

《漏網喁魚集》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
2008/5/28 葉冠孜 製作

一、《漏網喁魚集》一書(漏網喁魚集.xml)使用舊版造字 52 個，字頻 288 次，詳如附件一。這 52 個造字中，44 個可轉成 Windows XP 能顯示的字，字頻 240 次；另外 8 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 4134 的「衛」係「衛」字之異體，故以「衛」字取代。又編號 3899 的「担」字，係「擔」字之簡體字，亦以標準字的「擔」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 4242 的「鬪」字，Unicode 字型呈現為「鬪」，實際上仍為同一字。
- 3、待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 1908 的「衤  兩」。

附件一、《漏網喁魚集》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
297	厶	2	2	FBED			厶	
573	乚	1	1	FDC7			乚	
770	扌	1	1	FEEF			扌	
1908		1	1	9557			衤 厶 苜	待造字。
2749	枱	3	3	9AB1	67B1	枱		檯，異體字問題。
3021	忄	1	1	9C65			忄	
3789	冲	1	1	8154	51B2	冲		沖，異體字問題。
3799	刂	27	27	815E	5226	刂		劫，異體字問題。
3804	効	3	3	8163	52B9	効		效，異體字問題。
3814	却	1	1	816D	5374	却		卻，異體字問題。
3825	疊	8	8	8178	53E0	疊		疊，異體字問題。
3829	咏	1	1	817C	548F	咏		
3858	寃	6	6	81BB	5BC3	寃		冤，異體字問題。
3864	峯	3	3	81C1	5CEF	峯		峰，異體字問題。
3868	卮	1	1	81C5	5DF5	卮		
3883	徧	6	6	81D4	5FA7	徧		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3899	担	7	7	81E4	62C5	担		擔，異體字問題。
3938	櫬	3	3	824C	6AC8	櫬		凳，異體字問題。
3950	浜	7	7	8258	6D5C	浜		
3966	烟	35	35	8268	70DF	烟		煙，異體字問題。
3974	牀	2	2	8270	7240	牀		
3978	猪	7	7	8274	732A	猪		豬，異體字問題。
3996	疎	1	1	82A8	758E	疎		
4000	疴	1	1	82AC	75B4	疴		
4016	着	43	43	82BC	7740	着		著，異體字問題。
4028	稟	2	2	82C8	7980	稟		稟，異體字問題。
4034	穉	4	4	82CE	7A49	穉		
4036	窰	3	3	82D0	7AB0	窰		窯，異體字問題。
4038	竈	1	1	82D2	7AC8	竈		灶，異體字問題。
4039	疇	2	2	82D3	7ADA	疇		
4045	箒	1	1	82D9	7B92	箒		
4065	紮	1	1	82ED	7D25	紮		紮，異體字問題。
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4080	羣	10	10	82FC	7FA3	羣		群，異體字問題。
4090	脚	8	8	8347	811A	脚		腳，異體字問題。
4094	館	1	1	834B	8218	館		館，異體字問題。
4107	葱	1	1	8358	8471	葱		
4112	蔴	1	1	835D	8534	蔴		
4132		34	34	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	5	5	8373	885E	衛		衛，異體字問題。
4139	袴	1	1	8378	88B4	袴		
4146	覩	1	1	83A1	89A9	覩		
4173	駝	6	6	83BC	8EAD	駝		
4242	鬪	4	4	8442	9B2D	鬪		Unicode 呈現差異。
4447	捏	4	4	8572	63D1	捏		捏，異體字問題。
4702	場	1	1	86F6	5872	場		
5176		5	5	89F9			氵𠄎𠄎	沉，異體字問題。
5437	并	1	1	8BC4	5E77	并		并，異體字問題。
5472	瓦	3	3	8BE7			瓦	瓦，異體字問題。
5666	污	5	5	8D4D	6C5A	污		污，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5704	𣎵	1	1	8D73	6C0A	𣎵		
5706	敘	9	9	8D75	654D	敘		敘，異體字問題。