

《萬曆野獲編》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
2008/7/7 陳維君 製作

一、《萬曆野獲編》一書(萬曆野獲編.xml)使用舊版造字 319 個，字頻 3330 次，詳如附件一。這 319 個造字中，227 個可轉成 Windows XP 能顯示的字，字頻 2293 次；另外 92 個字必須轉成構字式，字頻 1037 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 5545 的「鬻」字，原字應為「鬻」。
- 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「血𠂇𠂇」係「眾」字之異體，故以「眾」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。
- 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 4242 的「鬪」字，Unicode 字型呈現為「鬪」，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 1633 的「𠂇𠂇𠂇」。

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 1359 的「襪」字，Unicode 編碼為 25725，由於 Windows XP 並不支援，仍須使用構字式「衤𠂔冀」。

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、《萬曆野獲編》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
12	陟	1	1	FA4B	28E97		卩𠂔走	
33	葛	1	1	FA60	85A5	葛		
38		1	1	FA65			王𠂔己	玨，異體字問題。
296	𠂔	22	22	FBEC			𠂔	
297	𠂔	42	42	FBED			𠂔	
298	𠂔	5	5	FBEE			𠂔	
299	𠂔	6	6	FBEF			𠂔	
300	𠂔	6	6	FBF0			𠂔	
304	𠂔	1	1	FBF4			𠂔	
463	𠂔	1	1	FCF6			𠂔	
510	𠂔	1	1	FD66			𠂔	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
521	茱	1	1	FD71			茱	
567	苧	1	1	FDC1			苧	
630	畝	1	1	FE41			畝	
652	鳥	1	1	FE57			鳥	
675	食	1	1	FE6E			食	
685	衤	1	1	FE78			衤	
715	彡	2	2	FEB8			彡	
738	亾	1	1	FECF			亾	
745	亾	1	1	FED6			亾	
750	冫	1	1	FEDB			冫	
759	广	1	1	FEE4			广	
768	艹	3	3	FEED			艹	
780	彡	2	2	FEF9			彡	
784	亾	1	1	FEFD			亾	
788	炝	2	2	8E42	708C	炝		
899	馳	2	2	8ED3	99DE	馳		
912	鷓	1	1	8EE0	9DC0	鷓		
923		2	2	8EEB			𠄎火产兼𠄎	待造字。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
932	兌	4	4	8EF4	7683	兌		
938	燮	1	1	8EFA	7215	燮		
985	慙	1	1	8F6A	6197	慙		
1000		5	5	8F79			王𠂇巳	玨，異體字問題。
1010	玨	2	2	8FA5	73CF	玨		玨，異體字問題。
1022	珺	3	3	8FB1	73E4	珺		
1024	虵	1	1	8FB3	8675	虵		
1039	稟	4	4	8FC2	69C0	稟		
1067	碁	1	1	8FDE	68CA	碁		碁，異體字問題。
1075	駟	2	2	8FE6	9A1F	駟		
1092	瑞	5	5	8FF7	747A	瑞		
1100	身	1	1	9040	8EAB	身		
1133	瓌	5	5	9061	74CC	瓌		
1201	昵	1	1	90C7	7724	昵		
1213	睽	1	1	90D3	7743	睽		
1359	襪	1	1	91C8	25725		襪 𠂇冀	
1381	穉	1	1	91DE	7A3A	穉		
1463	攤	1	1	9271	651E	攤		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
1467	纂	1	1	9275	25C67		𦰃 𠂇 秦	
1530	紕	1	1	92D6	7CFA	紕		
1559		1	1	92F3			𠂇 广 目 无 𠂇	廢，異體字問題。
