

外研社学术文库·翻译研究

# 基于语料库的 语言对比和翻译研究

CORPUS-BASED APPROACHES TO

CONTRASTIVE LINGUISTICS AND TRANSLATION STUDIES

Sylviane Granger  
(比) Jacques Lerot 编  
Stephanie Petch-Tyson



外语教学与研究出版社

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS

1010-2006-1010 : 字图标京

Original edition published in English under the title *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*  
© Erlbaum Robert B. V., Amsterdam-New York, NY, 2003

图书在版编目(CIP)数据

外研社学术文库·翻译研究

# 基于语料库的 语言对比和翻译研究

CORPUS-BASED APPROACHES TO

CONTRASTIVE LINGUISTICS AND TRANSLATION STUDIES

(比) Sylviane Granger  
Jacques Lerot 编  
Stephanie Petch-Tyson

出版人：张维成  
责任编辑：李亚平  
封面设计：李亚平  
出版发行：外语教学与研究出版社  
社址：北京市西城区  
网址：http://www.fltrp.com  
印 刷：北京市丰台区  
开 本：620×980  
印 张：14.75  
版 次：2012年8月  
书 号：I204.3/078

邮购电话：(010)8819232  
发行部电话：(010)8819232  
编辑部电话：(010)8819232

外语教学与研究出版社

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS

北京 BEIJING

京权图字: 01-2006-1010

Original edition published in English under the title *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*

© Editions Rodopi B. V., Amsterdam-New York, NY 2003

## 图书在版编目(CIP)数据

基于语料库的语言对比和翻译研究 = *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* / (比)格朗热(Granger, S.)等编. — 北京: 外语教学与研究出版社, 2012. 8

(外研社翻译研究文库)

ISBN 978-7-5600-8529-6

I. 基… II. 格… III. ①对比语言学—文集—英文 ②翻译—研究—文集—英文 IV. ①H0-53 ②H059-53

中国版本图书馆 CIP 数据核字 (2006) 第 154331 号

出版人: 蔡剑峰

责任编辑: 官亚平

封面设计: 袁璐

出版发行: 外语教学与研究出版社

社址: 北京市西三环北路 19 号 (100089)

网址: <http://www.fltrp.com>

印刷: 北京市鑫霸印务有限公司

开本: 650×980 1/16

印张: 14.75

版次: 2012 年 8 月第 1 版 2012 年 8 月第 1 次印刷

书号: ISBN 978-7-5600-8529-6

\* \* \*

购书咨询: (010)88819929 电子邮箱: [club@fltrp.com](mailto:club@fltrp.com)

如有印刷、装订质量问题, 请与出版社联系

联系电话: (010)61207896 电子邮箱: [zhijian@fltrp.com](mailto:zhijian@fltrp.com)

制售盗版必究 举报查实奖励

版权保护办公室举报电话: (010)88817519

物料号: 185290001

# 基于语料库的语言对比和翻译研究 导 读

外研社翻译研究文库

语料即现实生活中的语言，用它来研究语言并促进语言学习，由来已久。据 McEney (2001:1) 的考察，语料的系统运用可以追溯到 19 世纪。而语料库语言学作为一种方法论则出现在 20 世纪 60 年代。传统意义上的语料库规模有限、取材范围窄小，语料收集和应用的经济性及效率均难如人意。自上世纪下半叶以来，计算机技术突飞猛进，互联网开发和应用也日新月异，为语料收集、存储、加工、检索和提取提供了重要的资源和技术支持 (王克非 2004)。计算机语料库既提高了分类、标注和检索效率，又扩大了语料库的适用范围。基于信息技术的语料库发展至今天，人们有理由期望甚至相信计算机语料库能更好地辅助我们从事与语言相关的活动；或者说，语言研究者和语言学习者都可以从海量的语料存储中获取自己所期望的重要信息。

近二十年来，语料库语言学处在飞速发展之中，主要表现为：语料类型不断丰富，语料规模不断扩大，语料加工深度不断增加，语料标注和检索工具也在不断完善；而且方便实用的人工标注软件也渐为人们熟悉。有了这些技术保证，语料库研究者开始更多地注意语料库的应用问题。本书就是在这样一个背景下编写出来的。

本书是一个论文集，论文来自 1999 年在比利时鲁汶天主教大学举办的 Contrastive Linguistics and Translation Studies: Empirical Approaches 研讨会 (1999 年 2 月 5 日至 6 日)。这次会议所提交的论文一部分已经结集出版 (书名为 *Lexis in Contrast—Corpus-based Approaches*)，另一部分则编入本书。本书收录的论文更侧重双语、多语和类比语料应用于翻译教学和语言对比中的技术性问题。

