《莊子集解》舊版造字轉碼說明

中研院資訊所文獻處理實驗室 2007/12/18 陳建安 製作

一、 《莊子集解》一書(莊子集解.xml)使用舊版造字 124 個，字頻 1543 次，詳如附件一。這 124 個造字中，101 個可轉成 Windows XP 能顯示的字，字頻 522 次；另外 23 個字必須轉成構字式，字頻 1021 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、 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 2103 的「形辶吅亏⊡」字，原字為「逻」。

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⛰巫」係「眾」字之異體，故以「眾」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 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 797 的「扼」字，Unicode 編碼為 39D6 由於 Windows XP 並不支援，仍須使用構字式「扌⛰戹」。

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

四、 《莊子集解》中未確定字形「●」共 123 處，詳細處理內容可參考附件二的未確定字形手動取代表。附件二的手動取代表說明如下：

甲、 檔案原文摘錄：摘錄檔案(莊子集解.xml)中包含未確定字形「❷」的文句段落，其中紅字底線的部分，為需要手動取代的字形。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。

乙、 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有修改後結果。

丙、 備註：註明相同文句在檔案中出現的次數，未註明者表示只出現一次。

附件一、《莊子集解》造字分析表

編號	造字	頻次（txt）	頻次（xml）	Big5	Unicode	WinXP	構字式	備註
797	挜	1	1	8E4B	39D6		扌厄	
814	蟇	2	2	8E5C	87C7	蟇		
826	癱	1	1	8E68	7655	癱		
892	樨	1	1	8ECC	6A98	樨		
893	峩	3	3	8ECD	5CE9	峩		
932	皃	2	2	8EF4	7683	皃		
949	枏	1	1	8F46	678F	枏		
982	胷	1	1	8F67	80F7	胷		
994	楚	9	9	8F73	8E75	楚		

編號	造字	頻次（txt）	頻次（xml）	Big5	Unicode	WinXP	構字式	備註
1040	聃	46	46	8FC3	803C	聃		
1044	儛	2	2	8FC7	511B	儛		
1048	頟	1	1	8FCB	981F	頟		
1077	稟	2	2	8FE8	7A3E	稟		
1133	瓌	2	2	9061	74CC	瓌		
1226	罸	7	7	90E0	7F78	罸		
1383	穅	2	2	91E0	7A45	穅		
1646	脗	2	2	93AD	8117	脗		
2050	蹏	1	1	9648	8E4F	蹏		
2053	躧	1	1	964B	8EA7	躧		
2103		1	1	967D			形辶吅亏⊡	逻，舊漢籍錯字，可用程式全部取代。
2502	飈	2	2	98F4	98C8	飈		
2732	孚	2	2	9A7E	5BFD	孚		擘，舊漢籍錯字，可用程式全部取代。
2734	岀	1	1	9AA2	5C80	岀		恅，舊漢籍錯字，可用程式全部取代。
2739	恺	10	10	9AA7	607E	恺		蜊，舊漢籍錯字，可用程式全部取代。
2757	崨	1	1	9AB9	6B71	崨		儵，舊漢籍錯字，可用程式全部取代。
2770	絳	1	1	9AC6	7D98	絳		牀，舊漢籍錯字，可用程式全部取代。
2818	鰱	2	2	9AF6	29EF8		魚⚠逮	吡，舊漢籍錯字，可用程式全部取代。
2875	亼	7	7	9B70	4EBC	亼		蚤，舊漢籍錯字，可用程式全部取代。
3787	冢	9	9	8152	51A2	冢		
3791	况	1	1	8156	51B5	况		
3796	决	1	1	815B	51B3	决		

