

中央研究院歷史語言研究所會議論文集之二

中國境內語言

暨語言學

第二輯 歷史語言學

李壬癸
黃居仁 編輯
湯志真

中華民國 臺北

中華民國八十三年五月

中央研究院歷史語言研究所會議論文集之二

中國境內語言暨語言學 第二輯
歷史語言學

CHINESE LANGUAGES AND LINGUISTICS II
HISTORICAL LINGUISTICS

定價：精裝本新臺幣 950 元

(外幣定價按當時美金匯率換算，匯票每張另加匯兌費美金 10 元)

不准翻印

編輯者 李壬癸、黃居仁、湯志真

出版者 中央研究院歷史語言研究所
出版品編輯委員會

發行者 中央研究院歷史語言研究所
臺北市南港區

印刷者 福元印刷事業有限公司
臺北市雅江街 58 號

代售處 臺灣商務印書館
臺北市重慶南路一段 37 號

臺灣學生書局
臺北市和平東路一段 198 號

三民書局
臺北市重慶南路一段 61 號

中華民國八十三年五月出版

ISBN 957-671-231-9 (精裝)

出版說明

中央研究院歷史語言研究所語言組自從民國七十九(1990)年七月起每隔一、二年舉辦一次「中國境內語言暨語言學國際研討會」。第一屆會議的主題是漢語方言，會議論文集業已出版，所收的論文有25篇，外加1篇附錄。

第二屆在1991年8月召開，會議的主題是歷史語言學。報名參加的人數比第一屆還多，如期提出論文摘要的有108人。經過評審小組仔細篩選之後，錄取其中29篇在研討會上發表。全文修訂稿經過正常手續審查通過刊登的卻只有20篇，外加4篇專題演講稿共為24篇。少數論文因特殊緣故改在其他刊物刊登，因此未收入本論文集。這兩屆會議都在中央研究院召開。

這一屆的研討會主題既然是歷史語言學，因此所提有關這個領域的有多篇會議論文、專題演講、座談會也都是環繞著這個主題。論文內容有歷史的，也有現代的；有語音的，也有語法的；絕大多數為漢語，本來也有兩篇討論少數民族語言的論文，但卻改在較專門的期刊上發表了。

中國境內的語言分屬五大語系，而這兩屆的論文都集中在漢藏語系，尤其是漢語。這種現象並不令人感到意外。只是盼望大家今後能多關注少數民族的語言，特別是面臨滅絕危機的語言研究，包括阿爾泰語系、侗傣語系、南亞語系、南島語系的語言研究。

這一輯內容的安排，是把專題演講的4篇放在前面，依通論、語音、語法的次序排列。其餘20篇論文依歷史語音、語法、構詞以及現代語音、語法、其他兩大類的次序排列。此次會議的經過、各篇論文的主要內容、討論的要點等都已見於徐芳敏小姐(1992)的「第二屆中國境內語言暨語言學國際研討會紀要」(《漢學研究通訊》11.1: 14-20, 11.2: 121-127)或蘇宜青小姐(Su 1992, *Journal of Chinese Linguistics* 20.2: 337-342)。

本集論文大都附有摘要，並採用較本文為小的字體。摘要一般附於本文之前，唯摘要與本文不屬同一語言者，則改列在本文之後。為方便讀者的辨認和引

用，在目錄中並舉每一位作者的中文與英文姓名。

籌備會議期間，中研院史語所語言組的同仁都投入相當多的人力、物力、精神，實際負責事務的黃居仁先生和王幸真小姐更是不分晝夜、任勞任怨。會後論文集的編印工作由湯志真小姐負責，也是以極負責認真的態度完成這項艱難而又繁瑣的任務。本組研究人員本身的研究工作因此受到影響，少寫一些論文也在所不惜，只希望出席該次會議和閱讀本論文集的讀者都覺得很有收穫，就是我們最大的報償和願望了。

第三屆會議的主題是詞彙與構詞，於1992年7月初在清華大學召開。第四屆會議的主題是語言類型 (typology)，訂於1994年7月下旬在中央研究院召開，論文集也都將在歷史語言研究所出版，成爲一個系列。

李 壬 癸

1993, 8.14.

