

# 《齊東野語》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  
2008/6/19 葉冠孜 製作

一、《齊東野語》一書(齊東野語.xml)使用舊版造字 227 個,字頻 1104 次,詳如附件一。這 227 個造字中,179 個可轉成 Windows XP 能顯示的字,字頻 896 次;另外 48 個字必須轉成構字式,字頻 208 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 5702 的「厨」字,原字應為「廁」。
- 2、校對問題,未確定字形,必須人工取代:在舊版漢籍電子文獻中即存在的未確定字形,持續留存在新版漢籍電子文獻中,必須人工一一比對修改,如編號 17「●」,可參考附件二「●」字形手動取代表。
- 3、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 2938 的「簪」係「簪」字之異體,故以「簪」字取代。又編號 4067 的「綉」字,係「繡」字之簡體字,亦以標準字的「繡」字取代。
- 4、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。

5、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 1578 的「日𠄎冉」。

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 1520 的「𠄎」字，Unicode 編碼為 3B20，由於 Windows XP 並不支援，仍須使用構字式「日𠄎𠄎」。

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

五、《齊東野語》（齊東野語.xml）中有「●」字形 4 個，「●」字形的處理可參考附件二「●」字形手動取代表。附件二欄位說明如下：

- 甲、頁碼：「●」字形所在位置之原書頁碼。
- 乙、檔案原文摘錄：摘錄檔案中有「●」的文句。
- 丙、手動取代結果：「●」的正確原書字形，如為待造字或 Windows XP 無法顯示，則以構字式表示。
- 丁、備註：註明相同文句出現的次數（未標明者為只出現一次）及其他需註明事項。