1600	兪	3	3	935D	3E84		彡 𠂇 兪	
1618	睥	1	1	936F	2522D		目 𠂇 𠂇	
1622	艘	1	1	9373	8267	艘		
1631	燠	3	3	937C	24424		火 𠂇 熙	
1632	奩	2	2	937D	215E8		大 𠂇 互	
1633		1	1	937E			艹 𠂇 涯	待造字。
1634	忬	8	8	93A1	5FEC	忬		
1635	燻	13	13	93A2	3E05		火 𠂇 節	
1636	眈	1	1	93A3	6616	眈		
1644	謹	3	3	93AB	6F0C	謹		
1646	脬	3	3	93AD	8117	脬		
1668	臙	1	1	93C3	81C8	臙		
1675	𦰃	1	1	93CA	2319F		此 𠂇 旦	
1691	鬃	2	2	93DA	4BFC		鬃 𠂇 狄	
1704	擘	2	2	93E7	3BAE		木 𠂇 擘	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
1705	啓	2	2	93E8	21E58		攴△山	
1707	僂	1	1	93EA	349C		亻△與	
1708	梠	1	1	93EB	3B7F		木△昂	
1739	燼	5	5	944B	3DFF		火△僉	
1782		1	1	9476			土△龠	待造字。
1785	漬	3	3	9479	23E5F		氵△貢	
1796		3	3	94A6			王△ㄩ	待造字。
1812	穊	4	4	94B6	4185		禾△資	
1899	恠	1	1	954E	3995		忄△隆	
1902	顛	1	1	9551	29507		皇△頁	
1905	裨	1	1	9554	88E9	裨		
1912	襍	1	1	955B	894D	襍		
1919	僂	2	2	9562	2032F		亻△离	
1930	鼎	1	1	956D	4D7A		土△鼎	
1944	簪	1	1	957B	8A67	簪		
1981	讐	2	2	95C2	8B90	讐		讎，異體字問題。
1982	調	1	1	95C3	8B87	調		
2046	踪	1	1	9644	8E2A	踪		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
2050	蹠	1	1	9648	8E4F	蹠		
2069	軌	2	2	965B	8F04	軌		
2142	冢	1	1	96C6	7758	冢		Unicode 呈現差異。
2189	攴	1	1	96F5	23C3D		毛△瞿	毘，異體字問題。
2190	毳	1	1	96F6	6BFA	毳		Unicode 呈現差異。
2274	鎡	1	1	97AD	9348	鎡		
2298	殼	1	1	97C5	3C89		殼△卵	
2345	鑽	1	1	97F4	9404	鑽		
2371	玕	1	1	984F	3EB9		王△弁	
2408	鬧	3	3	9874	9599	鬧		鬧，異體字問題。
2409	闕	1	1	9875	95A7	闕		闕，異體字問題。
2440	璫	1	1	98B6	7499	璫		
2461	葱	1	1	98CB	44E4		艹△忽	葱，異體字問題。
2505	翻	1	1	98F7	98DC	翻		
2544	昏	2	2	995F	661A	昏		
2602	鰐	1	1	99BB	9C10	鰐		
2671	畱	1	1	9A41	7550	畱		
2681	步	1	1	9A4B	3C54		此△夕	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
2693		1	1	9A57			𠄎 ^竹 行缶示 ^口	籩，異體字問題。
2731	雀	5	5	9A7D			宀 ^隹	雀，異體字問題。
2834	女	1	1	9B47	6535	女		
2874	饴	1	1	9B6F	9939	饴		
3042	堵	1	1	9C7A	5835	堵		
3113	彡	1	1	9CE3	6245	彡		
3255	善	284	284	9DD4	5584	善		
3288		1	1	9DF5			王 ^厶 宏	待造字。
3770	个	5	5	8141	4E2A	个		個，異體字問題。
3787	冢	35	35	8152	51A2	冢		
3789	冲	3	3	8154	51B2	冲		沖，異體字問題。
3801	剗	3	3	8160	5257	剗		
3804	効	10	10	8163	52B9	効		
3813	邛	3	3	816C	536D	邛		
3814	却	38	38	816D	5374	却		卻，異體字問題。
3815	𡗗	3	3	816E	537A	𡗗		𡗗，異體字問題。
3816	厓	11	11	816F	5393	厓		
