



形聲字「近音諧聲」的發生機理

顧國林 2019-02-11

【撮要】統計《廣韻》得到結果：當音節字數是 35、25、20、15、10、5 時，諧聲不準（聲符音 ≠ 字音）的平均發生率是 25.7%、38.3%、48.4%、56.6%、66.1%，逐次增高。音節的字數越少，諧聲越容易不準。這證明了：1、聲符不準是造字時姑且使用近音聲符造成；2、形聲字在首次製造、第一次使用時，聲符已經不準。（小站之前曾討論漢字的假借、異讀也如此，它們的實質是相同的）

（一）poem 《聲形析微表》

有些作品對探索知識的影響極大，2014 年，poem 在網絡上共享發佈原創的《聲形析微表》，含廣韻全部字的聲符、義符分析，從此諧聲探索進入電腦統計時代。這篇小文的統計使用《聲形析微表》。

（二）原理

形聲字近音諧聲，指聲符和字音近似、不同音。

那你怎麼知道在上古造字之初，它們就不同音了呢？要證明它困惑頗多，至少有兩個揮之不去的陰影：一是方言融合，如果中古語文是由先秦多種方言、多國文字融合而成，那麼會產生字音和聲符不合的面貌；二是文字更革、漢字活用，也會造成聲符和字音的偏離，比如它即蛇，但後來前字轉為其他用途，不再念蛇了。

下文統計《廣韻》失諧率在各小韻中的分佈，提出「失諧率和同音字數呈顯著負相關」這個觀察結果，這是上古近音諧聲造字的證據。

（1）何為「失諧率和同音字數呈負相關」？

舉個例子，在一切條件相同的情況下，《廣韻》A 小韻轄 20 字（即有 20 個同音字），B 小韻轄 40 字（即有 40 個同音字），如果前者平均有 10 字失諧，那麼後者平均有幾字失諧呢？根據統計，後者失諧不是平行的 20 字，而是平均在 14 字左右，即字數越多，失諧率越低，且下降明顯。對此解釋是：在同音字越多的音節裏，越容易找到合適的同音聲符，失諧率越低；反之，在同音字越少的音節裏，越難找到合適的同音聲符，造字者採用近似音聲符的可能性越大，失諧率越高。

（2）失諧率、同音字數的約定

失諧率：失諧字數占總字數的比例，比如《廣韻》某小韻轄形聲字 100 個，其中 50 字和聲符不同音（包括聲調），那麼失諧率為 50%。同音字數：指失諧的統計範圍，下文指《廣韻》每個小韻的形聲字總數。如無特殊說明，下文的字數都是指形聲字字數。

（3）相關性方法

兩組數據，是否存在相關性（第一組增大時，對應第二組是否也相應增大或相應減少），統計上一般採用皮爾森相關性。由於相關性不等同因果性，對因果的定性應排除其他更直接可靠的原因，在第（四）節使用語言學和文字學常理去排除。

（三）失諧的分佈

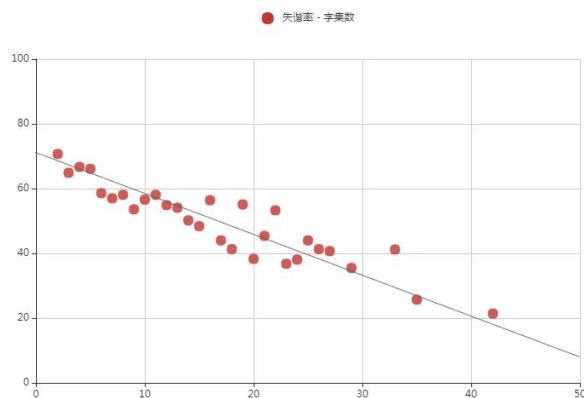


《廣韻》有 3867 個小韻（含重修者添加字時所增的重出小韻，數量不多），每個小韻轄 1-83 字不等，最多的一個小韻轄 83 字，以 5-20 字為常見。

如果該小韻只有一個字，那麼它極可能是失諧的，統計排除了單字小韻，對小韻的要求是至少有 2 個形聲字，這樣的小韻一共有 3194 個，分別對著 3194 個小韻計算各自失諧率。然後，把字數相同的小韻統計到一起。比如擁有 12 字的小韻數量有 96 個，把它們的字數加起來，一共得到 1152 字，把它們的失諧字數也加起來，一共得到 633 字，計算平均失諧率為： $633/1152=54.9\%$ 。也就是說：擁有 12 字的小韻平均失諧率是 54.9%。

