

《洗冤集錄校譯》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
2008/2/13 黃榮順 製作

一、《洗冤集錄校譯》一書(洗冤集錄校譯.xml)使用舊版造字 23 個，字頻 41 次，詳如附件一。這 23 個造字中，21 個可轉成 Windows XP 可顯示的字，字頻 39 次；另外 2 個字必須轉成構字式，字頻 2 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 1073 的「算△么」係「纂」字之異體，故以「纂」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 1350 的「褚」字，Unicode 字型呈現為「褚」，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。所以如造字編號 1331 的「嶸」字，Unicode 編碼為 3813，由於 Windows XP 並不支援，仍須改用構字式「山△蒙」。

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、「洗冤集錄校譯.xml」發現之未正確修改聲調符號共 88 處，修正情形詳列附件二，附件二「洗冤集錄校譯.xml」聲調符號修改表體例說明如下：

- 甲、編號：流水號，依原書出現順序排列。
- 乙、檔案原文摘錄：摘錄檔案(洗冤集錄校譯.xml)中包含該聲調符號的文句段落，其中紅字底線的部分，為需要手動取代的原文。
- 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。

附件一、「洗冤集錄校譯.xml」造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1073		1	1	8FE4			算△么	篡，異體字問題。
1331	嶢	1	1	91AC	3813		山△蒙	
1350	褚	1	1	91BF	7987	褚		Unicode 字型差異。
2567	稗	1	1	9976	7A21	稗		
3145	擗	2	2	9D44	64C0	擗		
3876	廻	4	4	81CD	5EFB	廻		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3966	烟	4	4	8268	70DF	烟		
3978	猪	2	2	8274	732A	猪		Unicode 字型差異。
3979	猫	1	1	8275	732B	猫		
3996	疎	1	1	82A8	758E	疎		
4015	眦	4	4	82BB	7726	眦		
4038	竈	1	1	82D2	7AC8	竈		
4118	藁	1	1	8363	8602	藁		
4134	衛	1	1	8373	885E	衛		衛，異體字問題。
4178	輓	1	1	83C1	8F2D	輓		
4184	迹	1	1	83C7	8FF9	迹		
4218	韭	2	2	83E9	97EE	韭		
4223	顚	1	1	83EE	984B	顚		
4256	鷄	3	3	8450	9DC4	鷄		
4447	捏	4	4	8572	63D1	捏		
4558	厖	1	1	8644	53A9	厖		Unicode 字型差異。
5704	羶	2	2	8D73	6C0A	羶		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5706	敘	1	1	8D75	654D	敘		敘，異體字問題。

附件二、「洗冤集錄校譯.xml」聲調符號修改表

編號	檔案原文摘錄	手動取代結果
1	牒(蝶 di <u>e</u>)	牒(蝶 di <u>é</u>)
2	那(挪 nu <u>o</u>)	那(挪 nu <u>ó</u>)
3	見(現 xi <u>a</u> n)	見(現 xi <u>à</u> n)
4	瘞(意 y <u>i</u>)	瘞(意 y <u>ì</u>)
5	寺觀(貫 gu <u>a</u> n)	寺觀(貫 gu <u>à</u> n)
6	卒(促 cu <u>u</u>)	卒(促 cu <u>ù</u>)
7	勘(堪 k <u>a</u> n)	勘(堪 k <u>ā</u> n)
8	又讀看 k <u>a</u> n)	又讀看 k <u>à</u> n)
9	齧(涅 ni <u>e</u>)	齧(涅 ni <u>è</u>)
10	睿(瑞 ru <u>i</u>)旨	睿(瑞 ru <u>ì</u>)旨
11	度(奪 du <u>o</u>)	度(奪 du <u>ò</u>)
12	詣(佚 y <u>i</u>)	詣(佚 y <u>ī</u>)
13	劄(札 zh <u>a</u>)	劄(札 zh <u>á</u>)
14	(余 y <u>u</u>)	(余 y <u>ú</u>)
15	罨(掩 y <u>a</u> n)	罨(掩 y <u>ǎ</u> n)
16	譴(遣 qi <u>a</u> n)	譴(遣 qi <u>ǎ</u> n)
17	攬(竄 cu <u>a</u> n)倒	攬(竄 cu <u>à</u> n)倒
18	挾(協 xi <u>e</u>)讎	挾(協 xi <u>é</u>)讎
19	屏(丙 b <u>i</u> ng)	屏(丙 b <u>ǐ</u> ng)
20	識(至 zh <u>i</u>)	識(至 zh <u>ì</u>)
21	將(江 j <u>i</u> ang)	將(江 j <u>ī</u> ang)
22	根勘(堪 k <u>a</u> n)	根勘(堪 k <u>ā</u> n)
23	賁(躋 j <u>i</u>)	賁(躋 j <u>ì</u>)
24	觸髅(獨樓 du—l <u>o</u> u)	觸髅(獨樓 dú—l <u>óu</u>)
25	櫟(舉 j <u>u</u>)	櫟(舉 j <u>ǔ</u>)
26	顛(凶 x <u>i</u> n)門	顛(凶 x <u>ìn</u>)門
27	筭(蒜 su <u>a</u> n)害	筭(蒜 su <u>à</u> n)害

