

六合叢書

编余问学录

徐文堪

 ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

六合叢書

编余问学录

徐文堪

图书在版编目(CIP)数据

编余问学录 / 徐文堪著. —杭州: 浙江大学出版社, 2013. 12

ISBN 978-7-308-11903-0

I. ①编… II. ①徐… III. ①社会科学—文集 IV.
①C53

中国版本图书馆CIP数据核字 (2013) 第170925号

编余问学录

徐文堪 著

策 划 周 运

责任编辑 王志毅

出版发行 浙江大学出版社

(杭州天目山路148号 邮政编码310007)

(网址: <http://www.zjupress.com>)

制 作 北京百川东汇文化传播有限公司

印 刷 北京中科印刷有限公司

开 本 880mm × 1230mm 1/32

印 张 11

字 数 227千

版 次 2014年1月第1版 2014年1月第1次印刷

书 号 ISBN 978-7-308-11903-0

定 价 35.00元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社发行部联系方式: (0571) 88925591; <http://zjdxcs.tmall.com>

目录

谈语言起源与古代语文的释读	1
谈《现代汉语词典》及词典的编纂	19
论印汉翻译史研究和佛教汉语词典编纂	36
《汉语大词典订补》后记	46
汉语词源词典的编纂	55
谈早期西方传教士与辞书编纂	66
马礼逊及其汉语研究	76
关于近现代汉语新词词典的编纂	87
张永言《语文学论集》读后	96
关于“身毒”、“天竺”、“印度”等词的词源	108
关于越南语的系属	112
人类起源问题上的中国声音	115
寻找北京人化石的新线索	122
基因、语言和族群起源	131
对“农作—语言扩散假说”的审视	141
略谈有关南岛语的几个问题	151
也谈“种族”	154
维也纳归来谈吐火罗学	160

谈语言起源与古代语文的释读

郑诗亮记录整理

乍听起来，语言起源这个问题似乎让人无从谈起。它是从什么时候开始得到关注的？到目前为止，有哪些值得关注的研究成果？

徐文堪：语言的起源是个研究起来非常困难的问题。很久以前就有人关注这个问题，比如达尔文就谈过，卢梭也写过一本《论语言的起源兼论旋律与音乐的摹仿》，国内有中译本，2003年出版。北京外国语大学的姚小平先生还翻译过德国学者赫尔德（J. G. Herder）所著《论语言的起源》。但因为这个问题实在没法解决，后来差不多就成为一个禁区了，1866年巴黎语言学会曾经禁止讨论这个问题。这种情况一直到20世纪90年代才有所改变，有不少语言学家、考古学家、心理学家、人类学家以及其他领域的专家做了很多努力，提出了很多假设，他们试图弄清楚语言是怎样产生的，为什么会产生，何时产生，在什么地方产生……这些问题现在都还没有定论，因为直接证据很

少，但语言的起源和演化成为了一个跨学科的问题，慢慢引起了不同领域学者的关注，而且近十年来取得了比较大的进展。

现在，语言起源理论可以划分为两种假说：连续性和非连续性假说。所谓连续性假说的基本思想认为，语言不可能突然之间形成，而是有一个过程，一定是由人类的灵长类的祖先早期的前语言系统演变而来的。非连续性的假说则认为，语言有一些独一无二的特征是在人类演化过程的某个时间段中突然出现的，跟人类基因演化和突变也是有关的。过去对连续演化研究得比较多，也有很多的假设。现在还在继续研究的人当中，有的认为人类祖先由姿势产生了跟语言有关的符号系统，有声的语言也随着人类的进化慢慢产生出来。也有人认为，语言是模仿自然界的聲音，或者是一种感叹。几十年前我还小的时候，受的教育是说，人是劳动创造的——这是恩格斯的话，人是能够制造工具的动物，语言也是由劳动创造的。比如，我们的祖先在协同劳动时需要交流，由比划产生了有声的语言。但因为直接的证据很少，对语言的起源要做实证研究是比较困难的。近十年以来，随着科学的进步，就可以用新的方法进行了。比如，可以由化石来看人类的解剖学特征，观察是如何演进的，也可以用计算机来模拟这个过程。也有的学者比较强调语言的社会性、人的社会文化认知对语言起源的影响。

