

面向中文信息处理的 现代汉语动名组合问题研究

王淑华 著

中西書局

面向中文信息处理的 现代汉语动名组合问题研究

王淑华 ● 著

中西書局

图书在版编目(CIP)数据

面向中文信息处理的现代汉语动名组合问题研究/
王淑华著. —上海: 中西书局, 2014. 9
ISBN 978-7-5475-0710-0

I. ①面… II. ①王… III. ①现代汉语-动名词-短
语-研究 IV. ①H146.2

中国版本图书馆 CIP 数据核字 (2014) 第 170773 号

面向中文信息处理的现代汉语 动名组合问题研究

王淑华 著

责任编辑 路征远

装帧设计 梁业礼

出版 上海世纪出版集团

中西书局(www.zxpress.com.cn)

地址 上海市打浦路 443 号荣科大厦 17F(200023)

发行 上海世纪出版股份有限公司发行中心

经销 各地新华书店

印刷 上海市印刷七厂有限公司

开本 700×1000 毫米 1/16

印张 21.75

版次 2014 年 9 月第 1 版 2014 年 9 月第 1 次印刷

书号 ISBN 978-7-5475-0710-0 / H.032

定价 60.00 元

本研究受到教育部人文社会科学研究青年基金项目(《基于模型的现代汉语动名组合问题研究》,09YJ740050)资助。

序

计算语言学是语言学中的一个新兴交叉学科,涉及到计算机科学、语言学、数学(数理逻辑和抽象代数)、心理学、认知科学等不同领域。当今社会,互联网普及,网络技术日新月异,机器翻译、信息获取与检索、语音合成与识别等各种应用技术迅猛发展,计算语言学事实上已经成为互联网技术的重要支撑。

当前,计算语言学的瓶颈主要不在于算法之类的关键技术上,而在于表达人类认识和思维的语言表达式与当代计算技术的中间衔接环节上,在于语言表达式这样的非结构化数据如何能转换表达为计算机可以理解和计算的结构化数据。结构主义理论不足以满足当代信息计算的需求,认知理论和语义模型理论研究的兴起和活跃,试图在人类认知概念与语言表达式的形式化之间寻找一种沟通,寻找合理的理据解释。

王淑华是我兼职于上海师范大学时带的博士,语言学出身,攻读博士学位期间,曾参与多项国家课题如“语义模型与语义计算”“现代汉语内涵逻辑语义词典”等项目的研究,深入到内涵概念语义这个十分艰难但又富于学术价值和实用价值的研究方向。博士毕业以后,她一直从事将内涵逻辑模型论思想应用于汉语语法的研究工作。现在,她在所主持的“教育部人文社会科学研究青年基金项目”研究报告的基础上,完成了这部专著,令我备感欣慰。著作的价值主要体现在以下几个方面:

一、清晰地表达了模型论思想,将动词的语义分为指称事件、行为、动作类的外延义和表示广义方式、过程、结果三因子的内涵义,在此基础上,区分动词的外延用法和内涵用法(外延指称事物,内涵表示组合成分关联特征),并找出了一些形式标志作为动词在具体出现时进行语义概念分析的依据。这为动词的语义分析提供了一种新的方法,是对于动词语义和动词用法的一种新解读。

二、较为成功地实现了模型论思想,构建了汉语动词模型原型,同时为一定数量的动词成功地构建了模型,覆盖了相当数量的动词与名词性宾语组合的概

念联系,对于动词带表示事件或属性的谓词性宾语的问题也有一定的解释力,克服了传统义素分析方法的局限性,在概念属性层次上体现了较高的概括力和解释力。

三、以叠置原理为基础,按照单一的意义标准,将汉语中的词分为原子词和叠置复合词两类,避免了传统语法采用多重标准定义“词”时所面临的一系列矛盾问题,从而形成复合词义的分析求解原则。建议编写人读词典和机读词典时采取不同的处理原则,将机读词典分为原子词表和叠置词表两个部分,这种方法可以更好地服务于中文信息处理中未登录词的处理。

