

李斌 著

动宾搭配的语义分析和计算

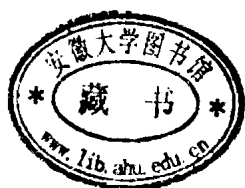
计算语言学研究系列

陈小荷 主编

世界图书出版公司

动宾搭配的语义分析和计算

李 斌 著



世界图书出版公司

北京·广州·上海·西安

图书在版编目(CIP)数据

动宾搭配的语义分析和计算/李斌著. —北京:世界图书出版公司北京公司, 2011. 11

ISBN 978-7-5100-4083-2

I. ①动… II. ①李… III. ①汉语—动宾—研究—现代
IV. ①H146.3

中国版本图书馆 CIP 数据核字 (2011) 第 223775 号

动宾搭配的语义分析和计算

著 者: 李 斌

责任编辑: 夏 丹

出 版: 世界图书出版公司北京公司

出 版 人: 张跃明

发 行: 世界图书出版公司北京公司

(地址: 北京朝内大街 137 号 邮编: 100010 电话: 64077922)

销 售: 各地新华书店和外文书店

印 刷: 三河市国英印务有限公司

开 本: 711mm × 1245mm 1/24

印 张: 11.25

字 数: 250 千

版 次: 2011 年 11 月第 1 版 2011 年 11 月第 1 次印刷

ISBN 978-7-5100-4083-2/H · 1259

定 价: 28.00 元

版权所有 翻印必究

《语言科技文库》总序

李葆嘉

当代语言学已经进入了一个科学与技术的互补时代，信息处理水平成为衡量国家现代化水平的重要标志之一。知识世界的载体是语符系统，信息处理的根本对象是语言信息处理。与计算机的出现使得语言符号有可能成为数据处理对象相似，神经科学实验仪器设备的应用，使得在大脑神经层面探讨语言机制成为可能。这些无疑都引导语言研究走向科技化，“语言科技新思维”（李葆嘉 2001）应运而生。所谓“语言科学”包括理论语言学、描写语言学、历史语言学、应用语言学等分支学科，所谓“语言技术”指语言研究的现代技术手段，包括语言信息处理、语音实验分析，以及语言的神经、心理和行为实验分析的技术手段等。就语言信息处理而言，又可以分为语料库研制技术、知识库研制技术、知识挖掘和抽取技术、句法信息处理技术、词汇信息处理技术、语音信息处理技术、语义信息处理技术、语用信息处理技术等。

2001年5月，南京师范大学文学院创办了史无前例的“语言科学与技术系”，率先迈出了从传统文科教育范型向现代科技教育范型转变的步伐。“十五”期间，南京师大“211工程”重点学科建设项目“语言信息处理与分领域语言研究的现代化”（陈小荷教授主持），以基础平台建设、资源建设和理论探索等为主，迈出了语言科技研究的一大步。

“十一五”期间，南京师大文学院、外国语学院和国际文化教育学院联袂申报“211工程”三期重点学科建设项目。该项目以“语言科技”为引导，以“多学科交叉、跨院系整合、开放型营运”为理念，建设具有前瞻性、原创性、成长性的语言科技高级工作平台。以典型课题的工作原理为核心，进行资源开发和系统研制，拓展语

音科技、二语习得的神经机制研究、言语能力受损儿童的语言能力研究等新方向。同时造就新一代学术领军人物和培养一批高层次复合型人才，以期形成一支高水平的交叉学科团队。该项目设计，体现了工作平台建设、理论创新、应用研究、人才培养、团队建设的学科发展一体化思路。其旨趣在于，加速语言研究从传统文科范型向现代科技范型的转变，以引领 21 世纪语言科技的新潮流。

