



圖書館與中文譯音使用原則

——記通用拼音與漢語拼音電腦轉換系統之建置

嚴鼎忠 國家圖書館研究組編輯

一、緣起

行政院於民國 91 年 8 月 22 日，函示准予備查教育部函報的「中文譯音使用原則草案」暨「推動中文譯音相關配套措施分工表」。在分工表中，有關與國際接軌——研發電腦轉換系統的分工有：

1. 有關通用拼音與漢語拼音國際接軌問題，急需研發電腦轉換系統以予解決，由教育部國語推行委員會負責辦理；
2. 有關國家圖書館配合中文譯音採用通用拼音系統在資料查詢時，增置提供電腦系統轉換功能，由國家圖書館負責辦理。

民國 91 年 10 月 4 日，教育部以台(91)語字第 91150627 號函，委託國家圖書館（以下簡稱，本館）辦理「通用拼音與漢語拼音的電腦轉換系統」研發工作，亦即上述兩項業務都交由本館負責辦理，完成期限訂在 92 年 5 月 31 日。

二、背景分析

「中文譯音使用原則」經政府公布施行後，我國的中文譯音以使用「通用拼音」為主，由於國際上較為通行的中文譯音為中國大陸所使用的漢語拼音，因此，在推動通用拼音政策時，必須考量通用拼音與國際接軌的相關問題，解決此一問題，所需辦理的事項包括：

1. 首要解決通用拼音與漢語拼音的對照問題，進而研發通用拼音與漢語拼音的電腦轉換系統，以應各種電子資料庫轉換上的需求。

2. 為使國內使用者能以通用拼音，查詢國內外圖書館館藏資料，圖書館提供查詢資料的電腦，必須建置中文通用拼音轉換成漢語拼音的輸入法；為使國外使用者能以通用拼音、漢語拼音、韋傑式（威妥瑪）羅馬拼音等，查詢國內圖書館的館藏資料，圖書館必須在現有的資料庫，增置各種拼音間相互轉換的電腦系統，提供便捷的查詢功能。

綜合上述，圖書館界因應中文譯音增置資料查詢功能所需要辦理的項目有四：

1. 依據教育部編印的「中文譯音使用原則」等為素材，研發通用拼音與漢語拼音間，以音節對音節轉換為主的電腦轉換系統。
2. 建置中文譯音轉換查詢系統，安裝於網站上供大眾上網查詢。
3. 提供全國各級圖書館有關通用拼音轉換成漢語拼音的輸入法，以供查詢國內外圖書館館藏資料之需；亦可提供外籍人士以通用拼音、漢語拼音、韋傑式（威妥瑪）羅馬拼音等，查詢國內圖書館的館藏資料。
4. 研發設計拼音轉換介面與機制，供圖書館管理系統或書刊文獻資料庫，於查詢端與接收端使用，以解決資料查詢上與國際接軌的問題，獲取較滿意的查詢結果。由於圖書館管理系統、各圖書館館藏目錄、各類型或主題的資料庫，欲達到以通用拼音作為查詢的基本項目，牽涉到系統的軟硬體設備與環境，基於差異性過於龐大，暫列為下個階段辦理項目。



三、辦理經過

拼音系統向來並非圖書館主要業務範圍，對於學校不教、又不熟悉的業務，於接獲命令後，只好藉助專家學者的力量，戮力完成。謹就辦理情形，分成研擬電腦轉換系統、審查系統需求規格、建置轉換系統網站、購置拼音輸入工具等四項工作，略述經過如下：

(一) 研擬電腦轉換系統

民國 92 年 1 月 14 日，教育部台 (92) 語字第 0910181073 號函，核定國家圖書館所提的計畫書，雙方並完成「教育部圖書館應用通用拼音與漢語拼音電腦轉換系統計畫契約書」的簽約程序。民國 92 年 3 月 17 日，本館委託交通大學柯皓仁教授研擬「圖書館通用拼音與漢語拼音電腦轉換系統規格書」。規格書以國語會《中文譯音使用原則》為系統設計的主要依據，是項研擬作業於 4 月 10 日完成初稿。

