



《廣韻》異讀近音假借層的探測方法

2018/12/20 顧國林 (2022/06/17 修改)

[撮要] 我層寫過一篇《異讀率和同音字數負相關現象所反映的廣韻多音字成因》，證明廣韻異讀的主體是近音假借，這篇將數據和簡帛假借材料直接對比，增加這個結論的可信度。

一、基本原理

用 A 字的音記錄 B 詞，如果兩者不同音，稱近音假借。近音假借會形成異讀，流傳在中古韻書裡。近音假借可以通過統計方法探測，使用這條原理：

同音字數量越多的音節，近音假借發生概率越低，反之越高。

用統計語言可表述為「近音假借發生率與音節字數呈負相關」。

這條原理的依據如下：

正常心理，古人使用記音字，同音字優先，在沒有合適同音字的情況下，才會姑且用近音字，使用近音字對閱讀不利。因此造成：在同音字越多的音節裏，越容易找到合適的字記音，用近音字的概率就越低。

異讀的其他成因沒有這個特點，下面來檢查一下各種可能性：

- (1) 權威語影響和文白讀。比如讀書音的進入造成一字兩音，如普通話剝、血等。
- (2) 方言接觸。比如兩個相鄰語言因居民頻繁來往而發生混雜，某些字詞兼兩個語言的發音。
- (3) 同源詞、詞義派生。比如「種」在種子、種地裏兩個音，屬同源詞分化。
- (4) 形態和詞根。比如動詞的使動用一個音，被動用另一個音，但字是同一個。
- (5) 同形字。兩個字碰巧一樣，造成一字兩音，如吳語片（片義）和古文字片（床義）。
- (6) 訓讀，讀為別的同義詞，僅是詞義關聯，語音毫無關係。
- (7) 錯讀，比如讀半邊，如普通話的墟，又如聲符的錯誤類推，如普通話的礦。
- (8) 其他數量較少的音變，如避諱、前後字感染音變、語法詞弱化、小稱兒化等。
- (9) 近音假借。

這些異讀成因在活語言裏都真實存在，但(1) - (8)對音節字數並不敏感，不適用以上的原理，只有(9)近音假借能匹配這個原理，因此這個探測方法具有精確的針對性。

文章分兩部分：(1)用實例證明這條原理是正確的，使用簡帛通假字材料；(2)對比《廣韻》異讀數據和簡帛通假字數據的異同，以判斷《廣韻》異讀的主體是否來自近音假借。

對於數據處理，判斷相關性（第一組增大時，對應第二組是否也相應增大或相應減少）一般採用皮爾森相關性係數，下文也採用這個方法，並同時使用分佈圖和表格數據，增加說服力。

二、簡帛材料的驗證

簡帛文字含有大量假借字和異文，為驗證以上原理提供了試驗場，使用白於藍《簡帛古書通假字大系》（以下簡稱《大系》）。用電腦對《大系》假借字的統計結果是：前後字（通假字和被通假字）都見於《廣韻》的有對 6455 對，這 6455 對裏屬於同音假借的有 1872 對，屬於近音假借的有 4583 對。

把 4583 對近音假借提取出來，選擇借入字（如女-汝中的汝），放到《廣韻》的 3867 個小韻^①中，觀察近音假借發生率和小韻字數的關係（小韻是《廣韻》同音字組，每個小韻可視為中古的一個音節ⁱ），圖 1 是它們的分佈情況：

