The Formation of *Hewen* 合文

Kulayb Razzak

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Arts

University of Washington

2020

Committee:

William Boltz

Zev Handel

Program Authorized to Offer Degree:

Asian Languages & Literature

University of Washington

Abstract

The Formation of *Hewen* 合文

Chair of the Supervisory Committee:
William Boltz
Department of Asian Languages & Literature

When one reads the literature on *hewen* 合文 (a pre-Han type of ligation formed from two or more Chinese characters) the majority of it will correctly state that these ”combined graphs” (as the term literally means) were used to save the scribe both space and effort. However, this does not tell us what a scribe considered before he combined graphs. Could all graphs potentially be formed into any type of *hewen*, whether adjunct, combination, or parturient? The answer to this question is ”no” when it comes to Chu orthography. So what were the principles behind the formation of *hewen*? Adjuncts were formed from any characters, as the components would only have to be squeezed together. Combinations would only be formed between characters that shared similar strokes, and what they shared could be as little as a single stroke or as much as an entire determiner. Parturients, being written as one standard character inside which the other could be found, did not require any special considerations. Whether a scribe ligated eligible graphs and how he chose to do so was a matter of personal choice, all things being equal. In other words, just because two graphs could be formed into a combination *hewen* does not mean that the scribe was obliged to make them into one; he could just as well form them into an adjunct or even do nothing to them. What served to reign in the whims of the individual was the system of principles undergirding the orthographic device of ligation.

Contents

1. The Formation of *Hewen* 合文

1.1 Introduction

The term *hewen* 合文 (’ligature, ligation’) refers to two or more graphs which have been combined into one. Chu 楚 manuscripts almost always have *hewen* marked with a symbol which resembles the modern ”equals” sign (=). This is called a *hewen hao* 合文號 (’ligated graph mark, ligature mark’), and it is typically placed to the bottom right of the ligature. Related to this is another category of ligated graphs called *chongwen* 重文 (’repeated graphs’), which refer to one or more graphs which have been given the same mark to indicate that they are to be duplicated when read. In other words, with *chongwen* the = functions like the Japanese ditto mark (々). 一 ₌ 人 ₌ and 君 ₌ are two examples of *chongwen* from the Shanghai Museum Chu manuscript collection, and they are read as 一人一人 and 君君 respectively. Standard *hewen* are treated as being distinct from *chongwen*, but because of the similarity between the two, it is worth comparing and contrasting them in this paper.

Much has already been written on the fact that the motivation behind the existence of *hewen* is to save the scribe both space and time, but little has been written on how they are formed. My thesis aims to lay out the principles by which this class of graphs was formed. Such discussions are necessary, as *hewen* seem arbitrarily formed when first encountered, and because they appear so rarely, it is difficult to grasp their underlying systematicity. It does exist, however, and since ligation is as much a part of the writing system as

anything else, it warrants a closer study.

And it is fortuitous that I am not the first person in the field of Sinology to have thought this. Scholars first became aware of *hewen* during the Song dynasty, and they have discussed them during every era since. Song scholar Zhao Yanwei 趙彥衛 (active in 1195), in his fifteen *juan* 卷 historiographical work *Yun lu man chao* 雲麓漫鈔 discusses an instance of *daifu* 大夫'grandee' written as one abbreviated graph on the Mount Yi stele[1]. The Qing scholars Sun Xingyan 孫星衍 (1753 –1818) and Sun Yirang 孫詒讓 (1848 –1908) performed a philological inquiry into the *hewen* 子孫 = that appears in bronze script. Additionally, besides secondary resources discussing *hewen*, we now also have the primary resources of ancient manuscripts themselves. In the wake of each new manuscript's discovery, scholars have also compiled accompanying *zibian* 字編, such as Sun Haibo's 孫海波 (1909 –1972) *Jiaguwen bian* 甲骨文編.

Hewen researchers generally divide their field of study into three areas of focus: Ligations that appear in the Shang oracle bones, the Zhou bronzes, and the Warring States silk bamboo manuscripts. The amount of research that has been done on ligations in each of these specific areas also corresponds to this ordering. That is to say, the greatest amount of work has been done with the oracle bones, an unsatisfactory amount with bronzes, and but a paltry amount with the silk and bamboo manuscripts; of note, however, is the fact that of all Warring States manuscripts, Chu manuscripts are the greatest in number[2]. The start of the Qin demarcates when ligatures began to disappear, although we do continue to see them into the Han dynasties. For example, we observe an example of *hewen* in the first two strips of an edict:

> 具上吏民壹切蒙恩勿治其罪者名，會今，罪別之，以贖行者，如詔書 = 到言.
>
> Assemble and submit the name-list of all officers and people who received clemency and were not punished for crimes. This is due now. Separate them according to crime and send it by courier. Carry

[1]That is, written as {夫 =}.
[2]Li Shoukui 李守奎 (2003) 12.

it out as in the edict document. When the document arrives, report.[3]

This document was sent to a military outpost in the Juyan region during the second year of the Shijianguo 始建國 period (10 AD).

In modern times, scholars not only discuss what *hewen* are, but how they are formed. Jin Xinxin 金欣欣 quotes a Chinese professor in his 2017 article on page 88 to explain that ligated graphs in the OBS were formed based on orthographic considerations:

> 写成合文的两个字, 有时可以共用偏旁或笔画。[...] 古文字中的‘合书’也是一种省减方式，这种方式是通过两字互借偏旁或笔划，再加上重文符号标示，以求得书写的简便。
>
> Two characters written together as a ligature may sometimes share a radical or some strokes. [...] The ’ligation’ seen in ancient graphs is also a method of economizing, and this method seeks to expedite writing via shared radicals or strokes and also the repeated graph mark.

On the whole, however, there is a dearth of explicit discussions on the principles of this orthographic phenomenon. It is my desire to join with those scholars who have done just this so as to help fill this need.

1.2 Research

The term *hewen* is a modern one and refers to the pre-Han practice of writing two or even three graphs as one, hence why this term literally translates to ”combined characters.” As we shall see, this character combining was achieved through one of three strategies:

1. by squeezing the graphs together into the space of one graph, but otherwise leaving them unmodified;

[3] Sanft 49.

2. by abbreviating one graph and combining it with another graph which shared the same strokes that were abbreviated, or by otherwise combining graphs where similarity of strokes between the two ensured that no graphical information was lost;

3. by wholly omitting one graph due to all its strokes being present in the other.

I will follow the naming conventions Li 2009 uses for each of these subtypes of *hewen*. Each of them are referred to as adjunct types, combination types, and parturient types respectively[4]. In the following pages, I will give examples of each to illustrate what exactly I mean by the above descriptions.

1.2.1 Adjunct *Hewen*

Adjunct *hewen* are formed in a way that is still familiar to us through the rules of modern Chinese orthography. The two component graphs of adjuncts are usually arranged one atop the other or sometimes beside each other. As an example of this, we can look at the first four adjuncts which appear in the *Chu xi jian bo wen* (CXJBW) *zibian* on pages 1261 & 1274:

Figure 1.1: 八曰：帛乙 3 (CXJBW)

Figure 1.2: 直（犆）牛：包二 222 (CXJBW)

As noted, the distinctive feature of adjunct *hewen* is that there are no shared strokes, as these examples show. Each component graph stands whole, and

[4]The Chinese names for these are *pinhe shi* 拼合式, *gongti shi* 共體式, and *baoyun shi* 包孕式.