依次類推，統計其他字數的小韻，共得 41 組數據，其中總字數大於 100 個的有 30 組，列為有效數據（見表 2）。除去小樣本是為了提高統計精度，因為小樣本的結論價值較低，比如在「硬幣正面朝上 50% 概率」的實驗中，如果只投了兩三次，所得結論是不可靠的。

小韻內字數	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
小韻數量	535	425	357	293	276	247	212	133	106	111	96	71	66	51	35
形聲字總數 A	1070	1275	1428	1465	1656	1729	1696	1197	1060	1221	1152	923	924	765	560
失諧總數 B	756	828	952	968	971	985	986	642	600	710	633	499	464	370	316
失諧率 B/A (%)	70.7	64.9	66.7	66.1	58.6	57	58.1	53.6	56.6	58.1	54.9	54.1	50.2	48.4	56.4
小韻內字數	17	18	19	20	21	22	23	24	25	26	27	29	33	35	42
小韻數量	27	21	13	20	13	11	11	13	8	4	7	7	5	3	3
形聲字總數 A	459	378	247	400	273	242	253	312	200	104	189	203	165	105	126
失諧總數 B	202	156	136	153	124	129	93	119	88	43	77	72	68	27	27
失諧率 B/A (%)	44	41.3	55.1	38.3	45.4	53.3	36.8	38.1	44	41.3	40.7	35.5	41.2	25.7	21.4



表：《廣韻》小韻內字數（橫軸）和平均失諧率（縱軸）的關係

最終結果顯示（圖 3），隨著字數的增加，失諧率相應降低，規律非常明顯，比如：

- (1) 字數為 5 的全部小韻，平均失諧率是 66.1%。
- (2) 字數為 10 的全部小韻，平均失諧率是 56.6%。
- (3) 字數為 15 的全部小韻，平均失諧率是 48.4%。
- (4) 字數為 20 的全部小韻，平均失諧率是 38.3%。
- (5) 字數為 35 的全部小韻，平均失諧率是 25.7%。

可以看到，它們的數值逐級下降，從最高的約 70% 一直下降是最低的約 20%，幅度是非常明顯的。對這 30 組數據進行相關性計算，皮爾森相關性數值為：-0.932，屬於「極強負相關」，可以認為兩者有非常顯著的線性下降關係。

(四) 繼續探討

相關性和因果性不能完全劃等號，有必要對造成形聲字失諧的原因考察一番，排除存在其他更直接的可能。造成失諧的原因可能有：

- ① 聲符借出。比如，莠是遘（遇）初文，古今字，莠後來借為它用，不再和遘同音，所有用莠作聲符的字，聲源會被誤判。觀察古文字，這類失諧有相當數量。
- ② 形聲字借出。形聲字借為它用，聲音改變，但用字未變，造成失諧。這類失諧數量比①要少得多。
- ③ 先借用再加形。本是借用字，後加形變成形聲字，這是形聲字的重要來源。借用指借音，如豆本為器皿，借音用為植物的果實，加形指添加形符，如簡帛裏陰、陽常寫𠂔、易，後來添加形符卩。桶字常寫甬，木旁是後加的，桶和甬的失諧，早期實際上是近音借用。
- ④ 近音諧聲。造字時使用非同音聲符，可分兩種：①其字有祖，即表義字添加聲符，形成形聲字，如鳳字，最初畫一只鳥，後來加聲符凡；②原生形聲字，使用義符和聲符造的全新字。
- ⑤ 詞義引申。微小改變語音派生新詞，即同源詞，但用字未變，造成失諧。這類失諧有一定數量，但容易被高估。
- ⑥ 形態和詞根。因形態造成一字多音，造字時用 A 音，後世只存 B 音，造成失諧（潘悟雲，2000）。
- ⑦ 方言影響、文白讀。A 方言造的字到 B 方言裏變成失諧，或字音念 A 方言、聲符念 B 方言造成失諧。
- ⑧ 文字訛變。非聲符訛為聲符，或聲符訛為非聲符，造成後人誤判。如盤的聲符是凡，因形似而訛為舟，據學者研究，凡是盤的象形初文，後來借為它用，再造盤字。
- ⑨ 訓讀。讀為另一個同義詞，訓讀在近古較少，在早期是否流行則不明。
- ⑩ 同形字。兩字碰巧一樣，如吳語有片字（意思是片），古文字亦有片字（意思是床），兩者音義毫無關係，用吳語片字解牆、將、妝等字，就會失諧。漢字同形字數量不多，這類失諧較少。
- ⑪ 其他原因，如：避諱、詞間感染音變、語法詞弱化等，此類原因造成的失諧不會太多。

