

《東觀漢記》舊版造字轉碼說明

中研院資訊所文獻處理實驗室 2007/9/14 陳建安 製作

一、《東觀漢記》一書(東觀漢記.xml)使用舊版造字 17 個，字頻 70 次，詳如附件一。這 17 個造字中，11 個可轉成 Windows XP 可顯示的字，字頻 59 次；另外 6 個字必須轉成構字式，字頻 11 次。未確認字形 2 個 (◻及●)，字頻 49 次，詳如附件二。

二、附件一的造字分析表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、造字：舊版造字

丙、字頻(txt)：造字在「.txt」文件的出現次數

丁、字頻(xml)：造字在「.xml」文件的出現次數

戊、Big5：造字的 Big5 碼

己、Unicode：造字所對應的 Unicode 碼

庚、WinXP：造字在 Windows XP 的對應字形

辛、構字式：Windows XP 若無對應字形，則改採用構字式

壬、備註凡例：

1、錯字：與原書字形不符，用正確的字取代原造字，如編號 5221 的「𠂔𠂔𠂔𠂔𠂔」字，原字應為「粵」。

2、取代：為減少造字數量，經比對原書為同字異構者，採用正體字取代之，亦便檢索，如編號 4507 的「吳」係「吳」係「吳」字之簡體，故以「吳」字取代。

3、Unicode 字型呈現差異：Unicode 字型與舊漢籍造字有些微差異，但只是字體風格差異，實際上仍為同一個字，因此仍取 Unicode 字型。如編號 4158 的「譌」字，Unicode 字型呈現為「譌」，實際上仍為同一字。

4、配合修改，手動取代：xml 字頻低於 txt 字頻，經比對 txt 及 xml 檔案內容後，發現一部分原舊漢籍造字之前已被人工修改，所以剩餘未修改的舊漢籍造字也配合修改。如編號 4132，舊漢籍造字字形為「血𠂔𠂔」，比對 txt 檔及 xml 檔後，發現原先在 txt 檔使用舊漢籍造字「血𠂔𠂔」者，在 xml 檔中部分已人工修改為構字式「血𠂔𠂔」，故配合修改，將編號 4132 的「血𠂔𠂔」尚未修改者，手動取代為構字式「血𠂔𠂔」。又如編號 853，舊漢籍造字字形為「焞」字，xml 檔部份修改成構字式「火𠂔育」，但由於已有 Unicode 字型，故將「火𠂔育」

改為「𤇗」，剩餘未修改的舊漢籍造字也配合修改。

5、重碼部件，不轉碼：同時使用舊漢籍造字及構字式，導致一個字碼代表兩種字形的重碼問題，必須一一比對原書字形，以判斷應取何者之字形，取構字式部件者不轉碼。如編號 5335，舊漢籍造字字形為「𤇗」，構字式部件為「𤇗」，比對原書發現書中皆取構字式部件字形「𤇗」，故不需轉碼。

6、待造字：舊版漢籍造字字形為電腦缺字，且漢字構形資料庫尚未收錄此字形，等待補造字中，所以「造字」欄空白無法看到字形。如編號 4132 的「𤇗𤇗𤇗」。

三、Unicode 目前收錄的漢字總數為 70194，分屬於三個不同區段，詳如表一。目前 Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以如造字編號 5225 的「𤇗𤇗𤇗」字，Unicode 編碼為 460F，由於 Windows XP 並不支援，仍須改用構字式「𤇗𤇗𤇗」。

表一、Unicode 的字數及編碼區段

Unicode	字集子集合	新增字數	新增編碼區段	總字數	WinXP
1.1 版	CJK 認同表意文字區	20902	4E00-9FFF	20902	支援
3.0 版	CJK 認同表意文字擴充 A 區	6582	3400-4DFF	27484	不支援
3.1 版	CJK 認同表意文字擴充 B 區	42710	20000-2A6D6	70194	不支援

四、《東觀漢記》(東觀漢記.xml)中部分字形顯示為未確認字形「❓」及「●」，已對照原書字形更正，詳細更正內容可參考附件二的未確認字形取代表。附件二的未確認字形取代表說明如下：

甲、檔案原文摘錄：摘錄檔案(東觀漢記.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。文句後方標示(全部取代)者，表示執行全部取代的動作，不一一列舉所有原文。

乙、手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。文句後方標示(全部取代)者，表示執行全部取代的動作，不一一列舉所有修改後結果。

五、造字分析表「備註」欄標示「手動取代」者，詳細取代內容可參考附件三的手動取代表。附件三的手動取代表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、 檔案原文摘錄：摘錄檔案(東觀漢記.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。

丙、 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有修改後結果。

六、 對於在缺字之外發現的錯字，其詳細內容可參考附件四的校出錯字表。附件四校出錯字表的說明同附件二。

附件一、《東觀漢記》造字分析表

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
853	煇	2	1	8EA5	7134	煇		改火𤇀育為煇，手動取代。
1077	稟	8	1	8FE8	7A3E	稟		改高𤇀禾為稟，手動取代。
3883	徧	3	1	81D4	5FA7	徧		改彳𤇀扁為徧，手動取代。
3996	疎	3	1	82A8	758E	疎		改正 𤇀束為疎，手動取代。
4080	羣	110	8	82FC	7FA3	羣		改君𤇀羊作羣，手動取代。
4132		218	9	8371			血𤇀𠂔	血𤇀𠂔，配合修改，手動取代。待造字。
4158	譌	5	5	83AD	8B4C	譌		Unicode 呈現差異
4507	吳	188	18	85D0	5434	吳		吳，取代。
4882	輻	1	1	884E	283DF		輻𤇀車	
5221		1	1	8A67			𠂔白米𠂔	粵，錯字。