作为新兴交叉学科项目，通过教育部组织的专家匿名评审，“语言科技创新及工作平台建设”（2008～2011）获批，总投入 1000 万元。总体而言，这一“语言科技创新”团队，分支学科齐全，专业知识互补。涵盖了理论语言学、计算语言学、语义科技、语音科技、实验方言学、历史语言学、神经语言学、二语习得研究、话语行为语言学等领域。这一期间，项目组成员获批的国家级基金项目达 20 多项。该项目理念之前瞻、实力之雄厚、工程之浩大、经费之保障，为学界瞩目。

2008 年秋，本项目以南京师范大学语言科技研究所为实施单位正式启动。主要有三大任务：建设一个领先性的语言信息科技实验室、建立一个独创性的语言科技工作平台、撰著一套有特色的语言科技文库。

从实验室方案设计到设备招标采购，再到实验室用房改造，经过 8 个月的努力，2009 年 12 月，语言信息科技实验室建成，为语言研究从传统范型向科技范型的转变提供了基本保障。该实验室划分为实验工作区、科研工作区和管理服务区。实验工作区建有语音实验与计算室、神经认知实验与计算室、课堂话语实录室三个专门实验室。科研工作区建有语义科技工作室、语音科技工作室、方言实验工作室、知识工程工作室 I（先秦词汇）、知识工程工作室 II（中古词汇）、知识工程工作室 III（敦煌俗语言文字）、语言习得神经机制工作室、语言习得中介机制工作室，以及参研工作室。管理区服务包括办公室、管理室、编辑室和交流室。出席“语言科技高层论坛暨语言信息科技实验室落成仪式”（2009 年 12 月 14 日）的专家认为，该实验室体现了语言学跨学科研究的当代性和先进性，具有整体性、科技型、开放型三个特点，处于全国领先地位，是“语言科技新思维”的又一体现。同时认为，该实验室的科研工作涵

盖了四个二级学科、四个博士学位点，有稳定明确的研究方向，有合理的设计规划和很好的科研基础；整体设计合理，功能齐备。以教育部重点实验室建设标准衡量，很多方面超过了指标。

语言科技工作平台是基于工作原理（课题定位—理论方法—技术路线—关键技术—评估方式）而建设的高级平台。一方面，从语言信息、语言知识和语言机制三个层面，围绕典型课题进行设备配置、资源建设和软件开发；一方面，将典型课题研究和工作平台建设融为一体，依据典型课题建设的子平台应具有解决同类课题的功能。

建设语言科技工作平台的目标是要实现语言研究手段的技术化和模型化，总体设计包括三个二级平台和八个子系统。

一、语言信息工作平台 1. 语义科技工作系统（李葆嘉教授主持）：基于词汇语义—句法语义的一体化研究思路，开发“人—机交互语义标注工具”，研制“深度语义标注信息库”；研制“幼儿（2~6）日常话语跟踪语料库”，完成幼儿语义系统和话语行为分析研究。2. 语音科技工作系统（顾文涛教授主持）：研制“多语言、多语境、多语用的语音语料库”，基于声学信号分析、感知实验和数学建模，完善语音韵律理论与相关技术应用。3. 方言实验工作系统（刘俐李教授主持）：完成“网络版汉语方言有声语料库”，拟定系统的可操作性语音、词汇、语法实验模型和研究方法，进一步完善新兴交叉学科“实验方言学”。

二、语言知识工作平台 1. 先秦词汇统计与知识检索系统（陈小荷教授主持）：研制“先秦文献语料库”、“专名知识库”、“汉语词汇档案库”等，开发先秦文献自动分词算法、古籍版本异文自动发现算法、同指专名检索软件工具等，完成“先秦汉语词汇统计与知识检索”。2. 中古词汇统计与知识检索系统（董志翘教授主持）：研制“中古文献语料库”、“专名知识库”、“中古汉语词汇档案库”等，开发中古文献自动分词和标注工具等，完成“中古汉语词汇统计与知识检索”。3. 敦煌俗语言文字统计与检索系统（黄征教授主持）：研制“敦煌文献资料库”、“敦煌文献俗语词语档案库”，开发相应工具，完成“敦煌文献资料与知识检索”。

