

# 《海國四說》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  
2008/5/19 陳維君 製作

一、《海國四說》一書(海國四說.xml)使用舊版造字172個，字頻1374次，詳如附件一。這172個造字中，132個可轉成Windows XP能顯示的字，字頻799次；另外40個字必須轉成構字式，字頻575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4750的「廿」字，原字應為「𠂇」。
- 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893的「峩」係「峨」字之異體，故以「峨」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4145的「羈」字，Unicode字型呈現為「羈」，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2985的「如𠂇馬」。

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 55 的「禩」字，Unicode 編碼為 2771E，由於 Windows XP 並不支援，仍須使用構字式「衤𠂔曼」。

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、《海國四說》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
33	藟	2	2	FA60	85A5	藟		
55	禩	1	1	FA76	2771E		衤𠂔曼	
296	𠂔	2	2	FBEC			𠂔	
297	𠂔	227	227	FBED			𠂔	
298	𠂔	1	1	FBEE			𠂔	
336	畝	1	1	FC55			畝	
569	蚕	1	1	FDC3			蚕	
624	畷	4	4	FDFA			畷	
820	烟	1	1	8E62	70F1	烟		
893	峩	1	1	8ECD	5CE9	峩		峨，異體字問題。
949	柎	15	15	8F46	678F	柎		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
997	嘍	1	1	8F76	568A	嘍		
1077	稟	1	1	8FE8	7A3E	稟		
1184	狻	1	1	90B6	72D3	狻		
1187	盖	1	1	90B9	76D6	盖		蓋，異體字問題。
1226	罰	7	7	90E0	7F78	罰		罰，異體字問題。
1278	碍	1	1	9155	788D	碍		礙，異體字問題。
1364		1	1	91CD			冗厶圭	罍，異體字問題。
1565	絲	1	1	92F9	7DDC	絲		
1642	脉	16	16	93A9	8109	脉		
1836		1	1	94CE			口厶丹	待造字。
1841	塿	2	2	94D3	5896	塿		
1854	啣	1	1	94E0	20E95		口厶留	
1893	袞	2	2	9548	886E	袞		
1911	嚟	3	3	955A	569F	嚟		
1914		1	1	955D			口厶思	待造字。
1926	覓	2	2	9569	8994	覓		覓，異體字問題。
1951	味	4	4	95A4	20C0C		口厶末	
1959	墳	1	1	95AC	213B4		土厶貢	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
1981	讐	2	2	95C2	8B90	讐		讐，異體字問題。
2046	踪	2	2	9644	8E2A	踪		
2140	迫	4	4	96C4	5EF9	迫		迫，異體字問題。
2305	鎖	1	1	97CC	93BB	鎖		
2409	闕	1	1	9875	95A7	闕		
2590	咩	1	1	99AF	54F6	咩		
2806	勒	1	1	9AEA	976D	勒		
2826	巡	15	15	9AFE	5EF5	巡		巡，異體字問題。
2863	菱	1	1	9B64			𠂇 𠂇友	菱，異體字問題。
2893	屬	1	1	9BA4	5C5E	屬		屬，異體字問題。
2985		1	1	9C41			如 𠂇馬	待造字。
2993	办	1	1	9C49	5205	办		
3003	嘆	82	82	9C53	20E04		口 𠂇英	
3020	禘	1	1	9C64			禘	
3048	叅	1	1	9CA2	53C5	叅		
3096	歲	4	4	9CD2	5D57	歲		歲，異體字問題。
3216	喊	12	12	9DAD	55B4	喊		