編號	造字	頻次（txt）	頻次（xml）	Big5	Unicode	WinXP	構字式	備註
3798	凷	1	1	815D			凷	
3816	厓	1	1	816F	5393	厓		
3817	厠	5	5	8170	53A0	厠		
3819	厮	1	1	8172	53AE	厮		
3840	坂	1	1	81A9	5742	坂		
3846	壻	1	1	81AF	58FB	壻		
3849	奬	1	1	81B2	596C	奬		
3865	瀦	3	3	81C2	24162		氵 ⿱亠⿱水⿱水	
3866	崐	2	2	81C3	5D10	崐		
3868	巵	1	1	81C5	5DF5	巵		
3869	帀	3	3	81C6	5E00	帀		
3883	徧	12	12	81D4	5FA7	徧		
3889	慤	4	4	81DA	6164	慤		
3891	慙	4	4	81DC	6159	慙		
3929	棨	1	1	8243	68C3	棨		
3935	槩	1	1	8249	69E9	槩		Unicode 呈現差異。
3974	牀	1	1	8270	7240	牀		
3976	犂	4	4	8272	7282	犂		
3978	猪	1	1	8274	732A	猪		Unicode 呈現差異。
3979	猫	4	4	8275	732B	猫		
3980	猨	8	8	8276	7328	猨		
3991	攜	1	1	82A3	64D5	攜		
4025	礴	2	2	82C5	7934	礴		
4034	穉	6	6	82CE	7A49	穉		
4036	窰	1	1	82D0	7AB0	窰		
4038	竈	7	7	82D2	7AC8	竈		

編號	造字	頻次（txt）	頻次（xml）	Big5	Unicode	WinXP	構字式	備註
4040	竝	6	6	82D4	7ADD	竝		
4041	竪	3	3	82D5	7AEA	竪		
4055	粃	2	2	82E3	7C83	粃		
4070	緤	2	2	82F2	7DE4	緤		
4080	羣	40	40	82FC	7FA3	羣		
4089	脇	2	2	8346	8107	脇		
4090	脚	5	5	8347	811A	脚		
4102	苣	5	5	8353	82E2	苣		
4104	莅	2	2	8355	8385	莅		
4132		144	144	8371			血⿱人丞	眾，異體字問題。
4134	衞	41	41	8373	885E	衞		
4140	袵	1	1	8379	88B5	袵		
4146	覩	17	17	83A1	89A9	覩		
4156	謟	1	1	83AB	8B1F	謟		
4158	譌	2	2	83AD	8B4C	譌		Unicode 呈現差異。
4160	謫	1	1	83AF	8B81	謫		
4184	迹	83	83	83C7	8FF9	迹		
4226	飡	2	2	83F1	98E1	飡		
4227	飣	2	2	83F2	98E6	飣		
4238	髦	2	2	83FD	9AE2	髦		
4239	髯	1	1	83FE	9AE5	髯		
4242	鬬	15	15	8442	9B2D	鬬		
4250	鰆	18	18	844A	9C0C	鰆		
4265	鼉	19	19	8459	9F03	鼉		
4266	鼂	1	1	845A	9F02	鼂		
4267	鼈	4	4	845B	9F08	鼈		

編號	造字	頻次（txt）	頻次（xml）	Big5	Unicode	WinXP	構字式	備註
4269	鼹	1	1	845D	9F39	鼹		
4286	从	3	3	846E	4ECE	从		
4396	蟨	1	1	84FE	87E9	蟨		
4455	劒	2	2	857A			僉⿰刀	
4458	勌	2	2	857D	52CC	勌		Unicode 呈現差異。
4489		1	1	85BE			口⿱亼亏	号，舊漢籍錯字，可用程式全部取代。
4496	叚	2	2	85C5	53DA	叚		
4507	吴	13	13	85D0	5434	吴		
4554	嗁	1	1	8640	55C1	嗁		
4578		1	1	8658			形⿱屮欠非⊡	形⿱屮欠非⊡，舊漢籍錯字，可用程式全部取代。
4588	鮅	6	6	8662	4C54		魚⿰攵攸	
4607	犹	3	3	8675	3E60		犭⿰屮屯	
4735	壠	1	1	8758	58E0	壠		
4870	韲	1	1	8842	97F2	韲		
4950	崘	2	2	88B4	5D18	崘		
4987	躅	2	2	88D9	4831		足⿰屮屬	
5091	鉶	1	1	89A4	92E3	鉶		
5108	惛	3	3	89B5	60FD	惛		
5112	园	2	2	89B9	56ED	园		
5133	憙	1	1	89CE	6199	憙		
5162	嵔	1	1	89EB	37EA		山⿰畏	
5173	漱	1	1	89F6	6F44	漱		
5200		2	2	8A52			形ノ冂米大⊡	奧，異體字問題。