四、根据动词模型,将述宾结构分为入格和入局两种类型。入格类型的名词性宾语可以直接和模型中的元名词匹配,是完整动词词义必须的内容。入局类型的名词性宾语不能跟模型中的元名词直接匹配,不是完整动词词义必须的内容,这类述宾结构其深层是两个语言结构的耦合,是在语言经济性原则作用下形成的一个浓缩型的单位。用这种方法定量地考察了《汉语动词用法词典》中的处所宾语,将“吃食堂、坐火车、教中学、唱堂会”从一般的处所宾语中剥离出来,显得比较合理,对于深入认识其他类型的非常规宾语也有比较积极的借鉴作用。

五、研究了三元动名复合结构 N_1VN_2 内部各组成成分联系的紧密程度,并以此作为概念组合理据解释的重要依据。给出了该结构的两种语义生成模式: $[(N_1V)N_2]$ 采用的是整体限制部分的方法,其深层组合模式是“事件—要素”; $[N_1(VN_2)]$ 则是 N_2 的外延经过两次限制的结果。方法上完全不同于结构主义理论所崇尚的动词中心论,突出了名词概念的地位,这是从刻画汉语特点出发反映多语种语言共性的成功之举。

由于动词和名词是人类最基本的词类范畴,值得深入的地方非常多,作者所进行的这项面向中文信息处理的基础研究只能是揭其一端,且实际研究成果离直接为计算机所用尚有一定距离。但若这本书能促进汉语本体研究同中文信息处理进一步结合,在一定程度上深化计算语言学相关研究,则功莫大焉。

是为序。

陆汝占

2014年7月

目 录

第1章 绪论 \ 1

1.1 研究背景 \ 1

1.2 研究现状概述 \ 4

1.3 本书研究思路 \ 5

1.4 相关问题说明 \ 6

1.4.1 资源的使用 \ 6

1.4.2 符号的使用 \ 8

第2章 内涵逻辑、叠置原理与动词模型 \ 10

2.1 内涵逻辑 \ 10

2.1.1 关于内涵逻辑 \ 10

2.1.2 动词的内涵和外延 \ 11

2.1.3 内涵和外延的转化 \ 11

2.1.4 内涵逻辑的语义描写原则 \ 13

2.1.5 汉语内涵逻辑的独特性 \ 14

2.2 叠置原理 \ 16

2.2.1 叠置原理的发展 \ 16

2.2.2 原子词与叠置词的区分 \ 18

2.2.3 区分原子词和叠置词的意义 \ 22

2.3 动词模型 \ 24

2.3.1 建立动词模型的必要性 \ 24

2.3.2 建立动词模型的过程 \ 25

- 2.3.3 入格与入局——两种不同的述宾关系 \ 27
- 2.3.4 动词模型的应用范围 \ 28
- 2.3.5 动词模型与格语法、义素分析法的比较 \ 30

第3章 单音节动语素构词问题研究 \ 32

- 3.1 前人研究概述 \ 32
- 3.2 通过模型求解未登录词的说明 \ 36
 - 3.2.1 语素组合时操作命令的选择 \ 36
 - 3.2.2 计算机理解未登录词的目标和步骤 \ 38
- 3.3 16个高频动语素构词情况的分析 \ 40
 - 3.3.1 考察对象的选择 \ 40
 - 3.3.2 动语素“发”的构词分析 \ 41
 - 3.3.3 动语素“流”的构词分析 \ 46
 - 3.3.4 动语素“动”的构词分析 \ 49
 - 3.3.5 动语素“回”的构词分析 \ 52
 - 3.3.6 动语素“交”的构词分析 \ 55
 - 3.3.7 动语素“过”的构词分析 \ 58
 - 3.3.8 动语素“见”的构词分析 \ 61
 - 3.3.9 动语素“落”的构词分析 \ 63
 - 3.3.10 动语素“失”的构词分析 \ 67
 - 3.3.11 动语素“来”的构词分析 \ 69
 - 3.3.12 动语素“立”的构词分析 \ 72
 - 3.3.13 动语素“收”的构词分析 \ 74
 - 3.3.14 动语素“进”的构词分析 \ 77
 - 3.3.15 动语素“游”的构词分析 \ 79
 - 3.3.16 动语素“接”的构词分析 \ 81
 - 3.3.17 动语素“选”的构词分析 \ 84
- 3.4 4个动语素和名语素的构词情况分析 \ 87
 - 3.4.1 考察对象的确定 \ 87