3817	廁	1	1	8170	53A0	廁		廁，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3818	廩	1	1	8171	53AB	廩		廩，異體字問題。
3825	疊	1	1	8178	53E0	疊		疊，異體字問題。
3826	叶	6	6	8179	53F6	叶		
3829	咏	18	18	817C	548F	咏		
3830	咤	1	1	817D	54A4	咤		
3835	噍	1	1	81A4	564D	噍		
3840	坂	18	18	81A9	5742	坂		
3843	堦	1	1	81AC	5826	堦		
3846	墀	53	53	81AF	58FB	墀		Unicode 呈現差異。
3849	獎	1	1	81B2	596C	獎		
3852	嫵	1	1	81B5	5A63	嫵		
3858	寃	1	1	81BB	5BC3	寃		寃，異體字問題。
3860	尠	6	6	81BD	5C20	尠		
3864	峯	55	55	81C1	5CEF	峯		峰，異體字問題。
3883	徧	33	33	81D4	5FA7	徧		
3886	悞	3	3	81D7	609E	悞		
3891	慙	11	11	81DC	6159	慙		慙，異體字問題。
3892	睺	1	1	81DD	3B0B		日 𠄎 侯	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3900	挖	2	2	81E5	62D5	挖		
3909	擡	3	3	81EE	64E1	擡		
3915	斃	1	1	8459	6585	斃		
3916	斃	4	4	81F5	656D	斃		
3921	暫	5	5	81FA	6673	暫		
3922	替	1	1	81FB	23293		執 [△] 日	
3924	菴	1	1	81FD	671E	菴		期，異體字問題。
3931	椀	3	3	8245	6900	椀		
3932	椀	2	2	8246	6936	椀		
3934	椀	2	2	8248	69C5	椀		
3935	槩	6	6	8249	69E9	槩		概，異體字問題。
3938	橈	3	3	824C	6AC8	橈		
3949	涖	16	16	8257	6D96	涖		
3959	潛	1	1	8261	6FF3	潛		潛，異體字問題。
3966	烟	19	19	8268	70DF	烟		煙，異體字問題。
3971	熿	11	11	826D	71BA	熿		
3974	牀	7	7	8270	7240	牀		
3975	牕	2	2	8271	7255	牕		Unicode 呈現差異。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3976	犁	5	5	8272	7282	犁		
3977	狗	6	6	8273	72E5	狗		
3978	猪	10	10	8274	732A	猪		猪，異體字問題。
3981	珉	9	9	8277	73C9	珉		
3990	甄	1	1	82A2	750E	甄		
3996	疎	19	19	82A8	758E	疎		
4000	疴	1	1	82AC	75B4	疴		
4010	皐	15	15	82B6	7690	皐		
4011	皦	2	2	82B7	76A1	皦		
4013	盃	5	5	82B9	76CC	盃		
4015	眦	5	5	82BB	7726	眦		
4016	着	23	23	82BC	7740	着		著，異體字問題。
4017	螻	1	1	82BD	880F	螻		蟹，異體字問題。
4024	礮	3	3	82C4	792E	礮		
4032	秣	1	1	82CC	2578A		禾𪎭未	秣，異體字問題。
4034	穉	3	3	82CE	7A49	穉		
4035	穽	2	2	82CF	7A7D	穽		
4036	窰	12	12	82D0	7AB0	窰		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4038	竈	4	4	82D2	7AC8	竈		
4039	竝	1	1	82D3	7ADA	竝		
4042	笋	5	5	82D6	7B0B	笋		
4044	筍	3	3	82D8	7B6F	筍		
4046		1	1	82DA			竹 𠂇 𠂇 𠂇 𠂇 𠂇 𠂇	筴，異體字問題。
4048		1	1	82DC			形 竹 广 虎 口	箴，異體字問題。
4055	糝	5	5	82E3	7C83	糝		
4058	糶	1	1	82E6	7CC9	糶		
4067	綉	2	2	82EF	7D89	綉		繡，異體字問題。
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。
4074	縑	3	3	82F6	7E27	縑		
4080	羣	95	95	82FC	7FA3	羣		群，異體字問題。
4081	翱	3	3	82FD	7FFA	翱		
4089	脇	3	3	8346	8107	脇		脅，異體字問題。
