

《國語》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
2008/3/28 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 964 的「𪗇」字，Unicode 編碼是 28EF3，由於 XP 並不支援，仍需使用構字式「卩𪗇焉」。

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：

《國語》使用舊版造字 63 個，字頻 307 次，這 63 個造字中，58 個可轉成 Windows XP 能顯示的字，字頻 299 次；另外 5 個字必須轉成構字式，字頻 8 次。

轉碼完成製作轉碼分析表，請參考附件一《國語》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 4134 的「衛」字，是「衛」的異體字，以標準字「衛」取代。
 2. 部份使用不同字形，手動修改：檔案中為同一舊版造字，校對時卻發現原書使用兩個以上不同的字形，可能部份舊版造字為錯字或原書使用兩種異體字形，導致同一舊版造字轉碼時對應到兩個以上的字形；將一部份手動修改，剩餘的頻次則用程式取代成對應的 Unicode 或構字式。例：造字編號 5591 的「杞」字頻次為 3，依書將國名的 2 字手動修改為「杞」，剩餘 1 字轉為「木△巳」。手動修改部份請參考附件二《國語》手動取代表。
 3. 錯字，以程式全部取代：與原書字形不符，並查詢教育部異體字字典確認非異體字後，皆歸類為錯字，以程式取代為正確字形。例：檔案中使用造字編號 5505 的「耆」字，而原書使用之字為「耆」，查詢教育部異體字字典確認「耆」並非「耆」之異體字，因此歸為錯字，以程式取代為正確字形「耆」。

附件一、《國語》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
827	鮫	8	8E69	9B8C	鮫		
938	燮	2	8EFA	7215	燮		
964	隲	1	8F55	28EF3		卩△焉	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
980	犖	7	8F65	72AB	犖		
981	豺	3	8F66	72B2	豺		豺，異體字問題。
988	犖	1	8F6D	7F8B	犖		
1040	聃	1	8FC3	803C	聃		
1044	儻	1	8FC7	511B	儻		
1158	咳	4	907A	7561	咳		
1565	絲	1	92F9	7DDC	絲		
1642	脉	3	93A9	8109	脉		脈，異體字問題。
1857	蚩	1	94E3	86A0	蚩		
2043	踳	1	9641	8E01	踳		
2146	蝸	1	96CA	8739	蝸		
2524	饒	1	994B	994D	饒		
2765	罇	1	9AC1	7AF1	罇		
2828	蝸	1	9B41	8744	蝸		
3770	个	2	8141	4E2A	个		
3787	冢	2	8152	51A2	冢		
3806	勳	1	8165	52E0	勳		
3849	獎	1	81B2	596C	獎		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3883	徧	12	81D4	5FA7	徧		
3889	慤	1	81DA	6164	慤		
3891	慤	1	81DC	6159	慤		
3911	攡	1	81F0	651F	攡		
3921	晷	3	81FA	6673	晷		
3924	碁	1	81FD	671E	碁		期，異體字問題。
3949	洩	1	8257	6D96	洩		
3974	牀	1	8270	7240	牀		
3976	犁	2	8272	7282	犁		犁，異體字問題。
3987	穀	1	827D	7474	穀		
4035	穽	1	82CF	7A7D	穽		
4038	竈	3	82D2	7AC8	竈		
4043	笄	1	82D7	7B53	笄		笄，異體字問題。
4045	箒	1	82D9	7B92	箒		
4079	待	1	82FB	38E5		彳厶侍	
4080	羣	30	82FC	7FA3	羣		
4104	莅	3	8355	8385	莅		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4114	藥	1	835F	8616	藥		
4132		87	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	58	8373	885E	衛		衛，異體字問題。
4142	褒	11	837B	8943	褒		褒，異體字問題。
4160	適	4	83AF	8B81	適		
4184	迹	1	83C7	8FF9	迹		
4203	隣	1	83DA	96A3	隣		鄰，異體字問題。
4217	鞞	1	83E8	97E0	鞞		
4252	鰓	2	844C			魚𠄎眾	鰓，異體字問題。
4256	鷄	3	8450	9DC4	鷄		雞，異體字問題。
4265	鼃	2	8459	9F03	鼃		
4267	鼈	7	845B	9F08	鼈		
4469	勺	2	85AA	5304	勺		
4913	塿	1	886D	58A2	塿		
5089		1	89A2			冗𠄎畢	畢，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5249	蝻	1	8AA5	873D	蝻		
5335	蚤	1	8AFB	459F		亡𠄎虫	
5437	并	1	8BC4	5E77	并		
5456		2	8BD7			韋𠄎未	韋，異體字問題。
5505	耆	2	8C49	264BE		老𠄎占	考 𠄎 口，錯字，以程式全部取代。
5564	槩	1	8CA6	6ABF	槩		
5583	殮	2	8CB9	98F1	殮		殮，異體字問題。
5591	杞	3	8CC1	233CC		木𠄎巳	部份使用不同字形「杞」，手動修改。
5703	𪗇	1	8D72	9F53	𪗇		
5706	敘	3	8D75	654D	敘		敘，異體字問題。

