

语言演化的三个尺度

王士元

苏东坡有一首诗我特别喜欢，我也常把诗末的两句改为“不识语言真面目，只缘身在言语中”，因为打从我们在母亲的肚子里起，就已经隐隐约约听得到说话的声音，于是我们一辈子都离不开语言文字，大部分的思想也一定需要靠大脑里的语言组织妥当，然后才能脱口而出或化为文字。要研究语言的真面目，不能不借助语言，但又难免会受到语言的限制，所以想要透彻明了语言的本质可谓难上加难！

古人一定对语言涌现有很大的兴趣，因此存在许多有趣的神话。圣经里的上帝让亚当为各种动物命名，就是一个家喻户晓的故事。

古代的哲学家也喜欢讨论语言涌现的问题，很巧合的是，古希腊的柏拉图和战国时代的荀子，对词的来源竟说过大同小异的话，他们都表示，词是以音表义，但一个词的音与义之间却没有什么固定的关系。因此“水”这样东西，可以叫做water，可以叫做mizu，也可

以叫做acqua等等，只要大家的用法一致，沟通起来就不会有障碍，用荀子的话说就是“名无固宜，约定俗成”。

我们当然无法把时间拨回十多万年前，来看看“约定俗成”是怎么回事，可是却可以用计算机做些简单的模拟实验，以证明一群个体的确能逐步发明一些词，用一些符号来代表事物(王士元 & 柯津云2001; Wang 2006a; Ke et al. 2008; Gong et al. 2010)。

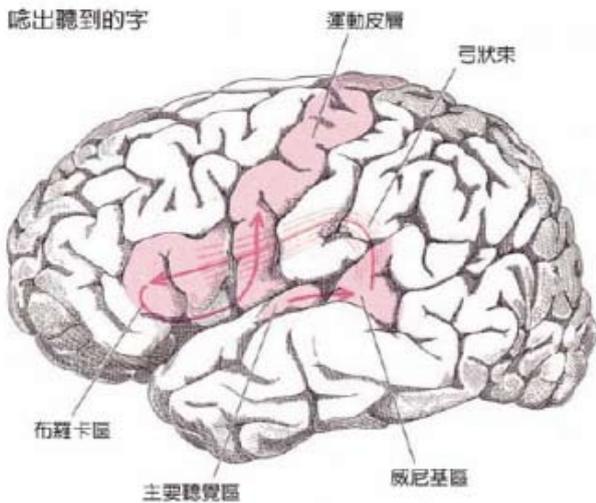
词是语言最核心的部分，要掌握一个语言，我们必须知道怎么把大量的词串起来，短则成句，长则成篇，并且要知道如何在不同的场合里适当地运用这些话。

每个语言都是个复杂适应系统(Complex Adaptive System, Wang 2006b)，当然不是瞬间就可以发明出来的，而是经过几万年的演化，随着人类的文明累积逐步逐步地发展而成，这就是演化语言学的宏观史(macroyhistory)。为了要了解语言涌现这个问题，德国的

科学院在1769年曾经设置一个奖项，得奖者是赫尔德(G. Herder)，他的获奖论文，已经有中文版可供参考(Herder 1772)。

此文有许多颇有见地之处，比方说，他注意到环境给我们的信息，是通过不同的器官传入大脑的。眼睛看到的是颜色、形状，耳朵听到的是不同的频率及节律，鼻子、嘴巴、手、皮肤等，每个器官都有自己独特的编码方式，而大脑却有能力把这些代码综合起来，融合成一个统一的感觉或概念，这种能力叫做“跨感官模式联结”(cross modal association)，是语言涌现不可或缺的基础。赫尔德认为这种能力就是生理学家所提出的sensorium commune，也就是英文常用的common sense一词的来源。

这种联结能力的神经基础，一直到20世纪中叶才被研究清楚(Geschwind 1976)。贾许温德(Geschwind)指出了弓状束(arcuate fasciculus)是大脑皮质语言区的一个组成部分，把颞叶、顶叶、额



