

现代汉语名词词义 组合分析

A Syntagmatic Study on Noun Senses
in Contemporary Chinese

王 惠 著



北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

现代汉语名词词义组合分析 / 王惠著. — 北京 : 北京大学出版社, 2004. 12
(语言学丛书)
ISBN 7-301-07835-8

I. 现... II. 王 III. 汉语—名词—现代—研究 IV. H146.2

中国版本图书馆 CIP 数据核字(2004)第 108411 号

书 名 : 现代汉语名词词义组合分析

著作责任者 : 王 惠 著

责任编辑 : 沈 岚

标准书号 : ISBN 7-301-07835-8 / H · 1152

出 版 者 : 北京大学出版社

地 址 : 北京市海淀区中关村北京大学校内 100871

网 址 : <http://cbs.pku.edu.cn> 电子信箱 : zpup@pup.pku.edu.cn

电 话 : 邮购部 62752015 发行部 62750672 编辑部 62752028

排 版 者 : 兴盛达激光照排中心 82715400

印 刷 者 :

发 行 者 : 北京大学出版社

经 销 者 : 新华书店

890 毫米 × 1240 毫米 A5 8.125 印张 240 千字

2004 年 12 月第 1 版 2004 年 12 月第 1 次印刷

定 价 : 14.00 元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,翻版必究

内 容 提 要

随着语言研究的深入以及语言应用需求的推动,词汇语义学成为当今国际语言学界一个备受关注的研究热点,其中,词义组合分析更是词义研究的前沿课题之一。本书对现代汉语名词的词义组合能力进行了大规模、系统的研究,在充分吸收语法研究成果的基础上,探讨词义特征与词的组合能力之间复杂的制约关系。本项研究的突出特点是理论性与实用性并重。一方面,采用组合分析方法来全面研究名词词义,这在汉语词汇学界还是第一次,在词汇语义学理论上有新发展、新突破;另一方面,该研究成果可直接用于辞书编纂、计算机自然语言信息处理等语言工程之中,具有广阔的应用前景。

本书可供对汉语词汇学、语义学、辞书学以及计算语言学等感兴趣的专业人员参考,也可作为高校本科和研究生课程的教学参考书。

ABSTRACT

The research on lexical meaning , especially the syntagmatic study of lexical meaning has great theoretical and practical significance. It has attracted more and more linguists ' attention all over the world since the 80s of the 20th century. This dissertation focused on the large-scale syntagmatic analysis of noun sememes in contemporary Chinese and indicated that there was close correlation between lexical meaning and its distribution.

Besides introduction this dissertation consists of seven chapters.

Chapter 2 briefly recounts the development of research methods about Lexical Semantics , and the current level of study on Chinese lexical meaning analysis. At present , the paradigmatic analysis of lexical-semantic system , including synonymy , antonym , hyponymy , has scored great successes , but the syntagmatic analysis still remains at the initial investigation and study levels. So the syntagmatic analysis of Chinese noun sememes is chosen as the subject of this thesis.

In Chapter 3 , based on a descriptive study on more than 4 ,000 Chinese nouns ' distribution , a distinction criterion between the Free meaning and Restrictive meaning is put forward firstly , their syntagmatic constrains in sentences are also defined in detail. At the same time , a multi-level framework of syntagmatic analysis is designed for describing the syntactic constrains and semantic constrains of any noun in Chinese.

Chapter 4 employs a wordnet-like noun thesaurus to analyze Chinese polysemous nouns. As a result , 87% polysemous words contain two or more than two sememes belonging to different semantic categories. Depending on the above analysis and the framework

proposed in Chapter 3, I survey all of these polyseme nouns, and find that different sememe has different and complementary syntagmatic features.

Chapter 5 makes a thorough investigation and study on some principles about how to set up a new sememe with the help of syntagmatic analysis. It is pointed out that a new lexical meaning would be regarded as a sememe only if the difference between its semantic features and the prime meaning's manifests at the syntax and/or collocation levels. Afterwards, I give a reasonable explanation for three vexed questions of Chinese noun meaning in terms of the influence of distribution.

