

孙道功 著

词汇—句法语义的衔接研究

语义语法学研究系列

李葆嘉 主编

世界图书出版公司

词汇—句法语义的衔接研究

孙道功 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

词汇—句法语义的衔接研究/孙道功著. —北京:世界图书出版公司北京公司, 2011. 12

ISBN 978-7-5100-4194-5

I. ①词… II. ①孙… III. ①自然语言处理—研究 IV. ①TP391

中国版本图书馆 CIP 数据核字 (2011) 第 258588 号

词汇—句法语义的衔接研究

著 者: 孙道功

责任编辑: 吴 叹

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

印 刷: 三河市国英印务有限公司

开 本: 711mm × 1245mm 1/24

印 张: 15

字 数: 335 千

版 次: 2011 年 12 月第 1 版 2011 年 12 月第 1 次印刷

ISBN 978-7-5100-4194-5/H · 1276

定 价: 38.00 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“语言科学及技术系”，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，将典型课题研究与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词语文字研究,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

前 言

20 世纪 60 年代, 欧美语言学开始了从语形研究到语义研究的历史性转向。罗斯 (L. Ross, 1972)、拉考夫 (G. Lakoff, 1973) 和麦克莱 (J. McCawley, 1974) 等为代表的生成语义学派提出句法特点取决于语义, 语义具有生成能力的观点。此后, 无论是切夫 (Chafe, 1994) 语法, 还是功能语言学和认知语言学, 无一不把语义纳入研究视野。一向被结构主义认为无须或无法研究的语义终于登堂入室。当今语言学已经进入了语义学为主导的时代。(Ю. Апресян, 1974)

世界范围内的自然语言处理研究, 经过字处理、词处理阶段之后进入句处理阶段。以往句处理研究的重点是句法结构和语义结构的自动分析与标注。实际上, 句处理的核心问题是让计算机理解并生成自然语言的句子。无论是基于规则还是基于统计的处理策略, 在结构主义背景下都面临“句法—语义接口”问题(陆俭明 2006、孙英杰 2007), 而这一问题至今尚未完满解决。计算机尚未具备理解和生成自然语言能力的原因, 就在于计算机需要语义网络的可计算系统。如何解决语义问题, 如何为计算机的理解生成提供可形式化、可计算化的语义网络, 成为语义研究的核心。

从国内外研究来看, 句法和语义的接口研究, 在研究思路(语义网络建构, 还是句子结构分析)、基本立场(语义为本, 还是语形为本、语义为辅)、衔接层面(词汇语义—语义角色, 还是句法成分—语义角色)等问题上存在若干根本性分歧。目前“句法—语义接口”研究, 通常基于语形语法为主而句法语义为辅的思路, 大都在题元结构层面展开。首先把一个句子分为语形结构和语义结构, 通过分析得到句法成分和语义角色, 再进一步考察语义角色如何投射为句法成分。本文称之为基于语形语法的“句法成分—语义角色”接口(以下简称“句法—语义角色”接口)研究。

目前国内的“句法—语义角色”接口研究，更多关注或局限于语义角色一句模层面。有时也采用语义成分分析法，即针对句中的关键词提取与句法密切相关的某些语义特征，以揭示某类角色在某一一句模中所呈现的某些特点或某些句法制约关系。尚未对一定规模的汉语常用词进行义场建构和义征分析的互动，以揭示某一词群或语义子场在句法结构中出现的多种可能性和制约性。

根据语义语法学理论，句法语义植根于词汇语义，词汇语义决定句法语义。语义性是语言的本质属性，语法形式只是语义的部分投影和外部标记。如果把语义结构的生成规则揭示清楚，语法结构就转换为语块排序问题。揭示词汇语义和语义角色的衔接机制或制约条件，揭示词汇进入句法语义层面的语义变化规则，才是“句法结构处理”（句法结构的理解与生成）的关键。基于语形语法的“句法—语义角色”接口研究，仅仅考虑二者之间的对应和投射关系，仍然没有触及句法结构处理的核心。本文认为，揭示“词汇—句法语义”的衔接机制，即词元单位如何实现为句法范畴尤其是语义角色的机制，应是句法结构生成研究的关键。