(二) 審查系統需求規格

民國 92 年 4 月 25 日，本館召開規格書審查會議，邀請資訊界與圖書資訊界等相關領域的專家學者，進行系統需求書、規格書等的審查作業，以確保思慮周延、執行無誤，落實有成。會中學者專家提出寶貴建議，交通大學復於 5 月 9 日完成修正定稿。

民國 92 年 5 月 17 日，教育部召開「研修中文譯音使用原則事宜會議」，討論修訂《中文譯音使用原則》。由於該原則的修訂，將會造成電腦轉換系統規格的變動，遂於 5 月 30 日，提出擬俟《中文譯音使用原則（修訂版）》公佈實施後，再據以修正電腦轉換系統規格書。6 月 18 日，教育部同意將「教育部圖書館應用通用拼音與漢語拼音電腦轉換系統計畫契約書」的履約期限往後延長至民國 92 年 9 月 30 日。

民國 92 年 7 月 24 日，國家圖書館以台 (92) 圖研字第 920002089 號函，將「圖書館應用通用拼音與漢語拼音電腦轉換系統招標規範（稿）」陳報教育部，並獲同意辦理招商事宜。

(三) 建置轉換系統網站

民國 92 年 9 月 15 日，「圖書館應用通用拼音與漢語拼音電腦轉換系統」採購案完成履約。該案經公開招商由網際智慧股份有限公司承攬製作，92 年

9 月 19 日通過國家圖書館初步驗收，復於 92 年 9 月 29 日完成於教育部電算中心主機的安裝作業。此一於教育部網站建置的中文譯音查詢系統，將可提供各界有關中文譯音的查詢，並推廣通用拼音。

(四) 購置拼音輸入工具

為解決使用者（以外籍人士居多）能以通用拼音、漢語拼音、韋傑式（威妥瑪）羅馬拼音等，查詢國內各圖書館的館藏資料等問題，圖書館必須在提供查詢資料的電腦上，安裝可輸入通用拼音的輸入法。民國 92 年 9 月 30 日，「自然輸入 2003 圖書館版全國圖書館使用授權」採購案完成履約，並將該輸入法寄發至全國 5,290 所圖書館，供安裝於全國各級圖書館建築物內的電腦使用，以滿足使用者對於通用等各種拼音輸入法的需求。

四、通用拼音與漢語拼音轉換系統

國立交通大學（註 1）接受本館委託辦理通用拼音與漢語拼音的電腦轉換系統規格書，針對本案僅摘錄重要的建議事項於下：

(一) 依循標準

通用拼音：以教育部國語推行委員會 91 年 9 月所公佈的「中文譯音使用原則」為參考依據。

漢語拼音：以國際化組織 ISO 所制定的漢語拼音版本 ISO 7098:1991 (<http://www.iso.org/>) 為參考依據。（大陸官方版本為 1958 年公佈的漢語拼音方案）。

依據教育部所公佈的「中文譯音使用原則」與 ISO 國際標準組織公告的漢語拼音方案，兩者在轉換上會有以下幾類情形發生：

1. 依據教育部公佈的《中文譯音使用原則》，以「注音」為轉換基礎，通用拼音與漢語拼音擁有標準的轉換方法。

建議事項：對於通用拼音以及漢語拼音中該縮寫與該省略的規則詳列於轉換系統的說明。

2. 通用拼音與漢語拼音特有與傳統注音不同的發音，並無特定的轉換原則。通用拼音擁有的部分客語、台語發音，例如 q、v、x 三個字母在拼寫台灣華語時並不使用，並無法在漢語拼音找到絕對相同的對應。

