

华中语学论库

◎邢福义

主编

第二辑

复句关系标记的 搭配研究

姚双云 著

华中师范大学出版社

华中语学论库(第二辑)

邢福义 主编

教育部人文社会科学重点研究基地基金资助

华中师范大学语言与语言教育研究中心成果

Fùjù Guānxì Biāoji de Dāpèi Yánjiū
复句关系标记的搭配研究

姚双云 著

华中师范大学出版社

2008·武汉

新出图证(鄂)字 10 号

图书在版编目(CIP)数据

复句关系标记的搭配研究/姚双云著. —武汉:华中师范大学出版社, 2008. 6

(华中语学论库(第二辑)/邢福义主编)

ISBN 978-7-5622-3738-9

I. 复… II. 姚… III. 汉语—复句—研究 IV. H146.3

中国版本图书馆 CIP 数据核字(2008)第 061822 号

Fùjù Guānxì Biāoji de Dāpèi Yánjiū

复句关系标记的搭配研究

© 姚双云 著

责任编辑:夏兴通

责任校对:王 炜

封面设计:罗明波

编辑室:文字编辑室

电话:027—67863220

出版发行:华中师范大学出版社

社址:湖北省武汉市珞喻路 152 号

电话:027—67863040(发行部) 027—67861321(邮购)

传真:027—67863291

网址:<http://www.ccupress.com>

电子信箱:hscbs@public.wh.hb.cn

经销:新华书店湖北发行所

印刷:湖北恒泰印务有限公司

督印:章光琼

字数:350 千字

开本:850mm×1168mm 1/32

印张:14

版次:2008 年 6 月第 1 版

印次:2008 年 6 月第 1 次印刷

印数:1—2000

定价:30.00 元

欢迎上网查询、购书

敬告读者:欢迎举报盗版,请打举报电话 027—67861321

序

邢福义

随着历史的发展，社会的进步，科技的发达，语言学在整个世界范围内越来越展示出强大的活力和能量。中国语言学是世界语言学的重要组成部分。为了对中国语言学事业有所推动，我们组织撰写“华中语学论库”。作为专用名称，这里的“语学”主要指汉语语言学，近期的15年时间里以现代汉语语法专题研究为重点。“语学论库”，这是汉语语言学的一个系统工程，如果将来主客观条件具备，在研究范围上可以不断扩大，在研究时间上可以无限延展，在研究队伍上可以辈辈交接，代代传承。“华中”一词，既跟研究队伍的华中群体相关，又跟华中师范大学出版社的名称相关。

汉语语言学源远流长。千百年来，特别是《马氏文通》出版以来，尤其是20世纪70年代之后，由于一代代学者的不懈努力，汉语语言学沿着“创业——拓新——发展”的轨道不断推进。目前，汉语语言学所统括的汉语语法学、汉语语音学、汉语方言学、汉语词汇学、汉语语用学等等学科，都已出现了初步繁荣的喜人局面。

但是，初步繁荣并不意味着已经成熟。对于语言学这样一门

社会科学来说,成熟与不成熟的突出标志,应该是学派或流派是否已经形成。在这一点上,科学跟艺术情况相同。比方说,我国的京剧表演艺术已经达到了成熟的高峰,最基本的表现就是形成的这“派”那“派”,只要一提到“梅派”和“程派”,稍有京剧表演艺术知识的人就会知道这是两个具有各自特点的著名流派。又比方说,我国的书法艺术早已达到了成熟的高峰,最基本的表现就是形成了这“体”那“体”,只要一提到“颜体”,稍有书法艺术知识的人就会知道它是不同于“柳体”“欧体”等的有独特风格的书写体,甚至还会知道颜真卿打破了“书贵瘦硬”的传统书风,开创了二王体系之外的新体。然而,汉语语言学的各门学科,即使是其中发展速度最快的现代汉语语法学,仍然缺乏显示成熟的任何标志,距离真正成熟实际上还十分遥远。