越来越多的学者认识到词语所携带的句法语义信息，在很大程度上决定了句法语义结构，从而把语法规则研究转移到词汇特征刻画上。所以在自然语言处理中，波拉德和萨格（C. Pollard and I. A. Sag, 1987）提出了中心词驱动的“短语结构文法”（Head-Driven Phrase Structure Grammar）。其原理是基于约束的词汇主义理念，把语法规则研究几乎全部转移到词汇特征刻画上，词语特征描写成为短语结构文法中提供语法信息的重要手段。陆俭明先生（2006）也指出：重视词语的具体意义对句子意思理解的影响，重视词语的特征研究与描写，这可能是解决“句法—语义”接口问题的大道。由此提出了两种研究思路：一是“由外往里”的思路，即探讨一个句子的意思由哪些意义编织而成；一是“由里往外”的思路，即考虑人们如何把对世界的认知通过言辞表达出来。“无论是由外往里的思路，还是由里往外的思路，都不可忽视词语特征的研究。……重视词语的句法、语义特征的研究与描写，将是解决好句法—语义接口问题的重要一步。”“重视词语的特征研究与描写”的这一观点，对词汇—句法语义的衔接研究有重要启发和指导价值。无论是“题

元结构—句法结构”投射的研究，还是“句法成分—语义角色”关系的配比，都难以揭示句法结构的生成机制。只有把研究颗粒度细化到语义特征，即析义元语言层面，打通词汇语义和句法语义，才能揭示词元与语义角色之间的衔接问题，才有可能解决句法结构的生成机制。

基于对 20 世纪中国转型语法学的理性反思和语言信息处理中语形研究的困境，李葆嘉先生（2001）提出面向信息处理的语义语法理论和元语言研究的新课题。语义语法学理论认为自然语言系统是建立在认知符号的实体范畴化（聚义性义场）和实体关联化（组合性义场）基础上的语义网络。元语言理论认为元语言系统存在词汇元语言、释义元语言、析义元语言和认知元语言四个层面的分野。衔接研究任务包括：在词汇语义层面，进行义征提取和义场建构；在句法语义层面，进行语义范畴的提取、描写和语义句模的建构，以及词汇语义和句法语义衔接研究。但词汇语义—句法语义（简称词汇—句法语义）的衔接研究不是凭空进行的，必须建立在词汇义征和句法语义范畴研究的基础上。衔接研究的基础工作包括：（1）对汉语常用词元进行义征分析和义场建构的互动，以建构《词汇义征标记集》；（2）对汉语常见语义角色进行义征分析和范畴建构的互动，以建构《句法范畴义征标记集》和《语义角色范畴标记集》。

词汇语义与句法语义关系的探讨，是建立语义网络不可或缺的条件，也是解决句法结构处理（句法结构的理解—生成）的关键。在这一思路的指引下，本课题从词汇语义和句法语义两个角度入手，寻找词汇—句法语义的衔接机制。

词汇—句法语义的衔接研究，其研究价值表现为两方面：一方面是为汉语信息处理提供基础研究。根据语义语法学理论，句法语义植根于词汇语义，词汇语义决定句法语义。语义性是语言的本质属性，语法形式只是语义的部分投影和外部标记。语言处理的最大瓶颈是语义而不是语形。句子的生成，需经过对象世界—认知图式—语义结构—句法结构四个阶段实现的，其中语义结构是核心环节。如果把语义结构的生成规则揭示清楚，语法结构就转换为语块排序问题。（李葆嘉 2008）揭示词汇—句法语义的衔接机制，揭示词汇语义进入句法语义层面的变化规则，才是“句法结构处理”的