图一：弓状束及与语言相关的大脑分区

叶等脑叶联系起来，见图一。贾许温德的研究发表了30多年后，又有科学家用最先进的扩散张量造影(DTI, Diffusion Tensor Imaging)技术，把这些和语言相关的大脑区域绘制出来(Catani et al 2005)。

图二所示的左大脑皮质上，作者标出了以19世纪两位伟大的神经学家布洛卡(Paul Broca)和韦尼克(Karl Wernicke)命名的区域，同时也加上了贾许温德区，因为他所强调的弓状束及角回(angular gyrus)，能把前人一些重要的发现综合起来。

可是18世纪的科学，还远不能解决语言涌现这个难题。虽然有些知名的思想家，如亚当·斯密(Adam Smith)、卢梭(Jean Jacques Rousseau)等，都曾着眼于这个问题，可是“巧妇难为无米之炊”，没有人类演化的知识，是不可能讲出什么好道理的。要解决这个难题，我们必须知道人类跟别的动物有哪些不同，因而导致语言为人类所独有，别的动物不但没有发展出像语言这样变化无穷的沟通工具，而且还根本无法学会我们现有的语言。

人类之所以能具备运用语言的能

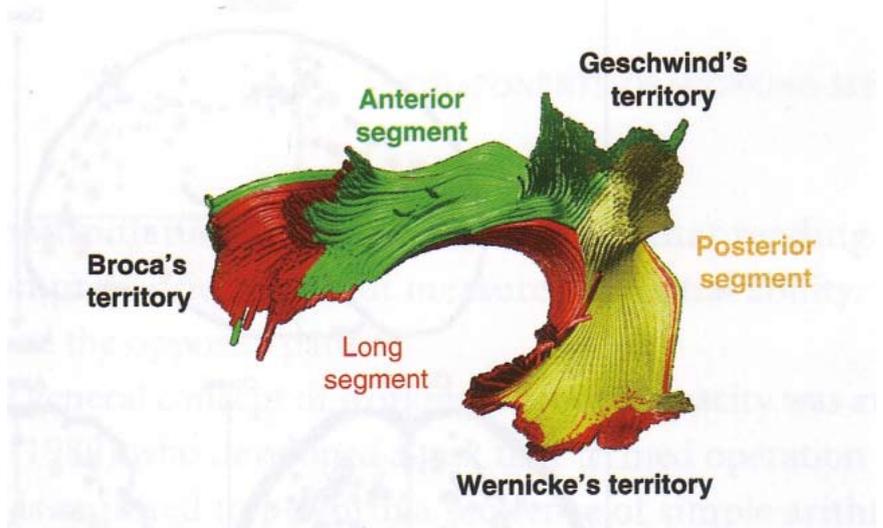
力，一方面要从不同的人类语言里去研究语言究竟是一种什么样的复杂适应系统，一方面要从人体上的演化，尤其是大脑的增长去探究，另一方面还要从人类群体生活的进展去思索，这三条思路前人都还未能很清楚地考虑过，可是无稽之谈的文章却越来越多。所以赫尔德的文章发表100年后，巴黎语言学会在1871年宣布，他们的集会不再接受讨论关于语言涌现的文章。像这样由学会公