目 錄

出版說明	i
語言變化的機理 (Mechanisms of Language Change)	
王士元 (William S. Y. Wang)	1
漢語上古音的元音問題	
丁邦新 (Pang-hsin Ting)	21
關於上古音的四個假設	
白一平 (William H. Baxter III)	41
唐代、宋代共同語的語法和現代方言的語法	
梅祖麟 (Tsu-Lin Mei)	61
Reconstructing Scenarios of Sound Change	
陳淵泉 (Matthew Y. Chen)	99
劉宋時期在漢語音韻史上的地位——兼論音韻史的分期問題	
何大安 (Dah-an Ho)	125
《日本寄語》所反映的明代吳語聲調	
木津祐子 (Kizu Yuko)	139
A Theory of the Bifurcation of the Middle Chinese Voiced Syllable-initial Stops and Affricates into Aspirates and Unaspirates after Devoicing	
徐雲揚 (Eric Zee)	173
Post-stopped Nasals and Lateral Flaps in the Zhongshan (Yue) Dialect: A Study of a Mid-eighteenth Century Sino-Portuguese Glossary	
陳潔雯 (Marjorie K. M. Chan)	203
陝北清澗話人稱代詞和指人名詞語尾〔·mi〕探源	
劉勳寧 (Xunning Liu)	251

Interaction of Syntactic Changes	
何萬順 (One-soon Her)	263
古漢語被動式的發展與演變機制	
魏培泉 (P'ei-chuan Wei)	293
The Grammaticalization of 'Bei' in Chinese	
張洪明 (Hongming Zhang)	321
On the History of Chinese Locative Prepositions	
貝羅貝 (Alain Peyraube)	361
On the Mechanisms and Constraints in Syntactic Change: Evidence from Chinese Dialects	
曹逢甫 (Feng-fu Tsao)	389
Synaesthetic Words in the Southern Min Dialect: Their Semantic Structure and Change	
連金發 (Chinfa Lien)	421
異形語的共存與淘汰——台灣話的輕聲與數量語	
鄭良偉 (Robert L. Cheng)	453
元音與聲調	
遠藤光曉 (Mitsuaki Endō)	487
Tone Sandhi and the Dissimilation of Phonation Types: Reflexes of the Beijing Mandarin Third Tone Sandhi Rule in Northern Chinese Dialects	
Deborah S. Davison	517
A Beat-counting Theory of Mandarin Foot Phrasing	
蕭宇超 (Yuchau E. Hsiao)	555
Verb Movement and Some Syntax-semantics Mismatches in Chinese	
黃正德 (C.-T. James Huang)	587
Wh-words as Polarity Items	
鄭禮珊 (Lisa Lai-Shen Cheng)	615

Conditions on the Distribution of Postverbal Duration and Frequency
Phrases in Chinese Revisited

湯志真 (Chih-Chen Jane Tang) 641

Discourse Explanations for the Choice of *Jiu* and *Cai* in Mandarin
Conversation

劉美君 (Mei-chun Liu) 671

語言變化的機理*

(Mechanisms of Language Change)

王 士 元

加州大學柏克萊校區

我們這個研討會的主題是歷史語言學，好像歷史語言學和共時語言學的界限是二十世紀初以後才分開的，當初畫這條線是有它的道理的，因為在畫這條線之前，根本就沒有共時語言學，所謂語言學就是歷史語言學，所以畫這條線以後，大家就可以比較放心地研究共時的現象。但是現在卻好像有點矯枉過正，歷史語言學似乎不再是語言學的主流了，這不見得是很好的現象，因為有些現象光從共時看好像亂七八糟，沒有什麼道理，但歷史語言學是研究變的，在共時上不見得能看出所以然的可以用歷時來解釋。所以我覺得中國語言學也許可以不把歷時、共時的界限分得那麼死，很多東西互相交流會有比較多的好處。

