



上古漢語有無形態的檢測方法和結果

2019/03/19 顧國林

[提要] 上古漢語到底是不是形態型語言？文章提出檢測上古漢語形態的方法：（1）找到能追蹤動詞和非動詞的有效標記；（2）觀察兩者語音變體的存世痕跡。文章選擇才、艸兩個部件做為動詞、名詞的標記（分別為 554 字和 938 字），統計結果顯示，前者的異讀率、形聲字失諧率、簡帛近音假借率比後者高 38.9%、66.2%、16.8%。這意味著，動詞的語音變體比名詞多，從而推斷上古漢語普遍存在形態。

[關鍵詞] 形態和諧聲 形態和異讀 古漢語構詞構形 上古音構擬

一、動詞的語音變體

形態有多種理解，下文指通過變換輔音、韻尾、聲調等元素或添加前綴、後綴等方式來實現詞的不同功能^①，特別是動詞。形態在藏緬語言裏較發達，上古漢語是否大量存在形態呢？類推並不是很科學的做法。

上古漢語的形態，在異讀上可能存有一些蛛絲馬跡，比如常用例子“折”，有清、濁兩種中古音，在今天語言裡仍然感受到它的辨義功能。但這樣的例子並不太多，原因是，在字書、韻書形成的中古時代，漢語形態已經不顯。

早期漢語是否是形態型語言，已很難用例證法獲知，下文採用統計法探測，材料為異讀、諧聲、簡帛通假字三種，方法為對比動詞、名詞對語音的敏感差異。

要證明形態，需設計一些探針，下文以“詞性對語音的敏感差異”來識別形態，使用艸部件字（938 字）和才部件字（554 字）為動詞、名詞的探針，得到以下結果：

- （一）《廣韻》異讀率：才字（動詞）比艸字（名詞）高 38.9%。
- （二）《集韻》異讀率：才字（動詞）比艸字（名詞）高 68.9%。
- （三）《廣韻》形聲字失諧率：才字（動詞）比艸字（名詞）高 66.2%。
- （四）簡帛通假字近音假借率：才字（動詞）比艸字（名詞）高 16.8%。

四組數據顯示：**動詞的語音活力顯著高於名詞、上古動詞擁有更多的語音變體。**

如果動詞、名詞出現對語音的敏感差異，那麼很值得懷疑是形態現象。在形態豐富的語言中，動詞的語音變體通常比名詞多。形態在漢語中消失後，動詞的語音變體仍然保留在異讀中，表現為較高的異讀率，同時給諧聲、假借帶來影響，表現為高失諧率、高近音假借率（從中古音角度觀察）。過去，潘悟雲提出形態現象在辭彙平面有選擇性^②，文章把它當檢測方法來使用。

二、檢測方法

下文假設並不知道是否存在形態，根據這條原則，通過檢測異讀率、諧聲失諧率、近音假借的差異，探索形態是否存在，以及估測其規模。檢測方法有兩部分組成：

- （一）找到識別詞性的有效標記，即如何確定某字在上古是動詞還是名詞。

^① 有學者提出的漢語具有通過詞的組合、虛詞應用等來實現語法功能的“廣義形態”，本文不涉及這類形態。

^② 潘悟雲，2000 年，《漢語歷史音韻學》123 頁，上海教育出版社。



(二) 量化不同詞性的語音敏感差異，觀察它們的差異。

對於(一)，普通方法是直接採用古文獻的詞性信息，但漢字詞性經常動詞、名詞兼類，一字多性，難以取捨。統計採用了“部件追蹤法”，用“扌”追蹤動詞，用“艸”追蹤名詞，主要考慮以下兩方面：

(a) 詞性指示性好。艸和扌不作聲符，只作形符，艸字大部分是和植物相關的名詞（如艾、芥、芋、蔥等），扌字大部分是動詞（打、撫、推、掩等）。

(b) 數量可觀，具備統計價值。《廣韻》數量排名前十的部件為：水、艸、木、扌、人、心、口、金、言、蟲。艸排第二位，共 938 字，扌排第四位，共 554 字。

對文字部件艸和扌的識別，取自《廣韻》字形，不細究文字訛變。比如萬字，古文字原是蠍子象形，上面的艸是蠍子的鉗足訛變而成，和植物沒有關係。這類訛變兩者都有一些，但數量很有限，不影響統計結論。

