

俄罗斯现代语言学丛书

主 编 吴国华 王铭玉 副主编 赵蓉晖

工程语言学

易绵竹 武 斌 姚爱钢 ⊙ 编著



上海外语教育出版社

SHANGHAI FOREIGN LANGUAGE EDUCATION PRESS

俄罗斯现代语言学丛书

主 编 吴国华 王铭玉
副主编 赵蓉晖

工程语言学

易绵竹 武 斌 姚爱钢 ○ 编著

图书在版编目(CIP)数据

工程语言学 / 易绵竹、武斌、姚爱钢编著.

— 上海: 上海外语教育出版社, 2007

(俄罗斯现代语言学丛书)

ISBN 978-7-5446-0261-7

I. 工… II. ①易…②武…③姚… III. 俄语—
应用语言学—研究 IV. H35

中国版本图书馆CIP数据核字(2006)第128693号

高等学校全国优秀博士学位论文作者专项资助项目
第 200313 号

出版发行: 上海外语教育出版社

(上海外国语大学内) 邮编: 200083

电 话: 021-65425300 (总机)

电子邮箱: bookinfo@sflap.com.cn

网 址: <http://www.sflap.com.cn> <http://www.sflap.com>

责任编辑: 岳永红

印 刷: 上海华文印刷厂
经 销: 新华书店上海发行所
开 本: 850×1168 1/32 印张 11 字数 290千字
版 次: 2007年3月第1版 2007年3月第1次印刷
印 数: 3 100 册

书 号: ISBN978-7-5446-0261-7 / H · 0116
定 价: 19.00 元

本版图书如有印装质量问题, 可向本社调换

序一

吴国华同志是全国外语界第一位“高等学校全国优秀博士学位论文”获得者，他用教育部为此而提供的专项资金汇集一批专家编写出一套“俄罗斯现代语言学”丛书，共分六部：《功能语言学》、《工程语言学》、《口语学》、《语言文化学》、《语言翻译学》和《人文修辞学》。这六部著作涉及当前俄罗斯语言学最具世界影响的六个学术领域。在付梓前得以拜读，获益良多。这套丛书对相关学科的性质、研究的历史、研究的对象、研究的方法、研究的内容做了详尽的阐述。

王铭玉等编著的《功能语言学》论述了俄罗斯及世界功能语言学的源流、发展和区别，尤其对俄罗斯功能语法各种流派进行了深入细致的梳理。全书共六章，其中用五章分别介绍了彼得堡功能语法学派邦达尔科的语法理论、佐洛托娃的交际语法、弗谢沃洛多娃的功能·交际句法、塔尔图大学教授舍利亚金的功能语法及梅里丘克的《意思·文本》模型理论。当代俄罗斯的功能语言学有其独到的特点，在许多方面可以与西方和中国的功能语言学相互借鉴。

易绵竹等编著的《工程语言学》着重分析了圣彼得堡工程语言学学术中心和计算语言学创始人皮奥特罗夫斯基在语言学、符号学、信息学、心理语言学、人工智能、言语的统计组合描写与文本的自动处理、机器翻译和语言教学的计算机优化等方面所取得的成就。同时，还介绍了莫斯科学术中心梅里丘克倡导建立的《意思·文本》语言学模型理论及其学术思想（参看《功能

语言学》第六章梅里丘克的《意思—文本》模型理论)。相信,这对我们学习和借鉴国外的学术成就以推动我国从20世纪50年代开始的机器翻译研究及近年发展起来的语料库语言学、语料库建设与词典编纂、计算机辅助外语教学多媒体网络技术的应用等都会是十分有益的。

赵蓉晖等编著的《口语学》共分两部分。第一部分为口语学概论(第一章);第二部分为俄罗斯口语学的主要流派及代表人物,具体又分:达里和俄罗斯民间口语研究(第二章)、雅库宾斯基和对话言语研究(第三章)、科仁娜和功能修辞学派的口语研究(第四章)、西罗季尼娜和萨拉托夫学派的口语研究(第五章)、什维多娃的口语学思想(第六章)、拉普捷娃和普希金学院的口头言语研究(第七章)、泽姆斯卡娅和科学院学派的都市口语研究(第八章)。