這個會是談中國境內語言，這方面有一篇文章我覺得很重要，就是李方桂先生 1937 年在 *Chinese Yearbook* 裡一篇「*Languages and Dialects of China*」。二十年前我們創辦中國語言學報時，曾經請李先生讓我們用他的這篇文章作為第一期的第一篇文章，李先生答應了 [E10]。當時我問過李先生，他將中國境內語言歸了大類，怎麼沒講到它們內部的關係如何，怎麼沒有畫樹？李先生回答說因為當時材料不夠豐富，工具、理論方法不夠成熟。這是五十多年前的情形，李先生如此形容那時候的情形，是非常恰當的。中國人講「工欲善其事，必先利其器」，所以要是器具不夠，要想談到其它的，不是罔然了嗎？可是現在已經過了

* 編者按：本文是邀請講席的演講紀錄稿，並經作者修改過。

半個世紀，材料豐富了，工具也很好，很多語言學上的東西有個微型電腦就很夠用了，比較缺的是理論方法的東西。今天要跟各位討論的就是是否我們能建立一些方法、理論，把中國的語言，包括少數民族語言，方言，的語言材料中的內部關係組織起來，看歷史、語言、和文化之間的關係。

要研究一群語言內部的關係，也就是 *subgrouping*，可以從不同的角度來看，比方音韻方面，就有很多人提過。像丁邦新先生 1982 年寫過一篇「漢語方言區分的條件」[C1]，謝信一先生 1973 年的「A New Method of Dialect Subgrouping」[E8]。丁先生的文章主要是以音韻上音變的時間先後作條件，謝先生的文章主要是以詞彙擴散理論，看哪些詞變了，哪些沒變，還有哪些方言中有相同變的詞，或相同不變的詞。最近鄭錦全先生寫了一篇「Quantifying Affinity among Chinese Dialects」[E5]，他所用的方法和我今天要講的就很接近。

我今天所要講的主要是報告一個小小的實驗，一個試探性實驗。我們用很小的一批詞彙，44 個詞目，看看是不是能看出方言中的種種內部關係 [C2]。我們用的是「漢語方言詞匯」裡的一頁詞目，比方「爺爺」或「祖父」在不同的方言裡有不同的稱呼法。在「漢語方言詞匯」裡一共有 44 個這種親屬稱謂。可是我們用的不是方言詞匯裡的方言點，而是用吳語材料。一九二八年趙元任先生寫過一本「現代吳語研究」[C9]，這大概是研究漢語方言的第一部，也是一本經典之作，我們用的就是趙先生所研究過的江蘇南部和浙江的三十三個方言點。中國的大家庭有很多親屬詞彙，其中有許多意料之外的奧妙。比方「爸爸」在有些地方指伯父，「媽媽」有的地方指伯母或祖母。我覺得親屬稱謂也可以是語意研究裡很有研究價值的一個領域。音韻學裡有所謂 *push chain*，和 *pull chain*，我想同樣的現象在語意詞彙系統中也是可以發生的，這是很值得研究的一個部分。

在研究方法方面，首先要做的是一個頻率表。比方「祖父」在蘇州是「阿爹」，宜興是「爺爺」，寧波是「阿爺」，上海是「老爹」，這些都是不同的詞，可以做成下面這樣的表：

	阿爹	爺爺	阿爺	老爹
蘇州	1	0	0	0
宜興	0	1	0	0
寧波	0	0	1	0
上海	0	0	0	1

很多人以前是這麼做的，我覺得這樣的做法不合理，因為其中還有很多相同的地方，如果用這樣的表示方法，那麼內部詞素的相同點就被埋沒了。所以我們第一個決定，就是 **Table of Shared Traits** 不應該用詞，而是用詞素，另外還要注意構詞方式，比方重疊，像宜興就有。做出來就像下面的表：