认为语言产生是非连续性的代表人物之一，就是乔姆斯基。大家知道，乔姆斯基是一个很有影响的语言学家，他也很关心政治，是著名的左翼政论家。可以说，在当代的语言学以

及相关领域来说，乔姆斯基的影响力是最大的。当然，也有不少人不同意他的意见，但不能不承认他的影响。我个人并不完全赞同他的观点，但有一些还是有道理的，应该重视。比如，他强调只有人类才有语言，其他任何动物都不可能具有语言能力，语言能力是人独有的。这个观点我觉得是很对的。上世纪六七十年代国外有些人做过很多实验，比方说，他们试图教猿猴讲话，但猿猴没办法发出人的声音，不止一个中心尝试过教猿猴辨认代表某个概念的词语，虽然也有一些进展，但最终都没有取得成功。这也说明动物不可能拥有人类这样的语言能力。

语言进化应该是存在一个发展过程的吧？

徐文堪：语言进化肯定是有有一个发展过程的，但我们今天几乎完全看不到这个过程。因为我们今天所接触的语言，包括很多濒危的语言在内，大概有七千种左右。所有这些语言都是高度发展的，现在世界上不存在什么原始语言了，无论这个群体——我们不用民族这个说法——经济文化发展水平是高还是低，并不是说经济文化水平高，语言就高级，水平低，语言就低级。通过全球范围内的语言调查，可以说明这一点。亚马孙丛林里有一种皮拉哈人，文化发展程度是比较低的，他们几乎没有数的概念，只能数到三，一般的原始部落总还有自己的故事、传说或是神话，皮拉哈人却没有。但他们也有自己独特的文化，他们的语言并不比其他的群体落后，在他们自己的文化系统内，语言是高度发达的，可以表情达意。就像有的语言学

家所说的，在一个荒山野地的原始部落里从事狩猎、采集的人，他讲的话，从语言的角度来说，不见得比亚里士多德、孔子、老子，或者爱因斯坦低级，完全可以处于同一个水平。这就说明人是生下来就有一种语言能力的，用乔姆斯基的话来说，是语言共性或者普遍语法——人生下来，头脑中就有一部语法了。当然，需要一定的环境才能表现出这种能力。一个小孩生在汉语环境中就能学会汉语，生在英语环境中就能学会英语，如果这个环境是双语的，那么他也能轻松掌握双语。而且，我们知道，学任何东西，有的人学得快，有的人学得慢，这种个体差异是存在的，这在艺术、体育方面表现得非常明显，但就语言来说，只要是正常的人，生理没有缺陷，他不需要父母刻意地教学，就能很自然地学会一门语言，或者双语，如果环境是双语的话。有的人口才好，有的人不善言辞，但这是表达能力的差异，而语言能力的高低则是另一回事，只要是正常人，掌握语言能力的水准是相当的。从这些地方可以看到，关于语言的起源和演化，一方面有社会文化因素，另一方面也有生物学因素，否则刚才说到的现象就没办法解释了。当然，也有学者更加强调文化因素的作用，不赞同普遍语法，例如这几天国外就出版了一部新书，认为语言演化源于文化，语言是文化的工具。

那么，研究语言起源究竟应该如何着手呢？

徐文堪：我觉得，语言起源问题需要从文化和生物两个方

面研究，这就需要跨学科的研究，不然很难讲清楚。过去我们可能比较忽视语言的生物学基础，这一点，通过现在的研究就比较清楚了。所以，现在包括乔姆斯基在内的一些学者，提倡生物语言学（Biolinguistics）的研究。大概在 2001 年到 2002 年间，就有报道说发现了跟语言相关的基因 FOXP2。这个问题我想不详说了，语言这么复杂的事物，当然不可能只由单个基因决定，但是语言和基因两者是有联系的，从这可以看到，语言的演化和基因的演化有关。作语言起源研究的时候，有多种方法，比如对灵长类动物进行观察研究，国外此类研究比较多，又比如，对婴儿语言习得的研究，现在还可以借助很多新的技术手段，比如功能性核磁共振，最近媒体也有报道，通过核磁共振看人类想读一个词语时的脑部活动，还有就是刚刚讲到的通过计算机建立模型对语言演化过程进行模拟。美国、澳大利亚、新西兰和欧洲都在做这方面的工作。

人什么时候具有语言能力，上限在什么时候，现在还说不清楚。有的人认为可能在直立人的时候，有的认为在能人的时候。我个人认为，这样一种能力的萌芽，可能在三四十万年前，但也不会太早，我们知道的北京人，过去称为中国猿人，实际上应该称直立人，可能就没有语言能力。真正具有比较完备的语言能力，可能要在智人形成以后，离现在也不过几万年，有的说五万年，有的说十万年。根据现在的解剖学研究或生理学研究，在智人产生之前，尼安德特人也可能具有一定的语言能力。