开禁止讨论某个话题，倒是学术界里极为罕见的一段趣事。1872年，伦敦的语言学会也采取了同样的行动，明文规定禁止谈论语言涌现。

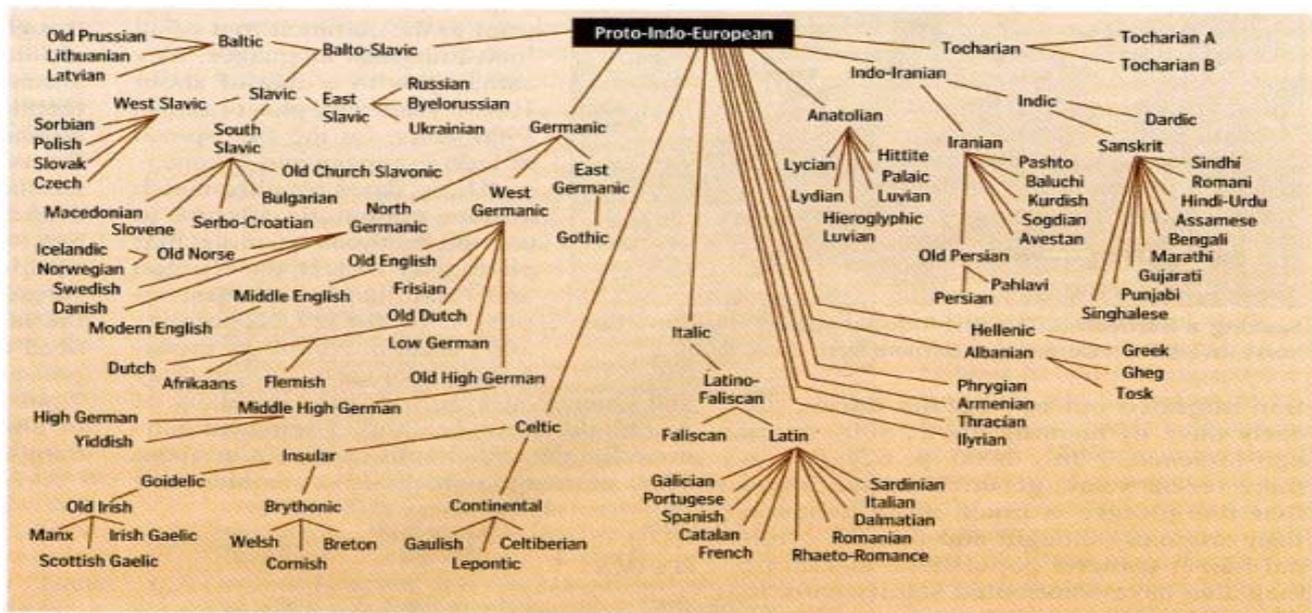
自从禁令颁布后，语言涌现这个话题于是销声匿迹了近100年。一直到1960年，才有美国的霍克特(Charles Hockett)在权威性的《科学美国人》(Scientific American)上发表了一篇引起诸多回响的文章，名叫The Origin of Speech (Hockett 1960)，文章里提出了一套设计特征(design features)，把人类语言和其他动物的沟通系统做了对比。

霍克特的文章发表后，纽约科学院在1976年召开了一场盛大的会议，邀请不同学科里多位大师出席，专门探讨语言涌现与演化的问题(Harnad 1976)。我很荣幸当时也应邀参与了那场盛会。巴黎语言学会于1871年颁布的禁令至此总算解除了。

这两件事后，下一个里程碑应该是1996年起在爱丁堡开办的EVOLANG研讨会，会议名字是Evolution of Language的缩写。EVOLANG每两年集会一次，2012年的第九届会议于日本京都召开，目前



图二：与语言相关的大脑分区



Say it in Indo-European. The 144 languages of this family descend from one ancient mother tongue.

图三：印欧语系的144种语言

尚未在中国举行过。

中国语言学界也已开始关注语言演化的问题。2009年起，在广州就举办了首次演化语言学会议，叫做Conference in Evolutionary Linguistics，简称为CIEL，恰好是法文“天”的意思，正如严复翻译的书《天演论》也用“天”字来解释生物演化之道。之后几届的CIEL会议，每年陆续在不同城市召开，并且特别着重语言研究的多学科性质（王士元2011a）。比方在上海的CIEL-3邀请了许多遗传学家共同与会，在北京刚刚落幕的CIEL-4则请到遗传学家、考古学家、数学家等共襄盛举。CIEL-5将在香港举行，计划结合神经科学的研究成果一同探索语言的演化。

当赫尔德等人在讨论语言涌现的问题时，有一位英国青年到了印度。威廉·琼斯(William Jones)很早就对语言有着浓厚的兴趣，曾经自学拉丁文、古希腊文及古波斯文。他在1786年于加尔各答(Calcutta)做了一场划时代的讲演，替语言研究奠定了科学基础。琼斯的伟

