

20世纪中国
语言学方法论
—— 研究 ——

陈保亚 著



商务印书馆
The Commercial Press

20世纪中国 语言学方法论 研究

陈保亚 著



创于1897

商务印书馆
The Commercial Press

2015年·北京

图书在版编目(CIP)数据

20世纪中国语言学方法论研究/陈保亚著. —北京:
商务印书馆, 2015

ISBN 978-7-100-11607-7

I. ①2… II. ①陈… III. ①语言学—方法论—
中国—20世纪 IV. ①H0-3

中国版本图书馆CIP数据核字(2015)第232523号

所有权利保留。

未经许可,不得以任何方式使用。

20世纪中国语言学方法论研究

陈保亚 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

北京市艺辉印刷有限公司印刷

ISBN 978-7-100-11607-7

2015年12月第1版 开本 787×960 1/16

2015年12月北京第1次印刷 印张 49 1/4

定价: 108.00元

本书是作者所主持的一系列研究项目的成果之一，感谢各基金会的支持：

1994：汉藏语系的发生学研究（国家社科基金项目，批准号 94 CYY 002）；

1996：二十世纪中国语言研究方法论（国家教委人文社会科学研究规划项目，批准号 1996—1998, 96JC740001）；

1998：近二十年来中国语言学理论建设研究（国家社会科学基金项目，批准号 98BYY001）；

2001：网络汉语研究（北京市社科基金重点项目，批准号 03 BJA WY 002）；

2003：汉越语音对应语素的时间层次（教育部博士点基金项目，批准号 03JB740008）；

2004：汉台对应语素数据库及其时间层次标注（国家社科基金重点项目，批准号 04A Y Y 001）；

2006：基于严格语音对应同源语素的汉藏语系数理谱系分类（教育部基地重大项目，批准号 2006JDXM007）；

2008：基于汉藏语言有序对应数据库的亲缘关系计量研究（国家自然科学基金项目，批准号 60773159）；

2011：基于系统语音对应的核心词分阶及建模研究（教育部基地重大项目，批准号 11JJD740004）；

2013：基于严格语音对应的汉语与民族语言关系字专题研究（国家社科基金重点项目，批准号 13AZD051）；

2014：基于中国语言及方言的语言接触类型和演化建模研究（国家社科基金重大项目，批准号 14ZBD102）。

目 录

引言	1
1 马建忠的语法理论	5
2 中国结构语言学	18
2.1 语法理论	19
2.1.1 词类论	19
2.1.1.1 分布观念的形成	19
2.1.1.2 一线制和双轴制	22
2.1.1.3 词类的意义标准和三品说	27
2.1.1.4 鉴定字的提出	31
2.1.1.5 分布理论的全面展开	38
2.1.1.6 鉴定字与英汉词类的有阶比较	45
2.1.1.7 语类划分的相对性与分布特征推导	61
2.1.2 单位论	69
2.1.2.1 同形替代法与词的提取	70
2.1.2.2 扩展法与词的提取	78
2.1.3 结构论	83
2.1.3.1 同构	84
2.1.3.2 句法结构的初始概念	87
2.1.3.3 结构关系的判定	95
2.1.4 层次论	104
2.1.5 词组本位论	109
2.2 音系理论	113
2.2.1 音位的相对性	113
2.2.2 音节中心观念与音节的层次	118

2.2.3 调位论	125
2.2.4 音系和语法的关系	128
2.3 中国结构语言学的地位和去向	136
3 基于语文学的汉语历时研究	149
3.1 系联法	151
3.2 对音与古音构拟	153
3.3 谐声原则	163
3.4 空格论与内部构拟法	164
3.5 右文说	168
4 汉藏语言发生学	174
4.1 汉藏语言谱系划分分歧的产生	175
4.2 同构标准的采用	176
4.3 对应标准的转向	180
4.4 同源标准的探索	184
5 异质语言研究的兴起:语义关系研究	191
5.1 异质语言观的转向	191
5.2 语义研究的必要性	194
5.3 语义结构关系	197
5.4 变换分析与转换层面	202
5.4.1 变换分析的必要性	204
5.4.2 变换分析的平行性原则	220
5.4.3 空语类和深层结构	224
5.4.4 汉语的支配与约束	230
5.5 语义结构关系的扩展	237
5.5.1 格语法和配价语法的研究	237
5.5.2 形容词和名词的配价	246
5.5.3 论元的判定	248
5.6 语义特征研究	250
5.7 语义指向分析	258
5.8 语法研究的三个层面	261