Chapter 6 takes the further step to demonstrate how to make full use of the syntagmatic analysis framework to compile or improve Chinese dictionaries and electronic dictionaries.

The main work in Chapter 7 is to discuss how to solve ambiguities of Chinese nouns by computer and establish an efficient WSD model based on syntagmatic features.

The last chapter concludes the main achievements and significance of the research done in this dissertation. Planning of further research is also proposed.

序

在组合中分析词义,我国古代学者“随文释义”就是这样做的。西方学者在语言符号系统性认识的基础上提出的聚合关系分析和组合关系分析,则是建立了一种科学理念,使组合分析上升到一般的方法论的高度。而以后学者的探讨则在不同领域、不同层面、不同问题上丰富了组合分析的内容、规则,乃至至于操作的细则。描写现代语言的词义分析也是如此。我国学者吸收这种理论,对现代汉语虚词、实词的词义分析进行了多种探索,取得了很多重要成果。

王惠的博士论文《现代汉语名词词义组合分析》,无疑是这方面研究的一个引人注目的硕果。王惠在北京大学中文系硕士学习阶段主攻语法,毕业后在北京大学计算语言学研究所工作,参加了《现代汉语语法信息词典》编纂、扩充以及相关项目的研究工作,能熟练地应用计算机进行语言文学信息处理的各种工作。博士学习阶段又以词汇词义学为专业方向。有这样的专业理论学习、专业研究为基础,才有条件较好地完成研究课题所提出的任务。王惠的研究值得重视的有几个方面:

1. 吸收语法学、词汇学研究的成果,提出了一个名词义位组合分析的框架,将名词的义位分为自由义、限制义,限制义又分组合中的语法限制和搭配限制,后者又再分为词汇搭配限制、特定格式限制、固定搭配、跨句搭配等不同情况。而这种分析是在考察两万名词的基础上,选定名词义位 4300 条,在 4500 万字的语料中进行的。所作的语法、语义属性分析、组合分析,都有量的统计,不少地方甚至是穷尽性的描写。在这个基础上所作的论证、论断,有很强的说服力。

2. 提出一个比较适合于现代汉语名词意义分析的分类,对不同意义类型名词义位组合的差异作了分析,使不同意义类型义位组合的差异条理化,对一定阶段的词义信息化处理来说,这种概括比从一个个具体的词的词义特征上说明组合的差异,显得简明,可以收以简

驳繁之效。

3. 作者根据论文的分析提出的改进已出版的现代汉语语文词典中名词义位组合说明的意见,现代汉语详解词典中名词义位组合分析的框架,以及信息处理中基于组合特征的消歧方法,都有新的理论上、应用上的阐发。

以“义位”指称词义的最小单位,把它作为组合分析的单位,从现在研究的分析平面来看,是能满足要求的。但是要看到,“义位”的命名是借鉴于“音位”的命名,在类比的同时要注意到其间的差别。一种语言的“音位”是分析该语言语音系统归纳出来的最小的语音单位。而词的“义位”是在一个词内部划分意义的单位。每个词都有意义单位划分的问题。有多个因素影响“音位”的划分,而影响词义单位划分的因素就更多。如应用变化的原因、认识发展的原因、分析详略的原因、不同性质词典有不同的处理的原因等等。因此,就词义单位和组合的关系而言,研究还有待于扩展和深入。

词义单位不同,词义特征就有差别。在组合分析中细化词义特征的分析不是本研究的重点。但它是词义单位和组合关系分析的重要方面,也有待于进一步探讨。研究好比走路,登上一个山头,而前面,路正长。

符淮青

2004.7 于北京大学

前 言

随着语言研究的深入以及语言应用需求的推动,词汇语义学成为当今语言学中一个备受关注的研究热点,其中,词义组合分析更是词义研究的前沿课题之一。本书的主要工作是对现代汉语名词各个意义上可能有的组合能力进行全面、系统的研究,探讨词义特征与词的组合能力之间复杂的制约关系,为大规模的汉语词义组合分析探索出一条路子,并使之在词典编纂和计算机词义自动分析等语言工程中发挥重要作用。全书共分8章。