- 3.4.2 实例分析 1——藏 \ 88
- 3.4.3 实例分析 2——飞 \ 93
- 3.4.4 实例分析 3——考 \ 101
- 3.4.5 实例分析 4——疗 \ 106
- 3.5 与 V_单、N_单 构词相关的问题 \ 109
 - 3.5.1 缩略词的意义求解模式 \ 109
 - 3.5.2 复合词的词性与语素性类的关系 \ 109
 - 3.5.3 词典释义时应该坚持的两个原则 \ 111
- 3.6 本章小结 \ 113

第 4 章 动词、名词组合问题研究 \ 116

- 4.1 前人研究概述 \ 116
- 4.2 我们的思路和方法 \ 118
- 4.3 50 个动词的模型 \ 120
- 4.4 46 个动词和名词性宾语的组分析 \ 125
 - 4.4.1 20 个单音节动词和名词性宾语的组分析 \ 125
 - 4.4.2 26 个双音节动词和名词性宾语的组分析 \ 143
- 4.5 语流中动名组合的边界确定和语义求解 \ 158
 - 4.5.1 实例分析 1——买 \ 158
 - 4.5.2 实例分析 2——卖 \ 166
 - 4.5.3 实例分析 3——加入 \ 174
 - 4.5.4 实例分析 4——进口 \ 181
- 4.6 本章小结 \ 188

第 5 章 三元动名组合 N₁ VN₂ 结构研究 \ 190

- 5.1 前人研究综述 \ 190
- 5.2 名词性结构 NV、VN、N₁ VN₂ 关系探究 \ 192
 - 5.2.1 A 类动词与 N₁ VN₂ 结构 \ 193
 - 5.2.2 B 类动词与 N₁ VN₂ 结构 \ 198

- 5.2.3 C类动词与 N_1VN_2 结构 \ 201
- 5.3 关于 N_1VN_2 结构的几点认识 \ 204
 - 5.3.1 N_1VN_2 结构的性质 \ 204
 - 5.3.2 N_1VN_2 结构内部的语义关系 \ 207
 - 5.3.3 N_1VN_2 结构的语义组合与层次划分 \ 209
 - 5.3.4 N_1VN_2 结构与相关结构的比较 \ 213
- 5.4 计算机理解 N_1VN_2 结构的实例分析 \ 221
 - 5.4.1 计算机理解 N_1VN_2 结构的过程 \ 221
 - 5.4.2 实例分析 1——“ N_1 调查 N_2 ” \ 222
 - 5.4.3 实例分析 2——“ N_1 零售 N_2 ” \ 228
 - 5.4.4 实例分析 3——“ N_1 修改 N_2 ” \ 234
 - 5.4.5 实例分析 4——“ N_1 创新 N_2 ” \ 239
- 5.5 本章小结 \ 244
- 第 6 章 “动词+处所宾语”结构研究 \ 246
 - 6.1 前人研究综述 \ 246
 - 6.2 处所宾语的范围与分类标准 \ 249
 - 6.2.1 处所宾语的范围 \ 249
 - 6.2.2 处所宾语的分类标准 \ 252
 - 6.3 处所宾语的分类考察 \ 256
 - 6.3.1 “奔”类处所宾语考察 \ 256
 - 6.3.2 “朝”类处所宾语考察 \ 260
 - 6.3.3 “戴”类处所宾语考察 \ 263
 - 6.3.4 “来”类处所宾语考察 \ 268
 - 6.3.5 “捡”类处所宾语考察 \ 271
 - 6.3.6 “蹲”类处所宾语考察 \ 275
 - 6.4 对《词典》中两类特殊情况的讨论 \ 278
 - 6.4.1 可以带处所宾语但被词典遗漏的动词 \ 278
 - 6.4.2 被词典误处理为带处所宾语的动词 \ 283