当今的汉语语言学,面临的主要问题是“二求”:一求创建理论和方法,二求把事实弄清楚。这是互补互促而又互成因果的两个问题。没有理论和方法的成熟,一门学科不可能是成熟的。而理论和方法的创建,是学者们长期深入研究的成果,是有效地进行群体性思考、独立性思考和开拓性思考的结晶。因此,必然带有鲜明的个性,带有学派的印记,反映一派学者的思想体系、研究特点和总体成就。另一方面,没有对事实的清楚了解,理论和方法的创建便成为空中楼阁。以现代汉语语法研究来说,之所以至今尚未成熟,自成体系的理论和方法之所以尚未创立起来,最根本的原因还是对事实的了解基本上仍然处于朦胧的状态。真正适合于我国语言文字的理论和方法,最终只能产生在我国语言文字的沃土之上。因此,应该强调“研究植根于汉语泥土,理论生发于汉语事实”。不然,我国的汉语语言学在世界语言学中就可能永远处于附庸的地位,就永远不会有跟国外理论对等交流的时候。

学术派别的产生,起码应该具备三个条件:第一,有特定的

学术领地，提示标帜性的理论和主张；第二，有鲜明的治学特点，形成一套自己的研究方法；第三，有良好的学风，形成一支富有活力的队伍。近年来的研究状况表明，我国的学者们已经或多或少地显示了各自的风格特点，但是，顶多只能说其中孕育着某些派别意识，或者顶多只能说预示了某种派别意识的萌芽。汉语语言学的真正成熟，需要经历很长很长的历史阶段，有赖于众多的学者群策群力，更有赖于一辈一辈的学者发扬愚公移山的接力精神。我们华中研究群体人数很少，力量单薄，起点不高，功力不足，对于汉语语言学的发展起不了多大的作用，但是，我们愿意跟在前辈学者的后头，跟在全国各地学者的后头，尽心竭力地做点力所能及的工作。如果把建设富于特色的汉语语言学比作建筑一座大厦，那么，我们组织撰写“华中语学论库”，便是想为这座大厦的建筑献上几根钢筋几块石头。通过参加大厦的建筑，使我们这支小小的队伍受到训练，这是我们的最大愿望。各部著作在内容上具有独立性，但我们希望，在出版了以上二十部之后可以看到研究风格上的某些特色和理论方法上的某种网络。

“华中语学论库”的撰写和出版，得到华中师范大学出版社的大力支持。年初，出版社社长朱峰先生和中文编辑室主任陈昌恒先生到我家，鼓励我牵头编写一套关于汉语语言学的丛书，要我拟订一个初步的计划。不久之后，新上任的总编辑王先需先生了解了有关情况，立即审定计划，并且从内容到选题都提出了好些中肯的意见。他们为发展学术事业所作的决策，他们在出版事业上的决心、魄力和历史责任感，不管是对我个人还是对华中语言研究群体的所有成员，都是极为有力的鞭策。

千里之行，始于足下！

贵在努力，贵在坚持！

1996年5月4日

目 录

序	邢福义(1)
导论	(1)
第一章 研究背景:“小句中枢理论”与复句信息工程	(10)
引言	(10)
第一节 “小句中枢理论”在中文信息处理中的应用	(10)
一、应用于词语的自动切分和标注	(11)
二、应用于短语的识别和标注	(12)
三、应用于歧义的消解	(15)
四、应用于语法实体复杂特征集的描写	(16)
第二节 “小句中枢理论”与复句信息工程	(18)
一、复句研究的重要意义	(18)
二、复句信息工程的研究内容、目标和开展	(21)
三、关于“汉语复句语料库”	(22)
第三节 “汉语复句语料库”的利用	(30)
一、利用统计数据了解关系标记的使用频率	

.....	(30)
二、利用复句语料库统计关系标记的搭配距离	(31)
三、利用标注语料库发现关系标记的搭配模式	(34)
四、利用标注语料库进行关系标记的识别	(37)
小结	(41)
第二章 词语搭配的理论体系与关系标记的搭配	(42)
引言	(42)
第一节 词语搭配理论体系的介绍	(43)
一、搭配的定义	(43)
二、搭配研究发展的几个重要阶段	(44)
三、国内搭配研究的开展及简单评价	(45)
第二节 关系标记搭配研究的意义、内容和手段	(48)
一、关系标记搭配研究的重要意义	(48)
二、关系标记搭配研究的主要内容	(50)
三、本研究中用到的基本统计方法	(51)
第三节 本书开展研究的语料准备	(64)
一、本书使用的两个语言资源	(64)
二、对研究语料的预处理	(67)
三、对本书统计数据的说明	(71)
小结	(74)
第三章 关系标记搭配的基本情况与相关分析	(75)
引言	(75)

