第一图书网, tushu007.com

<<日语语料库研究的理论与实践>>

图书基本信息

书名：<<日语语料库研究的理论与实践>>

13位ISBN编号：9787544612210

10位ISBN编号：754461221X

出版时间：2009-5

出版时间：上海外语教育出版社

作者：毛文伟

页数：191

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：http://www.tushu007.com

前言

依据实际语料对语言进行研究有着悠久的历史，最早可以追溯到中世纪。
即便从现代意义上的第一个语料库布朗语料库建成至今，也已经经历了近半个世纪。
20世纪90年代以来，随着个人电脑的普及和功能的不断增强，利用语料库进行语言研究在技术条件上已经逐渐成熟。
而随着研究的深入，人们也越来越清楚地认识到内省法研究的局限。
一些研究者开始逐步尝试利用报刊、小说等素材，对语言现象进行实证性研究，涌现出一大批优秀的科研成果。
综观该领域，我们可以发现，语料库研究法的有效性不仅在理论上得到广泛认同，在实际的研究活动中也得到了越来越广泛的应用。
但是，毋庸讳言，其中还存在着一些不尽如人意之处。
尤其是相对于英语语料库在架构、内容和研究手段方面的日臻充实、完善，日语语料库的建设和应用无论是在研究理念，还是在具体实施上都无疑是较为落后的。
首先，缺乏经过周密规划、语料来源广泛、具有较好代表性的大型日语语料库。
只有在强有力的理论支撑下，语料库才有可能具备较好的平衡性和代表性，全面、忠实地反映语言的本来面貌。
为此必须进行科学规划，并在语料的选择、录入、校对以及后期维护方面投入极大的人力、物力。

内容概要

《日语语料库研究的理论与实践》结合研究实例，从理论和实践两个方面探讨了语料库在语言研究中的具体应用。
内容涉及语料库研究的学科定位、理论意义、日语语料库建设的现状与不足以及语料库的设计规划等方面的诸多问题。
在此基础上，运用语料库信息自动处理技术，实现了语料自动赋码、信息自动筛选和提取等各项功能，并结合具体研究课题，对其进行了实践和拓展。
《日语语料库研究的理论与实践》内容翔实，紧贴语言研究的实际，具有一定的创新性和启发意义。

<<日语语料库研究的理论与实践>>

书籍目录

第一部 理论篇第1章 语料库研究概述1．1何谓“ 语料库 ”1．2关于语料库在语言研究中的价值的论争1．3“ 语料库语言学 ”和“ 计量语言学 ”的关系1．4“ 语料库语言学 ”质疑1．5小结 第2章 日语语料库建设的现状及问题2．J语料库的分类2．2日语语料库建设的现状2．3日语语料库存在的问题及原因分析2．4小结 第3章 语料库的设计与构建3．1语料库的内容设计3．2语料库的结构设计3．3语料的后期加工及赋码3．4小结 第4章 信息自动抽取技术的原理及实现4．1形态素自动分析和赋码4．2检索对象的定位4．3赋码的删除4．4小结 第5章 语料库素材对结论信度的影响5．1语料库的规模问题5．2语料库素材的类型5．3语料库素材的时代特征5．4关于网上素材的使用问题5．5小结 第6章 分析手法对结论信度的影响6．1如何进行科学取样6．2关注样本的数量问题6．3重视样本的时代特性6．4样本的分布与去伪存真6．5样本归纳的客观、合理6．6小结 第二部 实践篇第7章 共时语法研究领域的应用之一7．1先行研究7．2对机能辞功能的假设7．3关于先行词的限制7．4语气层面的考察7．5小结 第8章 共时语法研究领域的应用之二8．1先行研究8．2对先行词的初步观察8．3先行词的意义范畴8．4语气方面的考察8．5小结 第9章 历时语法研究领域的应用之一9．1先行研究9．2使用情况的初步分析9．3机能辞语法化时期的判断9．4功能的分析9．5各个形态使用频率的历史变迁9．6小结 第10章 历时语法研究领域的应用之二——对瞬间继起机能辞历史变迁的考察10．1表示瞬间继起关系的各机能辞间的竞争10．2各机能辞的核心功能10．3使用频率出现消长的原因分析10．4前后事项之间的时间关系10．5小结 第11章 词汇研究领域的应用11．1引入语料库信息自动处理技术的必要性11．2先行研究11．3例句自动筛选方案11．4自动筛选方案二的思路11．5对筛选结果的分析11．6小结 第12章 文体研究领域的应用——夏目漱石短篇小说的计量性研究12．1何谓“ 文体 ”12．2先行研究12．3夏目漱石短篇小说的统计结果12．4数据的换算12．5对于作品近似程度的观察12．6小结 附录1日本国立国语研究所研究报告一览附录2与语料库研究有关的部分参考文献 附录3本书使用的语料库素材一览 附录4本书各章 节与已发表论文的关系

章节摘录

正是由于基于问卷调查的研究方法存在着以上不足，自20世纪80年代以来，研究者开始尝试运用存储于计算机内的大量实际语料对语言现象进行考察，涌现出大量基于语料库（corpus）的研究成果。
这一方面是由于计算机技术日新月异，硬件性能不断提高，海量存储设备逐渐普及，软件在易用性方面也取得了长足进步，在很大程度上降低了研究者进入的门槛。
另一方面也是由于人们对于实际语料的价值有了新的认识。
这些研究的共同之处在于，它们都是在大量现有的出版物或口语素材中检索包含某种语言现象的实例。
通过进一步的观察、分析和统计，总结出其中蕴含的客观规律。
因此，从本质上看，都属于实证性研究方法。
近年来，基于语料库的实证性研究已经产生了许多成果，在语法、词汇以及句法等各个领域奠定了自己稳固的地位。
对于这种建立在观察分析自然语言现象基础上的实证性研究来说，语料库是不可或缺的研究基础。
实证性研究方法在语言学研究的各个领域广受推崇也使得语料库的重要性不断提高，逐渐成为研究者不可或缺的工具之一。
语料库一词来源于拉丁语Corpus，意为“资料的总体”。
英语中的corpus一词继承了拉丁语的原意。
但我们现在所说的语料库显然不再是简单的资料的集合，它被赋予了新的意义。
Crystal（1991）认为，语料库是语言资料的集合，其资料来源既可以是书面语篇，也可以是话语的记录脚本。
Sinclair（1991）将语料库定义为对自然发生的语篇的收集，目的在于描述一种状态或某种语言中的变化。
Bmer（1998）认为语料库是对自然语篇大宗的、有原则的收集。
顾日国（1998）则将语料库定义为放置语言材料的仓库，存放在计算机里的原始语料文本或经过加工后带有语言学信息标注的语料文本。

第一图书网, tushu007.com

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:http://www.tushu007.com