

《契丹國志》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
2008/6/26 葉冠孜 製作

一、《契丹國志》一書(契丹國志.xml)使用舊版造字 26 個，字頻 65 次，詳如附件一。這 26 個造字中，22 個可轉成 Windows XP 能顯示的字，字頻 60 次；另外 4 個字必須轉成構字式，字頻 5 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，新舊版缺字並存：漢籍在舊轉新的過程中，以人工修改方式取代舊漢籍造字，雖然修改的結果是正確的，但是只修改到部分造字，沒有全部修改到，導致舊漢籍造字在新版漢籍檔案中，一部分已被取代成 Unicode 或構字式，另一部分仍為舊漢籍造字的並存情形。處理方式為配合漢籍工作室的修改，將沒被修改到的舊漢籍造字，以對應的 Unicode 或構字式取代。如編號 3814，舊漢籍造字為「却」，比對 txt 檔及 xml 檔後，發現原先在 txt 檔使用舊漢籍造字「却」者，在 xml 檔中部分以人工修改為「卻」，故配合修改，將編號 3814 的「却」字，手動取代為「卻」。
- 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 3791 的「況」係「況」字之異體，故以「況」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 2343 的「𪗇」字，Unicode 編碼為 4633，由於 Windows XP 並不支援，仍須使用構字式「革𪗇占」。

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、 txt 檔與 xml 檔頻次不同且造字分析表「備註」欄有備註問題情況者，詳細取代內容可參考附件二的手動取代表。附件二的手動取代表說明如下：

- 甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
- 乙、 造字：舊漢籍造字。
- 丙、 取代：實際使用字形與舊漢籍造字不同時，於此欄列出。
- 丁、 檔案原文摘錄：摘錄檔案(契丹國志.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。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。
- 戊、 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。

- 五、 《契丹國志》一書(契丹國志.xml)中未確定字形「?」共出現 6 次，詳細取代內容可參考附件三未確定字形「?」手動取代表。附件三未確定字形「?」手動取代表的欄位說明如下：

- 甲、 檔案原文摘錄：摘錄檔案(契丹國志.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。
- 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。

附件一、《契丹國志》造字分析表

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-----|------------------------|
| 297 | 𠄎 | 16 | 1 | FBED | | | 𠄎 | 校對問題，新舊版缺字並存。 |
| 2343 | 鞞 | 4 | 1 | 97F2 | 4A5E | | 革𠄎占 | 校對問題，新舊版缺字並存。 |
| 3791 | 况 | 2 | 1 | 8156 | 51B5 | 况 | | 况，校對問題，新舊版缺字並存，以標準字取代。 |
| 3796 | 决 | 1 | 1 | 815B | 51B3 | 决 | | 决，異體字問題。 |
| 3805 | 勅 | 10 | 3 | 8164 | 52C5 | 勅 | | 校對問題，新舊版缺字並存。 |
| 3814 | 却 | 9 | 7 | 816D | 5374 | 却 | | 卻，校對問題，新舊版缺字並存，以標準字取代。 |
| 3824 | 叙 | 2 | 2 | 8177 | 53D9 | 叙 | | 敘，異體字問題。 |
| 3860 | 𠄎 | 1 | 1 | 81BD | 5C20 | 𠄎 | | |
| 3877 | 迺 | 1 | 1 | 81CE | 5EFC | 迺 | | 迺，異體字問題。 |
| 3891 | 慙 | 4 | 1 | 81DC | 6159 | 慙 | | 慙，校對問題，新舊版缺字並存，以標準字取代。 |
| 3937 | 檝 | 1 | 1 | 824B | 6A9D | 檝 | | |
| 3966 | 烟 | 2 | 1 | 8268 | 70DF | 烟 | | 煙，校對問題，新舊版缺字並存，以標準字取代。 |
| 4016 | 着 | 2 | 1 | 82BC | 7740 | 着 | | 著，校對問題，新舊版缺字並存，以標準字取代。 |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-----|--------------------------------|
| 4080 | 羣 | 33 | 3 | 82FC | 7FA3 | 羣 | | 群，校對問題，新舊版缺字並存，以標準字取代。 |
| 4132 | | 95 | 2 | 8371 | | | 血𠂇𠂇 | 眾，校對問題，新舊版缺字並存，以標準字取代。 |
| 4134 | 衛 | 47 | 1 | 8373 | 885E | 衛 | | 衛，校對問題，新舊版缺字並存，以標準字取代。 |
| 4146 | 覩 | 45 | 2 | 83A1 | 89A9 | 覩 | | 校對問題，新舊版缺字並存。 |
| 4158 | 譌 | 2 | 2 | 83AD | 8B4C | 譌 | | Unicode 呈現差異。 |
| 4184 | 迹 | 9 | 8 | 83C7 | 8FF9 | 迹 | | 跡，校對問題，新舊版缺字並存，以標準字取代。 |
| 4242 | 鬪 | 9 | 3 | 8442 | 9B2D | 鬪 | | Unicode 呈現差異。 校對問題，新舊版缺字並存。 |
| 4460 | 勅 | 1 | 1 | 85A1 | 52D1 | 勅 | | |
| 4878 | 宀 | 2 | 2 | 884A | 5B82 | 宀 | | 宀，異體字問題。 |
| 5219 | 妃 | 25 | 1 | 8A65 | | | 女𠂇巳 | 妃，校對問題，新舊版缺字並存，以標準字取代。 |
| 5437 | 并 | 51 | 10 | 8BC4 | 5E77 | 并 | | 并，校對問題，新舊版缺字並存，以標準字取代。 |
| 5704 | 穉 | 4 | 3 | 8D73 | 6C0A | 穉 | | 穉，校對問題，新舊版缺字並存。 |
| 5706 | 敘 | 11 | 5 | 8D75 | 654D | 敘 | | 敘，校對問題，新舊版缺字並存，以標準字取代。 |

