

《永和本薩天錫逸詩》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
2008/5/21 葉冠孜 製作

一、《永和本薩天錫逸詩》一書(永和本薩天錫逸詩.xml)使用舊版造字 79 個，字頻 235 次，詳如附件一。這 79 個造字中，67 個可轉成 Windows XP 能顯示的字，字頻 201 次；另外 12 個字必須轉成構字式，字頻 34 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 3864 的「峯」係「峰」字之異體，故以「峰」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 5405 的「叢」字，Unicode 字型呈現為「叢」，實際上仍為同一字。

三、「永和本薩天錫逸詩.xml」發現之缺字外錯字共 24 處，修正情形詳列附件二。附件二「永和本薩天錫逸詩.xml」缺字外錯字修改表體例說明如下：

甲、檔案原文摘錄：摘錄檔案(永和本薩天錫逸詩.xml)中包含錯字的文句段落，其中紅字底線的部分，為需要手動取代的字形。文句後方標示(全部取代)者，表示執行全部取代的動作，不一一列舉所有原文。

- 乙、 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。文句後方標示(全部取代)者，表示執行全部取代的動作，不一一列舉所有原文。
- 丙、 備註：數字表示原文在全書中出現的次數。

附件一、《永和本薩天錫逸詩》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	5	5	FBEC			𠄎	
297	𠄎	3	3	FBED			𠄎	
298	𠄎	3	3	FBEE			𠄎	
772	之	2	2	FEF1			之	
1187	盖	3	3	90B9	76D6	盖		蓋，異體字問題。
1278	碍	1	1	9155	788D	碍		礙，異體字問題。
1893	袞	1	1	9548	886E	袞		
1926	覓	1	1	9569	8994	覓		
1981	讐	1	1	95C2	8B90	讐		
1999	贊	1	1	95D4	8CDB	贊		贊，異體字問題。
2042	䟽	6	6	9640	47FD		足𠄎充	疏，異體字問題。
2232		1	1	9761			𠄎厂墨卩𠄎	廊，異體字問題。
2415	鬪	1	1	987B	95D8	鬪		鬥，異體字問題。
2505	翻	2	2	98F7	98DC	翻		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
2885	叫	1	1	9B7A	544C	叫		
3021	丩	1	1	9C65			丩	
3066	冫	1	1	9CB4	519D	冫		冫，異體字問題。
3204	體	1	1	9DA1	8EC6	體		
3213	塔	4	4	9DAA	366E		土 厶 答	塔，異體字問題。
3251	灯	2	2	9DD0	706F	灯		燈，異體字問題。
3255	善	1	1	9DD4	5584	善		
3770	个	1	1	8141	4E2A	个		個，異體字問題。
3791	况	6	6	8156	51B5	况		況，異體字問題。
3792	凉	4	4	8157	51C9	凉		涼，異體字問題。
3796	决	1	1	815B	51B3	决		決，異體字問題。
3814	却	15	15	816D	5374	却		卻，異體字問題。
3824	叙	1	1	8177	53D9	叙		敘，異體字問題。
3829	咏	11	11	817C	548F	咏		
3864	峯	2	2	81C1	5CEF	峯		峰，異體字問題。
3944	冰	4	4	8252	6C37	冰		冰，異體字問題。
3966	烟	22	22	8268	70DF	烟		煙，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3996	疎	1	1	82A8	758E	疎		
4008	癯	1	1	82B4	766F	癯		
4012	鼓	1	1	82B8	76B7	鼓		
4016	着	20	20	82BC	7740	着		著，異體字問題。
4038	竈	1	1	82D2	7AC8	竈		灶，異體字問題。
4042	笋	1	1	82D6	7B0B	笋		
4055	粃	2	2	82E3	7C83	粃		糝，異體字問題。
4073	纒	1	1	82F5	7E6E	纒		韁，異體字問題。
4074	縑	2	2	82F6	7E27	縑		條，異體字問題。
4080	羣	1	1	82FC	7FA3	羣		群，異體字問題。
4090	脚	3	3	8347	811A	脚		腳，異體字問題。
4091	臯	2	2	8348	81EF	臯		
4094	館	3	3	834B	8218	館		館，異體字問題。
4107	葱	4	4	8358	8471	葱		
4115	藁	1	1	8360	85C1	藁		
4132		4	4	8371			血𠂇𠂇	眾，異體字問題。
4188	邨	5	5	83CB	90A8	邨		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4193	鈎	2	2	83D0	920E	鈎		
4226	飡	2	2	83F1	98E1	飡		
4256	鷄	2	2	8450	9DC4	鷄		雞，異體字問題。
4286	从	1	1	846E	4ECE	从		
4307	罵	2	2	84A5	99E1	罵		罵，異體字問題。
4371	儻	1	1	84E5	50F4	儻		
4437	廩	1	1	8568	5EEA	廩		廩，異體字問題。
4507	吳	28	28	85D0	5434	吳		吳，異體字問題。
4528	嶼	1	1	85E5	5D60	嶼		
4577	爰	1	1	8657	7231	爰		
4702	場	1	1	86F6	5872	場		
4735	壠	1	1	8758	58E0	壠		
4768	妬	1	1	8779	59AC	妬		妒，異體字問題。
4785	贗	1	1	87AC	8D0B	贗		贗，異體字問題。
4888	荔	2	2	8854	8318	荔		荔，異體字問題。
4902	兗	1	1	8862	5156	兗		
4909	弥	2	2	8869	5F25	弥		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5095	來	1	1	89A8	6765	來		來，異體字問題。
5290	攜	1	1	8ACE	643A	携		攜，異體字問題。
5316	晉	1	1	8AE8	664B	晋		
5402	朮	3	3	8BA1	672F	朮		
5405	藪	1	1	8BA4	6A37	藪		Unicode 呈現差異。
5439	栢	1	1	8BC6	6822	栢		
5549	藁	1	1	8C75	85F3	藁		
5568	开	1	1	8CAA	5F00	开		開，異體字問題。
5697	煑	5	5	8D6C	7151	煑		煮，異體字問題。
5753	潜	1	1	8DC6	6F5C	潜		
5797	灑	2	2	8DF2	7054	灑		
5850	厶	1	1	C6C9			厶	

