



《詩經》叶韻音在《廣韻》異讀中的計算機檢測

2019/03/05 顧國林

[提要] 叶韻音在韻書中反映為一種異讀，通過計算《詩經》全部韻腳字、非韻腳字的《廣韻》異讀率差異，可檢測出叶韻音。計算機經過 10 萬次編程模擬，對誤差進行量化，從中可判斷出叶韻音的數量及可信範圍。

[關鍵字] 早期叶韻 詩經韻腳 廣韻異讀層次

一、合理的懷疑

《廣韻》是唐、宋兩代學者不斷加字加注的成果，異讀面貌複雜，是否收錄（或不慎收錄）叶韻音呢？下文採用統計法去探索這個問題，並給出較嚴格的估算答案。

叶韻，又作協韻、諧韻，指改讀韻腳，使古詩重新押韻，以宋代朱熹最有名。兩晉南北朝時，叶韻已萌芽，晉代徐邈、南梁沈重已有為韻腳字注殊讀音的現象，徐邈稱“取韻音”，沈重稱“協句”。陸德明對叶韻持謹慎態度，但其著作《經典釋文》也收錄不少叶韻音，據張文軒《試析陸德明的叶韻》整理^①，共收詩經叶韻音 25 例、23 字，這些用例陸書都明確標明“協韻”“協句”等叶韻標記。

叶韻有“恢復古音”的追求，但由於古人的認識局限，不少叶韻音是臆造音，不是真實存在的語言。目前上古韻部劃分，主要依靠詩經韻腳，如果《廣韻》存在叶韻，會給古音研究帶來麻煩。

二、處理思路

叶韻音很隱蔽，如果原著無注明，很難發現，統計方法的思路如下：

- (1) 叶韻如果被韻書記錄，那麼將表現為一種異讀（雖然異讀的來源是多樣的）。
- (2) 如果沒有叶韻，詩經韻腳字和非韻腳字的《廣韻》異讀率不應該有明顯差異，如果前者大於後者，那麼應該懷疑存在叶韻音。
- (3) 統計方法都有誤差，必須衡量這種誤差，結論才更加可信。

下面從這三個原則出發，計算詩經叶韻音在《廣韻》中的數量規模。

三、統計和誤差衡量

詩經 311 篇，如果合併重複字，正文用字 2795 個（統計不計標題），其中韻腳用字 1811 個，不作韻腳的字是 984 個。韻腳全部依據王力《詩經韻讀》^②，電子版《廣韻》字頭參考“小學堂”^③和“Poem 開源《廣韻》”制定。

^① 張文軒《試析陸德明的叶韻》載於《蘭州大學學報》1983 年第一期。

^② 統計不計標題，統計文本參考高亨《詩經今注》（正體字形），不見於《廣韻》的字形用全等異體字代替，使得全部字都能在《廣韻》中找到語音。

^③ <http://xiaoxue.iis.sinica.edu.tw>

(1) 結果值

1811 個韻腳字中，《廣韻》一共記錄 2486 個字音，異讀率為： $2486/1811=137.3\%$ ^④。

984 個非韻腳字中，《廣韻》一共記錄 1344 個字音，異讀率為： $1344/984=136.6\%$ 。

計算異讀率差值為： $137.3\%-136.6\%=0.7\%$ 。這個數值不算大，相當於 1811 個韻腳裏有 13 個叶韻音，它能否作為叶韻存在的依據？主要看統計方法的誤差。

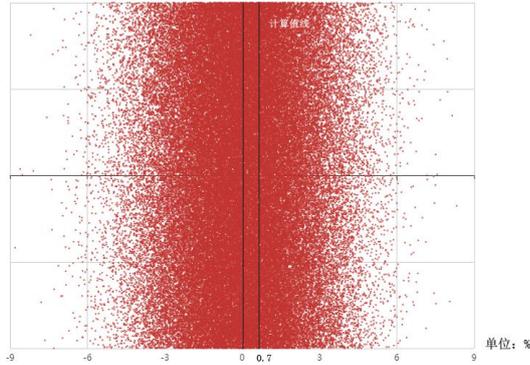


圖 1: 叶韻結果值及統計誤差分佈 (原始散點)

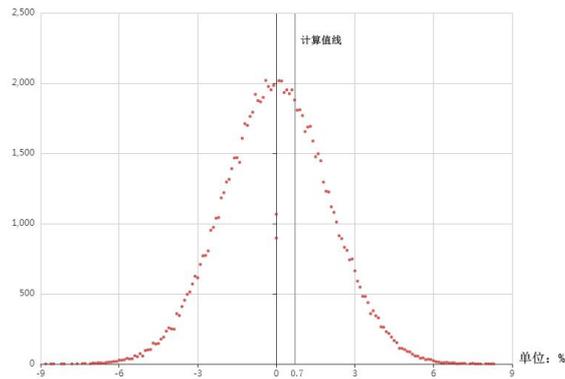


圖 2: 叶韻結果值及統計誤差分佈 (對散點統計)

