

# 《阿惟越致遮經》舊版造字轉碼說明

中研院資訊所文獻處理實驗室  
中央研究院語言所文獻語料小組  
2009/04/27 丁玟伶

轉碼主要工作是把檔案中的舊版造字轉換成 Windows XP 能支援的 Unicode 字形，Unicode 目前共收錄漢字 70194 個字，而 XP 只能支援 20902 個字（詳如表一），不支援之字將以構字式表達。例：造字編號 2042 的「䟽」字，Unicode 編碼是 47FD，由於 XP 並不支援，仍需使用構字式「𠃉𠃉𠃉」。

表一、Unicode 的字數及編碼區段

Unicode	新增字數	新增編碼區段	總字數	WinXP
1.1 版	20902	4E00-9FFF	20902	支援
3.0 版	6582	3400-4DFF	27484	不支援
3.1 版	42710	20000-2A6D6	70194	不支援

## 一、舊版造字轉碼分析：

《阿惟越致遮經》使用舊版造字 16 個，字頻 141 次，這 16 個造字中，14 個可轉成 Windows XP 能顯示的字，字頻 138 次；另外 2 個字必須轉成構字式，字頻 3 次。

轉碼完成製作轉碼分析表，請參考附件一《阿惟越致遮經》轉碼分析表，欄位說明如下：

- (一) 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
- (二) 造字：舊版造字。
- (三) 頻次：舊版造字在文件的出現次數。
- (四) Big5：造字的 Big5 碼。
- (五) Unicode：造字所對應的 Unicode 碼。

(六) WinXP：造字在 Windows XP 的對應字形。

(七) 構字式：Windows XP 無法對應字形改用構字式。

(八) 備註凡例：備註欄中記錄轉碼後字形及修改原因，凡例如下：

1. 異體字問題：為了使用者查詢和使用的方便，在處理異體字時最主要的方式是以標準字取代，除非是專有名詞或特殊情形，如：人名、地名等。例：造字編號 4768 的「妬」字，是「妒」的異體字，以標準字「妒」取代。
2. Unicode 字型呈現差異：Unicode 字型與舊版造字有些微差異，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 4773 的「姉」字，Unicode 字型呈現為「姉」，實際上仍為同一字。

附件一、《阿惟越致遮經》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2042	䟽	1	9640	47FD		足𠄎充	
3039	覩	68	9C77	89A9	覩		睹，異體字問題。
3804	効	1	8163	52B9	効		
3805	勅	1	8164	52C5	勅		
3824	叙	1	8177	53D9	叙		
3975	牕	1	8271	7255	牕		Unicode 字型呈現差異。
3992	嘗	32	82A4	751E	嘗		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4041	豎	3	82D5	7AEA	豎		
4184	迹	22	83C7	8FF9	迹		
4240	鬚	1	8440	9AF4	鬚		
4696	燿	2	86F0	3E0C		火 ㄩ 霍	
4738	睺	2	875B	777A	睺		
4768	妬	1	8779	59AC	妬		妒，異體字問題。
4773	姉	1	877E	59C9	姉		Unicode 字型呈現差異。
5300	鬪	1	8AD8	9B2A	鬪		
5316	晉	3	8AE8	664B	晉		晉，異體字問題。

## 二、手動取代表說明：

《阿惟越致遮經》檔案中出現「●」字形，共計12字，比對原書後以Unicode字形或構字式手動取代，並製作「●」字手動取代表，請參考附件二。欄位說明如下：

1. 頁碼：「●」字形所在位置之原書頁碼。
2. 檔案原文摘錄：摘錄《阿惟越致遮經》檔案有「●」字形的文句。
3. 原書字形：「●」在原書中之字形，以此字取代檔案原文的「●」。
4. 備註：記錄修改相同詞句的次數或其他事項。

附件二、《阿惟越致遮經》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 199	未央可●爾	竟	
p. 200	光明從軒●	窓	
p. 200	●一切菩薩	姦	
p. 200	天●軒牖刻	窓	
p. 201	鬼佛犍沓●	愁	
p. 201	神犍沓●	愁	
p. 202	億百千●	姦	有 4 處
p. 214	六十億●	姦	
p. 225	身當長一萬●	姦	