附件一、《齊東野語》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
8	賦	1	1	FA47	27E10		貝𠄎貳	
17	●	4	4	FA50			●	未確定字形，人工取代，參考附件二。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
64	颺	1	1	FAA1	98B7	颺		颺，異體字問題。
74	髻	2	2	FAAB	29B06		髻△𠂇	髻，異體字問題。
296	厶	9	9	FBEC			厶	
297	厶	19	19	FBED			厶	
298	厶	2	2	FBEE			厶	
299	形	9	9	FBEF			形	
300	口	9	9	FBF0			口	
493	胃	2	2	FD55	5190	胃		
639	丿	1	1	FE4A			丿	
685	丿	1	1	FE78			丿	
711	丿	1	1	FEB4			丿	
741	白	2	2	FED2			白	
746	丿	2	2	FED7			丿	
747	圭	1	1	FED8			圭	
759	广	2	2	FEE4			广	
768	++	1	1	FEED			++	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
772	之	1	1	FEF1			之	
773	厶	1	1	FEF2			厶	
780	彳	1	1	FEF9			彳	
803	烜	2	2	8E51	70A3	烜		
853	焞	1	1	8EA5	7134	焞		
893	峩	1	1	8ECD	5CE9	峩		峨，異體字問題。
899	馳	4	4	8ED3	99DE	馳		
979	躑	1	1	8F64	56CF	躑		
1010	珏	3	3	8FA5	73CF	珏		
1022	瑤	3	3	8FB1	73E4	瑤		
1067	碁	3	3	8FDE	68CA	碁		棋，異體字問題。
1076	高	1	1	8FE7	5368	高		Unicode 呈現差異。
1126	璫	1	1	905A	74B9	璫		
1133	瓌	2	2	9061	74CC	瓌		
1226	罰	1	1	90E0	7F78	罰		罰，異體字問題。
1261	庸	7	7	9144	5EAF	庸		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1269	臺	1	1	914C	21334		直厶土	
1281	儂	5	5	9158	349F		亻厶裏	
1334	眇	2	2	91AF	26548		耳厶少	
1394	啗	1	1	91EB	35D6		口厶啗	啗，異體字問題。
1432	瑋	12	12	9252	7438	瑋		
1435	箊	1	1	9255			𦏧厶甫	
1437	犮	2	2	9257	7265	犮		
1465	炒	1	1	9273	7C86	炒		
1518	佻	5	5	92CA	20201		𠃉亻尸匕口	佻，異體字問題。
1520	嚙	2	2	92CC	3B20		日厶煞	
1544	絙	6	6	92E4	7D59	絙		
1578		2	2	9347			日厶冉	待造字。
1603	倂	1	1	9360	5034	倂		
1642	脉	5	5	93A9	8109	脉		
1871	蝨	2	2	94F1	87C1	蝨		
1893	袞	18	18	9548	886E	袞		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1912	襍	2	2	955B	894D	襍		
1981	讐	17	17	95C2	8B90	讐		
1982	調	1	1	95C3	8B87	調		
2046	踪	1	1	9644	8E2A	踪		
2050	蹠	1	1	9648	8E4F	蹠		
2146	蝸	2	2	96CA	8739	蝸		
2234	攷	1	1	9763	38B8		弜𠂇女	
2325	睭	1	1	97E0	77A4	睭		
2408	鬧	1	1	9874	9599	鬧		鬧，異體字問題。
2409	闕	1	1	9875	95A7	闕		闕，異體字問題。
2487	髀	1	1	98E5	29A49		骨𠂇孝	
2524	饕	1	1	994B	994D	饕		
2592	贓	1	1	99B1	8D1C	贓		贓，異體字問題。
2598	鄒	1	1	99B7	9381	鄒		
2885	叫	2	2	9B7A	544C	叫		
2938	簪	1	1	9BD1	7C2E	簪		簪，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
2998	棧	1	1	9C4E			木厶戈	
3096	歲	60	60	9CD2	5D57	歲		歲，異體字問題。
3791	况	6	6	8156	51B5	况		况，異體字問題。
3792	涼	1	1	8157	51C9	涼		涼，異體字問題。
3794	湊	1	1	8159	51D1	湊		湊，異體字問題。
3804	効	1	1	8163	52B9	効		效，異體字問題。
3805	勅	2	2	8164	52C5	勅		敕，異體字問題。
3809	匱	1	1	8168	5333	匱		
3814	却	34	34	816D	5374	却		卻，異體字問題。
3816	厓	3	3	816F	5393	厓		
3817	廁	8	8	8170	53A0	廁		廁，異體字問題。
3824	叙	4	4	8177	53D9	叙		敘，異體字問題。
3829	咏	10	10	817C	548F	咏		
3830	咤	1	1	817D	54A4	咤		
3835	噍	1	1	81A4	564D	噍		
3846	壻	7	7	81AF	58FB	壻		Unicode 呈現差異。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3852	媼	2	2	81B5	5A63	媼		
3858	寃	27	27	81BB	5BC3	寃		冤，異體字問題。
3868	卮	1	1	81C5	5DF5	卮		
3877	迺	1	1	81CE	5EFC	迺		迺，異體字問題。
3883	徧	4	4	81D4	5FA7	徧		
3886	悞	2	2	81D7	609E	悞		誤，異體字問題。
3891	慙	6	6	81DC	6159	慙		慚，異體字問題。
3911	攬	1	1	81F0	651F	攬		
3916	敷	1	1	81F5	656D	敷		
3924	朞	5	5	81FD	671E	朞		
3933	桀	2	2	8247	6998	桀		
3935	槩	2	2	8249	69E9	槩		概，異體字問題。
3936	棹	1	1	824A	6A70	棹		
3944	氷	4	4	8252	6C37	氷		冰，異體字問題。
3966	烟	25	25	8268	70DF	烟		煙，異體字問題。
3968	煨	1	1	826A	7145	煨		