	阿	爹	爺	老	重疊
蘇州	1	1	0	0	0
宜興	0	0	1	0	1
寧波	1	0	1	0	0
上海	0	1	0	1	0

把這表填起來以後，現在比方有兩個方言 D1 和 D2，某個詞素可能兩個方言都有，可能一個有一個沒有，也可能兩個都沒有，怎麼把這些數據系統化呢？過去一般用相似度 (similarity)，比方雙有為 a，雙無為公 d，一有一無為 b 和 c，有四種可能性，可以用 $S = \frac{ad - bc}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$ 這樣的統計式來算。但是我覺得這個做法有個毛病，也就是如果我們再加入另一個方言，詞素便增加了，同時零也增加了。雖然原來方言的詞素沒變，但會受到這些增加的零的影響，原來關係密的方言會因此關係變薄，這樣就不穩定了。所以應該用一種新的相似性測量 (similarity measure)， $S = a / (a+b+c)$ ，這在統計上常用的，叫 Jaccard coefficient，也就是不該用雙無的那部分數據。語言學家在這方面沒有很深入的研究，但是生物學界因為要給很多不同的生物分類，這方面的學問是很發達的，這個 Jaccard coefficient 和一些數學的概念都可以在 Sneath 和 Sokal (1973) 「Numerical Taxonomy」書中看得到，所以我們要研究語言和語言內部的關係，可以向他們學一些方法。[E15a]

第三點就是做頻率表，過去的做法是把所有的 a，所有的 b 和所有的 c 作同樣的處理，不分詞目間的界限，我覺得這是很不合理的，因為要是「祖父」有三個詞素，「祖母」只有兩個，那麼照那個方法，「祖父」就比「祖母」重要，它的權重 (weighting) 就不平衡了，就傾向詞多的詞目。比方「太陽」在「漢語方言詞彙」裡有 8 個詞素，「風」只有 1 個詞素，「星星」有兩個，那麼「太陽」就有「風」的八倍那麼重要了，我覺得這是不合理的。要是把詞目的界限分開，是會影響統計結果的。過去我們做過一些數學的分析，把詞目的界限分開再算，會發現到以前沒看到的關係。以上談的是我們做頻率表時應注意的三點。

現在就可以看實在的數據了。假設有 n 個方言，就需要跟 (n-1) 個方言對比，但蘇州和上海比等於上海和蘇州比，所以再除以 2，因為有一半是沒用的。所以 33 個方言點，就有 528 個方言和方言中親屬稱謂的關係。參看圖一及 [C2]。

我想很簡單地說說三種不同的分析方法。我想要強調的一點是，我們在如何了解方言間的內部關係上，作一個初步的認識，所以每個方法都有它的優、缺點。在初步階段最好用很多不同的方法，看看每一個方法能提供什麼樣的訊息。

第一個是主分量分析法 (Principal Component Analysis)，這些在 Sneath & Sokal 的書中都講得很清楚。主要是將數據分成多維空間 (multidimensional space)，在我們具體的世界裡，一般是三維的 (three dimensional space)，但數學就沒有這種限制，所以數學就可以把這 528 個點分成三十多個多維空間，那麼在這個多維空間裡，就有很多的點，一共有 528 個點。主分量分析 (Principal Component Analysis) 是一種 curve-fitting，畫一條線，這條線能夠解釋這 528 個點當中有多少方差 (variance)。那麼第一條線自然解釋得最多。畫完這條線，再在一個與第一條線成直角的面上，再畫一條線，把剩下的方差再解釋出來，就是第二個分量 (component)。所以逐步把分量加上，分量研究完以後，方差就完全解釋了。一般作主分量分析時，只要用頭三個主分量就夠了。在圖 x 中第一分量表現為每一個方言點的高度，第二、第三分量分別由 X 和 Y 軸上表現。有趣的是，我們用電腦作主分量分析時，對相關係數表 (Table of Correlation) 作分析，