以上這些原因中，失諧率對同音字數敏感的有①②③④。⑤⑥（詞義引申、形態和詞根）可能會對詞性有選擇性，⑦（方言影響、文白讀）可能會對特定的聲、韻、調有選擇性，但都不會對同音字數敏感。⑧⑨⑩不僅不會對同音字數敏感，而且數量較少，可以忽略。

①②③④中的④是上文提到的近音諧聲，①②③都屬於文字借用（借音記詞），借用也存在同音優先原則，也會對同音字字數敏感，值得重點分析：

- ① 聲符借出。聲符借為它用，本音失傳，借用音傳世，有一定的數量，在常用獨體字裏比例很高。比如午，本是是舂米之杵的象形字，早期的午即杵，杵和午不能算失諧，但午後世借為「上午、下午」之午，杵和午就失諧了。基本聲符的借出，發生時間很早，在經典上古音時代（春秋詩經時代），基本聲符借出已成常態，其中很大比例是近音借用，在《說文》時代基本定型，表現在許慎對基本聲符字的釋義釋音會失誤，如韋、夬、凡、覃、各、犖、莠、吏、午等等。
- ② 形聲字借出。數量不多，原因是形聲字大量產生後，近音借用已經衰落。

- ③ 先借用再加形。這類字最隱蔽，文獻所能提供的例子很有限，需要出土古文字檢驗。在金文簡帛時代，漢字作為表音文字使用，假借字產能很高，這和現代普通人對漢字的看法不同。有一點是可以確信的：《廣韻》裏的某個形聲字，很難確定它在誕生的時候就是形聲字。

由此可見，以上三個因素中，①和②的數量可估，它們不是失諧字的主要部分，③的數量不可估，需要未來考古檢驗，本文假設了兩種結果：

(1) 中古形聲字大部分是原生的，那麼本文不需作改動，結論不變。

(2) 中古形聲字大部分是已有的假借字加形生成，那麼本文需要修改一下標題，把「近音諧聲」改為「近音諧聲和近音借用」，那麼結論依然有效。

最後說明一下《廣韻》同音字如何反映上古。上文統計失諧率的單位是小韻，這是中古音節的最小顆粒。用最小顆粒觀察失諧率有個好處，它可以反映上古音位。從上古音到中古音，雖然聲類、韻部都發生了分化、合併、重組等變化，但最小顆粒只需用合併去解釋就可以了。即上古音節的最小顆粒，通過合併形成中古音節的最小顆粒。在這個過程中，顆粒的規模也被繼承，即上古大顆粒（大指同音字多），到中古仍然是大顆粒，上古小顆粒（小指同音字少），到中古有可能是小顆粒，也有可能並為大顆粒。從統計上來說，它們仍有繼承效應。雖有變異，以《廣韻》小韻為最小字集單元，大體也可反映上古的字集框架。

（五）結語

漢字是原生文字，原生文字帶有草創階段的不足，早期漢字聲符的作用是提示發音，做不到精確表音，音節性聲符無法駕馭早期漢字音的複雜度，失諧是註定的。在最早期（甲骨文、早期金文）的結構裏，除了聲符、義符，還有較多區別符（加橫、加豎、加口等，如亼-今、兀-元等），區別符是對近音聲符的補償制度，告訴閱讀者：「這是聲符，但和原字音義有差異，請你注意！」聲符+區別符可理解為古人講的轉註，聲符+義符可理解為形聲。區別符後來被加形取代，形聲字成為主流。漢字這種不太完美的近似表音，由於區別符和形符的補償作用，一直延續了下來。聲符能大量精確表音，要等到漢字音節大量簡化的近古才實現。

總的來看，中古形聲字失諧有多種原因造成，但其主體、絕大部分的數量是上古近音諧聲造字。

參考文獻：

- 《聲形析微表》《廣韻字音表》poem 網絡發佈
《宋本廣韻》 2008 江蘇教育出版社
《上古音系》 鄭張尚芳 2008 上海教育出版社
《漢語歷史音韻學》 潘悟雲 2000 上海教育出版社
《金文形義通解》 張世超等 1996 中文出版社
《簡帛古書通假字大系》 白於藍 2017 福建人民出版社