

**Zeitschrift:** Asiatische Studien : Zeitschrift der Schweizerischen Asiengesellschaft = Études asiatiques : revue de la Société Suisse-Asie

**Herausgeber:** Schweizerische Asiengesellschaft

**Band:** 59 (2005)

**Heft:** 1: Methodological issues in the study of early Chinese manuscripts : papers from the second Hamburg tomb text workshop

**Artikel:** A corpus-based approach to palaeography : the case of the Houma covenant texts

**Autor:** Galambos, Imre

**DOI:** <https://doi.org/10.5169/seals-147670>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

**Download PDF:** 03.04.2025

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

# A CORPUS-BASED APPROACH TO PALÆOGRAPHY

## The case of the Houma covenant texts\*

Imre Galambos, The British Library

### *Abstract*

This article is an experiment with a corpus-based approach to the study of early Chinese manuscripts and inscriptions. This method is concerned with the statistical correlation between words and their graphical representation. It looks at writing habits and patterns of usage, identifying the dominant ways of writing words. Naturally, a statistical approach can only produce reliable results in the case of a relatively large corpus. The Houma covenant texts represent an ideal material for such analysis because they consist of hundreds of nearly identical copies of the same few basic texts in which only the name of oath taker varies. Thus the same word often occurs in the same context hundreds of times, excluding the possibility of semantic divergence. The article uses the cryptic formula *ma yi fei shi* 麻夷非是 to show how such a method could benefit one's understanding of a difficult passage.

Archaeological discoveries of the past few decades in China have yielded an unprecedented amount of manuscript material. An increasing number of these excavated texts are now becoming available to researchers in photographic format, providing a much-needed raw data for palaeographic analysis. In this paper, I would like to draw attention to the importance of the corpus-based approach in palaeography. I use the cryptic formula *ma yi fei shi* 麻夷非是 from the Houma covenant texts (*Houma mengshu* 侯馬盟書) to show how such a method could benefit one's understanding of a difficult passage.

An often disregarded issue in the study of character forms is that beside the occurrence of forms it is also useful to document the frequency of their occurrences. Many of the dictionaries and character compendia are still deficient in this respect, because they only list the various forms and thus give a skewed impression of the relevance of individual character forms. Nevertheless, despite the apparent orthographic variability in early manuscripts, it is possible to

\* I would like to thank the participants of the Second Hamburg Tomb Text Workshop, especially Matthias Richter, for their valuable suggestions leading to this article. I would also like to thank Gábor Kósa and Péter Vámos for their help and input.

establish the presence of a preferred or dominant form that was used by the scribes with the highest frequency. This dominant form can only be extracted from a statistically significant amount of data; solitary instances of orthographic variants are often not representative of general usage but simply examples of possible configurations.






Naturally, a basic prerequisite for any statistical analysis is that one is certain that the forms compared all stand for the same word. In early manuscripts, however, two visually identical character forms could represent two distinct words, and vice versa. In this respect, the Houma covenant texts<sup>1</sup> represent an ideal material because they consist of hundreds of nearly identical copies of the same few basic texts in which only the name of oath taker varies. Thus the same word often occurs in the same context hundreds of times, excluding the possibility of semantic divergence.

In order to document the patterns of variation, I first observe examples where the meaning and grammatical function of the word is unambiguous. These words are A) *zhi* “will, intention” (written in the modern script as 志); B) *fu* “to restore” (written in the modern script as 復); and C) *fu* “abdomen” (written in the modern script as 腹). The tables below show the frequency of occurrences of variant character forms used to write them.<sup>2</sup>

A) The word *zhi* “will, intention” appears in the context “should he dare to have the intention of restoring 敢有志復 [...]”. Since the context is identical in each case, one can be certain that all variants, regardless of how different they are graphically, stood for the same word. Table A shows the orthographic variants of the 232 legible character forms.










1 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. The texts on the tablets date to the beginning of the 5<sup>th</sup> century BC, sometime between 497 and 470. Because the texts were not incised but written with ink and a stylus or pen, they represent the earliest large-scale corpus of handwritten materials available today. Their texts are so-called “covenant texts” from the state of Jin 晉 that recorded the oaths of vassals swearing alliance and loyalty to a covenant lord. Starting from the 1930s, covenant texts were also unearthed in Wenxian 溫縣. 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.