對於(二)，統計採用《廣韻》《集韻》異讀率、諧聲失諧率、戰國秦漢簡帛近音假借率三項，覆蓋古音研究的三大材料。

三、統計結果

統計艸字和扌字所做的準備有：Unicode 擴展 A-F 區全部已編碼漢字的資料庫和輸入法，共 8.9 萬字^③，其中艸字有 3513 個，扌字有 2282 個。完備的字音材料包括：《廣韻》全字表（含全部 2.53 萬電子字頭）、《集韻》全字表（含全部 5.38 萬電子字頭）、白於藍《簡帛古書通假字大系》全字表（含全部 9655 對電子字頭）^④、《廣韻》全部字諧聲聲符表^⑤。以下數據計算通過編程實現。

(一) 異讀率對比

異讀即一字多音，《廣韻》《集韻》記錄了豐富的異讀，後者數量更多：

(1) 艸部在《廣韻》中共有 938 字，一共記錄 1222 個字音，異讀率為：30.3%。

(2) 扌部在《廣韻》中共有 554 字，一共記錄了 787 個字音，異讀率為：42.1%。

結論：在《廣韻》中，扌部的異讀率比艸部高出：38.9%（見表 1）。

(1) 艸部在《集韻》中共有 1475 字，一共記錄了 2584 個字音，異讀率為：75.2%。

(2) 扌部在《集韻》中共有 827 字，一共記錄了 1877 個字音，異讀率為：127%。

結論：在《集韻》中，扌部的異讀率比艸部高出：68.9%（見表 2）。

部件	廣韻音數 (A)	廣韻字數 (B)	廣韻異讀率 (A-B) / A	異讀率比值 (Y-X) / X
艸	1222	938	30.3% (X)	+ 38.9%
扌	787	554	42.1% (Y)	

表 1: 艸部和扌部《廣韻》異讀率比較

部件	集韻音數 (A)	集韻字數 (B)	集韻異讀率 (A-B) / A	異讀率比值 (Y-X) / X
艸	2584	1475	75.2% (X)	+ 68.9%
扌	1877	827	127.0% (Y)	

^③ 網址：www.guguolin.com/font.php。

^④ 文章中所用的《廣韻》《集韻》電子字頭和音韻資訊參考“漢字全息資源應用系統”和“小學堂”公開資源，《簡帛古書通假字大系》電子字頭參考“引得市”公開資源。

^⑤ 聲符分析是一項很專門的工作，聲符表在沈兼士《廣韻聲系》的基礎上，參考餘乃永《上古音系研究》、鄭張尚芳《上古音系》諧聲分析修訂，爭議字的聲符參考於省吾《甲骨文字詁林》、張世超《金文形義通解》、白於藍《簡帛古書通假字大系》三部著作，吸收古文字學主流考釋成果。



表 2: 艸部和才部《集韻》異讀率比較

(二) 諧聲失諧率對比

失諧指形聲字和聲符不同音（包括聲調），比如“許”字與其聲符“午”不同音，失諧是站在中古音的觀察，對比字音是《廣韻》音。如果該字是異讀，或其聲符是異讀，只要其中一個組合同音，那麼就不算失諧。

(1) 艸部在《廣韻》中共有 938 字，其中形聲字 906 個，這 906 個形聲字中，失諧 359 個，失諧率為： $359/906=39.6\%$ ，即平均每 100 個艸部形聲字，有 39.6 個字和聲符不同音。

(2) 才部在《廣韻》中共有 554 字，其中形聲字 550 個，這 550 個形聲字中，失諧 362 個，失諧率為： $362/550=65.8\%$ ，即平均每 100 個才部形聲字，有 65.8 個字和聲符不同音。

結論：才部的失諧率比艸部高出：66.2%（見表 3）。

部件	廣韻字數	諧聲字數 (A)	失諧字數 (B)	合諧字數	失諧率 B/A	失諧率比值 (Y-X) / X
艸	938	906	359	547	39.6% (X)	+ 66.2%
才	554	550	362	188	65.8% (Y)	

表 3: 艸部和才部諧聲失諧率比較

(三) 近音假借率對比

用 A 字的音寫 B 詞，稱為假借，如古文用“女”字寫第二人稱“汝”，簡帛用“昏”字寫問答的“問”。如果 A 和 B 同音，那麼屬於同音假借，如果 A 和 B 不同音，那麼屬於近音假借。古音很遠，一般不會發生假借。判斷同音假借還是近音假借，也是以中古音為依據，並不完全等同造字時的本貌。假借在先秦很發達，是形聲字大量出現前最重要的造字方法。