本书的三位主编都是比利时鲁汶天主教大学的研究人员。该大学有一个英语语料库研究中心，其主要研究方向是计算机学习者语料库和跨语言对比研究，主要研究领域有语料库语言学、学习者语料库、对比语言学和翻译研究、词汇学和词典编纂。近十年，尤其是 1998 年之后，其研究成果不断涌现。该中心主任 Sylviane Granger 是鲁汶天主教大学英语语

言和语言学教授，也是本书第一主编。她主攻学习者语料库和双语语料库，是1995年“国际英语学习者语料库”(International Corpus of Learner English)项目的发起人，有很多著述发表于学术期刊或收录于语料库研究文集。第二主编 Jacques Lerot 是鲁汶天主教大学荣誉教授，主要研究领域是语义学和信息处理。第三主编 Stephanie Petch-Tyson 是鲁汶天主教大学英语语料库语言学中心博士生，主要研究兴趣是二语习得、语料分析的心理语言学方法，以及学习者语料库、翻译语料库和双语语料库，已参编论文集多本，有多篇论文被其他文集收录。

全书分为三个部分，分别涉及语料库应用于对比语言学和翻译研究的理论和方法(第一部分)、语料库在对比研究和翻译中的实际应用(第二部分)，以及语料库检索工具的功能和应用(第三部分)。

### 第一部分 语料库应用的理论方法

本书主编 Sylviane Granger 指出，语料库方法适用于对比语言学和翻译研究。随着语料库尤其是多语语料库的创建和应用，对比语言学和翻译研究渐呈合流之势，主要表现为对比语言学和翻译研究都使用相似或相同的多语语料库从事研究，语料创建和使用时也遇到和处理大致相似的问题。要留心作者使用的术语：multilingual corpora 指平行语料库和类比语料库；monolingual corpora 则包括同一语言的原创文本和翻译文本，或者用母语撰写的文本和学习者文本，是单纯的对比语料库。语料库方法在跨语言对比研究上有其局限：其一是语料规模毕竟有限，无法断言它能反映语言的全貌；其二是语料检索和提取基于形式因素，如果没有人工标注，语义范畴或句法结构的提取就很困难。对语料库方法的这些局限性我们应有清醒的认识。

语料库在对比语言学中以何种方式有效应用始终是语料库语言学的—个核心问题，也是第二篇文章的主题。作者 Stig Johansson 从理论上探讨了英语—挪威语双向平行语料库在语言对比中的应用的可能方式。我们知道，平行语料库的创建促成了20世纪90年代对比语言学的复兴，越来越多的学者已认识到翻译语料库在语言对比研究中的重要作用，有人甚至说翻译是语言比较的上好平台，可以在此基础上得出对比研究的结论(P.35)。鉴于此，双向平行语料库可以成为综合对比的理想平台，它所提供的多向对比(P.39)有助于研究者深入了解两种语言的异同。作者指出，双语库的用途可体现在以下三个方面：1. 新型双语词典的编纂；2. 开发新的教学材料，如对比语法；3. 为译员培训提供针对性强的教学材料。

在翻译研究领域，基于语料库的研究为认识、研究和教授翻译提供了新的方法和思路。Laviosa 试图对语料库融入翻译研究引发的研究范式和方法的变化作出说明。作者认为，Gideon Toury 的历史—描写翻译研究和基于语料库的翻译研究原则有很多交合点。比如，二者都采用经验论视角，均依赖直接观察的真实语言素材；二者并不明确界定哪些文本或语料更适合研究，只要有共通的标准即可，而且文本选择通常是随机作出的。作者指出，描写翻译研究和基于语料库的翻译研究可以互相借鉴对方的研究发现、方法论甚至理论成果。可以相互借用假设、数据和研究成果，甚至可以研究相同的课题。此外，作者指出，基于语料库的翻译研究有助于加强翻译描写和翻译实践之间的关联，使翻译产品有效地服务于翻译教学、翻译理论研究和译员培训。

## 第二部分 基于语料库的案例分析

这一部分共收录文章六篇，前两篇关注语料库在对比语言研究的应用问题；第三篇有些特别，谈的是如何通过创建语料库来探讨手势语和日常语言之间的差异；第四和第五篇探讨语料库在翻译研究中的应用；而第六篇则谈及一种独特的语言对比对象：翻译语言和目标语言。

