

Essays on Evolutionary
Linguistics

演化语言学论集

王士元 著



商務印書館
The Commercial Press

創于1897

演化语言学论集

王士元 著



创于1897

商务印书馆

The Commercial Press

2013年·北京

图书在版编目(CIP)数据

演化语言学论集/王士元著. —北京:商务印书馆, 2013
ISBN 978-7-100-09511-2

I. ①演… II. ①王… III. ①语言演变—文集
IV. ①H0.09

中国版本图书馆 CIP 数据核字(2012)第 226707 号

所有权利保留。

未经许可,不得以任何方式使用。



演化语言学论集

王士元 著

商务印书馆出版

(北京王府井大街36号 邮政编码 100710)

商务印书馆发行

北京瑞古冠中印刷厂印刷

ISBN 978-7-100-09511-2

2013年8月第1版

开本 787×960 1/16

2013年8月北京第1次印刷

印张 26 1/4

定价: 58.00 元

前言

语言是一个系统，可是却是个很不规则的系统。这主要是因为，语言老是在演化，只要有人使用，它就会变，而且语言内部不同的成分变的程序并不一样。当我们学外语时，就会发现差不多每条规则都有不少例外，让我们觉得学习时特别费力。当然我们的母语里也存在着种种不规则的现象，只是因为天天都在说，所以就感觉不到不方便。

著名哲学家维特根斯坦的比喻很贴切：他说一个语言就像一座古老的城市，有保留了数千年的街道土墙、古井石碑等等的文化遗产，也有刚盖好的钢筋玻璃的高楼大厦。不同年代的建筑凑在一起，每来了一批新的人群，就会在旧的基础上添加新东西而混合在一起。我在本书的《语言演化的探索》一文里，做过一个类似的比喻：语言像一张永远拼贴不完的马赛克镶嵌图，不同的部分都有不同的来历，有的是一代一代地纵向传递而来，有的是模仿其他语言经由横向传递而来。

每种语言都有大量来自他种语言的词汇，同时语言的其他部分，如语音、构词、句法，也就或多或少地跟着这些外来词发生变化。过了几十年或几百年，这些外来的语言成分就被完全同化，跟原有的语言成分分不开了。比方说英语里大约有一半的词汇，是在公元 11 世

纪左右由法语来的。日语里或韩语里的词汇,也有很多是从更早的世纪前不同朝代的汉语而来。举个现代的例子,香港由于殖民历史,其粤语长期与英语接触,因此有不少英语借词,如士多(store)、车厘子(cherry)、士多啤梨(strawberry)等等。这些词在广州粤语里就不存在,但在香港人的心目中,也都成为母语里很自然的一部分了。

这些语言现象无论是规则的或不规则的,新造的或固有的,来自社会因素或生理认知因素的,都源于语言演化,因此应当用演化论的观点及方法来探讨。虽然演化论最早用于生物演化,可是它有许多概念与方法,都完全可以应用在语言上。

我开始朝这个方向考虑问题,是在1969年那篇《竞争性演变是残留的原因》的老文章里。因为音变其实就是一种个别的词的发音竞争,所以后续就有许多朋友把这种词汇扩散理论分析得更清楚。

郑锦全(1971)的研究,是从词汇扩散的观点上,把中古汉语的声母变化分析得最完整的一篇文章。接着陈渊泉(1972)讨论了词汇扩散与时间的关系。谢信一(1972)探讨了怎样把这个演化理论用在婴儿习得语言上,而连金发(1993)则提出这种扩散是双向的。直至沈钟伟(1997)用统计模型调查上海、温州音变时,我们对扩散的程序及步骤已经有所认识。

梅祖麟(2000)在1980年就体会到语法研究也可以采用这种扩散法。小仓美惠子(1993)研究英语里*Do*结构的发展,以及余霭芹(1993)讨论汉语语法变化的一些现象,都是基于这个理论。前几年我和柯津云、Minett(2004)曾经发表过一篇文章,把这些从演化论出发研究语言的成果做了个总结。

近年来,我与香港中文大学语言工程实验室的一些年轻朋友做

了些新的研究,发表过一些文章。我们也筹办了两次工作坊,并邀请国外志同道合的朋友来参加,彼此交流共同的兴趣。Minett 和我主编的两本书(2005,2009),就是两次工作坊的文集。而我们实验室自己发表的文章中,有的是比较一般性的讨论,有的是利用电脑来模拟演化,还有的则是运用心理或神经实验来探讨个别的课题。