附件二、《永和本薩天錫逸詩》校出錯字表

檔案原文摘錄	手動取代結果	備註
薩本「 ? 」作「餐」。(全部取代)	薩本「 淦 」作「餐」。(全部取代)	2
平生書劍 # 中師 (全部取代)	平生書劍 眾 中師 (全部取代)	2
洗 b 收龍帶雨還 (全部取代)	洗 金厶本 收龍帶雨還 (全部取代)	2

□說法如雲雨	△說法如雲雨	1
蒼茫半逐?聲散	蒼茫半逐雞聲散	1
蘚骨誰教0火侵	蘚骨誰教劫火侵	1
o攀楊柳弄春陰	戲攀楊柳弄春陰	1
「鐘」作「?」	「鐘」作「雞」	1
#山拱戴似兒孫（全部取代）	眾山拱戴似兒孫（全部取代）	2
西j隔屋送新酒	西鄰隔屋送新酒	1
竹?飛上繡春臺	竹雞飛上繡春臺	1
水宿風?鬢髮焦（全部取代）	水宿風淪鬢髮焦（全部取代）	2
馬上聞?夢早朝	馬上聞雞夢早朝	1
願學聞?舞劍雄	願學聞雞舞劍雄	1
j翁看竹借籃輿	鄰翁看竹借籃輿	1
五鳳樓前看??	五鳳樓前看鬥雞	1
?鳴犬吠人家近	雞鳴犬吠人家近	1
昂昂野鶴在?群	昂昂野鶴在雞群	1
毛本「壘」作「罍△中」	毛本「壘」作「罍△里」	1