

《柳南隨筆》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
2008/5/29 陳維君 製作

一、《柳南隨筆》一書(柳南隨筆.xml)使用舊版造字 124 個，字頻 606 次，詳如附件一。這 124 個造字中，107 個可轉成 Windows XP 能顯示的字，字頻 537 次；另外 17 個字必須轉成構字式，字頻 69 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 3858 的「寃」係「冤」字之異體，故以「冤」字取代。又編號 3966 的「烟」字，係「煙」字之簡體字，亦以標準字的「煙」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 2363 的「翮」字，Unicode 編碼為 4396，由於 Windows XP 並不支援，仍須使用構字式「曾𪛗羽」。

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、《柳南隨筆》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	3	3	FBEC			厶	
297	厶	8	8	FBED			厶	
298	厶	1	1	FBEE			厶	
449	斤	8	8	FCE8			斤	
556	商	5	5	FDB6	5547	商		
777	宀	1	1	FEF6			宀	
785	ノ	1	1	FEFE			ノ	
893	峩	1	1	8ECD	5CE9	峩		峨，異體字問題。
1278	碍	1	1	9155	788D	碍		礙，異體字問題。
1642	脉	5	5	93A9	8109	脉		
1668	臙	1	1	93C3	81C8	臙		
1809	困	1	1	94B3	56E6	困		
1893	袞	1	1	9548	886E	袞		袞，異體字問題。
1981	讐	1	1	95C2	8B90	讐		讎，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
2044	踈	1	1	9642	8E08	踈		
2046	踪	1	1	9644	8E2A	踪		
2363	翮	1	1	9847	4396		曾 𪗇 羽	
2408	鬧	2	2	9874	9599	鬧		鬧，異體字問題。
2524	饑	1	1	994B	994D	饑		
3780	儁	1	1	814B	5101	儁		
3789	冲	8	8	8154	51B2	冲		冲，異體字問題。
3791	况	14	14	8156	51B5	况		况，異體字問題。
3794	湊	2	2	8159	51D1	湊		湊，異體字問題。
3798	𠂇	3	3	815D			𠂇	
3802	劓	1	1	8161	529A	劓		
3804	効	1	1	8163	52B9	効		
3805	勅	2	2	8164	52C5	勅		
3814	却	26	26	816D	5374	却		卻，異體字問題。
3817	廁	3	3	8170	53A0	廁		廁，異體字問題。
3819	廡	3	3	8172	53AE	廡		廡，異體字問題。
3820	厦	1	1	8173	53A6	厦		厦，異體字問題。
3825	疊	5	5	8178	53E0	疊		疊，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3826	叶	6	6	8179	53F6	叶		
3829	咏	2	2	817C	548F	咏		
3843	堦	3	3	81AC	5826	堦		
3846	堦	12	12	81AF	58FB	堦		Unicode 呈現差異。
3858	寃	1	1	81BB	5BC3	寃		冤，異體字問題。
3864	峯	19	19	81C1	5CEF	峯		峰，異體字問題。
3868	卮	1	1	81C5	5DF5	卮		
3869	币	1	1	81C6	5E00	币		
3876	廻	1	1	81CD	5EFB	廻		迴，異體字問題。
3883	徧	16	16	81D4	5FA7	徧		
3886	悞	2	2	81D7	609E	悞		
3895	淑	2	2	81E0	6E86	淑		淑，異體字問題。
3921	皙	1	1	81FA	6673	皙		晰，異體字問題。
3936	棹	1	1	824A	6A70	棹		
3938	橈	1	1	824C	6AC8	橈		
3961	灾	3	3	8263	707E	灾		災，異體字問題。
3966	烟	20	20	8268	70DF	烟		煙，異體字問題。
3974	牀	22	22	8270	7240	牀		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3996	疎	12	12	82A8	758E	疎		
4000	疴	1	1	82AC	75B4	疴		
4011	皞	1	1	82B7	76A1	皞		
4016	着	6	6	82BC	7740	着		著，異體字問題。
4028	稟	1	1	82C8	7980	稟		稟，異體字問題。