畫出來的圖與地圖有很密切的關係，真是出乎意料之外的配合得好。所以用主分量分析可以探討地理、空間上的關係。參看圖二及 [C2]。

在 Cavalli-Sforza 和我 1986 年發表在 *Language* 上的一篇文章「*Spatial Distance and Lexical Replacement*」[E2] 裡研究 Micronesia 地區北太平洋上的一些小島，結果也很相似，就是地理上相近的地方，詞彙也相近，而且相近的程度相當的成比例。這和十九世紀的波浪理論 (Wave Theory) 很相近，也就是一個語言的特徵像波浪一樣傳播出去，比較近就多受到影響，早受到影響，比較遠影響就比較少，比較遲。單從 33 個方言點，44 個詞目就能分析出這樣的結果，是原先沒有想到的。第二種分析分法是最短距離樹 (Minimum Spanning Tree)，這也是 *Sneath & Sokal* 書中有的。把 33 個方言點，哪些方言和哪些別的方言關係最近的把它連在一起，跟剛才的主分量分析不一樣，是另外一種統計的方法。圖 X 中有個很有趣的現象，就是偏僻的地方和別的方言點的關係比較少，比方溫州，只有和另外一個方言點有關係；又比方上海，和很多別的地方都有關係，有時候關係拉得很遠，這也是很合理的，因為我們知道上海現在雖然是世界性的大都市，它的歷史不超過一兩百年，主要是鴉片戰爭以後，很多各地來的人在舊上海方言的基礎上融合成新的上海方言，所以上海方言有很多別的方言的成份在內。由於有這些成份在裡頭，它就和很多別的方言點有關係。參看圖三及 [C2]。

再講最後的一個方法，這個方法就是要畫樹。但是我們應該了解，畫樹在邏輯上到底是怎麼一回事。假設有 a、b、c 三個方言，樹有兩種，一種是有根的 (rooted)，一種是無根的 (unrooted)，如果是畫一棵有根的樹，那已經表達很多信息了，所以我們先畫一棵無根樹 (unrooted tree)，那麼 3 個方言一共只能畫一棵無根樹。現在如果再加一個方言，有一個 d，這個 d 可以有三種加上去的方法。所以有四個方言時，我們可以有三棵無根樹。要是再加一個 e，就更複雜了。所以要是我們把原則看出來，原則是相當簡單的，就是所有的單數，3 個語言是一棵樹，4 個語言是 1×3 ，5 個語言是 $1 \times 3 \times 5$ ，再加一個語言，就整個再乘以 7 了，再加一個，再乘以 9，所以樹的數目增加得非常快。大家要是對這有興趣，可以看 Meyers & Wang 1963 「*Tree Representation in Linguistics*」

[E11]，這是我還在俄亥俄州立大學時做的一個研究。所以我們就可以知道這些樹假使語言多的話增加得非常快。

我們爲什麼要算這個呢？第一是我們要知道它的科學價值，它的信息。要是可能性不多的話，它的信息就不多，要是有很多樹，那麼從這裡頭選一棵樹，這棵樹的信息就豐富得多，這是一點。另外一個原因呢，就是當材料越豐富時，計算起來就越繁瑣，一定要用計算機來做，因此要預先知道大概要花多少時間。有了這些概念之後，我們可把 33 個方言點畫成一棵樹，參看圖四。這棵樹和我們剛才談的主分量分析 (Principal Component Analysis) 的聚類基本上是相同的，但是也有些比較有意思的地方。比方江蘇的一些方言點距離都比較近，因爲它們地理上的距離也比較近，在樹上的距離也比較近；但是浙江的部分距離比較遠，交通比較不方便，所以在樹上的距離也比較遠。而且浙江的方言點是不成群的。

這種樹有個好處，就是我們一般在語言學上看到的樹，每一個據點跟它的根的距離是一樣長，這是不合理。因爲要是從根到每一個據點的距離一樣長，等於說每一個語言變遷的速度 (the rate of linguistic change) 是一樣的，我覺得我們現在還沒有資格作這樣一個假定，我們應該先查出是不是這些語言的變化速度是一樣的，這是一個查的方法。就用樹的長短來告訴我們變化的程度，如果根離得遠就長一點，根比較近的話就短一點。可是這種樹還有個好處，就是畫這種樹的時候，可以把樹再換成一個數據矩陣，就是送進一個輸入矩陣，畫出一棵樹，再把它換成一個輸出矩陣。所以一開始送進去的矩陣，是實際觀察到的。你畫出樹之後，把它再換成一個矩陣，那是一個輸出矩陣。把兩者對比一下，有時會發現某些方言群與其它方言群的關係，其實比預測的 (prediction) 要相近得多。在這種情形之下，我們也許可以假設這些「額外相似度」 (additional similarity)，是因爲接觸，借用之類的相互影響造成的。這些都是初步嘗試的階段，所以我想也許這樣做下去，可能將來會有一點希望，可以把縱向傳遞 (vertical transmission (inheritance)) 和橫向傳遞 (horizontal transmission (contact)) 分出來。這是歷史語言學非常重要的一個課題。