三、语言机制工作平台 1. 二语习得的神经机制研究系统（倪

传斌教授主持):研制“英语受蚀词汇库”等,基于行为学、脑成像和脑电三维度模型,进行中国人英语习得与磨蚀的神经机制研究,完成“基于神经机制的英语个性化学习分析系统”。2. 二语习得的中介机制研究系统(肖奚强教授主持):研制“留学生汉语口语中介语语料库”,基于中介语理论、对比分析理论、偏误分析理论以及二语习得影响因素等,完成“留学生汉语习得的中介机制研究”。

这一工作平台,既是科技研究平台,也是人才培养平台,即一个现代化的科学研究和人才培养工作体系。

作为本项目的文本成果,《语言科技文库》包括计算语言学研究、语义语法学研究、汉语方言学研究、古代汉语学研究、语言教学与研究、语言新专题研究六个系列。其总体特征为:领域的开拓性、理论的原创性、选题的新颖性、方法的交叉性、考据的精审性、成果的应用性。在研究过程中,除了数据采集分析、资源建设和软件开发,更重要的还是要有新思路、新理论和新材料。陈小荷提出的先秦文献信息处理新方法,从先秦典籍注疏文献中挖掘出用于自动分词和词义消歧的知识,再注入已开发的古汉语分词和词性标注工具中去,所取得的先秦古籍版本异文自动发现、先秦词汇知识自动挖掘等成果均具开拓性。李葆嘉提出的语义语法学理论和话语行为理论,基于研制专用语料库或语义信息库和技术手段,开拓了语义网络建构、深度语义分析和话语行为研究等新的领域。刘俐李建构的实验方言学理论和方法,为方言学向现代科技方法的转型研究提供了新路,并取得了一系列新成果。黄征多年来从事敦煌文献及其俗语文字研究,古代汉语学研究系列中的敦煌文献校录整理,以及敦煌写本字词考释、以古佚和疑伪经为中心的敦煌佛典词语和俗字研究、两汉声母系统研究等新见迭出。肖奚强基于汉语中介语语料库的二语习得研究,在对外汉语教学研究界已经产生了影响。钱玉莲的汉语介词与相应英语形式比较研究等专著各有亮色。倪传斌依据语言测试和认知实验等数据,从行为学、生理学和语言学三个层面分析影响中国英语学习者外语磨蚀的相关因素。刘宇红基于隐喻的理论探讨,对各类隐喻形式的结构、特性和解读规律进行了多视角的深入探讨。

《语言科技文库》所收论著,由作者在2008年12月申报选题,

2011 年始逐步完稿。系列主编审读了书稿，主要就其学术价值、章节安排、内容关联、行文表述、图表绘制等方面，提出审阅意见。此后，作者们对书稿又进行了修改和润色。《语言科技文库》的作者，大多数是具有博士学位的年轻教师。对于我们这些 20 世纪 80 年代走进语言学研究领域的而言，出版论著可能已不足为道。然而，对于年轻学者而言，其论著的出版既是几年来研究的结晶，也是对其继续探索的促进。换言之，“211 工程”重点学科建设的目的之一，就是为年轻教师搭建一个可持续发展的科研和教学平台。学科带头人的主要任务之一就是提携后进。

尽管从根本上来说，科学或学术研究是一种个人的探索行为，然而复杂问题的研究，无疑需要群体协作。“学科建设”或团队合作模式，是 20 世纪 90 年代后期出现的一个新概念。这种模式涉及总体规划、多方协调，是需要付出精力和心血的。2008 年，通过投票方式推举我担任该项目总负责时，就意识到自己成了一个“劳动班委”。2009 年，前往安徽大学拜访黄德宽教授时，曾谈到“学科负责人的任务就是规划设计，争取项目经费和提供科研设备设施”，得到黄教授的赞许。2010 年，申报江苏省高校哲学社会科学重点研究基地时，评审专家柳士镇教授提问的“作为一个交叉学科项目，各学科之间的协调是怎么考虑的，有什么做法”，可谓一语中的。作为后学，深知交叉研究之艰、学科整合之难。相关学科之间的整合协调需要借助行政机制，但凭借行政方式并非就能完成。当时的回答是，目前做到的是建成了一个可以合作研究的场所，至于学科之间的进一步沟通合作应有较长过程。有一点很明确，只有通过交叉项目，相应学科才能渗透，合作者才能逐步磨合。我们只是在一步步探索。