2 For a more detailed analysis of these three words in the Houma covenant texts see Galambos 2002.

A1.	221		之+心
A2.	6		之+寸 (= 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

B) The word *fu* “to restore” appears in the same context “should he dare to have the intention of restoring 敢有志復 [...]”, referring to the restoration and support of the enemy clan.<sup>3</sup> Table B shows the orthographic variants of the 207 legible character forms.<sup>4</sup>

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

- 3 The same character also occurs as writing *fu* in the adverbial sense of “again, repeatedly”. Being a different part of speech with a distinct semantic sphere, it qualifies as a separate word.
- 4 At this level of analysis, I chose to consider the component 复 as a single unit without dividing it any further. However, the apparent orthographic variability of this component could be interesting in its own right.

C) The word *fu* “abdomen” appears in the context “should he dare not to exert his body (=abdomen) and mind in serving his lord 敢不剖其腹心以事其主”. Table C shows the orthographic variants of the 225 legible character forms.

C1.	74		复+肉	C12.	2		复+攴
C2.	46		复+肉+彳	C13.	2		复+肉+冫
C3.	27		复+止+彳	C14.	3		复+肉+彳+冫
C4.	14		复+肉+攴	C15.	2		复+肉+彳+攴
C5.	13		复+肉+止	C16.	2		复+彳+攴
C6.	9		复+彳	C17.	2		复+彳+止+攴
C7.	8		复+肉+口	C18.	1		复+冫
C8.	5		复+止	C19.	1		复+彳+心+冫
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

In these examples, graphic variability cannot be interpreted on the basis of semantic considerations; the only possible reason behind it is the flexibility of the writing habits on the part of the people who produced these manuscripts. At the same time, despite the high degree of orthographic variability the examples demonstrate the presence of a dominant form: the most common way of record-

ing a particular word. One can suspect that if more material came to light from the same region and era, the current dominant form would still prove to be the most frequent one. Naturally, the larger one's initial pool of sample data is, the more accurate such assumptions would be. In the case of the words *zhi* “will, intention” and *fu* “abdomen”, the structure of the dominant form matched the way these words are written in the modern script: 志 and 腹. One cannot fail to notice that out of the sometimes numerous structurally different configurations (e.g. at least 22 variants in the case of the word *fu* “abdomen”) it is the dominant form that corresponds to the modern one. Naturally, in the course of the evolution of the script dominant forms had a higher chance of prevailing, whereas less common forms disappeared.<sup>5</sup>

In the case of the word *fu* “to restore”, however, the structure of the dominant form in the Houma corpus (彳+复+止) is not identical to the modern one (彳+复). The difference lies in the presence of the component 止 in the Houma form or, depending on our point of view, its absence from the modern form. Thus in this case, the usual way of writing the character differs from its modern counterpart. This observation is important with respect to the identification of *jiajie* 假借 loans, because even though from the perspective of modern script it is form B3 that corresponds to the structure of the character 復 in modern orthography, it would be incorrect to regard B1, which was the typical representation of the same at that time, as a loan for the less common B3 form. Instead, one should accept that B1 was the common way of writing the word *fu* “to restore”, regardless of its dissimilarity with the modern form.

Another pattern apparent in the three examples above is that nearly every form retains the phonetic element.<sup>6</sup> Thus the variations mostly consist of the

5 This is not to say that new non-dominant forms did not arise. Much to the contrary, manuscripts from later periods (Han, Tang, etc) confirm that character variants continued to co-exist.

6 The one exception from this in all of the above examples is form A4 (𠄎) “substituting” the character 志. From the point of view of the context, the word *xin* “heart” (𠄎) is synonymous with the word *zhi* “will” (志), which seems to suggest that this case of structural variation could be interpreted as a character level variation caused by lexical variation – in this case one between synonyms. Needless to say, the structural connection between the two characters also influenced the choice of the synonym. The possibility that this is simply a case where due to the erosion of the jade tablet and the ink, the upper part of the character cannot be seen anymore can be ruled out on the basis of the spacing between the characters.

addition or omission of semantic components.<sup>7</sup> Once again, although some of the character forms might appear to be *jiājiē* loans (e.g. form C22, 復, denoting the word *fu* “willful, stubborn” in the modern script), looking at the variant forms as a group reveals that those cases are merely instances of structural variation which “overlap” with different characters; the structure of most of the other variant forms does not exist in modern script.