建議事項：初期目標以 Big 5 中文字集所包含



的中文字為轉換的內容。遠程目標可以制定常用客語、台語與漢語的對應規則，也可以透過電腦系統建議類似的拼音。

(二) 通用拼音及漢語拼音與中文字詞轉換

1. 通用拼音及漢語拼音在轉換成中文字詞的時候，會有一對多的情形發生。舉例而言，Hua（通用拼音以及漢語拼音），會對應到許多的中文字，例如花、話、畫等。

建議事項：在轉換拼音跟中文字詞遇到一對多情形發生時，應提供使用者選擇字詞，或者由系統參考「常用名詞詞庫」，挑出符合的字詞，抑或利用人工智慧選字的方式根據前後文由系統推薦出適合的字詞。「常用名詞詞庫」應以教育部所公佈「常用語詞」調查報告書所公佈內容為主。

2. 中文字詞轉換到通用拼音及漢語拼音的時候會有特殊名詞的考量。舉例而言，不管任何拼音對於台北的表示應該以常用的 Taipei 為主。

建議事項：建立「特殊名詞轉譯對照表」，除了包含國際通用及慣用翻譯、縣級以上行政區域名稱、行政區域劃分、國際機場港口的譯寫外，並能提供動態增加新詞。「特殊名詞轉譯對照表」應以教育部所公佈的「中文譯音使用原則」內所列範圍為主。

3. 完整的通用（漢語）拼音包含聲調、分隔符號、以及特有字元，將可以找到對應的漢語（通用）拼音。而為了要提供方便的中文譯音輸入環境，抑或輸入環境的特殊限制，系統必須提供不完整的拼音輸入及輸出，例如省略聲調、忽略分隔符號以及無法輸入特殊字元等，將會造成拼音轉換上的困擾與錯誤。

建議事項：中文譯音轉換系統應提供輸入包含音調，應提供輸出包含音調以及不包含音調輸出的轉換系統。

(三) 通用拼音與漢語拼音電腦轉換系統

1. 電腦介面無法提供使用者輸入標準拼音符號，因此造成拼音的混淆。舉例而言，漢語拼音的 u 無法讓使用者辨別是ㄨ還是ㄩ，造成轉換上的混淆。

建議事項：提供使用者選擇字詞，或者由系

統參考「常用名詞詞庫」，挑出符合的字詞，抑或利用人工智慧選字的方式根據前後文由系統推薦出適合的字詞。

2. 使用者若無輸入分隔符號，將會造成翻譯上的混淆。舉例而言，通用拼音的女兒 nyuer 若沒有以分隔符號隔開，將會造成斷詞上 nyu 及 er，或者 n、yue 及 r 的混淆。

建議事項：在使用者輸入畫面提供使用者利用大小寫、破折號、空白或者單引號等作為字詞分開的方式，並由電腦系統判斷使用者採用的方式。

3. 在拼音與中文字詞的轉換當中，無法輸入聲調符號，造成轉換上存在同一譯音對應到多個中文字詞的情形出現。

建議事項：提供使用者選擇字詞，或者由系統參考「常用名詞詞庫」，挑出符合的字詞，抑或利用人工智慧選字的方式根據前後文由系統推薦出適合的字詞。透過輔助數字，ba0（吧），ba1（八）或者 ba（省略不輸入聲調視為一聲），ba2（拔），ba3（把），ba4（爸）的輸入，經過系統的轉譯即可得出聲調，輸出時再行轉換為正確拼音。

4. 通用拼音與漢語拼音的特殊字元在電腦系統上顯示的問題。（註：特殊字元是指除了標準 26 個英文字母之外無法以鍵盤輸入的字元，包含變形的英文字母以及聲調符號）。此類問題需考量所使用的系統介面為何？若是以網頁方式呈現，World Wide Web Consortium (<http://www.w3c.org/>) 所訂定之 HTML 的 Character Entity Reference (ISO-Latin-1)，僅包含部分通用拼音以及漢語拼音的特殊字元例如：「ê」、「ü」。

