

漢語相關語言間音韻對應規律的統計方法

駱嘉鵬

台灣清雲科技大學講師・國立新竹教育大學台語所博士生

luo@cyu.edu.tw・g9438102@mail.nhcue.edu.tw

【提要】

語音的演變，具有一定的規律和方向。同族系的兩種語言，經過千百年的演變，也許會讓彼此的差異大到無法互相通話的地步，但其音韻之間，必然存在或多或少的對應規律。至於原本不同系屬的語言，也有可能因為接觸的頻繁，拉近彼此的距離，而形成另一種形式的音韻對應規律。

雙語之間音韻對應規律的統計分析，是比較語言學的重要工作項目。精確的對應規律統計，對於分析雙語音韻異同、擬測共同祖語音韻、考求本字和同源詞、語言(方言)系屬的劃分、語言(方言)調查字表的製定、第二語言學習，甚至相關語文翻譯的研究，都具有決定性的參考價值。

然而，對應規律的統計，是一項煩瑣而又複雜的工作。例如多音字的配對選擇、音韻層次的分析……等等，任何一個環節的疏失，都有可能降低統計結果的精確性，進而影響它的價值。

本文擬針對漢語相關語言間的音韻對應規律的統計課題，從字音資料庫的建立，到統計軟體的設計，以至於統計結果的應用，提出個人一得之愚，以分饗同道，並就教於方家。

【關鍵詞】 字音對應規律、比較語言學、計算機語言學、資料庫檔案管理、
樞紐分析表

1.前言

雙語之間音韻對應規律的統計分析，是比較語言學的重要工作項目。精確的對應規律統計，對於分析雙語音韻異同、擬測共同祖語音韻、考求本字和同源詞、語言（方言）系屬的劃分、語言（方言）調查字表的製定、第二語言學習，甚至相關語文翻譯的研究，都具有決定性的參考價值。

根據不同的研究目的，我們可以進行兩種不同方式的對應關係的統計。無論如何，這是一件煩瑣而又複雜的工作。借助資訊科技的運用，可以提高我們的工作效率，減少人為的誤差。本文擬針對這個課題，分享個人一些經驗，並就教於方家。

2.對應規律的由來

語音的發展，具有一定的規律和方向。同族系的兩種語言，經過千百年的演變，也許會讓彼此的差異大到彼此無法互相通話的地步，但其音韻之間，必然存在或多或少的對應規律。至於原本不同系屬的語言，也有可能因為接觸的頻繁，拉近彼此的距離，而形成另一種形式的音韻對應規律。前者例如歷代韻書和現代漢語方言，後者例如現代中日、中韓、中越漢字音，以至於這些相關語言所衍生出來的各種方言之間，都可以察覺整齊對應的痕跡。儘管德國語言學家格里姆 (Jacob Grimm) 指出「任何規律都免不了例外」，青年語法學派則認為「任何例外都有其規律」，兩派觀點雖然有所差異，但對於規律和例外現象的掌握，卻都是不容迴避的首要課題。

3.兩種處理模式

雙語之間的音韻對應規律，是由個別字音對應彙總統計得來的。由於漢字在各語言（方言）間普遍存在一字多音的現象，關於多音字對應方式的研究目的和方法的不同，就會導致統計結果的差異。

就語言習得方法而言，要透過雙語之間的對應規律來學習相關的第二語言，往往不遑深究歷史來源問題。我們可以把雙語之間同一個字的所有字音一一對應，以免有所遺漏。例如「重」字，現代漢語普通話有 chong2、zhong4 兩個音，閩南話有 tiong5、tiong7、ting5、tang7 四個音¹，兩者相乘，便可以得到八組對應關係。

這種窮舉式的對應關係，可以避免因為主觀的誤判而造成的疏漏，對於第二語言的習得而言，既可以掌握規律，又能夠列舉例外，自然具有一定的參考價值。然而，就歷史語言研究的觀點來看，卻又顯得含糊籠統、條理不清。我們都知道「重」字普通話的 chong2 音相對應的閩南話讀音是 tiong5 跟 ting5，而和普通

¹ 參見普通話閩南方言詞典(1982)，p102、p1019，閩南話注音採用台灣語言拼音方案。

話 zhong4 音相對應的閩南話讀音是 tiong7 跟 tang7，如此就可以把上列八組對應關係簡化為四組，這是在進行歷史語言比較時必須要做的一項篩選工作。經過這樣的篩選之後所得出來的對應規律才能夠準確的解釋語言的發展脈絡。