二、手動取代表說明：

《國語》檔案中需手動修改的情況分為以下幾種：

- (一) 《國語》檔案中之舊版造字，因字形部份不同而無法全部以程式取代，皆手動修改成Windows XP能顯示的字或構字式，並製作手動取代表，請參考附件二。欄位說明如下：

1. 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
2. 造字：舊版造字字形。
3. 取代：實際使用字形與舊版造字不同時，於此欄列出。
4. 檔案原文摘錄：摘錄檔案《國語》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。
5. 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。
6. 備註：記錄修改相同辭句的次數或其他事項。

附件二、《國語》手動取代表

編號	造字	取代	檔案原文摘錄	手動取代結果	備註
5591	杞	杞	<u>庠</u> 、繒由大姒	<u>杞</u> 、繒由大姒	
			有夏雖衰， <u>庠</u> 、鄩猶在	有夏雖衰， <u>杞</u> 、鄩猶在	

(二)《國語》檔案中出現「●」字形，共計 24 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，請參考附件三。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《國語》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。

附件三、《國語》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 74	不亦●姓矣乎	嬪	
p. 109	佐●者嘗焉	食 ㄩ 雍	

頁碼	檔案原文摘錄	原書字形	備註
p. 178	水虞於是禁罍罍●	罍△鹿	
p. 178	麇●，鳥翼毳卵	鹿△天	
p. 205	懼●季孫之怒	忤	
p. 209	康子往焉，●門與之言	門△為	
p. 223	戎車待遊車之●	前△衣	
p. 228	及寒，擊●除田	菓	
p. 228	首戴茅蒲，身衣襪●	襪	
p. 239	輕罪贖以●盾一	鞮	
p. 241	地南至於●陰	食△甸	
p. 472	今忤日而●歲	汜△歇	
p. 494	鄭人擊我，吾伏弑●血	略	
p. 507	鄔、弊、補、舟、依、●、 歷	黑△柔	
p. 514	其在芊姓乎？芊姓●越	首△冫△歧	
p. 516	收經入，行●極	姦	
p. 538	若合而●吾中	召	
p. 549	亡△虫●之既多	維△虫	
p. 567	日月會于龍●	豕△尫	
p. 594	執箕箒以●姓於王宮	咳	
p. 597	於其心也●然	佻	
p. 602	盛以鷗●，而投之於江	鷗	
p. 605	將夾溝而●我	广△侈	
p. 612	自●於客前以酬客	亞△丩	

(三)《國語》檔案中出現「☆」字形，共計 15 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，

請參考附件四。欄位說明如下：

1. 檔案原文摘錄：摘錄《國語》檔案有「☆」字形的文句。
2. 原書字形：「☆」在原書中之字形，以此字取代檔案原文的「☆」。
3. 備註：記錄修改相同詞句的次數或其他事項。

附件四、《國語》「☆」字手動取代表

檔案原文摘錄	原書字形	備註
穀、洛☆	鬪	
佐☆者傷焉	鬪	
王將防☆川以飾宮	鬪	
是飾亂而佐☆也	鬪	
遂☆而死	鬪	
師少於我，☆士眾	鬪	
☆士是故眾	鬪	
偃也聞之：戰☆	鬪	
濟可☆。若有違質	鬪	
楚將顧夷，莫有☆心	鬪	
魯孟獻子有☆臣五人	鬪	
子常問蓄貨聚馬☆且論其必亡	鬪	
☆且廷見令尹子常	鬪	
昔☆子文三舍令尹	鬪	
自☆伯比以來	鬪	

(四)《國語》檔案中，對於舊版造字以外發現的錯字，以手動修改，

並製作錯字修改表，請參考附件五。欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《國語》中的文句段落，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。

2. 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，不一一列舉所有原文。
3. 備註：記錄修改相同詞句的次數或其他事項。

