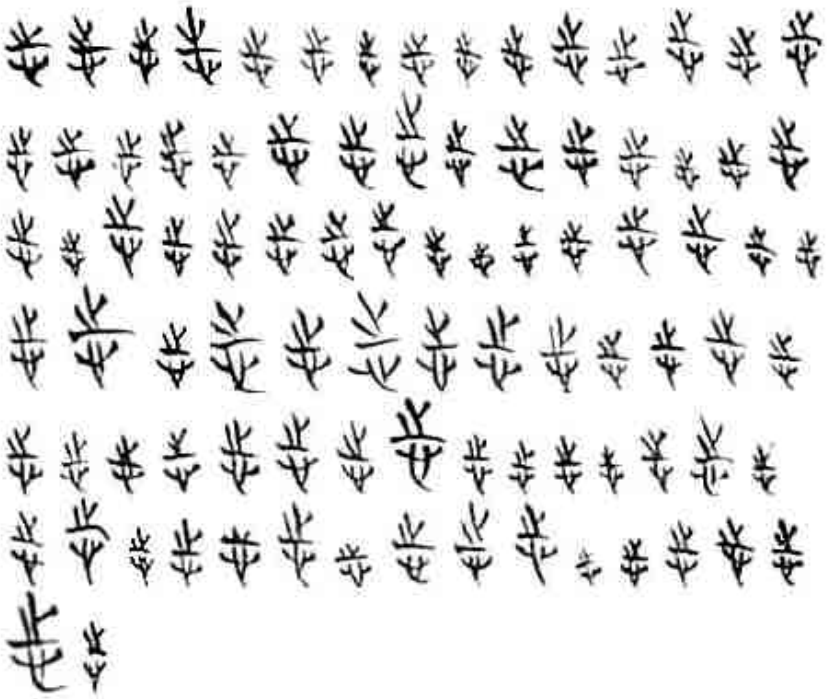
APPENDIX II.

CHARACTER FORMS SELECTED FROM THE HOUMA COVENANT TEXTS

Zhi “will, intention” (志)

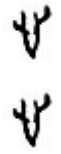
1.  之 + 心 221 occurrences





2.  寸 + 心 6 occurrences

3.  之 + 心 + 寸 4 occurrences

4.  心 1 occurrence

Fu “to restore” (復)

1.  　　复 + 彳 + 止　　153 occurrences

2. 复 + 肉 + 彳 19 occurrences

3. 复 + 彳 14 occurrences

4.  复 + 彳 + 口 5 occurrences

5.  复+彳+止+ 口 5 occurrences

6.  复 3 occurrences

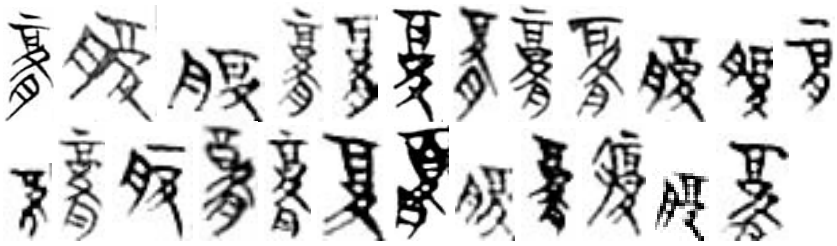
7.  复 + 肉 2 occurrences

8.  复 + 止 1 occurrence

9.  复 + 口 1 occurrence

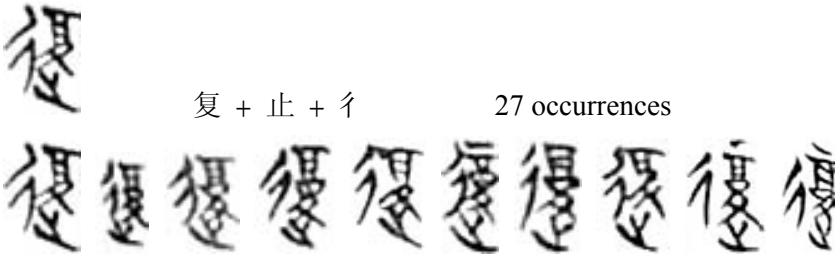
Fu “belly” (腹)

