



中国社会科学院老学者文库

亚欧语言基本词比较研究

卷一（通论）

吴安其◎著



中国社会科学出版社



中国社会科学院老学者文库

亚欧语言基本词比较研究

卷一（通论）

吴安其◎著

中国社会科学出版社

图书在版编目 (CIP) 数据

亚欧语言基本词比较研究: 全5卷 / 吴安其著. —北京: 中国社会科学出版社, 2017. 1

(中国社会科学院老学者文库)

ISBN 978-7-5161-7911-6

I. ①亚… II. ①吴… III. ①比较词汇学-南印度语系 ②比较词汇学-印欧语系 ③比较词汇学-高加索语系 ④比较词汇学-芬匈语系 IV. ①H620.3 ②H703 ③H650.3 ④H660.3

中国版本图书馆 CIP 数据核字 (2016) 第 070530 号

出 版 人 赵剑英
责任编辑 王 茵 马 明
责任校对 朱妍洁
责任印制 戴 宽

出 版 中国社会科学出版社
社 址 北京鼓楼西大街甲 158 号
邮 编 100720
网 址 <http://www.csspw.cn>
发 行 部 010-84083685
门 市 部 010-84029450
经 销 新华书店及其他书店

印刷装订 北京君升印刷有限公司
版 次 2017 年 1 月第 1 版
印 次 2017 年 1 月第 1 次印刷

开 本 710×1000 1/16
印 张 176.25
字 数 2420 千字
定 价 638.00 元 (全五卷)

凡购买中国社会科学出版社图书, 如有质量问题请与本社营销中心联系调换
电话: 010-84083683

版权所有 侵权必究

前 言

20 世纪 30 年代布拉格学派的俄国学者特鲁别茨科伊 (Trubetzkoy) 考虑到语言之间语词的借用情况说：“由此看来，所有印欧语言都由同一语言分化而来的设想便不能令人们信服了，因为印欧语的祖先可能原本不一样，只是通过长时间的互相接触与借用才有今天相似的面貌。在讨论印欧原始语问题的时候，我们应时刻考虑到这种可能性。单一原始语的设想在当下已成为主流，正如当年中国‘罢黜百家，独尊儒术’一样：失去百家争鸣的同时，我们也已经失去了纯粹的语言学精神，假以时日，这种精神将灰飞烟灭。”^①

亚、欧两地主要分布着汉藏、阿尔泰、南岛、南亚、达罗毗荼、印欧、高加索、芬兰-乌戈尔和亚非九个语系的语言，这些语言的基本词往往有跨语系的词源对应关系，一定程度上反映了更早以前语言的渊源关系。事实上没有一个语系是孤立的，也没有一种语言是孤立的。在我们这部《亚欧语言基本词比较研究》的书中读者们将会看到印欧语不仅不是孤立的，还与亚洲的语言有密切的关系。

语言及方言的扩散、迁徙和转用的研究是有关语源的，词源学以往是关于文献中出现的词的历史，是角度不同的观察。现在看来这两个方面研究的范畴还可以再扩大一些。结合这两个方面进一步的研究可以发现更多的问

① 特鲁别茨科伊：《关于印欧语言问题的设想》，邹一帆译，石锋、张洪明主编《语言学译林》，世界图书出版公司 2013 年版。

题，对一些久而未决的问题提出新的解释。

下文诸卷在讨论亚欧人群和语言历史的基础上讨论了亚欧语言 275 项基本词的词源关系，诸篇以独立文本的形式探讨一些词或词根跨语系对应的情况，并以印欧语与东亚太平洋语言意义相关的基本词的词源对应关系为基础，进一步探讨亚欧和非洲、美洲、巴布亚、澳大利亚原住民语言的词源和语源关系。所涉及的东亚太平洋语言的分布和归类已在第五卷中说明。为了讨论的方便我们沿用传统的分类，并不刻板地认为这些分类中的语系或语族的语言都是从某一种原始语中分裂出来的。

参考考古、体质人类学的结论，根据东亚太平洋语言基本词的词源关系的情况，可以推测最近的四五万年东亚有四次大规模的迁徙。一是末次冰期以前南方人群北上抵北亚，南、北不同人群的结合；二是末次冰期北方居民南下；三是末次冰期结束之后距今一万年前开始的南方居民的北上；四是五千多年前到两千年前的这段时间里阿尔泰人向西和向南迁移，藏缅人沿西南民族走廊南下，沿海地区南岛语居民南下并向太平洋岛屿迁移。

