

行政院國家科學委員會補助專題研究計畫  成果報告  
 期中進度報告

人文研究與電腦科技整合計畫一

漢字古今音電腦檢索系統（五）現代方音（2）

計畫類別： 個別型計畫  整合型計畫

計畫編號：NSC 90-2411-H-002-039-A9

執行期間：90年8月1日至92年5月31日

計畫主持人：楊秀芳

共同主持人：何大安

計畫參與人員：

成果報告類型(依經費核定清單規定繳交)： 精簡報告  完整報告

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、  
列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年  二年後可公開查詢

執行單位：台灣大學中文系

中華民國 92 年 6 月 6 日

# 行政院國家科學委員會專題研究計畫成果報告

## 人文研究與電腦科技整合計畫— 漢字古今音電腦檢索系統（五）現代方音（2）

計畫編號：NSC 90-2411-H-002-039-A9

執行期限：九十年八月一日至九十二年五月三十一日

主持人：楊秀芳 教授 國立台灣大學中國文學系

共同主持人：何大安 研究員中央研究院語言學研究所籌備處

計畫參與人員：

### 一、中文摘要

本計畫在執行上分為兩個部分，一個是建立語音的資料庫，一個是設計檢索系統。建立語音資料庫方面，本年度輸入漢語現代方言共包括陽西塘口、陽春三甲、信宜思賀、信宜錢排、高州新垌、電白沙琅、化州新安、廉江石角、廉江青平等九個粵西客家方言點的資料，加上過去五年所建立的中古後期、中古前期、上古，以及數十種現代方音資料，目前資料庫語料已具相當規模。本資料庫除了建有古今字音以供檢索外，並建有輸入所據的研究或調查報告書目，以及我們對這些研究或調查資料所作的校訂說明。

在設計檢索系統方面，本年度有重大突破。電腦工程師克服萬難，於今已開發完成一個可以上網檢索的系統，近期內將可提供各界使用。

本計畫從民國 86 年 6 月開始執行，目前已經完成第五期的工作。第一期輸入隋唐音，第二期輸入南北朝音，第三期輸入上古音及若干現代方音，第四期輸入現代方音(1)，第五期輸入現代方音(2)。本件為第五期的成果報告，隨件附上磁片一張，內含本年度輸入的 9

種現代方音資料。

### 關鍵詞：

上古音系、南北朝音系、隋唐音系、現代方言、語音資料檢索系統。

### Abstract

Readings of thousands of selected words from nine dialects have been examined and recorded during the last year. An annotated bibliography has been included in the database as well. The database is to be on line after the website is set up soon.

### Keywords:

Old Chinese, Middle Chinese, Dialect, Reading Search System

### 二、緣由與目的

本計畫是要建立一個漢字古今字音的資料庫，以作為漢學研究上字音查

詢的工具。這個資料庫除了可以依字頭查詢古今字音外，還有檢索系統可以讓使用者依自己的需要，設計不同的檢索條件，做大量資料的閱讀和分析。

這個資料庫的主要用途可分三方面來說：(一)用來查詢漢字古音，以賞析古詩詞的聲韻之美，或訓詁解讀古典文獻。(二)用來查詢現代方言的音讀。(三)用來從事古今字音演變的研究。本資料庫的交叉檢索功能，可以提供學術界做各種有關古今字音的研究。

### 三、結果與討論

本計畫在執行上分為兩個部分，一個是設計檢索系統，一個是建立語音的資料庫。

我們最初利用現成的資料庫軟體建立語音資料庫，這個資料庫需要一個檢索系統才能提供各界使用。四年前我們聘請電腦工程師設計這個檢索系統，當時規劃的是配合單機使用的環境；後來為因應電腦網路使用的趨勢，於三年前開始嘗試讓資料庫上網的可能性。由於本資料庫不僅有大量的罕用字、又有許多特殊的國際音標符號、又要提供網上自由檢索的功能，種種這些條件，使得這成為一個極具挑戰性的工作。經過三年的努力，我們的電腦工程師克服種種困難，現在已經可以讓使用者上網檢索資料。這是本年度最重要的成績之一。