(1) 一般軟體若以 ASCII 為文字編碼方式，超過編號 127 的 extend ASCII 編碼方式，可以顯示部分特殊符號。

(2) Unicode 所能顯示的特殊字元比 ASCII 多，但僅能顯示部分拼音的特殊字元。

（註：Windows98 不支援 Unicode）

建議事項：初期目標建議採用圖案來顯示特殊字元。遠程目標以協調各相關國際組織與軟體大廠加入通用拼音特殊字元的編碼。



(四) 其他建議事項

1. 建議制定通用拼音正詞法基本規則，制定通用拼音分詞連寫、外來語拼寫、人名地名拼寫、標點符號以及語調、移行等規則。方便將來通用拼音撰寫時，擁有更完整的拼音原則，在電腦轉換系統也能夠有更多的準則可以依循。
2. 若要制定通用拼音與漢語拼音電腦轉換系統並開放程式碼供各界使用，必須考量拼音轉換中，必須運用到的智慧型字詞轉換所牽涉的智慧財產權問題。依據實際可行的原則，建議開放程式碼部分以教育部所公佈的「中文譯音使用原則」為範圍。
3. 在拼音與中文字詞轉換方面，初期目標建議以教育部公佈「常用名詞詞庫」為轉換的主要內容。
4. 在拼音轉換所參照的「特殊名詞轉譯對照表」，初期目標建議以教育部公佈的「中文譯音使用原則」為參考依據。

五、中文譯音查詢系統

為樹立通用拼音的標準規範，方便全國民眾多加利用，設立查詢網站，是為可行的作法之一。遂於教育部國語推行委員會的網站中，架設一「中文譯音查詢系統」網頁 (<http://140.111.34.69:8080/national-library/index.jsp>)，以為推廣的根源。該系統包含：中文譯音拼音查詢、拼音轉換查詢與檔案下載三部分。

查詢系統分成中文譯音拼音查詢、拼音轉換查詢與檔案下載三部分。

1. 中文譯音拼音查詢
 - (1) 中文單字譯音查詢：使用者輸入中文單字，將其轉換為通用拼音、漢語拼音、韋傑式（威妥瑪）羅馬拼音、國音一式（注音）以及國音二式。如中文單字有多種拼音應一同顯示（破音字情形）。
 - (2) 中文文句譯音查詢：使用者輸入繁體中文文句，系統必須將之轉換為通用拼音文句及漢語拼音文句。中文文句轉換應以單字為轉換基礎，如有多種拼音應一同顯示。

(3) 中文地名譯音查詢：由於許多地名的譯音已經沿用多年，為求其一致性，不以通用拼音方式譯出，乃依據內政部 92 年 6 月 30 日公布的「地名譯寫原則」建置專屬的地名譯音查詢網頁。

(4) 通用拼音轉中文：使用者輸入通用/漢語拼音文字或者文句，將其轉換為中文文字或者文句。轉換應以拼音單字為轉換基礎，如有多種轉換的對應一同顯示。

2. 拼音轉換查詢

(1) 通用拼音轉漢語拼音：使用者輸入通用拼音，將其轉換成漢語拼音。拼音顯示如有非英文 26 字母的符號（例如 e、? 等）以最接近的英文單字來表示。

(2) 漢語拼音轉通用拼音：使用者輸入漢語拼音，將其轉換成通用拼音。漢語拼音轉通用拼音，在不輸入聲調的情形下會有多種拼音對照的結果產生，所以轉譯結果如果發生一對多的情況，應另外提供聲調選擇選項讓使用者挑選正確的通用拼音。

3. 檔案下載

研發完成通用拼音與漢語拼音的電腦轉換系統，並公開所取得的原始碼及程式集（資料庫）方式，提供各界運用。內容為：(1) 通用拼音轉漢語拼音（不含音調）；(2) 漢語拼音轉通用拼音（不含音調）；(3) 通用拼音轉漢語拼音（含音調）；(4) 漢語拼音轉通用拼音（含音調）等四種。提供下載的檔案包括：(1) WinAPI 執行檔；(2) WinAPI 程式碼。