本书的作品可归成三大组。A 组是从前为好友们祝寿所写的文章,皆收录在不同的庆祝文集里。在《索绪尔与雅柯布森——现代语言学历史略谈》里我讨论了 20 世纪语言学历史的一些问题,虽然索绪尔的贡献的确不容忽视,可是他的负面影响使语言学成为过于孤立的领域;相对来说,雅柯布森视语言研究为一种多学科的课题,就像赵元任先生,他们的视野广阔得多,才是我们应该继承的传统。《迷上语言》简述了我选择走上语言学这条路的心路历程,可算是我个人的简短自传,也谈了一些这几十年来我对语言研究的体会和领悟。

B 组文章则刊载于好似风马牛不相及的各类学报里,有的是数学性的,如《从模拟仿真角度看词汇与句法的共同演化》;有的是物理性的,如《语言变化和社会网络》;有的可归于神经学,如《大脑半球偏侧化受双语状态和词的合成性的影响》;当然也有归于传统语言学的,如《语言演变中的横向传递与纵向传递》。这些学报分别由美国、欧洲、中国大陆和台湾等地出版,能让这些原本相隔几大洋的文章出现在论点相同的同一部书中,相信会引起更多的反响。B 组牵涉到的课题,有的已有更完整的论述,例如建模方面,龚涛(2009)的论文是很详尽的介绍与分析,又如大脑方面的探讨,除了本书收录的彭刚译《大脑半球偏侧化受双语状态和词的合成性的影响》一文之外,还

有我们实验室近来出版的彭刚, Minett, 王士元(2010); 彭刚, 郑洪英等(2010)及郑洪英等(2011), 都论及了些非常有意思的发现, 希望读者们参考指教。

C组里有四篇文章都取自《行为与大脑科学》(*Behavior and Brain Science*)。这个学报的特点在于它希望激励学者们互相讨论切磋与交流意见, 所以每篇篇幅较长的主题文章后, 就有好几十篇评论的短文。《凭空出现的语言器官》批判语言器官这个概念, 因为语言是一种行为, 类似音乐, 而不是个像眼睛或肝脏一样的器官。我在这篇文章里所用的标题 *Organum ex machina*, 来自古希腊戏剧里常见的一种手法, 就是每逢遇上一个人间解决不了的难题, 舞台上的机器里就会有神出现, 以妥善解决这个问题, 这种手段就叫作 *Deus ex machina*。在我看来, 把语言能力当作一种生理器官, 实在太过牵强, 是在耍同样的把戏。

《拟构人类演化历史》、《〈人类沟通的起源〉中文版序》、《〈语言演化生态学〉中文版序》介绍了三本非常有启发性的书, 是从不同的学科来谈语言。Cavalli-Sforza 是群体遗传学里的权威, 他把语言学、遗传学、考古学共冶一炉以讨论人类的演化, 的确贡献卓著。Tomasello 把婴儿与黑猩猩的认知发展对比, 得到了宝贵的第一手材料, 协助我们理解语言演化和语言习得之间的关系。Mufwene 则从演化论的观点来讨论语言接触, 这三本书共有的跨学科研究精神, 非常值得我们学习。

最后两篇本来是一本书的序言, 这本书是我替美国的《科学美国人》(*Scientific American*) 主编的《语言涌现: 发展与演化》(*The Emergence of Language: Development and Evolution*)。该书 1991

年出版后,很快就销售一空,也因此绝版了。所幸林幼菁博士将该书译为中文,并由台湾“中研院”语言学研究所于2008年发行中文版。由此两篇序言中,可以反映出这17年来语言学界的新进展。

以上这些作品散见于不同的文集或学报里,既然不太容易完整取得,也就更谈不上能一并观之而看出贯穿这些篇章的中心思想。所以,当石锋教授建议把这些材料收集在一起编辑成书时,我欣然同意了。非常感谢石锋教授负责统筹翻译及校正事宜,也谢谢几位朋友把一些英文文章翻译成中文,尤其是蔡雅菁小姐把全部稿子重新修订校阅了一遍,力求做到关键词汇和学术用语全书统一,让整本书更容易阅读,以避免读者因多位译者翻译专有名词时的不一致而徒增困惑。

最后我要由衷感谢香港中文大学校长沈祖尧教授及信兴高等工程研究所所长程伯中教授。多亏他们对我们实验室的信心与支持,我们多年来才能持续专注地投入研究中。香港研究资助局几次提供我们的资金,也发挥了积极的作用,在此一并致谢。

王士元

2011年7月于香港

参考书目

1. 陈渊泉. 1972. The time dimension: contribution toward a theory of sound change. *Foundations of Language* 8, 457-98.
2. 郑锦全, 王士元. 1971. Phonological change of Middle Chinese initials. *Tsing Hua Journal of Chinese Studies* 9, 216-70.
3. 龚涛. 2009. *Computational Simulation in Evolutionary Linguistics: A Study on Language Emergence*. Taipei: Academia Sinica.