Figure 1.3: 一夫：包二 3 (CXJBW)

Figure 1.4: 一月：帛乙 3 (CXJBW)

the only thing changed about them is their spatial relationship to each other, as they have been squeezed next to each other into one character space. This is why these types of ligated graphs are called adjuncts: One graph is written next to the other, rather than being made a part of it, as is done in the other two categories of *hewen*. Hence, it needs to be noted that they are not ligations in the strict sense, so this category will be treated as being distinct from the others, which are true ligations.

But with that said, figures 1.2 and 1.3 seem to be exceptions, since the two component graphs are technically connected. So have they not been combined? Let us note that this connection does not affect how the graphs are written in any way. The one point of intersection the components share in these tokens may as well have come about accidentally. The *hewen* which are ligations in the strict sense are those whose component graphs share strokes and can therefore be written into one graph to reduce the number of strokes needed to write either. Neither of these *hewen* have had the strokes of their component characters reduced.

1.2.2 Combination *Hewen*

I will next discuss combinations, which are our first example of true ligation. Combinations stand in between the previously discussed adjuncts and the more radically ligated parturients. The distinguishing feature of this *hewen* category is a shared stroke or even classifier between the two component

graphs.In other words, there are one or more strokes unambiguously contributing to the graphic structures of *both* graphs. Below are four examples taken from the SHMM *zibian* of combinations made by a single point of combination.

Figure 1.5: 上下：孔子詩論 4 (SHMM)

Figure 1.6: 上帝：東大王泊旱 6 (SHMM)

Figure 1.7: 少人（小人）：周易 8 (SHMM)

Figure 1.8: 之所：從甲 9 (SHMM)

These graphs demonstrate true ligation, i.e. ligation which makes meaningful use of shared strokes in order to merge one graph into another. I will also note that the stroke whereby the two component graphs are conjoined need not be identical to the stroke normally seen in either character; it need only be judged to be similar to it by the scribe.

Take, for example, figure 1.9. The 口 of {君} and the ”head” of {子} are not the same in Chu orthography:  and  . Nonetheless, they are structurally similar, and for that reason we find the word *junzi* 君子’gentleman’ formed into a combination *hewen*:

Figure 1.9: 君子：緇衣 16 (SHMM)

Besides the above examples, there are combinations which are made by nesting one character within another, so that their structures conjoin via more than one stroke.

Figure 1.10: 二十：包二 277 (CXJBW)

Figure 1.11: 教學：郭語叢 61 (CXJBW)

Figure 1.12: 里社：郭六德 22 (CXJBW)

Figure 1.13: 馬鹿：天策 (CXJBW)

It is interesting to note that the combination *hewen* of {二} and {十} (figure 1.10) has survived to the present day unchanged: 廿. That and figure 1.13 are the least complicated to unravel, as they have both been combined via sharing and extending multiple strokes. Figures 1.11 and 1.12 are more complicated, because it is in examples like these two that we begin to see just how much creative liberty a scribe was allowed when ligating characters.

We can see that the *jiao xue* 教學'teach and study' combination was formed by fusing the phonophoric of {教} together with {學}, a move made possible by the fact that in Chu orthography, *jiao* is written as  , and *xue* as  . They both share a 子 and have something written above it, and both these components are stacked in the ligature, with the 攴 still written off to the right.

Figure 1.12 is still more complicated and has proven difficult to identify. Lei Liming 雷黎明 2009 identifies it as writing *sheji* 社稷'grain and earth gods' altars'. I personally am not convinced of this, however, as the context in which this ligated graph appears in is:

> 上共下之宜 [義]，以 [土 + 八 + 介]◯，胃 [謂] 之孝... [5]
>
> The act of propriety for which those above cooperate with those below in order to...is referred to as filial piety...

[5]Cai Min 蔡敏, 1998, *Guodian Chu mu zhu jian* 郭店楚墓竹簡, 200.

Religious rites for the grain and earth gods do not have anything to do with filial piety and are therefore probably not what we should read into the *Liu de*'s 六德 definition for it, especially as it relates to those above and those below cooperating to achieve some end. The CXJBW alternatively glosses  as *li she* 里社, as I wrote above; both of these words are nouns that refer to an administrative unit of 25 households[6]. Although this gloss is at least closer to the semantic sphere of *xiao* 孝 due to indirectly involving families, I still find it unsatisfactory.

Besides semantic concerns, however, there is the composition of this graph to consider. The first detail that jumps out at us is the fact that the nested component graph does not look like any of the Chu graphs in the suggested glosses — neither 示 nor any subcomponent in 稷. *Hewen* may combine and modify components, but they may not discard components. The nested graph could conceivably be a 示, but if that is the case, it is an uncharacteristically poorly drawn one. The next detail to notice is that the subcomponent identified as 里 actually has a second stroke in the 田. We could attribute this to an error on the scribe's part, but that would make it an exceptionally cleanly done error. For these reasons, I do not find either of the proffered interpretations convincing. However, for our purposes the exact reading of this ligation is not important, so I have chosen to use the CXJBW *zibian*'s gloss.

The above concludes my discussion on combination *hewen*. This subtype refers to combined graphs whose ligation is done with either one or more strokes. When multiple strokes create the ligation, parts of a component graph may be modified or combined with those of the other component graph, but whole subcomponents may not be deleted.

1.2.3 Parturient *Hewen*

The final category of ligated graphs are those which exhibit the greatest amount of abbreviation. These graphs are the logical conclusion to ligation,

[6] "... 以二十五家為里，故知二十五家為社也。" *Ci yuan* 詞源, 1998, 1227.

because parturients are those *hewen* which completely omit a component graph due to it being fully represented in the other graph. Hence the name: One graph is ”pregnant” with the other.

Figure 1.14: 六馬：曾侯乙墓 171 (CXJBW)

Figure 1.15: 之時：容成氏 51 (SHMM)

Figure 1.16: 之志：李庚子問於孔子 7 (SHMM)

The {六} of figure 14 is actually written as {馬＋六} in the manuscript, so the second graph, *ma* 馬’horse’, is identified as being contained within the preceding number[7]. In figure 15, the {時} is written as 𠂤, with the top component graph being the combining form of the graph {之}. Finally, in figure 16, we see the same thing: I.e. that the top component graph was written as 屮 (之).

[7]Writing the number making up a team of horses with the horse semantophoric and a number is a common pre-Qin orthographic practice. Another example of this is *si* 駟’four (horses)’.

1.2.4 *Chongwen* 重文

With the above examples in mind, I would like to now talk about *chongwen* 重文 (lit. ”repeated graphs”). *Chongwen*, unlike most *hewen*, write the same graph twice and can be used with phrasal units. These are treated as distinct from the graphs that I have heretofore discussed[8]. I would like to suggest in this thesis that they are actually a distinct subcategory of parturients. For now, though, let us first get acquainted with *chongwen*. Towards that end, I will be presenting them in accordance with how they are typically discussed in academia.

Figure 1.17: 出之出之：內禮 6 (SHMM)

Figure 1.18: 行行：緇衣 16 (SHMM)

Figure 17 is interesting for being a *hewen chongwen*: The repeated graph is the Chu graph for 出 and here is a parturient; the Chu graph consists of {之+止}, written from top to bottom respectively. What we have here, then, is 出 ₌ 出 ₌, which in turn is read as 出之出之.