附件五、《國語》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
子常問蓄貨聚馬 <u>鬥</u> 且論其必亡（全部取代）	子常問蓄貨聚馬 <u>鬪</u> 且論其必亡（全部取代）	有 2 處
子常問蓄貨 <u>蜜</u> 馬	子常問蓄貨 <u>聚</u> 馬	
問蓄貨 <u>蜜</u> 馬	問蓄貨 <u>聚</u> 馬	
令尹問蓄 <u>蜜</u> 積實	令尹問蓄 <u>聚</u> 積實	
夫古者 <u>蜜</u> 貨不妨民衣食之利	夫古者 <u>聚</u> 貨不妨民衣食之利	
王不聽， <u>鉞</u> 征之	王不聽， <u>遂</u> 征之	
<u>鉞</u> 為踐土之盟	<u>遂</u> 為踐土之盟	
狄人 <u>鉞</u> 入	狄人 <u>遂</u> 入	
武子 <u>鉞</u> 不敢對而退	武子 <u>遂</u> 不敢對而退	
<u>鉞</u> 假道於陳	<u>遂</u> 假道於陳	
王 <u>鉞</u> 不賜	王 <u>遂</u> 不賜	
稱 <u>鉞</u> 共工之過	稱 <u>遂</u> 共工之過	
王室 <u>鉞</u> 卑	王室 <u>遂</u> 卑	
節之鼓而行之，以 <u>鉞</u> 八風	節之鼓而行之，以 <u>遂</u> 八風	
<u>鉞</u> 田于大陸	<u>遂</u> 田于大陸	
公不聽， <u>鉞</u> 如齊	公不聽， <u>遂</u> 如齊	
弗聽， <u>鉞</u> 躋之	弗聽， <u>遂</u> 躋之	

檔案原文摘錄	手動取代結果	備註
弗從· <u>鉞</u> 作中	弗從· <u>遂</u> 作中	
不如往也·」乃 <u>鉞</u> 行	不如往也·」乃 <u>遂</u> 行	
「將使鼈長而後食之·」 <u>鉞</u> 出	「將使鼈長而後食之·」 <u>遂</u> 出	
<u>鉞</u> 逐之·五日	<u>遂</u> 逐之·五日	
擇其善者而業用之； <u>鉞</u> 滋民	擇其善者而業用之； <u>遂</u> 滋民	
「諾·」 <u>鉞</u> 修舊法	「諾·」 <u>遂</u> 修舊法	
桓公親見之， <u>鉞</u> 使役官	桓公親見之， <u>遂</u> 使役官	
犧牲不略，則牛羊 <u>鉞</u>	犧牲不略，則牛羊 <u>遂</u>	
一戰帥服三十一國· <u>鉞</u> 南征伐	一戰帥服三十一國· <u>遂</u> 南征伐	
<u>鉞</u> 北伐山戎	<u>遂</u> 北伐山戎	
<u>鉞</u> 下拜，升受命	<u>遂</u> 下拜，升受命	
軍譚、 <u>鉞</u> 而不有也	軍譚、 <u>遂</u> 而不有也	
公弗聽， <u>鉞</u> 伐驪戎	公弗聽， <u>遂</u> 伐驪戎	
被羽先升， <u>鉞</u> 克之	被羽先升， <u>遂</u> 克之	
太子 <u>鉞</u> 行，克霍而反	太子 <u>遂</u> 行，克霍而反	
太子 <u>鉞</u> 行，狐突御戎	太子 <u>遂</u> 行，狐突御戎	
乃 <u>鉞</u> 之狄	乃 <u>遂</u> 之狄	
乃 <u>鉞</u> 之梁	乃 <u>遂</u> 之梁	
是以 <u>鉞</u> 於逆命	是以 <u>遂</u> 於逆命	
夫二國士之所圖，無不 <u>鉞</u> 也	夫二國士之所圖，無不 <u>遂</u> 也	
公曰：「非鄭之所知也·」 <u>鉞</u> 不予	公曰：「非鄭之所知也·」 <u>遂</u> 不予	
置而不 <u>鉞</u> ，擊而不勝	置而不 <u>遂</u> ，擊而不勝	
亦不克救， <u>鉞</u> 止于秦	亦不克救， <u>遂</u> 止于秦	