編號	檔案原文摘錄	手動取代結果
28	胖(膨 p <u>e</u> ng)	胖(膨 p <u>a</u> ng)
29	頰(科 k <u>e</u>)	頰(科 k <u>ē</u>)
30	舊讀孩 h <u>ai</u>	舊讀孩 h <u>ái</u>
31	賺朐(廉刃 li <u>a</u> n r <u>e</u> n)	賺朐(廉刃 li <u>án</u> r <u>èn</u>)
32	(秋 q <u>i</u> u)	(秋 q <u>iū</u>)
33	胖(p <u>e</u> ng)	胖(p <u>àng</u>)
34	九竅(俏 q <u>i</u> ao)	九竅(俏 q <u>iao</u>)
35	掇(弱 ru <u>o</u>)	掇(弱 ru <u>ó</u>)
36	又讀挪 nu <u>o</u>	又讀挪 nu <u>ō</u>
37	爿(脇 xi <u>e</u>)	爿(脇 xi <u>é</u>)
38	沃(握 w <u>o</u>)	沃(握 w <u>ō</u>)
39	攢殯(攢殯 cu <u>a</u> n-b <u>i</u> n)	攢殯(攢殯 cu <u>án</u> -b <u>ìn</u>)
40	祿(團 tu <u>a</u> n)	祿(團 tu <u>àn</u>)
41	發(廢 f <u>ei</u>)	發(廢 f <u>èi</u>)
42	辟(批 p <u>i</u>)	辟(批 p <u>ì</u>)
43	搵(問 w <u>e</u> n)	搵(問 w <u>èn</u>)
44	頻(貧 p <u>i</u> n)	頻(貧 p <u>ín</u>)
45	(闐 t <u>a</u>)	(闐 t <u>à</u>)
46	白僵乾(甘 g <u>a</u> n)	白僵乾(甘 g <u>ān</u>)
47	搭(答 d <u>a</u>)	搭(答 d <u>ā</u>)
48	(底 d <u>i</u>)骨	(底 d <u>ǐ</u>)骨
49	揸(擺 b <u>ai</u>)骨	揸(擺 b <u>ǎi</u>)骨
50	贛(贛 b <u>i</u> n)骨	贛(贛 b <u>ìn</u>)骨
51	(喟 ku <u>i</u>)	(喟 ku <u>ì</u>)
52	枋(方 f <u>a</u> ng)	枋(方 f <u>ā</u> ng)
53	桁(恒 h <u>e</u> ng)	桁(恒 h <u>é</u> ng)
54	扎磨(m <u>o</u>)	扎磨(m <u>ó</u>)
55	水搵(文 w <u>e</u> n)死	水搵(文 w <u>èn</u>)死
56	攞(鹵 l <u>u</u>)	攞(鹵 l <u>ǔ</u>)
57	胖(膨 p <u>e</u> ng)	胖(膨 p <u>àng</u>)
58	撙(扎 z <u>a</u>)	撙(扎 z <u>ā</u>)
59	自斡(握 w <u>o</u>)	自斡(握 w <u>ò</u>)
60	斫(酌 zhu <u>o</u>)	斫(酌 zhu <u>ó</u>)
61	透斡(桿 g <u>a</u> n)	透斡(桿 g <u>ǎn</u>)

編號	檔案原文摘錄	手動取代結果
62	贅(聊 li <u>ao</u>)膜	贅(聊 li <u>áo</u>)膜
63	(脫 tu <u>o</u>)	(脫 tu <u>ō</u>)
64	揞(俺 <u>yan</u>)	揞(俺 <u>ǎn</u>)
65	(闢 t <u>a</u>)皮	(闢 t <u>à</u>)皮
66	醞(鹽 y <u>an</u>)	醞(鹽 y <u>àn</u>)
67	黦(慘 c <u>an</u>)黑色	黦(慘 c <u>ǎn</u>)黑色
68	葷(訓 x <u>un</u>)	葷(訓 x <u>ùn</u>)
69	饋(分 f <u>en</u>)	饋(分 f <u>ēn</u>)
70	羸(雷 l <u>ei</u>)	羸(雷 l <u>éi</u>)
71	口眼喎(<u>歪</u> w <u>ai</u>)	口眼喎(<u>快</u> w <u>āi</u>)
72	(派 p <u>ai</u>)	(派 p <u>ài</u>)
73	(脫 t <u>uo</u>)	(脫 t <u>ūo</u>)
74	(墊 di <u>an</u>)	(墊 di <u>àn</u>)
75	撙(扎 z <u>a</u>)	撙(扎 z <u>ā</u>)
76	兌厶皮(脫 t <u>uo</u>)	兌厶皮(脫 t <u>ūo</u>)
77	龕(粗 c <u>u</u>)	龕(粗 c <u>ū</u>)
78	窳(印 y <u>in</u>)	窳(印 y <u>ìn</u>)
79	撚(粘 ni <u>an</u>)	撚(粘 ni <u>án</u>)
80	搦(懦 nu <u>o</u>)	搦(懦 nu <u>ò</u>)
81	翕(系 x <u>i</u>)	翕(系 x <u>ì</u>)
82	喝(荷 h <u>e</u>)	喝(荷 h <u>è</u>)
83	或謁 y <u>e</u>	或謁 y <u>è</u>
84	(敢 g <u>an</u>)	(敢 g <u>ǎn</u>)
85	魘(演 y <u>an</u>)	魘(演 y <u>ǎn</u>)
86	中(仲 zh <u>ong</u>)	中(仲 zh <u>òng</u>)
87	(午 w <u>u</u>)	(午 w <u>ǔ</u>)
88	卒(猝 c <u>u</u>)	卒(猝 c <u>ù</u>)