



语言类型学名篇译丛

语言类型学的 基本方法与理论框架

戴庆厦 汪锋 主编

YUYAN LEIXINGXUE DE

JIBEN FANGFA YU

LILUN KUANGJIA



商务印书馆
The Commercial Press

014038573

H003
13

语言类型学名篇译丛

语言类型学的 基本方法与理论框架

戴庆厦 汪 锋 主编



商务印书馆
The Commercial Press

2014年·北京



北航

C1726555

H003
13

014038273

图书在版编目(CIP)数据

语言类型学的基本方法与理论框架/戴庆厦,汪锋
主编. —北京:商务印书馆,2014
(语言类型学名篇译丛)
ISBN 978-7-100-09675-1

I. ①语… II. ①戴… ②汪… III. ①类型学(语言学)—文集 IV. ①H003-53

中国版本图书馆 CIP 数据核字(2012)第 309122 号

所有权利保留。

未经许可,不得以任何方式使用。

语言类型学的基本方法与理论框架

戴庆厦 汪 锋 主编

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 978-7-100-09675-1

2014年2月第1版

开本 787×1092 1/16

2014年2月北京第1次印刷

印张 27

定价: 59.00 元

前 言

2009 年春天,语言类型学协会(Association for Linguistic Typology)希望在中国推广语言类型学,有一系列计划。其中一项是翻译出版经典的类型学论文,计划出成系列。这一想法得到商务印书馆的支持,我们商量将这一系列定名为《语言类型学名篇译丛》。由时任协会主席的 Anna Siewierska 教授和罗仁地(Randy LaPolla)教授担任学术顾问,帮我们遴选论文篇目。第一辑的文章主要是基础性的类型学论文,便于对类型学有兴趣的研究者把握类型学研究的基本方法和理论框架。书中的人名、语言名称等的翻译尽量遵循学术界的惯例,未做硬性统一。2011 年 7 月译稿基本完成,提交给商务印书馆。Anna Siewierska 教授很高兴看到定稿,并愿意代为联络版权事宜。其时,她正在亚洲休假。我们收到她的最后一封邮件是 7 月 22 日,她答应 8 月中回到英国后即开始着手办理。不幸的是,她于 8 月 6 日在越南的一场车祸中遇难。这是世界语言学界的重大损失。这两年多来,她对于类型学和中国给予了特别的关注,每每念及她的热心和敬业,不禁泫然。我们决定,谨以此丛书纪念我们的良师益友——Anna Siewierska 教授。

《语言类型学名篇译丛》在选文、翻译、审校过程中,得到了很多同道的支持,尤其是北京大学李子鹤同学和商务印书馆朱俊玄编辑承担了后期的编辑整理工作,贡献尤多。本辑收入的绝大多数论文都得到授权免费使用,这是嘉惠学界的功德。中央民族大学“985 工程”中国少数民族语言文化教育与边疆史地研究创新基地和北京大学中国语言学研究中心提供了出版资助和论文版权费用。在此,我们一并表示感谢。

这是《语言类型学名篇译丛》第一辑,后续几辑也在紧张有序的组织中,希望能尽快跟大家见面。

戴庆厦 汪锋

2012 年 8 月 20 日

目 录

回顾与前瞻

- 21 世纪类型学的主要进展 Balthasar Bickel 著;黄成龙译;汪锋校 3
- 十年来类型学与语言学理论之我见 William Croft 著;黄成龙译;潘露莉校 16

原则与方法

- 什么是类型学的共性? Edith A. Moravcsik 著;汪锋译;黄成龙校 29
- 论在无预设范畴下的语言描写和类型学研究
..... Martin Haspelmath 著;黄成龙译;潘露莉校 42
- 语言的取样 Jan Rijkhoff and Dik Bakker 著;林幼菁译 55
- 区域类型学原则 Östen Dahl 著;项梦冰译;戴忠沛校 97
- 走近词汇类型学 Maria Koptjevskaja-Tamm 著;李榕译;乐耀校 117

标记与类型学

- 核心词标记与从属词标记语法 Johanna Nichols 著;孙文访译;姚瑶校 161
- 主动/施动者格标记及其动机 Marianne Mithun 著;林幼菁译 229
- 作格与特征层级 Michael Silverstein 著;孙文访译;李榕校 272
- 区别性格标记的不对称与有生性 ... Andrej L. Malchukov 著;马云霞译;陆丙甫校 335