編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3971	熇	4	4	826D	71BA	熇		
3974	牀	10	10	8270	7240	牀		
3978	猪	1	1	8274	732A	猪		猪，異體字問題。
3984	瑠	1	1	827A	7460	瑠		
3989	璿	3	3	82A1	74C8	璿		
3996	疎	10	10	82A8	758E	疎		
4008	癯	2	2	82B4	766F	癯		
4011	皞	2	2	82B7	76A1	皞		
4013	盃	1	1	82B9	76CC	盃		
4015	眦	1	1	82BB	7726	眦		
4016	着	10	10	82BC	7740	着		著，異體字問題。
4019	矻	1	1	82BF	77F4	矻		
4025	礪	1	1	82C5	7934	礪		
4031	杭	1	1	82CB	79D4	杭		
4035	穿	1	1	82CF	7A7D	穿		
4038	竈	3	3	82D2	7AC8	竈		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4044	筋	3	3	82D8	7B6F	筋		
4067	綉	1	1	82EF	7D89	綉		繡，異體字問題。
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。
4074	縑	1	1	82F6	7E27	縑		
4080	羣	49	49	82FC	7FA3	羣		群，異體字問題。
4081	翱	1	1	82FD	7FFA	翱		
4085	耻	1	1	8342	803B	耻		恥，異體字問題。
4089	脇	2	2	8346	8107	脇		
4090	脚	12	12	8347	811A	脚		腳，異體字問題。
4091	臯	7	7	8348	81EF	臯		
4097	芑	1	1	834E			艹 厶 巳	芑，異體字問題。
4098	芪	1	1	834F	82AA	芪		
4107	葱	4	4	8358	8471	葱		
4112	蕪	1	1	835D	8534	蕪		
4114	藜	1	1	835F	8616	藜		
4115	藁	3	3	8360	85C1	藁		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4118	藁	8	8	8363	8602	藁		
4124	蝨	1	1	8369	8771	蝨		
4129	虬	1	1	836E	866C	虬		
4131	劬	2	2	8370	8842	劬		
4132		74	74	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	17	17	8373	885E	衛		衛，異體字問題。
4135	衛	1	1	8374	885C	衛		
4138	表	2	2	8377	88A0	表		
4140	衽	1	1	8379	88B5	衽		
4146	覩	12	12	83A1	89A9	覩		
4158	譌	6	6	83AD	8B4C	譌		Unicode 呈現差異。
4176	輒	7	7	83BF	8F19	輒		
4184	迹	5	5	83C7	8FF9	迹		
4186	遡	3	3	83C9	9061	遡		
4190	邳	1	1	83CD	90C4	邳		
4193	鈎	7	7	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		
4213	鞞	2	2	83E4	97BE	鞞		
4218	萑	2	2	83E9	97EE	萑		
4219	頰	1	1	83EA	982E	頰		
4220	頰	4	4	83EB	9834	頰		頰，異體字問題。
4221	頰	2	2	83EC	9833	頰		
4222	頰	1	1	83ED	9847	頰		
4225	顛	1	1	83F0	9866	顛		
4237	髡	3	3	83FC	9AE0	髡		髡，異體字問題。
4240	髡	3	3	8440	9AF4	髡		
4242	鬪	7	7	8442	9B2D	鬪		Unicode 呈現差異。
4247	鮎	1	1	8447	9B8E	鮎		
4254	鷓	1	1	844E	9D7E	鷓		
4256	鷄	4	4	8450	9DC4	鷄		雞，異體字問題。
4267	鼈	4	4	845B	9F08	鼈		
4275	义	1	1	8463	4E49	义		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4293		2	2	8475			∞束△火	棘△火，舊漢籍錯字， 可用程式全部取代。
4309	刼	8	8	84A7	523C	刼		劫，異體字問題。