本計畫最初規劃時是配合單機使用的環境，後來嘗試升級轉型，希望讓資料庫能夠上網。這個脫胎換骨的變化，使得原先輸入的國際音標符號不再適用於新環境，因此我們必須改用網路允許的辦法，為每一條語料重新輸入國

際音標符號。這個工作，我們目前利用有限的人力和經費，逐步更新代換中。

在將語音資料逐一輸入《廣韻》資料庫的時候，我們必須先確認音字的關係無誤。而《廣韻》常見一字多切，因此我們不止要判斷語料的音字關係是否正確，事實上還要判斷各條語音資料所對應的反切為何。這個工作非常重要，因為如果將語音資料輸入到非對應的反切底下，它將誤導得出錯誤的音變規則。因此本計畫需要用許多時間作學理的判斷，計畫進度常常因此而受到影響。

檢討音字關係需要利用歷史語言學的知識，關於這個，我們撰寫方言本字考證相關理論方面的研究論文，例如楊秀芳 1999、2000、2001 等，這些論文將登上網頁提供使用者參考。

### 四、計畫成果自評

執行這個計畫，我們在講求品質和效率的要求下，期許它在語音資料的呈現上要做到正確無誤，在電腦技術上要做到功能強大而且檢索方便。以這樣的目標來看，我們在語音資料的輸入上已經盡量做到正確無誤，這包括檢討每一筆資料的音字關係，務使語音資料都能放在正確的反切欄位下；同時還包括對於所依據的研究成果或調查報告的檢討，例如所據調查報告若在內容上有前後不一致的現象，我們必須仔細檢討訂正；最後還經過至少兩次的仔細校對。至於電腦技術的問題，經過多年的努力，我們的電腦工程師克服種種困難，現在已經可以讓使用者上網檢索資料。

整體說來，本資料庫不僅在電腦技術的開發上不容易，在輸入資料之前，我們投入許多心血，對語料逐條檢討，以確認音字關係無誤，因此建立本資料庫的難度，遠非一般全文檢索或電子書

所可比擬。

## 五、參考文獻

丁邦新

1982 〈漢語方言區分的條件〉。  
《清華學報》14.1,2: 257-273, 新竹。

丁邦新、張雙慶編

2002 《閩語研究及其與周邊方言的關係》。香港，中文大學出版社。

中國社會科學院語言研究所

1998 《方言調查字表》。北京，  
商務印書館。

中國社會科學院、澳大利亞人文科學院

1987 《中國語言地圖集》。香港，  
朗文出版社。

北京大學中國語言文學系語言學教研室

1989 《漢語方音字匯》(第二  
版)。北京，文字改革出版社。

1995 《漢語方言詞匯》。北京，  
語文出版社。

江敏華

1998 《台中縣東勢客語音韻研究》。  
台灣大學中文研究所碩士論文。

呂嵩雁

1999 《閩西客語音韻研究》。台  
灣師範大學國文研究所博士論文。

李榮

1966 《古今字音對照手冊》。太

平書局，香港。台北廣文書局翻印《國  
音中古音對照表》。

李珍華、周常楫

1993 《漢字古今音表》。北京新  
華書店。

李如龍、張雙慶

1992 《客贛方言調查報告》。廈  
門大學出版社。

李如龍等

1999 《粵西客家方言調查報  
告》。廣州，暨南大學出版社。

袁家驊等

1989 《漢語方言概要》(第二  
版)。北京，文字改革出版社。

張光宇

1996 《閩客方言史稿》。台北，  
南天書局。

張琨

1990 〈客家方言中《切韻》全濁  
上聲字的讀法〉。《王力先生紀念論文  
集》: 407-419, 北京，商務印書館。

張維耿

1995 《客家話詞典》。廣東人民  
出版社。

黃雪貞

1995 《梅縣方言詞典》。江蘇教  
育出版社。

梅祖麟

1995 〈方言本字研究的兩種方

法)。*《吳語和閩語的比較研究》*: 1-2。

楊秀芳

1999 〈方言本字研究的探義法〉。In Alain peyraube and Chaofen Sun eds., *Linguistic Essays in Honor if Mei Tsu-Lin: Studies on Chinese Historical Syntax and Morphology*: 299-326。Paris: ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES, Centre de Recherches Linguistiques sur l'Asie Orientale.

