

# 《嘯亭雜錄》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  
2008/5/15 陳建安 製作

一、《嘯亭雜錄》一書(嘯亭雜錄.xml)使用舊版造字 196 個，字頻 1442 次，詳如附件一。這 196 個造字中，154 個可轉成 Windows XP 能顯示的字，字頻 1012 次；另外 42 個字必須轉成構字式，字頻 430 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、校對問題，舊版漢籍錯字，可用程式全部取代：在舊版漢籍電子文獻中即存在的錯字，因校對時的疏漏而未更正，持續留存在新版漢籍電子文獻中；若該造字的所有頻次，皆屬於錯誤使用的錯字情形，可以用程式全部取代為正確字形。如編號 3159 的「足△蘭」字，原字為「躡」。
- 2、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 4132 的「血△叢」係「眾」字之異體，故以「眾」字取代。又編號 1641 的「胆」字，係「膽」字之簡體字，亦以標準字的「膽」字取代。
- 3、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 3935 的「槩」字，Unicode 字型呈現為「槩」，實際上仍為同一字。
- 4、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 5749 的「片△庸」。

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 1123 的「𠄎」字，Unicode 編碼為 23C25，由於 Windows XP 並不支援，仍須使用構字式「毛△蒙」。

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、《嘯亭雜錄》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
296	△	19	19	FBEC			△	
297	△	13	13	FBED			△	
298	△	1	1	FBEE			△	
299	形	1	1	FBEF			形	
300	□	1	1	FBF0			□	
308	∞	13	13	FBF8			∞	
461		1	1	FCF4				※，校對問題，舊版漢籍錯字，可用程式全部取代。
489	隶	1	1	FD51			隶	
610	彡	1	1	FDEC			彡	
671	白	1	1	FE6A			白	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
675	食	2	2	FE6E			食	
691	牛	1	1	FE7E			牛	
766	糸	1	1	FEEB			糸	
821	脊	1	1	8E63	7718	脊		
868	休	1	1	8EB4	201F1		亻 厶 未	
957	遨	1	1	8F4E	906C	遨		
1010	珏	4	4	8FA5	73CF	珏		
1025	城	2	2	8FB4	73F9	城		
1092	瑞	3	3	8FF7	747A	瑞		
1123	毳	1	1	9057	23C25		毛 厶 蒙	
1124	恚	1	1	9058	3963		奎 厶 心	
1133	瓌	1	1	9061	74CC	瓌		
1353	糖	3	3	91C2	799F	糖		
1354	禩	9	9	91C3	79A9	禩		
1360	禩	1	1	91C9	79B5	禩		
1565	繇	1	1	92F9	7DDC	繇		
1622	艘	1	1	9373	8267	艘		
1641	胆	1	1	93A8	80C6	胆		膽，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1847	嶺	1	1	94D9	5D90	嶺		
1885	蠱	4	4	9540	882D	蠱		
1940	袂	1	1	9577	4109		衤 𠂔我	
1955	伺	1	1	95A8	201F6		亻 𠂔同	
1963	訖	1	1	95B0	8A5D	訖		
1964	還	3	3	95B1	8B5E	還		
1981	讐	1	1	95C2	8B90	讐		
2046	踪	6	6	9644	8E2A	踪		
2064	聃	1	1	9656	8065	聃		
2087	霽	1	1	966D	291A9		雨 𠂔 樹	
2131	弼	2	2	96BB	5F3B	弼		弼，異體字問題。
2345	鑽	1	1	97F4	9404	鑽		
2401	𧯂	1	1	986D	475F		豸 𠂔 契	
2434	隨	1	1	98B0	968F	隨		隨，異體字問題。
2487	髀	3	3	98E5	29A49		骨 𠂔 孝	
2885	叫	1	1	9B7A	544C	叫		
3004	籬	1	1	9C54	7BD0	籬		
3016	闐	1	1	9C60	28D91		門 𠂔 臭	闐，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3096	歲	1	1	9CD2	5D57	歲		歲，異體字問題。