结构与类型学

- 被动和相关的结构:一个原型分析
..... Masayoshi Shibatani 著;李子鹤、董理译;姜望琪审 359
- 主宾语一致关系的分布与语序类型
..... Anna Siewierska and Dik Bakker 著;李计伟、姚晓东译;龚涛校 390

回 顾 与 前 瞻

21 世纪类型学的主要进展*

Balthasar Bickel 著

黄成龙(中国社会科学院民族学与人类学研究所)译

汪锋(北京大学中文系/北京大学中国语言学研究中心)校

1. 类型学作为一门学科

在过去的一个世纪里,类型学与生成语法共同的目标之一,就是为探寻可能的人类语言的范围提供另一种研究方法,从而对语法的共性理论有所贡献。这种研究范式(paradigm)把语言学上不可能的,而把逻辑上可能的现象排除在绝对共性法则之外,例如,一种语言只在第一人称单数上区分性范畴。

过去十年来,类型学开始从这个目标中解放出来,并从一种方法变成一门成熟的学科,它有自己的研究议题(research agenda)、自己的理论、自身的问题。已经占据类型学中心议题的是了解语言的多样性,而类型学的新目标是解释为什么语言的多样性就是如此——这个目标最初是由尼科(Nichols 1992)提出来的,所谓的群体类型学(population typology),它与群体生物学(population biology)平行。我们不再询问“什么现象有可能?”越来越多的类型学者在追询“什么现象?在哪儿?为什么?”(what's where why)。追询“什么现象在哪儿”是把目标锁定在犹如地理或谱系偏差(geographical or genealogical skewings)的共性优先选择(universal preferences)上,引起概率论正确呈现采样分布(sampled distributions)。追询“为什么”建立在两个前提下:(i) 类型分布是历史产物,(ii) 类型分布与其他分布之间的相互关系。

要理解分布作为历史产物至少可追溯到格林伯格(Greenberg 1965、1978)和葛翁

* 本文最初于 2005 年在美国加州奥克兰举行美国语言学会议后,由 Larry Hyman, Lynn Nichols 和 Johanna Nichols 组织的“美国语言学的类型学-田野评鉴会”上应邀做过报告。我对会议组织者、其他参与者以及进行热烈讨论的听众表示感谢。Dik Bakker, Juliette Blevins, Joan Bresnan, Martin Haspelmath, Edith Moravcsik 和 Sabine Stoll 对初稿提出了非常有帮助的意见,在此我也向他们深表谢忱。这里表达的是我个人的观点,如果表达有误,我个人负全责。

(Givón 1979)早期提出的类型学的历时探求与共时分布的方法。无论共性优先选择还是地理聚群(geographical clusterings)大体上都被看成类型转化(type transitions)和历时过程的产物。^①(也参见 Bybee 1988 和 Hall 1988 对此很有说服力的论证)如今争论的焦点是共性优先选择是由引导(或“选择”)历时结果的优先原则引起的(如 Nettle 1999; Kirby 1999; 或 Haspelmath 1999 所假设的),还是由局部性的优先变化路径引起的(如 Croft 2000; Bybee 2001; Blevins 2004 的论著以及语法化理论的一些论著)。就这两种观点看来,把现时的分布理解成历史的产物,探究的宗旨是变化的概率以及它们背后的原则。