5.9 IA 与 IP: 汉语与中国结构语言学	269
6 非线性音系研究	274
6.1 对非线性的认识: 基本形式和派生形式	275
6.2 声调的调域和音高	277
6.3 连读变调的复杂条件	284
6.4 字调与词调	287
6.5 节律研究	289
6.5.1 莫拉、音节和音步	290
6.5.2 韵律词和韵律短语	295
6.5.3 节律层阶的表现形式	297
6.5.4 节律和句法研究	299
6.6 语调研究	303
6.7 优选论与汉语音系学	304
7 字本位论	308
7.1 语素与字	311
7.1.1 对比原则的必要性	312
7.1.2 字与语素的切分	314
7.1.3 字与语素的归并	324
7.2 字化现象	327
7.2.1 分析型字化	327
7.2.2 综合型字化	330
7.3 字与词	333
7.3.1 自由标准与词的提取	334
7.3.2 平行周遍对比与规则字的提取	336
7.3.3 单层单位说的不充分性	344
7.3.4 语法单位的类型差异	346
7.4 平行周遍原则	349
7.4.1 自由与平行周遍条件	349
7.4.2 自由与规则的关系	354
7.4.3 平行周遍条件与类比	356

7.4.4 汉语词法和句法的对齐	361
7.5 语义语法	366
8 词汇扩散论	381
8.1 词汇扩散与音变方式	382
8.2 词汇扩散与音变条件	385
9 叠置式音变论	392
9.1 音系叠置与音变规律	393
9.2 音系叠置和差异原则的冲突	400
9.3 语法系统的叠置	412
10 自组织论	415
10.1 结构协合与变异的序	419
10.2 组合协合与聚合协合	431
10.3 多层次干扰	436
10.4 结构与音变规律	444
11 语言接触的无界有阶论	453
11.1 母语匹配与接触规律性	455
11.2 语言转换与语言转型	463
11.3 互协同构与无界有阶性	468
11.4 对应:同源的必要条件而非充分条件	473
12 语素关联对应与层次	478
12.1 关联对应分析的必要性	478
12.2 关联对应的单位	480
12.3 建立对应的基本形式:语素与字	480
12.4 对应的语义实证原则	481
12.5 完全对应和不完全对应	483
12.6 完全对应的周遍性	487
12.7 一致对应	499
12.8 多重对应与时空层次区分	501
12.9 普遍对应与非普遍对应	504
12.10 完全一致对应与普遍一致对应	505

12.11 核心一致对应	507
12.11.1 核心字、核心语素、核心语符和核心词	507
12.11.2 核心一致对应语素集	508
12.11.3 核心一致对应语素集的重要性	510
12.11.4 声调在核心一致对应中的层次判定作用	511
12.12 有序对应关联表	512
13 关系词有阶分析	521
13.1 关系词比较的基础:现代语言还是构拟语言	521
13.2 绝对有阶分析	524
13.2.1 绝对有阶分析程序	525
13.2.2 绝对有阶分析的可行性与限度	535
13.2.3 核心词自动分阶问题	538
13.3 相对有阶分析	540
13.3.1 单向相对有阶分析	541
13.3.2 双向相对有阶分析	545
13.3.3 内核语符相对有阶分析	550
13.4 语素类聚有阶分析	554
13.4.1 考古证据和语素文化类聚分阶	556
13.4.1.1 类聚语素的对应和考古分布	556
13.4.1.2 语素的时空判定和传播过程	561
13.4.1.3 语素类聚考古有阶分析方法	565
13.4.2 自然序列和语素类聚分阶	566
13.4.2.1 数词对应的语源问题争论	566
13.4.2.2 汉台数词类聚的有阶分布	567
13.4.2.3 数词分布语序的有阶差异	569
14 历史比较与恰当重构	573
14.1 比较的充分性	574
14.2 弱重构、过度重构与恰当重构	579
14.3 构拟的系统格局和音变机制	581
15 基于异质性的《切韵》研究	585