4. 漢字音韻資料庫的建立

爲了探討漢語相關語言之間的字音對應關係，個人嘗試自行開發一套漢字音韻管理系統，目前共收錄 22,692 筆紀錄。其中包括上古漢語字音紀錄 6,791 筆²，中古漢語 (廣韻) 字音紀錄 15,170 筆，現代漢語普通話字音 14,468 筆，台灣閩南語字音 11,916 筆，台灣四縣客家話字音 11,712 筆。除了提供使用者各種不同的檢索方式之外，目前正在積極改寫系統功能，以期符合音韻對應規律研究的種種需求。個人規劃的要點如下：

4.1 開放性的資料庫

漢語相關語言研究的學者遍及全球，各擅所長，每位學者所掌握的語料和研究旨趣也有所差異。本系統除了提供個人所建置的語料庫之外，將開放使用者自行匯入個人建立的各種語料，包括歷代韻書、各種漢語方言、少數民族語言，以至於域外漢語借音——如韓漢音、漢越音和日本吳音、漢音、唐音等，以便使用者進行比較研究。

4.2 系統自動比對、建立關聯

對於使用者所匯入的語料，除了將其與系統內建的各音系語料並列以供查詢之外，並能依據使用者需求，提供任何兩種語料之間字音自動配對的功能。如前所述，窮舉式的對應和兼顧源流發展的合理對應各有不同的參考價值。本系統將以窮舉式的對應爲基礎，先將雙語之間相同漢字的所有字音一一對應，其次再進行合理對應的分析，推薦系統認定的合理對應紀錄，並附上標記，最後再交由使用者核對確認。

4.3 以樞紐分析模式提供統計結果

任何一個音系的字音資料匯入系統，並經過系統輔助，完成了雙語字音對應資料表之後，便可以進一步使用系統提供的樞紐分析功能，輕易的統計出雙語之間各式各樣的音韻對應關係表。

5. 本系統統計法的特點

5.1 兼具兩種對應統計的功能

使用者可以隨時根據需求，進行窮舉式的音韻對應規律統計，以做為第二語

² 此以「字+音」合併計算，凡一字多音則以多筆計算，但同字同音則合計爲一筆。

言教學的輔助教材。也可以進行合理對應規律統計，以做為歷史比較或其他相關研究的參考依據。

5.2 可避免重複紀錄造成的誤差

一份多語並列的字音對應資料表固然可以一目了然，對於了解個別漢字在各音系之中的音讀有其便利性，例如：

漢字	普通話	閩南話	客家話	反切
長	zhang3	tiang2	cong2	知丈
長	zhang3	tiong2	cong2	知丈
長	zhang3	tiunn2	cong2	知丈
長	zhang4	tiong7		直亮
長	chang2	chiang5	chong5	直良
長	chang2	tiang5	chong5	直良
長	chang2	tio5	chong5	直良
長	chang2	tiong5	chong5	直良
長	chang2	tng5	chong5	直良

上表臚列「長」字在中古和現代普通話、閩南話、客家話的讀音，但若要以此資料來進行上述任何兩個音系的各種對應規律的統計，顯然有所不足。除非我們把所有可能組合相乘並列，否則極有可能有所遺漏。相乘並列的對應資料表除了導致資料的重複激增之外，在進行任何雙語之間的對應規律統計時，又勢必將因為其他音系的需要而區分的紀錄合併計算，否則一樣不能得到精確的統計數據。

本系統在個別音系語料匯入之後，僅依據使用者特定需要製做雙語之間的字音對應關係表。經過系統半自動化的輔助校定之後，便能夠精確掌握該雙語之間的各種對應關係，既不會有掛漏之失，也不致造成相同資料重複累計的錯誤。

5.3 可隨機設定統計條件

語音的演變，往往有其特定的條件；對應規律的分析，自然也不能忽略音節內部音韻成分相互制約的功能。在特定條件限制下所整理出來的對應規律，往往更能說明語音演變的原理。因此，本系統將提供使用者依據各種音韻成分及其特徵來設定統計條件的功能，以便掌握其在語言研究上的價值。