4091	臯	8	8	8348	81EF	臯		
4097	芑	1	1	834E			艹 𠂇 巳	芑，異體字問題。
4107	葱	1	1	8358	8471	葱		
4114	蘖	4	4	835F	8616	蘖		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4115	藁	14	14	8360	85C1	藁		
4129	虬	3	3	836E	866C	虬		
4131	劬	3	3	8370	8842	劬		
4132		147	147	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	313	313	8373	885E	衛		衛，異體字問題。
4138	裘	1	1	8377	88A0	裘		
4139	袴	10	10	8378	88B4	袴		
4140	衽	3	3	8379	88B5	衽		
4142	裛	48	48	837B	8943	裛		
4145	羈	8	8	837E	898A	羈		Unicode 呈現差異。
4146	覩	29	29	83A1	89A9	覩		
4156	謠	1	1	83AB	8B1F	謠		
4160	謫	1	1	83AF	8B81	謫		
4167	賫	11	11	83B6	8CEB	賫		
4176	輒	3	3	83BF	8F19	輒		
4184	迹	12	12	83C7	8FF9	迹		
4186	遡	5	5	83C9	9061	遡		
4188	邨	8	8	83CB	90A8	邨		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4190	邳	3	3	83CD	90C4	邳		
4193	鈎	10	10	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	13	13	83D2	9262	鉢		
4198	鑛	9	9	83D5	945B	鑛		
4220	穎	1	1	83EB	9834	穎		穎，異體字問題。
4221	頰	1	1	83EC	9833	頰		
4232	隲	3	3	83F7	96B2	隲		隲，異體字問題。
4237	髡	11	11	83FC	9AE0	髡		髡，異體字問題。
4238	髡	1	1	83FD	9AE2	髡		
4240	髡	1	1	8440	9AF4	髡		
4242	鬪	38	38	8442	9B2D	鬪		Unicode 呈現差異。
4251	緞	1	1	844B	9C15	緞		
4254	鴟	1	1	844E	9D7E	鴟		
4259	麇	3	3	8453	9E95	麇		Unicode 呈現差異。
4260	麇	2	2	8454	9E9E	麇		
4262	麪	9	9	8456	9EAA	麪		
4263	麪	1	1	8457	9EAF	麪		
4265	鼃	1	1	8459	9F03	鼃		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4266	鼉	1	1	845A	9F02	鼉		
4268	鼈	1	1	845C	9F17	鼈		
4278	亘	8	8	8466	4E98	亘		
4307	罵	21	21	84A5	99E1	罵		罵，異體字問題。
4353	菟	2	2	84D3	83AC	菟		
4374	漼	3	3	84E8	3D36		氵 厶 隹	
4439		1	1	856A			形 疋 巳 心 口	跽，異體字問題。
4447	捏	10	10	8572	63D1	捏		捏，異體字問題。
4461	勛	8	8	85A2	52D4	勛		
4477		2	2	85B2			形 艹 勹 一 口	芍，異體字問題。
4482	廛	1	1	85B7	53AA	廛		廛，異體字問題。
4493	特	1	1	85C2	3E40		牛 厶 孛	犝，異體字問題。
4495	麩	1	1	85C4	9EA8	麩		
4504	呶	1	1	85CD	541A	呶		
4515	呪	14	14	85D8	546A	呪		
4533	唵	1	1	85EA	552B	唵		
4546	鷓	3	3	85F7	9D76	鷓		
4550	疋	12	12	85FB	24D13		疋	疋，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4603	憇	2	2	8671	6187	憇		
4619	但	1	1	86A3	4F39	但		
4657	竣	2	2	86C9	57C8	竣		
4667	埙	20	20	86D3	57E8	埙		
4755	摯	2	2	876C	66AC	摯		
4768	妬	26	26	8779	59AC	妬		妒，異體字問題。
4769	憇	2	2	877A	60E5	憇		Unicode 顯示差異。
4785	贗	9	9	87AC	8D0B	贗		贗，異體字問題。