- 6.5 “动词+处所宾语”结构的理解与生成 \ 299
 - 6.5.1 “动词+处所宾语”结构的理解方式 \ 300
 - 6.5.2 “动词+处所宾语”结构的生成方式 \ 302
- 6.6 本章小结 \ 303

第7章 结语 \ 306

附录1 本书所考察的动词的频次表 \ 310

附录2 含未登录词的部分例句 \ 312

附录3 由动名语素构成的非动名类复合词 \ 318

附录4 本书中出现的图表索引 \ 320

主要参考文献 \ 322

后记 \ 333

第1章 绪论

1.1 研究背景

早在计算机出现之前,英国数学家图灵就天才地预见了计算机与自然语言之间的联系:“我们可以期待,总有一天机器会同人在一切的智能领域里竞争起来。但是,以哪一点作为竞争的出发点呢?这是一个很难决定的问题。许多人以为可以把下棋之类的极为抽象的活动作为最好的出发点,不过,我更倾向于支持另一种主张,这种主张认为,最好的出发点是制造出一种具有智能的、可用钱买到的机器,然后,教这种机器理解英语并且说英语。这个过程可以用仿效小孩子说话的那种方法来进行。”(冯志伟,2001,P334)自计算机诞生以后,人们更是在不断地思索以下问题:计算机或者机器是否能够成为认识的主体?机器的形式化过程与人的思维在本质上是相同的吗?如果承认计算机有智慧,这种智慧能够超过人类吗?……(玛格丽特·博登,2001,P8)其实,要想使计算机像图灵所主张的那样——“理解英语并且说英语”,人类要想深刻地认识上述问题,仅仅依靠计算机科学一门学科显然是不够的,更需要语言学的大力参与。让计算机理解并且生成自然语言,前提是人类本身对自然语言已经有了充分的认识和理解,并且能运用一定的技术手段,把人理解语言结构的过程一步步地描述出来,然后把这种认识和理解以计算机可以读懂的方式“告诉”计算机,这样计算机才有可能根据这些知识做进一步的推理和加工。值得庆幸的是,越来越多的专家和学者认识到了计算机科学和语言学联手的重要性,并且积极推动这两门学科的合作,取得了可观的成绩。其中,边缘交叉学科——计算语言学的兴起与蓬勃发展就是证明。

常宝宝、詹卫东(2003)给计算语言学定义如下:计算语言学指的是这样一门学科,它通过建立形式化的数学模型来分析、处理自然语言,并在计算机上用程序来实现分析和处理的过程,从而达到以机器来模拟人的全部或者部分语言能力的目的。同时指出,在宽泛的意义上,有时也以“自然语言处理”来指“计算语

言学”。当处理的对象是汉语时,我们又经常称为“中文信息处理”。

纵观中文信息处理的历史,到目前为止,大约经历了三个阶段:字处理阶段、词处理阶段、句处理阶段。字处理阶段主要解决的是汉字在计算机上的输入输出、语音的输入输出问题;词处理阶段主要关心的是词频统计、通用词库、电子词典、自动分词、词性标注等;句处理阶段的核心问题是句子的理解,首先是单句,然后是复句,接下来是篇章的理解。就目前来说,字处理阶段的问题已经基本解决,词处理阶段也取得了不少成果,市场上已经有了很多自动分词和词性标注的实用软件,现在是处于“句处理”阶段。句处理一般包括两个部分的内容,一是让计算机准确地理解自然语言中每个句子的意思;二是让计算机生成符合自然语言规则的句子。首先要研究解决的是前一个问题。理解一个句子,包括以下几个方面:对句子做出正确的句法分析,根据句子的内容回答相关的问题,能把句子翻译为另一种语言或变换成近义格式等。但是,仅在第一步句法分析上,就遇到了诸多问题。众所周知,汉语没有严格的形态变化,词类与句子成分之间没有对应关系,同一种词性序列可以表达多种句法关系,同一种句法格式可以包含多种语义内容。汉语的这些特点,给自动句法分析增添了难度。目前供计算机使用的多数语言理论是以短语语法为基础,再补充一定的语义信息,但不能满足自然语言处理的需要,即使加上概率也不行。因为概率求得的只是最优解,并不一定是正确解。