編號	造字	頻次（txt）	頻次（xml）	Big5	Unicode	WinXP	構字式	備註
5244	孌	1	1	8A7E	3738		薛⿱亼女	
5261	挩	2	2	8AB1	639C	挩		
5302	昏	2	2	8ADA	662C	昏		
5306	簡	1	1	8ADE			𥳑⿱亼閒	
5310	鳥	1	1	8AE2	8204	鳥		
5335	蚩	4	4	8AFB	459F		亡⿱亼虫	
5356	蠫	1	1	8B51	221CD		蠫	
5406	尻	2	2	8BA5	51E5	尻		
5416	頗	2	2	8BAF	2947A		冉⿰丷頁	
5470	棅	2	2	8BE5	68C5	棅		
5508	鬐	1	1	8C4C	9B10	鬐		
5516		1	1	8C54			○○束	棘，舊漢籍錯字，可用程式全部取代。
5546	樾	4	4	8C72	6A36	樾		
5583	飧	1	1	8CB9	98F1	飧		
5585	寓	3	3	8CBB	3762		宀⿱亼禹	
5704	氊	1	1	8D73	6C0A	氊		
5860	无	837	837	C6D3	65E0		无	
5863		1	1	C6D6				辵，舊漢籍錯字，可用程式全部取代。

附件二、《莊子集解》未確定字形手動取代表

檔案原文摘錄	手動取代結果	備註
連叔曰：「然。●	連叔曰：「然。瞽	
●然疲役而不知其所歸	艹⿱亼尔然疲役而不知其所歸	
●，當作茶	艹⿱亼尔，當作茶	
熒，●之借字	熒，⿱炏亼目之借字	

檔案原文摘錄	手動取代結果	備註
●，惑也	烝𠆢目，惑也	
眴、恂，並●之假借	眴、恂，並𠆢之假借	
●，驚辭也	𠆢，驚辭也	
甕●大瘿說齊桓公	甕央𠆢瓦大瘿說齊桓公	
甕●，大瘿貌	甕央𠆢瓦，大瘿貌	
與當是●之借字	與當是走𠆢與之借字	
●，安行也	走𠆢與，安行也	
跰●而鑑於井	跰足𠆢鮮而鑑於井	
跰●，曳疾貌	跰足𠆢鮮，曳疾貌	
死為決●潰癰	死為決疒𠆢丸潰癰	
●，胡亂反	疒𠆢丸，胡亂反	
若●癰決潰	若疒𠆢丸癰決潰	
●萬物而不為義	形𠆢次韭⊡萬物而不為義	
●，碎也	形𠆢次韭⊡，碎也	
說文作●，亦作	說文作形𠆢次韭⊡，亦作	
執●之狗來藉	執𦍒之狗來藉	
罪讀為●，說文作●	罪讀為山𠆢罪，說文作山𠆢辠	
故以●乎形容之	故以山𠆢罪乎形容之	
正字通弟作●	正字通弟作弟	
●然而封戎	忄𠆢分然而封戎	
攫。』說文作●	攫。』說文作扌𠆢寒	
蒿是●之叚字	蒿是目𠆢雀之叚字	
●，目明	目𠆢雀，目明	
是●為望視之貌	是目𠆢雀為望視之貌	
必為●然遠望	必為目𠆢雀然遠望	
●與蒿，古音相近	目𠆢雀與蒿，古音相近	
文選景福殿賦作『●●』	文選景福殿賦作『白𠆢雀白𠆢雀』	
蒿之通●	蒿之通目𠆢雀	
●矣。」</qn>	白𠆢雀矣。」</qn>	
崔本穀作●	崔本穀作殼𠆢孚	
孺子曰●	孺子曰殼𠆢孚	
跳也，字書作●	跳也，字書作馬𠆢垂	
『喫，同●。』●	『喫，同口𠆢𣪍。』口𠆢𣪍	
執留，本又作●	執留，本又作犭𠆢留	