第一节 关系标记的分布情况与使用频率	(75)
一、各类关系标记的分布情况	(75)
二、单个关系标记的使用频率	(135)
第二节 关系标记的搭配强度与语义的衍生	(139)
一、关系标记搭配强度的统计分析	(139)
二、常见的搭配强度大、共现次数多的搭配结构	(145)
三、从搭配强度与意义的衍生看“语表”对“语里”的影响	(148)
第三节 关系标记的搭配距离与相关句法特征	(155)
一、关系标记的搭配距离分析	(155)
二、搭配距离与相关格式的“表—里—值”分析	(158)
第四节 关系标记的语义韵	(172)
一、“语义韵”的界定与关系标记的“语义韵”	(172)
二、原因标记“多亏”的“语义韵”分析	(174)
小结	(179)
第四章 递进标记的搭配、关联与易位	(181)
引言	(181)
第一节 递进标记搭配的基本情况与个案分析	(181)
一、递进标记搭配的基本情况	(181)
二、反递格式“不仅/不但……反而”前搭配项	

“表—里—值”的分析	(183)
第二节 递进层级句式的关联与易位	(191)
一、递进层级句式的语义类型	(191)
二、递进层级句式的关联	(195)
三、递进层级句式的易位	(202)
小结	(216)

第五章 连词“结果”的语法化与共时平面的搭配

.....	(217)
引言	(217)
第一节 “结果”的语法化及语义类型分析	(219)
一、“结果”在历时纵线上的两个不同形式及其用法	(219)
二、“结果”的语义演变及“结果”虚化过程	(226)
三、连词“结果”虚化的动因	(230)
四、连词“结果”虚化的特点	(235)
五、“结果”虚化的认知基础与类型学意义	(237)
六、连词“结果”的语义分析	(240)
第二节 “结果”的连词典型性问题及其共时平面的使用情况	(252)
一、“结果”的连词典型性聚类分析	(252)
二、连词“结果”的分布、句法位置与搭配	(262)
第三节 “结果”和“所以”的搭配差异	(273)
一、关联句类的差异	(273)
二、与因果标记搭配强度的差异	(276)
小结	(277)

第六章 关系标记的连用与异类搭配	(279)
引言	(279)
第一节 关系标记的连用	(279)
一、关系标记连用的分类	(281)
二、关系标记连用基本情况的统计与分析	(282)
三、影响关系标记连用的重要因素	(313)
四、关系标记连用的语用价值	(322)
第二节 关系标记的异类搭配	(326)
一、关系标记组合的三种形式	(326)
二、异类标记搭配的基本模式	(330)
三、异类标记的搭配强度及语用价值	(337)
四、异类标记的搭配与复句层次的划分	(339)
小结	(340)
全书总结	(342)
一 本书的结论	(342)
二 本书的意义	(344)
三 本书的不足及今后的工作	(348)
附录一 “《人民日报》语料样本”265 个关系标记的 出现次数	(350)
附录二 “《人民日报》语料样本”常见搭配模式的 MI 值统计数据	(354)
附录三 “《人民日报》语料样本”常见搭配模式的 Z 值 统计数据	(361)
附录四 “《人民日报》语料样本”中共现次数、 MI 值和 Z 值均在前 95 位的 54 组搭配格式	(368)

附录五	“《人民日报》语料样本”中常见搭配标记的 平均搭配距离	(370)
附录六	“因为……所以”与“之所以……是因为”搭配 距离对比检索语料	(373)
附录七	90 个含有标记“多亏”的检索例句	(398)
附录八	122 个“不仅……反而”的检索例句	(405)
附录九	关系标记的异类搭配部分语料	(417)
参考文献		(424)
后记		(435)

导 论

搭配 (collocation) 是语言学领域最为重要的概念与研究内容之一, 词语搭配研究经过半个多世纪的探索和发展, 已经形成了一套完整的研究体系, 有一系列基本的研究手段和方法。特别是近年来随着语料库语言学的蓬勃发展, 搭配研究能够得到大规模语料库的数据支持, 更加展现出了研究方法的独特魅力。搭配研究在语言教学、词典编撰、自然语言处理等诸多领域中具有重要的应用价值。