十一五期间的“211 工程”建设项目即将完成，但是学科建设的任务并没有结束。2010 年，“语言信息科技研究中心”被评审为江苏省高等学校哲学社会科学重点研究基地，为“语言科技”这一交叉领域注入了新的建设活力。重点研究基地建设，除了“跨院系整合、多学科交叉、开放型运行”理念，需要凸显“合作性攻关”。围绕交叉性项目，实施计算语言学、语音科技、神经语言学、语义科技等力量的联合攻关计划。只有通过全面开放以及和与国内外同

行的合作交流，才有望建成具有影响的语言科技研究、人才培养和学术交流基地。

十年前，我（2001）曾写道：“语言科技”的内涵是以理论研究为指导，以描写研究为基础，以应用研究为枢纽，促使语言研究向计算机应用、认知科学和现代教育技术领域等延伸，沟通文理工相关学科以实现语言研究过程及其成果的技术化。“语言科技”的外延为语言工程科技、语言教育科技和语言研究科技。其中，“语言研究科技”是将语言研究活动与资源建设、软件开发相结合，其目标是实现语言学自身的科技化。还应包含语言实验、数据处理这些实验语音学、神经语言学研究的科技手段。

虽然语言学家不可能也不必要都转向语言计算或实验研究，尽管描写、考据和内省始终是最基本的方法，但是具有一定的语言科技意识却非常必要。语言学家只有了解有哪些可供利用的资源、软件或仪器，才能提高其研究深度、精度和效率。语言学家也只有了解到信息处理的语言研究需求，才有可能为之提供可资应用或参考的基础成果。“语言科技”是21世纪语言学研究的潮流。

此为出版缘起。是为总序。

2011年8月谨识于南都

冯 序

李斌博士的《动宾搭配的语义分析和计算》一书就要出版了。他要我写一个序言，我一直关注他对于语义的研究，就欣然应允了。

本书对《现代汉语辞海》中的1462个动词的38119条动宾搭配的选择限制(selectional restriction)进行了较为系统的考察、分析和计算，指出选择限制理论中的语义限制在动词覆盖面和语义约束方面存在局限性，他以认知和计算的双重视角，对影响动宾搭配的语义层面的因素做了个案考察，形成了“语义类别，属性(值)，动态性、主观性，认知事件框架，转喻和隐喻”六个层面的综合分析框架。其中，语义类和语义属性(语义特征)是传统语义研究中常用的两个分析因素，后四个层面是过去研究较少的。本书在这个大框架下展开分析和计算，探讨了六个层面的合理性和可计算性。首先指出了搭配限制中语义类、语义特征的作用范围，用搭配限制的主观因素和认知事件框架来分析宾语的多样性，还特别考察了主观因素中的褒贬指向问题，最后对常见的显性转喻进行了理解性计算，对搭配范围极大又是明喻标记词的“像”做了比喻的自动判定和比喻成分的自动识别。本书成功地解释了诸如“排车票”、“挤公交”、“解决住房”、“吃大碗”、“吃食堂”、“感谢红领巾”、“听贝多芬”等动宾短语中的语义制约关系，这样的解释思路非常新颖，同时又具有说服力。这是一本难得的现代汉语语义研究的著作。