[8]The SHMM *zibian*, Lei 2009, and Zhang 2015 all categorize *chongwen* distinctly from *hewen*.

The next graph worth looking closer at is figure 19.

Figure 1.19: 一人一人：從政甲 3 (SHMM)

Although there is nothing special about the individual graphs here, this *chongwen* actually shows us that there is more than one way to read these special graphs. After learning about figure 17, you might expect this one to be read as 一一人人, but the whole counter phrase itself is repeated, yielding 一人一人. In other words, the *chongwen hao* 重文號[9] only indicates that repetition has happened and nothing more. How exactly to read the repetition of graphs will be context dependent. This makes *chongwen* similar to parturients, so it is worth comparing them against each other to understand why *chongwen* are distinct from *hewen*.

Recall that parturients are defined as those ligatures which fully omit another character due to its strokes being fully present within another one that is written (e.g. writing 夫 = for 大夫). Now with this definition in mind, consider the following examples of *chongwen*:

Figure 1.20: 君君：昔者君老 1 (SHMM)

[9]Just as the = mark which indicates a ligature is called a *hewen hao*, the mark which indicates repetition is called a *chongwen hao* in the literature.

Figure 1.21: 善才善才：三德 5 (SHMM)

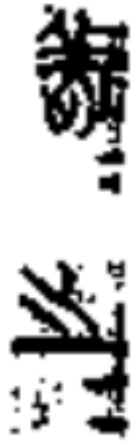

Figure 1.22: 敬之敬之：三德 3 (SHMM)

Above we have three different examples of *chongwen*. This lineup was chosen to show that a variety of morphemes may be repeated, whether they be noun or verb phrases. Each graph is marked individually with a *chongwen hao*, even in phrases. This mark is exactly the same as the *hewen hao* 合文號. This is because = generally served as a mark of abbreviation in Chu manuscripts, whether that abbreviation was for an entire character or for only a part of one[10]. Another point to notice is that when more than one character comprises a *chongwen* unit, its elements are never read in an AABB pattern, but instead in an ABAB pattern.

Chongwen being duplicates means that although the logic behind their formation is similar to the logic behind parturients, the details ultimately make them distinct from *hewen*, because *hewen* are a class of abbreviations formed on the basis of graphic shape. What this means is that parturients can only be

[10]Zhang 2015

formed when one graph has the same number of strokes in the same configuration as another graph which is nonetheless *distinct*. On the other hand, *chongwen* can always be formed from any characters, because once you have written a character, you can simply add a ditto mark to it. No graphic modification or consideration of the writing system's principles is necessary. Thus, although *chongwen* are similar to parturients, there are real, practical differences between the two which justify recognizing *chongwen* as distinct in academic discussion. Just like how all squares are quadrilaterals, but not all quadrilaterals are squares, so too is it the case that all *chongwen* are parturients, but not all parturients are *chongwen*.

2. Research Findings

2.1 Observations on Usage

Now that the key terms and concepts have been established, I can move into the primary discussion of my thesis: The principles by which Chu 楚 scribes formed *hewen* 合文. This topic concerns itself with the ”how” of *hewen*: How were they formed? Discussions of this sort are more difficult to find than those on the ”why” of *hewen*: Why were they formed? Lei Liming 雷黎明 2009, for example, writes on the ”why” of *hewen*, stating:

> 楚簡合文興盛的主要原因和重文一樣，都是書者趨簡心理的驅使所致。不管二字合寫是否真的減省筆畫，楚簡書寫者主觀上認爲，將二字合寫，總會在一定程度上減少筆畫，於是，合文頻頻出現。
>
> The primary reason for the rise in popularity of ligation in Chu bamboo strips is the same as that for repeated graphs: That is, it is due to the scribe’s psychological drive for simplification. Regardless of whether the two graphs’ being written together actually reduced the strokes, it was the subjective feelings of the one writing the Chu bamboo strips that by ligating the two graphs, he was over all reducing the strokes to a certain degree. As such, ligatures appeared frequently.

In brief, ligation was done in order to economize effort. Although this scholar also describes the different *hewen* categories in this article, he does not delve into the principles behind their formation. We might contest the claim that

ligations which do not result in less strokes were nevertheless considered by the Chu scribes ”to save on strokes to a certain degree” (在一定程度上減少筆畫), the over all explanation is sound. And as I will mention in more detail later in this thesis, Wang 2017 also talks about economizing time and effort with ligation.

But suppose we take this ”how” explanation and make it into a ”why” explanation. Perhaps *hewen* were employed whenever a scribe felt short on room and wished to ensure that he could fit his writing into one strip of bamboo. If that were the case, we could expect to see adjuncts located terminally on the bamboo strips, since these types of *hewen* could be freely formed and thus employed when one found oneself short on room.

In order to test this hypothesis, I decided to look at all the adjuncts given in the Shanghai Museum Manuscripts *zibian* 字編. This *zibian* covers five of the seven volumes of the manuscript collections which the Shanghai Museum has published. This choice is due to the fact that I have complete access to these manuscripts. My findings for the placement of these adjuncts is listed in the table below.

Adjunct	Source	Position
日月	III 中 19	Medial
古之	III 中 21	Medial
一日	III 中 2	Medial
一人	IV 曹 26	**Terminal**
少人（小人）	V 李 7	Medial
君子	I 孔 12	Medial
君子	I 性 28	Medial
君子	II 從甲 4	Medial
君子	II 從乙 5	**Initial**
君子	V 李 1	Medial

The source is given in the following format: Volume number, initial character of manuscript title, strip number. Because there are so many tokens of adjunct

{君子} (18 all total), I have only listed one token from each manuscript[1]. There was a total of 23 adjunct tokens, and as can be seen above, only two of them occurred in a non-medial position. We must reject the idea, then, that *hewen* were only a tool for squeezing more text into less space.

That being the case, let us consider an alternative explanation. Wang 2017 provides us with one in his article on oracle bone script (OBS) ligation:

> 首先，合文是出现在文字系统还没有稳定的时期，造字体系也还未完善因而容易出现两字合并为合字的现象 [...] 其次，合字往往是出现在人名、数量、时间这样相对固定的特殊词汇之中，在生产力水平较低的殷商时代，甲骨文的制作较为困难，为了方便制作，工匠们便把这些卜辞之中常出现的固定名词进行合文，从而降低契刻难度。最后，我们也不能排除“合文”中暗含了殷商先民造字过程中“会意法”的思维雏形。
>
> First, ligatures appeared during a time when the writing system was still unstable. Character formation had also yet to be perfected, making it easy for the phenomenon of two characters being combined into a ligature to occur [...] Ligated characters often happened with names, numbers, times, and other such relatively set phrases. During the Yin Shang 殷商 period when the level of productivity was comparatively low, production of oracle bone writing was rather difficult, so in order to simplify the process, craftsmen ligated these commonly seen set phrases in divination, thereby reducing the difficulty of carving. Finally, we cannot disregard the hints we see in the ligatures of the Yin Shang people's rudimentary understanding of *huiyi* 會意 character formation.