理解类型分布与部分源自其他分布之间的相互关系,要从更广的领域(从社会到认知再到遗传结构上)来了解反映出语言结构中的研究结果趋向于语言结构之间以及与人类学模式之间的系统性的联系。有些发现引起一些理论来预测语言结构的共性优先选择与认知交际共性优先选择之间的密切关系(例如,处理优先选择,正如被 Hawkins 2004 广泛讨论的)。这些优先选择已经处在类型学研究的最先考虑之处,然而,与分布更新、更广的视角相一致的局部结构与局部优先选择在认知或社会相互作用中的相互关联日益得到关注——多数情况下,通过全身心地进行跨学科研究。这首先可以通过过去十年在如下几方面快速增多的研究来说明:语言相对论的研究[如 Lucy 1992; Gumperz & Levinson (eds.) 1996; Niemeier & Dirven (eds.) 2000; Levinson 2003; Bickel 2003; Gentner & Goldin-Meadow (eds.) 2003; Roberson et al. 2005],以及由当地文化支撑具体语言结构的分析[例如 Bickel 2000; Enfield (ed.) 2002; Evans 2003]。此外,心理语言学已经开始通过严格的比较来系统研究和控制处理过程和习得中的类型学变异(对神经语言学的研究,参见 Bornkessel et al. 2005; Bornkessel & Schlesewsky 2006;对语言习得的研究,参见 Berman & Slobin (eds.) 1994; Bowerman & Levinson 2001; Lieven & Stoll in press)。最后且同样重要的是,绝大多数类型分布显示出不同的地理区域模式,而这些模式只能被理解为有悖于人口迁移模式和语言接触,它们通过历史人类学、遗传学、考古学所了解的知识而系统性地透露出来(如 Nichols 1992、1997; Fortescue 1998; Bickel & Nichols 2005、2006; Dunn et al. 2005)。

为了获取和测试分布,类型学者发展了衡量语言之间相似性与差异性的类型变量(typological variables)^②。类型变量(如果是精雕细琢的)以形式上精确的方式适用于跨语言研究,对语言具体结构的分析做清楚的预测,并对相似性和差异性(a tertium comparationis——中间对比项)的本体进行明确的定义。为了解释为什么这些变量值以它们自身的方式分布于世界,类型学家发展了区域偏差理论(theories of areal skewings),或者基于人类学(包括心理学)不同领域中共性优先选择的论点。正如 Dryer (1997)所讨论的,在类型学中发展的变量,或者解释性理论,对特定语言的诸多结构和对这些结构之间

可观察的相似性存在着本体性偏见(ontological commitments)。然而,不像那些旨在定义人类语言绝对条件的工作,类型学并不需要承诺语法中最普遍的设计特征(design feature)之外的普遍性实体(“普遍语法”,简称UG)。此外,类型学理论关注的是分布的概率,不是可能性,因此,它们远远超越了普遍语法所定义的目标“什么是可能的”(Bell 1978; Newmeyer 1998、2005; Dryer 1998; Haspelmath 2004; Nichols 2007)。

如果普遍语法的参项可以直接预测所观察到的分布,例如,通过绝对共性的蕴涵链表现出的条件概率,那么现代类型学与普遍语法不就更有共同之处了吗?如贝克(Baker 2002)提出当三组双向参项有特定值时,从这三组双向参项定义的动词只能在句首位置出现,预测非普遍性 VSO 词序,就违背 SVO 和 SOV 词序,其他词序($2^3 - 1$)与其他参项逻辑上可能的组合相一致。尽管这种方法仍然是有待解决的问题,我还是同意 Baker & McCloskey (2007)的意见。我们从近年来的类型学论著中觉察到其方法存在本质上的缺陷:所观察到的分布实质上受人口历史的影响,通过单独的语法原则试图预测所观察到的分布绝对是有问题的。仍以贝克的例子为据,动词居首词序在世界范围的分布取决于人口历史,因此,显示出明显的区域模式(areal patterns):在太平洋周围地区有一个统计学上高度显著的、历史上相当稳定的频率在增长(Nichols 1992; Bickel & Nichols 2006),在东非大裂谷(Rift Valley)和欧洲西北内陆有些更狭窄的高峰(后者很可能与北非的区域历史连接在一起,见 Gensler 1993)。这使得不仅地理分布,而且纯粹的基本频率都是由历史过程造就的,超越了可能的绝对普遍语法参项模式。在贝克的例子里,如果欧亚与美洲人群有不同的进程,动词居首语言如今就很可能是默认的共性。