现代汉语动宾短语的语义关系极为复杂，在研究这个问题的時候，似乎是从零数到无限大，越是往前研究，问题就越来越多，就越觉得前面充满了险阻，困难重重，有看不到尽头的感觉。所以，有的专家指出：“动词和宾语的关系是说不完的”，这是有深入研究体会的经验之谈。李斌博士在这本书中的研究，敢于碰这种“从零数到无限大”、“说不完”的困难问题，而且取得了很好的成果，这是值得高兴的。

自然语言处理的范围涉及众多的部门，我认为，这些部门可以归纳为如下四个大的方向：

- **数据处理方向**：这个方向把自然语言看成一种“资源”（resources），对于“语言资源”（language resources）进行数据处理。这一方向早期的研究有术语数据库的建设、各种机器可读的电子词典、“词网”和“知网”的开发。近年来随着大规模语料库的出现，语料库的建设和加工显得越来越重要，成为获取语言学知识的主要手段，通过机器词典和语料库的数据，可以帮助自然语言处理研究者诱导或分析出大量有用的语言知识，从而弥补研究者主观“内省”的不足。

- **语言工程方向**：这个方向把自然语言处理作为面向实践的、工程化的语言软件开发来研究。这一方向的研究一般称为“人类语言技术（Human Language Technique，简称 HLT）”，或者称为“语言工程”（Language Engineering）。如语音自动识别、语音自动合成、机器翻译、人机对话、信息检索、信息抽取、文本分类、文本数据挖掘、自动文摘、文字识别，等等。这个方向的研究可以取得很好的经济效益和社会效益，成为自然语言处理的研究热点。

- **语言本体方向**：这个方向把自然语言处理作为语言学的一个分支来研究，主要使用计算机来研究语言的本体，探索语言本身的内在规律。这个研究方向的最重要的研究领域是语音、词汇、语法和语义形式化理论和自然语言处理的数学理论。

- **智能认知方向**：这个方向把自然语言处理作为在计算机上实现人类语言能力的学科来研究，探索自然语言理解的智能机制和认知机制。这一方向的研究与人工智能、心理语言学、认知科学有着密切的关系。

在我国的自然语言处理研究中，数据处理方向成绩卓著，语言工程方向战果累累，可是，语言本体方向和智能认知方向却显得非常薄弱，几乎是一片空白。李斌博士的著作就属于研究语言本体方向和智能认知方向的，他的研究与单纯的语言学定性分析和自然语言处理的工程应用不同的，他力求在清晰地描写语言现象的基础上，对于语言现象寻求具有说服力的认知和语义解释，而且这些描写和解释是建立在一定规模语料的考察之上的，其理论形式具有可计算

性，并可以在具体的计算中进行验证、修正和应用。这是难能可贵的。

李斌博士在本书中指出了选择限制在动词覆盖面、语义约束和比喻等方面的局限性，我认为，除了李斌指出的这些局限性之外，还应当注意如下两方面的局限性：

(1) 当在否定句子中的时候，否定关系明显地违反了选择限制，但是，句子的语义却是合法的。

例如，People realized you can't eat gold for lunch if you're hungry. (人们认识到，当你饥饿的时候，你不会把金子当做午餐来吃的。)

句子中的 eat gold 显然违反了 eat 的选择限制，因为 gold 不具有可食性。但是，由于有否定词 can't，这个句子的语义却是完全合法的。

(2) 当句子描述的事件是不寻常的事件时，尽管违反了选择限制，句子的语义仍然是完全合法的。

例如，In his two championship trials, Mr. Kulkirni ate glass on an empty stomach, accompanied only by water and tea. (在他的两次冠军比赛中，库尔基尔尼先生空腹吞食玻璃，吞食的时候只是喝点水和茶。)

句子中 glass (玻璃) 是不具有可食性的，违反了 eat 的选择限制，可是，这个句子的语义仍然是合法的，因为库尔基尔尼先生具有“吞食玻璃”这种不寻常的本事。