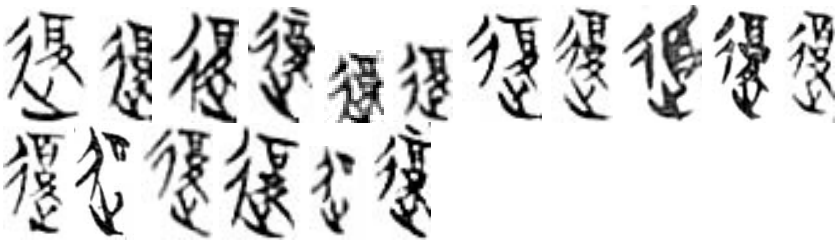
1.  复 + 肉 74 occurrences



2. 复 + 肉 + 彳 46 occurrences

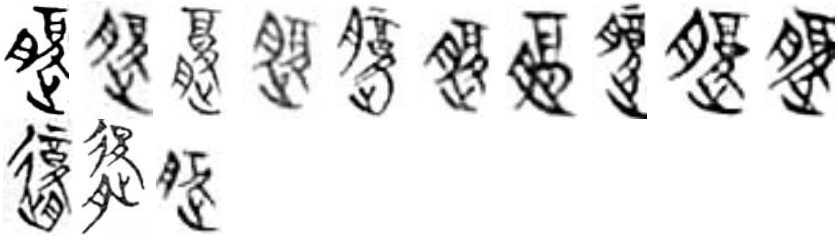
3. 复 + 止 + 彳 27 occurrences





4.  复 + 肉 + 攵 14 occurrences

5.  复 + 肉 + 止 13 occurrences



6.  复 + 彳 9 occurrences

7.  复 + 肉 + 口 8 occurrences

8.  复 + 止 5 occurrences

9.  复 + 彳 + 口 4 occurrences

10.  复+彳 +口+ 止 4 occurrences

11.  复 4 occurrences

12.  复 + 攵 2 occurrences

13.  复 + 肉 + ⺄ 2 occurrences

14.  复+ 肉 + 彳 + ⺄ 2 occurrences

15.  复 + 肉 + 彳 + 攵 2 occurrences

16.  复 + 彳 + 攵 2 occurrences

17.  复 + 彳 + 止 + 攵 2 occurrences

18.  复 + ⺆ 1 occurrence

19.  复 + 彳 + 心 + ⺆ 1 occurrence

20.  复 + 口 1 occurrence

21.  复 + 彳 + 心 1 occurrence

22.  复 + 心 1 occurrence

BIBLIOGRAPHY

1. Traditional sources

Ban Gu 班固. *Hanshu* 漢書. Commentary by Yan Shigu 顏師古. Beijing: Zhonghua, 1975.

Chen Shou 陳壽. *Sanguozhi* 三國志. Beijing: Zhonghua shuju, 1959 (1973 printing).

Du You 杜佑. *Tongdian* 通典. Taibei: Taiwan shangwu yinshuguan, 1983.

Fan Ye 范曄. *Houhanshu jijie* 後漢書集解. Changsha: Shangwu yinshuguan, 1940.

Guanzi 管子. Shanghai: Shangwu, 1926.

Guo Zhongshu 郭忠恕. *Hanjian* 汗簡. Shanghai: Tianshizhai, 1883.

Jiao Hong 焦竑. *Sushu kanwu* 俗書刊誤. Taibei: Taiwan shangwu yinshuguan, 1983.

Laozi Daodejing 老子道德經. Taibei: Taiwan shangwu yinshuguan, 1967.

Liu Xiang 劉向. *Zhanguoce* 戰國策. Shanghai: Shanghai guiji, 1978.

Lunyu 論語. *Shisanjing zhushu fu jiaokan ji* 十三經注疏附校勘記. Taibei: Yiwen, 1976.

Wei Zheng 魏徵. *Suishu* 隋書. Beijing: Zhonghua shuju, 1973 (1991 printing).

Ouyang Xiu 歐陽修. *Xin Tangshu* 新唐書. Shanghai: Shangwu yinshuguan, 1936.