3159	躡	1	1	9D52	28206		足𠂔蘭	躡，校對問題，舊版漢籍錯字，可用程式全部取代。
3780	儻	1	1	814B	5101	儻		
3787	冢	1	1	8152	51A2	冢		
3789	冲	7	7	8154	51B2	冲		冲，異體字問題。
3794	湊	1	1	8159	51D1	湊		湊，異體字問題。
3805	勅	5	5	8164	52C5	勅		
3814	却	7	7	816D	5374	却		卻，異體字問題。
3815	𡗗	3	3	816E	537A	𡗗		
3816	厓	2	2	816F	5393	厓		
3817	廁	1	1	8170	53A0	廁		廁，異體字問題。
3820	廈	6	6	8173	53A6	廈		廈，異體字問題。
3824	叙	1	1	8177	53D9	叙		敘，異體字問題。
3825	疊	1	1	8178	53E0	疊		疊，異體字問題。
3826	叶	1	1	8179	53F6	叶		
3833	噉	1	1	81A2	5611	噉		
3835	噉	2	2	81A4	564D	噉		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3836	喻	1	1	81A5	564F	喻		
3846	壻	15	15	81AF	58FB	壻		Unicode 呈現差異。
3849	獎	1	1	81B2	596C	獎		獎，異體字問題。
3858	寃	19	19	81BB	5BC3	寃		冤，異體字問題。
3864	峯	32	32	81C1	5CEF	峯		峰，異體字問題。
3866	崐	3	3	81C3	5D10	崐		
3868	卮	2	2	81C5	5DF5	卮		卮，異體字問題。
3883	徧	1	1	81D4	5FA7	徧		
3884	忽	6	6	81D5	6031	忽		
3886	悞	1	1	81D7	609E	悞		
3899	担	2	2	81E4	62C5	担		擔，異體字問題。
3916	敷	1	1	81F5	656D	敷		
3921	皙	5	5	81FA	6673	皙		皙，異體字問題。
3939	礮	1	1	824D	6AFD	礮		
3955	澗	4	4	825D	3D4E		汜 㒰 澗	澗，異體字問題。
3956	渚	3	3	825E	6F74	渚		Unicode 呈現差異。
3966	烟	9	9	8268	70DF	烟		煙，異體字問題。
3969	煊	1	1	826B	714A	煊		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3971	熇	1	1	826D	71BA	熇		
3974	牀	11	11	8270	7240	牀		
3976	犁	46	46	8272	7282	犁		
3977	狗	4	4	8273	72E5	狗		
3978	猪	4	4	8274	732A	猪		Unicode 呈現差異。
3988	璦	2	2	827E	7482	璦		
4000	疴	1	1	82AC	75B4	疴		
4008	癰	7	7	82B4	766F	癰		
4015	眦	1	1	82BB	7726	眦		
4016	着	14	14	82BC	7740	着		著，異體字問題。
4024	礫	1	1	82C4	792E	礫		
4027	裨	1	1	82C7	7986	裨		裨，異體字問題。
4028	稟	4	4	82C8	7980	稟		稟，異體字問題。
4029	褱	1	1	82C9	799D	褱		
4036	窰	1	1	82D0	7AB0	窰		
4041	豎	1	1	82D5	7AEA	豎		
4044	筋	1	1	82D8	7B6F	筋		
4069	綑	1	1	82F1	7DB3	綑		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4074	縑	1	1	82F6	7E27	縑		
4080	羣	42	42	82FC	7FA3	羣		群，異體字問題。
4085	耻	4	4	8342	803B	耻		
4089	脇	1	1	8346	8107	脇		
4090	脚	2	2	8347	811A	脚		脚，異體字問題。
4094	館	1	1	834B	8218	館		
4097	芑	1	1	834E			艹 厶 巳	芑，異體字問題。
4104	莅	7	7	8355	8385	莅		
4107	葱	6	6	8358	8471	葱		
4109	葛	1	1	835A	84AD	葛		
4115	藁	1	1	8360	85C1	藁		
4129	虬	1	1	836E	866C	虬		
4131	劬	2	2	8370	8842	劬		
4132		278	278	8371			血 厶 亼	眾，異體字問題。
4134	衛	255	255	8373	885E	衛		衛，異體字問題。
4139	袴	3	3	8378	88B4	袴		
4146	覩	3	3	83A1	89A9	覩		
4149	覩	1	1	83A4	89BC	覩		