附件二、《契丹國志》手動取代表

| 編號 | 造字 | 取代 | 檔案原文摘錄 | 手動取代結果 |
|------|----|-----|------------------------------|------------------------------|
| 297 | 厶 | | 鄧詩熙君及王女__秀婷君助成 | 鄧詩熙君及王女厶秀婷君助成 |
| 2343 | 鞞 | 革厶占 | 衣服如義襪紫__鞞 | 衣服如義襪紫革厶占鞞 |
| 3791 | 况 | 况 | 何 <u>奎</u> 成書及刊行的時間 | 何 <u>况</u> 成書及刊行的時間 |
| 3805 | 勅 | | 宋祕書丞葉隆禮奉 <u>冒</u> 撰 | 宋祕書丞葉隆禮奉 <u>勅</u> 撰 |
| | | | 若此書誠為承應孝宗 <u>冒</u> 命所撰 | 若此書誠為承應孝宗 <u>勅</u> 命所撰 |
| | | | 臣奉 <u>冒</u> 命 | 臣奉 <u>勅</u> 命 |
| 3814 | 却 | 卻 | 上引提要結尾處 <u>屨</u> 云 | 上引提要結尾處 <u>卻</u> 云 |
| | | | <u>屨</u> 又在同書同卷中說 | <u>卻</u> 又在同書同卷中說 |
| | | | 然而余嘉錫先生 <u>屨</u> 根本懷疑此書為隆禮所撰 | 然而余嘉錫先生 <u>卻</u> 根本懷疑此書為隆禮所撰 |
| | | | <u>屨</u> 不見于遼史 | <u>卻</u> 不見于遼史 |
| | | | <u>屨</u> 極為具體、生動 | <u>卻</u> 極為具體、生動 |
| | | | 但韓延徽 <u>屨</u> 尚未至契丹 | 但韓延徽 <u>卻</u> 尚未至契丹 |
| | | | 但 <u>屨</u> 非如蘇天爵所說「絕不可信」 | 但 <u>卻</u> 并非如蘇天爵所說「絕不可信」 |
| 3891 | 慙 | 慙 | 淺短自 <u>刈</u> | 淺短自 <u>慙</u> |
| 3966 | 烟 | 煙 | 採取隆禮 <u>汤</u> 兩樓和朱南傑韻一首 | 採取隆禮 <u>煙</u> 兩樓和朱南傑韻一首 |
| 4016 | 着 | 著 | 上 <u>进</u> 手撰次此書 | 上 <u>著</u> 手撰次此書 |