Mengzi 孟子. Ruan Yuan 阮元, ed. *Shisanjing zhushu fu jiaokan ji* 十三經注疏附校勘記. Taibei: Yiwen, 1976.

Shangshu 尚書. Ruan Yuan 阮元, ed. *Shisanjing zhushu fu jiaokan ji* 十三經注疏附 校勘記. Taibei: Yiwen, 1976.

Shen Yue 沈約. *Songshu* 宋書. Beijing: Zhonghua shuju, 1977.

Shi You 史游; Commentary by Yan Shigu 顏師古. *Jijiupian* 急就篇. Taipei: Taiwan shangwu, 1983.

Sima Qian 司馬遷. *Shiji* 史記. Beijing: Zhonghua shuju, 1959 (1975 printing).

Wang Yuzhi 王與之. *Zhouli dingyi* 周禮訂義. *Yingyin wenyuange siku quanshu*, vol. 93–94. Taibei: Taiwan shangwu, 1983.

Xia Song. 夏竦. *Guwen sishengyun* 古文四聲韻. Taipei: Xuehai, 1978.

Zhouli 周禮. *Shisanjing zhushu fu jiaokan ji* 十三經注疏附校勘記. Taibei: Yiwen, 1976.

Xu Shen 許慎. *Shuowen jiezi: Fu jianzi* 說文解字：附檢字. Hongkong: Taiping shuju, 1966.

Yan Shigu 顏師古. *Kuangmiu zhengsu* 匡謬正俗. Shanghai: Shangwu yinshuguan, 1936.
Yan Yuansun 顏元孫. *Ganlu zishu* 干祿字書. *Gujingjie huihan: Fu xiaoxue huihan* 古經解彙函・附小學彙函. Guandong: Yuetong, 1873.
Zhang Shen 張參. *Wujing wenzi* 五經文字. Taibei: Taiwan Shangwu yinshuguan, 1983.
Zhang You 張有. *Fugupian* 復古篇. Taibei: Taiwan Shangwu yinshuguan, 1983.
Zheng Qiao 鄭樵. *Liushu lue* 六書略. Taibei: Yiwen yinshuguan, 1976.
Zuozhuan 左傳. Ruan Yuan 阮元, ed. *Shisanjing zhushu fu jiaokan ji* 十三經注疏附校勘記. Taibei: Yiwen, 1976.

2. Modern sources

Bagley, Robert
1999 "Shang Archaeology." In Loewe, Michael and Shaughnessy, Edward L. (eds.), *Cambridge History of Ancient China: from the origins of civilization to 221 B.C.* Cambridge, UK: New York: Cambridge University Press.

Barnard, Noel
1965 "Chou China: A Review of the Third Volume of Cheng Te-k'un's *Archaeology in China*." *Monumenta Serica* XXIV., pp. 307–459.
1979 "The nature of the Ch'in 'Reform of the Script' as Reflected in Archeological Documents Excavated under Conditions of Control." T. Roy and Tsien Tsuen-hsuin, eds. *Ancient China: Studies in Early Chinese Civilization*. Hong Kong: The Chinese UP: 181–213.

Beijing tushuguan jinshi zu 北京圖書館金石組編 (ed.)
1989 *Beijing tushuguan cang Zhongguo lidai ketaben huibian* 北京圖書館藏中國歷代刻拓本匯編. Zhengzhou: Zhongguo guji.

Bodde, Derk
1938 *China's First Unifier: A Study of the Ch'in Dynasty as Seen in the Life of Li Ssu (280?–208 B. C.)*. Leiden: E. J. Brill.

Boltz, William G.
1994 *The Origin and Early Development of the Chinese Writing System*. New Haven, Connecticut: American Oriental Society.

Boodberg, Peter A.
1937 "Some Proleptical Remarks on the evolution of archaic Chinese." *Harvard Journal of Asiatic Studies* 2:329–372.

1995 *Huangdi sijing jinzhu jinyi: Mawangdui Han mu chutu boshu* 黃帝四經今註今譯：馬王堆漢墓出土帛書. Taibei: Taiwan shangwu yinshuguan.