Once again, the claim is that *hewen* were used to simplify writing, although importantly, this scholar adds the qualification that it was high frequency technical words and phrases that were ligated. And indeed, names for royal ancestors total to a third (127) of the ligations Bao Huifang 暴慧芳 found in his

[1] If there had been any tokens appearing in a non-medial position, I would have listed those as well.

analysis of 376 OBS *hewen* samples[2]. The next greatest category of ligations found was phrases of quantity at 70 tokens, followed by words for time at 55[3].

So let us ask if this explanation might apply to the Warring States corpus. Perhaps it was technical vocabulary that was frequently made into *hewen*. According to the analysis that Bao Huifang performed on a sample of 95 *hewen* taken from Warring States silk and bamboo manuscripts, most of the words ligated are nouns (66.3%), and the majority of them are ordinary nouns (38 of the 63)[4]. In contrast, only 17 specialized nouns were found[5].

And upon taking a more detailed look into our own Chu data, we find results contrary to this supposition. To investigate whether the semantics of the written words might influence the decision to ligate or not, I selected three words to look for across the 28 different manuscripts provided in the 5 SHMM volumes: *Kongzi* 孔子'Confucius', *junzi* 君子'gentleman', and *xiaoren* 小人'petty man'. These three words were chosen on the assumption that they would be both frequent and specialized enough (in the case of the second and third words) to allow us to answer this question. What I found is that only Confucius's name was the only one that is always ligated. In the case of *junzi* and *xiaoren*, the non-ligated forms occur the most — 35 to 24 and 4 to 1, respectively. It would seem then that we cannot say semantics dictated which words were made into ligatures and which were not.

This result should not be surprising, however, as centuries of time separate the OBS and Warring States corpuses. The scribes, culture, and even the language are all different. But if these hypotheses for how Chu scribes decided to ligate graphs do not explain what we observe, then let us now turn to the proposal that it was the graphic shape of the characters which was of principal importance.

[2]Bao 12.
[3]Bao 12.
[4]Bao 46.
[5]Bao 46.

2.2 Explanation for Our Data

In order to evaluate the hypothesis that graphic shapes were what determined whether a scribe chose to ligate, we should read what Jin 2017 writes on OBS ligation so as to understand what we mean by this claim. He states:

> 写成合文的两个字, 有时可以共用偏旁或笔画。[...] 古文字中的‘合书’也是一种省减方式，这种方式是通过两字互借偏旁或笔划，再加上重文符号标示，以求得书写的简便。
>
> Two characters written together as a ligature may sometimes share a radical or some strokes. [...] The ’ligation’ seen in ancient graphs is also a method of economizing, and this method seeks to expedite writing via shared radicals or strokes and also the repeated graph mark.

That is to say, *hewen* served to reduce strokes by means of shared components between the graphs being ligated. This specification of the means by which they operated is the crucial detail which gives the hypothesis predictive power. Because by definition, combinations and parturients (the true ligatures) can only be formed between graphs whose components are either partially or fully shared, if the data supports this hypothesis, then the characters composing adjuncts should not have any components amenable to being combined or fused into each other.

The SHMM has only five adjuncts, so I will first list each of these out for investigation. The CXJBW, however, has 41 adjuncts in it, so I will only consider ten from it, taking care not to select any that are also found in the other *zibian*. The purpose of listing out all of these characters is to examine each one and offer an orthographically motivated explanation for why each has been combined in the manner we observe.

2.2.1 Shanghai Museum Manuscript Adjuncts

Going in the order in which the graphs are sorted in the *zibian* itself, the first *hewen* we shall look at is that for the phrase *yi ren* 一人'one person'.

Figure 2.1: 一人：曹沫之陳 26 (SHMM)

Upon seeing the Chu graph for {人}, one might ask why it is that scribes did not simply straighten out one of the strokes to represent {一} and make a combination *hewen*. The answer to that is that this would be a non-standard way of forming characters. Just as is the case today, there were conventions that scribes followed in order for their writing to be regarded as finely done (to say nothing of being legible). Despite there being more ways in which a word could be written, that does not mean that there was not a well-defined scope for what could be written, as this thesis seeks to show. Ligation did not give a scribe license to write however he pleased, so long as the resulting ligature still looked like a Chinese character. A {人} with a straightened out 丿 stroke just looks like a poorly written {人}. Compare the above graph with this ligation for {junzi} 君子'gentleman' that we have seen before: At first, one

Figure 2.2: 君子：緇衣 16 (SHMM)

might not understand why this ligation could be formed when the two component graphs do not share any identical strokes between themselves. As was already noted, just as is the case in the writing system today, the 

and  graphs, despite looking like they have similar subcomponents, do not, strictly speaking, share any graphical structures. However, shared structures being present is not *per se* important, but rather is Important because that fact allows for the preservation of graphical information. A graph needs to be recognizable and readable, which in turn naturally entails that it be considered well-formed. In the case of this *junzi hewen*, the Chu writing system allowed the two structures to be formed into a combination due to both requiring a ”circular”[6] shape in them. Furthermore, there were not yet exacting standards regarding how a character should look as there are today (e.g. as with the difference between a 丿 and a 丨 stroke), so there was a little bit more margin for liberality in orthography. The above combination is aesthetically pleasing for conforming to what one would expect a proper Chinese character to look like, and it is also orthographically well made due to utilizing a shared shape without adding or subtracting anything from either character. Ligations had to be formed within the boundaries of the writing system.

This same principle we just established applies to the next *hewen* that we will look at.

Figure 2.3: 一日：中弓 24 (SHMM)

As discussed above, it would not be acceptable to combine the middle stroke of {日} with {一}, as this would result in a poorly formed ligation, one whose graphs are indistinctly and improperly put together — there does not exist a character which looks like a {日} but with an extended middle stroke. Hence, they can only be combined into an adjunct.

[6]Let us not get into a geometry debate due to the fact that one graph is written with a more oval-like shape. Such minutiae are trivial for the discussion at hand.

The third *hewen* that we will look at is a variant of a ligation we have seen before.

Figure 2.4: 尛人（小人）：李庚子問於孔子 7 (SHMM)

When the *xiaoren* 少人'petty person' ligation was first introduced, it appeared as a combination. Now we see that scribes could also choose to write it as an adjunct. The existence of this *hewen* variant is problematic for the stroke reduction hypothesis, because here we have evidence that scribes could and would choose to form a *hewen* without reducing even one stroke.

On the other hand, it poses no problem if we insist that *hewen* also followed the same principles of orthography as other characters. There was significantly more character variation in the Warring States era than there is today, after all; that we should see variation among *hewen* as well is not surprising. And here, we can see that the scribe perfectly formed each of the component characters. The 尛 and the 人 are distinct from each other, as we would expect of a *hewen* consisting of two graphs which have no graphic similarities.

There exists another adjunct variant for *xiaoren*, only this one uses the same form of the graph 少 as was previously discussed.

Figure 2.5: 少人（小人）：包山楚墓 125 (CXJBW)

The two graphs are stacked one atop the other, and the final, bottom stroke of {少} is facing to the left, the same direction as the "arm" of {人}. Yet no combination takes place, and the graphs are left simply squeezed together. In

the SHMM, there is an even distribution of three adjunct variants and three combination variants. Across all the manuscripts covered by the CXJBW, however, there are twice as many combination variants as there are adjuncts (fourteen and seven). These adjunct forms have the advantage of being more readabe than their combination variants. This was likely the strongest motivation behind forming them.