另值得一提的是，本系統招商開發規範書中，訂定了「測試標準」，對於拼音轉換系統與通用拼音與漢語拼音轉換網站，需要檢測的項目、內容與方法都事先加以規範，廠商完成系統後，需自行依據測試標準進行測試，並提出測試報告。這對於系統開發的品質控制，有了一定的保證，是值得推廣的作法。

六、自然輸入 2003 圖書館版授權使用

為推廣普及「通用拼音」，全國各級圖書館所使用的電腦，都需要具備通用拼音的中文輸入法。由



於部分外籍人士學習中文，採行的是漢語拼音，因此，必須兼備通用拼音與漢語拼音兩者的中文輸入法，才是圖書館所需要的輸入法。於是選定「自然輸入二〇〇三圖書館版」成為全國各級圖書館使用輸入法。

(一) 主要規格

1. 支援注音、倉頡、漢語拼音及通用拼音方式輸入中文，讓國人與外國人都可以方便輸入中文。
2. 注音輸入支援標準鍵盤、倚天鍵盤、許氏鍵盤，使用者也可自行定義注音鍵盤位置。
3. 提供至少兩種以上放大尺寸螢幕鍵盤輸入功能，方便弱視朋友輸入中文。
4. 提供人工智慧自動選擇同音字（適用注音、通用拼音、漢語拼音輸入）或同碼字（適用倉頡輸入）功能，可依據使用者所輸入的前後文，自動選擇最恰當的同音字或同碼字，使用者不必再自己選擇。
5. 提供近音查詢功能，利用快速鍵可尋找近音字，協助注音或拼音不標準的使用者可輕鬆輸入中文。
6. 由輸入法輸出中文時，可選擇或切換輸出繁體字或簡體字，或選擇輸出中文的通用拼音拼碼、漢語拼音編碼、注音符號編碼及點字碼，以方便查詢國內外各種中文介面或非中文介面的資料庫系統或應用於專業文件編輯。
7. 由輸入法輸出中文時，可另外以視窗同步顯示該字的通用拼音編碼、漢語拼音編碼、注音符號編碼、倉頡編碼、點字編碼等，做為輔助查詢使用。
8. 由輸入法輸出中文時，可以同步朗讀該字讀音，亦可以由使用者自行選擇文章段落進行整篇發音，發音時可以自動根據前後文自動判斷破音字、數字，選擇最恰當的唸法，以支援視障朋友查閱各種電子書或使用圖書資料庫系統。
9. 提供片語輸入功能，可將一段文字或片語用數字或英文簡碼自行定義，輸入時只需鍵入代碼即可輸入所定義的文字或片語（例：定義『NL』為『國家圖書館』，輸入時只需輸入

『NL』即可出現『國家圖書館』）。可輸入片語數 6 萬個以上，並提供片語匯出及匯入功能。

10. 提供部首查詢輸入功能，利用部首查詢不知如何發音或拼碼的生字。

(二) 系統需求

1. 作業系統：Windows 2000、XP、2003。（Windows 98、ME、NT 需另外下載安裝程式，才可進行安裝）
2. 個人電腦：Pentium II 以上的電腦，64MB 以上記憶體
3. 硬碟空間：10MB（不含語音模組）；16.5MB（含語音模組）

(三) 授權使用對象

1. 臺閩地區圖書館暨資料單位名錄 (<http://book.ncl.edu.tw/libdir/search.asp>) 收錄的所有圖書館。於授權期間中若圖書館數量有增加，仍需提供相關服務。
2. 於授權使用期間內收錄於該名錄的圖書館都可以獲得完整的使用授權。
3. 廠商函送至各圖書館，並附上產品光碟片及授權書，及試用十五天的授權；各圖書館需向廠商所提供的專屬網頁註冊後，才能取得正式使用的序號
4. 限於全國各級圖書館建築物（室）內的電腦安裝使用，非圖書館服務或工作區域內的電腦，不在授權使用範圍之內。（註 2）