4. 谢信一. 1972. Lexical diffusion: Evidence from child language acquisition. *Glossa* 6. 89-104.
5. 连金发. 1993. Bidirectional diffusion in sound change revisited. *Journal of Chinese Linguistics* 21. 255-76.
6. 梅祖麟. 2000. 《三朝北盟会编》里的白话资料. 《梅祖麟语言学论文集》. 28-61. 商务印书馆. 原文刊于: 1980. 《中国书目季刊》14. 2. 27-52.
7. Minett, 王士元(eds). 2005. *Language Acquisition, Change and Emergence: Essays in Evolutionary Linguistics*. Hong Kong: City University of Hong Kong Press.
8. Minett, 王士元(eds). 2009. *Language, Evolution, and the Brain*. Hong Kong: City University of Hong Kong Press.
9. 小仓美惠子. 1993. The development of periphrastic do in English: A case of lexical diffusion in syntax. *Diachronica* 10. 51-85.
10. 彭刚, Minett, 王士元. 2010. Cultural background influences the liminal perception of Chinese characters: An ERP Study. *Journal of Neurolinguistics* 23(4). 416-26.
11. 彭刚, 郑洪英等. 2010. The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics* 38. 616-24.
12. 沈钟伟. 1997. Exploring the dynamic aspect of sound change. *Journal of Chinese Linguistics*. Monograph No. 11.
13. 王士元, 柯津云, Minett. 2004. Computational studies of language evolution. *Computational Linguistics and Beyond, Frontiers in Linguistics 1, Language and Linguistics Monograph Series B*, ed. by C. R. Huang & W. Lenders, 65-108. Taipei: Academia Sinica, Institute of Linguistics.
14. 余蔼芹. 1993. The lexicon in syntactic change: Lexical diffusion in Chinese syntax. *Journal of Chinese Linguistics* 21. 213-54.
15. 郑洪英等. 2011. The impact of tone systems on the categorical perception of lexical tones: An event-related potentials study. *Language and Cognitive Processes*. doi:10.1080/01690965.2010.520493.

目 录

语言的涌现与传播.....	1
索绪尔与雅柯布森	14
语言演化的探索	30
迷上语言	78
演化语言学的新进展.....	105
语言学习与大脑:一些最新进展	123
从模拟仿真角度看词汇与句法的共同演化(合著).....	155
语言演变中的横向传递与纵向传递(合著).....	184
演化语言学中的电脑建模.....	210
语言是一个复杂适应系统.....	221
语言马赛克及其生物基础.....	241
语言变化和社会网络(合著).....	262
演化论与中国语言学.....	284
演化语言学的演化.....	308
大脑半球偏侧化受双语状态和词的合成性的影响(合著).....	347

凭空出现的语言器官·····	365
演化原则和句法的产生(合著)·····	368
语言的异质性和自我组织意识(合著)·····	373
在智能主体中产生感知范畴及其对实践研究的指导(合著)·····	377
拟构人类演化历史·····	385
《人类沟通的起源》中文版序·····	392
《语言演化生态学》中文版序·····	397
《语言涌现、发展与演化》原序·····	402
《语言涌现、发展与演化》导论·····	404
附录:本书论文出处一览·····	414

语言的涌现与传播

引言

爱德华·萨丕尔 1921 年曾推测语言的某种形式在人类制造石器之前就已经存在了。众所周知,这个时期要追溯到二百万年以前。把坚如石头的材料做成一组不同的工具,是人类演化的特殊阶段,它需要有心智能力来控制意象,其准确度和复杂性都远远地超过了任何其他物种的智力水平。