第一章“引论”,简要介绍本书的研究课题和方法特点。

第二章对中外词义研究理论方法以及现代汉语词义分析研究的发展进行综合评述,指出目前对词义系统内的聚合关系(包括同义、反义、上下位关系等)研究得比较透彻,但词义组合分析仍处于较为零散的理论探讨和举例性分析阶段,取得的实际成果是有限的,具体分析方法也不成熟。以此为背景,确定了本书的研究对象是数量庞大的现代汉语名词义位,尝试着在一个较大范围内通过穷尽性的微观分析、比较,力图得到名词义位与其分布之间关系的一些规律性认识。

第三章在4300余个名词义位的语法属性描写基础上,明确定义了现代汉语名词自由义和非自由义的功能划分标准,并详细刻画了它们在组合中所受到的句法限制、词汇搭配限制的类型。提出一个分层次的名词义位分析框架,为进一步进行大规模的词义组合分析提供了统一的观察窗口和操作平台。

第四章深入到多义名词内部,利用词义组合分析框架和一个简要的名词语义类体系,对多义词不同义位的组合能力进行细致分析。结果发现,不同义位之间存在的词义特征差异一定会在语言的实际用法中有所表现,使它们在语法功能或词汇搭配层面上具有彼此不同、并且互为补充的组合特征。

第五章从词义组合分析的角度对汉语名词义位划分原则作了深入探讨,指出词义差别在词的组合能力中会得到体现,不同词义具有各自不同的分布空间,可以成为确立义位的重要根据。这样,不仅可以增强义位划分的科学性,而且也可以对目前争议较多的几类名词义位划分情况作出合理的解释。

第六章以具体实例说明词义组合分析在语文词典编纂的义位划分、例证选择、用法说明,以及汉语语法信息词典的改进等方面具有很大的实用价值。

第七章以词义自动消歧(WSD)为例论述了词义组合分析在自然语言信息处理领域也同样可以发挥重要的作用。

第八章“结语”对本书的研究工作进行总结,指出进一步的研究方向。

本书的研究工作可以说是现代汉语词汇语义学领域内的一次全新探索,内容基本上来自于作者的博士学位论文。除了个别章节有所补充、修订外,基本框架和体例都仍保持了论文原貌。在本书脱稿之际,回顾这几年的学习与探索,我从内心深处感谢各位老师、专家、朋友所给予我的指导、帮助与关心。

首先我要向我的导师——北京大学中文系符淮青教授致以最诚挚的谢意。正是他的悉心指导和言传身教把我一步一步带入了词汇语义学的科学殿堂,使我由5年前对它的陌生好奇变为今天的充满热爱与迷恋,这个广阔而奇妙的领域深深地吸引了我,使我立志要在今后继续投身于词义研究。就论文本身的写作而言,从选题到动笔,从初稿到一遍遍地修改、定稿,从大的理论框架到小的具体细节表达,无处不渗透着先生的心血。先生严谨求实的治学作风、敏锐独到的学术目光,以及诚恳正直的品格,对我今后的研究与为人处世都将产生直接而深刻的影响。

“词义研究要同语法研究相结合,理论开拓要同工程实践相结合”的研究理念一直贯彻在本书的写作之中。而这种理念的形成,则要归功于北京大学计算语言学研究所陆俭明教授和俞士汶教授。这些年来,正是在他们的直接指导下,我作为主要完成人参加“现代汉语语法信息词典”的开发工作,并主持了“现代汉语语义词典”的设计

与研制。通过对7万词语的语法、语义属性的详细描述,我不仅积累了大量的第一手材料,而且对中文信息处理和辞书编纂等语言工程的实际需求有了深切的体会与认识。所有这些,为本书的顺利完成奠定了扎实的基础。

在我的博士论文选题到完成答辩过程中,北京大学中文系王洪君教授、袁毓林教授、沈阳教授、蒋绍愚教授,中国社科院语言所谭景春教授、国家语委鲁川教授、北京语言大学张旺熹教授,都提出了许多宝贵的富有建设性意见,并给予我热情的指导和帮助。在此谨表诚挚的谢意。

