

《鏡湖自撰年譜》舊版造字轉碼說明

中研院資訊所文獻處理實驗室
中研院史語所漢籍電子文獻工作小組
2008/5/28 陳維君 製作

一、《鏡湖自撰年譜》一書(鏡湖自撰年譜.xml)使用舊版造字 56 個，字頻 496 次，詳如附件一。這 56 個造字中，47 個可轉成 Windows XP 能顯示的字，字頻 343 次；另外 9 個字必須轉成構字式，字頻 153 次。

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、異體字問題：新版漢籍考量到使用者檢索及使用時的便利性，將用字原則改為除專詞等特殊情形之外，一律改用標準字呈現。如編號 3858 的「寃」係「冤」字之異體，故以「冤」字取代。又編號 3966 的「烟」字，係「煙」字之簡體字，亦以標準字的「煙」字取代。

2、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 4242 的「鬪」字，Unicode 字型呈現為「鬪」，實際上仍為同一字。

附件一、《鏡湖自撰年譜》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
296	厶	1	1	FBEC			厶	
297	厶	1	1	FBED			厶	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
750	丩	1	1	FEDB			丩	
1077	稟	1	1	8FE8	7A3E	稟		
1187	盖	2	2	90B9	76D6	盖		蓋，異體字問題。
3043	儼	11	11	9C7B	510E	儼		
3161	壘	2	2	9D54	58DC	壘		
3789	冲	4	4	8154	51B2	冲		沖，異體字問題。
3791	况	1	1	8156	51B5	况		況，異體字問題。
3792	凉	1	1	8157	51C9	凉		涼，異體字問題。
3794	凑	1	1	8159	51D1	凑		湊，異體字問題。
3796	决	1	1	815B	51B3	决		決，異體字問題。
3804	効	3	3	8163	52B9	効		
3814	却	16	16	816D	5374	却		卻，異體字問題。
3831	喆	17	17	817E	5586	喆		
3842	堃	3	3	81AB	5803	堃		
3858	寃	15	15	81BB	5BC3	寃		冤，異體字問題。
3864	峯	3	3	81C1	5CEF	峯		峰，異體字問題。
3883	徧	7	7	81D4	5FA7	徧		
3909	擡	8	8	81EE	64E1	擡		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3938	橈	1	1	824C	6AC8	橈		
3945		1	1	8253			汙 𠄎义	汙，異體字問題。
3949	涖	2	2	8257	6D96	涖		
3966	烟	5	5	8268	70DF	烟		煙，異體字問題。
3974	牀	2	2	8270	7240	牀		
3996	疎	1	1	82A8	758E	疎		
4015	眦	1	1	82BB	7726	眦		
4016	着	76	76	82BC	7740	着		著，異體字問題。
4028	稟	11	11	82C8	7980	稟		稟，異體字問題。
4036	窰	2	2	82D0	7AB0	窰		
4038	竈	4	4	82D2	7AC8	竈		
4080	羣	2	2	82FC	7FA3	羣		群，異體字問題。
4090	脚	26	26	8347	811A	脚		腳，異體字問題。
4094	館	16	16	834B	8218	館		
4105	菓	1	1	8356	83D3	菓		
4112	蔴	2	2	835D	8534	蔴		
4132		134	134	8371			血 𠄎 𠄎	眾，異體字問題。
4134	衛	15	15	8373	885E	衛		衛，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4165	獾	2	2	83B4	8C9B	獾		
4173	駝	1	1	83BC	8EAD	駝		
4193	鈎	2	2	83D0	920E	鈎		鈎，異體字問題。
4242	鬪	1	1	8442	9B2D	鬪		Unicode 呈現差異。
4256	鷄	2	2	8450	9DC4	鷄		
4259	麇	1	1	8453	9E95	麇		Unicode 呈現差異。
4309	刼	29	29	84A7	523C	刼		刼，異體字問題。
4447	捏	1	1	8572	63D1	捏		捏，異體字問題。
4603	憇	16	16	8671	6187	憇		
4735	壠	2	2	8758	58E0	壠		
4904	尅	3	3	8864	5C05	尅		
5019	畧	5	5	88F9			畧	畧，異體字問題。
5176		4	4	89F9			氵 厶 元	沉，異體字問題。
5472	瓦	3	3	8BE7			瓦	瓦，異體字問題。
5591	杞	3	3	8CC1	233CC		木 厶 巳	杞，異體字問題。
5659	派	15	15	8D46	6D71	派		
5666	污	1	1	8D4D	6C5A	污		污，異體字問題。
5706	敘	4	4	8D75	654D	敘		敘，異體字問題。