本书将搭配研究的基本理论和方法运用到汉语复句研究领域, 主要研究关系标记的搭配问题。关系标记的搭配研究具有非常重要的意义, 汉语重句法, 复句在汉语句法中具有重要的地位, 因此, 要深入研究复句就必须重视对关系标记的研究。关系标记是小句联结的重要手段, 是复句和句群的重要构件单位, 复句、句群乃至篇章的句法、语义分析都离不开关系标记。关系标记在汉语的本体研究和面向中文信息处理的应用研究中都具有重要的意义。

本书的研究主要使用了两个语言资源, 一个是华中师范大学语言与语言教育研究中心开发的“汉语复句语料库” (the Corpus of Chinese Compound Sentences, 简称 CCCS), 另外一个笔者利用《人民日报》连续文本制作的语料统计样本。全书的撰写完全建立在语料库的基础上, 从这个角度来说, 本书的研究属于语料库语言学的范围。

一、书名的阐释

正如书名《复句关系标记的搭配研究》所言，本书的研究范围是汉语复句的关系标记，研究内容是考察关系标记的使用情况，发掘其中的搭配规律，并对相关现象作出合理的解释。

本书对关系标记的定义是：在复句中能够标明分句间的语义关系，主要句法功能是关联分句、构成复句句式的一种连接性成分。

表达复句连接成分的概念有不少种，如黎锦熙（1924）的“关系词”、王力（1943）的“联结词”或“关系末品”、朱德熙（1982）的“关联词语”、邢福义（1985）的“关系词”、廖秋忠（1986）的“连接成分”、王国文（1988）的“逻辑联系语”等等。

本书沿用了邢福义先生的“关系标记”的概念。邢福义先生（1993 a）在《〈红楼梦〉中的“因 p，因 q”》一文里使用了“关系标记”的说法，在专著《汉语复句研究》中也多次使用了“标记”“无标”“有标”“空标”等概念或类似的概念。我们认为，与其他的术语相比，“关系标记”的提法比较合理。“关系标记”的说法涵盖范围广，更具有灵活性。在汉语的复句研究，特别是面向中文信息处理领域的复句研究中使用“关系标记”的概念具有两个明显优点：

其一，“关系标记”表达的概念清晰，不会与现有的“词”的概念混淆。

通常情况下，使用“关联词”或类似的说法容易与“词”的概念混淆。复句关联词不等同于词^①，汉语的词类体系中也没有

^① 邢福义先生（2002a：195）指出，大体上说，关系词有四种，分别为句间连词、关联副词、助词和超词形式。

“关联词”的地位。使用“关联词”容易与词类混淆，在概念上纠缠不清。许多情况下，为了表述的严谨，需要花费很多的精力来解释“关联词”与“词”的区别。如果使用“关系标记”就不会有这方面的困扰了。“标记”涵盖的范围很广，语素、词、短语乃至句法结构都可以纳入“标记”的范围。

其二，使用“关系标记”便于扩大研究的范围，使句法语义分析更细致。

传统语法研究中，“关联词”主要是针对虚词来说的，很少涉及意义实在的词或者短语，因此，“关联词”一类的说法就成了句间连词的代名词，这样一来，研究范围会相对狭窄。实际上，一些实词以及句法功能与实词相当的短语在句子中可以作为分句之间的句法与语义关系的标志。比如，动词“多亏”“亏得”有时可以作原因标记使用。例如：“多亏有少校，才捞回一点面子”（王朔《看上去很美》）；“亏得我警惕性高，没上当”（雪克《战斗的青春》）。这两个句子中，“多亏”“亏得”既是动词，又是一种语义标记。它们所在的分句与后面的分句之间具有因果关系，前面的分句表示原因事件，后面的分句表示结果事件，这两例中的“多亏”“亏得”都可以分析为原因标记。扩大关系标记的范围有助于使句法语义分析更加细致化。在这方面，英语的句法研究为汉语提供了一个很好的范例，英语中的语义标记涵盖范围很广。Kennedy（1998：167）指出，英语中标记原因关系的有词汇化的成分、句法结构甚至小句，比如 cause, result in, arise from, because, thanks to, give rise to, since, responsible for 等。英语中标记原因关系的多达十几个，其中不乏动词或动词性成分。在一些词法分析程序中，in spite of, in order to, because of 等通常当成一个单独的词位。LOB 语料库使用了类似的策略，because of 这样成对的词都被标记为单独的词性。（参看 Manning, 2005：79）