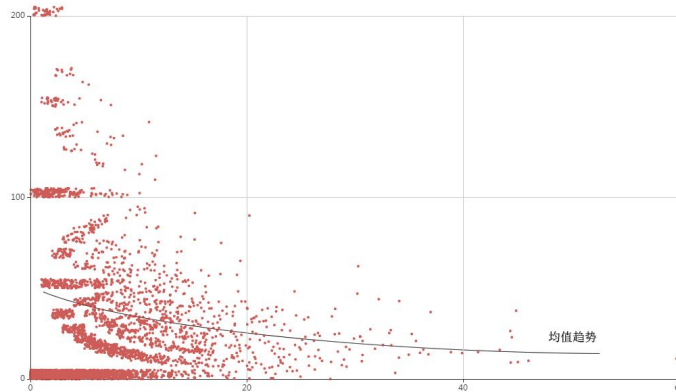


圖 1: 簡帛近音假借發生率對小韻字數的下降式分佈^②

圖 1 中的每個點，對應《廣韻》的一個小韻，橫軸表示小韻字數（1-87 個），縱軸表示該小韻「近音假借發生率」，計算方法是：該小韻內近音假借字數/該小韻的字數。比如某個小韻，簡帛 4583 個近音假借字落入該小韻的有 7 個，小韻字數是 20 個，那麼「近音假借發生率」是 $7/20=35\%$ ，形成圖 1 的一個點。《大系》部分借入字有多個借出字，以上數字 4583 是記次的字數，所以「近音假借發生率」可以超過 100%。

從圖 1 可以看出，隨著小韻字數的增加，「近音假借發生率」的均值呈明顯下降趨勢，從而驗證上面的原理是正確的。為了更清楚展示小韻字數和「近音假借發生率」的關係，對相同字數的小韻進行合併，得到表 1：

小韻字數	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
總轄字數 A	537	1018	1242	1440	1565	1614	1785	1592	1476	1300	1298	1284	1066	854	1035
總近音假借數 B	201	370	413	408	434	421	443	352	341	286	362	282	256	183	209
平均近音假借發生率 B/A(%)	37.4	36.3	33.2	28.3	27.7	26.1	24.8	22.1	23.1	22	27.9	21.9	24	21.4	20.2
小韻字數	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
總轄字數 A	848	595	450	361	440	357	352	322	240	200	182	189	112	174	180
總近音假借數 B	154	109	100	62	108	61	62	32	44	35	39	16	22	33	47
平均近音假借發生率 B/A(%)	18.2	18.3	22.2	17.2	24.5	17.1	17.6	9.9	18.3	17.5	21.4	8.5	19.6	19	26.1
小韻字數	31	32	33	34	35	36	37	38	40	42	44	45	47	60	87
總轄字數 A	93	64	132	102	70	108	74	38	40	42	88	180	47	60	87
總近音假借數 B	16	9	31	13	22	23	9	5	4	5	15	30	3	5	4
平均近音假借發生率 B/A(%)	17.2	14.1	23.5	12.8	31.4	21.3	12.2	13.2	10	11.9	17.1	16.7	6.4	8.3	4.6

表 1: 簡帛近音假借發生率對小韻字數的下降式分佈（合併同字數小韻）

表 1 以第五列數據舉例，小韻的字數是 5，它們總共轄字 1565 個，簡帛近音假借的總字次是 434，計算所得「平均近音假借發生率」是 27.7%。「平均近音假借發生率」是對字數相同的小韻的平均統計，可以避免樣品少造成的如圖 1 那樣的波動，使結論更具有說服力。下面，對表 1 中小韻總字數超過 200 字^③的數據繪製成分布圖 2：

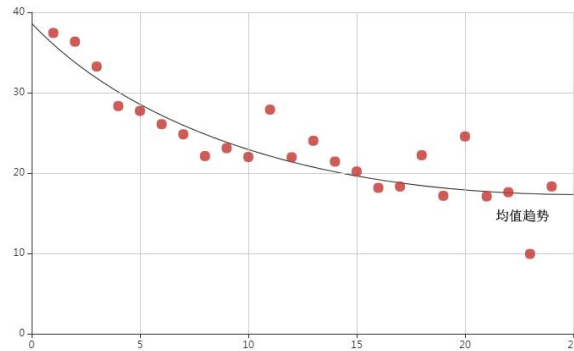


圖 2：簡帛近音假借發生率對小韻字數的下降式分佈（合併同字數小韻）

圖 2 橫軸是小韻字數，縱軸是平均「平均近音假借發生率」，從中可以看出明顯的下降趨勢。對圖 2 數據計算的皮爾森相關性數值^④是：-0.856，顯示為「極強負相關」。