彭文钊等编著的《语言文化学》从语言国情学到语言文化学的建立谈起,描述了语言文化学的发展历程,对语言文化学对象和任务、研究方法论(第一、二、三章)等做出了说明。全书第四、五、六章分别介绍了沃罗比约夫语言文化学的理论、普罗霍罗夫的社会文化定型理论、科斯托马罗夫的语言信息单位理论、克拉斯内赫与库德科夫先例现象研究。特别是最后还拿出一部分章节对隐喻、象征与文化伴随意义、语言图景作为语言文化学的研究对象、空间关系的认知语义等进行了分析和阐释。

杨仕章编著的《语言翻译学》以翻译学的理论和语言翻译学的基本理论为出发点,对翻译与翻译学、翻译过程、翻译单位、翻译标准及翻译的技艺进行了论述。尤其是语言翻译学的基本理论一章概括了翻译的普通语言学理论、翻译的语义研究、翻译的语用学研究、翻译修辞学研究、翻译的文本学研究、翻译的社会语言学研究等问题。

樊明明等编著的《人文修辞学》由两部分组成:俄罗斯现代人文修辞学简史及学科综述、俄罗斯现代人文修辞学主要流派及观点。前者论述了俄罗斯现代人文修辞学简史及俄罗斯现代人文修辞学概况,后者论述了莫斯科大学学派、莫斯科师范学院学派、莫

斯科国际大学学派、彼尔姆学派、克拉斯雅尔斯克学派、萨拉托夫学派、沃罗涅什学派及叶卡捷林堡学派主要流派的观点。

综合来看,这套丛书有两点值得注意:第一,把各门学科的对象、方法、任务总结出来,使我们对该学科有一个了解;第二,该丛书详细地介绍了各种学派的观点,使我们能了解各学派之长。这一浩繁的工作来之不易,意义重大。为此,我们应该感谢这套丛书的作者。

王福祥

2005年7月12日

序二

我国的“外国语言学”，就研究的对象而言，有两个方面，一个是对外国的语言（英语、俄语、日语等）的研究，一个是对外国（英国、俄国、日本等）如何研究语言的研究。这后一种研究属于宽泛的“语言学史”的范畴。以史为鉴，国外语言研究中提出的理论和采用的方法，在了解和吸收其中的精华之后，可以为我所用，用来改进我们的语言研究。“语言学”作为一门学科是不分国界的，各国的语言学家互相交流、互相学习，才能共同为语言学的发展作出贡献。

4
上世纪五六十年代，我国对前苏联的语言研究状况是比较重视的，对他们的有影响的语言学家、语言理论、研究方法等都做了许多介绍和评论。之后我们关注的重心就逐渐转向了欧美，对前苏联和俄罗斯语言学的了解就比较隔膜了。每个国家都有自己的语言学传统和特色，我对俄罗斯的语言研究状况知道得很少，不敢妄加评说，只是印象中他们词汇学的研究特别发达，对语言和思维之间关系的研究也很深入，但是我相信，有悠久传统和富有特色的俄罗斯语言学一定有许多值得我们去了解、研究和借鉴的方面。

现在呈现在读者面前的这套“俄罗斯现代语言学”丛书，是我国俄语学界近年来一些年轻有为的学者的研究成果。我浏览了一遍书稿，发觉俄罗斯现代语言学的内容确实很丰富，涉及语言学的方方面面，既有传统的影子，又有当代的特色，有许多理论和方法是值得我们重视和借鉴的。我们如果眼睛只盯住欧美，忽视

了俄罗斯,那我们的外国语言学研究就有很大的缺陷,因此要感谢这套丛书的主编为我们了解现代俄罗斯语言学提供了方便,我也期望今后能见到更多的研究、介绍、评介俄罗斯现代语言学的著作问世。

沈家煊

2005年7月18日



俄罗斯语言学历历史悠久、基础雄厚，世纪之交的俄罗斯语言学流派更是异彩纷呈，研究成果丰硕，发展势头强劲。但目前中国语言学界和外语学界对俄罗斯现代语言学诸分支学科的理论成就还相当陌生，这不能不说是中西文化贯通中的一个缺憾。为此，我们组织了16名学有所长的博士，用了3年时间展开了“俄罗斯现代语言学”课题的研究工作，集中精力梳理了俄罗斯语言学的研究成果，客观评介了俄罗斯现代语言学在世界语言学中的学术地位，勾勒出了俄罗斯现代语言学的学科概貌。并在此基础上，选择了6个既有俄罗斯特点又有国际影响力的语言学研究领域进行重点挖掘，其成果形成了目前这套“俄罗斯现代语言学丛书”。该丛书由6部书组成，分别是《功能语言学》、《工程语言学》、《口语学》、《语言文化学》、《语言翻译学》和《人文修辞学》，全套书共计约150万字。