15.1	从《切韵》同质论到《切韵》异质论	585
15.2	比较法的作用与限度	591
15.3	语文学和比较法的互补	595
15.4	构拟的协调性和充分性	600
15.4.1	赋值的充分性	605
15.4.2	内转与外转	606
15.4.3	等	607
15.4.4	开合	610
15.4.5	摄	611
15.4.6	《切韵》音系韵母协调构拟表	611
16	基于异质性的上古音构拟	623
16.1	上古音和中古音的关系	625
16.2	上古韵母	626
16.2.1	韵腹和押韵原则	626
16.2.2	介音和等	628
16.2.3	重组	629
16.2.4	上古韵部拟音修订	634
16.3	上古声母	635
16.3.1	复辅音	635
16.3.2	腭化	640
16.3.3	日母和泥母的对立	642
16.3.4	上古声母表	643
16.3.5	上古声母及其和中古声母的关系	644
16.4	上古声调	645
16.4.1	从音变规则看上古四声	645
16.4.2	阴入通押	646
16.4.3	去声	647
16.4.4	上古类声调及其和中古声调的关系	649
17	汉藏区域语言的谱系关系	651
17.1	汉藏关系字关联对应与有阶分布	651

17.2 澳台关系字关联对应与有阶分布	659
17.2.1 首音、腹音、尾音的切分和关联对应	659
17.2.2 傣印词根关联对应表	662
17.2.2.1 傣印首音对应	662
17.2.2.2 傣印腹音对应	663
17.2.2.3 傣印尾音对应	665
17.2.3 傣印完全对应有阶分布	666
17.2.4 傣印宽式不完全对应有阶分布	667
17.2.5 同源关系的可传递性和澳台语源关系	670
17.3 汉藏区域语言谱系关系	673
17.4 谱系分类的原则	674
17.4.1 共享创新法	674
17.4.2 词源统计法	676
17.4.3 严式词源统计法和谱系树绘制	679
17.4.4 谱系分类存在的问题	681
结语	683
符号和术语说明	689
中西人名对照	697
人名译名对照西文排序	697
人名译名对照汉语排序	701
参考文献	704
后记	773

引言

一种方法论就是一种理论，语言学方法论就是语言学理论。本书所理解的中国语言学方法论指以汉语、中国境内民族语言以及跨境民族语言为研究对象所涉及的方法论或理论。

中国的语言学研究有个传统，很多受过严格理论训练的学者并不轻易正面展开方法论或理论的讨论，但在材料的观察和分析过程中都带有特定的方法论。比如赵元任的《现代吴语研究》(1928)就是这方面的典型，它的方法论基础就是描写语言学和历史比较语言学。实际上对任何材料的观察和分析都有特定的方法论。材料需要理论和方法来照亮，从语言学家和被调查人对材料的解释过程中就可以看出这一点。被调查人往往都比调查者更熟悉所面对的语言，但通常都不可能像语言学家那样能够提取出材料的规律。扎实而有规律的描写往往蕴涵了坚实的理论基础。