这些问题，还有进一步研究的必要。

语言学中“搭配”(collocation)的研究，从1957年英国伦敦学派的弗斯(Firth)开始，已经有50多年了，在这50多年中，很多有才气的语言学家，如韩礼德(Halliday)、辛克莱(Sinclair)、乔姆斯基(Chomsky)、卡茨(Katz)、福多(Fodor)都对搭配的研究做出了自己的贡献，有的语言学家还编写了各种搭配词典；近年来，神经语言学和认知语言学也开始探索搭配的问题。自然语言处理中，对于搭配获取、搭配识别也开展了很多有意义的研究工作。这些事实说明，搭配问题确实是一个非常引人瞩目而又十分困难的语言学问题。

面对这样的困难问题，我们自然语言处理研究者决不能望而却

步，而应当以大无畏的追求真理的精神，知难而进。自然语言处理的研究者就像侦察兵，对于侦察兵来说，没有路才是最好的路，如果道路都铺好了，还要我们这些侦察兵干什么？我们需要有披荆斩棘的勇气，以克服困难为荣，以忍受艰苦为乐，在没有路的荆棘丛中开出一条金光大道来。在自然语言处理的艰苦探索中，尽管我们这些侦察兵常常有“山重水复疑无路”的困惑，但是，在披荆斩棘之后，也会感受到“柳暗花明又一村”的愉快。一旦这条充满荆棘的艰险道路被铺开了，前面就是一马平川的金光大道。

我1957年在北京大学读书的时候，就开始对自然语言的计算产生兴趣，1959年毅然从理科转到语言学，专门从事自然语言计算机处理的研究，而今已经50多年了。50多年来，我自己从一个不谙世事的幼稚青年变成了一个白发苍苍的垂垂老人了。然而，我毕生所钟爱的自然语言处理却丝毫没有变老，它仍然显得非常年轻，仍然是一门新兴的学科，充满了新鲜的活力。科学之树常青，而我们个人的生命与科学这个常青的大树相比，的确是太渺小了，犹如沧海之一粟。科学事业需要几代人持续不断的努力，才有可能取得明显的成绩。因此，我把自己对于自然语言处理的热爱之情寄托于年轻人，我相信，有更多的像李斌博士这样的年轻人会把自然语言处理研究的接力棒一代一代地传下去，从而使我们这个学科更加成熟。

希望李斌博士再接再厉，在自然语言处理的艰苦探索中创造出更多、更新的成绩。

冯志伟

2011年4月13日

于德国海德堡

陈 序

动宾搭配这个题目很多人做过，但大多是从句法骨架分析的角度去做的。李斌的博士论文《动宾搭配的语义分析和计算》则另辟蹊径，以框架语义学为理论基础，结合认知语言学的思想，从语义角度来分析动词述语对非施事名词宾语的选择限制，取得了重要的研究成果。

过去人们一般认为，动宾搭配的合法性只取决于动词与宾语名词的词义，作者通过自动标注和人工标注两个途径统计证实，动词对于宾语名词有较强语义选择限制的只占一半左右，动词对宾语的制约更多地体现在主观性和认知事件框架上。作者进一步分析，动宾搭配的语义类限制之所以显得薄弱，是有认知逻辑上的深层原因的：

一、宾语位置上可以出现不同的认知事件角色。例如“锻炼年轻人”、“锻炼意志”、“锻炼年轻人的意志”，宾语分别是“宿主”、“属性”和“宿主+属性”，而“宿主”和“属性”都是“锻炼”这一事件框架中的认知角色，换言之，理解什么是“锻炼”，这两个角色都是应有之义。

二、事件框架可以叠加，叠加后一个动词的宾语变成了另一个动词的宾语。例如“挤公交”应理解为“挤着乘公交汽车”。

三、名词可以转喻其他事物。例如“红领巾”转喻“少先队员”，这就可以解释“遇到一个红领巾”之类的动宾搭配。

以上种种发现，表明作者在语言学理论上有着厚实的基础和敏锐的洞察力，对框架语义学有所丰富和发展。其中名词转喻方面已经找到了可行的推测其本体的方法，表明了他在语言计算方面的实力。