4349	𧈧	1	1	84CF	878A	𧈧		
4455	劒	9	9	857A			僉△刀	劍，異體字問題。
4460	勅	6	6	85A1	52D1	勅		
4485	廔	1	1	85BA	3553		厂△墨	廔，異體字問題。
4507	吳	155	155	85D0	5434	吳		吳，異體字問題。
4515	呪	3	3	85D8	546A	呪		
4550	疋	1	1	85FB	24D13		疋	
4603	憇	1	1	8671	6187	憇		
4607	狫	1	1	8675	3E60		豸△屯	
4702	場	2	2	86F6	5872	場		場，異體字問題。
4735	壠	1	1	8758	58E0	壠		
4768	妬	7	7	8779	59AC	妬		妒，異體字問題。
4790	賺	1	1	87B1	81C1	賺		
4799	顛	1	1	87BA	9856	顛		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4817	寶	6	6	87CC	5BF3	寶		寶，異體字問題。
4902	兗	1	1	8862	5156	兗		
4907	橐	2	2	8867	69D6	橐		
4910		3	3	886A			氵 厶 畧	淫，異體字問題。
4957	崑	3	3	88BB	5D53	崑		
4980	藝	1	1	88D2			藝 火	
4996	斲	1	1	88E2	65B5	斲		
5008	幪	1	1	88EE	5E5E	幪		
5013	彳	2	2	88F3			彳	
5019	畧	4	4	88F9			畧	略，異體字問題。
5031	闕	1	1	8946	95B4	闕		闕，異體字問題。
5074	尫	2	2	8971	5C2B	尫		尫，異體字問題。
5075	彳	1	1	8972			彳	
5111	旃	1	1	89B8	65BE	旃		旃，異體字問題。
5163	憂	2	2	89EC	621E	憂		憂，異體字問題。
5164	肇	1	1	89ED	8088	肇		肇，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5176		11	11	89F9			冫 厶 冗	沉，異體字問題。
5225	卹	1	1	8A6B	460F		血 厶 卩	卹，異體字問題。
5241	鶩	2	2	8A7B	9D5E	鶩		鵝，異體字問題。
5290	携	19	19	8ACE	643A	携		
5302	昏	6	6	8ADA	662C	昏		昏，異體字問題。
5327	嶧	1	1	8AF3	5D2A	嶧		
5442	枿	1	1	8BC9	67BF	枿		
5447	躡	1	1	8BCE	8E83	躡		
5482	璽	2	2	8BF1	8812	璽		
5483	眡	1	1	8BF2	7721	眡		
5484	楛	4	4	8BF3	6973	楛		
5488	霽	1	1	8BF7	4A34		雨 厶 對	
5549	藁	8	8	8C75	85F3	藁		
5551	賈	2	2	8C77	8CF7	賈		
5558	麓	1	1	8C7E	9E84	麓		
5559	欸	2	2	8CA1	6B35	欸		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5569	瓊	3	3	8CAB	74AE	瓊		
5575	峇	4	4	8CB1	8A14	峇		
5578	弛	1	1	8CB4	38AE		弓厶屯	弛，異體字問題。
5579	宐	3	3	8CB5	5D08	宐		
5582	鞣	1	1	8CB8	97B5	鞣		
5587	歛	3	3	8CBD	6B58	歛		
5590	朶	1	1	8CC0	6736	朶		朶，異體字問題。
5591	杞	1	1	8CC1	233CC		木厶巳	杞，異體字問題。
5621	泪	1	1	8CDF	6CEA	泪		淚，異體字問題。
5702	厨	1	1	8D71	53A8	厨		廁，舊漢籍錯字，可用程式全部取代。
5704	羶	2	2	8D73	6C0A	羶		
5706	敘	18	18	8D75	654D	敘		敘，異體字問題。
5721	頓	1	1	8DA6	9814	頓		
5764	籩	1	1	8DD1	7C5E	籩		
5773	滌	9	9	8DDA	6FDA	滌		
5806	膾	2	2	8DFB	4408		月厶直	



編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5860	无	1	1	C6D3	65E0	无		

附件二、《齊東野語》「●」字形手動取代表

頁碼	檔案原文摘錄	手動取代結果	備註
176	時趙京尹與●委官調停	時趙京尹與●委官調停	2次。 無法處理，維持原狀。
176	「●委」	「●委」	無法處理，維持原狀。
169	趙與●	趙與●	無法處理，維持原狀。