笔者发现东亚太平洋地区语言的基本词，如近指代词“这”、形容词“圆的”“小”等的说法可分别归为有词源关系的三个或四个组，这意味着东亚太平洋地区的语言较早前有着多个源头。

比较亚欧语言的基本词的词源关系，可以看到跨语系分布的基本词中，第一人称代词的词根可以作为标记。亚洲、欧洲、非洲、澳大利亚和美洲印第安人语言的第一人称单数代词词根的词源关系主要有 *ni(*ne、*no、*na)，*g^wa(*go、*gu、*ku、*ŋa) 和 *mi(*bi、*me) 三类，与包括第二人称在内的其他基本词有一定程度一致的分布。

亚洲、欧洲和非洲基本词词源关系的分布为：从非洲、欧洲到北亚及太平洋地区；从非洲、中东到南亚和东亚太平洋地区。

东亚语言和印欧语农业文明相关的词有渊源关系，可能是末次冰期之后东亚向中亚、中东的移民，把自己与农业相关的词汇传给前印欧语系的语言。

根据苏米尔语和阿尔泰语基本词和形态的对应关系可以推测早期的苏米尔人可能是末次冰期后从东亚经中亚来到中东的。

比较的材料涉及高加索语系的格鲁吉亚语等，芬兰-乌戈尔语系的芬兰语、爱沙尼亚语和匈牙利语，巴斯克语，亚非语系的希伯来语和阿拉伯语，尼罗-撒哈拉语系的卡努里语（Kanuri）和扎尔马语（Zarma），尼日尔-科尔多凡语系班图语支的祖鲁语（Zulu）、科萨语（Xhosa）和斯瓦西里语（Kiswahili），科伊桑语系的科洪语（!Khoong）和科伊科伊语（Khoekhoe）等，爱斯基摩语和几种印第安语。涉及的印第安语有：美国新墨西哥州阿巴齐语（Apache）、明尼苏达州达科他语（Dakota）、苏语（Sioux）、俄克拉荷马州车罗科语（Cherokee）、墨西哥那瓦特尔语（Nahuatl）和玛雅语。我们眼界的开阔和工作效率的提高得益于对现代电子语料库和互联网的利用，我们感谢所有为此做出贡献的学者和工作人员。

各篇古汉语读音的解释可参见笔者《汉藏语同源研究》的有关章节，南岛语的语音对应和演变的解释可参见笔者《南岛语分类研究》的有关章节，有关印欧语的构拟参照这一方面专门的研究，不同之处另有说明。全书共五卷，第一卷为通论，第二、第三卷有关名词，第四卷有关动词，第五卷有关形容词、副词、代词和数词。多有疏漏，敬请指正！

卷一目录

语源和词源.....	(1)
一 语言的承传和词的对应	(1)
二 语言的演变和传承	(9)
三 人称代词的分布和语言的迁移.....	(26)
东亚语言的历史关系	(49)
一 阿尔泰语的历史	(49)
二 南岛语的历史	(58)
三 汉藏语的历史	(71)
四 南亚语的历史	(94)
五 东亚太平洋语言的语音和形态.....	(105)
商周汉语的词源关系	(143)
一 商周汉语的语言文化背景.....	(143)
二 商周汉语基本词的词源关系.....	(146)
三 汉语和印欧语基本词的对应.....	(180)

亚欧语言的历史关系	(199)
一 亚欧地区早期的文明及其语言	(199)
二 印欧语的语源和词源的关系	(215)
三 芬兰—乌戈尔语基本词的词源关系	(263)
四 高加索语系语言基本词的词源关系	(282)
五 巴斯克语基本词的词源关系	(292)
六 苏美尔语基本词的对应关系	(299)
七 达罗毗荼语的比较	(327)
 亚欧语言和非洲语言的历史关系	 (343)
一 闪米特语的比较	(343)
二 撒哈拉语族卡努里语的比较	(362)
三 班图语支祖鲁语的比较	(374)
四 科伊桑语系他语的语音和词汇	(384)
五 非洲语言与亚欧语言词源的比较	(394)