第一篇文章聚焦于英语和荷兰语中的中动结构。作者指出，现有的分析（夺格分析和及物性分析）并没有充分考察语料，所以只能解释某些典型的中动结构。比如，夺格分析错误地断言所有的中动结构都可识别为有两个参与者；而实际情况是英语中的不及物、及物、夺格过程都可以在中动结构中使用，不能局限于词类层面的解释。及物性分析的基本信念是中动结构是主动形式表示被动意义。大致思路是表层只出现一个参与者，而实际有两个。这样，及物性分析就无法说明不及物中动结构，也不能解释材料类主语构成的结构中为什么可以带直接宾语。作者指出，中动结构中使用主动语态动词在语义上的解释还不令人满意，而且受影响 (affected) 这一性质也被过分夸大了。以上难以解释的现象都是在考察语料的基础上发现的，这要比单纯内省式解释更能反映语言使用的真实情形。

第二篇文章使用自制语料库来探讨英语动名词和挪威语中的 det+ 小句组合 (group)。这篇论文使用的研究方式值得注意：语料库自制，规模不大，由一位英语作者的九部小说和它们的挪威语对译文本构成。主要使用以下方法实施对比研究：寻找对应结构、寻找翻译对应项；通过考察原文在译文中的处理方式来说明两种语言里某类结构的对应程度。

第三篇文章也是对比语言分析,但对比的是(佛兰芒)手势语和(荷兰语)书面语。该文旨在通过实验方式验证以下假设:失聪儿童在书面语文本中的某些错误可以推定为受佛兰芒手势语干扰所致。该文主要考察的是指称追踪方式,即叙述者识别某一被叙述人物的方式。观察结果显示:受试通常会在使用荷兰语写作时使用手势语结构。这项研究的实验方式和实验过程值得注意。受事的选择、实验设计、用于对比的叙述语料的收集都有较为严格的计划和程序,实验结果都以语料的方式存储和标注。这项研究的主要发现是:手势语使用者在指称一个物体时使用手势指点,不像荷兰语那样有人称、性和数的差异。换个角度说,这种差异的弥补方式就是指向某一地点,根据被指示物体所处的位置来确定所指。通过类比语料分析,手势语和书面语的差异以及手势语对书面语的影响可以得到全面的反映。

第四篇文章探讨了法语直陈式过去时在英译中的损益问题。一般认为,法语的直陈式过去时和英语的一般过去时有明显差异。通过语料分析,该文却得出了不同的结论:这两种过去时在复杂性、多义性和体意义不确定性上有很多相似的特征。语料分析还说明,法语直陈式过去时译成英语时,直陈式过去时所表达的持续性和重复性在英语译文中往往要通过显性体标记来实现;有时甚至要保留原文中体的不确定特征,即使用一般过去时来翻译。该文的研究基于自制的法英平行语料库,只包含文学文本和新闻报道文本。

第五篇文章依据原型理论探讨了马来语和英语习语的互译问题,试图观察语料库能够以什么方式辅助建立英语—马来语习语翻译的基本原则。这项研究使用的是类比语料分析。作者指出,尽管两种语言在同一对比单位如熟语(phrasology)上有不同的涵盖范围,语料库方法还是可以应用。具体做法是找到独立于具体语言的范畴。语料库可以说明哪种词组类型在两种语言中存在,并能帮助我们建立最相关的范畴,以供比较分析。为此,该文在概念隐喻基础上使用原型理论来研究两种语言间习语的对比和翻译问题。该文进行的语料分析显示,在翻译习语时,要把表层形式在多大程度上可靠,即比较相关语言在多大程度上有相似的概念基础。相似程度越高,直译的可接受程度就越高;反之,翻译时就要建立一个共有概念范畴。该项研究表明,平行语料库能帮助我们区分哪些是两种文化共有的概念隐喻,哪些是文化特有的、概念化了的语言现象;而且,语料库也能帮助我们确定歧义性习语中字面意义和比喻意义的相

对频率。

第六篇文章探讨的是翻译语言的特征和翻译共相 (translation universals)。所分析的案例是芬兰语儿童文学中的非定式结构。这篇文章做的是语言对比,但不是两种语言的对比,而是某一语言里两种变体的对比:芬兰语翻译文本和芬兰语原创文本。对比的内容是两种文本在句法上的异同。相关的研究方法在文中有比较详细的说明,这里不赘述。