(2) 誤差衡量

統計方法一般都存在隨機誤差，樣品數量越少誤差越大。比如投硬幣正反面的概率，如果只投了 3 次，那麼很可能算得 66.7%，和真實的 50% 相差甚遠，如果投 100 次，那麼會很接近 50%，如果投 1000 次，那麼將更精確。對於 1811 個韻腳、984 個非韻腳的樣品數量來說，它們的異讀率差值隨機誤差有多大呢？下面用電腦模擬並計算出這個誤差。

在詩經全部 2795 字中，用程式隨機抓出 1811 字和 984 字^⑤，形成 A 組和 B 組，分別求取它們的異讀率，並計算差值。由於是隨機抓取，從理論上來說 A 組和 B 組的異讀率是一樣的，即理論差值是零。但由於隨機誤差的存在，實際並不為零，這個隨機誤差是我們需要的。

用電腦重複以上操作 10 萬次，得到 10 萬個差值，並處理如下：

1、將 10 萬個異讀率差值 (A 組異讀率減 B 組異讀率) 繪製於圖 1，這樣可以看到誤差的分布。橫軸是異讀率差值，縱軸是對橫軸的空間拉開，以便於觀察，縱軸數值無實際意義。可以看到，異讀率差值的均值是 0，但有誤差分布，核心部分落在[-3%,3%]範圍內，最密集範圍是[-1%,1%]。

^④ 这是近似数，精确到小数点后一位，下文同类计算亦用近似数，不再一一指出。

^⑤ 计算程序采用 PHP 语言，用 mt_rand() 函数对诗经 2795 字进行随机抓取。



2、對圖 1 進行統計得到圖 2，具體做法是：將橫軸以 0.1%為單位分段，得到 180 個區間段，將 10 萬個數據放入 180 區間段中，縱軸是落入該區間段的數據數量，誤差分佈可一目了然。這是一個常見的正態分佈，計算得標準差 σ 為 2.4%（意味著數量約為 68.3%的數據落入此範圍）。

（3）結論

由此看，統計方法的誤差可以被計算（標準差 σ 為 2.4%的正態分佈），計算的結果值也是明確的（0.7%），兩者結合，可以推斷出“真值”的範圍。打個比方，一把尺子的誤差是 2mm，用它丈量的結果是 100mm，那麼真實值很可能在 98mm-102mm 之間。

根據正態分佈的特點，數據落在一個標準差（2.4%）內的概率是 68.3%，落在兩個標準差（4.8%）內的概率是 95.5%，落在三個標準差內（7.2%）內的概率是 99.7%，因此可以得到結論見表 1。

結論	概率（可信度）	界定依據
詩經叶韻音在《廣韻》中存 0-54 字	68.3%	$\sigma + 0.7\%$
詩經叶韻音在《廣韻》中存 0-99 字	95.5%	$2\sigma + 0.7\%$
詩經叶韻音在《廣韻》中存 0-143 字	99.7%	$3\sigma + 0.7\%$

表 1：詩經叶韻音存於《廣韻》的數量和可信範圍

（4）解讀

統計可得詩經叶韻音在《廣韻》中的數量規模，雖然不能精確到具體數值，但可確定數量範圍和概率，從中可以得到兩個信息：

1、“《廣韻》不存在詩經叶韻音”和“《廣韻》存在少量詩經叶韻音”都是有可能的，但“《廣韻》存在大量詩經叶韻音”是不可能的，叶韻發生率小於 3.1%的概率是 68.3%。

2、如果《廣韻》存在詩經叶韻音，那麼其數量規模由表 1 限定，它指出了叶韻音的上限。

由此可知，《廣韻》記音基本真實可靠，叶韻音數量很有限（如果有的話），不足以對韻部研究造成較大干擾。

四、結語

最後總結一下用計算機方法檢測和量化叶韻的要點：首先需求取所有韻腳字和非韻腳字的異讀率差值（結果值），再求取統計方法的誤差範圍，如果結果值（正值）遠大於誤差範圍，那麼可以判定叶韻音的存在，如果結果值在誤差範圍內，雖然不能判斷叶韻是否存在，但仍然可以計算出叶韻數量的上限。

其中，誤差範圍的計算涉及到《廣韻》異讀的組成（雙音、三音、四音等），是比較複雜的計算，用傳統方法很難求取，而用編程方法模擬則可輕鬆獲得，這是現代電腦處理誤差的優勢。

參考文獻：

- 祝敏徹 張文軒 《論早期叶韻》 載於《蘭州大學學報》 1982 年 第 1 期
 張文軒 《論叶韻和讀破的關係》 載於《蘭州大學學報》 1984 年 第 4 期
 張民權 《王夫之詩經叶韻辨述評》 載於《語言研究》 2000 年 第 1 期
 駱瑞鶴 《毛詩叶韻補音研究》 2005 年 武漢大學博士論文 編號 10486
 高亨 《詩經今注》2009 年 上海古籍出版社
 北宋官修 《宋本廣韻》 2008 年 江蘇教育出版社出版
 王力 《詩經韻讀·楚辭韻讀》 2012 年 中國人民大學出版社出版