Appropriately and interestingly enough, the other *hewen* which shows this categorical variance in its form is *xiaoren*'s antonym: *junzi* 君子'gentleman'.

Figure 2.6: 君子：孔子詩論 12 (SHMM)

The first *hewen* shown is the more commonly encountered one in the Shanghai Museum manuscripts. {子} is squeezed in beside the {口}.

Figure 2.7: 君子：緇衣 16 (SHMM)

And we can observe that for the second *hewen*, the combination variant, The "head" of {子} has instead taken the place of the {口}, something that is made possible by the fact that they share similar strokes in the orthography, thus ensuring that there is no loss of graphic information.

When we look at the raw numbers for this *hewen*, however, we find a curious situation in regards to the data distribution. Each one of the nine adjunct variant tokens found in the CXJBW comes from the Shanghai Museum manuscripts, and all 17 of the combination variants come from the rest of the

Chu corpus. When we in turn look at the SHMM, we find that there are only six combination variants, while there are 26 adjuncts. Another rarity the SHMM records is this adjunct *junzi hewen*:

Figure 2.8: 君子：緇衣 3 (SHMM)

This variant in which the characters are stacked one atop the other shows up no where else in our current Chu corpus (as defined by the CXJBW). What might be the motivation for ligating the two component characters like this? Having a second type of adjunct variant suggests to me that Chu scribes preferred being able to form ligatures into adjuncts, even when combining the component graphs was common practice. Stacking them like this complies even more nicely with the orthography, so it is not a complete mystery as to why these combinations exist. And as previously noted, adjuncts are easier to read. The existence of these 君子 and 小人 combinations is perhaps a result of the tension between the orthographic considerations unique to *hewen* and the more general ones.

Finally, the last adjunct *hewen* is refreshingly simple to present.

Figure 2.9: 古之：中弓 21 (SHMM)

Here we see that the top graph has been nested inside the lower one. {之} normally looks like this:  . The two graphs have no strokes that work well with each other — that is, there are no shared or complementary ones. {之} consists entirely of straight lines, only one of which is not at a diagonal angle.

{古}, on the other hand, has curves and straight lines. As we would expect, then, the only way these two characters can be ligated is by simply squeezing them together into a tighter space.

2.2.2 *Chu xi jian bo wen* 楚系簡帛文 Adjuncts

I wish to now turn my attention to the other *zibian* to consider four more adjuncts, so as to supplement my discussion with *hewen* from other manuscripts. Although the first one was selected to further reinforce a point, the other three serve as new examples to consider. After this section, I will have presented all the prerequisite information for the conclusions that I will draw.

As already stated, the first character for discussion is one that we have already seen before. This time, however, more will be said about its specific graphic configuration.

Figure 2.10: 一夫：包二 3 (CXJBW)

Once more, we see that {一} cannot be fused into other strokes. Again, we see the formation principle at work that the constituent graphs must remain readily identifiable by following the contemporary rules of orthography. Shared strokes are not the only factor that was considered. Since there is no other way for these graphs to be ligated together, the scribe simply squeezed them together into one.

Now, the second character we are looking at is a rare example due to consisting of three characters instead of the more standard two.
There are two details that should be pointed out. The first and most notable one is that the *hewen hao* 合文號 is not placed in the lower right corner as is standard, but in the middle of the ligature. This is undoubtedly because of the

Figure 2.11: 一十二：仰二五 21 (CXJBW)

strokes of {二} occupying that space. The second detail is that the graphs do not have their order changed from what it would be were they not ligated (i.e. top down, one after the other). This detail is in fact commonly seen in all adjuncts[7].

The next adjunct is another interesting one because it can serve as a good demonstration of the principles presented herein.

Figure 2.12: 七十：唐 26 (CXJBW)

Once again, how {一} is handled proves to be informative. Chu orthography allows the horizontal stroke of *shi* 十'ten' to be made either as a horizontal stroke  or as a dot  . We already know not to expect the full horizontal stroke form to be used in a ligature with *qi* 七'seven', since it was written the same way (), but because these two graphs are so similar, they cannot be made into adjuncts, and as such, are written together as a combination ligature. We have two tokens of this. The first is the above character, and the second is written thus:  . It should be apparent now why the other ligature is the one that I have chosen to focus on: The scribe modified the *shi* by transforming the horizontal stroke into a curved one. This is similar to the variation we saw with *xiaoren* and *junzi*. One was written as *hewen* conventions dictated, while the other was tweaked in order to facilitate

[7]Of course, that is not to say that it is a rule, however

recognizing the components. And if facilitating readability were a legitimate concern Chu scribes had, then it give us a third pressure, which would help explain these variants.

But we need to ask if curving the middle stroke of {十} is orthographically sound, as I am suggesting that all *hewen* must be. I would argue that it is, because we do have two other examples of this character being written with a curved stroke within the Chu corpus, in the SHMM *zibian*. Both come from the *Rongcheng shi* 容成氏 manuscript and are written as follows:  , and  . Although these curves are upward facing instead of downward facing like in the *hewen*, I would say that the direction of the curved stroke is unimportant, since in both cases we have a variant way of realizing the same stroke. The point is that writing {十} with a curved stroke was acceptable; hence, the *hewen* under consideration can still be said to have been formed in compliance with the rules of both ligation and standard orthography.

The final ligature we will look at is being presented here in order to further comment on what it means for a *hewen* to be properly formed.

Figure 2.13: 之日：望一 156 (CXJBW)

As anyone who is familiar with premodern orthography will know, this graphic shape is one way to write the word *shi* 時'time', so this could potentially be writing the phrase *zhi shi* 之時'time of'. This ligation can therefore be said to be ambiguous. It is only the context for this particular graph[8] which makes it clear that it is to be read as *zhi ri* 之日'day of'. So being well formed and legible does not preclude one *hewen* having the same shape as another graph. When I bring up these evaluations, I am referring exclusively to how well the graph complies with the contemporary orthography.

[8]"辛未～(Teng 1265)."

2.3 Preliminary Conclusions

So far I have defined the key concepts of my research and presented the data that I have looked at. Now it is time to draw conclusions about that data. This section will bring all my explanations and suppositions together into one place to conclude with why an internally based explanation for *hewen* formation is the most feasible.

Recall that the reason adjuncts became central to this thesis is because they are the only category of ligatures which might allow us to determine whether scribes were motivated by factors external or internal to the writing system when they ligated characters. This is because adjuncts are formed by squeezing characters together, potentially allowing scribes to form *hewen* out of any two characters they want, however they want. If ligation is only a tool for convenience, then we should not expect to see any systematicity beyond what the more standard writing system provides.

One need only look back over all the adjuncts I have listed here to realize that this is not what we see in the Chu corpus. There are six adjuncts in the SHMM, all of which are formed using a top-down configuration. There are 37 adjuncts in the CXJBW, and only twelve of them use a configuration other than top-down. It seems safe to conclude, then, that the top-down configuration of graphs is standard for adjunct ligatures. But then how do we explain the twelve exceptions to this? That matter is actually simple, because it relates to a previously mentioned observation: When a certain configuration would result in a more standard arrangement of graphs, then that arrangement is to be preferred.