从目前所了解的灵长类社会的线索中,我们有理由相信原始直立人的各种级别的沟通方式也是十分丰富的,而且已经开始使用准语言的变体形式。这些沟通方式可能包括面部表情及身体姿势所表现出的视觉信号,以及由韵律特征(音高、音强和音质)的调节中所编码的声音信号。所有的人在言语中都或多或少地使用手势,并且所有语言都有语调。

一、通往真正语言的三个阈限

当早期人类偶然地发现,即使信号和结果没有因果关系,这些信号仍然在听话人身上产生一连串的作用时,人类超越了第一个

阈限。^① 我们可以把这个阈限称为符号化。它的第一次实现很可能是在模仿自然声音或无意中发出的噪音,但只有符号完全从原来的语境中抽象出来,并能自由使用,这个阈限才能越过。然而,还要跨过几个阈限之后,这种准语言才能演化成为准确的、强有力的、我们今天使用的智能工具。

与面部或身体活动产生的视觉信号不同,声音信号拥有明显的优势:首先,它是全方位的,并可长距离传送;其次,它可以越过视觉障碍,可以在黑暗中接收到;再者,它可以在做其他动作的同时发出,如,与打猎时需要的动作相配合。然而,在灵长类呼叫中使用的音韵特征的数量是非常少的,而且每一个单位时间内所携带的信息量也极为有限。准语言下一个要跨越的主要阈限是把声音信号组成链状的音节,这便是音段音系的创立。^②

第二个阈限的跨越使语言获得两个决定性的优势。一个优势是大大地扩展了语言材料中信息单位的数量,有些现代语言的单个音段已超过几十个。这个优势允许构建庞大的词汇库,有的语言拥有上万条词汇。^③ 另一个优势与发出音段的高速率有关,有时其速度可达每秒 12 个音段。以上两个优势使准语言得以超越人的短时记

① 参见 Helen Keller 的说明。她曾描述过自己如何第一次领悟到,当她的老师用手指在她的一只手上拼写出字母时,这个符号原来代表了在另一只手上流动的水。

② 元音开口和辅音闭口的交替是有节奏的咀嚼需要的功能,正如话语本身是呼吸和咀嚼需要的功能一样。元音提供声源动力,辅音提供听感差异的基础。

③ 然而,正如达尔文在他的早期著作《人类的遗传》中所描述的,人类有一个可以保存主动性词汇记忆的上限。记忆提供一种类似“推挤式储藏”的方式接受新词,淘汰旧词。这种上限可能在所有语言中都能粗略地进行比较。

忆限度,^①从而为句法的组织奠定了基础。

词是按时间顺序一个接着一个发出的,因而顺序就不可避免,所以顺序在某种意义上受使用者自由支配。下一个发明就是创建带有层次功能的顺序,以便让词的序列可以在词汇之间获得一套系统的关系。就像 Herbert Simon 在复杂结构的讨论中所观察到的那样,那些达到特定复杂程度的系统,可以预期地会出现层级组织,语言也不例外。一旦发明了句法,真正的语言也随之涌现,即为词序赋予功能。这是三个阈限的最后一个。

像符号化和音段音系化一样,语法的充实也一定是一个渐进的过程。一方面,词的构建只允许有限的递归:如“losslessnessness...”一词,但句法中的递归却是赋予语言具有“无限地使用有限的手段”功能的主要方式。家喻户晓的儿歌“This is the home that Jack built.”就是清楚的例证。递归法可能存在于一切语言之中。

另一方面,那些涉及构成成分(constituent)之间复杂关系的句法手段,在语言演化的过程中,可能出现较晚。其中一些句法手段可能得利于现有的书面语。因此,这些手段,特别是在前文字时期的社群中,可能不具有普遍性。一个具有说服力的例子就是句中话题与评价语进行间接地比较。如,

She is more beautiful than he is rich.

另一个有力证据就是文言词,如“respectively”,在句中的名词之间表现出超越序列的关系。例如,

① 我们发现当电话号码的叙说中带有较长的干扰性停顿时,人们记忆号码就越来越困难。当句子中的主要构成成分用太多的词分开时,也出现相同的困难。

Tom, Dick and Harry are the doctor, lawyer and chef respectively.