以上從簡帛假借字證實文章開頭提出的原理：同音字越多的音節，近音假借的發生概率越低，反之越高。

有人可能會有疑問：《廣韻》是中古音，和假借字量產期的上古音是不同的，結論是否有效？其實不必當心，《廣韻》小韻是中古最小同音字組，它繼承自上古最小同音字組，雖然音值發生了變化，期間亦有兼併，但大小規模仍有繼續性。

三、《廣韻》異讀的檢測

在簡帛通假字對原理的驗證後，把這條原理應用於《廣韻》，探測《廣韻》異讀是否有這個現象。對《廣韻》的每個小韻單獨計算異讀率，得到 3867 組數據，它們對小韻字數的分佈見圖 3：

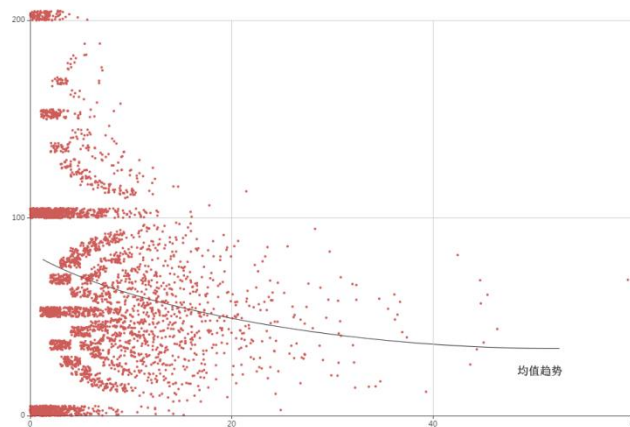


圖 3：《廣韻》異讀率對小韻字數的下降式分佈^⑤

圖 3 每個點對應一個小韻，橫軸是小韻字數（1-87 個），縱軸是該小韻的異讀率，異讀率的計算為：該小韻中所有字的異讀總次數/小韻字數，由於異讀記次數，所以異讀率可以超過 100%。比如某小韻有 20 字，這 20 個字在《廣韻》裏有 12 次異讀，那麼異讀率是：12/20=60%。

從分佈圖上可以看出，隨著小韻字數的增加，異讀率逐漸下降，可證明「近音假借」這個層次的存在。同樣，為了更好地觀察趨勢，把小韻合併，得到表 2：

小韻字數	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
總轄字數 A	537	1018	1242	1440	1565	1614	1785	1592	1476	1300	1298	1284	1066	854	1035
總異讀數 B	440	758	866	962	1026	986	1073	938	844	769	709	679	527	431	497
平均異讀率 B/A(%)	81.9	74.5	69.7	66.8	65.6	61.1	60.1	58.9	57.2	59.2	54.6	52.9	49.4	50.5	48
小韻字數	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
總轄字數 A	848	595	450	361	440	357	352	322	240	200	182	189	112	174	180
總異讀數 B	399	261	217	187	222	171	139	149	118	77	78	78	45	100	66
平均異讀率 B/A(%)	47.1	43.9	48.2	51.8	50.5	47.9	39.5	46.3	49.2	38.5	42.9	41.3	40.2	57.5	36.7
小韻字數	31	32	33	34	35	36	37	38	40	42	44	45	47	60	87
總轄字數 A	93	64	132	102	70	108	74	38	40	42	88	180	47	60	87
總異讀數 B	33	24	58	29	23	53	39	15	4	33	25	97	19	40	20
平均異讀率 B/A(%)	35.5	37.5	43.9	28.4	32.9	49.1	52.7	39.5	10	78.6	28.4	53.9	40.4	66.7	23

表 2：《廣韻》異讀率對小韻字數的下降式分佈（合併同字數小韻）

表 2 以第 10 列數據舉例，小韻字數是 10 個，它們總共轄字 1300 個，總共存在異讀 769 次，計算平均異讀率是 769/1300=59.2%。平均異讀率是對字數相同的小韻的平均統計，可以避免樣品少造成的波動，使結論更具有說服力。