Chen Mengjia 陳夢家
1966 "Dong Zhou mengshi yu chutu zaishu" 東周盟誓與出土載書. *Kaogu* 5:271–281.

Chen Songchang 陳松長
1996 *Mawangdui boshu yishu* 馬王堆帛書藝術. Shanghai: Shanghai shudian.

Cheng Te-k'un
1959 *Archaeology in China*. Cambridge, UK: W. Heffer.
1971 "The inconstancy of character structure in Chinese writing." *Journal of the Institute of Chinese Studies* (Hong Kong University): 137–170.

Connery, Christopher Leigh
1998 *The Empire of the Text: Writing and Authority in Early Imperial China*. Lanham, Md.: Rowman & Littlefield Publishers.

Cornell Kimball
NA "A Study of Some of the Most Commonly Misspelled Words." http://www.barnsdle.demon.co.uk/spell/error.html.

Cui Renyi 崔仁義
1998 *Jingmen Guodian Chujian* Laozi *yanjiu* 荊門郭店楚簡《老子》研究. Beijing: Kexue.

D. C. Lau (tr.)
1970 *Mencius*. Baltimore, Maryland: Penguin Books.
1979 *Confucius: The Analects*. Maryland: Penguin Books.

Deng Dexiu 鄧德修
1979 *Wei shijing chutan – Wei shijing guzhuan zidian* 魏石經初探・魏石經古篆字典. Taibei: Xuehai.

Dong Lianchi 董蓮池
2000 "Ershi shiji Zhongguo xuezhe de Zhanguo wenzi yanjiu" 二十世紀中國學者的戰國文字研究. *Gudai hanyu yanjiu* 4:11–14.

Dong Zuobin 董作賓
1956 *Xiaotun dierben Yinxu wenzi yibian* 小屯第二本殷虛文字乙編. Beijing: Kexue chubashe.

Duan Yucai 段玉裁
1815 *Shuowen jiezi zhu* 說文解字注. Rpt. of 1815 ed. Taibei: Yiwen yinshuguan, 1955.

Huang Yongwu 黃永武 (ed.)

1981–1986 *Dunhuang Baocang* 敦煌寶藏. Taibei: Xin wenfeng.

Galambos, Imre
2005 "A Corpus-Based Approach to Palaeography: The Case of the Houma Covenant Texts." In Matthias Richter (ed.), *Asiatische Studien*. LIX. I. 2005: 115–130.

Gao Ming 高明
1980 "Guwenzi de xingbang ji qi xingti yanbian" 古文字的形傍及其形體演變. *Guwenzi yanjiu* 4:41–90.
1987 *Zhongguo guwenzixue tonglun* 中國古文字學通論. Beijing: Wenwu.

Guo Ruoyu 郭若愚
1994 *Zhanguo Chujian wenzi bian* 戰國楚簡文字編. Shanghai: Shanghai shuhua.

Henricks, Robert G. (tr.)
Lao-Tzu Te-Tao Ching: A New Translation Based on the Recently Discovered Ma-wang-tui Texts. New York: Ballantine Books.

Huang Shengzhang 黃盛璋
1981 "Guanyu Houma mengshu de jige wenti" 關於侯馬盟書的幾個問題. *Zhongyuan wenwu* 2:27–33.

Huang Xiquan 黃錫全
1990 *Hanjian zhushi* 汗簡注釋. Wuhan: Wuhan daxue.

Hubei sheng Jingsha tielu kaogudui 湖北省荆沙鐵路考古隊
1991 *Baoshan Chu jian* 包山楚簡. Beijing: Wenwu.

Hubei sheng wenwu kaogu yanjiu suo 湖北省文物考古研究所 and Beijing daxue zhongwenxi 北京大學中文系
1995 *Wangshan Chujian* 望山楚簡. Beijing: Zhonghua.

Hucker, Charles
1985 *A dictionary of official titles in imperial China*. Stanford: Stanford University Press.

Jennings, William (tr.)
1985 *Confucian Analects: A translation, with annotation and an introduction by William Jennings*. London, New York: Routledge and Sons, Ltd.