所以在這種情形之下，幾年前我做過一個嘗試，那時在台大人類學系講一些

語言的關係，主要用的是 Tsuchida 的數據材料 [E16a]，把台灣一些高山族的語言，畫出這樣一棵樹。（參看圖五）雖然看起來每一個語言離根都一樣遠，但是實際上是不一樣的，它的長度可以在數字上看出來。那麼照我剛才所說的那個方法，我們發現某一組的兩個語言，跟另一組的兩個語言有某種關係。所以這種樹不只可以畫出縱向傳遞，也許將來還可以畫出橫向傳遞，也就是借用。（請參看 E10a 及 E17）

這當然只是一個初步的嘗試，我們需要具體研究，一次一次的對，看是不是有這樣的預測能力。我們希望將來不只是研究幾個很近的方言點，也能夠研究很大的人群。

最近東京大學人類學系有一位齋藤成也 (Saito Noruya) 先生，他把中國人類，尤其是少數民族的血型，做一個很仔細的，很先進的遺傳分析 [E14]。從這裡看出來所謂「漢」，漢方言，漢語，漢人是一個很複雜的概念。比方徐州的漢，跟西藏人、蒙古人在生物化學上更接近，跟北京來的漢人距離反而遠（參看圖六）。況且我們可以看出來秦淮以北有一個聚合 (cluster)，秦淮以南一個聚合，漢人的北方人跟韓國人、西藏人、蒙古人的關係比較密切，而廣西、廣東、海南島的漢人就跟那裡的少數民族，比方苗人、傣人、壯人、侗人的關係比較密切。圖六中的信息也可以畫成一個有根的樹，參看圖七。我希望將來語言學不要老孤立在那裡，歷史語言學是要研究我們的過去，我們的歷史的，而語言是人群的工具，所以我們將來語言的歸類，親屬關係，和人群的歸類一定要連得起來，要互相支持。如果當中有衝突的話，要了解為什麼遺傳學家、人類學家那麼說；為什麼我們語言學家這麼說，當中是什麼過程讓它們變得相互不合了。

我今天就講到這裡，現在稍微重覆一下。第一，我今天講的這個，不是一個成果，只是一個試探性實驗，因為我們只用了 33 個方言點，44 個詞目，真正要做這樣的研究，材料要豐富得多，花的時間也多得多。可是我沒有預料到，單用這 44 個詞目，能夠得到這麼合理的方言內部關係。這讓我想到他們生物學家說的，你只要給他皮膚上的一個細胞，因為細胞裡頭所有基因的信息都在，他差不多可以仿製一個人出來，英文裡這個字叫 clone。雖然語言和生物有很大的不同，

但是從這樣一小部分的詞彙，能夠擠出這麼多信息來，這是出乎我意料之外的。這是很好的預兆，我們可以這麼做下去。

第二點，就是我們不應當只像我們的前輩，只做樹，樹固然很重要，況且做距離樹 (distance tree)，樹枝不同，不同長短的那種樹，特別重要。但是我們也應當探索一些別的方法，比方最短距離樹 (Minimum Spanning Tree)，主分量分析 (Principal Component Analysis)，這些東西有時候不同的分析方法給我們不同的信息。我們在初步的時候，什麼樣的信息，只要是合理的，我們都希望能探討出來。我們用我們知道答案的材料做，可以證實我們的方法。方法做出來，如果每次做出來都是合理的，那麼將來我們用我們的方法來做大問題，做我們不知道答案的材料，就比較有說服力了。

這兩三天聽了很多很有價值，很有內容的東西，學到了很多，也覺得很興奮，因為這一兩百年來，語言學主要是跟著印歐語系語言走的，他們出了什麼理論，出了什麼方法，我們就跟著走。聽了這些文章以後，我覺得也許情形會改變了，也許在不久的將來，中國語言學能夠走到前頭，能夠帶一些路，那麼我們就不辜負老前輩給我們這些光輝的傳統了。