关于语言形成的问题，我们还要注意到，随着现代人从非

洲迁徙分布到全球以后，世界上不同地区的语言互相接触、互相影响，也会形成新的语言。上海有所谓的洋泾浜语，也就是外语在本族语影响下产生了变种。其实，这种情况不光是上海有，在比上海更早的通商口岸，比如广州也有这个现象。这种现象称为皮钦语（Pidgin），在全世界范围内都有，现在就有所谓中国式英语，比如 People Mountain People Sea。如果在皮钦语的基础上，把本地语和外来语结合起来，就会形成一种混合语，这种混合语约定俗成以后，就叫它“克里奥尔语”（Creole Language）。世界上有不少语言就是这样形成的。例如我国四川甘孜藏区有一种特殊语言叫倒话。两百多年前，有一批清兵被派往今天甘孜州雅江县河口镇镇守，汉族船夫也被征来经营渡口，他们逐渐与当地人通婚，世代居住，繁衍成为如今说倒话的居民。倒话最主要的特点在于：词汇主要来自汉语，但语法结构上又与藏语语序和语法范畴有高度同构关系。所以，倒话是一种汉藏混合语。

至于语言是在什么地方产生的，因为直立人也好，智人也好，都是从非洲走出去的，所以现在认为非洲是语言的故乡的学者要多一些，但这也不是定论。

最近《文汇报》的一篇报道说，复旦专家的最新研究成果表明，如果全世界的语言有一个扩散中心，这个中心最有可能在亚洲里海南岸，与传说中“通天塔”的位置不谋而合。

徐文堪：这个情况，我想稍微作一些介绍。今年2月10日，

复旦现代人类学教育部重点实验室的李辉教授课题组与中文系的陶寰副教授合作，在《科学》(Science)杂志上发表了一篇文章。在此之前，新西兰有一位名叫阿特金森(Q. D. Atkinson)的学者，他2006年在新西兰的奥克兰大学心理学系获得博士学位，之后，到英国牛津大学还有其他一些大学作了三年研究，现在是奥克兰大学心理学系的高级讲师。他研究的范围是比较广的，他本人并不是语言学家，但他2003年就和他的老师，一位叫格雷(Russell Gray)的学者合写了一篇讨论印欧语系的文章，在《自然》(Nature)杂志发表。我们知道，两百年来，对印欧语系的研究已经很深入了，阿特金森和格雷用一种新的随机统计的方法研究印欧语的故乡和它的各支语言分化的年代。他们得出结论，印欧语系的年代距今已有八千五百年左右，起源地点可能是在安纳托利亚，属于今天的土耳其。这篇文章发表后引起了很多关注，引用率非常高。到2005年、2006年的时候，这两位学者又对方法作了一些改进，同时，除了对印欧语进行研究，也研究了其他一些语言。大家知道，新西兰属于大洋洲，他们对大洋洲、对东南亚的语言，都比较关注。他们研究的就是南岛语系(Austronesian family)的语言，从台湾到非洲的马达加斯加岛、美洲的复活节岛，都说这种语言，这也是一个大的语系。他们也建立了一个语料库，研究非洲的班图语。后来阿特金森也研究其他问题，比如宗教的起源等，但他主要还是关注认知领域。2011年4月15日，阿特金森在《科学》上发表了一篇文章，主要是通过对全球五百多种的语言音素进

行分析，建立模型，推断人类最早的语言产生于非洲。这篇文章发表以后也引起了不少关注，很多媒体进行了报道。

李辉教授的专业好像不是语言学，而是生命科学？

徐文堪：复旦生科院的李辉博士虽然专业不是语言学，但他对这个问题确实关注了很久。在2008年的《现代人类学通讯》上，李辉发表过一篇《人类语言基本元音体系的多样性分析》，提出把音位转化为数字，世界语言的元音音系的多样性，可以用遗传学中的线粒体多样性的网络分析方法进行分析。这个课题组作了大量的统计和分析，去年也投了一篇文章给《科学》，今年《科学》刊载了他们针对阿特金森观点的文章，主要是和阿特金森商榷，也就是不赞同阿特金森的观点。文章不是作为论文，而是作为评注刊登的。同时刊登的有三篇评注，其中一篇是两位美国学者写的，一篇是在德国和荷兰工作的学者写的，再有就是李辉、陶寰他们的文章；这篇文章还有两位作者是王传超、丁琦亮，李辉是通讯作者；最后，刊登了阿特金森的回应。《科学》在同期发表四篇文章，实在非同寻常，可见对此十分重视。李辉等人得出的结论是，语言演化如果有一个中心，应该是在里海南岸——这是媒体报道的说法，文章里说是在中亚。这些年轻学人，花了很长的时间，用了很大的精力，对这么一个前沿问题进行探讨，这是很好的。阿特金森自己也对进一步的讨论和研究表示欢迎，但他认为这些还不足以推翻他的结论——用他的术语来说，叫作系列奠基者效应，指