关系标记既是复句句式中重要的句法标记，又是重要的语义关系标记。对于形态不丰富的汉语来说，发掘和整理这些可被计算机利用的形式标记就显得非常重要。随着句法语义研究的深入，对句子的语义关系标记的研究必然会更加细致。使用“关系标记”的概念，可以扩大研究范围，满足对句法语义进行细致分析的需要。比如“由此可见”“总的来说”“换句话说”都可以归纳为解注标记类成员，“与此相反”“恰恰相反”等可以归纳为对照标记成员^①。诸如此类的标记应该还有许多，以后可以慢慢地发掘，逐渐添加到关系标记库中。添加新的关系标记时，不需要过多地考虑它们在别的著述中是否已经被当作关联词来处理了。不提“词”而提“标记”更容易被大家认可、接受。

二、本书运用的理论和方法

本书的研究有两个理论背景。一个是“小句中枢”理论；另一个是“词语搭配”理论。本书力图将两种理论有机地结合，用来解决汉语的一些实际问题。研究过程中综合运用了“两个三角”研究方法、搭配研究的基本手段和方法、统计测量以及聚类分析等跨学科研究方法，希望在弄清楚汉语事实、解释好其背后规律的同时，能够做一些研究方法上的探索。

“小句中枢”是邢福义先生1995年提出来的。邢福义先生(1995)在《中国语文》上发表了题为《小句中枢说》的文章，全面阐述了“小句本位”的学术主张。1996年，邢先生出版了《汉语语法学》专著，实践了“小句本位”的学术主张。“小句中枢”的思想符合汉语的实际情况，是适合于汉语语法研究的语法描写理论。“小句中枢说”的基本内容包括五个方面：其一，认

^① 在邢福义先生的复句分类体系里面，解注关系与对照关系都是属于并列关系类的。

定小句在汉语各类各级语法实体中占中枢地位；其二，认定小句有“成活律”、“包容律”和“联结律”，了解“小句三律”，可以加深对汉语语法系统中小句的中枢地位的认识；其三，认定汉语语法重视“句管控”，即小句在中枢地位上对汉语语法规则的方方面面发挥其管束控制的作用；其四，认定汉语句法结构具有兼容性和趋简性。结构形式的趋简，导致结构语义的兼容；语义兼容的可能性，又提供结构趋简的可能性；其五，认定要在研究工作中贯彻和实践“小句中枢说”。主要的思路和方法是“两个三角”，即“表一里一值”小三角和“普一方一古”大三角。^①

本书的另一个理论背景是“词语搭配”理论，搭配(collocation)是语言学领域最为重要的概念与研究内容之一。1957年Firth正式把搭配作为语言学术语提出来，词语搭配的概念自提出来至今已有半个多世纪。20世纪60年代，新弗思学派(neo-Firthians)发展了词语搭配的概念，建立了词语学研究体系。20世纪70年代中期，琼斯和辛克莱(Jones & Sinclair)完成了第一个基于语料库的词语搭配研究报告，提出了一系列研究原则和基本手段。20世纪80年代至今，基于语料库或语料库驱动的词词语搭配研究广泛展开，开发出了面向搭配研究的检索、统计软件和搭配程序。词语搭配的研究蓬勃发展，并取得了丰富的成果。词语搭配研究在语言教学、自然语言处理、词典编撰等领域发挥了重要的作用。

从研究方法来说，本书综合运用了多种方法。本书运用了“两个三角”的方法来分析复句搭配格式的“语表”“语里”和“语值”的特点。将搭配强度作为搭配格式的语表形式，搭配能力跟一定格式的句式义相关，并影响到格式的语用价值。本书还将关系标记的搭配距离看作语表形式，关系标记的搭配距离也与

^① 参看邢福义(2004: 48)。