檔案原文摘錄	手動取代結果	備註
司馬云：『●，竹鼠也	司馬云：『犭⿰留，竹鼠也	
說文：「●，竹鼠也。」	說文：「鼠⿰卯，竹鼠也。」	
竹●。」郭璞山海經	竹鼠⿰卯。」郭璞山海經	
非竹●，特竹●之音	非竹鼠⿰留，特竹鼠⿰留之音	
香、●、腐	香、鯉、腐	
●萬物而不為戾	形弗次韭⊡萬物而不為戾	
其知憯於●蠆之尾	其知憯於厂⿱蠆蠆之尾	
●、蠆，皆蝎之異名	厂⿱蠆、蠆，皆蝎之異名	
蠆之言●	蠆之言虫⿰旦	
毒、●、瘌，痛也	毒、虫⿰旦、瘌，痛也	
蜊與●，古同聲	蜊與厂⿱蠆，古同聲	
●、蜊實一字	厂⿱蠆、蜊實一字	
猶蜊之通作●矣	猶蜊之通作厂⿱蠆矣	
●淳散朴	氵⿰梟淳散朴	
●，本亦作澆	氵⿰梟，本亦作澆	
一足●踔而行	一足跨踔而行	
●踔，跳躑也	跨踔，跳躑也	
鰌，本又作●	鰌，本又作足⿰酋	
苦弔反。說文作●	苦弔反。說文作敦⿱手	
『褚，裝衣也。』字或作●	『褚，裝衣也。』字或作衤⿰宁	
裝衣曰●	裝衣曰衤⿰宁	
得水則為●	得水則為⿰齊	
此●乃繼字	此⿰齊乃繼字	
即莊子所謂●也	即莊子所謂⿰齊也	
故以●為名	故以⿰齊為名	
「●」或誤「斷」	「⿰齊」或誤「斷」	
得水為●	得水為⿰齊	
蝤，蝤●	蝤，蝤齊⿱虫	
蝎，蝤●也	蝎，蝤齊⿱虫也	
●，●●也。」「●，●●也	齊⿱虫，齊⿱虫曹⿱蚰也。」「⿱蚰，齊⿱虫⿱蚰也	
羊奚比乎不●	羊奚比乎不⿱草	2
謂引取之也。說文作●	謂引取之也。說文作⿰扌	
以黃金注者●	以黃金注者歹⿰昏	

檔案原文摘錄	手動取代結果	備註
一作●。說文：『●，瞀也。』	一作歹⿰民。說文：『歹⿰民，瞀也。』	
吾將三月●汝	吾將三月扌⿰豢汝	
●音患	扌⿰豢音患	
雽音戶。又作●	雽音戶。又作雽	
直當為●。●	直當為鹵。鹵	
墮其天●	墮其天失⿱衣	
●，束囊也	失⿱衣，束囊也	
●無質，是能有無矣	曜無質，是能有無矣	
故以是非相●也	故以是非相⿰形弔次韭⿰⊙也	
●，子兮反	⿰形弔次韭⿰⊙，子兮反	
日中穴●	日中穴阦	
●，牆也	阦，牆也	
傅奕本作●	傅奕本作憂⿰欠	
●於志，鏌●為下	⿰形忄 牦曰⿰⊙於志，鏌鋣為下	
毒也。」字或作「●」	毒也。」字或作「⿰形忄 牦曰⿰⊙」	
有生，●也	有生，黻也	
一名栩，一名●	一名栩，一名予⿱木	
芋、●、杼三字通	芋、予⿱木、杼三字通	
委蛇攫●	委蛇攫⿰形扌 叉爪⿰⊙	
●，本又作搔	⿰形扌 叉爪⿰⊙，本又作搔	
徐本作●，七活反	徐本作⿰形扌 叉木⿰⊙，七活反	
攫●，騰擲也	攫⿰形扌 叉木⿰⊙，騰擲也	
桔梗也，雞●也	桔梗也，雞广⿰雍也	
背●可以休老	背女⿰戚可以休老	
●，本亦作摵	女⿰戚，本亦作摵	
居於●畝之中	居於亩⿰犬畝之中	
吾嘗食於十●	吾嘗食於十將⿱食	
●讀曰漿	將⿱食讀曰漿	
而五●先饋	而五將⿱食先饋	
而●其所患。<qn>釋文：「●	而⿰形弔次韭⿰⊙其所患。<qn>釋文：「⿰形弔次韭⿰⊙	
●有釀意。一說●與齎同	⿰形弔次韭⿰⊙有釀意。一說⿰形弔次韭⿰⊙與齎同	
夫●特為食羹之貨	夫將⿱食特為食羹之貨	

檔案原文摘錄	手動取代結果	備註
故以賣●之事推之	故以賣將⿱亼食之事推之	
子為●粉夫	子為⿰形屮次韭⿰⿸⿱⿰⿰⿰⿰粉夫	
而不免於●斷	而不免於魠斷	
輐、●音義同也	輐、魠音義同也	
玉篇引作『●馳』	玉篇引作『僻馳』	