“俄罗斯现代语言学丛书”是我国俄语界第一套较为系统全面引介并阐释俄罗斯现代语言学思想的著述。它具有以下几个鲜明的特点：（1）注重前沿。在世界语言学发展史上，俄罗斯语言学一直是一支生力军，对人类语言学的贡献是不言而喻的。比如，功能语言学的多元化进程，工程语言学的建模与模块化构筑，口语学的多维探索，语言文化学的独创精神，语言翻译学的交叉融合，人文修辞学的推陈出新，无不反映俄罗斯语言学的与时俱进主旨，体现前沿性的思想精髓。所以，这套丛书的推出是对世界语言学研究的一个有益

的补充。(2)突出特色。俄罗斯语言学研究历来有自己的传统,虽然在其发展过程中也不排斥外来思想的融入,但总的说来,俄罗斯语言学有一定的“原生性”,其主要思想渊源是自己先辈思想的自然延伸,其发展道路与西方语言学的发展道路并不完全相同。我们选定的6个学科正是最具俄罗斯语言学特色的学科,是一套独具风味的“俄罗斯大餐”,相信它对中国和西方语言学具有较大的借鉴意义。(3)坚持客观。我们的研究所遵循的方法是:以某一语言学领域的重要学术流派为纲,广泛搜集各学派代表人物的经典论著和相关文献,通过仔细研读原著原文,厘清各学派的起源、分期、研究方向、代表人物等有关学科发展背景的线索和脉络,原汁原味地提取各学派独树一帜的学术思想和方法论原则,以中国学者的视角对各流派思想的优缺点进行客观评述,并最终形成该研究领域的相对完整的理论体系。(4)体现浓缩。一个学科的发展是丰富多采的,想通过一本书“和盘托出”是非常困难,也是不可能的。但为了方便读者,使其在有限的时间内能够“窥其全貌”,我们纵横兼顾,既有对学说史的挖掘,更注重对现代思想进行缩微式聚合分析,力争达到纲举目张之目的。

当然,囿于有限的学术水平和视野,在这套丛书中会有一些不尽如人意,甚至谬误之处,我们真诚地恳请语言学界各位专家、学者和广大读者批评指正。

吴国华 王铭玉

2005年7月6日



绪 论 工程语言学学科概述

- 1 第一节 工程语言学的学科背景
- 6 第二节 工程语言学的学科形态
- 13 第三节 工程语言学的研究范围
- 17 第四节 工程语言学的应用领域
- 22 第五节 俄罗斯工程语言学的两大学术中心及其代表人物和主要成就

第一章 工程语言学的基本理论与研究方法

- 25 第一节 工程语言学与理论语言学
- 31 第二节 工程语言学悖论
- 52 第三节 语言自动机的构造及工作原理

第二章 俄罗斯工程语言学理论模型举隅

- 84 第一节 工程语言学的建模原则
- 90 第二节 《意思⇌文本》语言学模型
- 104 第三节 信息—语言学模型

第三章 工程语言学分支学科研究述要

- 137 第一节 俄语自动形态分析技术
- 144 第二节 俄语自动句法分析技术
- 150 第三节 俄语自动语义分析技术
- 167 第四节 俄语语料库研究与应用

第四章 机器翻译

- 179 第一节 机器翻译的历史与现状
- 188 第二节 机器翻译的基本问题与方法
- 202 第三节 苏俄著名机译平台介绍
- 216 第四节 机器翻译研究新景观

第五章 信息检索

- 236 第一节 信息检索的基本问题
- 238 第二节 信息检索语言
- 246 第三节 信息检索系统的类型
- 249 第四节 信息检索系统的示例:УИС

第六章 自动文摘

- 257 第一节 文摘的种类和文摘杂志
- 258 第二节 自动文摘
- 269 第三节 摘要翻译

第七章 计算机语言教学法通论

- 273 第一节 CALL 的基本情况
- 273 第二节 教学软件的制作原则与分类
- 282 第三节 教学软件的评测
- 286 余 论 俄罗斯工程语言学前景展望
- 296 附录 1 俄罗斯工程/计算语言学在线资源
- 311 附录 2 外国人名索引
- 314 参考文献