3791	况	1	1	8156	51B5	况		况，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
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。
3796	決	1	1	815B	51B3	決		決，異體字問題。
3804	効	6	6	8163	52B9	効		
3805	勅	17	17	8164	52C5	勅		
3807	驗	5	5	8166	9A10	驗		
3814	却	7	7	816D	5374	却		卻，異體字問題。
3819	廡	1	1	8172	53AE	廡		廡，異體字問題。
3820	廈	7	7	8173	53A6	廈		廈，異體字問題。
3825	疊	2	2	8178	53E0	疊		疊，異體字問題。
3826	叶	1	1	8179	53F6	叶		
3829	咏	3	3	817C	548F	咏		
3833	嘍	7	7	81A2	5611	嘍		
3835	噍	1	1	81A4	564D	噍		
3849	獎	12	12	81B2	596C	獎		獎，異體字問題。
3864	峯	2	2	81C1	5CEF	峯		峰，異體字問題。
3876	廻	1	1	81CD	5EFB	廻		迴，異體字問題。
3883	徧	2	2	81D4	5FA7	徧		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3886	悞	9	9	81D7	609E	悞		
3891	慙	1	1	81DC	6159	慙		慙，異體字問題。
3914	敏	1	1	81F3	6542	敏		
3923		2	2	81FC			晶𠄎宜	疊，異體字問題。
3933	槩	2	2	8247	6998	槩		矩，異體字問題。
3966	烟	31	31	8268	70DF	烟		煙，異體字問題。
3974	牀	9	9	8270	7240	牀		
3977	狗	1	1	8273	72E5	狗		
3978	猪	2	2	8274	732A	猪		豬，異體字問題。
3979	猫	2	2	8275	732B	猫		貓，異體字問題。
3985	璫	1	1	827B	7447	璫		Unicode 呈現差異。
3996	疎	9	9	82A8	758E	疎		
4013	盃	1	1	82B9	76CC	盃		
4014	蓋	35	35	82BA	8462	蓋		蓋，異體字問題。
4016	着	12	12	82BC	7740	着		著，異體字問題。
4024	礮	65	65	82C4	792E	礮		
4028	稟	49	49	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
4040	竝	1	1	82D4	7ADD	竝		並，異體字問題。
4041	豎	3	3	82D5	7AEA	豎		豎，異體字問題。
4080	羣	39	39	82FC	7FA3	羣		群，異體字問題。
4085	耻	2	2	8342	803B	耻		
4089	脇	1	1	8346	8107	脇		脅，異體字問題。
4090	脚	7	7	8347	811A	脚		腳，異體字問題。
4094	館	2	2	834B	8218	館		
4095	舩	1	1	834C	8229	舩		
4105	菓	4	4	8356	83D3	菓		
4107	葱	4	4	8358	8471	葱		蔥，異體字問題。
4114	藪	1	1	835F	8616	藪		
4130	衅	2	2	836F	8845	衅		釁，異體字問題。
4132		117	117	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	22	22	8373	885E	衛		衛，異體字問題。
4136	衰	1	1	8375	887A	衰		
4137	袜	2	2	8376	889C	袜		襪，異體字問題。
4140	衽	4	4	8379	88B5	衽		衽，異體字問題。
4145	羈	3	3	837E	898A	羈		Unicode 呈現差異。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4146	覩	6	6	83A1	89A9	覩		
4166	賄	2	2	83B5	8CC9	賄		
4167	賈	19	19	83B6	8CEB	賈		
4173	駝	4	4	83BC	8EAD	駝		
4176	輒	32	32	83BF	8F19	輒		
4184	迹	7	7	83C7	8FF9	迹		
4188	邨	2	2	83CB	90A8	邨		