非常感谢北京大学对外教育学院副院长李红印教授,与他进行的多次讨论使我深受启发;而且,作为我的师兄,在我的整个写作过程中,他给予了我很多无私帮助和殷切鼓励,使我少走了不少弯路。

十分感谢北大中文系郭锐教授、北大计算语言学研究所段慧明高级工程师、清华大学电子系李涓子博士,他们提供的大规模真实语料与软件检索工具,大大便利了本书的写作。

特别感谢我曾工作过的北大计算语言学研究所这个文理结合、充满朝气的集体。在每周一次的计算语言学讨论班上,不同的观点相互碰撞所产生的思想火花,每次都能从不同角度给我新的启迪,我从中汲取了丰富的营养。孙宾副教授、詹卫东副教授、王厚峰副教授、于江生副教授、胡俊峰博士、常宝宝博士、吴云芳博士都对本书写作提出了很多新颖且颇富见地的提议,在此深表谢意。

近两年,本书中的部分章节在不同的国际语言学研讨会上先后报告过,夏威夷大学李英哲教授、台湾黄居仁教授、台湾交通大学刘美君教授、台湾大学苏仪文教授、台湾清华大学连金发教授、中正大学竺家宁教授、香港大学单周尧教授、香港语文学会汪惠迪先生等从不同角度提出了宝贵的修正意见,进一步拓宽了作者的学术视野。这里一并致以深深的谢意。

特别感谢新加坡国立大学中文系主任李焯然教授对我学术研究的一贯支持与鼓励,特别感谢李子玲副教授、黄贤强副教授、刘宏副教授、苏瑞隆副教授、徐杰副教授、石毓智博士、梁秉赋博士等同事对我的关照与帮助,使我可以很快地融入新环境,心情愉快地从事自己

挚爱的汉语词汇学的教学研究工作的。同时,也感谢新加坡国立大学学术基金的资助,使得该课题得以在原有基础上继续拓宽广度、挖掘深度。

本书能够列入北京大学出版社的出版计划,则要尤其感谢北京大学中文系王洪君教授的热情推荐,尤其感谢郭荔女士的大力支持与辛勤付出,她所具有的优秀专业能力与执著的敬业精神都使本书增色不少。

尽管有幸承受了如此之多的指导与关爱,书中的内容也几经调整、修改,但由于个人水平所限,错误疏漏之处仍在所难免,恳请广大读者谅解,也敬请各位前辈、时贤以及青年才俊们多多批评指正!

目 录

序	(1)
前言	(1)
第一章 引论	(1)
1.1 课题的提出.....	(1)
1.2 研究任务与方法、材料	(4)
1.2.1 研究任务	(4)
1.2.2 研究方法	(5)
1.2.3 研究材料	(5)
1.3 本书的结构安排.....	(6)
第二章 词义研究的理论方法	(7)
2.1 国外的词义研究.....	(7)
2.1.1 结构语义学	(7)
2.1.2 80年代以后的词汇语义学研究	(10)
2.2 中国境内的词义研究.....	(14)
2.2.1 传统训诂学中的词义分析	(14)
2.2.2 19世纪末期至20世纪60年代的词义 分析	(16)
2.2.3 八九十年代以来的现代汉语词义研究	(17)
2.3 小结.....	(25)
第三章 现代汉语名词义位的组合能力	(28)
3.1 词义的最小单位——义位.....	(28)
3.1.1 什么是义位	(28)
3.1.2 义位与义项的关系	(30)
3.1.3 本书的考察范围	(31)