* * * * *

李壬癸：我有兩個問題，一個是人種跟語言之間的關係，我們知道常常語言跟人種是一致的，但有時候也會遇到語言跟人種不一致的地方，一個語系裡有很多不同的種族。剛才王先生提到北方的漢人跟當地的非漢人體質上有很多相近的地方，而華南的漢人就和南方的非漢人有相近的地方。這是一個問題，就是歷史上不一定一個族的人一定要講同一個語言，一直傳下去，往往一個種族會因為周圍優勢族的影響而放棄自己的語言，採用別的語言，因此在歷史上語言不曉得產

生過多少變化。這種現象在南島民族也有，南島民族基本上也是分成南北兩大種族，但語言上都屬於所謂的南島語言，語言之間的親屬關係是非常肯定的，但在赤道以北的北太平洋一帶，和南太平洋的是不同的關係。我沒答案，但聽了王先生演講，想到這是一個問題。

第二個問題是您所得的高山族語言的樹圖，基本上對我來說是很熟悉，沒有什麼意見，但是其中一部分關於魯凱族的，我1977年曾經發表一篇文章，畫的樹圖也是很接近，是根據cognate的多少及sound change幾個大的現象所得的結果。最近鄭錦全指導一位學生用quantifying的方法，重新檢驗我的資料，結果是Mantauran是最早分出來的，跟您的結果有出入。

王士元：謝謝李先生的comment，我想這裡頭有些東西不是那麼三言兩語可以講清楚的，諸位有興趣的話，我們現在所談的這種distance tree，是Fitch & Margoliash 1967「Construction of Phylogenetic Trees」[E6]，這種方法自然有它的好處也有它的壞處，它是在生物界所發明出來的，也許用在語言上會有些不恰當的地方。主要的李先生剛剛講的頭幾句話，我覺得聽起來相當投機。照理說我們的語言是從父母來的，我們的基因也是從我們的父母來的，從最簡單的情形來看，語言和基因的分布應該是相當一致的，但是在很多情形下這當中有出入，所以要是我們能把語言的分布和基因的分布對比一下，能夠把這作了解其出入的初步資料，我覺得是一個能把各種歷史綜合在一起的重要階段。

赤松祐子：像吳語區裡南京基本上屬於北方官話而不是吳語，您怎麼處理這種dialect island？

王士元：要是已經知道我就不需要怎麼處理了。我想一般大都市的情形和鄉下是不同的，大都市一般講官話，標準語，我們研究的33個方言點，除了少數像杭州之外，都是相當標準的吳語。

羅仁地：請問Saito先生畫的這棵樹究竟是憑哪些特徵？

王士元：根據血型，尤其是HLA，請參考Sun, et al (1986)。

曹逢甫：基本上我很贊同您現在改用語素 (morpheme)，而不用詞來計算，這種情形尤其在漢語我想是更合適。不過即使在漢語裡，suffix 和 stem，affix 我們都算它是一個 morpheme，不過是不是 affix 跟 stem carry 同樣的 weight，想必是值得進一步研究的問題。

王士元：我完全同意，況且一個 suffix 和 reduplication 是不是也同樣 weight 也是一個問題。所以我們第一步把它分出來，其餘的問題，詞素和詞素是不是同樣地位還有待進一步研究。

梅祖麟：我對於 Saito 的東西非常有興趣，二十年前和 Jerry Norman 寫過一篇文章，主要是說閩語有個底子，其中有包括一些借詞。在寫那篇文章時我們去查資料，看到一個法國人做的也是血型的研究，他的結論是說廣東人在血型方面跟越南人相近的程度遠超過於廣東人跟華北人相近的程度，華南的中國人跟越南人更類似而跟華北的人不類似。

王士元：Saito 的這篇文章只是講中國的一部分人類。Cavalli-Sforza 在 1994 年會出版一本很大的書，叫「History and Geography of Human Genes」Princeton University Press. [E3]，在這本書裡把中國人的基因上的分析做得精密得多。

(本文於民國八十二年二月十八日通過刊登)