由此形成的一般情形是普遍语法研究大体上与现代类型学无关,反过来,如 Newmeyer (1998、2004)和 Haspelmath (2004)所补充的,类型学也与普遍语法研究无关。在普遍语法某些概念下可能如此,但语法的共性理论如果不以类型的充分性(如 Dik 1978; Van Valin & LaPolla 1997; Baker & McCloskey 2007; 以及一般意义上的优选论)为目的,那么在类型学变量的本体意义(而不是它们的分布)上来说,应该把它们从理论体系里除掉。有时也有人建议共性理论的主张要在多种语言中仔细检验(如 Van Valin & LaPolla 1997; Baker & McCloskey 2007),对共性的许多断言的确被新近描写的语言所修正了。[见 the Konstanz Universals Archive and *Raritätenkabinett*, 参照 Plank & Filimonova 2000; Plank & Mayer 2006; Plank (ed.) 2003 所记录的例子]然而,通过分析未描写的语言扩充我们的语料集,不是也不应该只是类型学家的任务——在大量语言消亡的这个时代尤其不应该。

最后,类型学已经发展了一套变量和概率论来解释这些变量的分布,借此从普遍语法研究所用的一种方法转变为回答“什么现象?在哪儿?为什么?”的一门学科。但不是所

有的事情都改变了:最明显的是,过去,类型学家已经发现尽可能接近所观察到的语料和接近田野调查^⑤的变量是很有用的。这首先是由实用性决定的,因为定义非常抽象的变量很难在足够大的语言样本中调查得到,而样本经常只能通过做进一步的田野调查才能完成。然而,假设一些抽象变量理论上也可行,因为抽象变量的定义通常与自身寻求抽象、远离语言多样性的一些普遍语法的模型联系在一起,而不顾有用的人类学的假设^⑥。

2. 是什么把我们引到这儿的

类型学正成为一门成熟学科的明显标志就是这个领域现在已经有了专门的期刊、专业学会、讲座教授以及研究中心。然而,引领类型学新视野的知识发展是什么呢?我提出两个主要的发展源头:即所谓的共性区域性(universal areality)的发现以及方法论上的主要进展。

2.1 共性区域性

自20世纪80年代后期以来,尤其是通过Dryer (1989、1992)和Nichols (1992)的论著,我们已经很清楚,几乎任何类型变量及其结合,在世界上都没有均衡分布。多数分布易受非偶然地理偏差(non-accidental geographical skewing)的影响。举个例子:我测试过一个假设,“动词居尾或词序自由”与“用格或用介词标志区分及物动词的主语和宾语”之间的关联性(参照Konstanz Universals Archive #447)有违于大语料集($N = 383$)世界语言结构地图(WALS)(Comrie 2005 和 Dryer 2005)和AUTOTYP(欧洲类型学)(Bickel & Nichols 2007)^⑦结合的信息,其相关性(association)是显著的[费希尔精确测试(Fisher's exact test),条件P值 $p = .003$]。当通过运用Dryer(1989)或Bickel (2007)所提出的方法调整谱系膨胀效应(genealogical inflation effects)的数据之后,也是如此。这正好与Dryer (2002)早先用一个部分不同且更小的语料集($p < .001$, $N = 257$)所得出的结果相一致。但是,如果我们一个大陆一个大陆地检验语料,这一相关性只在欧亚大陆显示较为显著($p = .04$, $N = 48$),这是一个已知有独立区域性的地区(Jakobson 1931; Nichols 1992)。其他地区的情况则可以通过两个变量的基本频率来预测,这样的例子很容易增加,这也凸显了Dryer's (1989)的告诫,除非由地理因素(以及别的不清楚的因素,见下)所操控,统计上的相关性并不需要支持某个共性优先假设。

即使当一种模式表面上有稳固的共性支撑,如OV词序与后置词之间,或者VO词序和关系子句后置(postnominal relative clauses)之间的相关性,所观察到的分布也只能通过探索区域发展才能被完全理解。因为对这一共性趋势的一些偏离(如伊朗语中OV与前置词,见Stilo 2005,或汉语中VO与关系子句前置,见Dryer 2003),再次倾向于显示区域模式。它们一般集中在Stilo(1987、2005)所谓的缓冲带(buffer zones),即类型特征

相异的扩散区域之间的地带。(Masica's 1976 的伊朗语族案例中的“印度-阿尔泰”对东南亚地区;汉语案例的中印度-阿尔泰对东南亚地区)。因为它在历史发展中与语言内部因素[如分析简化(parsing ease)]竞争,所以,这样的接触压力会影响分布。