关于 20 世纪中国语言学方法论已经有不少论著做了介绍，本书将把重点放在对方法论的研究和论证方面。由于中国语言学研究把方法隐藏在材料背后的这种特殊传统，分析中国语言学方法论就显得非常困难。这就需要有面向材料分析的评价标准，而不是简单的介绍和梳理。只有经过大量材料的分析，我们才能够断定隐藏在材料背后的方法论的得失。断定一种方法的得失可以有很多角度，可以看这种方法和某个著名语言学家的说法是否一致，可以看这种方法和某个重要理论是否一致，也可以看这种方法是否跟当前趋势一致。这些都是值得参考的角度，但本书评判一种方法的得失主要是观察这种方法对材料的解释力，这就需要在讨论一种方法的得失时分析很多材料，这样做可能会使某些问题的叙述过分专门化，过分烦琐，但这样展开分析可以使方法论进展的线路更具有实证性，同时也便于说明我们提出不同方案或模型的理由。

方法论就是找出材料的规律的理论。任何一种方法都试图用规则和模型有序地解释材料。比如汉语中“差点儿没”有两种相反的意思：

肯定意义：

差点儿没买着。（买着了）

差点儿没考上大学。（考上了）

差点儿没修好。（修好了）

差点儿没赶上车。（赶上了）

差点儿没见着面。（见着面了）

否定意义：

差点儿没摔死。（没摔死）

差点儿没掉进水里。（没掉进水里）

差点儿没输了。（没输）

差点儿没错过机会。（没错过机会）

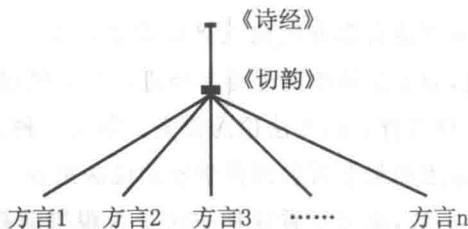
差点儿没落榜。（没落榜）

仅仅描写“差点儿没”有两种相反的意思是不够的，还需要给出理论上的解释，即在什么情况下是肯定意义，在什么情况下是否定意义。在很长一段时间，人们对“差点儿没”都没有合理的解释。后来朱德熙（1959.9）用说话者“是否企望发生”的规则解释了这个问题：企望发生的是肯定，不企望发生的是否定。提出这种解释本身就是方法论的进展，因为这一解释涉及了心理原则，以前没有人做过这样的解释。

有时候面对相同的材料可以提出不同的方法来解释。怎样判定其中哪一种方法解释力强？可以拿《切韵》同质论和《切韵》异质论来说明这个问题。

《切韵》是联系上古、近代音系和现代方言音系的重要著作。中国地域广袤，《切韵》记录的是何处的语音系统？这是历史语言学家面临的一个根本问题。面对这一问题形成了两种解释：一种把《切韵》看成一时一地的音系，可以称为同质论；一种把《切韵》看成包含了南北方言特点的综合音系，可以称为异质论。

同质论假定《切韵》音系是从《诗经》音系发展下来的，现代方言又是从《切韵》发展下来的，可以概括成这样的模式：



Karlgren 是《切韵》同质论的主要代表，他在《中国音韵学研究》（高本汉，1915—1926）中认为《切韵》代表了公元 6 世纪左右的长安音系。这种观念也反映在 Maspero 的《唐代长安方音》（1920）以及 Dragunov 的《对于中国古音重订的贡献》（龙果夫，1928）等论文中。按照这种观点，以《诗经》为代表的上古音发展成为以《切韵》为代表的中古长安音，中古长安音又分化成现代的诸多方言。

同质论在方法论上会碰到两个难以解释的问题。

从《切韵》往上看,即把中古的《切韵》和上古的《诗经》比较,同质论会和语音演变的规律性产生矛盾。根据《诗经》的押韵,最多可以分成 31 个韵部。这是不计算声调的统计。但《切韵》系统的韵部,即使不算声调,也有 61 部。也就是说,从《诗经》到《切韵》,韵部主要是一个分化过程,而且分化的数量相当大,因此韵部增加了很多。问题还不在于分化的数量。如果分化是有条件的,分化得再多也符合语音演变的规律性。同质论的矛盾就在于,从《诗经》到《切韵》,很多分化是找不到条件的。如果只是几个韵部找不到分化的条件,尚可说我们的研究深度不够,条件尚未找到。现在的难点是,从《诗经》到《切韵》,很多韵部的分化都找不到条件。这可能不是条件尚未找到的问题。找不到分化条件,而说从《诗经》到《切韵》产生了分化,这和语音演变的规律性有矛盾。