(四) 授權使用期限

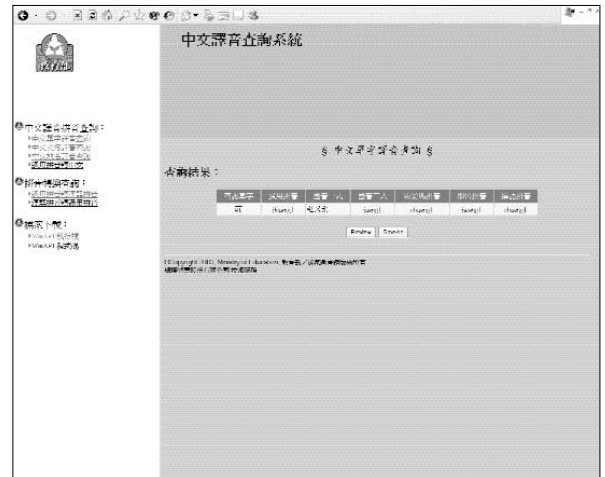
1. 民國 92 年 9 月 30 日起至民國 94 年 3 月 31 日止。
2. 授權期滿後，各圖書館取得授權期間內最後版本的永久使用權。

(五) 授權使用服務

1. 軟體送達：民國 92 年 9 月 30 日前，由廠商以掛號郵件寄發產品光碟片、授權書及相關說明文件至「臺閩地區圖書館暨資料單位名錄」內的各圖書館。
2. 專屬網頁：建置專屬網頁，提供產品簡介、安裝設定、使用說明、答客問、軟體下載、版本更新、最新消息等相關資訊與服務。
3. 服務管道：提供專線電話、傳真、電子郵件



中文譯音查詢系統——文句譯音查詢



中文譯音查詢系統——單字查詢

【專題報導】

等多元服務管道，服務時間必須至少比照政府部門正常上下班時間。

4. 使用手冊：提供線上快速簡版及詳細版的使用手冊，可供使用者自行下載運用。
5. 版本更新：授權使用期限內，免費提供自然輸入的線上更新版本。

七、結語

圖書館基於保存人類知識的重責大任，不能也無法拒收任何資料，因此，面對通用拼音與漢語拼音兩者間，誰臨駕誰的爭端？不會有任何主觀的認知與考量。因為，凡進到圖書館中的資料，都得處理、保存與傳布。本計畫也是基於此一立場，站在歷史的角度上思考，執行完成公開原始碼的通用拼音與漢語拼音的電腦轉換系統、建置中文譯音查詢系統網頁及提供全國各級圖書館使用的電腦得以通用拼音方式輸入中文等三項成果，冀望在推動政府既定政策之際，亦能兼顧圖書館與國際接軌的實際需求。



自然輸入 2003 圖書館版光碟片

註釋

1. 「圖書館通用拼音與漢語拼音電腦轉換系統規格書」之專案研究計畫由國立交通大學圖書館柯皓仁教授主持，研究人員為：黃夙賢、林忠億、葉鎮源。
2. 如遇使用單位需要將使用範圍升級至同單位之不同部門，或擴及全單位。其相關規定說明如下：
 - (1) 費用：以目前市場銷售之授權（自然輸入 2003 專業版）建議價格 4 折優惠提供。
 - (2) 價格有效期間：同授權時間。
 - (3) 各級單位之授權價格如下：

教育授權	建議售價	優惠價格
中小學全校授權	9,900	3,960
完全中學全校授權	15,000	6,000
高中職全校授權	12,000	4,800
大專院校全校授權	29,000	11,600
大學科系全系授權	12,000	4,800
大學全校授權	49,000	19,600

企業授權	建議售價	優惠價格
授權 30-49 人，每人價格	330	132
授權 50-99 人，每人價格	290	116
授權 100 人以上，每人價格	250	100
授權 300 人以上，每人價格	130	52
全體授權不限人數	79,999	31,999