Jingmen shi bowuguan 荆門市博物館
1998 *Guodian Chumu zhujian* 郭店楚墓竹簡. Beijing: Wenwu.

Karlgren, Bernhard
1923 *Analytic Dictionary of Chinese and Sino-Japanese*. New York: Dover Publications, 1974.

Kathman, David

NA “The Spelling and Pronunciation of Shakespeare’s Name.” http://shakespeareauthorship.com/name1.html.

Katsu Eikai 葛英會 and Ho Ko 彭浩 (eds.)
1992 *So kanpaku moji hen* 楚簡帛文字編. Tokyo: Toho shoten, 1992.

Knoblock, John and Riegel, Jeffrey
2000 *The Annals of Lü Buwei: A Complete Translation and Study*. Stanford: Stanford University Press.

Lao Gan 勞榦, ed.
1992 *Juyan Hanjian* 居延漢簡（圖版之部）. Taipei: Institute of Philology, Academia Sinica.

Ledderose, Lothar
2000 *Ten Thousand Things: Module and Mass Production in Chinese Art.* Princeton: Princeton University Press.

Li Cunzhi 李存智
1995 *Qin Han jiandu boshu zhi yinyunxue yanjiu* 秦漢簡牘帛書之音韻學研究. PhD. dissertation, Taiwan National University.

Li Xueqin 李學勤
2002 “Shi shuo Zhangjiashan jian *Shilu*” 試說張家山簡《史律》. *Wenwu* 2002, 4, pp. 69–72.

Lin Suqing 林素清
1976 *Xian Qin guxiyin wenzi yanjiu* 先秦古璽印文字研究. MA thesis, Taiwan National University.

Lin Yun 林澐
1983 “*Li feng* bian” 豊豐辨. *Guwenzi yanjiu* 8:181–186.

Luo Fanzhen 羅凡㪚
2001 “Guodian Chujian yitizi yanjiu 郭店楚簡異體字研究.” *Guoli Taiwan Shifan daxue Guowen yanjiusuo jikan* 國立台灣師範大學國文研究所 集刊 45:363–608.

Luo Zhenduan 呂振端
1981 *Wei Santi shijing canzi jizheng* 魏三體石經殘字集證. Taipei: Xuehai.

Luo Zhenyu 羅振玉
1930 *Han Xiping shijing canzi jilu* 漢熹平石經殘字記錄. n.p.

Ma Chengyuan 馬承源
2001 *Shanghai Bowuguan cang Zhanguo Chu zhushu (yi)* 上海博物館藏戰國楚竹書（一）. Shanghai: Shanghai guji.

Ma Zonghuo 馬宗霍

1959 *Shuowen jiezi yin fangyan kao* 說文解字引方言考. Beijing: Zhongguo kexueyuan.

Mandarin Promotion Council, Ministry of Education, R.O.C 中華民國教育部國語推行委員會
2001 *Yitizi zidian* 異體字字典. http://140.111.1.40.

Matsumaru Michio 松丸道雄
1980 *Sei-shuu seidouki to sono kokka*. 西周青銅器及其國家 Tokyo: Tokyo daigaku.

McIntosh, Angus, M. L. Samuels, Michael Benskin
1986 *A linguistic atlas of late mediaeval English*. New York: Aberdeen University Press.

Morohashi Tetsuji 諸橋轍次
1955–1960 *Daikanwa jiten* 大漢和字典. Tokyo: Taishukan Shoten.

Morris, William, ed.
1980 *The American Heritage Dictionary.* Boston: Houghton Mifflin Co.

Perelomov, Leonard Sergeevich
1993 *Konfutsii: zhizn', uchenie, sud'ba*. Moskva: Nauka.

Qi Guiyan 戚桂宴
1979 "*Ma yi fei shi* jie" 麻夷非是解. *Kaogu* 3:272, 230.

Qin Gong 秦公
1995 *Guang bei biezi* 廣碑別字. Beijing: Guoji wenhua. (With Liu Daxin 劉大新.)