再从《切韵》往下看,把中古《切韵》和现代方言比较。如果《切韵》代表了中古音,而中古音又是长安音,就必须假定现代中国的几大方言都是从长安音分化出来的。但是从方言存在的一些特殊音类对立看,闽方言可能在中古以前就形成了,不太可能是由长安音分化出来的。

张琨在《原始汉语韵母系统与〈切韵〉》(1972)中开始对《切韵》的性质做了新的解释,可以把这种解释作为《切韵》异质论的代表,其含义可以概括如下:



即《切韵》不完全是由《诗经》发展而来的,早在《诗经》以前,原始汉语就分化成为北方汉语和南方汉语,《切韵》既包含了《诗经》的北方汉语系统,也包含了《诗经》以外的南方的汉语系统,《切韵》所反映的语音系统是综合系统,是一个异质系统。这样比较容易解释《切韵》和《诗经》的复杂关系,也比较容易解释《切韵》和现代方言的复杂关系。

面对相同的材料,两种方法的解释力可能相同,这时需要诉诸简明性原则。比如对汉语儿化词的分析有两种方法。一种是线性切分,即把 kar^{51} (盖儿) 这样的词进行线性切分,得到 ka^{-51} 和 r 两个成分, ka^{-51} 是 kai^{51} (盖) 的语素变体,分布

在-r 前面,-r 是“体词·小称”语素。这是美国结构语言学描写北京话时采取的方法。另一种分析方法是先不直接做表层的线性切分,而用改写规则描写,即先假定底层语素 kai⁵¹(盖)和 r 的存在,当这两个底层形式结合而变成可以观察到的表层形式时,有一条语音规则控制着改写规则:

$$-ai + r \rightarrow ar$$

这是很多中国学者描写北京话时隐藏在背后的一种方法。显然,这两种方法都可以描写儿化,但线性切分法不符合简单性原则,因为儿化词很多,线性切分法必然面临大量语素变体,而用改写规则来描写儿化词要简单得多。

评价一种方法的强弱还涉及可操作性问题。在词类理论中,分布(distribution)标准比意义标准好,就在于分布标准有可操作性,意义标准可操作性较低。比如^①:

金	银	铜	铁	锡
* 金是金属	* 银是金属	铜是金属	铁是金属	锡是金属
* 一块金	* 一块银	一块铜	一块铁	一块锡
* 买金	* 买银	买铜	买铁	买锡

如果根据意义把“金、银、铜、铁、锡”都看成事物,把它们都看成名词,就不能说明为什么“金、银”在组合上要受到限制,而“铜、铁、锡”不受限制。坚持意义标准的人可以说“金、银”的语法意义和“铜、铁、锡”的语法意义有区别,但怎样判定这种区别,没有可操作性。如果根据分布,就有可操作性,“金、银”可归入区别词,“铜、铁、锡”可归入名词。

在解释力强、简单性、可操作性等基本标准下,我们将展开具体材料分析,研究 20 世纪中国语言学方法论,说明建立在中国境内语言研究基础上的语言学方法对普通语言学的意义,揭示存在的主要问题,并尽可能在我们自己已有的研究基础上提出自己的研究方法或思路。