关键。此外，衔接关系的考察中，把语义分析的颗粒度细化到了义征层面，能够更加清楚的凸显词元之间的细微语义差异，为提高词义消歧的精度提供了帮助。

另一方面，可以丰富语义语法学理论。语义语法学强调语言的语义本质性，认为句法结构本质是相关义场的语义关联或相关词语的义征缠绕。“语义统一场”假说是语义语法理论的核心思想，语义语法学包容语形语法学（语义结构法的标记化）研究。在语义语法的框架内，传统词汇学与句法学的畛域彻底泯灭。（李葆嘉 2007）义征标记集的建构可以为词汇语义系统的构建和组合选择提供元语言基础，为汉语词汇研究提供义类分析框架。同时，句法范畴义征的提取，语义角色标记集的建构，以及词汇一句法语义衔接机制的揭示，无疑进一步丰富了语义语法学理论。

笔者从 2005 年秋开始着手这一领域的探索，至今已六年之久，本书即将展现本研究的过程及初步结论。除前言与结论外，全书共分 10 章。书末附有术语表和析义元语言标记集，也是本文的有机组成部分。

孙道功

目 录

第一章 词汇语义和句法语义的研究现状	1
第一节 词汇语义的研究现状	1
第二节 句法语义的研究现状	14
第三节 语义研究存在的问题	38
第二章 词汇一句法语义衔接的研究思路	43
第一节 衔接研究的理论背景	43
第二节 衔接研究的思路方法	46
第三节 衔接研究的对象受限	48
第四节 衔接研究的关键问题	54
第五节 衔接研究的语料来源	57
第三章 《词汇义征标记集》的建立	59
第一节 已有词汇义征的汇集和梳理	59
第二节 词汇义征的验证范围和方法	76
第三节 词汇义征的验证过程和增删	79
第四章 词汇义征的标注效能评估	85
第一节 对验证语料的选择和处理	85
第二节 对词元的义征分析和验证	86
第三节 对词汇义征标记集的评估	113
第五章 词汇义征的特性与元句法模式	117
第一节 词汇义征的人类性和民族性	117
第二节 词汇义征的精确性和模糊性	123
第三节 义征分析的多维度和层次性	129
第四节 义征分析的元句法模式研究	135
第六章 组合结构中的义征缠绕与变化	156
第一节 组合结构中的义征缠绕分析	156
第二节 组合结构中的义征变化分析	166

第七章 《句法语义范畴标记集》的建立	176
第一节 句法语义范畴的界定及其范围	176
第二节 动核结构的语义角色范畴研究	182
第三节 名核结构的属性语义范畴研究	208
第八章 语义角色的特性及其统计	224
第一节 语义角色的主要特性	224
第二节 语义角色的出现频度	229
第三节 语义角色的语形分布	238
第九章 句法范畴义征的提取与定义	245
第一节 句法范畴义征的研究现状	245
第二节 句法范畴义征的提取研究	252
第三节 句法范畴义征的验证增补	270
第四节 动核与角色范畴义征及其界定	278
第十章 词汇语义与句法语义的衔接	282
第一节 “词汇—句法语义”衔接研究的思路	284
第二节 “穿着”场词汇—句法语义衔接例说	288
第三节 “进食”场词汇—句法语义衔接例说	296
第四节 “居住”场词汇—句法语义衔接例说	304
第五节 扩展动核结构词汇—句法语义衔接	311
结论	323
基本术语注释表	330
附录1 词汇义征标记集 (A—H)	333
附录2 词汇义征频度表 (1716 个)	337
参考文献	342
后记	350

