

# 《石林燕語》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  
2008/6/25 葉冠孜 製作

一、《石林燕語》一書(石林燕語.xml)使用舊版造字 66 個，字頻 286 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 238 次；另外 14 個字必須轉成構字式，字頻 48 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 3789 的「冲」係「沖」字之異體，故以「沖」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 1153 的「坳」字，Unicode 編碼為 362D，由於 Windows XP 並不支援，仍須使用構字式「土幻」。

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、《石林燕語》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	2	2	FBEC			厶	
297	厶	8	8	FBED			厶	
299	形	1	1	FBEF			形	
300	口	1	1	FBF0			口	
770	扌	1	1	FEEF			扌	
778	、	1	1	FEF7			、	
821	脊	1	1	8E63	7718	脊		
972	犇	4	4	8F5D	7287	犇		
1077	稟	2	2	8FE8	7A3E	稟		
1133	瓌	2	2	9061	74CC	瓌		
1153	坳	2	2	9075	362D		土厶幻	坳，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1318	甌	4	4	917D	24B41		同𠄎瓦	
1565	繇	1	1	92F9	7DDC	繇		
1572		1	1	9341			𠄎革口王	𠄎，異體字問題。
1893	袞	13	13	9548	886E	袞		
1906	袞	2	2	9555	88E6	袞		
1981	讐	1	1	95C2	8B90	讐		
3789	冲	6	6	8154	51B2	冲		冲，異體字問題。
3792	凉	3	3	8157	51C9	凉		凉，異體字問題。
3794	湊	1	1	8159	51D1	湊		
3796	决	1	1	815B	51B3	决		决，異體字問題。
3804	効	1	1	8163	52B9	効		効，異體字問題。
3805	勅	4	4	8164	52C5	勅		
3813	邛	8	8	816C	536D	邛		
3814	却	1	1	816D	5374	却		卻，異體字問題。
3824	叙	2	2	8177	53D9	叙		叙，異體字問題。
3826	叶	1	1	8179	53F6	叶		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3846	墇	9	9	81AF	58FB	墇		Unicode 呈現差異。
3883	徧	9	9	81D4	5FA7	徧		
3924	朥	1	1	81FD	671E	朥		
3931	椀	2	2	8245	6900	椀		
3935	槩	1	1	8249	69E9	槩		概，異體字問題。
3996	疎	1	1	82A8	758E	疎		
4010	皐	2	2	82B6	7690	皐		
4028	稟	7	7	82C8	7980	稟		稟，異體字問題。
4080	羣	29	29	82FC	7FA3	羣		群，異體字問題。
4090	脚	2	2	8347	811A	脚		腳，異體字問題。
4104	蒞	1	1	8355	8385	蒞		蒞，異體字問題。
4115	藁	1	1	8360	85C1	藁		
4132		16	16	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	23	23	8373	885E	衛		衛，異體字問題。
4176	輒	5	5	83BF	8F19	輒		
4184	迹	3	3	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
4195	鉢	1	1	83D2	9262	鉢		
4213	鞞	3	3	83E4	97BE	鞞		
4237	髡	1	1	83FC	9AE0	髡		髡，異體字問題。
4258	麇	1	1	8452	9E90	麇		
4272	龐	2	2	8460	9F8E	龐		龐，異體字問題。
4294	仔	1	1	8476	4F03	仔		
4460	勅	7	7	85A1	52D1	勅		
4507	吳	43	43	85D0	5434	吳		吳，異體字問題。
4750	卅	2	2	8767			卅	
4766	揜	1	1	8777	6366	揜		
4768	妬	1	1	8779	59AC	妬		妒，異體字問題。
4812	羗	1	1	87C7	26351		羗△、	羗，異體字問題。
4878	冗	1	1	884A	5B82	冗		冗，異體字問題。
4943	扌	1	1	88AD			扌	
5008	幞	1	1	88EE	5E5E	幞		
5019	畧	7	7	88F9		畧		略，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5398	滄	1	1	8B7B	6E7B	滄		滄，異體字問題。
5423	楛	1	1	8BB6	6810	楛		
5542	鑠	1	1	8C6E	93C1	鑠		
5590	朶	5	5	8CC0	6736	朶		朶，異體字問題。
5705	廐	2	2	8D74	5ED0	廐		廐，異體字問題。
5706	敘	12	12	8D75	654D	敘		敘，異體字問題。
5720	埴	3	3	8DA5	578D	埴		