Let us take a look now at three of those twelve exceptions. The first one is noteworthy, because at first glance it seems to be a top-down adjunct, but once one gives it a second look, one then notices that the two graphs have been all but combined into one.

As can be seen, it would not be correct to describe this ligature as merely

Figure 2.14: 八十：包二 140, verso (CXJBW)

being a top-down adjunct. The {八} has been placed squarely within the space that the {十} is occupying. This was most likely done because this configuration looks better than having {八}, with its central white space, rest above {十}, with its lateral white space. In other words, this combination was likely done out of a sense of orthographic aesthetic.

The next character represents a subcategory of adjuncts that we can observe, which are composed of horizontally placed characters.

Figure 2.15: 夫人：新甲三 176 (CXJBW)

Unlike with the previous example, the component characters could be organized vertically and look pleasing enough to the eye, yet the scribe chose to place {人} in a position more natural to it — that is, off to the left. This arrangement is preferable even to our modern character creation sensibilities. This decision not only does nothing to save on effort, but also goes against the typical placement of an adjunct's compositional characters. This ligature strongly suggests that orthographic considerations were indeed important in the formation of *hewen*.

Finally, the third subcategory of adjuncts is those which nest one graph within another. And again, the principles of character formation motivating these kinds of ligatures are still present today.

Here we can see that {犬} has been placed under the *yi* 邑 variant's 宀 and to

Figure 2.16: 邑犬：包二 145 (CXJBW)

the right. The only discernible reason this arrangement might be chosen over simply stacking the characters atop each other is to abide by character formation principles. To explain in more detail: The 宀 component allows multiple characters to be arranged underneath it, making it a more proper ”home” for a sub-graph than having it float in nearby space outside it. No effort is saved by ligating the characters this way, but it is more aesthetically pleasing as far as the Chinese writing system is concerned. Hence why we see this ligature in its current composition.

So, what can we conclude after this investigation into adjuncts? Clearly we must reject any explanation that is external to the writing system *per se*. The above exceptions to the principle that adjuncts are formed with vertical placement admit no other explanation for their existence than orthographic considerations. The many examples we have seen with {一} also support this, because we do not see characters modified to have lengthened horizontal strokes to form combinations with it. The reason for this must be that doing so is inadmissible by Chu orthographic standards. These would not be well formed Chinese characters, even though they would be more convenient for the scribe. Although ligation allows for graphs to be written in ways that they would not otherwise be written, these modifications are kept to a minimum so as to preserve a character’s graphic shape as much as possible, and they are done as the writing system allows. *Hewen* do not exist outside of the standard writing system, but instead are a sub-system within it.

But how do we account for the variants noted above for *junzi* and *xiaoren*? The existence of both a combination variant and an adjunct variant is a perfect example of the tension at play between seeking to form true ligations (thereby

saving effort and space) and seeking to adhere to the conventions of Chinese writing. Consider that {君} itself consists of two components arranged in a nested vertical layout. Putting another graph into that space to the right of {口} is doing nothing more than utilizing a standard operation for character formation. This arrangement of the *hewen*'s three graphical components is the most optimal one, because it places the {子} at the intersection of spaces where both {君} and {口} can have a component added. This being the best organization is supported by the fact that the variant which simply places {子} vertically underneath {君} is barely attested. On the other hand, with *xiaoren*, {少} can take a graph underneath it. And concerning the variant with {⺌}, since there was no strict codification for writing as there is now, the scribe was free to choose between the options available to him. That is all that has happened.

Hewen are methods of writing characters into the space of one character — metacharacters, if you will. This differentiates them from *chongwen*, which are instead dittos. *Chongwen* are formed the same way as parturient *hewen*, but parturients never consist of two characters that are exactly the same. In conclusion, adjuncts show us that *hewen* are as much a part of the Chinese writing system as the more standard graphs are. They have their graphic components arranged in the same way as other graphs, but unlike those other graphs, neither semantic nor phonetic considerations play a role in their formation.

3. Comparison with Other Scripts

As previously noted, the textual phenomenon of *hewen* 合文 can be observed across the span of time starting from the first confirmed instance of Chinese writing all the way to the Han dynasty. Scribes used different scripts on different writing media throughout this span of time, and the writing practices associated with each medium (bone, bronze, and bamboo) naturally also influenced how *hewen* were formed. Despite this variation, the ligations found in each medium are more similar to each other than they are different. This, of course, is due to the fact that they are all part of the same writing tradition. Nonetheless, there are unique differences that can be observed, so it is worth looking at the *hewen* of bronze script and oracle bone script (OBS) as well in order to better understand what distinguishes Chu 楚 ligations in the history of the Chinese writing system.

3.1 Oracle Bone and Bronze Script Data

The first detail we shall discuss is the spatial composition of the ligations. The majority of the ones found in the OBS, bronze script, and the Warring States writing are composed of two characters, according to the findings of Lee and Kim (2012). They account for 93.74% of the OBS *hewen* in the *Xiao zheng jiabuwen bian* 校正甲骨文編[1], 91.82% of those in the bronze script[2], and 97.42% of the *hewen* found in the *Zhanguo guwen zidian* 戰國古文字典[3]. In every case,

[1]There are 363 total ligations that they looked at (Lee & Kim 90).
[2]There are 110 total ligations that they looked at (Lee & Kim 90).
[3]There are 233 total ligations that they looked at.

two character ligations are over 90% of those present. Three character *hewen*, in contrast, are less than 10% of ligations present in the sources the researchers looked at. In the OBS, bronze script, and Warring States writing, they account for: 5.99%, 6.36%, and 2.58%, respectively.

Lee & Kim (2012) suggest on page 91 that the reason for this is simply due to the fact that the contemporary languages had more commonly used disyllabic lexemes than it did trisyllabic ones. If a scribe wished to ligate commonly used special vocabulary, then it makes sense that rare trisyllabic morphemes are a minority.

Next, we can broadly divide *hewen* into three types based on how their components are arranged next to each other: Left-right, up-down, and inclusionary. Within each of these character arrangement categories, the reverse orientation is also possible, i.e. right-left and down-up. What Lee & Kim call ”inclusionary” refers to not only those ligations which have one component nested inside another[4], but also *hewen* constructions such as  (OBS 五千), which are similar to combination *hewen* but differ due to not being formed with shared strokes. Instead, combinations proper are classified by the researches under two different categories: Shared strokes and shared radicals[5].

Left-right oriented two character *hewen* are 42.78%, 7.27%, and 15.88% of the OBS, bronze script, and Warring States writing, respectively[6]. Only the OBS has any trisyllabic phrases in this category, and they consist of just 0.54% of the corpus[7]. Right-left oriented ones are 9.54% of OBS *hewen*, 0.91% of bronze script *hewen*, and 8.15% of Warring States writing *hewen*[8]

Up-down oriented disyllabic lexemes present an unusual pattern when we look at them across these three mediums. Although in the OBS, they make up less than half of the *hewen* with their 38.69% presence[9], they become the

[4]E.g. as 甫 is nested inside 勹 in the modern character 葡.
[5]*Gongxiang bihua* 共享筆畫 and *gongxiang pianpang* 偏旁, respectively.
[6]Lee & Kim 92.
[7]Lee & Kim 92.
[8]Lee & Kim 94.
[9]Lee & Kim 91.

majority when we look at the other two scripts: A whole 80.91% in the bronze inscriptions, and 59.23% in the Warring States texts[10]. And as expected, trisyllabic words or phrases written this way are the minority: 0.54% of the OBS *hewen*, 5.45% of the bronze script, and 2.58% of the Warring States writing. Ligations that are read inversely account for 1.9% of the OBS ones and 0.42% of the Warring States writing ones[11].