下面，選取表中小韻總轄字數超過 200 字^⑥的組繪製分佈圖 4：

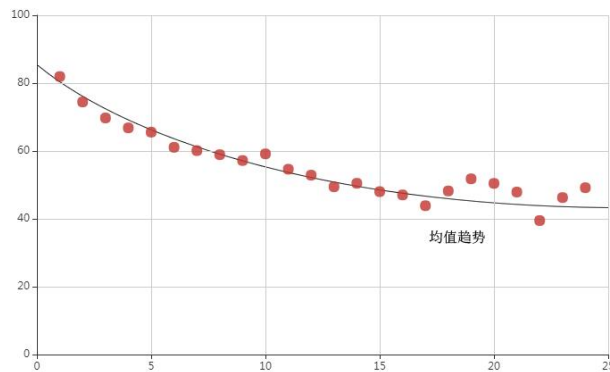


圖 4：《廣韻》異讀率對小韻字數的下降式分佈（合併同字數小韻）

圖 4 橫軸是小韻字數，縱軸是平均異讀率，從中可以看出明顯的下降趨勢。對圖 4 數據計算的皮爾森相關性數值是：-0.896，顯示為「極強負相關」，甚至比簡帛通假字還要顯著、規則。

至此，結論已經明確：《廣韻》異讀率對小韻字數的分佈和簡帛近音假借字一樣，呈明顯的下降式分佈，證明《廣韻》異讀存在一個明顯的「近音假借」層次。



四、結論

簡帛通假字具有很高的可信度，《廣韻》異讀與之進行統計比較，可知「近音假借」層次的存在。這個層次的規模有多大？從分佈圖 2、圖 4 可以看出，《廣韻》異讀率對小韻字數的變化斜率，和簡帛「近音假借發生率」對小韻字數的變化斜率非常接近，前者從約 80% 下降到約 40%，下降約一半，後者從約 36% 下降到約 18%，下降也約一半，程度相當，這可以推證：近音假借是《廣韻》異讀的主體，而不是一小部分。如果近音假借只占《廣韻》異讀的一小部分，那麼圖 4 的「均值趨勢」的傾斜會變緩，如果沒有近音假借層，「均值趨勢」會變成水平直線。

雖然不反對《廣韻》異讀可能有多種來源層次，如方言、古文白讀、詞義派生、形態、訓讀、錯讀、避諱等，但主體層（數量最多的部分）來自近音假借是可以確信的。

參考資料：

- 《宋本廣韻》 江蘇教育出版社 2008 年
李方桂 1980 年 《上古音研究》 商務印書館
鄭張尚芳 2008 年 《上古音系》 上海教育出版社
潘悟雲 2000 《漢語歷史音韻學》 上海教育出版社
周法高等 1975 年 《金文詁林》 香港中文大學
張世超等 1996 年 《金文形義通解》 中文出版社
白于藍 2017 年 《簡帛古書通假字大系》 福建人民出版社

- ①小韻數包含《廣韻》增修重出小韻，一般位於該韻最後，但數量較少，在統計效應中可以忽略。
- ②在字數很少時，縱軸數值只有固定幾個，會造成大量點重疊在一起，為了更好地展示每個點，圖中對縱軸數據進行了加[0,10]的亂數偏移，對橫軸數據進行了加[0,1]的亂數偏移，這樣就能把重疊在一起的點在小範圍內擴散開，便於觀看。該圖週邊還有極少量數據點，因佔用版面太大沒有列出。
- ③以 200 字為界捨棄小樣品，因為樣品小，隨機波動大，價值低。
- ④皮爾聲相關性計算的結果值範圍是[-1,1]，正數表示 A 列數據增大時 B 列也增大，負數則反之，絕對值是 1 時，表示所有數據點在一條直線上，是理想的線性關係，絕對值為 0 時，表示數據雜亂無規律，可視為沒有相關性，常用的數值參考如下：0.8-1.0（極強相關）、0.6-0.8（強相關）、0.4-0.6（中等相關）、0.2-0.4（相關）、0.0-0.2（弱相或無相關），更精確的應用還需考慮標本數量。
- ⑤數據進行了微小的亂數偏移處理，見圖 1 的說明，該圖週邊還有極少量數據點，因佔用版面太大而沒有列出。
- ⑥以 200 字為界捨棄小樣品，因為樣品小，隨機波動大，價值低。