的是从非洲开始，离起源地远，语音就趋向简化，换句话说，语言在发源地非洲音素非常复杂，越远就越简单。实际上，在阿特金森的文章发表以后，有一些专业语言学家也在专业刊物上和他商榷。这些不同意见可以分为两类：一类是认为他的方法本身有问题；另一类则认为，方法本身大体可行，但材料处理存在一些瑕疵，得出结论过程中的一些推断也是有问题的。李辉博士他们的工作我想可能是属于后一类，并不完全否定阿特金森的研究路径，但是用的材料更加精细。德国有个很大的科学研究机构叫马克斯－普朗克学会，它有一个进化人类学研究所，其下有一个语言学学部，这个学部编了一部《世界语言结构图册》，既有牛津大学出版社 2005 年的纸质本，也有网上的电子版，书中收罗了世界很多语言，对它们的语法、形态、词汇作了分析。阿特金森和李辉他们都参考了这部书，但是李辉团队也做了不少工作，搜集了不少材料。现在这个讨论还在继续。3 月 1 日出版的《科学》又发表了美国罗切斯特大学和麻省理工学院三位学者的一篇评注和阿特金森的回应。我个人希望，李辉博士他们能继续把这个课题做下去，做得更深、更细，也希望他们能和国外有更多的交流。3 月份在京都就有一个语言演化的国际会议，阿特金森的老师格雷是要做大会报告的，其中肯定也会谈到这方面的问题。实际上，这个会议也有华裔学者参加，香港大学一位叫龚涛的年轻学者就要作特邀报告。

您提到，李辉他们用于语言学的材料更加精细了，这些材料是如何获取和积累的？

徐文堪：他们参考了有关文献，也进行了大量统计和计算。实际上，现在作关于语言起源和演化的研究，还要仰仗于考古学和人类学的发现。顺带一提，《文汇报》登的关于李辉博士的报道中提到有这么一句话：“课题组最终发现，人类语言起源于中东里海南岸，而这一地点位置与传说中的‘通天塔’的位置不谋而合，也是黑格尔在1863年《自然创造史》中所描述的现代人类种群扩散的地点。”这里说的“黑格尔”实际上是德国学者海克尔。毛泽东对海克尔是很有兴趣的。说来也巧，2月18日《文汇报》登了一篇《马君武与〈宇宙之谜〉》，《宇宙之谜》的作者就是海克尔。马君武没有全译《宇宙之谜》，译出来的部分登在当时的刊物上，青年毛泽东读到了。海克尔是一个很有名的生物学家，贡献很大，他是德国最早的进化论者。他的书除了马君武译过，刘文典等人也译过（《上海书评》曾于2008年7月13日，10月12日，10月26日刊发过高山杉和袁志英关于海克尔的介绍。——《上海书评》编者）。因为毛泽东对《宇宙之谜》有兴趣，1974年、1975年左右，作为一个政治任务，一些搞德语的老师把这本书翻译出来了，足足印了四十多万册，也送给毛泽东看了。为什么要翻译呢？毛泽东在接见施特劳斯时说：有四个德国人对我的世界观形成有重要影响。前三个是黑格尔、马克思、恩格斯，这大家都知道，但最后一个海克尔，翻译不知道是谁，还是译成了“黑格尔”。德国人很奇怪，

为什么要把“黑格尔”说两遍呢？毛泽东也意识到了这个问题，摆手示意不对，在座的一位读过这部书的苏联问题专家才会过意来。海克尔写的书非常多，《宇宙之谜》是他的代表作；“通天塔”的这篇报道里提到的《自然创造史》是一部科普著作，影响也是相当大的。海克尔是进化论者，他认为任何物种都有产生发展和灭亡的过程，不断地变化，不断地消亡。毛从他那儿得到了不小的启发，“不断革命”的思想事实上是受海克尔的影响。1967年，毛泽东在和刘少奇的最后一次谈话中，对刘说要好好学习，推荐了三本书，一本是《淮南子》，一本是费尔巴哈的书，还有一本就是海克尔的书，但究竟是哪一本，现在也很难说了，有人说就是《宇宙之谜》，所以“文革”后期这本书被大量印刷。我查了一下，这里提到的《自然创造史》，马君武也有一个译本，倒是全译本，1935年由商务印书馆出版。