① 左上方带“*”的言语片段是不可接受的言语片段。

1. 马建忠的语法理论

《马氏文通》是中国第一部语法著作,尽管语料的时间跨度很大,但仍然是一部共时研究著作。共时研究主要描写要素在同一时间中的关系,而不管要素在不同时间中的关系。

任何学科的初级阶段都需要把所研究的对象限定在一个确定的范围内,使研究的对象有一个稳定的、静态的、没有变异的、可观察的基础。这就是研究对象同质化。同质化的目的是使研究者容易达成共识。语言研究的初级阶段需要对语言同质化,中国语言研究的初级阶段也需要对语言同质化。在同质化的基础上对语言所做的研究可以称为同质语言研究。

马建忠的《马氏文通》(1898)标志着中国现代语言学的开始,同时也标志着中国同质语言研究的开始。后来 Karlgren 的《中国音韵学研究》属于历时研究范畴(高本汉,1925—1926)。在共时研究中,为了使研究对象高度同质化,《马氏文通》把研究对象限制在汉代及汉代以上,汉代以后只取韩愈的文章,因为韩愈的文章和汉代以前的文章比较接近。在历时研究中, Karlgren 的《中国音韵学研究》把研究对象集中在《切韵》上,认为《切韵》代表了中古音,是一个内部一致的系统。

马建忠和 Karlgren 分别从共时和历时的角度开创了中国同质化语言研究的范式。20 世纪上半个世纪的中国语言研究,正是在同质化过程中展开的。

共时语法研究的根本任务是找到有限的单位和将有限的单位组合成句子的规则。马氏找出了最基本的单位:字。为了说明字的组合,马氏引入了三个相当重要的概念:相当于词类的字类、相当于语义格(论元)的“词”、相当于句子成分或语法格的“次”。因此,马氏在方法论上的重要特点是:

1. 提出了基本单位“字”。
2. 提出了汉语字类的学说。
3. 初步区分了语义结构关系和语法结构关系,提出了两套句子成分。

纵观百年来语法理论的变迁,各种语法体系都是围绕这几个基本问题展开的。体系总是建立在少数几条基本原则的基础上的。很多人认为马建忠模仿拉丁语法体系,这只是从语法体系上去看问题。从方法论上看,马建忠一开始就有自己的特点。把“字”作为基本单位,提出汉语的字类学说,并不完全是模仿印欧语法。

马氏的字类学说对后来的汉语语法研究影响很大。在马氏看来,字类问题是语法的核心问题。马氏认为:

惟字之在句读也必有其所,而字字相配必从其类,类别而后进论夫句读焉。(《马氏文通·例言》,P15)

这里的论述涉及了 20 世纪语法研究方法论的两个根本概念:单位的分布(位置)和单位的类(语类)^①。这两个概念也是 Boas(1911)、Bloomfield(1926; 1933)、Harris(1946; 1951)等结构语言学家甚为关心的问题。语类问题在马氏的著作中体现为字类问题。马氏认为字类是组字造句的条件(字字相配必从其类),这是马氏观察到的一个非常重要的语法事实。句法的核心问题就是说明组合关系,而组合关系是以聚合关系为条件的。当然马氏并没有弄清楚字类和分布的复杂关系,尽管他认识到了“惟字之在句读也必有其所”这样一个分布问题,但并没有处理好分布和字类的关系。在实际划分字类的时候,既根据意义,又根据位置(分布),有时候两个标准会产生矛盾。马氏首先划分了实字和虚字,所根据的标准就是意义及跟意义相联系的功能:

界说一:凡字有事理可解者,曰实字。无解而惟以助实字之情态者,曰虚字。(《马氏文通·正名》,P19)

马氏的虚字主要是从位置和功能来分类的。《马氏文通·正名》有如下界说:

界说七:凡虚字以联实字相关之义者,曰介字。

界说八:凡虚字用以为提承展转字句者,统曰连字。

界说九:凡虚字用以煞字与句读者,曰助字。

界说十:凡虚字以鸣人心中不平之声者,曰叹字。

马氏实字的划分在界说时根据的是意义(《马氏文通·正名》,P20—21):

界说二:凡实字以名一切事物者,曰名字,省曰名。

① 基于不同的体系,目前有字类和词类的说法,我们统称为语类。