

# 《春明退朝錄》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  
2008/5/27 葉冠孜 製作

一、《春明退朝錄》一書(春明退朝錄.xml)使用舊版造字 42 個，字頻 107 次，詳如附件一。這 42 個造字中，34 個可轉成 Windows XP 能顯示的字，字頻 81 次；另外 8 個字必須轉成構字式，字頻 26 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 4134 的「衛」係「衛」字之異體，故以「衛」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 3846 的「壻」字，Unicode 字型呈現為「壻」，實際上仍為同一字。

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。所以造字編號 3788 的「慤」字，Unicode 編碼為 22871，由於 Windows XP 並不支援，仍須使用構字式「殼心」。

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

附件一、《春明退朝錄》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
296	𠄎	5	5	FBEC			𠄎	
297	𠄎	4	4	FBED			𠄎	
713	𠄎	1	1	FEB6			𠄎	
1000		1	1	8F79			王𠄎已	玦，異體字問題。
1026	珣	1	1	8FB5	73E6	珣		
1077	稟	2	2	8FE8	7A3E	稟		
1226	罰	1	1	90E0	7F78	罰		罰，異體字問題。
1893	袞	4	4	9548	886E	袞		
3788	慤	1	1	8153	22871		殼𠄎心	
3814	却	3	3	816D	5374	却		卻，異體字問題。
3817	廁	1	1	8170	53A0	廁		廁，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3819	廝	1	1	8172	53AE	廝		廝，異體字問題。
3824	叙	2	2	8177	53D9	叙		敘，異體字問題。
3846	壻	1	1	81AF	58FB	壻		Unicode 呈現差異。
3877	迺	1	1	81CE	5EFC	迺		迺，異體字問題。
3916	敷	1	1	81F5	656D	敷		
3961	灾	1	1	8263	707E	灾		災，異體字問題。
3985	璫	3	3	827B	7447	璫		Unicode 呈現差異。
3996	疎	1	1	82A8	758E	疎		
4010	阜	1	1	82B6	7690	阜		
4080	羣	5	5	82FC	7FA3	羣		群，異體字問題。
4107	葱	2	2	8358	8471	葱		
4118	藥	1	1	8363	8602	藥		
4132		8	8	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	12	12	8373	885E	衛		衛，異體字問題。
4142	褒	3	3	837B	8943	褒		褒，異體字問題。
4180	臯	1	1	83C3	8FA0	臯		
4184	迹	4	4	83C7	8FF9	迹		跡，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4220	穎	1	1	83EB	9834	穎		穎，異體字問題。
4242	鬪	1	1	8442	9B2D	鬪		Unicode 呈現差異。
4278	亘	1	1	8466	4E98	亘		
4460	勅	2	2	85A1	52D1	勅		敕，異體字問題。
4507	吳	14	14	85D0	5434	吳		吳，異體字問題。
4698	瘞	1	1	86F2	5850	瘞		
4871	嬾	1	1	8843	5B4F	嬾		
4902	兗	3	3	8862	5156	兗		
5439	栢	1	1	8BC6	6822	栢		
5476	楷	5	5	8BEB	3B7C		木𠂔沓	
5582	鞣	1	1	8CB8	97B5	鞣		
5645	𠂔	1	1	8CF7	4836		𠂔	射，異體字問題。
5704	羶	1	1	8D73	6C0A	羶		
5706	敍	2	2	8D75	654D	敍		敘，異體字問題。