### 第三部分 跨语言语料研究工具及应用

这一部分介绍了跨语言语料研究工具的功能和应用。共收录论文四篇,谈的都是语料库软件工具在教学、翻译培训和词典编纂方面的应用问题。

关于语料库检索软件,第一篇文章介绍了两个常用语料库检索软件 Paraconc 和 Multiconcord。Paraconc 由 Barlow 制作(该软件的演示版可从网上下载,国内已有人使用),该软件的特点是可以进行多达四种语言的同时检索,或者是一个原文的三个译本的检索。这一软件的特点是可以灵活定义语言、索引行的大小、标注符号的隐显,而且支持通配符检索。Multiconcord 也是一个在 Windows 窗口下运行的软件。这种软件和 Paraconc 在检索功能上相似,但检索结果在呈现方式上不同。另外,Paraconc 可以检索纯文本格式的文件;Multiconcord 则需要一个 Minimark 程序来最低程度地标记文本,如 <p> (段落)和 <s> (句子)。

作者指出,语料库在语言学习中运用的一种主要方式是设计学习点 (learning points),然后通过检索软件在语料库中提取并呈现与这些学习点(即教学内容)相关的索引,供学生练习和讨论,教师随后作总结。该文还简要介绍了平行语料库在翻译行为研究(译者的个性和偏好)、词典编纂和语法研究中应用的可能性。这些都有助于我们更多地了解平行语料库检索软件的用途和适用范围。

语料库在翻译培训上有多大作为?这是第二篇文章关心的问题。该文介绍了两类语料库在这方面的用途。一是 CCBT (corpus created by translators),即出自译者(翻译受训学员)的译语文本库。库中的语料可供受训译员调用,使用检索软件便能同时提取原文和多个目的语对应译文。这类语料还可按照主题检索和调用,甚至还可以建立一个跟踪性 CCBT,以考察受训译员的翻译能力在某段受训期内发生了多大程度的发展和变化。另一类语料库是 CCFT (corpus created for translators)。注意,这类语料实际上没有译文语料,它只提供相似题材的资料供翻译参考。



该文作者通过实验说明,使用 CCFT 要比单纯使用传统资源(如词典)对翻译培训更有用。基本原理是在 CCFT 中使用通配符检索,找到目标词在目的语文本中具体使用的情形,这样无疑使译文遣词更为地道,更适应目的语语境。这实际上相当于一个有专门用途、可以即时调用和参考的类比语料库。

就翻译教学和译员培训而言,我们通常所说的类比语料库主要起评估作用,帮助翻译教员对学生译文质量作出评估,并将译文修改建议反馈给学生;当然,它还可以用来探究翻译过程。

双语语料用于翻译教学也是语料库应用的一个重点和难点。第三篇文章探讨了如何让双语语料库作为一个平台实施词汇教学。这篇文章关注的是计算机科学文献里层出不穷的动词,这些动词在教材甚至在专业词典里常常被忽视。作者收集语料的方法值得借鉴:通过互联网收集电子文本建立语料库,语料收集来源多、渠道广、省时省力。语料的收集、来源、对齐、检索和处理在文中都有详细的说明,读者可以从中得到不少启发。这项研究完全可以延伸到普通的词汇教学过程,甚至可以考虑以对比索引为素材设计填空练习。

最后一篇文章介绍了习语词典的设计和编纂。该文研究有两个背景:习语研究越来越受重视;计算机语料库在语言研究的作用日渐突出。在这种背景下,语料库就成了词典编纂者不可或缺的工具。这项研究像是一个基于语料库的词典编纂计划,编纂内容是法语的习语。其特别之处是习语的来源:一是法国本土法语,二是魁北克法语。习语词典的编纂基于新闻文本语料,每一习语词条的编码内容涉及非功能变体、功能变体和使用频率分析。这种编纂方式很特别,想必能够方便同一语言不同变体的使用者有效地交流信息。

总之,这本书展现了近些年来在语料库技术支撑下对比语言学和翻译研究的新进展,书中的多层面和多向度研究说明,对比语言学和翻译研究有很强的互补性,它们可以使用同一个语料库,也可以相互借鉴各自的研究发现。这一领域的研究散见于期刊、文集和专著里。除 Johansson 和 Oksefjell 主编的《语料库和跨语言研究》(1998)和王克非等著的《双语对应语料库:研制与应用》(2004)外,专门将这类研究结集出版的还不多见。从这个意义上说,本书的出版在一定程度上满足了学界在这方面的需求。本书把语内变体(目的语原创文本和翻译文本;地域变体)、手势语和书面语都列入对比语言研究的范围,相信这能帮助我们扩大语言

