

《玉壺清話》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
2008/6/23 葉冠孜 製作

一、《玉壺清話》一書(玉壺清話.xml)使用舊版造字 105 個，字頻 517 次，詳如附件一。這 105 個造字中，76 個可轉成 Windows XP 能顯示的字，字頻 414 次；另外 29 個字必須轉成構字式，字頻 103 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。

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 5315 的「浚」字，Unicode 編碼為 3CC2，由於 Windows XP 並不支援，仍須使用構字式「浚公」。

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、《玉壺清話》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	7	7	FBED			厶	
298	厶	2	2	FBEE			厶	
299	形	1	1	FBEF			形	
300	口	1	1	FBF0			口	
336	亩	1	1	FC55			亩	
432	業	1	1	FCD7			業	
674	乚	1	1	FE6D			乚	
711	禛	1	1	FEB4			禛	
745	衣	1	1	FED6			衣	
770	扌	1	1	FEEF			扌	
772	辵	1	1	FEF1			辵	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
792	帑	1	1	8E46	5E2C	帑		
842	蚪	3	3	8E78	3AB7		蚪	
1000		1	1	8F79			王𠂇巳	玨，異體字問題。
1013	紆	4	4	8FA8	7D5A	紆		
1056	陔	1	1	8FD3	9641	陔		
1067	碁	1	1	8FDE	68CA	碁		碁，異體字問題。
1151	劄	1	1	9073	5273	劄		
1187	盖	1	1	90B9	76D6	盖		蓋，異體字問題。
1331	嶠	3	3	91AC	3813		山𠂇蒙	
1364		1	1	91CD			冗𠂇圭	罍，異體字問題。
1565	絲	1	1	92F9	7DDC	絲		
1893	袞	3	3	9548	886E	袞		
2230	銍	1	1	975F	92E5	銍		
2505	飜	5	5	98F7	98DC	飜		
2524	饒	1	1	994B	994D	饒		
2811		1	1	9AEF			亞𠂇田	留，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
2816	備	2	2	9AF4	50C3	備		
3789	冲	2	2	8154	51B2	冲		沖，異體字問題。
3813	邛	1	1	816C	536D	邛		
3814	却	1	1	816D	5374	却		卻，異體字問題。
3824	叙	3	3	8177	53D9	叙		敘，異體字問題。
3829	咏	2	2	817C	548F	咏		
3836	喻	3	3	81A5	564F	喻		
3840	坂	1	1	81A9	5742	坂		
3846	壻	1	1	81AF	58FB	壻		Unicode 呈現差異。
3858	寃	1	1	81BB	5BC3	寃		冤，異體字問題。
3864	峯	5	5	81C1	5CEF	峯		峰，異體字問題。
3868	卮	1	1	81C5	5DF5	卮		
3883	徧	9	9	81D4	5FA7	徧		
3891	慙	4	4	81DC	6159	慙		慚，異體字問題。
3916	敷	3	3	81F5	656D	敷		
3937	檝	1	1	824B	6A9D	檝		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3945		1	1	8253			汙 𠄎义	汙，異體字問題。
3949	泣	1	1	8257	6D96	泣		
3955	澗	1	1	825D	3D4E		汙 𠄎間	澗，異體字問題。
3966	烟	1	1	8268	70DF	烟		煙，異體字問題。
3968	煨	1	1	826A	7145	煨		
3974	牀	5	5	8270	7240	牀		
3978	猪	1	1	8274	732A	猪		豬，異體字問題。
3991	攜	1	1	82A3	64D5	攜		
3996	疎	5	5	82A8	758E	疎		
4010	臯	2	2	82B6	7690	臯		
4013	盃	3	3	82B9	76CC	盃		
4014	蓋	23	23	82BA	8462	蓋		蓋，異體字問題。
4015	眦	1	1	82BB	7726	眦		
4038	竈	1	1	82D2	7AC8	竈		
4040	竝	9	9	82D4	7ADD	竝		
4042	笋	1	1	82D6	7B0B	笋		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4044	筋	1	1	82D8	7B6F	筋		
4045	箒	1	1	82D9	7B92	箒		
4080	羣	30	30	82FC	7FA3	羣		群，異體字問題。
4094	館	2	2	834B	8218	館		
4095	舩	1	1	834C	8229	舩		
4115	藁	1	1	8360	85C1	藁		
4132		34	34	8371			血𠂇𠂇	眾，異體字問題。
4142	裛	1	1	837B	8943	裛		裛，異體字問題。
4146	覩	2	2	83A1	89A9	覩		
4184	迹	18	18	83C7	8FF9	迹		跡，異體字問題。
4195	鉢	2	2	83D2	9262	鉢		
4198	鑛	1	1	83D5	945B	鑛		
4219	頰	1	1	83EA	982E	頰		
4222	頰	1	1	83ED	9847	頰		
4240	鬚	1	1	8440	9AF4	鬚		
4242	鬪	3	3	8442	9B2D	鬪		Unicode 呈現差異。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4260	麤	1	1	8454	9E9E	麤		
4267	鼈	1	1	845B	9F08	鼈		
4278	亘	2	2	8466	4E98	亘		
4455	劒	1	1	857A			僉 刀	劍，異體字問題。
4469	勺	1	1	85AA	5304	勺		
4507	吳	164	164	85D0	5434	吳		吳，異體字問題。
4558	廐	3	3	8644	53A9	廐		廐，異體字問題。
4603	憇	1	1	8671	6187	憇		
4607	狫	2	2	8675	3E60		犭 屯	
4805	瀆	3	3	87C0	51DF	瀆		瀆，異體字問題。
4817	寶	31	31	87CC	5BF3	寶		寶，異體字問題。
4955	龜	6	6	88B9	9F9C		龜	龜，異體字問題。
4957	崑	6	6	88BB	5D53	崑		
5019	畧	1	1	88F9			畧	略，異體字問題。
5028	昇	3	3	8943	662A	昇		
5225	邨	2	2	8A6B	460F		血 卩	邨，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5293		2	2	8AD1			冗厶幕	幕，異體字問題。
5306	簡	22	22	8ADE			𠂇厶閒	簡，異體字問題。
5315	汧	1	1	8AE7	3CC2		汧厶公	
5381	脛	2	2	8B6A	8103	脛		Unicode 呈現差異。
5447	躡	1	1	8BCE	8E83	躡		
5463	岢	1	1	8BDE	65F9	岢		
5542	鑠	1	1	8C6E	93C1	鑠		
5549	藁	2	2	8C75	85F3	藁		
5555	迥	3	3	8C7B	9008	迥		迥，異體字問題。
5587	歛	1	1	8CBD	6B58	歛		
5591	杞	2	2	8CC1	233CC		木厶巳	杞，異體字問題。
5705	廐	1	1	8D74	5ED0	廐		廐，異體字問題。
5706	敘	5	5	8D75	654D	敘		敘，異體字問題。