第一章 词汇语义和句法语义的研究现状

从时间看,语义研究可以分为三个时期:训诂学、传统语义学、现代语义学。训诂学是为解经服务的,附属于经学。而传统语义学属于词汇学的一部分,以原子主义观为指导,主要关注词源、词的理据、词义的发展和演变、词语解释等。其研究方法是单向的,仅研究了语言中词义纵向的聚合而忽视了组合研究,词义被作为一元整体,没有把词义离析出更微观的元素。传统语义学的缺陷,导致20世纪30年代现代语义学的产生。现代语义学的主要特点是系统论和特征论,研究内容上已经涵盖词汇语义学、句法语义学、篇章语义学等。其中的篇章语义学需要考虑语境、话语功能等语用因素,涉及内容也要复杂得多。对词汇语义和句法语义研究现状的梳理,理应立足于现代语义学,从词汇语义和句法语义两个角度进行。

第一节 词汇语义的研究现状

现代语义学的第一个流派—结构语义学产生于20世纪30年代,主要贡献表现在两方面:语义场理论和义素分析理论。语义场是德国语言学家伊普生(G. Ipsen, 1924)提出的,特利尔(J. Trier, 1934)进一步发展了语义场理论。语义场理论的提出,突破了传统语义学原子论传统,证明了语义单位是成系统的。而义素分析理论的提出,突破了传统语义学词义的一元整体观,研究内容扩展至词义内部的语义特征等微观元素。语义场和义素分析理论与词汇语义密切相关。此后从转换生成语法阵营中分裂形成的解释语义学和生成语义学等语义学派,其研究重点都是句法语义。

一、国外学者的词汇语义研究

词汇语义现状的梳理和研究,应以语义场和词汇义征的相关成

果为重点。语义场理论提出以后,国外诸多语言学家都做过研究,如波尔齐格(W. Porzig, 1934)、乌尔曼(S. Ullmann, 1952)、莱昂斯(J. Lyons, 1977)等,主要贡献表现在五个方面:(1)构建并分析了语言中的颜色、军阶等典型义场;(2)提出了义场中的成员只有在义场中才能确定意义的观点;(3)词和词相互联系,可以形成系统;(4)波尔齐格提出了组合义场概念;(5)语义场中成员以共性特征相聚合、以个性特征相区别。^[1]20世纪80年代后,语义场的专题研究逐步淡出了西方语言学研究领域,被语义分类研究所替代。其原因主要包括两方面:一是义素分析和义场建构是互动的过程,没有大规模义素分析作为基础,只能停留在颜色、军阶、空间、时间、动物、植物等典型上位义场的罗列上,最终无法进行大规模义场的建构;二是注重了词与词之间的外部制约关系,而忽视了词所代表的概念及其语义特征形成的经验基础。所以,自语义场理论提出至今,义场建构的任务并没有完成。

Wordnet (G. Miller, 1993)是20世纪80年代后国外语义分类描写的代表。Wordnet是在基于人类词汇记忆的心理语言学理论推动下产生的,其分类是从语义关系出发,系统中的名词、动词和形容词都类聚为代表某一基本词汇概念的同义词集合,并在这些同义词集合之间建立起各种语义关系。Wordnet包括近95600个不同的词型(51500个简单词和44100个搭配词),这些词组成70100个词义(或者说同义词集合)。其特色之处是试图根据词义而不是词形来组织词汇信息。利用同义、反义、上下义以及部分关系,把众多的词汇组织起来,从而形成了词汇矩阵模型。^[2]WordNet初始分类依据是语法词类,包括名词、动词、形容词、副词和功能词,但是目前的语义分析只处理了前四类。在语法分类的基础上,对每类词语再作初始语义分类。名词的初始概念包括25个,经过概括归并为11个:(1) entity (实体);(2) abstraction (抽象);(3) psychological feature (心理特征);(4) natural phenomenon (自然现象);(5) activity (活动);(6) event (事件);(7) group (集体);(8) location (位置);(9) possession (所属);(10) shape (外形);(11) state (状态)。动词包括动作和状态两大类。动作动词分为14类:(1)身体动作动词(verbs of bodily functions and care);(2)变化动词