| 編號 | 造字 | 取代 | 檔案原文摘錄 | 手動取代結果 |
|------|-----|----|--------------------------------|--------------------------------|
| 4080 | 羣 | 群 | 即遍檢館閣 <u>倉</u> 官 | 即遍檢館閣 <u>群</u> 官 |
| | | | <u>倉</u> 牧之政 | <u>群</u> 牧之政 |
| | | | 又謂 <u>倉</u> 下曰 | 又謂 <u>群</u> 下曰 |
| 4132 | 血𠂇𠂇 | 眾 | 阿保機為 <u>𠂇</u> 所立 | 阿保機為 <u>眾</u> 所立 |
| | | | <u>𠂇</u> 皆以為此必有所據 | <u>眾</u> 皆以為此必有所據 |
| 4134 | 衛 | 衛 | 亦儀 <u>𠂇</u> 志所遺 | 亦儀 <u>衛</u> 志所遺 |
| 4146 | 覩 | | 耶律余 <u>𠂇</u> 諸傳 | 耶律余 <u>覩</u> 諸傳 |
| | | | 耶律余 <u>𠂇</u> 及大實等六人 | 耶律余 <u>覩</u> 及大實等六人 |
| 4184 | 迹 | 跡 | <u>攢</u> (全部取代) | <u>跡</u> (全部取代) |
| 4242 | 鬪 | | 又如興宗與其生母聖宗 欽哀皇后的 <u>𠂇</u> 爭 | 又如興宗與其生母聖宗 欽哀皇后的 <u>鬪</u> 爭 |
| | | | 透露出這場權力 <u>𠂇</u> 爭的 一些端倪 | 透露出這場權力 <u>鬪</u> 爭的一 些端倪 |
| | | | 少述 <u>𠂇</u> 爭之激烈 | 少述 <u>鬪</u> 爭之激烈 |
| 5219 | 女𠂇巳 | 妃 | 卷十三為后 <u>𠂇</u> | 卷十三為后 <u>妃</u> |
| 5437 | 并 | 并 | <u>泉</u> (全部取代) | <u>并</u> (全部取代) |
| 5704 | 𠂇 | | <u>廡</u> 廡遂安於徙宅 | <u>𠂇</u> 廡遂安於徙宅 |
| | | | 在穹廡中覆之以 <u>廡</u> | 在穹廡中覆之以 <u>𠂇</u> |
| | | | <u>廡</u> 中枯骨 | <u>𠂇</u> 中枯骨 |

| 編號 | 造字 | 取代 | 檔案原文摘錄 | 手動取代結果 |
|------|----|----|-----------------|-----------------|
| 5706 | 敍 | 敍 | 提要所 <u>𠄎</u> | 提要所 <u>敍</u> |
| | | | 然其 <u>𠄎</u> 次筆法 | 然其 <u>敍</u> 次筆法 |
| | | | <u>𠄎</u> 述之際 | <u>敍</u> 述之際 |
| | | | <u>𠄎</u> 述鬪爭之激烈 | <u>敍</u> 述鬪爭之激烈 |
| | | | 馬擴茆齋自 <u>𠄎</u> | 馬擴茆齋自 <u>敍</u> |

附件三、《契丹國志》未確定字形「?」手動取代表

| 檔案原文摘錄 | 手動取代結果 |
|-----------------|-----------------|
| 由此士無 <u>?</u> 志 | 由此士無 <u>鬪</u> 志 |
| 其中番漢雜 <u>?</u> | 其中番漢雜 <u>鬪</u> |
| 人皆死 <u>?</u> | 人皆死 <u>鬪</u> |
| 皆無 <u>?</u> 志 | 皆無 <u>鬪</u> 志 |
| 由是各無 <u>?</u> 志 | 由是各無 <u>鬪</u> 志 |
| 戰 <u>?</u> 而後得 | 戰 <u>鬪</u> 而後得 |