那么出路在什么地方呢?许嘉璐(2000)指出,“当前这类研究基本上都是在语料——主要是词——的统计概率的基础上进行的。许多专家已经感觉到,统计概率的路已经走到尽头,必须另辟蹊径,这‘蹊径’就是语义,以词义为基础,与句法规则结合,以句为突破的单位”。事实上,以语义为突破口深入语言内部,寻找新径走出自然语言处理的困境,已经成为国内外众多学者的共识。早在1963年,Katz and Fodor就提出用词语意义分解来描述动词对论元的选择限制,Chomsky(1965)指出在词项插入过程中需要加入选择限制,动词对论元的语义选择限制是生成语法词库中对动词描述的一项重要内容。Resnik(1993)基于WordNet的名词分类体系,采用信息熵的方法来研究动词对论元的语义选择限制,称之为语义优选。有越来越多的研究者重视句法语义接口问题的研究,重视词汇本身的语义因素对句法表现产生的影响,在这样的背景下,陆汝占等(1996)提出用内涵逻辑模型论思想来处理汉语,即在内涵逻辑思想的指导下,深入到语义层面,用模型论的思想对汉语中大小不同的语言单位进行解释,努力寻求句法

和语义之间的协调一致,使改进过的汉语语法知识更适合中文信息处理实践的需要。

语言类型学的研究表明,任何语言都有动词和名词这两个词类范畴。现代语言学的调查结果也证实,“名词与动词的区别存在于古往今来的所有语言之中”(陈平,1998)。可以说,动词和名词是词类中最重要的两种类别。吕叔湘(1985)指出:“世界上任何语言,不管它划分的词类是多还是少,都不能没有动词和名词。这是有道理的……同样,构成句子的最根本的词是动词和名词。”“在句子里边,动词和名词互相倚赖,互相制约。”“动词是句子的中心、核心、重心,别的成分都跟它挂钩,被它吸住。”就汉语中的句子来看,虽然不像英语那样,每个句子中都要求有一个和主语保持一致关系的主要动词,但是动词谓语句在汉语中依然占有很大的优势。研究动名组合,对于进一步充分研究汉语中的句子,将起到重要的作用。

在中文信息处理中,虚词可以起到标记的作用。通过虚词的提示,可以对句子进行预处理,把一个比较长的句子切分成若干个较短的词语串。例如,副词、介词、连词、助词等都有提示短语边界的作用。打一个比喻,实词好像头发,千丝万缕,头绪繁多,而虚词相当于木梳,梳一梳,可以帮助理顺实词之间的关系。但是,实词和实词直接组合,可以表示多种句法关系,且没有什么显性的形式标志可以利用,所以计算机对此是一筹莫展。以一个动词和一个名词组合的情况为例,动词在前名词在后可以形成述宾关系、定中关系,有时还可以形成主谓关系、同位关系等。动词在后名词在前也可以形成主谓、定中和状中等多种句法关系。如果遇到多个动词、名词连用,情况就更为复杂。

冯志伟(2001)结合汉英机器翻译的实际,把汉语自动分析中关键的特点和难点归纳为22个问题,其中,词汇分析方面8个,结构分析方面10个,语义分析方面4个。他认为动名组合问题是语义分析的难点之一,即“汉语的基本句式是NP+VP+NP,然而,NP与VP之间却有着各种复杂的句法关系和语义关系”。在随后列出的22种汉语常见的同形歧义结构中,涉及动名组合的就有11种:VP+的+是+NP、全部(部分)+VP+的+NP、VP+数量结构+NP、 V_1+V_2+NP 、 NP_1+NP_2+VP 、 $V+N$ 、 $N+V$ 、 $VP_1+VP_2+的+N$ 、 $VP+N_1+的+N_2$ 、 $V+N_1+N_2$ 、 $N+V+NP+AP$ 。