5.4 可隨意設定統計項目

如所周知，語音的演變，並不是孤立的現象，具有相同音韻特徵的一組音位，往往會朝著相同的方向發展。因此，對應規律的統計自然不能僅以臚列個別音位的對應現象為滿足。本系統在使用者精確輸入各漢字在各音系的完整音讀之後，將分析其音韻的組合成分，明辨各音位的辨義徵性，並以此做為使用者設計統計

項目和限定統計條件的依據。例如聲母的音讀、發音部位、發音方式、清濁、送氣徵性，介音的類型、開合、洪細，韻腹的高低、前後，韻尾的發音部位及其方式，鼻化的有無，以至於聲調的陰陽、四聲，都可以做為統計的項目，並且可以任意組合，以探索對應規律隱含的意義。

5.5 可隨時檢視所有字例

本系統將採用自行開發的樞紐分析工具，除了方便使用者任意設定統計條件及項目之外，一如 Excel 的樞紐分析表，在各類統計數字格上點擊，即可查閱符合該項對應關係的所有字例。³這不僅是統計結果完成後運用語料的一項利器，在資料建置階段，也能夠提醒使用者注意到一些特殊的對應現象，減少人工手誤的機率。

6. 結語

語言學是社會學科之中最適合用電腦輔助處理資料的一門科學；在語言學之中，又以音韻材料的處理，最具有客觀性。任何語言都具有大量的音韻素材，這些語料錯綜複雜又具有一定的條理。利用電腦科技，協助我們抽離出具有研究意義的素材，分析它的條理，解釋它的規律，這是我們必須要走的一條康莊大道。本文謹就個人目前研究的半成品，提出未來繼續加強改進的方案，敬請各位專家學者不吝賜教。

【參考文獻】

- 何大安，1988，《規律與方向：變遷中的音韻結構》，台灣台北：中央研究院歷史語言研究所。
- 徐通鏘，1991，《歷史語言學》，北京：商務印書館。
- 海柳文，2005.8，〈漢語方言民族語言語音材料處理軟件設計〉，《廣西民族學院學報》，第 11 卷第 3 期，p60-64。
- 袁焱，2001.5，《語言接觸與語言演變》，北京：民族出版社。
- 陳秀泉，2005.12，〈音韻學在柳州方言區普通話教學中的運用〉，《柳州師專學報》，第 20 卷第 4 期，p52-54。
- 陳保亞，1996.7，《語言接觸與語言聯盟》，北京：語文出版社。
- 廈門大學中國語言研究所漢語方言研究室主編，1982.10，《普通話閩南方言詞典》，香港：三聯書店香港分店。
- 駱嘉鵬，2005.10，《漢語音韻史教學網站》，網址：<http://w3.cyu.edu.tw/luo/941/>。

³ 參見駱嘉鵬(2005)，《漢語音韻史教學網站》<http://w3.cyu.edu.tw/luo/941/>，〈課程單元六〉中各種字音對應關係表的呈現方式。

Statistical methods on the corresponding patterns of the sound and sound system among Chinese dialects

Jiapeng, LUO

Lecturer of Ching Yun University in Taiwan

Ph D Student of National Hsinchu University of Education

luo@cyu.edu.tw, g9438102@mail.nhcue.edu.tw

Abstract:

Phonetic sound and sound system evolves toward certain direction. The differences of two dialects of the same race might be so huge after hundreds of thousands of years that mutual understanding is not possible. Yet, more or less, there must be certain corresponding patterns among the sound and sound system of them. Certain kinds of corresponding patterns of sound and sound system will form even among languages of different races, since continuous exposure with each other always leads to closer relationship.

Statistical analyses of the corresponding patterns of the sound and sound system of two languages have been a main focus in comparative linguistics. Precise statistics of the corresponding patterns is of crucial importance and values in analyzing the similarities and differences of the sound and sound system of two languages; examining the mutual origin sound and sound system; discerning the original word and homologous words; determining the language (dialect) clan; establishing language (dialect) word table; second language acquisition, and even translation researches of the related languages.

Yet, statistics of the corresponding patterns is a complicated task. Even slightest neglect in a link will decrease the accuracy and the values of the research, for example, matching polyphones, and analyzing of levels of sound and sound system, etc.

In this paper, I seek to share my years of experience and study in the statistical corresponding patterns of the sound and sound system among Chinese dialects. My discussion starts with setting up character-sound database, expands to planning software and ends at applying the statistical results. Your opinion is welcomed.

Keyword: corresponding patterns of characters and sounds, comparative linguistics, computational linguistics, database management , pivot table