Inclusionary *hewen* are the rarest orthographic orientation. They only occur with disyllabic lexemes and account for 2.45% of OBS *hewen*, 6.36% of bronze script *hewen*, and 8.15% of the Warring States *hewen*[12]. Those that have nested components and are read from the outside to the inside are 0.81%, 4.55%, and 5.15% of the OBS, bronze script, and Warring States writing *hewen*, respectively[13]. Those read inversely, from the inside to the outside, are 1.63%, 1.82%, and 3% of them[14].

The last category of graphs to compare is combinations. Those formed only through the use of shared strokes (e.g. OBS *shang xia* ) account for 11.82% and 2.14% of the bronze script and Warring States writing ligations. Combinations formed with shared radicals account for 2.73% and 6.43% of them. This category has no significant presence in the OBS[15].

3.2 Comparison against Chu Manuscript Data

We should now ask how this data compares to our own. The biggest discrepancy in trends across all three orthographies is that adjunct *hewen* are the majority, even in the Warring States writing sample. This is because the researchers were making use of more than Chu manuscripts, as I am limiting myself to here. The next major deviation from what we observe in the Chu manuscripts is that there is much more variation in how the graphs may be

[10]Lee & Kim 91.
[11]Lee & Kim 94.
[12]Lee & Kim 92.
[13]Lee & Kim 95.
[14]Lee & Kim 95.
[15]Lee & Kim 93.

arranged in the ligation. For example, we find in the OBS the graph (*wushi* 五十'fifty'), which is read from bottom to top instead of the more common top to bottom reading direction. The more common orientation is what we find in the Chu manuscripts . Indeed, we can say that there are few exceptions to the statement that there is little variation in the orthographic shape of Chu *hewen*. Aside from having more variation in arrangement, OBS orthography also allowed for more freedom in where component graphs could be placed.

Figure 3.1: 大丁：32385 (Bao 7)

Take this character for example: {丁} has been placed in the upper left corner. This is another orthographic decision that is inadmissible by Chu orthographic standards. Since writing still had little in the way of systemiation at this time, we naturally see the same variation in the *hewen* that we do in the more typical characters.

In contrast, the orthography used for bronzes reached a level that we can call systematized. Bronze script orthography demanded that graphs be placed either top, bottom, left, or right, but it still allowed more freedom in how combination *hewen* could be formed than was allowed in Warring States silk and bamboo manuscripts.

Figure 3.2: 二月：芹伯簋（Bao 26)

This graph is a good example of this. The {月} has been modified so that {二} can have its bottom stroke serve as one of its own strokes. This sort of modification is likely not admissible in Chu manuscripts, because as I have

already discussed, forming a combination by straightening out one of the component graph's strokes is not something we observe, as that was judged to violate the integrity of the graph. Another major difference is that 一 strokes could in fact be used to make a combination.

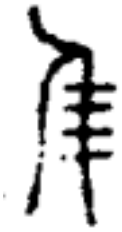

Figure 3.3: 四千：叔尸鎛 (Bao 25)

Here we can see that the {四} has been written onto the "body" of the other graph. We have already discussed at length how this is not possible in the Chu writing tradition, so that discussion will not be revisited.

Another difference one will be quick to notice is that there are no ligation marks on either the OBS or bronze characters. Although the = mark is sometimes observed in the OBS corpus, that is not a ligation mark, but a repetition mark[16].  (OBS 又 =) is therefore to be read as "*you you*." It is only in the Spring and Autumn period that we begin to see this mark used in the bronze script[17], and it is not used consistently except in the Warring States bamboo and silk manuscripts.

Of course, the Chu manuscripts are not without their own idiosyncrasies. The Spring and Autumn period brought with it not only the practice of consistently using the ligation mark, but also the practice of abbreviating characters[18], and we observe character abbreviation also occurring within contemporary *hewen*. Consider these two ligations which each have {馬} as a component character:  (*ma lu* 馬鹿 'horses and deer') and  (*cheng ma* 乘馬 'to ride a horse'). The former shows the full shape of the graph, and the latter shows the abbreviated shape. Such abbreviated allographs are not

[16]Bao 13.
[17]Bao 26.
[18]Bao 49.

uncommon. Furthermore, we find in the Chu corpus that non-adjuncts are the majority — they constitute 85.7% of the *Shanghai Museum Manuscript zibian* 字編 and 59% of the *Chu xi jianbo wen* 楚系簡帛文 *zibian*. This is due in part to another innovation of the Warring States period: Parturients, those *hewen* which are formed by virtue of one graph already being wholly present in another.

This chapter's purpose was to put the Chu orthographic practices relating to the formation of *hewen* into consideration from a historical perspective by looking briefly at the *hewen* of other orthographies (i.e. the oracle bone script and the bronze script). In doing so, we can see that while each orthography can be uniquely described, they all largely share the same principles of character formation. *Hewen*, like all Chinese characters, are formed according to the prevailing orthographic rules of the scribe's time.

4. Conclusion

> 子以四教：文，行，忠，信。
>
> The Master taught four things: Composition, etiquette, devotion, and trustworthiness.
>
> — *Lunyu* 論語 7.25

When one reads the literature on *hewen* 合文, the majority of it will correctly say that they were used to save on space and effort, but this does not tell us what a scribe considered before he combined two graphs. Could all graphs be combined? And could all graphs potentially be formed into any type of *hewen*? Clearly the answer to both these questions is ”no” when it comes to Chu orthography. So what were the principles behind the formation of *hewen*?

When a Chu scribe decided that he wanted to ligate two (or three) graphs, he then had pressure from two different systems acting on him: One being the more typical orthography, and the other being *hewen* orthography. The kind of ligation being done would determine which set of rules would be more relevant. Adjuncts were formed according to the more standard orthography, and combinations would be formed according to *hewen* orthography. Parturients, being written as only one standard character, did not have any special considerations for the scribe to take into account.

Combinations would only be formed between characters that shared similar strokes. The strokes shared could be as little as a single one or a whole determiner. When combining two characters together, the scribe was allowed

to either omit strokes that were *already represented* by one of the characters or modify strokes. He was not allowed to subtract strokes in order to form a combination, nor was he allowed to form a combination that went against the rules for how characters may normally be written together. These rules were active within the framework of the broader writing system, meaning that the resulting *hewen* would still have to be deemed acceptable per the rules of the more standard orthography. This is to say that ligation was not a license to write malformed characters[1]. Once this was done, a *hewen hao* 合文號 was added to the ligation in order to indicate to the reader that he was looking at a special character.