大成就，我曾在2001年给本刊读者（王士元2001）介绍过。

他演讲的基本内容是说，印度的梵文与欧洲的拉丁文、希腊文，以及亚洲西部的波斯文，都有许多相似的地方。这些语言的相似，不可能只出于偶然或巧合，他觉得唯一可信的假设就是，这些语言都应该从古代的同一种语言逐渐演化出来，而这个已经不存在的语言，叫做原始印欧语(Proto-Indo-European)。这是一种突破性的思想，前人所讨论的语言涌现，使用最宽的尺度去研究语言的宏观史，因为没有确实的数据，争论了几百年仍没能取得可靠的成果。而琼斯却缩短了尺度，在宏观史内去研究语言的中观史(mesohistory)，由此创造了一门新的科学，就是语言学。

琼斯开始用科学的比较方法来构拟数千年前的原始印欧语——也就是众多印度及欧洲语言的老祖宗之后，就有一系列顶尖的学者继承了他的基础，持续不断地从事累积性的研究。有位著

名的美国生理学家曾说，远在达尔文之前，语言学家就已经在做了了不起的演化论工作(Thomas 1981)。19世纪的语言学大师施莱歇尔(August Schleicher)，也曾很自豪地说：“I was a Darwinian before Darwin，”一语道出他在达尔文之前已经做了不少语言演化的研究。语言学对科学所做出的最有价值的贡献，可以说正是在这个领域里。

关于语言的中观史，文献相当丰富，这里我不再赘述，读者若参考2004年《科学》(Science 303: 1323)内的树图（见图三），便可以看到144个不同的语言，都源自数千年前同一个原始印欧语(Balter 2004)。其实施莱歇尔早在100多年前，就画过这个语系最早的树图了，虽然当时他以有限的材料只画出9个语言。

这一套构拟语言演化的科学方法，于20世纪由西方传入中国，这其中，赵元任、李方桂、罗常培三位语言学大师合译瑞典汉学家高本汉(Bernhard Karlgren)的作品《中国音韵学研究》

(1939)功不可没。汉语的历史文献非常丰富,方言和民族语言亦多姿多彩,而明末清初中国有许多非常杰出的小学家,如陈第、顾炎武、钱大昕等,研究古书里的音韵及文字,取得了丰硕的成果,这些成就都是中国语言学的一大本钱。可是我们对中观史的研究,却远不及印欧语言研究上取得的扎实成果,见Wang 1995。要了解中国语言的演化史,尤其是史前史,我们必需特别注意许多民族语言相互接触所引发的横向传递,见陈保亚1996。

每个中国人身体里的基因,都是几千年的族群迁徙和融合的结果。同样的,我们说的汉语,也包含了许多方言及民族语言所传入的词汇及语法结构,我们一定要谨记这一点,才能看清中国语言演化的全貌。

两三千年前的古汉语,在当今的普通话里经历了最多的变化,可是很多古汉语的遗迹,却都保留在方言里。例如汉语的入声字,本来都有-p、-t、-k这些塞音作韵尾,唐朝柳宗元的诗《江雪》,若用普通话来念,三个句末的入声字“绝”jué、“灭”miè、“雪”xuě都已经并入别的声调了,可是广州话里,这三个字还保留着唐朝-t的韵尾,因此读起来仍是押韵的。再举一个声母的例子,很多方言里,“饭”这个字都是f-的音开头的,是个清的唇齿辅音。可是上海话里,同样一个字却读作/ve/,用v-开头,保留了古代的浊辅音。而厦门话里,“饭”是/pŋ/,这说明f-是由p-变出来的,就像英文的“脚”foot是由印欧语的ped演变出来的。而汉语里p-变成f-的这个谜,就是由清儒钱大昕所解开的。

时间上越古老,语言的中观史研究就越困难。我们知道汉语和亚洲的几百个其他语言都是同源的,包括藏语,可是很多藏语方言都没有声调。汉语的

声调究竟何时兴起,是个耐人寻味的问题。除了音系,语法上也有很多难以解答的问题,一个待解的谜就是,汉语里动词出现在宾语前面,例如“骑马”,是典型的动宾结构,而藏语的语法却应该说“马骑”,是宾动结构。如果要知道汉人及藏人的老祖先在五六千年前的词序如何,是说“马骑”还是“骑马”,尚需要许多研究及考证。