绪论 工程语言学学科概述

第一节 工程语言学的学科背景

一、工程语言学的学术史考察

对任何学科进行学术史的考察,其作用就在于“辨章学术,考镜源流”,由此不仅可以意识到学术嬗变的契机,而且能够获得学科发展的方向感。我们追溯工程语言学的起源及学科形成的动因,必然需要考察机器翻译的早期设想和实验。从俄罗斯的学术发展背景来看,早在 20 世纪初期,苏联发明家斯米尔诺夫—特洛扬斯基(П. П. Смирнов-Троянский)就曾提出机器翻译的设想。1933 年 9 月 5 日,他在莫斯科登记一项专利,其内容是要求保护“制造一台在从一种语言翻译成另一种语言或多种语言时能选择和打印词汇的机器”的专利(патент на машину для подбора и печатания слов при переводе с одного языка на другой)(<http://www.langust.ru/etc/history.shtml#steps>)。这实际上是一台查字典的机器,其思想已经使计算机介入到语言翻译环节。但是将机器翻译的设想付诸语言工程实践并取得初步的应用成果,则首推美国乔治敦大学(Georgetown University)于 1954 年 1 月 7 日进行的世界上第一次俄英机器翻译实验。同年 9 月,苏联学者潘诺夫(Д. Ю. Панов)撰文“借助机器从一种语言到另一种语言的翻译:关于第一次成功实验的报告”(Перевод с одного языка на другой при помощи машины: Отчет о первом успешном испытании),该文的发表被认为是苏联机器翻译历史发轫的原点。紧随美国之后,苏联拉开了机器翻译大会战的序幕,着眼于应用性的工程化机器翻译研究被提到议事日程上,由此迎来了结构语言学、应用语言学和数理语言学的“白银时代”(Серебряный век



структурной, прикладной и математической лингвистики) (Успенский, 1998)。作为应用语言学的重要分支学科之一,工程语言学得到了快速发展。

在 <http://www.tagnet.ru/> 网站上发布的重要文献“俄罗斯计算机:辉煌成就 56 年”(Русский компьютер: 56 лет триумфа), 以类似纪年体的形式详细记录了发生在苏联及解体以后的俄罗斯具有里程碑意义的重大事件中。其中,1979 年的标志性事件就是列宁格勒师范学院(现俄罗斯赫尔岑国立师范大学)工程语言学实验室*研制的机器翻译理论。该理论关注的焦点并非翻译的语言学问题,而是翻译过程的数学建模,其理论实质是将自然语言文本理解为连贯的要素(词位)链。这种机器翻译系统采用“翻译转换的扩展网络技术和神经网络形式化技术”(Технологии расширенных сетей переходов и формализма пейронных сетей),并假定机器如果能够推测出这个网络中的基本要素且在解码过程中不出错,那么就得到合乎逻辑的、语义正确和语法合格的连贯翻译文本。正是在这一年,工程语言学实验室主任皮奥特罗夫斯基(Р. Г. Пиотровский)的专著《工程语言学和语言理论》(Инженерная лингвистика и теория языка)正式出版,它无疑有着教科书性质和里程碑意义,标志着工程语言学的学科体系已基本成型。顺便指出,由于在俄语文献中下列学科术语之间存在同义性 («компьютерная лингвистика», «вычислительная лингвистика», «автоматическая лингвистика», «инженерная лингвистика», «машинная лингвистика»),而作为英语学科名称“computational linguistics”的俄语对等翻译术语«компьютерная лингвистика»得到普遍认可,因此计算语言学学科名称逐步取得“公民权”(право гражданства)。鉴于皮奥特罗夫斯基在该领域所做

* 该实验室后来更名为“工程语言学与翻译教学实验室”(Учебная лаборатория инженерной лингвистики и перевода),主要以机器翻译和计算机语言教学法(компьютерная лингводидактика)为研究重点。(参见:<http://toia.hyperlink.ru/labs.html#ulidip>)