Qiu Xigui 裘錫圭
1974 "Cong Mawangdui yihao Han mu 'qiance' tan guanyu guli de yixie wenti" 從馬王堆一號漢墓"遣冊"談關於古隸的一些問題. *Kaogu* 1.
1992 *Guwenzi lunji* 古文字論集. Beijing: Zhonghua shuju.
1996 "Jiaguwen zhong de jian yu shi" 甲骨文中的見與視. *Jiaguxue bainian lunwen mulu* 甲骨學百年論文目錄.
URL: http://www.ntnu.edu. tw/ch/ob2.htm.
2000 *Chinese Writing*. Berkeley: The Society for the Study of Early China and The Institute of East Asian Studies, UC Berkeley.

Serruys, Paul L.-M.
1984 "On the System of the Pu Shou in the Shuo-wen chieh-tzu." *Bulletin of the Institute of History and Philology* 55:651–754.

Shanxi sheng wenwu gongzuo weiyuanhui 山西省文物委員會
1976 *Houma mengshu* 侯馬盟書. Shanghai: Wenwu.

Shi Anchang 施安昌

1992 *Yan Zhenqing shu Ganlu zishu* 顏真卿書干祿字書. Beijing: Zijincheng chubanshe.
2001 *Shanben beitie lunji* 善本碑帖論集. Beijing: Zijincheng chubanshe.

Shuihudi Qinmu zhujian zhengli xiaozu 睡虎地秦墓竹簡整理小組, ed.
2001 *Shuihudi Qinmu zhujian* 睡虎地秦墓竹簡. Beijing: Wenwu.

Teng Rensheng 騰忍生
1995 *Chu xi jianbo wenzi bian* 楚系簡帛文字編. Wuhan: Hubei jiaoyu.

Tsien, Tsuen-Hsuin
1993 "*Chan kuo ts'e*" 戰國策. Michael Loewe, ed. *Early Chinese Texts: A Bibliographical Guide*. Berkeley: The Society for the Study of Early China and the Institute of East Asian Studies, University of California, Berkeley.

Wang Hui 王輝
1990 *Qing tongqi mingwen biannian jishi* 青銅器銘文編年集釋. Xi'an: Sanqin.

Wang Li 王力
1982 *Tongyuan zidian* 同源字典. Beijing: Shangwu yinshuguan.
1984 *Wang Li wenji* 王力文集. Jinan: Shanding jiaoyu.

Wang Yuxin 王宇信 and Yang Shennan 楊升南
1999 *Jiaguxue yibai nian* 甲骨學一百年. Beijing: Shehui kexue wenxian.

Wang Zuoxin 王作新
2000 "'Fan zheng wei fa' jie" "反正爲乏"解. *Guhanyu yanjiu* 3:86-87.

Weld, Susan R.
1997 "The Covenant Texts from Houma and Wenxian." Shaughnessy, Edward L., ed. *New Sources of Early Chinese History: An Introduction to the Reading of Inscriptions and Manuscripts*. Berkeley: The Society for the Study of Early China and The Institute of East Asian Studies, UC Berkeley.

Wu Hung 巫鴻
1979 "Qin quan yanjiu" 秦權研究. *Gugong bowuyuan yuankan* 故宮博物院院刊 4:33–47.

Yang Haiqing 陽海清, Chu Peiyu 褚佩瑜, Lan Xiuying 蘭秀英.
2002 *Wenzi yinyun xungu zhijian shumu* 文字音韻訓詁知見書目. Wuhan: Hubei renmin.

Zhang Yongquan 張涌泉
1998 *Hanyu suzi yanjiu* 漢語俗字研究. Changsha: Yuelu shushe.

Zhao Ping'an 趙平安
1993 *Libian yanjiu* 隸變研究. Baoding: Hebei Daxue chubanshe.

Zhou Cezong 周策綜

1968 Shuo ‘shi zhi quewen’” 說《史之闕文》. *Dalu zazhi* 37:104–118.

Zhou Fengwu 周鳳五
1994 “Houma mengshu niandai wenti chongtan” 侯馬盟書年代問題重探 . *Zhongguo wenzi* 19:111–135.

Zhu Dexi 朱德熙 and Qiu Xigui
1972 “Zhanguo wenzi yanjiu (liuzhong)” 戰國文字研究 (六種). *Kaogu xuebao* 1:73–89.