语源和词源

◇ — 语言的承传和词的对应

1. 承传和分化

亚欧大陆不同人种体质和文明的差别四五万年前的旧石器时代已出现。距今四五万年前的广西柳江人已具有蒙古人种的特征，同一时期欧洲出现了持有莫斯特文化的高加索人种的居民。世界各地不同人种的体质和语言特征的差异是人群分隔和语言相对独立演变的结果。

同一地区的不同社团方言通常是新老交替，一代一代地传下去。分处两地的相近方言随着一代又一代各自的新老派社团方言的交替，最终成为差别较大的地域方言。

语言分处两地经历了各自的演变则成为不同的方言。索绪尔解释道：“这些差别是什么造成的呢？如果认为那只是空间造成的，那就受了错觉的欺骗。空间本身是不能对语言起什么作用的。殖民者在离开 G 在 G' 登陆的第二天所说的语言跟前一天晚上所说的完全一样。人们很容易忘记时间的因素，因为它没有空间那么具体。但是实际上，语言的分化正是由时间因素引起的。地理差异应该叫做时间差异。”^①

① 索绪尔：《普通语言学教程》，高名凯译，商务印书馆 2001 年版，第 277 页。

人类社会的近亲通婚必定带来遗传疾病不利于种族的延续，史前社会的条件下要逃脱氏族消亡的命运只能与其他群体结合，这一情况摩尔根在《古代社会》一书中已经说得很清楚。东亚古代的氏族和部落社会中使用不同语言或方言的人群通婚，应是最普通的情况。早期人类处于不同氏族或不同部落交错和相互通婚的状态，不同语言或方言的接触，观念上的相互影响和词的大量借用是难免的。

在不同氏族或不同部落交际语的传播中精通自己原来的方言和熟悉交际语的双语人起着关键作用，他们往往在一定的圈子里讲一种以自己原来的方言或母语为基础的交际语。如民国时期和清晚期的国语，即今天的“普通话”，就是这样在官场和知识界形成的。如今在北京除了本地话，还流行着带儿化韵和不带儿化韵的两种“普通话”。“普通话”夹杂着一些明代的南京话的词和清代的东北话的词，还受到河北等周边地区移民的影响，其源头可追溯至唐宋时期的中原通语，以及两千多年前诗经时代的方言。

2. 迁徙和传播

（1）人群的迁徙和语言的分布

据地质学和古气象研究，在近二百万年中全球气候经历了许多次冷、暖交替的重大变化。每一次寒冷的气候持续约一万年，这一时期称为“冰期”。距今四万三千年到三万八千年前的这段时间里，全球环境与今天差不多。距今二万二千年到一万五千年前气候最为寒冷，称为末次冰期。距今一万二千年到一萬年前开始，气候转暖。

旧石器时代人群迁徙频繁，但语言的分布仍有大范围的格局，其迁徙的方向与气候变化对植被等的影响有关。东亚的阿尔泰、南岛、汉藏和南亚诸语系语言基本词有对应关系外，它们的一些基本词还分别与今天分布于南亚次大陆达罗毗荼语系语言有对应关系，应是旧石器时代南亚和东南亚居民向

东亚各地移民的结果。

新石器时代农业出现，人群从迁徙到定居，农耕民族分布于河边台地或冲积平原，渔猎民族分布于草原森林，语言的分布也趋于稳定。亚、欧两地不同语系的语言分别在一定的范围内连绵分布。

人群缓慢迁徙，则带来语言或方言的逐渐推移。如我国的南方很多地方不同村寨的居民使用相近但有一定差异的方言土语。相邻村寨之间的居民较多交往，方言土语也更接近，相隔越远差异越大。

人烟稀少的地区和居民稠密地区的语言传播不同，江南十里不同音和城市语音混杂的情况不同，游牧民族和农耕民族的语言传播不同。不同人群有自己的交际语。部落联盟时代，以部落之间的沟通为目的形成了部落交际语，城邦时代城乡之间以城镇的口语为交际语。总的说来交际语的形成以不同人群的需要和一定的语言或方言为基础，并经历一定过程的演变以适应需要。