的开拓性工作,他被誉为“前苏联计算语言学的创始人”(создатель компьютерной лингвистики в СССР)(Трошин, 2003)。

探究工程语言学的学术发展脉络与走向,尤其值得关注的是,1959年在皮奥特罗夫斯基的倡导下,组建了全苏“言语统计”(Статистика речи, СтР)研究组,它联合苏联一些科研院所,并于1991年扩展为一个国际性的学术团体,其成员分布在俄罗斯、白俄罗斯、摩尔多瓦共和国、德意志联邦共和国、以色列、加拿大等国。近半个世纪以来,皮奥特罗夫斯基领导研究组成员对语言学建模与符号学、信息学与计算技术、神经及心理语言学、人工智能、言语统计组合描写与文本自动处理(包括机器翻译和语言教学的计算机优化)等诸多领域的理论和应用问题展开深入研究,凸显了研究课题学科交叉的多样性(Пиотровский, 1999a, b)。基于普通语言学、心理语言学和信息统计学的基本原理,СтР研究组研制开发了多主题、多语种的实用化机器翻译系统 МУЛТИС/ СИЛОД,当前俄罗斯一些商品化机器翻译系统如 ПРОМТ/ СТАЙЛОС, САМАРА/СОКРАТ, ПАРС等正是以此为基础设计建立的。总之,该学术团体在工程语言学和机器翻译领域的研究成果赢得了世界声誉,他们研制的系统能够对欧洲语言和东方语言科技与时政文献进行快速高效的自动翻译、摘要和检索,其优质的翻译性能与算法引起了国际学术界的广泛关注(Зайцева, 2001; <http://www.spb-business.ru/show.php?directory=431>)。

二、工程语言学的成因剖析

在皮奥特罗夫斯基看来,自然语言作为人们表达思想、进行交际的有力工具,实际上是一个“不精确的”、却运行良好的超级复杂的语义系统(сверхсложная семантическая система)。为了分析和设计这种超级复杂系统,现代语言学广泛采用启发式研究方法(эвристические приемы),因而语言学的数学化(математизация)和技术化(технизация)色彩也更加鲜明(Пиотровский, 1979:3)。随着科技进步和社会经济的快速发展,信息产业对语言学的需求越来越高,工程语言学应运而生。皮奥特罗夫斯基把促成工程语言

学学科形成的主要因素归纳为以下 5 种 (Питровский, 1979: 5--7):

1. 随着语言学研究领域的拓宽和深化, 必然需要建立一套可靠的实验方法, 借助于这套方法不仅能够对可见层次上的语料进行采集并加以解释, 而且还能揭示非直观的语言背后的深层隐性联系。为此, 就应当运用依据数量数学和非数量数学手段建立起来的语言学建模方法 (приемы лингвистического моделирования)。

2. 数学家试图利用言语材料以使其新思想得以外显和应用。比如, 马尔可夫链 (Марковские цепи) 研制的最初阶段就涉及语言学问题, 信息学奠基人申农 (C. Shannon) 也曾对言语的信息结构进行数学研究, 而乔姆斯基 (N. Chomsky) 的转换生成语法规则与算法论、图论、终端自动机理论等数学思想更是有着渊源关系。

3. 语言学研究广泛引入符号学和系统论方法, 其目的是建立人工智能系统, 以模拟人类言语活动以及复杂系统和超级复杂系统的运行机制。

4. 随着计算机技术的发展, 工业、交通、军队的自动化管理被提到议事日程上来, 这就迫切需要研制能有效处理以各种语言为载体科技和经贸文献信息的实用系统。

5. 为了改进母语及外语教学手段, 有赖于对教学过程进行工程-统计学意义上的优化 (инженерно-статистическая оптимизация), 它实际上涉及计算机辅助语言教学问题。

三、社会信息化与语言工程实施

工程语言学学科之形成与理论框架之建构, 必然会受到时代文化潮流和社会经济发展的深刻影响。在计算机文化日益普及的信息化社会里, 语言学在现代科学体系中的地位发生了急剧的变化, 它不仅被认为是哲学和人文科学发展的突破口, 而且是社会科学、自然科学与思维科学的接合部, 成了一门带头的科学。中国中文信息学会前理事长陈力为院士 (1995: IX) 认为, 语言学之所以会发生这种变化, 固然由于人们对语言所具有的文化本性, 但也