4038	竈	9	9	82D2	7AC8	竈		
4039	竚	1	1	82D3	7ADA	竚		
4041	豎	1	1	82D5	7AEA	豎		豎，異體字問題。
4042	笋	1	1	82D6	7B0B	笋		
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。
4074	縑	1	1	82F6	7E27	縑		
4080	羣	19	19	82FC	7FA3	羣		群，異體字問題。
4090	脚	9	9	8347	811A	脚		腳，異體字問題。
4091	臯	7	7	8348	81EF	臯		
4104	莅	1	1	8355	8385	莅		
4107	葱	2	2	8358	8471	葱		蔥，異體字問題。
4129	虬	12	12	836E	866C	虬		
4132		26	26	8371			血会亼	眾，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4134	衛	9	9	8373	885E	衛		衛，異體字問題。
4139	袴	4	4	8378	88B4	袴		
4145	羈	1	1	837E	898A	羈		Unicode 呈現差異。
4146	覩	5	5	83A1	89A9	覩		
4158	譌	3	3	83AD	8B4C	譌		Unicode 呈現差異。
4167	賁	1	1	83B6	8CEB	賁		
4172	蹙	2	2	83BB	8E77	蹙		蹙，異體字問題。
4173	駝	2	2	83BC	8EAD	駝		
4176	輒	4	4	83BF	8F19	輒		
4184	迹	8	8	83C7	8FF9	迹		
4188	邨	7	7	83CB	90A8	邨		
4190	邨	1	1	83CD	90C4	邨		
4193	鈎	2	2	83D0	920E	鈎		鈎，異體字問題。
4198	鑛	1	1	83D5	945B	鑛		
4203	隣	2	2	83DA	96A3	隣		鄰，異體字問題。
4213	鞞	1	1	83E4	97BE	鞞		
4216	鞞	1	1	83E7	97C8	鞞		
4232	隲	1	1	83F7	96B2	隲		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4242	鬪	1	1	8442	9B2D	鬪		Unicode 呈現差異。
4256	鷄	1	1	8450	9DC4	鷄		
4259	鷹	1	1	8453	9E95	鷹		Unicode 呈現差異。
4265	鼃	4	4	8459	9F03	鼃		
4266	鼃	1	1	845A	9F02	鼃		
4267	鼃	6	6	845B	9F08	鼃		
4307	罵	6	6	84A5	99E1	罵		罵，異體字問題。
4428	躡	1	1	855F	8E70	躡		躡，異體字問題。
4507	吳	128	128	85D0	5434	吳		吳，異體字問題。
4566	礪	3	3	864C	78F5	礪		
4608	嚳	1	1	8676	7581	嚳		
4702	場	1	1	86F6	5872	場		
4768	妬	1	1	8779	59AC	妬		妒，異體字問題。
4785	贗	1	1	87AC	8D0B	贗		
4856	毳	1	1	87F3	3CB2		疊△毛	
4902	兗	2	2	8862	5156	兗		兗，異體字問題。
4927	岞	1	1	887B	5C9E	岞		岞，異體字問題。
4944	颯	2	2	88AE	4B19		風△彡	颯，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5013	汜	2	2	88F3			汜	
5019	畧	1	1	88F9			畧	略，異體字問題。
5066	瑣	1	1	8969	24A0F		王𠂔賁	瑣，異體字問題。
5126	暮	1	1	89C7	8B29	暮		
5173	漱	1	1	89F6	6F44	漱		漱，異體字問題。
5218		8	8	8A64			𠂔一覓、𠂔	寬，異體字問題。
5225	卹	1	1	8A6B	460F		血𠂔卩	卹，異體字問題。
5470	棟	1	1	8BE5	68C5	棟		
5549	藁	1	1	8C75	85F3	藁		
5552	弃	1	1	8C78	5F06	弃		
5559	欸	10	10	8CA1	6B35	欸		款，異體字問題。
5583	殮	1	1	8CB9	98F1	殮		殮，異體字問題。
5586	𡗗	1	1	8CBC	6B57	𡗗		
5588	恠	1	1	8CBE	6061	恠		
5590	朶	1	1	8CC0	6736	朶		朶，異體字問題。
5619	繚	1	1	8CDD	470C		繚	
5706	敘	11	11	8D75	654D	敘		敘，異體字問題。
5742	澁	3	3	8DBB	6F81	澁		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5782	瀟	2	2	8DE3	701E	瀟		
5797	灑	1	1	8DF2	7054	灑		