

# 《海角續編》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  
2008/5/28 葉冠孜 製作

一、《海角續編》一書(海角續編.xml)使用舊版造字 29 個，字頻 74 次，詳如附件一。這 29 個造字中，26 個可轉成 Windows XP 能顯示的字，字頻 65 次；另外 3 個字必須轉成構字式，字頻 9 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、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 4134 的「衛」係「衛」字之異體，故以「衛」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。
- 2、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。

附件一、《海角續編》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
297	厶	1	1	FBED			厶	
573	乚	1	1	FDC7			乚	
746	亅	1	1	FED7			亅	
1544	絇	1	1	92E4	7D59	絇		
1784	萼	2	2	9478	855A	萼		
2749	枱	1	1	9AB1	67B1	枱		檯，異體字問題。
3799	刂	2	2	815E	5226	刂		劫，異體字問題。
3814	却	5	5	816D	5374	却		卻，異體字問題。
3825	疊	1	1	8178	53E0	疊		疊，異體字問題。
3842	堃	1	1	81AB	5803	堃		
3864	峯	1	1	81C1	5CEF	峯		峰，異體字問題。
3938	橈	1	1	824C	6AC8	橈		凳，異體字問題。
3950	浜	4	4	8258	6D5C	浜		濱，異體字問題。
3966	烟	3	3	8268	70DF	烟		煙，異體字問題。
4016	着	4	4	82BC	7740	着		著，異體字問題。
4034	穉	11	11	82CE	7A49	穉		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4080	羣	4	4	82FC	7FA3	羣		群，異體字問題。
4090	脚	8	8	8347	811A	脚		腳，異體字問題。
4094	館	1	1	834B	8218	館		館，異體字問題。
4112	蔴	1	1	835D	8534	蔴		
4132		7	7	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	2	2	8373	885E	衛		衛，異體字問題。
4193	鈎	1	1	83D0	920E	鈎		鈎，異體字問題。
4203	隣	1	1	83DA	96A3	隣		鄰，異體字問題。
4242	鬪	2	2	8442	9B2D	鬪		Unicode 呈現差異。
4256	鷄	2	2	8450	9DC4	鷄		雞，異體字問題。
5311	眈	2	2	8AE3	6644	眈		
5556	隸	2	2	8C7C	96B7	隸		隸，異體字問題。
5706	敘	1	1	8D75	654D	敘		敘，異體字問題。