要是我们把时间的尺度更加缩短,从宏观史缩到中观史,再缩到微观史(microhistory),就可以直接观察到目前正在发生的演化现象。举一个我自己亲身经历过的例子。我小时候的40年代,在当时说的上海话中,“冷”和“狼”,“朋”和“旁”这两组字的读音,很容易区分清楚。可是这两组字在现在年轻人的上海话里,已经完全同音了,换句话说,ā和ā这两个鼻化韵母业已合并。这个上海话的微观演化例子,沈钟伟做过仔细的研究,调查了多位不同年龄的受试者的语言,并且运用统计方法来构拟这个合并音变的来龙去脉(Shen 1997)。他这种详实、量化的研究方法,足堪作为以实证态度研究语言的楷模。

我们知道每一个活的语言总是在演化,因此微观史的研究材料,要比中观史以及宏观史的丰富得多。世界上不会有两个人长得一模一样,也不会有两个人的语言是完全相同的。既然同一个语言也会存在许多变异,这便是语言演化的原动力。有些变异如昙花一现,有些却会一代代传下去。我在香港工作的这几年,发觉此地的粤语变化得特别快,很可能是由于香港汇聚了许多语言和方言,增加了横向传递的机会。一个很明显的音变就是,许多带有n-声母的词,都已经变成l-的音,如“你”一词,老年人也许还说/nei/,可是年轻人都说/lei/了。在餐厅里常见的饮品

“奶茶”,也已经从/naits' a/变为/laits' a/。

语言的演化与婴儿的语言习得关系密切,随着近年来认知神经科学的一些突破,这个领域的研究也有了日新月异的进展。90年代,意大利科学家发现了镜像神经系统,帮助我们更了解学习的生理基础,见曾志朗2006的介绍。最近,一组欧洲学者用fMRI技术,报道了刚出生两天的婴儿大脑里,已经具备了学习语言所需的神经网络(Perani et al. 2011)。这些都是微观史研究里特别振奋人心的成果,势必会给我们带来许多语言演化的新见解!

以上我们谈了三个尺度(Wang 1978)。语言演化当然是个连续的过程,变化是经由个别的词或结构一个一个扩散出去的,打个比方,我们现有的知识,不过是一些零星的片段,但借着累积微观史的断片,如ā和ā怎么合并,我们希望能把几十年的演化与几百年或几千年的中观史演化联系上,至于是否有一天终能把从中观史研究所得的知识,再推论到几万年前的宏观史上,目前还不得而知。

无论如何,从德国科学院在1769年提出语言涌现的问题,至今已过了将近250年。那时候只有欧洲人在关注这个话题,但现在国际学术界的许多学科,都对此问题有极大的兴趣。各个学科都不约而同的携手探究同样的议题,其实是个理所当然的现象,而今年在北京大学甫圆满落幕的CIEL-4,就是个绝佳的例子,说明中国语言学界也不愿自外于语言演化的研究,见王银凤(2012)、曾志朗(2012)、汪锋(Wang 2013)的报告。对于语言演化出来后的这段也许长达十万年之久的漫长岁月,我们目前已经掌握了若干零散分布于各个时间尺度的快照。比起赫尔德的年代,我们现在对历史、史前史、语言结构及类型

学、横向及纵向传递等，也都有了更充分的认识。其中最重要的，莫过于近年来学界对大脑的积极探索。演化语言学的一个愿景便是，我们终将理解那些快照所给予我们的简单印象，如何拼合成一幅连贯的画面，以便用来更全面地解释人类最珍贵的财产——语言是怎么演变、成长的。既然语言为人类所独有，能够用科学方法探索语言，也才能对“人是什么”这个根本问题有更深入透彻的理解！

参考文献：

王士元. 2001. 孟德尔与琼斯，道不同不相为谋？. 科学中国人 11.28-31.

王士元. 2006a. 演化语言学中的电脑建模. 北京大学学报（哲学社会科学版）.43.17-22.

王士元. 2006b. 语言是一个复杂适应系统. 清华大学学报（哲学社会科学版）.21.6.5-13.