2000 〈方言本字研究的觀念與方法〉。*《漢學言究》*第18卷特刊「臺灣語言學的創造力專號」: 111-146。

2003 〈從方言比較論吳閩同源詞「撼」〉。*《語言暨語言學》*4/1: 167-196。

楊時逢

1957 *《台灣桃園客家方言》*。中央研究院歷史語言研究所單刊甲種之二十二。台北，中研院史語所。

1971 〈台灣美濃客家方言〉。*《中研院史語所集刊》*42/3: 405-465。

1982 〈江西方言的內部紛歧現象〉。*《清華學報》*新14: 307-326。

董同龢

1948 〈華陽涼水井客家話記音〉。*《中研院史語所集刊》*19: 81-201。

劉綸鑫

1999 *《客贛方言比較研究》*。北京，中國社會科學出版社。

2001 *《江西客家方言概況》*。南昌，江西人民出版社。

劉鎮發

2001 *《香港客粵方言比較研究》*。廣州，暨南大學出版社。

龍岩市地方志編纂委員會

1993 *《龍岩市志》*。中國科學技術出版社。

橋本萬太郎

1972 *《客家語基礎語彙集》*。東京，亞非言語文化研究所。

藍小玲

1999 *《閩西客家方言》*。廈門大學出版社。

Norman, Jerry

1989 What is a Kejia Dialect? Paper presented in Second International Conference on Sinology, Academia Sinica, Taipei, Taiwan, 29-31 Dec. 1986. Published in the *Proceedings of 2<sup>nd</sup> International Conference on Sinology: Section on Linguistics & Paleography*. Taipei: Academia Sinica, 1989:323-344.

Sagart, Laurant

1998 On Distinguishing Hakka and Non-Hakka Dialects. *JCL* 26/2: 281-302。

## 可供推廣之研發成果資料表

 可申請專利 可技術移轉

日期: 92年6月6日

國科會補助計畫	計畫名稱：人文研究與電腦科技整合計畫—漢字古今音電腦檢索系統 計畫主持人：楊秀芳、何大安 計畫編號：NSC 90-2411-H-002-039-A9 學門領域：語言學
技術/創作名稱	漢字古今音電腦檢索系統
發明人/創作人	楊秀芳、何大安、劉偉祥、黃宗瀚、郭樹欣、李慧菁、嚴立模、陳靜芳等
技術說明	<p>中文：</p> <p>本計畫研發「漢字古今音電腦檢索系統」資料庫，執行時間自86年至92年，主要以《廣韻》為基本架構，根據學者的研究成果，輸入漢字各期古音及現代數十種方言的語音資料，以提供學者檢索之便。本計畫最初是以單機使用的環境來規劃，後來因應電腦網路使用的趨勢，於三年前開始嘗試讓資料庫上網的可能性。由於本資料庫不僅有大量罕用字、又有許多特殊的國際音標符號、又要提供網上自由檢索的功能，種種這些條件，使得這成為一個極具挑戰性的工作。經過三年的努力，我們的電腦工程師克服萬難，現在已經可以讓使用者上網檢索資料。本資料庫不僅電腦技術開發不易，在輸入資料之前，我們投入許多心血，對語料逐條檢討，以確認音字關係無誤，因此建立本資料庫的難度，遠非一般全文檢索或電子書所可比擬。</p> <p>本資料庫為國科會資助的專案計畫成果，免費提供各界使用。唯為維護智慧財產權，並避免遭人竊取轉為商用，因此有必要申請專利。</p> <p>英文：</p> <p>The purpose of this project is to establish a database of the readings of Chinese characters in different times and dialects. By using Kuang-Yun as the framework, we have coded readings from dozens of dialects and phonological reconstructions of different periods. We also developed an on-line searching system for the database in the past three years.</p> <p>A database with searching system like this cannot be successful without various computer technologies as well as intellectual efforts. For that reason, we'll apply for the patent before it is formally released.</p>
可利用之產業 及 可開發之產品	

