

《老學菴筆記》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
2008/6/24 葉冠孜 製作

一、《老學菴筆記》一書(老學菴筆記.xml)使用舊版造字 115 個，字頻 432 次，詳如附件一。這 115 個造字中，89 個可轉成 Windows XP 能顯示的字，字頻 304 次；另外 26 個字必須轉成構字式，字頻 128 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 3791 的「况」係「況」字之異體，故以「況」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 4158 的「譌」字，Unicode 字型呈現為「譌」，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 5514 的「𪔐」字，Unicode 編碼為 25095，由於 Windows XP 並不支援，仍須使用構字式「執𪔐」。

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、《老學菴筆記》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
38		1	1	FA65			王𠂔已	玳，異體字問題。
296	𠂔	16	16	FBEC			𠂔	
297	𠂔	16	16	FBED			𠂔	
298	𠂔	3	3	FBEE			𠂔	
299	𠂔	5	5	FBEF			𠂔	
300	𠂔	5	5	FBF0			𠂔	
379	支	8	8	FCA2			支	
461	※	2	2	FCF4			※	
507	遺	1	1	FD63			遺	
695	且	2	2	FEA4			且	
714	卜	2	2	FEB7			卜	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
720	夕	1	1	FEBD			夕	
761	夕	2	2	FEE6			夕	
762	夕	2	2	FEE7			夕	
768	夕	8	8	FEE8			夕	
773	夕	3	3	FEF2			夕	
899	馳	3	3	8ED3	99DE	馳		
1076	高	2	2	8FE7	5368	高		Unicode 呈現差異。
1077	稟	1	1	8FE8	7A3E	稟		
1091	灑	3	3	8FF6	58CD	灑		
1126	璫	1	1	905A	74B9	璫		
1133	瓌	1	1	9061	74CC	瓌		
1226	罰	2	2	90E0	7F78	罰		罰，異體字問題。
1350	褚	1	1	91BF	7987	褚		褚，異體字問題。
1354	禩	1	1	91C3	79A9	禩		
1565	繇	1	1	92F9	7DDC	繇		
1781	蓋	1	1	9475	854B	蓋		

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		
2142	裛	1	1	96C6	7758	裛		Unicode 呈現差異。
2151	𠂔	1	1	96CF	961D	𠂔		
2157	酌	1	1	96D5	9167	酌		
3791	况	1	1	8156	51B5	况		况，異體字問題。
3797	凭	1	1	815C	51ED	凭		
3800		2	2	815F			杀 𠂔 𠂔	刹，異體字問題。
3814	却	18	18	816D	5374	却		卻，異體字問題。
3817	廁	3	3	8170	53A0	廁		廁，異體字問題。
3819	廡	2	2	8172	53AE	廡		廡，異體字問題。
3826	叶	1	1	8179	53F6	叶		
3830	咤	1	1	817D	54A4	咤		
3831	詰	1	1	817E	5586	詰		
3846	壻	5	5	81AF	58FB	壻		Unicode 呈現差異。
3858	寃	2	2	81BB	5BC3	寃		寃，異體字問題。
3883	徧	1	1	81D4	5FA7	徧		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3891	慙	2	2	81DC	6159	慙		慙，異體字問題。
3911	攡	1	1	81F0	651F	攡		
3955	澗	2	2	825D	3D4E		汜 𡗗 澗	澗，異體字問題。
3966	烟	6	6	8268	70DF	烟		煙，異體字問題。
3968	煨	1	1	826A	7145	煨		
3971	熇	11	11	826D	71BA	熇		
3974	牀	6	6	8270	7240	牀		
3975	牕	4	4	8271	7255	牕		Unicode 呈現差異。
3978	猪	2	2	8274	732A	猪		豬，異體字問題。
3979	猫	10	10	8275	732B	猫		貓，異體字問題。
3990	甄	4	4	82A2	750E	甄		
3996	疎	3	3	82A8	758E	疎		
4016	着	6	6	82BC	7740	着		著，異體字問題。
4024	礮	4	4	82C4	792E	礮		
4028	稟	3	3	82C8	7980	稟		稟，異體字問題。
4034	穉	1	1	82CE	7A49	穉		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4036	窰	2	2	82D0	7AB0	窰		
4038	竈	1	1	82D2	7AC8	竈		
4041	豎	1	1	82D5	7AEA	豎		豎，異體字問題。
4080	羣	21	21	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		
4114	藥	1	1	835F	8616	藥		
4118	藥	5	5	8363	8602	藥		
4132		30	30	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	4	4	8373	885E	衛		衛，異體字問題。
4139	袴	4	4	8378	88B4	袴		
4145	羈	2	2	837E	898A	羈		Unicode 呈現差異。
4146	覩	1	1	83A1	89A9	覩		
4158	譌	1	1	83AD	8B4C	譌		Unicode 呈現差異。
4184	迹	9	9	83C7	8FF9	迹		跡，異體字問題。
4193	鈎	5	5	83D0	920E	鈎		鈎，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		
4214	鞞	1	1	83E5	97C0	鞞		
4232	隲	1	1	83F7	96B2	隲		
4242	鬪	1	1	8442	9B2D	鬪		Unicode 呈現差異。
4262	麪	8	8	8456	9EAA	麪		麵，異體字問題。
4265	鼃	3	3	8459	9F03	鼃		
4271	齧	2	2	845F	9F69	齧		
4275	义	1	1	8463	4E49	义		
4307	罵	9	9	84A5	99E1	罵		罵，異體字問題。
4450	勑	1	1	8575	524F	勑		勑，異體字問題。
4460	勅	21	21	85A1	52D1	勅		
4461	勔	1	1	85A2	52D4	勔		
4495	麪	1	1	85C4	9EA8	麪		
4507	吳	31	31	85D0	5434	吳		吳，異體字問題。
4575	繼	2	2	8655	7E7F	繼		
4603	憇	1	1	8671	6187	憇		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4816	亻	4	4	87CB			亻	
4878	宀	1	1	884A	5B82	宀		宀，異體字問題。
4888	荔	4	4	8854	8318	荔		荔，異體字問題。
4902	兗	1	1	8862	5156	兗		
4966	癩	1	1	88C4	764E	癩		
4996	斲	1	1	88E2	65B5	斲		
5008	幪	1	1	88EE	5E5E	幪		
5075	彳	1	1	8972			彳	
5176		4	4	89F9			氵 厶 宀	沉，異體字問題。
5218		5	5	8A64			宀 艹 覀	寬，異體字問題。
5230		1	1	8A70			宀 土 尙 且 口	壇，異體字問題。
5289	愨	1	1	8ACD	6164	愨		愨，異體字問題。
5434	湟	1	1	8BC1	6E7C	湟		湟，異體字問題。
5437	并	1	1	8BC4	5E77	并		并，異體字問題。
5463	岄	1	1	8BDE	65F9	岄		
5464	煥	1	1	8BDF	7157	煥		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5514	塾	1	1	8C52	25095		執𠂇皿	
5542	鑠	2	2	8C6E	93C1	鑠		
5558	龕	2	2	8C7E	9E84	龕		
5582	鞣	2	2	8CB8	97B5	鞣		
5590	朶	1	1	8CC0	6736	朶		朶，異體字問題。
5704	羶	2	2	8D73	6C0A	羶		
5706	敘	13	13	8D75	654D	敘		敘，異體字問題。
5795	媵	1	1	8DF0	3715		女𠂇連	