統計採用的假借字來自白於藍《簡帛古書通假字大系》（以下簡稱《簡帛》），是戰國秦漢出土簡帛假借字最詳盡集成，共計 9655 對，其中借字、被借字都見於《廣韻》者有 6455 對，出土文字保留了未經篡改的原始面貌。

(1) 艸部在《簡帛》6455 對有效假借字中，共出現 516 對（不管是通假字，還是被通假字，只要出現艸字都計入，下同），其中同音假借 196 對，近音假借 320 對，近音假借率為： $320/516=62.0\%$ 。

(2) 才部在《簡帛》6455 對有效假借字中，共出現 362 對，其中同音假借 100 對，近音假借 262 對，近音假借率為： $262/362=72.4\%$ 。

結論：才字的近音假借率比艸字高出：16.8%（見表 4）。

部件	見於簡帛假借字數量 (A)	近音假借數量 (B)	同音假借數量	近音假借率 B/A	近音假借率比值 (Y-X) / X
艸	516	320	196	62.0% (X)	+ 16.8%
才	362	262	100	72.4% (Y)	

表 4: 艸部和才部簡帛近音假借率比較

四、統計解讀

統計顯示，動詞具有更多異讀，在形聲字中更容易發生失諧，在簡帛假借中更容易不同音，它們為何會這樣？對三者的影響解釋如下：

(一) 更多異讀。這是動詞構詞功能的痕跡，在形態語言中，動詞的功能變體，遠多於名詞。

(二) 更多形聲字失諧。潘悟雲曾用形態解釋部分諧聲失諧現象，認為形態會造成一字多音，造



字時用 A 音，而後世只流傳 B 音，造成失諧[®]，上文的數據支持這個觀點（但失諧現象並不全是形態原因）。

（三）更多近音假借。原理同失諧，也是審音性的，漢語的形態淡出後，一部分異讀被遺忘。

異讀、失諧、近音假借有相似性，它們都有多種來源，形態是其中一個來源。這個來自形態的層次可以由詞性檢測出來，它的特點是：詞性不同，異讀率、失諧率、近音假借率也不同。

五、結語

以上結論的分析，尚未考慮漢字早期假借的影響，先秦假借很多是近音，會增加異讀。如果艸、才字兩者假借數量不平衡，會影響到對形態的判斷。

通過考察，早期艸字的假借範圍稍大於才字，對上文的結論有利。比如，艸在上古較多假借為姓氏、地名等專名，而才很少假借給非動詞。《廣韻》中艸字借為姓氏的有“蔣、莉、蕓、苴、莞、蓋、蔞、薰、蔓、萬、芒、艾、蘆、葉”等，以上假借常帶來多音，而才借為姓氏則很少。艸字多為植物，古人喜聞樂見，借音用為姓氏、地名等專名的接受度高。才字多為動作，假借為非動詞的接受度低。

使用統計方法，可以在不分析形態例子的實例下，探索形態，彌補例證法的不足。從 38.9%、66.2%、16.8% 三個很顯著的數值看，早期漢語動詞存在較多形態變化、功能變體是可信的，可為“上古漢語較多存在形態”假說提供證據。

參考文獻：

- 白於藍 《簡帛古書假借字大系》，福建人民出版社 2017 年
李方桂 《上古音研究》，商務印書館 1980 年
鄭張尚芳 《上古音系》，上海教育出版社 2008 年
於省吾 《甲骨文詁林》，中華書局 1996 年
周法高等 《金文詁林》，香港中文大學 1975 年
張世超等 《金文形義通解》，中文出版社 1996 年
金理新 《上古漢語形態研究》，黃山書社 2006 年
梅祖麟 《上古漢語動詞清濁別義的來源》，《民族語文》2008 年 第 3 期
潘悟雲 《上古漢語使動詞的屈折形式》，《溫州師範學院學報》1991 年 第 2 期
吳安其 《漢藏語的使動和完成體首碼的殘存與同源的動詞詞根》，《民族語文》1997 年 第 6 期

[®] 潘悟雲，2000 年，《漢語歷史音韻學》，上海教育出版社，125 頁。