A perceptible feature among the variant forms is that characters that appear on the same manuscript can influence each other’s structure. For example, form A5 appears earlier in the same text as the second character in the place name Pingsi. Thus the component 阜 in form A5 is there as a direct influence of the structurally otherwise identical place name.<sup>8</sup> The same phenomenon can be seen in the structure of the variant forms of characters 復 and 腹. Although the component 肉 in forms B2 and B7 could have some semantic relevance to restoring someone’s clan but the prominent use of the components 彳 and 止 in forms C2, C3, C5, C6, C8, C9, etc. can only be explained as an influence of character 復. Another example is forms C21 and C22 which include the component 心, probably because they appear in the phrase *fluxin* “body and mind” (腹心) and are thus influenced by the following character.<sup>9</sup>

With respect to the structure of other variant forms, the addition or omission of semantic components was not entirely random. While the structure appears somewhat haphazard on the level of an individual character form, on the level of the whole corpus their sum total carries a certain amount of additional semantic information. The dominant components among the different character forms used for writing the same word are in accordance with what we are used to interpreting as semantic determinatives of these characters, and they fit the context well. It will, however, hardly be possible to recover in each specific case the semantic information possibly encoded in the choice of components absent from the modern standard character for the respective word.

For example, in group B, 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

7 In naming components “phonetic” or “semantic”, I am merely following a convention. At the same time, I fully agree with Boodberg’s (1937: 335) argument that even when a component is used for its semantic value, it also carries a “weak” phonetic value, and vice versa.

8 The same influence can also account for the presence of the component 寸 in form A3.

9 On the manuscripts, the character 心 follows both forms C21 and C22, proving that these are not simply cases of two characters being joined together.

seen in forms B1 and B8, typically appears in the modern script as the component 辵/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”).<sup>10</sup> Therefore, the choice of the ancient scribe to write the forms B1, B5 and B8 with the components 辵 + 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.<sup>11</sup>

The above examples served to demonstrate some of the basic patterns behind the variability of characters representing known words; there is no ambiguity for the modern reader regarding the meaning and usage of these words in context. By keeping in mind these patterns, one has the potential to disambiguate words in manuscripts or inscriptions that are unknown or obscure today. The significance of this approach is that it relies on the information gathered from the corpus itself, rather than trying to massage individual cases to fit one’s own understanding of early Chinese language and writing. Thus one favors the manuscript data versus linguistic knowledge which is primarily derived from transmitted literature.

Most of the Houma covenant texts end with the words *ma yi fei shi* 麻夷非是.<sup>12</sup> This formula usually appears at the end of the following or a similar pledge:

[...] 而敢有志復 [...] 晉邦之地者, 及群乎盟者, 吾君其明殛視之, 麻夷非是.

10 In some modern characters the components 辵 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.

11 Forms B4, B5, and B9 include the component 止 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 止 was added not as a separate component but as a component linked to the 辵 (appearing in form B5), together forming the component 辵. In forms B4 and B9, which do not include the component 止, the component 止 could simply be the abbreviation of 辵. The component 辵 could signify a “base” (as suggested by the component 止) or movement (as suggested by the component 辵). The presence of the component 止 in group C is likely a result of an influence of group B.

12 For an overview of the different types of covenant texts, see Weld 1997.



[...] and should he dare to have the intention of restoring [...] (here comes a list of enemies and their 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, *ma yi fei shi*.

The context makes it clear that the formula represented something detrimental, mostly fatal that would happen to the person in violation of the covenant. A number of similar examples are found in transmitted literature (e.g. *Zuozhuan* 左傳, *Gongyangzhuan* 公羊傳, *Sanguozhi* 三國志), all ending with some sort of curse upon those who dare to breach the covenant.

Despite the consensus regarding the general meaning of the formula, there has been some disagreement in its reading and interpretation. In 1966, based on only a couple of examples, Chen Mengjia 陳夢家 transcribed it as 麻夷我是, claiming that it meant 滅亡我氏 (“annihilate my clan”).<sup>13</sup> Thus he identified the third character as 我, undoubtedly in an attempt to try to render the formula comprehensible.

