

# 《東齋記事》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  
2008/5/26 葉冠孜 製作

一、《東齋記事》一書(東齋記事.xml)使用舊版造字 60 個，字頻 144 次，詳如附件一。這 60 個造字中，48 個可轉成 Windows XP 能顯示的字，字頻 93 次；另外 12 個字必須轉成構字式，字頻 51 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 3791 的「况」係「況」字之異體，故以「況」字取代。又編號 1513 的「糧」字，係「糧」字之簡體字，亦以標準字的「糧」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 3975 的「牕」字，Unicode 字型呈現為「牕」，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 2830 的「眈」字，Unicode 編碼為 4022，由於 Windows XP 並不支援，仍須使用構字式「目𠃉矢」。

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、《東齋記事》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	4	4	FBEC			𠄎	
297	𠄎	2	2	FBED			𠄎	
777	宀	2	2	FEF6			宀	
794	隋	1	1	8E48	9657	隋		Unicode 呈現差異。
842	𠄎	5	5	8E78	3AB7		𠄎	
1067	碁	2	2	8FDE	68CA	碁		碁，異體字問題。
1226	罰	2	2	90E0	7F78	罰		罰，異體字問題。
1513	糧	1	1	92C5	7CAE	糧		糧，異體字問題。
1544	絙	1	1	92E4	7D59	絙		
1660	賺	1	1	93BB	8181	賺		
1893	袞	1	1	9548	886E	袞		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
2104	緇	1	1	967E	260D3		纟 厶 留	
2472	靄	3	3	98D6	9741	靄		
2524	饌	1	1	994B	994D	饌		膳，異體字問題。
2630	鏹	1	1	99D7			金 厶 朶	剝，異體字問題。
2830	眈	1	1	9B43	4022		目 厶 矢	
2975	逖	1	1	9BF6	8FEF	逖		
3791	况	1	1	8156	51B5	况		况，異體字問題。
3796	决	1	1	815B	51B3	决		決，異體字問題。
3801	剗	1	1	8160	5257	剗		
3804	効	1	1	8163	52B9	効		效，異體字問題。
3805	勅	3	3	8164	52C5	勅		敕，異體字問題。
3813	邛	1	1	816C	536D	邛		
3814	却	6	6	816D	5374	却		卻，異體字問題。
3817	廁	1	1	8170	53A0	廁		廁，異體字問題。
3829	咏	1	1	817C	548F	咏		
3843	堦	1	1	81AC	5826	堦		
3846	堦	2	2	81AF	58FB	堦		Unicode 呈現差異。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3858	寃	3	3	81BB	5BC3	寃		冤，異體字問題。
3864	峯	2	2	81C1	5CEF	峯		峰，異體字問題。
3883	徧	1	1	81D4	5FA7	徧		
3894	筋	1	1	81DF	4225		竹 厶 勛	筋，異體字問題。
3916	敷	1	1	81F5	656D	敷		
3923		1	1	81FC			晶 厶 宜	疊，異體字問題。
3973	徧	1	1	826F			宀 厶 扁	徧，異體字問題。
3975	牕	1	1	8271	7255	牕		Unicode 呈現差異。
3979	猫	1	1	8275	732B	猫		貓，異體字問題。
3996	疎	4	4	82A8	758E	疎		
4016	着	2	2	82BC	7740	着		著，異體字問題。
4028	稟	3	3	82C8	7980	稟		稟，異體字問題。
4033	稽	1	1	82CD	7A2D	稽		
4040	竝	1	1	82D4	7ADD	竝		
4044	筋	1	1	82D8	7B6F	筋		
4080	羣	13	13	82FC	7FA3	羣		群，異體字問題。
4090	脚	1	1	8347	811A	脚		腳，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4105	菓	1	1	8356	83D3	菓		
4132		29	29	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	1	1	8373	885E	衛		衛，異體字問題。
4167	賈	1	1	83B6	8CEB	賈		
4242	鬪	1	1	8442	9B2D	鬪		Unicode 呈現差異。
4243	𩵑	4	4	8443	9B34	𩵑		
4278	亘	1	1	8466	4E98	亘		
4507	吳	8	8	85D0	5434	吳		吳，異體字問題。
5290	携	1	1	8ACE	643A	携		攜，異體字問題。
5582	鞣	1	1	8CB8	97B5	鞣		
5590	朶	1	1	8CC0	6736	朶		朶，異體字問題。
5591	杞	3	3	8CC1	233CC		木𠄎巳	杞，異體字問題。
5627	涑	2	2	8CE5	6D13	涑		
5702	厨	1	1	8D71	53A8	厨		厨，異體字問題。
5797	灑	2	2	8DF2	7054	灑		