檔案原文摘錄	手動取代結果	備註
若弗忘，是再施不 <u>鉞</u> 也	若弗忘，是再施不 <u>遂</u> 也	
納而不 <u>鉞</u> ，廢而不起	納而不 <u>遂</u> ，廢而不起	
受而載之· <u>鉞</u> 適齊	受而載之· <u>遂</u> 適齊	
將焉用之？」 <u>鉞</u> 行	將焉用之？」 <u>遂</u> 行	
<u>鉞</u> 如楚，楚成王以周禮享之	<u>遂</u> 如楚，楚成王以周禮享之	
彼己之子，不 <u>鉞</u> 其媾	彼己之子，不 <u>遂</u> 其媾	
公出救火而 <u>鉞</u> 殺之	公出救火而 <u>遂</u> 殺之	
二子求公不獲， <u>鉞</u> 如河上	二子求公不獲， <u>遂</u> 如河上	
王入于成周， <u>鉞</u> 定之于邾	王入于成周， <u>遂</u> 定之于邾	
<u>鉞</u> 伐曹、衛，出穀戍	<u>遂</u> 伐曹、衛，出穀戍	
敗楚師于城濮，於是乎 <u>鉞</u> 伯	敗楚師于城濮，於是乎 <u>遂</u> 伯	
<u>鉞</u> 與荊人戰於鄆陵	<u>遂</u> 與荊人戰於鄆陵	
於是乎國人不蠲， <u>鉞</u> 弑諸翼	於是乎國人不蠲， <u>遂</u> 弑諸翼	
於是乎 <u>鉞</u> 伯·	於是乎 <u>遂</u> 伯·	
不克而死·公 <u>鉞</u> 逐羣賊	不克而死·公 <u>遂</u> 逐羣賊	
是 <u>鉞</u> 威而遠權	是 <u>遂</u> 威而遠權	
<u>鉞</u> 令於國人曰	<u>遂</u> 令於國人曰	
出奔曲沃， <u>鉞</u> 刺欒盈	出奔曲沃， <u>遂</u> 刺欒盈	
<u>鉞</u> 不視·楊食我生	<u>遂</u> 不視·楊食我生	
<u>鉞</u> 施邢侯氏	<u>遂</u> 施邢侯氏	
而殺智伯于師， <u>鉞</u> 滅智氏	而殺智伯于師， <u>遂</u> 滅智氏	
王 <u>鉞</u> 置之	王 <u>遂</u> 置之	
<u>鉞</u> 不用	<u>遂</u> 不用	
椒舉奔鄭，將 <u>鉞</u> 奔晉	椒舉奔鄭，將 <u>遂</u> 奔晉	
以夏姬行， <u>鉞</u> 奔晉	以夏姬行， <u>遂</u> 奔晉	

檔案原文摘錄	手動取代結果	備註
又不圖也，乃 <u>鉞</u> 奔晉	又不圖也，乃 <u>遂</u> 奔晉	
<u>鉞</u> 趨而退，歸	<u>遂</u> 趨而退，歸	
<u>鉞</u> 去王·王歸，又求見	<u>遂</u> 去王·王歸，又求見	
不從， <u>鉞</u> 使為白公	不從， <u>遂</u> 使為白公	
<u>鉞</u> 伐齊·齊人與戰於艾陵	<u>遂</u> 伐齊·齊人與戰於艾陵	
<u>鉞</u> 自殺·將死	<u>遂</u> 自殺·將死	
王去其國， <u>鉞</u> 至于郢	王去其國， <u>遂</u> 至于郢	
吾謂吳王將 <u>鉞</u> 涉吾地	吾謂吳王將 <u>遂</u> 涉吾地	
我 <u>鉞</u> 踐其地	我 <u>遂</u> 踐其地	
奔 <u>鉞</u> 可出	奔 <u>遂</u> 可出	
王 <u>鉞</u> 出，夫人送王	王 <u>遂</u> 出，夫人送王	
王 <u>鉞</u> 出，大夫送王不出檐	王 <u>遂</u> 出，大夫送王不出檐	
右軍乃 <u>鉞</u> 涉而從之	右軍乃 <u>遂</u> 涉而從之	
越師 <u>鉞</u> 入吳國，圍王臺	越師 <u>遂</u> 入吳國，圍王臺	
<u>鉞</u> 使之行成於吳	<u>遂</u> 使之行成於吳	
越君其次也， <u>鉞</u> 滅吳	越君其次也， <u>遂</u> 滅吳	
<u>鉞</u> 興師伐吳，至於五湖	<u>遂</u> 興師伐吳，至於五湖	
不傷越民， <u>鉞</u> 滅吳	不傷越民， <u>遂</u> 滅吳	
<u>鉞</u> 乘輕舟以畜於五湖	<u>遂</u> 乘輕舟以畜於五湖	
夫唯能下其羣 <u>臣</u>	夫唯能下其羣 <u>臣</u>	