词汇计量及实现

苏新春 / 著

创于1897 商務印書館 The Commercial Press

词汇计量及实现

苏新春　著

商務印書館
2010年·北京

图书在版编目(CIP)数据

词汇计量及实现/苏新春著. —北京:商务印书馆,2010
ISBN 978-7-100-06860-4

Ⅰ.词… Ⅱ.苏… Ⅲ.关系数据库—数据库管理系统—应用—汉语—词汇—计量 Ⅳ.H13

中国版本图书馆 CIP 数据核字(2009)第 213237 号

CÍHUÌ JÌLIÀNG JÍ SHÍXIÀN
词汇计量及实现
苏新春 著

商 务 印 书 馆 出 版
(北京王府井大街 36 号 邮政编码 100710)
商 务 印 书 馆 发 行
北京市白帆印务有限公司印刷
ISBN 978-7-100-06860-4

2010 年 4 月第 1 版 开本 850×1168 1/32
2010 年 4 月北京第 1 次印刷 印张 12¼
定价:27.00 元

目　录

第一章 绪 论

一、撰写目的

本书是理论的书。它想探讨的是词汇计量研究的观念、性质、定位及方法。对词汇计量研究的理论问题作了纵横思考，可它并不追求理论阐述的系统化。它只是希望能帮助人们培养起在实际研究中自觉使用计量手段的意识，并知道从何入手来实现计量的目的。从更大点的角度来说，是想宣传一种理念，即如何把人文科学的语言学做得更形式、更全面、更精致，更具有可测性。

本书是操作的书。它对数据库作了较多的具体介绍，具体到一个命令、一个命令地讲，一个步骤、一个步骤地演示，可并没有把数据库当做独立、完整的学习对象，只是关心那些与词汇计量有密切关系的功能，重点在对语料的描写、筛选、查询、挖掘、统计上；没有深入到数据库的内部，介绍它的原理与内部结构，关注的只是与读者直接接触的使用层面。数据库技术发展到今天，其应用已经覆盖到了社会各个领域，其功能已到了极其强大、几乎是无所不能的地步。本书当然不可能全部包括这些，仅仅是寻找它与词汇研究的结合点，为语言的学习者、研究者，为词汇的学习者、研究者提供数据库应用的入门和桥梁，为实现词汇计量研究目的掌握一种好用些的工具和手段。词汇计量研究有许多工具和手段，这只是其中的一种，当然是比较好用的一种。掌握计量手段的最终目的

仍是为了达到对词汇本质和规律的认识。工具与软件只是一种手段、一种方法。如此看来，本书又似乎不仅仅是一本操作的书。

这样本书就具有了两个目标：第一，介绍词汇计量研究的基本理论、原则、特点与发展，主要在第二、三两章。希望通过这些让读者对汉语词汇计量工作有一个概貌的认识。

第二，讲解在数据库中实现词汇计量研究的一些操作方法。词汇计量是一门实践性很强的工作，要获得计量数据，都要依靠一定的手段和方法。

这两个目标，明确，却不那么纯粹；独立，可又紧密连接。

二、适用对象

本书是为对汉语词汇文字的计量研究感兴趣的学生、教师、研究人员及有关读者而写的，特别是不懂编程而希望学会使用数据库的人员。他们具备一定的电脑使用能力，学习和研究中有使用数据库的要求，想学习但又缺乏编程能力。这大概是出身文科背景人的通病。本书最初就是为中文系学生们上课用的讲义。为学生的需求而开课，为教学的进程而写稿。最初是十来页纸的提纲，后来是几万字的短课程教材。希望既有助于克服学生对计量方法的畏惧，又能满足他们学习计量方法的愿望。这成为本书在寻找合适的表述方式时考虑最多的地方。

要对语言符号、文字符号进行精确描写，现在有许多软件可以做到这一点。本书介绍的主要是 Microsoft Access，它是 Microsoft Office 办公套装软件中的子件。个别地方还附带介绍了 Microsoft Excel 的某些功能。本书结合词汇计量研究的需要，展示了它们在词汇研究与教学中的应用价值。故在内容安排上，

除了第四章是正面对软件作了一些概括性的介绍外，其他都是把词汇问题的讨论作为章节纲目的。

三、写作特点

准确地说，“写作特点”应是“撰写要求”，是对自己写作时提出的要求。

1. 突出应用。数据库的功能极其丰富，可运用于社会生活的几乎所有方面。写作时将完全根据词汇计量研究的实际需要，把如何实现常见的词汇计量功能作为主要内容。尽量结合自己十余年来从事词汇计量的实践，融入使用数据库的体会和心得。里面谈到的方法可能不是最好的，但应该是可行有效的。

2. 突出方便。尽量利用软件的可视化界面来实现相关功能。Access 内带有 SQL 语言，SQL 功能强大、表达简洁、操作简便，但对这套语句仍需要一定的学习。而它的视图界面，连“语言”也不需要，只要掌握相关的按键和操作步骤就可以了。为了最大程度地方便读者，也为了使可视化界面与 SQL 语言各显其利，本书将先列出视图界面的操作方法，再辅之以 SQL 界面下的操作。这样读者既可以选一舍一，也可以相互对应。或以视图界面为主，兼学 SQL 语句；或以学 SQL 语句为主，以视图界面来作佐证，收到事半功倍的效果。

3. 突出实践。全书有练习库，以供在讲解具体方法时使用。涉及一些特殊功能时还会随时使用到其他一些实例。每章后附有思考练习题，以复习本章所学的基本内容。全书最后一章为综合练习，设计了若干专题的完整处理过程，以达到把各个具体方法贯通融合的目的。

第二章　词汇计量研究的语言观

对世界的语言研究历史人们常常会做出不同角度的归类。习惯上分成古代的传统语文研究、历史比较语言学研究、结构主义语言学研究、转换生成语法学研究、功能语言学研究，这是就整个研究历史过程的大势而言。粗线条些则可以在结构主义语言学的发展阶段分出布拉格学派、哥本哈根学派、美国结构主义学派。（刘润清，1995）再细点还可以在此之外再分出配价语法、格语法、蒙塔鸠语法等学派。（冯志伟，1999）如果是在一时之学、一国之学的内部再分出若干又可细而又细了。如对20世纪的美国语言学，就有人根据不同的学术观点与活动空间，分出“人类语言学派”、“耶鲁派”、“密歇根派”、“麻省派”、“加州派”、“共性语言学派”、“纽约语言学派”七种学派。（赵世开，1989）

有人从研究者追求目标的不同来分类，认为一百多年来西方现代语言学界各种各样的学派，不外乎是在语言描写与语言解释二者之间作出选择。结构主义学派是描写，它追求的是对语言符号系统的真实把握；转换生成是解释，阐释的是语言的内部生成机制、人的语言能力获得机制；功能语法派是解释，阐释的是语言的表达功能；人类语言学、文化语言学也是解释，阐释的是语言与人文生态环境的关系。（陈平，1987）这样，语言学各家各派又归成了描写派与阐释派两大阵营。

从语言研究的哲学基础角度来观察，又会发现语言研究有着

理性主义与经验主义两大阵营的差别。理性主义在方法上主张的是演绎法，经验主义主张的是归纳法。

一、语言研究的归纳派与演绎派

（一）任何一种语言研究都有自己的哲学观基础

关于语言的研究，凡是有着深入研究、自成系统的，大都会涉及语言的一些本原问题，如语言的起源，语言与客观世界的关系，语言形成的动因，语言与人的发展，这实际上已经进入了哲学的范畴。语义学是语言学的各个部门探讨问题最深入的一个学科分支，所以它的哲学味也最浓，以致哲学界研究语义问题比语言学界显得更为热闹。“如果我们的兴趣在于哲学语义学，可以基本上不考虑语言学家的观点，但是如果我们的兴趣在于语言学中的语义学，则非得了解哲学家、逻辑学家的语义研究不可。”（徐烈炯，1995:1）在语言与客观世界的关系上，各种语义学派大体上可以归为唯物主义与唯心主义。如指称论就是典型的唯物主义，它认为语言的意义来源于客观世界，语言是用来指称事物的，从而使语言与世界联系起来。而意念论则可视作典型的唯心主义，它认为语言的意义来源于心灵。“思维皆源于心胸，埋藏着无法让别人看到，而且无法显露出来。没有思想交流便不会有社会带来的舒适和优越，所以人们有必要找些外表能感知的符号，以便让别人知道构成自己思想的意念。”①

而在语言机制、语言能力的形成与获得上，又形成了经验主义与理性主义两大派。20 世纪结构主义的两大学派的领袖人物布

① ［英］洛克，《论人类理解》，引自徐烈炯《语义学》（修订本），第 20 页。