实际上，海克尔的思想很有影响，他认为人类起源地是在亚洲，具体在印度洋、东南亚一带。我们知道，印度尼西亚曾是荷兰的殖民地，有一位叫做杜布瓦的荷兰学者看了海克尔的书，决心到东南亚寻找人的祖先，要把缺失的一环给补上。为了到印度尼西亚去，杜布瓦读了医学院，不惜签订八年的合同，到印尼去做军医。他在印尼找到了几件猿人的化石，把化石带回了欧洲。但欧洲人看到他的化石，并不相信，认为是猿的化石。杜布瓦是很执着的，依然认为这是猿人的化石。到了后来，北京周口店也有化石发现，1936年，有一个比杜布瓦年轻的、叫做孔尼华的德国人类学家到印尼去发掘化石，也有发现。孔

尼华认为，他所发现的材料肯定是人，用今天的术语来说，叫作直立人——这个术语就是杜布瓦定的。根据孔尼华和魏登瑞的合作研究，这些化石已经是人，而不是猿了，但杜布瓦依然坚持认为这是猿人。从这件事可以看到，从事研究是需要一代又一代人努力的。最近我还看到，在印尼工作的美国学者乔昆(R. L. Ciochon)等去年在《人类进化学报》上发表论文，介绍新发现的一百五十万年前的爪哇人的化石。有一本书《龙骨山：冰河时代的直立人传奇》就是他和另一位学者合写的，复旦的陈淳老师和他的学生译了出来，由上海辞书出版社出版。这些材料都有助于从考古学和人类学的角度来说明语言起源问题，材料是越来越多了，慢慢可以取得比以前更加深入的认识。

刚刚您谈的是语言起源问题，语言起源是在离我们比较遥远的史前时代，而文字的产生只有几千年历史。我们如何释读古代的语文？

徐文堪：古代语文的释读也是很引起人们关注的一个问题。我们要释读中国古代的文字，如甲骨文、金文等，有一个前提，就是按照汉字的规律来释读。其他中国境内发现的少数民族语文，如西夏文、契丹文、女真文等文字，记录的就不是汉语了，但由于受到汉文化的影响，还是汉字的衍生文字，还是汉字那种方块字的形式。在释读这些文字的时候，最开始我们也是不认识的，但因为汉文化的影响，其中有不少与汉文化相关的内容，如年号、帝王称号等，肯定是依照汉文化的习惯。这些来

自汉语的借词可以作为突破口，把这些专名辨认出来，整个文献也就慢慢可以通读了。有的时候还有汉语和少数民族语文对照的材料，西夏人自己就编过《番汉合时掌中珠》这一类词典，也是按照中国传统汉文词典那种方式来编纂的，从中可以得到很多线索，我们可以用这种办法来通读少数民族语文。

事实上，要解读一种未知的文字，和解读一种以前不知道的语言，这两者是有区别的。比方说，我们要解读甲骨文，是在肯定它是汉语的前提下、按照汉语的方式来解读。如果说记录的根本就不是汉语，那可能就困难得多了，因为没有什么线索。刚才举的契丹、女真、西夏等少数民族文字都属于汉字这个系统，西南和南方有些少数民族还有越南也模仿汉字造了他们自己的民族文字，越南有喃字，壮族有古壮字，还有白族的白文等等，都是汉字系统的。云南纳西族很著名的东巴文，不属于汉字系统，留下的写本很多，大约始于南宋，但到近现代当地仍有人能读。至于新疆，也就是古代狭义西域，有很多语文都不是汉字系统的。我们知道，19世纪末、20世纪初，有很多外国的探险队、考古队到中国新疆、敦煌活动，他们发现了大量各种语文的文献，当然其中最多的是汉文的，其次是藏文的，还有其他许多语文的文献，如突厥文、中古伊朗语、吐火罗语的文书。新疆发现的语种就特别多，当时去的，英国有斯坦因，法国有伯希和，俄国有奥登堡，日本有受大谷光瑞派遣的橘瑞超，瑞典有斯文·赫定，德国有勒柯克和格伦维德尔，这些大家都是耳熟能详的。德国获得的非汉文文书实际上是最