A few years later, Zhu Dexi 朱德熙 and Qiu Xigui 裘錫圭 corrected Chen Mengjia’s transcription, arguing that the third of the four characters was unquestionably 非, not 我.<sup>14</sup> They equated the formula with the phrase *mei zhi bi shi* 昧雉彼視 in the *Gongyangzhuan* (Xiang 襄 27) and interpreted both of them as 滅夷彼氏 (“annihilate that clan”). The *Gongyangzhuan* phrase likewise occurs as a closing remark of a covenant text: “Should one step on the land of Wei or eat the millet of Wei, *mei zhi bi shi* 苟有履衛地, 食衛粟者, 昧雉彼視”. The exact meaning of the closing phrase has been troubling readers and commentators for centuries; in lack of a better solution, He Xiu 何休 (129–182) offered the following explanation: “昧 means to slaughter (割). At the time people concluded covenants by slaughtering a pheasant. It is like saying, ‘Watch that pheasant being slaughtered: should someone violate this covenant, he will be just like that.’ 昧, 割也。時割雉以爲盟, 猶曰視彼割雉, 負此盟則如彼矣。” Zhu Dexi and Qiu Xigui showed that the character 昧 did not occur in early dictionaries and was probably a mistake for the character 昧 which was close in pronunciation to and thus interchangeable with the character 蔑. The character 蔑, in turn also had an attested meaning of “destruction, annihilation”.

In the same year, Guo Moruo 郭沫若 also confirmed the same 滅夷彼氏 reading.

13 Chen Mengjia 1966: 276.

14 Zhu Dexi 1972: 73.

Following the publication of the *Houma mengshu* 侯馬盟書 volume<sup>15</sup> with photographs and tracings of hundreds of covenant plaques, Qi Guiyan 戚桂宴 raised doubts regarding the above reading.<sup>16</sup> He pointed out that within the Houma corpus there were hundreds of examples of the phrase 非是, not one of them written as 彼氏. He also called attention to the context of the *Gongyangzhuan* passage in which the covenant was concluded between a husband and his wife, wherefore having one's clan wiped out was hardly an appropriate punishment for infringing the contract. Qi Guiyan proposed to read the first two characters as Wuyi 無夷, the name of Hebo 河伯, the spirit of the Yellow River, who was also referred to as Bingyi 冰夷 or Fengyi 馮夷. Thus he read the formula as 無夷非是, in the sense that Wuyi, the Yellow River spirit would impose sanctions upon those who violated the covenant.

A few years later, Li Yumin 李裕民 published an article contesting former readings and putting forward his own one.<sup>17</sup> He claimed that the first three characters of the formula were synonymous with each other and could be read in the sense of 滅滅滅氏. As a solution for the apparent grammatical awkwardness of this structure, he claimed that this was an abbreviation of a more complete version of the formula. In support of his argument, he cited plaques 1.41 and 1.42 in which the formula appeared as 麻夷之非是.<sup>18</sup> Thus he argued that 麻夷之 meant the death of the person who violated the covenant, whereas 非是(氏) referred to the annihilation of his entire clan.

Of the above interpretations, the most widely accepted one is still that of Zhu Dexi and Qiu Xigui. Indeed, there can be little doubt about the correctness of equating *ma yi fei shi* with the *mei zhi bi shi* formula in the received text of the *Gongyangzhuan*. However, reading it as 滅夷彼氏 poses several problems. First, they assume two *jiatie* substitutions in the second half of the formula: they read the character 非 as a phonetic loan for 彼, and 是 for 氏. While phonetically both of these substitutions are valid and attested, one has to take into account that there are hundreds of examples where such a substitution did not take place.