本书以汉语学界、中文信息处理学界关于动名组合问题的相关研究为基础,在蒙塔古语义学、现代词汇语义学等理论的指导下,对汉语中的部分动词进行考

察,以动词的词汇意义为出发点,以句法语义的同构为目标,寻求各种动名组合体(如动词性语素和名词性语素的组合、动词和一个或多个名词的组合等)的内部组合方式和组合规则,探讨动词在组合时的语义选择问题以及动词的词汇意义对句法表现产生的影响,希望能总结出不同层级上汉语动名组合的一些语义组合模式,并对相关的研究有所借鉴。

1.2 研究现状概述

现代汉语中动名组合问题的研究,总的来说,呈现出如下特点:

1. 本体多,应用少。本体研究领域关于动名组合的研究成果比较丰富,对动词和名词组成的基本结构 VN、NV 以及扩展结构 NVN、VVN、VNN、NNV 等从不同的角度,采用不同的理论(如格语法、配价语法、转换生成语法)进行研究。仅就 VN 而言,李晋霞(2002)指出其研究的切入点多种多样,包括句法位置、句法功能、语义关系、音节模式、短语理解、标记模式、词类地位、配价、转换、“的”的问题、指称、歧义、复合词等。就这样的一种基本结构,汉语语法学界已经从十几个角度进行研究。

相对于本体研究领域的缤纷多彩,应用领域较为冷清。本体研究的成果里面可供中文信息处理所用的也是寥寥无几,目前所见到的相关研究成果主要有:傅承德(1990)《论现代汉语动名短语的句法结构和语义关系的自动识别——兼论自然语言机器理解的策略》、孙宏林(1997)《从标注语料库中归纳语法规则:“V+N”序列实验分析》、陈小荷(1998)《动宾组合的自动获取与标志》、刘慧清(2000)《信息处理用现代汉语“名词+动词”词语串研究》、高建忠(2000)《汉语动宾搭配的自动识别研究》、李晋霞(2002)《面向计算机的“V_双+N_双”结构类型研究》、傅爱平(2004)《黏合式名词短语结构关系的考察和分析》、王霞(2005)《汉语动宾搭配自动识别研究》、周卫华(2007)《面向中文信息处理的现代汉语动宾语义搭配研究》等。

形成这种局面既有主观原因,也有客观原因。汉语语法研究起步虽然较西方为晚,但是也已经有了上百多年的历史,与动词相关的部分一贯是汉语语法学界研究的重点。而中文信息处理是一门新兴的学科,再加上汉语自身的特殊性,前期着重处理的是字、词问题,真正关心句法问题,其实才二十几年的历史。

2. 静态多,动态少。在分析动名组合的时候,对静态的组合能力和动态的实现能力区分得不够。也就是说句法层面的分析多,句子层面的分析少。多数是

分析动词名词组合成短语的情况,很少分析这种组合能力在现实句子中的实现情况。更确切地说,是对于静态的可以组合的动词、名词之间的语义关系分析比较多。但是,没有把结构本身放到句子、语篇中进行考察。

3. 孤立研究多,整体联系少。汉语中的动词、名词直接组合在一起,比较常见的是组成 NV、VN、NVN 这三种词语串。目前多数是以单篇文章讨论其中某一种结构的某一个方面的问题^①,少有把这些结构放在一起研究,这样割裂了三种结构之间的内在联系。实际上大部分的 NVN 结构都可以提取出 NV 和 VN 两种结构。邢福义(1994)认为 NV、VN 可以看作 NVN 的简省形式,这种观点对本研究有很大的启发作用。