4805	瀆	2	2	87C0	51DF	瀆		瀆，異體字問題。
4823	窑	1	1	87D2	7A91	窑		
4845		1	1	87E8			女𠃉間	待造字。
4864	鎬	1	1	87FB	9461	鎬		
4904	尅	9	9	8864	5C05	尅		
4907	橐	19	19	8867	69D6	橐		橐，異體字問題。
4934	髮	1	1	88A4	9B09	髮		
4961	盼	1	1	88BF	80B9	盼		
4965	橄	1	1	88C3			木𠃉救	
4996	斲	14	14	88E2	65B5	斲		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5008	幪	4	4	88EE	5E5E	幪		
5028	昇	2	2	8943	662A	昇		
5032	焜	1	1	8947	3B0E		焜	
5074	尫	1	1	8971	5C2B	尫		尫，異體字問題。
5086		5	5	897D			冗厶卓	罩，異體字問題。
5101	馱	1	1	89AE	4B7E		馬厶犬	馱，異體字問題。
5116	侑	1	1	89BD	4F37	侑		
5122	褻	1	1	89C3	88F5	褻		
5154	璿	1	1	89E3	7468	璿		
5159	鵠	1	1	89E8	9D60	鵠		
5174		1	1	89F7			𠃉刀冗、𠃉	兔，異體字問題。
5176		26	26	89F9			氵厶冗	沉，異體字問題。
5194		3	3	8A4C			𠃉扌畝且𠃉	擅，異體字問題。
5195		8	8	8A4D			𠃉木畝且𠃉	檀，異體字問題。
5209	摑	1	1	8A5B	6386	摑		
5218		60	60	8A64			𠃉宀寬、𠃉	寬，異體字問題。
5219	妃	455	455	8A65			女厶巳	妃，異體字問題。
5221		45	45	8A67			𠃉白米𠃉	粵，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5230		49	49	8A70			𠄎土尙且𠄎	壇，異體字問題。
5241	鶯	3	3	8A7B	9D5E	鶯		鶯，異體字問題。
5289	慤	1	1	8ACD	6164	慤		慤，異體字問題。
5292	粵	3	3	8AD0	65C9	粵		
5312		2	2	8AE4			𠄎女乚冈比𠄎	媿，異體字問題。
5316	晉	2	2	8AE8	664B	晉		晉，異體字問題。
5338	暎	2	2	8AFE	668E	暎		
5378	巉	1	1	8B67	5DD9	巉		
5392	息	2	2	8B75	60A4	息		
5396	鎔	1	1	8B79	938B	鎔		
5434	湟	3	3	8BC1	6E7C	湟		湟，異體字問題。
5437	并	297	297	8BC4	5E77	并		并，異體字問題。
5439	栢	1	1	8BC6	6822	栢		
5442	枿	1	1	8BC9	67BF	枿		
5451	玳	1	1	8BD2	7398	玳		
5452	轆	1	1	8BD3	8F5C	轆		
5478	飴	1	1	8BED	4D77	飴		
5479	蠅	1	1	8BEE	9F03		圭厶黽	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5491	槓	7	7	8BFA	6967	槓		
5540	檨	4	4	8C6C	6A53	檨		
5541	蠹	2	2	8C6D	8827	蠹		蠹，異體字問題。
5545	鬻	1	1	8C71	4C1E		鬻	鬻，校對問題，舊版漢籍錯字，可用程式全部取代。
5551	賫	5	5	8C77	8CF7	賫		
5558	麓	1	1	8C7E	9E84	麓		
5567	嶼	15	15	8CA9	3817		山 ㄩ 奧	嶼，異體字問題。
5583	殮	4	4	8CB9	98F1	殮		殮，異體字問題。
5585	寓	1	1	8CBB	3762		宀 厶 禹	
5591	杞	7	7	8CC1	233CC		木 ㄩ 巳	杞，異體字問題。
5666	污	17	17	8D4D	6C5A	污		污，異體字問題。
5670	暈	1	1	8D51	66C5	暈		
5704	羶	6	6	8D73	6C0A	羶		
5705	廐	2	2	8D74	5ED0	廐		廐，異體字問題。
5706	敘	64	64	8D75	654D	敘		敘，異體字問題。
5744	萑	1	1	8DBD	84A6	萑		
5764	籊	2	2	8DD1	7C5E	籊		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5772	澁	38	38	8DD9	6FD9	澁		
5806	膾	1	1	8DFB	4408		月 厶 直	
5854	幺	1	1	C6CD	5E7A	幺		