这种现象并不令人感到惊奇,我们知道语言接触中语言会转换它们的类型特点(从语言转用到双语现象),语言接触是已知的近似普遍性现象。但是,接触与类型扩散(type spread)当然不是造成分布的唯一因素,另一个重要的因素在主要扩散的边缘有相对孤立的分布[类型学飞地(typological enclaves),见 Bickel & Nichols 2003],这也可能导致偏离共性趋势[或就此而论,宏观区域(macro-areal)],这种偏离常常具有引人注目的历史持久性。Nettle (1999)将增加人口规模作为支持稀有现象的一个因素,假定变量在人口较少的群体中能更快稳定下来。

区域性效应的普遍发现已经在上述提到的类型学的历史转折(historical turn)里产生了两个直接的影响。

首先,遍布的区域性效应清楚表明现今许多类型分布不管在当地还是在全球都只能理解为实际的历史或史前史的结果。反过来看,类型分布提供了从人类学和历史学科比较中发现的大量的历史信号,等待着学者们来解释这些历史信号。

其次,对统计上显著的宏观区域(macro-areas),如太平洋周围地区,或者欧亚地区(Nichols 1992、1997; Nichols & Peterson 1996; Fortescue 1998; Bickel & Nichols 2003、2005) 现有最合理的解释表明,这些宏观区域是早期大规模人口迁移和语言扩散之后幸存下来的分布遗存。但是,如果分布与某些现象一样能长久幸存下来,正如 Maslova(2000)所讨论的那样,基于语言的本质因素,而不是基于人类人口历史的本质因素,这将大大提高我们准备接受共性的门槛。变量的相关性(例如,NP 与 PP 之间的顺序)不仅应在代表性样本[®]统计上显著且独立于已知的地域和谱系关系(Dryer 1989; Perkins 1989),而且必须证明可能存在的显著偏差独立于早期(或甚至是起始)阶段。换句话说,只要相关性能被证明历史上不稳定,在现时所观察的语料里静止分布,我们就可以假定相关性能反映严格的语言共性。这又需要从根本上从历时角度理解类型分布是由什么东西引起的,即不同类型转化的概率性。

2.2 方法论的进展

由于一般技术的进步,毫无疑问,过去十年里大型类型学数据库大量涌现,如今世界范围内有 20 多个研究团队参与开发数据库(欧洲进行的项目,见语言类型学资源中心(LTRC):<http://www.lotschool.nl/Research/ltrc>),以及世界语言结构地图(WALS)[Haspelmath et al. (eds.) 2005],项目背后大量的国际合作已经孵化成更多的数据库成果。大语料集(large dataset)几乎无一例外的显露出语言共性有例外现象,这与新描写的

语言大幅度增长以及借助于重要概念的讨论(如 Dryer 1998; Croft 2002: Chapter 8)密不可分,实际上已经废除了绝对共性概念和不可能性。现代类型分布的研究包含统计学的方法,从相关性测试(association tests)(参照 Cysouw 2005)到多变量测量方法(multivariate scaling methods)(如 Levinson 等 2003; Croft & Poole 2004)。在更为重要的历史研究方面,类型学已经引介了新语料聚群方法(new data aggregation methods)[如等值线方法(the isopleth method),见 Van der Auwera 1998]、统计测试手段(statistical testing strategies)[如预测区域性理论(predictive areality theory),见 Bickel & Nichols 2006]、稳定性矩阵(stability metrics,见 Wichmann & Kamholz 2007)以及来自生物学的种系发生方法(phylogenetic methods from biology)。目前尝试结合地理信息系统(geographical information systems)为此领域将来的进一步研究带来希望。

所有这些方法都有一个共同的特点就是他们对变量采用狭义定义,而不是用古典整体类型学(classical holistic typology)的粗略类型(如主动语“active language”、黏着语“agglutinative language”)或者语言联盟(*Sprachbund*)等范畴概念。一般假设如果语言结构之间,或者语言结构与地理之间有大规模的关联,这些关联存在于独立衡量的变量之间的概率关系[因此,例外缠身(exception-ridden)],则这些相关性并不需要从绝对的定义或者“理想”的类型来推断。在类似的层面上,现代类型学已经远离整个语言的类型化,相反,是把个体结构模式(结构、规则、限制等)拿来作为比较对象。语言多样性不是从大的类型概念,而是从很多套精细的变量来把握。