1.3 本书研究思路

前文已经提到,从词汇语义入手研究句法语义接口问题,已成为当前的研究趋势。本书从动词的词汇意义入手研究动名组合问题是内涵逻辑模型论思想应用于汉语的一个实践。考虑到即使是对于英语这样形态比较丰富的语言来说,句法和语义上的完全同构也是不可能的。因此,对于汉语,我们把目标定在句法和语义的协调上。句法上组合一次,语义上能给出一个相对适合的解释,反过来,语义又能够验证句法组合的可能性和现实性,这样,句法走一步,语义走一步,句法和语义的相互协调成为可能,期望研究成果能对中文信息处理有所助益。

具体地说,本书是以内涵逻辑方法论为指导,以叠置原理为理论基础,以动词模型为工具,以汉语中的 VN、NV、N₁ VN₂、“动词+处所宾语”等结构为研究对象,努力从 V、N 本身来实现 VN、NV、N₁ VN₂、“动词+处所宾语”等不同层级结构单位语义理解过程的形式化。在研究过程中,我们把词法和句法统一起来考虑,打破汉语研究中词法和句法的界限,同时尽量利用通行的工具书,汲取本体研究成果,实现规则的形式化。

本书由以下7章组成:

第1章,绪论。介绍研究背景、研究现状、研究思路,并对使用的资源和符号进行说明。

第2章,内涵逻辑、叠置原理与动词模型的阐述。本章是全文的理论基础,

^① 如讨论名词性的 VN、NV 结构中动词的词类地位、配价、动词和名词之间的语义关系等,讨论 NVN 结构的性质、生成问题等。

阐明汉语内涵逻辑语义思想的语义描写方法和描写原则,创造性地运用叠置原理,把汉语中的“词”分为“原子词”和“叠置词”两种,说明建立动词模型的必要性、建立模型的过程、模型的使用范围和在求解动名组合结构中的作用。

第3章,单音节动语素构词问题研究。主要讨论单音节动词性语素的构词问题,重点关注单音节动词性语素和名词性语素的构词问题。首先全面考察16个高频动词性语素居于首位和末位时不同的构词情况,然后以动词模型为工具,以“藏、飞、考、疗”4个动词性语素为考察对象,从已登录于词典的词语中总结出动语素与名语素的组合规律和组合条件,并尝试性地应用于未登录词的处理。

第4章,动词名词组合问题研究。分3个层次讨论动词和名词的组合问题:第一层是以词典的释义为基础,给50个动词建立起动词模型;第二层较第一层深入一步,在给46个动词建立模型的基础上,描写了这46个动词在语料中常带的名词性宾语的不同类型,并找出这些名词性宾语与模型中元名词之间的联系;第三层较第二层更进一步,以“买、卖、加入、进口”4个动词为对象,通过模型提供的信息帮助计算机判断动名语串的外部边界,确定合法语串内部的句法关系,求解组合后结构的语义,并针对多个动词名词连用的情况提出优先选择规则。

第5章,三元动名组合 N_1VN_2 结构研究。 N_1VN_2 是目前本体研究中的一个热点结构。本章首先讨论能用于NV、VN、 N_1VN_2 三种名词性结构的动词之间的关系,然后在对大量 N_1VN_2 结构进行分析的基础上,探讨其性质、内部语义关系、语义组合方式和内部层次划分等问题,并与相关的近似结构进行比较,最后以“创新、调查、零售修改”4个动词为对象,展示计算机处理多元动名语串 N_1VN_2 的过程。

第6章,“动词+处所宾语”结构研究。主要考察“动词+处所宾语”结构,对《动词词典》中的处所宾语进行全面考察,根据名词性宾语是否是完整动词语义的必需成分,将处所宾语分成入格和入局两种类型。除列举出《动词词典》中遗漏的可以带处所宾语的动词以外,还分类列举不是处所宾语但被词典误处理为处所宾语的各种情况。

第7章,结语。

1.4 相关问题说明

1.4.1 资源的使用

本书主要使用了以下几种资源: