

<<语料库语言学>>

图书基本信息

书名：<<语料库语言学>>

13位ISBN编号：9787100033640

10位ISBN编号：7100033640

出版时间：2002-4

出版时间：商务印书馆

作者：黄昌宁

页数：288

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>

## 内容概要

语料库语言学是20世纪80年代才崭露头角的一门交叉学科，它研究自然语言文本的采集、存储、加工和统计分析，目的是凭借大规模语料库提供的客观翔实的语言证据来从事语言研究和指导自然语言信息处理系统的开发。

《语料库语言学》将系统地介绍语料库语言学这门新兴学科的研究内容和方法，既包括用计量方法来研究词汇学、词典编纂学、语法学和语义学等方面的内容，也包括语言素材的计量与计算的运用，开发众多的自然语言信息处理系统。

主要章节有：语料库语言学研究的现状与展望，语料库的设计与开发，语料库的加工和管理技术，基于语料库语言学研究，语料库方法在计算语言学中的应用。

书籍目录

第1章 绪论第一节 什么是语料库？

什么是语料库语言学？

第二节 语料库语言学的发展历史第三节 语料库语言学的发展方向及前景第四节 计算机在语料库语言学中的作用第五节 语料库语言学的研究内容第六节 本书的编排

第2章 语料库的设计与开发第一节 语料库设计和编纂中的问题第二节 建设一个语料库第三节 语料库的类型第四节 国外语料库介绍第五节 汉语语料库的建设

第3章 语料库的加工和管理技术第一节 语料的索引及其应用第二节 语料库语言学中的统计第三节 逐词索引软件及其应用第四节 语料库标注

第4章 基于语料库方法的语言学研究第一节 语言研究中的语料库方法第二节 现代汉语句型统计与研究第三节 词典学研究第四节 汉语名词的语义分类研究第五节 词汇—语法问题调查第六节 语域变体（register variation）研究

第5章 语料库方法在计算语言学中的应用第一节 汉语文本中交集型切分歧义的研究第二节 汉语基本名词短语识别研究第三节 基于结构词义空间的汉语词义排歧模型附录 词性标记集参考文献

## 章节摘录

第1章 绪论 “语料库语言学已经成为语言研究的主流。基于语料库的研究不再是计算机专家的独有领域，它正在对语言研究的许多领域产生愈来愈大的影响”。

这是汤姆斯（Thomas）等人1996年为祝贺语料库语言学的主要奠基人和倡导者里奇（Leech）六十诞辰而编纂的语料库语言学研究论文集的开场白。

近年来，对语料库语言学类似的说法频频见于导论和方法论的专著及教科书中，它不仅仅是语料库语言学家的自誉，而且正在成为整个语言学界的共识。

## 第一节 什么是语料库？

什么是语料库语言学？

语料库（corpus）顾名思义就是存放语言材料的仓库（或数据库）。

传统上，语言学家用语料库这个术语表示可作为语言研究基础的、大量自然出现的语言数据。这些语料库可以由书面语和(或)口语的样本组成，并通常被用来代表一种特定的语言或语言变体。在计算机出现之前，研究者——特别是词典编纂者，也有语料库，只是规模小、范围窄，因而难以在学术界形成气候。

近40年以来，语料库这个术语通常指以电子形式保存的语言材料，并被广泛用于语言研究和语言工程

。随着计算机功效的成倍增长，语料库在规模、多样性和使用方便等方面都发生了剧烈的变化。

与此同时，为了存取和加工语料库所拥有的信息，已经开发了大量专用的软件。

计算机语料库迅速成为语言研究的一种普遍资源，现在世界上已经建立了许多规模较大的语料库，有些是国家级的，有些是大学和词典出版商联合研制的。

另外，由于个人电脑的迅猛发展，存储数据的硬磁盘造价持续下降，研究者个人也开始建立适合自己研究兴趣的小型语料库。

.....

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>