这种变量分析方法引起了与其他历史人口科学共同存在的统计问题——最明显的是,我们只能接触到人类使用过的所有语言中 1% 的语言,因此,现有的人口与其历史成长分布偏向始终在我们的采样中被过度呈现。此外,在类型学采样中,除了相对新近的遗传效应外,我们通常尝试对已知谱系多样性的规模做穷尽性和平衡性的涵盖,以便尽可能获知共性优先选择或者区域人口史的信息。然而,穷尽的采样使古典的统计方法(classical statistical methods)在数学上毫无意义。类型学家对这些问题的答复是,现在采取了蒙特卡罗(Monte-Carlo)方法和精确方法(exact methods),第一步先对编码进行随意性基础上的可靠性测试(Janssen et al. 2006)。不像古典方法从正态分布的群体(normally distributed population)中预先假定随机采样,这些方法不支持对人类所有语言潜在的群体进行统计推断。所有统计推断仅限于身边现有的样本,零假设测试(the null hypothesis of the test)就是在某些语料集里所观察到的偏差可以从相同语料集的差额总数(the margin totals)中预测——并不是所观察到的语料集毫无偏差地在一个总的群体里进行随机采样。

然而,这正好适合新近强调的“什么现象?在哪儿?”这也再次向运用类型学探索人类

语言的绝对范围发起了挑战。

3. 未来挑战：定量类型学和个体结构之分析

大规模定量研究经常受到正在寻求细致地比较少量语言里个体语言结构的类型学家的怀疑。然而,对我而言似乎是,如今定量研究的进展事实上引起了意想不到的兴趣。

如上所述,研究类型学分布中的区域和谱系信息需要数据库中设置更多的以及更细致的变量组。然而,传统的描写变量,如“有组并”(incorporation)与“无组并”,分类太粗糙,把大的变量组归并在一起,它们自己常常有独立的分布和历史特性,例如,关注所指的特指性(specificity)、话题性(topicality)以及所指的基本性(cardinality),形态与音系的一致性、短语扩展的可能性、语法关系状态、论元身份(argumenthood)等等。这与田野语言学家试图确定新的现象时反复得出的结论完全相同。一个相当普遍的经验是,新现象并不完全与另一个语言中,或教科书上所谓的“组并”(incorporation)相符。事情看起来常常相似,但是决不完全相同。^⑦

除了完全放弃比较研究,传统上对此问题的回答主要在于试图敲定对“组并”(或诸如此类)的普遍意义上的(技术上称作“跨语言适用”)定义,这常常引起热烈的争论。但是,如果这样的定义不从每个人都同意的第一原则(如任何语言描写的有名的理论中提出的第一原则)来推断得出,那么在这样的定义里还感觉保留了任意性,更糟的一种感觉是,学识渊博的语言学家碰巧研究的那些语言最后成为最接近普遍性定义(然后常常说提供了某种现象的“原型”)的。还有更糟的是,通过将许多不同的语言中的结构在同一定义下归入一类,类型学不去检验多样性,而以减少多样性告终。最终,这个学科击败了自己的目的。

走出这条死路,一个很有希望但仍在探索中的出路与定量类型学所需要的东西恰恰相同:精细的变量(fine-grained variables)。要确定 L 语言里 S 结构是不是组并(incorporation),我们将一组最大化精细的变量编码为 S_L ,正好能足够收集我们对 S_L 所知道的全部内容[如有无泛指参照(generic reference)、是否禁止移位(prohibits permutation)、有无词干类次范畴化(subcategorizes for a stem class)、是否一定与另一个词干毗邻、是否满足论元位置、是否控制一致关系,或如果某个语言开始没有一致关系则“未知”等等]。因为这些是最大化精细的变量,这些概念争议不大。例如,我们不用“词缀”这个术语,如果用这个词,就需要查阅这个词引起的所有凌乱的内涵,用诸如“词干类次范畴化”,或者“与毗邻成分进行音段交互作用的成分”等术语完成我们的定义。这样的变量允许获取多样性,而不是忽略多样性,并且,这样的变量更有机会在许多不同语言中以可复制的形式编码。精细的变量形成了适于研究各种结构在世界上如何分布的恰当输入材料,同时,这