編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4176	輒	1	1	83BF	8F19	輒		輒，異體字問題。
4184	迹	5	5	83C7	8FF9	迹		
4188	邨	1	1	83CB	90A8	邨		
4193	鈎	5	5	83D0	920E	鈎		
4195	鉢	3	3	83D2	9262	鉢		
4203	隣	1	1	83DA	96A3	隣		鄰，異體字問題。
4213	鞞	2	2	83E4	97BE	鞞		
4220	穎	8	8	83EB	9834	穎		穎，異體字問題。
4221	頰	1	1	83EC	9833	頰		
4226	飡	1	1	83F1	98E1	飡		
4237	髡	2	2	83FC	9AE0	髡		髡，異體字問題。
4242	鬪	2	2	8442	9B2D	鬪		Unicode 呈現差異。
4259	麇	1	1	8453	9E95	麇		Unicode 呈現差異。
4262	麪	2	2	8456	9EAA	麪		
4263	麴	1	1	8457	9EAF	麴		
4267	鼈	1	1	845B	9F08	鼈		
4277		1	1	8465			金𠂔义	釵，異體字問題。
4283	戲	1	1	846B	622F	戲		戲，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4293		1	1	8475			∞束△火	棘△火，異體字問題。
4307	罵	13	13	84A5	99E1	罵		罵，異體字問題。
4348	惚	1	1	84CE	506C	惚		
4432		1	1	8563			形絲厂墨	纏，異體字問題。
4437	廩	2	2	8568	5EEA	廩		廩，異體字問題。
4460	勅	2	2	85A1	52D1	勅		
4502	俎	3	3	85CB	723C	俎		俎，異體字問題。
4504	伊	1	1	85CD	541A	伊		
4507	吳	188	188	85D0	5434	吳		吳，異體字問題。
4515	呪	3	3	85D8	546A	呪		
4518	响	1	1	85DB	54CD	响		響，異體字問題。
4558	廐	7	7	8644	53A9	廐		廐，異體字問題。
4631	垚	1	1	86AF	5788	垚		
4702	場	2	2	86F6	5872	場		場，異體字問題。
4735	壠	2	2	8758	58E0	壠		
4742		8	8	875F			衣△執	褻，異體字問題。
4750	卅	1	1	8767			卅	**，校對問題，舊版漢籍錯字，可用程式全部取代。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4768	妬	1	1	8779	59AC	妬		妒，異體字問題。
4785	贗	4	4	87AC	8D0B	贗		贗，異體字問題。
4803	腭	1	1	87BE	816D	腭		
4805	瀆	1	1	87C0	51DF	瀆		瀆，異體字問題。
4816	亻	2	2	87CB			亻	
4878	宀	3	3	884A	5B82	宀		宀，異體字問題。
4889	毳	1	1	8855	3CAA		登𠂇毛	
4907	橐	2	2	8867	69D6	橐		橐，異體字問題。
4910		2	2	886A			汙𠂇畷	淫，異體字問題。
4930	岑	1	1	887E	5CBA	岑		岑，校對問題，舊版漢籍錯字，可用程式全部取代。
4940	韵	1	1	88AA	97F5	韵		
4966	癩	1	1	88C4	764E	癩		
5022	厲	1	1	88FC	5EBD	厲		
5046	彝	2	2	8955	5F5C	彝		
5101	馱	5	5	89AE	4B7E		馬𠂇犬	馱，異體字問題。
5111	旃	1	1	89B8	65BE	旃		旃，異體字問題。
5163	憂	8	8	89EC	621E	憂		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5173	漱	3	3	89F6	6F44	漱		漱，異體字問題。
5176		1	1	89F9			氵 厶 冗	沉，異體字問題。
5189	冗	2	2	8A47	5B82	冗		冗，異體字問題。
5216	鞣	1	1	8A62	97CA	鞣		
5218		43	43	8A64			形 宀 覓、口	寬，異體字問題。
5225	卹	4	4	8A6B	460F		血 冫 卩	卹，異體字問題。
5241	鷺	1	1	8A7B	9D5E	鷺		鷺，異體字問題。
5263		2	2	8AB3			彳 厶 力	務，異體字問題。
5289	慤	23	23	8ACD	6164	慤		慤，異體字問題。
5298	昨	2	2	8AD6	65FF	昨		
5306		1	1	8ADE			竹 冫 閒	簡，異體字問題。
5408	絳	1	1	8BA7	7D4D	絳		
5439	栢	4	4	8BC6	6822	栢		
5447	躡	1	1	8BCE	8E83	躡		
5541	蠹	1	1	8C6D	8827	蠹		
5559	欸	3	3	8CA1	6B35	欸		欸，異體字問題。
5584	愞	1	1	8CBA	611E	愞		
5587	欸	2	2	8CBD	6B58	欸		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5590	朶	11	11	8CC0	6736	朶		朶，異體字問題。
5591	杞	5	5	8CC1	233CC		木𠂇巳	杞，異體字問題。
5595	塚	3	3	8CC5	579C	塚		塚，異體字問題。
5629	淦	2	2	8CE7	6D24	淦		
5702	厨	3	3	8D71	53A8	厨		廚，異體字問題。
5705	廡	4	4	8D74	5ED0	廡		廡，異體字問題。
5706	敍	21	21	8D75	654D	敍		敍，異體字問題。
5749		1	1	8DC2			片𠂇庸	待造字。
5764	籩	4	4	8DD1	7C5E	籩		