We can now understand why we see the graphic forms that we do. In this

Figure 4.1: 上帝：東大王泊旱 6 (SHMM)

graph for the word the deity name *Shangdi* 上帝, for example, we can now talk specifically about what it means for a ligated graph to be ”well formed.” Although the orthography is different from what we are used to today, the two component characters of {上} and {帝} should still be evident: The former sits atop the latter, and they are conjoined at their bottom and topmost strokes respectively. Both characters are clearly identifiable, hence the ligature is legible. This is the most important factor in a graph (ligated or otherwise) being well formed. Compare how the two component graphs look when written independently against their above ligated form:  , and  . Although the two *di* graphs in this comparison are not exactly the same, they are similar enough to allow us to say that they are both intended to write the same graph. Finally, let us note the third requirement for a ligature to be well formed which we can see here: No strokes are missing. {帝} in Chu orthography does not

[1]The person determiner was not allowed to be placed in the upper left corner of the character space during this time, for example.

have the topmost 丶 stroke that the modern graph has[2]. These three facts must hold true for any combination *hewen* to be properly formed.

Let us consider a second example. The graphs {少} and {人} could be

Figure 4.2: 少人（小人）：周易 8 (SHMM)

ligated in multiple ways, as we have seen: In addition to the above method, they could also be put together as adjuncts. Additionally, this word, *xiaoren* 小人'petty person', could be written with the graph {心} as an adjunct. The adjunct forms are the majority of what we see across the two *zibian* 字編 that I have been working with. Adjuncts are in fact the majority of *hewen* found in Warring States manuscripts, and this ligation might be a perfect example of why that is. Compare the above character against its two component graphs:  and  . Both graphs are easily read in the ligation, they are formed as one would expect, and no graphical information has been lost. There is therefore no question that this is a well made ligation. And yet because of how simple both are, they cannot help but begin to get lost in each other. {人} has only two strokes composing it, so when one is shared with another graph, that means the reader must look over the ligation more carefully. More strokes mean a graph has more it can share while still allowing the reader to quickly identify it, because it has more graphical information. The weakness in the *hewen* subsystem, then, is that even when a ligation is allowed to be formed, it can still produce something which is more difficult to parse. In contrast to combinations (the only category of truly ligated graphs), adjuncts do not have this problem.

That is not all, however. On the opposite side of things, a ligation could be

[2]It is, however, often written with an extra stroke above its topmost stroke. As was previously noted, this is a stylistic feature of Chu orthography rather than something inherent to the character itself. One may think of this extra stroke as being similar to a serif in Latin typography.

formed in too complex a manner. If adjuncts represent the simplest way of putting two characters together within the same graphical space, then ligations which intermingle their components represent the most complicated way.

Figure 4.3: 教學：郭語叢 61 (CXJBW)

The graphical forms used to write the components of this ligation for *jiao xue* 教學'teach and learn' are the following:  ,-  . The expected graphical information is full present in a form readily identifiable, but some liberty has been taken in order to combine the two {子} graphs. Although this is a properly formed ligation, it requires the reader to "disentangle" the two characters in order to read them. Again, this is more work then reading graphs that haven been vertically squeezed together (as was most commonly done), or even graphs that violated the reading order to be squeezed horizontally together. What the scribe saves in effort, the reader loses in ease of comprehension, and this goes against the purpose of writing — i.e., the graphical dissemination of information.

Contrast this with adjuncts. Adjuncts could be formed from any characters, as the components would only have to be squeezed together. The spatial placement of the characters was determined by the conventions of the more standard orthography. If a character was usually written in a certain spot when it served as a determiner (e.g. {人} is always placed to the left), it would assume its normal spot within the *hewen* as well. If a character's spatial orientation was not already set, then the scribe was free to compose the *hewen* in a way that he found most pleasing. This usually meant in a top-to-bottom orientation, because this was the one that could accept the most

characters[3]. After this was done, the *hewen hao* was added. Thus, adjuncts were an orthographic device that saved the scribe effort without sacrificing much, if any, convenience for the reader. We should therefore expect that if more Chu manuscripts are found, then adjuncts will come to occupy the majority of the corpus, just as they do for the larger Warring States corpus.

This concludes my discussion on the principles by which *hewen* are formed in Chu orthography. This is the ”how,” as distinct from the ”why,” which is to economize on space and effort. While nothing in this paper may be groundbreaking, because *hewen* are a little researched topic in Western sinology, it was my desire to provide a resource which would gather the research that has already been done and bring it all into one place in order to answer a question which I had: What rules did a scribe have to consider before he could ligate two or more graphs? What considerations did he have which prevented him from taking advantage of this convenience more often? If my methodology is sound and I have not misinterpreted my data, then the answer to that would seem to be the fact that language, even in non-verbal form, is social and meant to achieve some end in an effective manner. Writing which fails to communicate fails to be writing. Writing and the graphs used for it are not only tools, but arts that have never failed to inspire admiration and appreciation by all cultures which acquire them. Thus it is my hope that I may contribute to our understanding and appreciation of both.

[3]Consider how {不} and {用} are combined to write the modern contraction *beng* 甭’need not’.

Bibliography

Bao Huifang 暴慧芳 (2009). “Hanyu gu wenzi hewen yanjiu”. 漢語古文字合文研究. 西南大學.

Jin Xinxin 金欣欣 (2017). “Jiaguwen de hewen suo tixian chu de ji ge tedian”. 甲骨文的合文所体现出的几个特点. In: 淮南师范学院学报 19.6, pp. 88–93.

Lee Kyoo-kap 李圭甲 (2014). “Jiaguwen yu jinwen ji Zhanguo wenzi hewen zixing bijiao fenxi”. 甲骨文與金文及戰國文字合文字形比較分析. In: 中國文字研究 19, pp. 90–96.

Lee Kyoo-kap 李圭甲 and Kim Eun-hee 金殷嬉 (2012). “Jiaguwen hewen yu Zhanguo wenzi hewen bijiao fenxi”. 甲骨文合文與戰國文字合文比較分析. In: 中國文字研究 16, pp. 45–53.

Lei Liming 雷黎明 (2009). “Lun Chu jian hewen”. 論楚簡合文. In: 寧夏大學學報（人文社會科雪板）31.6, pp. 7–12.

Li Shoukui 李守奎 (2003). “Chutu Chu wenxian wenzi yanjiu zongshu”. 出土楚文献文字研究综述. In: 古籍整理研究学刊 1, pp. 9–17.

Richter, Matthias (2013). *The Embodied Text: Establishing Textual Identity in Early Chinese Manuscripts*. Leiden: Brill.

Sanft, Charles (2019). *Literate Community in Early Imperial China: The Northwestern Frontier in Han Times*. Albany: State University of New York Press.

Teng Rensheng 滕壬生 (2008). *Chu xi jian bo wen zibian (zengdingben)*. 楚系簡帛文字編（增訂本）. 武漢: 湖北教育出版社.

Wang Jing 王景 (2017). “Jiaguwen ”hewen” xianxiang chutan - yi *Shang Zhou gu wenzi duben - Yinxu jiagu keci* 38 pian wei li”. 甲骨文“合文”现象初探——以《商周古文字读本——殷墟甲骨文刻辞》38 片为例. In: 黑河學刊 3, pp. 39–40.

Yu Peng 禹鵬 (2012). “Zhanguo Chu jian hewen yanjiu”. 戰國楚簡合文研究. 吉林大學.

Zhang Feng 張峰 (2015). “Chu jian sheng xing fu hao ”=” ji xiangguan zi lüe shuo”. 楚簡省形符號“=”及相關字略說. In: 古文字研究 6.141, pp. 112–116.