王士元. 2008. 演化论与中国语言学. 南开语言学刊.2.1-15.

王士元. 2011a. 演化语言学的演化. 当代语言学.13.1-21.

王士元. 2011b. 语言、演化与大脑. 北京：商务印书馆.

王士元，柯津云. 2001. 语言的起源及建模仿真初探. 中国语文.282.195-200.

王银凤. 2012. 2012演化语言学国际研讨会在北京大学召开. 科学中国人.24.16-27.

陈保亚. 1996. 论语言接触与语言联盟：汉越(侗台)语源关系的解释. 北京：语文出版社.

曾志朗. 2006. 牵动你我神经—镜像神经为什么重要？. 科学人.58.72-75.

曾志朗. 2012. 石钢铁，猪狗鸡羊马，葱蒜黄瓜都有话说！. 科学人.130.6-7.

Balter, Michael. 2004. Search for the Indo-Europeans. Science 303.5662.1323-6.

Catani, Marco, Derek K. Jones & Dominic H. ffytche. 2005. Perisylvian language networks of the human brain. Ann

Neurol 57.8-16.

Geschwind, N. 1976. The development of the brain and the evolution of language. Selected Papers on Language and the Brain. 86-104.

Gong, Tao, J.W. Minett & W.S.-Y. Wang. 2010. A simulation study exploring the role of cultural transmission in language evolution. Connection Science 22.1.69-85.

Harnad, Stevan R., Horst D. Steklis & Jane Lancaster, eds. 1976. Origins and Evolution of Language and Speech. Vol. 280. New York: Annals of the New York Academy of Sciences.

Herder, J.G. 1772. Ursprung der Sprache. 姚小平译. 1998. 论语言的起源. 北京：商务印书馆.

Hockett, C.F. 1960. The origin of speech. Scientific American 203.88-96.

Karlgren, Bernhard. 1939. Etudes Sur La Phonologie Chinoise 中国音韵学研究. 昆明：赵元任、李方桂、罗常培译.

Ke, Jinyun, Tao Gong & W.S.-Y. Wang. 2008. Language change and social networks. Communications in Computational Physics 3.4.935-49.

Perani, Daniela, Maria C. Saccuman, Paola Scifo, Alfred Awander, Danilo Spada, Cristina Baldoli, Antonella Poloniato, Gabriele Lohmann & Angela D. Friederici. 2011. Neural language networks at birth. PNAS 108.16056-61.

Shen, Zhongwei. 1997. Exploring the Dynamic Aspect of Sound Change. Journal of Chinese Linguistics, Monograph No.11.

Thomas, Lewis. 1981. Debating the unknowable. Atlantic Monthly. 248.49-52.

Wang, Feng. 2013. Report of Conference in Evolutionary Linguistics (2012). Journal of Chinese Linguistics.

Wang, W.S.-Y. 1978. The three scales of diachrony. In Linguistics in the Seventies: Directions and Prospects, edited by B.B. Kachru, 63-75.

Wang, W.S.-Y., ed. 1995. The Ancestry of the Chinese Language, Journal of Chinese Linguistics Monograph No.8. 李葆嘉主译. 2005. 汉语的祖先. 北京：中华书局.



作者简介：

王士元，香港中文大学研究教授，主要在语言学及现代语言学系和电子工程学系，并兼任翻译学系教授。1960年于密西根大学取得博士学位，博士论文研究主题为声学语音学。取得博士学位后，曾先后任职于M.I.T. 电子研究实验中心(Research Laboratory of Electronics)与I.B.M. Yorktown Heights Research Center。之后他在俄亥俄州立大学创办语言学系及东亚语言文学系。1965年被聘为伯克利加州大学语言学系正教授，并任教授该系三十余年退休。

从开始执教至今，指导过将近40个博士生，包括多位目前在语言学界知名的教授。王士元先生共发表论著200多篇。

王士元先生来到香港后，致力于濒危语言的研究，曾屡次前往云南、广西，调查当地的少数民族语言。曾在美洲、亚洲、欧洲各地的大学做过多次演讲，包括Princeton的Institute for Advanced Studies、巴黎的Collège de France、京都的国际高等研究所等。