D1. 291  非

Table D: Frequency of variant forms of character #3 in the formula

15 Shanxi sheng wenwu gongzuo weiyuanhui 1976.

16 Qi Guiyan 1979.

17 Li Yumin 1983.

18 In reality it is plaques 1.40 and 1.41 where this version of the formula appears.

The frequency table of variant forms for the character 非 consists of a single form (Table D). This means that on every single covenant tablet where the character 非 occurs legibly, it was invariably written as 非, never as 彼 or anything else.<sup>19</sup> Thus there are nearly three hundred manuscript examples against a single received version in the *Gongyangzhuan*. If anything, one should consider the transmitted version a *jiajie* loan, rather than the other way around. Alternatively, one would have to demonstrate that 1) the word written in the modern script with the character 彼 appeared in the form of 非 in early manuscripts and inscriptions and that 2) 非 can often represent the word *bi* “that”. This, however, is not the case. Although the character 彼 is not seen in early manuscripts, there are many examples where the form 皮 is read as 彼. At the same time, to my knowledge the character 非 is never read as 彼 in any of the early manuscripts outside the Houma corpus.<sup>20</sup> The fact that despite the relatively high degree of orthographic variability in the Houma covenant texts the character 非 shows absolutely no variation is in itself an indication that it indeed represented the same word as it usually does.




E1.	257		是
E2.	7		氏
E3.	2		正

Table E: Frequency of variant forms of character #4 in the formula

The fourth character of the formula, on the other hand, appears in more than one structurally different form (Table E). In the covenant texts, the respective word is written as 是 in over two hundred fifty cases, and only half a dozen times as 氏. It is true that in pre-Qin and even Han writing the two characters were

19 There is one instance on tablet 203.11 where a fragment reads 殺是 but due to the erosion of the ink it is hard to judge whether this occurs as part of the formula in question. If the context were certain, this fragment could serve as an important key in deciphering the meaning of phrase 非是.

20 Contrary to archaeological evidence, there are a few examples of interchangeability between 匪 and 彼 in transmitted literature, i.e. the phrase 彼交 from the *Shijing* (Mao 215 and 222) is quoted as 匪交 in the *Taiping yulan* 太平御覽 (498), the *Xunzi* (“Quan xue 勸學”) and the *Hanshi waizhuan* 韓詩外傳 (4).

frequently used in place of each other,<sup>21</sup> at the same time, it is useful to point out that it was the character 是 that was much more commonly used for 氏 than the other way around. In other words, the word *shi* “clan” was written often as both 氏 and 是, whereas the word *shi* “this; right” was written as 氏 less frequently. Therefore, from a statistical point of view it is probable that the graph 是 was not a loan for another graph but stood for itself.

The two instances of form E3, equivalent to the modern character 正, are yet another indication that the character should be read as 是. Xu Shen 許慎 defined 正 as 是 (正, 是也), showing that the two characters were closely related. This kinship is also attested in their structure, 是 being a composite of the components 正 and 日.<sup>22</sup> On the other hand, the character 正 has no phonetic connection with the character 氏, suggesting that its meaning in this case should be “correct, right”, for which the ancient scribes used the words *shi* “correct” (written with the characters 是 and 氏) and the synonym *zheng* “correct” (written with the character 正).

Yet another consideration in favor of this reading is that the previous character (#3 in the formula) could be confidently identified as 非. The combination 非氏 could be understood as “not a clan” or “to negate the clan”, either of which would be problematic in this context. The combination 非是, on the other hand, is also attested in transmitted literature and would be easy to interpret in a conventional way: “not right”, “not this”, “to negate the correctness of something”, “to repudiate this”, etc.





F1.	268		麻
F2.	8		亡

Table F: Frequency of variant forms of character #1 in the formula

Of the first two words in the *ma yi fei shi* formula, the first one (#1 in the formula) is written in two distinct graphic forms (Table F). The dominant form is identical to the modern character 麻, save the dot on the top which does not

21 I intentionally avoid using the word “interchangeable” here, because it implies a false sense of arbitrariness in usage, disregarding the pattern of the dominant form.

22 This composition is also present in the *Shuowen jiezi* 說文解字 (hereafter: *Shuowen*) where Xu Shen says that “the character 是/晿 means upright; it is composed of the characters 日 and 正 (是, 直也. 从日正.)”.

appear on the Houma forms. In this structure, the character appears only in the Houma corpus, which shows that elsewhere the same word either did not occur or was written with a different character. Even in the Wenxian 濇縣 covenant texts, where the same formula appears at the end of each text, this particular character occurs in a distinct structure written as . The difference is that the Wenxian form has the dot on the top (not in every instance though) and two horizontal lines across the lower part of the character.<sup>23</sup> The non-dominant form (F2) in the Houma corpus is structurally identical to the modern character 亡 used in late Spring and Autumn and Warring States times to write the adverb *wu* “have not” which appears in transmitted literature as the character 無. Among the archaeological material, the structure of the Wenxian form also occurs on a bronze object from Zhongshan 中山 in the form of  which in context reads as the character 靡. The character 靡 is glossed in the *Erya* 爾雅 as 無 (i.e. 靡, 無也), in which sense it also often occurs in the *Shijing* 詩經.<sup>24</sup>