对比研究和翻译研究的视野。此外，书中的案例分析方法和语料库工具软件的介绍和应用也能为语料库研究和应用提供有益的参照。

### 参考文献

- Johansson, Stig & Signe Oksefjell (eds.). 1998. *Corpora and Cross-linguistic Research: Theory, Method and Case Studies*. Amsterdam: Rodopi.
- McEnery, Tony & Andrew Wilson (eds.). 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- 王克非等, 2004. 双语对应语料库: 研制与应用. 北京: 外语教学与研究出版社。

**秦洪武** 曲阜师范大学外语学院教授, 北京外国语大学博士

**柯 飞** 北京外国语大学中国外语教育研究中心研究员、博士

## Contents

List of contributors	7
Preface	9
<i>Sylviane Granger, Jacques Lerot &amp; Stephanie Petch-Tyson</i>	
<b>I. Theoretical Approaches</b>	
The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies?	17
<i>Sylviane Granger</i>	
Contrastive linguistics and corpora	31
<i>Stig Johansson</i>	
Corpora and Translation Studies	45
<i>Sara Laviosa</i>	
<b>II. Corpus-based Case Studies</b>	
On the middle construction in English and Dutch	57
<i>Kristin Davidse &amp; Liesbet Heyvaert</i>	
English gerund clauses and Norwegian <i>det</i> + infinitive / <i>at</i> clause constructions	75
<i>André Hantson</i>	
Flemish Sign Language and Dutch: Syntactic differences due to a different modality as exemplified in reference tracking	91
<i>Mieke Van Herreweghe &amp; Myriam Vermeerbergen</i>	
Loss and gain in English translations of the French <i>imparfait</i>	105
<i>Hélène Chuquet</i>	
A prototype based approach to the translation of Malay and English idioms	123
<i>Jonathan Charteris-Black</i>	
Nonfinite constructions in Finnish children's literature: Features of translationese contradicting translation universals?	141
<i>Tiina Puurtinen</i>	

### III. Cross-linguistic Tools and Applications

- Parallel concordancing and its applications 157  
*Philip King*
- Corpus-based applications for translator training: Exploring the possibilities 169  
*Lynne Bowker*
- Teaching English verbs with bilingual corpora: Examples in the field of computer science 185  
*Natalie Kübler & Pierre-Yves Foucou*
- Towards a new dictionary of idioms 207  
*Elizabeth Dawes*

## Contents

List of contributors	7
Preface	9
<i>Sylviane Granger, Jacques Lerot &amp; Stephanie Petch-Tyson</i>	
<b>I. Theoretical Approaches</b>	
The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies?	17
<i>Sylviane Granger</i>	
Contrastive linguistics and corpora	31
<i>Stig Johansson</i>	
Corpora and Translation Studies	45
<i>Sara Laviosa</i>	
<b>II. Corpus-based Case Studies</b>	
On the middle construction in English and Dutch	57
<i>Kristin Davidse &amp; Liesbet Heyvaert</i>	
English gerund clauses and Norwegian <i>det</i> + infinitive / <i>at</i> clause constructions	75
<i>André Hantson</i>	
Flemish Sign Language and Dutch: Syntactic differences due to a different modality as exemplified in reference tracking	91
<i>Mieke Van Herreweghe &amp; Myriam Vermeerbergen</i>	
Loss and gain in English translations of the French <i>imparfait</i>	105
<i>Hélène Chuquet</i>	
A prototype based approach to the translation of Malay and English idioms	123
<i>Jonathan Charteris-Black</i>	
Nonfinite constructions in Finnish children's literature: Features of translationese contradicting translation universals?	141
<i>Tiina Puurtinen</i>	



### III. Cross-linguistic Tools and Applications

- Parallel concordancing and its applications 157  
*Philip King*
- Corpus-based applications for translator training: Exploring the possibilities 169  
*Lynne Bowker*
- Teaching English verbs with bilingual corpora: Examples in the field of computer science 185  
*Natalie Kübler & Pierre-Yves Foucou*
- Towards a new dictionary of idioms 207  
*Elizabeth Dawes*





**Corpus-based Approaches  
to Contrastive Linguistics  
and Translation Studies**