（2）社会组织 and 语言的变化

我们先有氏族、部落、部落联盟，后有民族。氏族、部落的兴衰带来语言的兴衰。不能想象早期的氏族语言在人烟稀少的环境中能够一直保留着自己的传统。没有一种语言的词汇不是混合的，只是不同语言的情况不同。东亚诸语言的历史是数万年氏族、部落语言的传播和转用的历史。

氏族社会在东亚延续了相当长的时间，母系氏族社会在东亚和美洲多数地区大约持续到末次冰期以后，至气候转暖，农业出现，人口迅速增加，才出现部落社会，然后才有部落联盟的时代。母系氏族家庭中语言按照母系氏族社会传播的特征传播，其中包括男人们为了寻找异性，母系家庭为了避免近亲婚姻接纳外来的男子造成的语言的广泛接触。

五千多年前黄河流域的不同氏族结合成部落，出现了父系社会。部落是氏族的结合，希腊和美洲印第安人的部落也都是如此。每一个部落都有自己的地域和自己的方言，于是出现了语言或方言和地域联系在一

起的情况。印第安人的一个部落通常不超过两千人，部落中如果出现不同方言，则是部落兼并的结果，而且方言相近的部落总是相邻的。^① 后来的部落联盟必定以某一语言为交际的基础。^② 古代的部落交际语应是不同部落之间交流的主要语言，并可能成为一些人群的母语，原有语言的底层得以保留。

部落联盟时代部落原本有不同的方言，以某一种方言为基础形成的通语中难免带有各种方言词。如果各部落的语言原本是不易相通的，则可能以某一种部落语为基础形成一种通语。不同语言的居民在使用部落交际语或地区共同语时改变了这些交际语，便使之成为不同风格的地区性“普通话”。早期的交际语在不同地区被改造，以缓慢的或跳跃式的形式向外传播。在中东和欧洲，宗教语言的传播和相隔形成不同的亲属语。如拉丁语族的语言来自通俗拉丁语，斯拉夫语是在古教堂斯拉夫语的基础上形成的。

交际语对土语的改造使它们成为相近的方言或亲属语。古代语言的底层经历了多次的语言转用之后，仍可以保留在今天的语言中。

末次冰期结束之后，阿尔泰语主体形成于今内蒙古地区，其分化的方言与北亚及北上的南岛语系的语言结合形成突厥、蒙古、满通古斯、朝鲜和日语等阿尔泰语系的语言。此后因相邻的语言进一步接触，如突厥语和蒙古语，蒙古语和满通古斯语，满通古斯语和朝鲜语，朝鲜语和日语等分别有更多相似的特点。

（3）人群的迁徙和语言的接触

末次冰期之后，沿海北迁的南岛语在日本沿海、朝鲜半岛和中国北方的沿海和阿尔泰语接触，南下后又与南方的语言结合。台湾的南岛语保留着与阿尔泰语有过密切关系的特点，如人称代词的一致和某些语音平行演变的

① 路易斯·亨利·摩尔根：《古代社会》，商务印书馆1964年版，第102、107页。

② 同上书，第122页。

特征。

巴布亚和新几内亚是波利尼西亚和新科里多尼亚南岛语扩散过程中的中间站。据 19 世纪的记录,分布在巴布亚的莫图部落的语言有莫图(Moto)、柯勒布努(Kelepunu)、阿罗玛(Aroma)、南岬(South Cape)、卡巴地(Kabadi)、玛伊瓦(Maiva)和莫图莫图(Motumotu)等方言。^① 20 世纪的调查者发现可以把莫图语分为真莫图语(True Motu)和洋泾浜莫图(Pidgin Motu)。前者的使用者有一千五百人,后者已有二十五万人在使用(1987)。真莫图语有东、西两种方言。

从大陆出发的南岛语经由马来半岛和印度尼西亚诸岛屿到达巴布亚新几内亚,并与较早前到达那里的非南岛系的土著语言接触。如莫图诸方言是在不同语言的基础上发展起来的,这些不同的语言与印度尼西亚(印尼语、爪哇语)、新喀里多尼亚(塔希提语、嫩戈内语)、巴布亚新几内亚本地的语言有密切关系,带有不同于中国台湾、菲律宾和马来半岛诸南岛语系语言的一些词。