When Zhu Dexi and Qiu Xigui interpreted the character 靡 in the formula as 滅 (“to destroy”), they relied on the *Guangya* 廣雅 definition (i.e. 靡, 滅也) and the *Fangyan*’s 方言 definition of the character 靡 (i.e. 靡, 滅也). However, the reading of the Zhongshan form as 靡 is a strong argument in favor of reading the Wenxian and Houma forms the same way. This interpretation is further supported by the fact that 靡 is compatible with both forms F1 and F2. Therefore, I read the first character of the formula as 靡, in a sense identical to that of 無.

23 Despite the apparent differences, He Linyi’s 何琳儀 (1998: 888) dictionary of Warring States character forms includes both the Houma and Wenxian forms under the same 靡 entry. In addition, he also includes the Zhongshan form.

24 For example, Mao 39: “my heart is in Wei; there is not a day I do not think of it 有懷于衛靡日不思”; Mao 45: “and I swear that till death I will have no other 之死矢靡它”; Mao 132: “while I do not see my husband, my sad heart has no joy 未見君子, 憂心靡樂”. (All translations in this paper, unless otherwise indicated, are from Legge 1960, vol. 4.)





G1.	270		夷 + 土
G2.	5		夷 + 土 + 歹
G3.	1		夷 + 土 + 彳
G4.	1		夷

Table G: Frequency of variant forms of character #2 in the formula

The second word in the *ma yi fei shi* formula appears in the corpus with four structurally distinct character forms (Table G). Form G1, which is by far the most common one in the group, matches the structure of the modern character 夷 but has an additional 土 component. Despite this, transcribers and interpreters unanimously identify it as 夷.<sup>25</sup> Except for Qi Guiyan, who sees it as part of the name of Wuyi, the Yellow River spirit, scholars believe that the character stands for the verb meaning “to wipe out, destroy”. The *Shuowen* definition has the following explanation: “The character 夷 means to flatten; it is composed of the characters 大 (‘great’) and 弓 (‘bow’). [It also refers to] the people from the East. 夷, 平也, 从大从弓. 東方之人也.” With respect to the second meaning in Xu Shen’s explanation (i.e. the Yi people from the East), it is important to point out that in Spring and Autumn and Warring States inscriptions the word was generally written as 尸. Therefore, in this case, the character 夷 is likely to refer to the first meaning in the *Shuowen* definition. This reading is further corroborated by the presence of the additional 土 component, a reference to the semantic domain of “flat; flattening”.

In transmitted literature, the word *yi* “to flatten” did not necessarily refer to destruction. In the *Shijing*, for example, it usually appears in the sense of “peace” or “pacification”. For example: “And my heart will then be at peace 我心則夷” (Mao 14); “But I have seen my husband, And should I but feel at rest?”

25 The *Gongyangzhuan* version of the formula has the character 雉 in this place. Zhu Dexi and Qiu Xigui explained the connection between the characters 夷 and 雉 on phonetic grounds, citing examples of interchangeability in transmitted texts. This interpretation is also logical from a statistical point of view, since the manuscript version with hundreds of occurrences is taken as the primary form and the singular instance in the transmitted version as a loan character.

既見君子，云胡不夷” (Mao 90); “The Hëen-yun are pacified 獵狁于夷” (Mao 168). However, in the *Zuozhuan* the word clearly refers to destruction: “destroy our crops, wipe out our borderland 芟夷我農功，虔劉我邊陲” (Cheng 成 13). Similar usages also occur in other early texts: “If one man is guilty, his kinship of three generations will be annihilated” 一人有罪，而三族皆夷 (*Xunzi* 24, “Junzi 君子”).

In the Houma formula, the context would suggest a negative sense of some sort of destruction or annihilation. This reading is confirmed by the composition of form G2 that has the component 歹, which is often part of characters with a meaning related to destruction or death (e.g. 死, 歿, 殃, 殄).