5225	邨	1	2	8A6B	460F		血𠂔卩	與土重碼，重碼部件，不轉碼。
5335	𧈧	4	5	8AFB	459F		亡𠂔虫	與𧈧重碼，重碼部件，不轉碼。
5439	栢	1	1	8BC6	6822	栢		
5558	麓	3	7	8C7E	9E84	麓		與冗重碼，重碼部件，不轉碼。
5575	𡵓	1	1	8CB1	8A14	𡵓		
5706	敍	29	7	8D75	654D	敍		改余𠂔支為敍，手動取代。
5850		1	1	C6C9				厶，錯字。

附件二、《東觀漢記》未確認字形取代表

檔案原文摘錄	手動取代結果	備註
趙王庶兄胡子進狗●馬醢	趙王庶兄胡子進狗𠂔𠂔葉馬醢	
啟●鴻化	啟迪鴻化	
●襲（全部取代）	𠂔牙𠂔木襲（全部取代）	有 2 處
其名曰●	其名曰𠂔𠂔𠂔	
礪●（全部取代）	礪石𠂔弟（全部取代）	有 4 處
「●」與「礪」通	「石𠂔弟」與「礪」通	
●書案破之	𠂔𠂔音書案破之	
鈎牧車屏●泥（全部取代）	鈎牧車屏𠂔𠂔屏泥（全部取代）	有 2 處
舉手●掄	舉手𠂔𠂔虎𠂔欠掄	
「●」，說文作「𡵓」	「𠂔𠂔虎𠂔欠」，說文作「𡵓」	
「●掄」，即「𡵓瘡」	「𠂔𠂔虎𠂔欠掄」，即「𡵓瘡」	
●目（全部取代）	目𠂔葉目（全部取代）	有 2 處
●哉是翁	日𠂔𠂔哉是翁	
荊州人曰●	荊州人曰𠂔𠂔𠂔	
「●」與「𠂔」字同	「𠂔𠂔𠂔」與「𠂔」字同	
任●（全部取代）	任𠂔𠂔𠂔（全部取代）	有 9 處
●家舊有東觀漢記（全部取代）	𠂔𠂔𠂔家舊有東觀漢記（全部取代）	有 2 處

其所●漏	其所 <u>里</u> 漏	
狩谷●齋	狩谷 <u>掖</u> 齋	
●奚△鳥走馬	<u>鬪</u> 奚△鳥走馬	
張鳴與敖戲爭●	張鳴與敖戲爭 <u>鬪</u>	
且●且行	且 <u>鬪</u> 且行	
伯升作攻城●車	伯升作攻城 <u>鬪</u> 車	
坐●車上	坐 <u>鬪</u> 車上	
四夷雲集龍●野	四夷雲集龍 <u>鬪</u> 野	
兩虎安得私●	兩虎安得私 <u>鬪</u>	
由是無爭道變●（全部取代）	由是無爭道變 <u>鬪</u> （全部取代）	有 2 處
彰子普坐●殺游徼	彰子普坐 <u>鬪</u> 殺游徼	
聞戰●之聲	聞戰 <u>鬪</u> 之聲	
●爭殺堅得	<u>鬪</u> 爭殺堅得	
持兵而●（全部取代）	持兵而 <u>鬪</u> （全部取代）	有 2 處
走馬●橐駝	走馬 <u>鬪</u> 橐駝	

附件三、《東觀漢記》手動取代表

編號	檔案原文摘錄	手動取代結果	備註
853	蔣光火△育藏清抄本	蔣光 <u>煇</u> 藏清抄本	
	蔣光__清抄本	蔣光 <u>煇</u> 清抄本	
1077	蓋與蔡邕殘__	蓋與蔡邕殘 <u>稟</u>	
	高△禾（全部取代）	<u>稟</u> （全部取代）	
3883	閱考之不得	<u>徧</u> 考之不得	
	彳△扁（全部取代）	<u>徧</u> （全部取代）	
3996	殊為許略	殊為 <u>疎</u> 略	
	正△束（全部取代）	<u>殊</u> （全部取代）	
4080	倉（全部取代）	<u>羣</u> （全部取代）	
	君△羊（全部取代）	<u>羣</u> （全部取代）	
4132	△（全部取代）	<u>血△平</u> （全部取代）	待造字
4507	垂（全部取代）	<u>吳</u> （全部取代）	
	口△天（全部取代）	<u>吳</u> （全部取代）	
5706	丩（全部取代）	<u>敍</u> （全部取代）	
	余△支（全部取代）	<u>敍</u> （全部取代）	

附件四、《東觀漢記》校出錯字表

檔案原文摘錄	手動取代結果
車駕東 <u>广△土</u> 狩	車駕東 <u>巡</u> 狩

逡 <u>广△土</u> 進退	逡 <u>巡</u> 進退
轉司 <u>馬△束</u> 校尉	轉司 <u>隸</u> 校尉