

# 《巾箱說》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  
2008/5/22 葉冠孜 製作

一、《巾箱說》一書(巾箱說.xml)使用舊版造字 66 個，字頻 152 次，詳如附件一。這 66 個造字中，52 個可轉成 Windows XP 能顯示的字，字頻 127 次；另外 14 個字必須轉成構字式，字頻 25 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 1067 的「碁」係「棋」字之異體，故以「棋」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。
- 3、待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 4654 的「土 $\Delta$ 廷」。

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 1347 的「𠄎」字，Unicode 編碼為 21E05，由於 Windows XP 並不支援，仍須使用構字式「 $\square$ 山弓女 $\square$ 」。

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、《巾箱說》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	2	2	FBEC			𠄎	
297	𠄏	3	3	FBED			𠄏	
299	𠄑	2	2	FBEF			𠄑	
300	𠄒	2	2	FBF0			𠄒	
969	插	4	4	8F5A	7250	插		
1010	珏	1	1	8FA5	73CF	珏		
1067	碁	1	1	8FDE	68CA	碁		碁，異體字問題。
1226	罰	1	1	90E0	7F78	罰		罰，異體字問題。
1347	𡗗	1	1	91BC	21E05		𠄑山弓女𠄒	
1416	埃	3	3	9242	7AE2	埃		
1981	讐	3	3	95C2	8B90	讐		
2189	𡗗	1	1	96F5	23C3D		毛𠄎瞿	
2190	𡗗	1	1	96F6	6BFA	𡗗		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
2212	襪	3	3	974D	277CC		衤 ㄇ 羅	
3792	涼	4	4	8157	51C9	涼		涼，異體字問題。
3814	却	4	4	816D	5374	却		卻，異體字問題。
3826	叶	1	1	8179	53F6	叶		
3829	咏	3	3	817C	548F	咏		
3846	壻	1	1	81AF	58FB	壻		Unicode 呈現差異。
3858	冤	1	1	81BB	5BC3	冤		冤，異體字問題。
3883	徧	1	1	81D4	5FA7	徧		
3886	悞	2	2	81D7	609E	悞		誤，異體字問題。
3930	憵	1	1	8244	6901	憵		
3931	椀	5	5	8245	6900	椀		
3961	灾	1	1	8263	707E	灾		
3966	烟	9	9	8268	70DF	烟		煙，異體字問題。
3974	牀	1	1	8270	7240	牀		
3996	疎	1	1	82A8	758E	疎		
4039	竝	2	2	82D3	7ADA	竝		
4055	粍	5	5	82E3	7C83	粍		粍，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。
4080	羣	6	6	82FC	7FA3	羣		群，異體字問題。
4090	脚	1	1	8347	811A	脚		腳，異體字問題。
4091	臯	1	1	8348	81EF	臯		
4115	藁	2	2	8360	85C1	藁		
4132		4	4	8371			血𠄎𠄎	眾，異體字問題。
4146	覩	2	2	83A1	89A9	覩		
4167	賈	1	1	83B6	8CEB	賈		
4188	邨	8	8	83CB	90A8	邨		
4195	鉢	1	1	83D2	9262	鉢		
4203	隣	2	2	83DA	96A3	隣		鄰，異體字問題。
4242	鬪	1	1	8442	9B2D	鬪		Unicode 呈現差異。
4256	鷄	1	1	8450	9DC4	鷄		雞，異體字問題。
4262		3	3	8456	9EAA	麪		麵，異體字問題。
4267	鼈	1	1	845B	9F08	鼈		
4307	罵	1	1	84A5	99E1	罵		罵，異體字問題。
4450	劬	1	1	8575	524F	劬		創，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4507	吳	9	9	85D0	5434	吳		吳，異體字問題。
4531	攢	3	3	85E8	6505	攢		
4654		1	1	86C6			土𠂔廷	待造字。
4902	兗	9	9	8862	5156	兗		
4940	韻	3	3	88AA	97F5	韻		韻，異體字問題。
4996	斲	1	1	88E2	65B5	斲		
5019	畧	1	1	88F9			畧	略，異體字問題。
5225	邨	1	1	8A6B	460F		血𠂔卩	
5241	鷺	1	1	8A7B	9D5E	鷺		鵠，異體字問題。
5277	遑	2	2	8AC1			形辶山王	
5290	携	2	2	8ACE	643A	携		攜，異體字問題。
5312		1	1	8AE4			形女辵冈比	媿，異體字問題。
5551	賈	1	1	8C77	8CF7	賈		
5555	迥	1	1	8C7B	9008	迥		迥，異體字問題。
5588	恠	1	1	8CBE	6061	恠		
5591	杞	1	1	8CC1	233CC		木𠂔巳	杞，異體字問題。
5697	煑	1	1	8D6C	7151	煑		煮，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5702	廚	2	2	8D71	53A8	廚		廚，異體字問題。
5704	𩛱	4	4	8D73	6C0A	𩛱		