In view of the above considerations, I have to disagree with reading the formula as 滅夷彼氏 (“annihilate that clan”). The overall consistency of the formula in the corpus indicates that the last two characters should be not only transcribed as 非是 but also interpreted as the words these two characters stand for in modern standard orthography. Regarding the first two characters of the formula, I suggest the 麻夷 transcription to be read as 靡夷, in the sense of “no destruction; no devastation”. In modern script, this meaning could also be glossed as 無滅. Thus I transcribe the entire formula as 麻夷非是 and read it as 靡夷非是. Naturally, even though a corpus-level analysis of the four characters leads to this transcription, their exact meaning is still not unproblematic. At this point, I can see two possible ways of rendering the formula into English. The first possibility is that “there shall be no destruction inappropriate”. In this sense, the structure would be similar to the *Shijing* (Mao 192) phrase 靡人弗勝 (“and there is none whom it will not overcome”). The second possibility is that “it would not be right if no destruction [befell the covenant breacher]”. In other words, he must meet destruction no matter what. In this sense, the structure would be similar to the *Guliangzhuan* 穀梁傳 (Cheng 成 8 and Zhuang 莊 1) statement 錫命，非正也 (“issuing the mandate is not right”). Either interpretation would fit the larger context and could be interpreted as 1) “may the bright spirits of our [former] lords punish and scrutinize him with no [degree of] destruction being inappropriate”; or 2) “may the bright spirits of our [former] lords punish and scrutinize him; it would not be right if destruction did not [befall him]”.<sup>26</sup>

26 Naturally, these readings are only a tentative attempt to interpret the formula and could be revised in the future when more information becomes available.

In sum, the analysis of orthographic variations across an entire corpus permits one to make observations regarding the usage of character forms in context. This approach is to a certain degree independent of phonetic considerations, because it is concerned with the statistical correlation between words known to be identical and their different graphic representation. It looks at writing habits and patterns of usage, identifying the dominant ways of writing particular words. Naturally, a statistical approach can only produce reliable results in the case of a relatively large corpus. The Houma covenant texts, where orthography can be analyzed in an unchanging context that guarantees an identical grammatical function, provide a rare opportunity for disambiguating usage and meaning based on the inconsistency (or stability) of character forms as a group.

## Bibliography

BOODBERG, Peter A.

1937 “Some Proleptical Remarks on the Evolution of Archaic Chinese.” *Harvard Journal of Asiatic Studies* 2: 329–372.

CHEN Mengjia 陳夢家

1966 “Dong Zhou mengshi yu chutu zaishu 東周盟誓與出土載書.” *Kaogu* (1966.5): 271–281.

GALAMBOS, Imre

2002 *The Evolution of Chinese Writing: Evidence from Newly Excavated Texts (490–221 BC)*. Berkeley: Ph.D. Dissertation.

HE Linyi 何琳儀

1998 *Zhanguo guwen zidian: Zhanguo wenzi shengxi* 戰國古文字典 – 戰國文字聲系. Beijing: Zhonghua shuju.

LEGGÉ, James

1960 *The Chinese Classics; With a Translation, Critical and Exegetical Notes, Prolegomena, and Copious Indexes*. 5 vols. Hong Kong: Hong Kong University Press [London: Oxford U.P., 1861–72].

LI Yumin 李裕民

1979 “Gu zi xin kao 古字新考.” *Guwenzi yanjiu* 10 (1983): 109–121.

QI Guiyan 戚桂宴

1979 “Ma yi fei shi jie 麻夷非是解.” *Kaogu* (1979.3): 272, 230.

SHANXI SHENG WENWU GONGZUO WEIYUANHUI 山西省文物工作委員會(ed.)

1976 *Houma mengshu* 侯馬盟書. Shanghai: Wenwu chubanshe.



WELD, SUSAN R.

1997 "The Covenant Texts from Houma and Wenxian." *New Sources of Early Chinese History: An Introduction to the Reading of Inscriptions and Manuscripts*. Shaughnessy, Edward L. (ed.) Berkeley: The Society for the Study of Early China and the Institute of East Asian Studies. 125–160.

ZHU Dexi 朱德熙; QIU Xigui 裘錫圭

1972 "Zhanguo wenzi yanjiu (liu zhong) 戰國文字研究 (六種)." *Kaogu xuebao* (1972.1): 73–89.