些精细的变量正好提供了分析个体结构非常好的工具,远远超越了无用的命名练习。

此外,一旦有这样详细的编码可利用,就可以用其他学科,从生物学到经济学已经使用了数十年的相同的工具来直接检验多样性:从语言具体结构所编码的一系列变量的一组矩阵中,我们很容易测试变量之间的共变,以及计算结构之间的相似性/非相似性矩阵。这一矩阵可用于日益丰富的聚类库(arsenal of clustering)和比例运算技术(scaling techniques),这一矩阵如今也可用于其他学科(见 Cysouw 2006 的最新研究),这样便于发现一些语言特定的结构之间是否比其他更相似,例如,形成跨语言组群(crosslinguistic clusters)。如果有这样的组群(聚类),它们就配得上基于经验的“原型”的头衔,而且很显然,需要理论上做出解释(但是,当然无法保证一定有组群)。

这就将责任从争论普遍性定义转向发展日益精确的描写性变量。最后,田野调查的任务就完全融入到先前每个有关大规模分布的研究项目中。这也确证了早先进行的观察,类型学中发展起来的变量是而且应该是密切对应于对田野调查基础上的基本分析的那些需求。

4. 结语

现代类型学是以发展变量来获取语言内部或者跨语言的结构相似性和差异性的一门学科(定性类型学),它探索这些变量的分布中所存在的组群和偏差(定量类型学),并且提出解释这些组群和偏差的理论(理论类型学)。类型学的最终目标就是理解“什么现象?在哪儿?为什么?”(what's where why?)这就要弄清楚类型学主要的贡献并不局限于狭义理解的认知科学。21 世纪类型学的目标牢牢根植于更广阔的人类学视野:帮助理解一个关键的社会组织机构的变体在世界上如何分布,什么普遍原则和什么意外事件是这些分布的历史根源。

附注

① 一些类型学者(如 Plank & Schellinger 2000)保留历时共性(共性优先选择)这个术语,把蕴涵共性直接翻译成历时[例如:OV 型语言偏向于后置词(而不是前置词)翻译成“O-V 偏向于发展成 NP-P(而不是 P-NP)”。对照起来,类型转化导致从共性不优先选择到优先选择模式的所有历史因素(例如从 VO 顺序与前置关系子句到 OV 顺序带后置关系子句)。

② 根据出版物,这是类型学研究中最显著的日常事务,同时,也是研究其他现象的基础,参见第 3 节。

③ 我们所操作的标准决定在一个特定语言里某些结构的可观察的手段,这并不意味着“表层”对立干“底层”。类型学犹如其他语言学学科一样与结构模式有关,而不是未分析的表层列表,或者粗糙的意义臆测。(参见 Nichols 2007; Hyman 2007)

④ 去除抽象相似的趋势,普遍语法派生的讨论或语言结构的衡量也可以在心理语言学以及语言人类学的特征中观察得到。

⑤ 就语料个案而言,在 WALS 和 AUTOTYP(欧洲类型学)的 100 种语言里,只有一种语言不相符。在语

序方面,238 种语言里有 9 种语言不相符,即低于 4%。这里所报告的结果在不相符的个案中选择哪种编码并不敏感。

⑥ 其前景并不像 20 世纪所普遍相信的那么明朗:在世界语言结构地图(WALS) [Haspelmath et al. (eds.) 2005]里,贝克(Bakker 2004)发现变量中所有逻辑上可能相关的大体上低于 1%,这在统计上重要,但只有小部分在语言学上有意义。