没有理由认为这三项发明,即符号化、音段音系化和句法化,是按照一个整齐的线性顺序创建,或是沿着单一的人类世系演变。可能在这数千年中,曾有过许多现已消失得无影无踪的失败的尝试。另外,早期部落间的接触也可能刺激并促进语言的演化,一个部落所取得的概念认识上的进步,很快会被另一个部落接受,成为优选的文化革新的时尚。

从严格的概率上讲,语言最有可能同时在许多不同的地点出现,是多源的而不是单源的(Freedman and Wang 1996)。诚然,并不排除可能有这样一种情景,即在人类演化的早期,除原始人类语(Proto-Sapiens)之外,所有其他的语言都被剔除了,原始人类语便是所有现代语言的祖先。

人类学家和遗传学家们,目前对于人类起源于非洲的假设所做的探讨,都证明了这种构想。^①最近的线粒体DNA和Y染色体显示其出现时间大约距今150,000年。探索全球性语源(即那些被认为可追溯到原始人类语的词汇)的语言学家提出的、令人满意的构想,具有更短的时期。因为很明显词汇要比DNA更替得快,所以如果原始人类语存在时间距离人类大迁徙的时间越近,就越有可能发现全球性的共同语词。

人们推测大约在距今40,000年左右,在艺术、宗教、工具组合和长距离航海方面,发生了一场“创造性的大爆炸”(Pfeiffer 1982),这无疑给全球性语源的研究带来了生机。在早期的讨论中(Wang

① 近期较为流行的是Stringer & MacKie (1997)的说法。

1976),我区分了跨越三个阈限之前的涌现(emergent)状态和语言的结构相对完整之后的稳定状态。人类语言进入一个稳定的状态之后,它们大多是保持平衡状态,进行周期性的变化,在一个周期中,语言中一个成分的简化通过另一个成分的增加得到补偿(Hodge 1970)。这样的互补关系最终可以在复杂适应系统的框架中予以分析(Gell-Mann 1992)。

距今 40,000 年似乎是状态转换发生的理想时期。严格地讲,统一性原则不适用于语言演化中历时分析的三个尺度(Wang 1978),因为语言涌现期和稳定期的传播机制不同。

即使有这样一个相对比较近的时代,追溯语言祖先的工作仍然是艰巨的。4 万年中,发生了几次重要的冰川活动。为了适应剧烈的气候变化,或逃避瘟疫和其他灾难,人群大量地迁移,他们的基因和语言随之在很大程度上相互混合。探寻全球语源毕竟有别于追溯特定的祖先,因此探寻语源的研究有可能更容易取得突破。

二、横向传播与融合

如果亚洲和欧洲的早期人类那时有语言的话,这些语言的一些特征很可能在非洲征服者的语言中留下痕迹。即使如此,今天也几乎不可能探测到这些痕迹。Cavalli-Sforza & Feldman(1981)认为:来自祖先语言的特征是通过纵向传播下来的,与其他语言接触而同化的特征是通过横向传播的。

对不同时期的语言进行研究的中心问题,是要分清由两种不同传播方式产生的特征。要想准确定位语言的历史,我们需要区分这两种不同的特征。但是,这种区分是极其困难的,特别是当所涉及的

语言相互间长期保持紧密接触时,这种区分几乎是不可能的。大多数汉语方言正是这种情况,二千年来,它们从地位较高的北方方言中反复吸收了大量的词汇。

用下面的表给出的广东潮州方言的材料来加以说明(Wang & Lien 1993:369)。如表所示,每个汉语音节可以分成三个成分:声母、韵母和声调。表中的“L”代表“书面语”,是来自北方方言的语言特征。与“L”相对应的“C”代表口语,从传统上讲,指那些被认为是该方言本来就有的语言特征。^①

也就是说 C 特征是经纵向传播的,而 L 特征是经横向传播而来的。在表的第一行,“饲”这个词的三个成分来自文言层,而“誓”的三个成分却来自口语。与这两个纯源词形成对照的“闹”是一个混源词,其中声母是口语,韵母和声调是文言。

声母	韵母	声调	例字			
1. L	L	L	饲	su	2b	to feed
2. C	C	C	誓	tsua	3b	oath
3. C	L	L	闹	lau	2b	noisy
4. L	C	C	露	lou	3b	dew
5. L	C	L	露	lou	2b	dew
6. C	L	C	谢	tsia	3b	thank
7. C	C	L	量	nio	2b	quantity

^① 中国语言学界把词汇读音分为两层,即文白异读。考虑到中国方言和语言之间上千年的接触,这种分法过于简单。根据借词的时间和地域,更细致地分为几层是比较真实的,如日语语言学中的汉音(kan-on)、吴音(go-on)、闽音(min-on)、唐音(too-on)。