4193	鈎	6	6	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	1	1	83D2	9262	鉢		
4198	鑛	2	2	83D5	945B	鑛		
4203	隣	30	30	83DA	96A3	隣		鄰，異體字問題。
4220	穎	1	1	83EB	9834	穎		穎，異體字問題。
4242	鬪	7	7	8442	9B2D	鬪		Unicode 呈現差異。
4253	鳶	1	1	844D	9CF8	鳶		Unicode 呈現差異。
4256	鷄	5	5	8450	9DC4	鷄		
4262	麪	1	1	8456	9EAA	麪		
4278	亘	1	1	8466	4E98	亘		互，異體字問題。
4283	戲	2	2	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
4286	从	3	3	846E	4ECE	从		從，異體字問題。
4309	刼	10	10	84A7	523C	刼		劫，異體字問題。
4310		1	1	84A8			革厶健	待造字。
4318		1	1	84B0			𠃉畝且毛𠃉	𦉳，異體字問題。
4374	漼	1	1	84E8	3D36		氵厶雀	
4432		1	1	8563			𠃉纒厂墨𠃉	纏，異體字問題。
4437	廩	1	1	8568	5EEA	廩		廩，異體字問題。
4447	捏	3	3	8572	63D1	捏		捏，異體字問題。
4455	劔	1	1	857A			僉厶刀	
4482	廛	1	1	85B7	53AA	廛		廛，異體字問題。
4496	段	2	2	85C5	53DA	段		
4507	吴	8	8	85D0	5434	吴		吳，異體字問題。
4547		6	6	85F8			口厶美	待造字。
4570	隣	2	2	8650	5652	隣		
4591	嚨	36	36	8665	56D2	嚨		
4643	徃	1	1	86BB	5F83	徃		往，異體字問題。
4661	埦	4	4	86CD	57DD	埦		
4670		1	1	86D6			口厶氏	待造字。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4702	場	1	1	86F6	5872	場		
4750	卅	1	1	8767			卅	卅，校對問題，舊版漢籍錯字，可用程式全部取代。
4768	妬	1	1	8779	59AC	妬		妒，異體字問題。
4805	瀆	4	4	87C0	51DF	瀆		瀆，異體字問題。
4817	寶	1	1	87CC	5BF3	寶		寶，異體字問題。
4844		2	2	87E7			𦉰𦉱凡	𦉰，異體字問題。
4870	壘	1	1	8842	97F2	壘		
4885	糝	3	3	8851	7CDA	糝		
4904	尅	1	1	8864	5C05	尅		尅，異體字問題。
5005	幫	1	1	88EB	5E2E	幫		幫，異體字問題。
5019	畧	64	64	88F9			田𦉰各	畧，異體字問題。
5026		1	1	8941			𦉰广回示𦉰	廩，異體字問題。
5046	彝	1	1	8955	5F5C	彝		彝，異體字問題。
5075	彳	6	6	8972			彳	
5104	礼	1	1	89B1	793C	礼		禮，異體字問題。
5176		20	20	89F9			氵𦉰冗	沉，異體字問題。
5194		3	3	8A4C			𦉰扌𦉰且𦉰	擅，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5195		5	5	8A4D			𣎵木亩且	檀，異體字問題。
5225	卹	3	3	8A6B	460F		血𠂇卩	卹，異體字問題。
5241	鶩	1	1	8A7B	9D5E	鶩		鵝，異體字問題。
5290	携	4	4	8ACE	643A	携		
5306	簡	3	3	8ADE			竹𠂇間	簡，異體字問題。
5376	麋	1	1	8B65	9E96	麋		
5439	栢	1	1	8BC6	6822	栢		
5493	迆	1	1	8BFC			迆	
5515	軌	1	1	8C53			車𠂇丸	軌，異體字問題。
5542	鑠	1	1	8C6E	93C1	鑠		
5555	迥	1	1	8C7B	9008	迥		迥，異體字問題。
5559	欸	22	22	8CA1	6B35	欸		欸，異體字問題。
5566	𠂇	2	2	8CA8	27D29		貝𠂇八	
5697	𦉳	1	1	8D6C	7151	𦉳		煮，異體字問題。
5704	𣎵	26	26	8D73	6C0A	𣎵		
5706	𣎵	5	5	8D75	654D	𣎵		𣎵，異體字問題。
5768	濶	10	10	8DD5	6FF6	濶		濶，異體字問題。
5799	𣎵	1	1	8DF4			𣎵𠂇炎𠂇	𣎵，異體字問題。