⑦ 近年来“半组并”(quasi-incorporation)的真实例子,见贝克尔(Bickel 等 2007)。

参考文献

- Baker, Mark C. (2002) *The Atoms of Language: The Mind's Hidden Rules of Grammar*. Oxford: Oxford University Press.
- Baker, Mark C. & Jim McCloskey (2007) On the relationship of typology to theoretical syntax. *Linguistic Typology* 11: 285—296.
- Bakker, Dik (2004) LINFER and the WALS database. Paper presented at the “Workshop on Interpreting Typological Distributions”, Leipzig, December 2004.
- Bell, Alan (1978) Language samples. In Greenberg, Ferguson & Moravcsik (eds.) (1978). 123—156.
- Berman, Ruth Aronson & Dan I. Slobin (eds.) (1994) *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, N. J.: Erlbaum.
- Bickel, Balthasar (2000) Grammar and social practice: The role of “culture” in linguistic relativity. In Niemeier & Dirven (eds.) (2000). 161—191.
- (2003) Referential density in discourse and syntactic typology. *Language* 79: 708—736. *Typology in the 21st century* 249.
- (2007) A refined sampling procedure for genealogical control. *Sprachtypologie und Universalienforschung*.
- Bickel, Balthasar & Johanna Nichols (2003) Typological enclaves. Paper presented at the 5th biennial conference of the Association for Linguistic Typology, Cagliari, September 2003. <http://www.uni-leipzig.de/~autotyp/download>.
- (2005) Inclusive/exclusive as person vs. number categories worldwide. In Elena Filimonova (ed.) *Clusivity*. 47—70. Amsterdam: Benjamins.
- (2006) Oceania, the Pacific Rim, and the theory of linguistic areas. *Berkeley Linguistics Society* 32. Available at <http://www.uni-leipzig.de/~bickel/research/papers>.
- (2007) The AUTOTYP database. Electronic database. <http://www.uni-leipzig.de/~autotyp>.
- Bickel, Balthasar, Martin Gaenszle, Arjun Rai, Prem Dhoj Rai, Shree Kumar Rai, Vishnu S. Rai & Narayan P. Sharma (Gautam) (2007) Two ways of suspending object agreement in Puma: Between incorporation, antipassivization, and optional agreement. *Himalayan Linguistics* 7: 1—18.
- Blevins, Juliette (2004) *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Bornkessel, Ina & Matthias Schlesewsky (2006) The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113:

- 787—821.
- Bornkessel, Ina, Stefan Zysset, Angela D. Friederici, D. Yves von Cramon & Matthias Schleesewsky (2005) Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage* 26: 221—233.
- Bowerman, Melissa & Stephen C. Levinson (eds.) (2001) *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press.
- Bybee, Joan (1988) The diachronic dimension in explanation. In Hawkins (ed.) (1988). 350—379.
- (2001) *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Comrie, Bernard (2005) Alignment of case marking. In Haspelmath et al. (eds.) (2005). 398—405.
- Croft, William (2000) *Explaining Language Change: An Evolutionary Approach*. Harlow: Longman.
- (2002) *Typology and Universals*. Cambridge: Cambridge University Press.
- Croft, William & Keith T. Poole (2004) Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. Manuscript, Center for Advanced Studies in the Behavioral Sciences, Stanford.
- Cysouw, Michael (2005) Quantitative methods in typology. In Gabriel Altmann, Reinhard Köhler & Rajmund G. Piotrowski (eds.) *Quantitative Linguistics: An International Handbook*. 554—578. Berlin: Walter de Gruyter.
- (2006) New approaches to cluster analysis of typological indices. Manuscript, Max Planck-Institut für evolutionäre Anthropologie, Leipzig.
- Dik, Simon C. (1978) *Functional Grammar*. Amsterdam: North-Holland.
- Dryer, Matthew S. (1989) Large linguistic areas and language sampling. *Studies in Language* 13: 257—292.
- (1992) The Greenbergian word order correlations. *Language* 68: 81—138.
- (1997) Are grammatical relations universal? In Joan Bybee, John Haiman & Sandra A. Thompson (eds.) *Essays on Language Function and Language Type Dedicated to T. Givón*. 117—143. Amsterdam: Benjamins.
- (1998) Why statistical universals are better than absolute universals. *Chicago Linguistic Society* 33: 123—145.
- (2002) Case distinctions, rich verb agreement, and word order type (comments on Hawkins' paper) *Theoretical Linguistics* 28: 151—157.
- (2003) Word order in Sino-Tibetan languages from a typological and geographical perspective. In Graham Thurgood & Randy J. LaPolla (eds.) *The Sino-Tibetan Languages*. 43—55. London: Routledge.
- (2005) Order of subject, object, and verb. In Haspelmath et al. (eds.) 2005. 330—334.
- Dunn, Michael J., Angela Terrill, Ger P. Reesink, Robert A. Foley & Stephen C. Levinson (2005) Structural phylogenetics and the reconstruction of ancient language history. *Science* 309: 2072—2075.