3.2	现代汉语名词义位的语法分布.....	(34)
3.2.1	91.1%的名词可以单独作主语 96.2%的可以 单作宾语	(34)
3.2.2	74.4%的名词可以作定语(直接修饰名词) ...	(34)
3.2.3	92.8%的名词可以作定中结构的中心语	(35)
3.2.4	少数名词可以作状语(占0.05%).....	(39)
3.2.5	名词一般不作谓语,但有些条件下也可以.....	(39)
3.2.6	现代汉语名词义位的句法功能统计	(40)
3.3	现代汉语名词的自由义和非自由义.....	(40)
3.3	现代汉语名词非自由义在组合中的语法限制.....	(44)
3.3.1	作主语有限制	(44)
3.3.2	作宾语有限制	(47)
3.3.3	作定语有限制	(49)
3.3.4	作中心语有限制	(51)
3.4	现代汉语名词义位的搭配限制.....	(55)
3.4.1	词汇搭配限制	(55)
3.4.2	特定格式限制	(57)
3.4.3	固定搭配	(58)
3.4.4	跨句搭配	(59)
3.5	现代汉语名词义位的组合分析框架.....	(60)
3.6	小结.....	(65)
第四章	现代汉语多义名词的组合分析	(66)
4.1	现代汉语多义名词统计分析.....	(66)
4.2	现代汉语名词的语义分类.....	(67)
4.2.1	Wordnet 中名词的语义类	(68)
4.2.2	现代汉语名词语义分类	(70)
4.3	名词内部各义位分别属于不同的语义大类.....	(74)
4.3.1	具体物→抽象物	(74)
4.3.2	抽象物→具体物	(78)
4.3.3	抽象物→人	(81)
4.3.5	具体物→人	(84)

4.4	名词内部各义位属于不同的语义小类.....	(87)
4.4.1	植物~构件	(87)
4.4.2	“构件”种类不同	(93)
4.4.3	时点~时段	(98)
4.5	名词内部各义位的语义类相同	(101)
4.5.1	表示特指.....	(101)
4.5.2	表示泛指.....	(107)
4.5.3	名词内部各义位之间没有上下位关系.....	(111)
4.6	小结	(119)
第五章	从词义组合特征看现代汉语名词义位划分.....	(121)
5.1	语法组合能力不同,宜作为不同的义位.....	(122)
5.1.1	词类不同.....	(122)
5.1.2	名词子类不同.....	(128)
5.2	子类相同,但词汇搭配限制不同,宜作为不同的 义位	(138)
5.3	现代汉语中争论较多的几类名词义位划分	(146)
5.3.1	“整体~部分”义.....	(146)
5.3.2	特指义.....	(152)
5.3.3	言语义.....	(159)
5.4	小结	(163)
第六章	名词义位组合分析在词典编纂中的应用.....	(168)
6.1	现代汉语中型语文词典编纂	(168)
6.1.1	确立义项应重视词义的组合能力.....	(168)
6.1.2	例证应能揭示出词义的典型组合特征.....	(175)
6.2	现代汉语用法词典编纂	(190)
6.3	现代汉语详解词典编纂	(197)
6.4	《现代汉语语法信息词典》的改进	(201)
6.4.1	语法信息词典对多义词的处理.....	(201)
6.4.2	语法属性描述的改进.....	(206)
6.5	小结	(213)

第七章	基于义位组合分析的现代汉语名词词义消歧.....	(215)
7.1	词义自动消歧(WSD)概述	(215)
7.1.1	词义消歧的知识源.....	(215)
7.1.2	现代汉语词义消歧研究.....	(218)
7.2	基于组合特征的现代汉语词义消歧策略	(219)
7.2.1	词性不同的多义词消歧.....	(220)
7.2.2	词性相同的多义词消歧.....	(220)
7.2.3	自由义和非自由义在词义消歧系统中的 作用.....	(224)
7.3	小结	(226)
第八章	结语.....	(228)
8.1	本书研究工作的总结	(228)
8.2	本书研究工作的实用价值	(231)
8.3	进一步的研究计划	(232)
	参考文献.....	(233)