印度尼西亚语和亚齐语等,是较晚离开大陆的南岛语,在马来半岛和印度尼西亚地区覆盖了早先的美拉—密克罗尼西亚语族的语言。爪哇语保留着早期美拉—密克罗尼西亚语族语言的某些词汇底层。

无论是南岛语、阿尔泰语、汉藏语,还是印欧语系的语言,从某地向外扩展成为分布于广袤地区的亲属语,总非一日之工。如山东半岛的南岛语直到西周之后才消失,而东南亚地区的南岛语早在商代之时就开始向太平洋地区的岛屿迁徙。即使我们假定一个语系的语言是从最初的一种语言,一再传播一再演变的,她们语音、语法和词汇也会有不同历史层次的变化。因此历史比较中把一个语系的语言不加区分地放到一个平面上比较的方法是不可取的。

① *Grammar and Vocabulary of Language Spoken by Motu Tribe (New Guinea)*, by Rev. W. G. Lawes, F.R.G.S., 1888.

3. 语言的底层和系统的变化

(1) 语言系统的相互作用

两种语言 A_1 和 A_2 ，她们的接触如果产生新的语言或方言 A_3 ，必定与 A_1 或 A_2 比较接近，较少出现混合语。假定语言 A_1 的语音系统为 S_1 ，语法系统为 G_1 ，词汇为 L_1 。跟 A_1 这三个系统分别接触的是语言 A_2 的是 S_2 、 G_2 和 L_2 。

A_3 的语音系统一定与其中的一种，譬如说 A_1 的比较接近，基本词中仍有相当一部分来自 A_2 ，东亚太平洋许多语言保存早期底层成分这一事实已经证明了这一点。当 A_1 、 A_2 、 A_3 与 B_1 、 B_2 、 B_3 两组语言穿插接触影响数千年之后，在它们分布的区域有可能出现几百种不同的语言。这就是美洲印第安人的语言和巴布亚新几内亚的语言显得复杂而难以给予语言发生学归类的原因。

我们已经看到，亲属语和方言之间的语音系统往往具有一定对应关系的相似特点，而不具有同构性，相邻的不同语系的语言之间有着一定的对应关系。

语言的基本词中，或所谓的“核心词”，第一人称代词的单数形式和“眼睛”“舌头”“手”等身体部位词最为稳定，在欧亚和非洲诸语系中我们都会看到这一类基本词作为底层保留在转用的语言中。

较为普遍的是，不同人群的交际语比起产生这种交际语的语言或方言来说语音或形态要简单一些。如北京本地话有一套儿化韵，双音节和多音节复音词有轻声，而官场和知识界流行的普通话没有儿化韵，北京以外地区的普通话没有轻声。北京电台和北京电视台的一些播音员讲的普通话接近于北京土话，通常有较多的儿化韵，也有轻声。东亚的语言中有小舌音的语言、有圆唇辅音的语言和复辅音系统较为复杂的语言通常分布在较少受其他语言干扰的地方。底层语言还对传播语言的系统加以改造，往往使它们的语音系统和形态系统变得比较简单，越是广为传播的语言越是如此。如非洲班图语

和科伊桑语系的语言一样，有一系列的吸气音，但比它们的简单（缺少某些系列），一些基本词有对应关系。上海话的声调系统就比周围的吴方言简单得多，美国黑人社会所讲的英语就不太讲究语法。

（2）语言的底层和谱系关系

语言一旦得以传播，尤其是为原本使用另外的语言的人所使用，马上就成为另一个样子，即接触中语言的系统可能发生调整。

交际语不完全的学习和不完全的交际语的传播使得基本结构来历不同的语言看起来像亲属语。词源关系的研究中我们可以暂时不区别这些情况，把以往有关语言的归属的争论搁置起来，先考虑词的对应及其传播的途径。我们尤其关注那些相隔遥远并有一些语词对应的语言，这种情况大多是早期的迁徙造成的。折中的观点是把这些长期分布在一个地区相互影响有密切关系的语言称为“语言联盟”（allying of languages），该提法实际上还是承认构成语言联盟的语言没有发生学关系。

许多基本词的来历比传统历史语言学所认为的语言历史要悠久得多。一个词从一种语言开始使用就有隐藏的语音、形态和语义的特征，词源的比较可以发现语词在不同语言中的历史，以及从一个区域向另一个区域传播的情况。那些过去在分类上排除在外的语言以及以往认为没有发生学关系的语言的历史关系可以从一些词（不是个别词）的传播历史中得到认识。