<p><b>技術特點</b></p>	<p>本資料庫在技術上重要的突破有二：</p> <p>(一) IPA 符號顯示: 在 IE 瀏覽器版本 4 以上，使用者已安裝 IPA 相關字型前提下，在判斷欲顯示 IPA 符號的 CODE 下加入 &lt;FONT FACE&gt; 的 TAG，來指定使用相關字型顯示。</p> <p>(二) IPA 符號輸入: 製作一 IPA 符號的表格，使用者經由點選相關符號產生要查詢符號的編碼，以此方式進行資料庫檢索</p>
<p><b>推廣及運用的價值</b></p>	<p>(一) 用來查詢漢字古音，以賞析古詩詞的聲韻之美，或訓詁解讀古典文獻。</p> <p>(二) 用來查詢現代方言的音讀。</p> <p>(三) 用來從事古今字音演變的研究。本資料庫的交叉檢索功能，可提供學術界做各種有關古今音的研究。</p>

- 1. 每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送 貴單位研發成果推廣單位（如技術移轉中心）。
- 2. 本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。
- 3. 本表若不敷使用，請自行影印使用。

## 可供推廣之研發成果資料表

 可申請專利 可技術移轉

日期：92年6月6日

國科會補助計畫	計畫名稱：人文研究與電腦科技整合計畫—漢字古今音電腦檢索系統 計畫主持人：楊秀芳、何大安 計畫編號：NSC 90-2411-H-002-039-A9      學門領域：語言學
技術/創作名稱	漢字古今音電腦檢索系統
發明人/創作人	楊秀芳、何大安、劉偉祥、黃宗瀚、郭樹欣、李慧菁、嚴立模、陳靜芳等
技術說明	<p>中文：</p> <p>本計畫研發「漢字古今音電腦檢索系統」資料庫，執行時間自86年至92年，主要以《廣韻》為基本架構，根據學者的研究成果，輸入漢字各期古音及現代數十種方言的語音資料，以提供學者檢索之便。本計畫最初是以單機使用的環境來規劃，後來因應電腦網路使用的趨勢，於三年前開始嘗試讓資料庫上網的可能性。由於本資料庫不僅有大量罕用字、又有許多特殊的國際音標符號、又要提供網上自由檢索的功能，種種這些條件，使得這成為一個極具挑戰性的工作。經過三年的努力，我們的電腦工程師克服萬難，現在已經可以讓使用者上網檢索資料。本資料庫不僅電腦技術開發不易，在輸入資料之前，我們投入許多心血，對語料逐條檢討，以確認音字關係無誤，因此建立本資料庫的難度，遠非一般全文檢索或電子書所可比擬。</p> <p>本資料庫為國科會資助的專案計畫成果，免費提供各界使用。唯為維護智慧財產權，並避免遭人竊取轉為商用，因此有必要申請專利。</p> <p>英文：</p> <p>The purpose of this project is to establish a database of the readings of Chinese characters in different times and dialects. By using Kuang-Yun as the framework, we have coded readings from dozens of dialects and phonological reconstructions of different periods. We also developed an on-line searching system for the database in the past three years.</p> <p>A database with searching system like this cannot be successful without various computer technologies as well as intellectual efforts. For that reason, we'll apply for the patent before it is formally released.</p>
可利用之產業 及 可開發之產品	



<p><b>技術特點</b></p>	<p>本資料庫在技術上重要的突破有二：</p> <p>(一) IPA 符號顯示: 在 IE 瀏覽器版本 4 以上，使用者已安裝 IPA 相關字型前提下，在判斷欲顯示 IPA 符號的 CODE 下加入 &lt;FONT FACE&gt; 的 TAG，來指定使用相關字型顯示。</p> <p>(二) IPA 符號輸入: 製作一 IPA 符號的表格，使用者經由點選相關符號產生要查詢符號的編碼，以此方式進行資料庫檢索</p>
<p><b>推廣及運用的價值</b></p>	<p>(一) 用來查詢漢字古音，以賞析古詩詞的聲韻之美，或訓詁解讀古典文獻。</p> <p>(二) 用來查詢現代方言的音讀。</p> <p>(三) 用來從事古今字音演變的研究。本資料庫的交叉檢索功能，可提供學術界做各種有關古今音的研究。</p>

- 1. 每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送 貴單位研發成果推廣單位（如技術移轉中心）。
- 2. 本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。
- 3. 本表若不敷使用，請自行影印使用。