第一章 引论

词汇学课题的提出

随着语言研究的深入以及语言工程实践的推动,词汇语义学成为当今语言学中一个备受关注的研究热点,其中,词义的组合分析更是词义研究的前沿课题之一。

从理论上讲,这首先是词汇语义学自身发展的必然趋势。我们知道,词义研究有着悠久的历史,上古的哲学家,如古希腊的苏格拉底、柏拉图和我国的荀子就在他们的著作中涉及了词义问题。欧洲传统语文学和我国的训诂学也一直对词源和个别词义的演变考证感兴趣。但它们都没有涉及词义系统问题。应该说,真正现代意义上的科学的词义研究始于19世纪。结构语言学的奠基人索绪尔首次明确区分了词的意义(指称)和价值(指称)。词的意义是词所指示的概念,价值是词与同一语言系统内相邻近的词之间的关系(组合关系和聚合关系),价值决定意义。在此理论直接影响下形成的结构语义学(指称),开始自觉地关注词义系统问题,提出了从关系出发的语义场理论和从词义分解出发的语义成分分析法,对同义、近义、反义、下义关系等进行细致的描述。但总的来说,结构语义学的词义分析主要集中于语义场内部词义聚合关系的研究,对于词义的组合关系分析做得很少。

20世纪年代以来,对语言符号使用更加精密的要求促使人们深入到词义系统内部,进一步揭示词义之间的细微差别,考察词在实际运用中的表达功能、词义变化、分布特点和组合限制,于是,词义的组合分析日益引起人们的关注。英国著名语义学家莱考夫明确指出“任何词义,无论组合受严格限制与否,都应包括该词的替换关系和组合关系”(1975: 100)。八九十年代出现的语境语义学、认知语义学、概念语义学和框架语义学的高潮也都是主张在组合关系中研究词义。中

国内地学者符淮青(员怨缘,员怨苑,员怨怨)、苏宝荣(员怨怨,圆园园)、张志毅、张庆云(员怨源,圆园园),谭景春(员怨怨,圆园园),李红印(圆园园),中国台湾学者黄居仁等(员怨怨)、蕴志耘(圆园园,圆园园)等对汉语词义组合分析也进行了一系列有益的探索。

词义组合分析的另一个强劲动力来自于语法研究的深入。愿年代以后,为了加强对句法规则的限制,词汇在语法理论(如 愿月语法、词汇功能语法、中心词驱动的短语结构文法、范畴语法等)中的地位越来越重要。人们发现,很多语法现象都可以从论元结构和论元的词义属性方面得到解释,于是开始对谓语动词的词义进行细致分析。这种分析为词义研究提供了一个新的方法,使词义成分的确立不再仅仅依赖于直觉判断,而是充分利用词在句法结构中的不同表现来观察词义特征的细微差别。比如,为什么可以说“悦(员怨源)贼(圆园园)”,但不能说“悦(员怨源)贼(圆园园)”?就是因为动词“精(圆园园)贼(圆园园)”有动作、接触、使动义,而“遭(圆园园)贼(圆园园)”则只有使动义(员怨源,圆园园)。这个时期,语义特征分析也成为汉语语法研究的一个重要方法,语法学家们对语义特征的性质,以及语义特征对句法结构的制约作用和解释作用进行了深入的研究,取得了丰富的成果。

语言工程实践的需求也在呼唤词义的组合分析。一方面,随着词典编纂的现代化,词典在分析词义的过程中所追求的目标也越来越高,不仅要给出词义的准确定义,而且要揭示词义的使用特点,同时义项的确立也要求更加科学、客观。这样,就需要通过词的组合能力来考察词义的具体存在状况,把词义研究与词的用法分布结合起来。

另一方面,语言研究与计算机科学相结合是信息时代语言研究的一个突出特点。随着计算机科学的迅速发展,词义分析的重要性与迫切性也为越来越突出。自然语言理解系统必须依赖于大规模的词汇知识库,而当前这些系统面临的重大问题恰恰就是“词义瓶颈”。《中文信息学报》员怨怨年 猿期为此专门发表了一个编者案:“汉语信息处理在字处理、词处理两个方面已经陆续取得了相对的突破,目前,句处理阶段的攻关已经开始。作为知识工程或智能化系统已经进入了实质性的知识获取和智能模拟的阶段。在这一阶段,