亚欧诸语系的亲属语基本词差异甚大，如印欧语的许多基本词来历不同，不能以原始共同语分化的模式解释。如：

	英语	丹麦语	法语	梵语	波斯语
太阳	sun	sol	soleil	divakarah	mihr
月亮	moon	maane	lune	tʃandrah	mah
水	water	vand	l'eau	dʒalam	ab

按照语言衰变理论（语言历史年代学）推测，有共同来历的语言经过二

三万年之后，语言的核心词几乎丧失殆尽，无法说明它们之间是否有亲缘关系。我们观察到的是，随着语言的变化，基本词的形式和代表的意义可能发生变化，而不是消失。我们怎么敢断言一个语言系的语言数万年以来只有一个源头。

历史语言学认为创新是划分方言或亲属语与语系及其亲属语亲疏关系的条件，事实上分布在相邻地区的不同语言和有着共同底层的语言都可能产生使我们认为历史上有共同创新的现象。语言发生学关系的表述以谱系理论为指导，其共同语（common language）分化的模式很难较好地解释亲属语的分歧。

现代语言的谱系关系所要反映的是古代语言的发生学关系，事实上因理论上的认识不同，所依据的材料不同，对语言历史推测的方法不同。特鲁别茨科伊认为：“印欧语言都由同一语言分化而来的设想便不能令人们信服了，因为印欧语的祖先可能原本不一样，只是通过长时间的互相接触与借用才有今天相似的面貌。在讨论印欧原始语问题的时候，我们应时刻考虑到这种可能性。”^①

另外，即使真的有一种原始印欧语存在于五千五百年前的伏尔加河地区，后来扩散到欧洲各地和印度的北部，她也并不是真的原始语。索绪尔以来的结构主义的历史语言学一开始用一种语言单纯分化的设想解释语言的历史，后来的学者索性把这一设想贯彻到底，使人产生许多误会。语言的传播既有欧洲殖民主义时代从欧洲大陆到非洲和美洲这样强势的语言传播，也有诸欧亚大陆的部落和部落联盟时代部落交际语的传播和分化，而在此前的氏族社会时代更多的传播形式应像是殖民主义时代的传播方式。

我们已经有足够的证据说明语言跨语系之间的渊源关系，就有必要突破

^① 特鲁别茨科伊：《关于印欧语言问题的设想》，邹一帆译，石锋、张洪明主编《语言学译林》，世界图书出版公司 2013 年版。

结构主义历史语言学关于诸语系语言原始语的分化的语言传播理论，更多地考虑一下早期的语言接触和相互影响给今天我们所观察到的不同语言或方言带来一致和不一致的情况。

◇◇ 二 语言的演变和传承

1. 语音的历史演变

语言的语音、语法和词汇的传播是相对独立的，微小的变化是语境促成的。言语中的语音变化，如同化、异化、合音、脱落、换位、误听、借用、创新等，可感染整个社团方言乃至语言，诱发语音不同方向的历史演变。语法同样是先有语境的语用，后有方言的变化。

亚欧、非洲各地的语言经历一代又一代的沿用，其语音结构可能从简单变得复杂。在语言转用的过程中又往往从复杂变得简单，后起的语音特点往往容易把原本的对应关系淹没。最令人惊讶的是非洲科伊桑语系语言除了我们常见的一套音系，还有一套各个部位的吸气音，合成一个音系。

(1) 对语音演变的认识

语音的描写都是建立在一定模式上的，注重心理或物理方面的差异。文字的表达，忽略了较多语音信息，以求书写的方便。拼音文字所提取的语音信息也相当有限，如古埃及的文字只表示音节的辅音。文字的读法，是不同程度的还原。历史语言学是依据保留的信息来说明语言历史的因果关系的。构拟，是在一定假设基础上的解释，包括对历史上曾经存在的语音系统的假设。对古代语言的认识是否接近事实，则在于对它们的解释是否符合语言类型学所描写的状态。

阿尔泰、汉藏、南岛和南亚诸语系的语言，有的音系有小舌系列，有的没有。没有小舌音的语言分别以舌根音或喉塞音与亲属语的小舌音对应。那