不过，动宾搭配的问题，距离严整的理论形态和计算应用，还有很多方面值得探索。我们希望，作者不只是提出问题，能够在将来的研究中解决问题，特别是工程应用中的现实问题。比如（1）找

到可行的方法来自动判别宾语位置上的认知角色，(2) 发现事件框架叠加的各种具体模式，(3) 动词叠加后的隐含动词发现方法，等等。除此之外，我想强调的是，动宾搭配在短语中往往是一种述谓结构的形式，并不一定遵循“动-宾”的线性序列，如名词短语“解决的问题”和“问题的解决”。这些都需要作者更多的研究和探索。

语言学的研究，早已走上了多学科交叉、文理相济的发展道路。就我本人来说，从计算的角度研究语言的兴趣，已经超过了单纯的理论研究，在语料库上尽可能地地毯式地描写和解释语言现象，设计算法和程序，乐在其中。这并不是说理论研究在学科交叉的背景下被漠视，反倒是在描写和计算的过程中，可以感觉到语言理论的局限性、主观性和片面性。希望作者在今后的研究工作中不要偏废，要继续扩展视野，把语言学置于金字塔的塔尖，用厚实的多学科基底支撑其持续发展。值得欣慰的是，在论文答辩后的两年里，作者没有止步不前，完善了实验，发表了相关论文，书稿也有相应调整。祝愿他的语言学研究之路越走越宽广。

陈小荷

2011年4月

于金陵龙江白云园

目 录

引 言	1
0.1 课题的提出	1
0.2 研究意义	4
0.3 问题的定义和研究思路	5
0.4 论文结构和使用资源	6
0.5 研究成果和结论	8
第 1 章 动宾搭配的研究现状	10
1.1 词语搭配的定义及类型	10
1.2 汉语固定搭配的特殊性	11
1.3 词语搭配及动宾搭配的研究现状	12
1.4 小结	28
第 2 章 动宾搭配的分析体系	30
2.1 动宾搭配的界定	30
2.2 语义选择限制的主要因素和分析策略	31
2.3 动宾搭配的分析体系	39
2.4 基于知网的辅助分析软件平台的设计实现	40
2.5 小结	46
第 3 章 动宾搭配的语义选择限制考察	47
3.1 问题的提出	47
3.2 相关研究	47
3.3 数据资源	48
3.4 义原自动标注算法	50
3.5 宾语语义类的考察	52
3.6 宾语义原的考察	56
3.7 动宾搭配的手工标注及分析	59
3.8 从宾语看主观性限制的作用	62

3.9	小结	65
第4章	动宾搭配和动词认知事件框架	66
4.1	问题的提出	66
4.2	认知事件框架的定义	67
4.3	认知事件框架角色的认知关系	70
4.4	认知事件框架和认知语境	77
4.5	认知事件框架和多义动词	79
4.6	认知事件框架与动词的叠加	80
4.7	小结	85
第5章	动词及其他褒贬词语的褒贬指向	87
5.1	问题的提出	87
5.2	褒贬指向的界定	88
5.3	词语褒贬指向的分析	89
5.4	词语褒贬指向的理论意义和实用价值	96
5.5	说话人和主观性	98
5.6	词语褒贬指向的自动获取	99
5.7	小结	102
第6章	名词宾语转喻的自动理解	104
6.1	问题的提出	104
6.2	搭配限制和名词转喻	105
6.3	转喻本体的自动理解	106
6.4	自动理解的计算结果	108
6.5	英语名词转喻的自动理解	111
6.6	小结	113
第7章	“像”的比喻框架识别	115
7.1	问题的提出	115
7.2	比喻义的界定和语料考察	116
7.3	实验及分析	119
7.4	小结	124
结 语	126
附 录	133
附录1	《辞海》的动词列表	133