



外研社翻译研究文库



语料库与译者培养

CORPORA IN TRANSLATOR EDUCATION

(意)

Federico Zanettin

Silvia Bernardini

Dominic Stewart

编



外语教学与研究出版社

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS

◀ 外研社翻译研究文库 ▶

语料库与译者培养

CORPORA IN TRANSLATOR EDUCATION

(意) Federico Zanettin
Silvia Bernardini 编
Dominic Stewart

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 - 1838

© St. Jerome Publishing 2003

All Rights Reserved, including those of translation into foreign languages. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without either the prior written permission of the Publisher or a license permitting restricted copying issued by the Copyright Licensing Agency (CLA), 90 Tottenham Court Road, London, W1P 9HE. In North America, registered users may contact the Copyright Clearance Center (CCC): 222 Rosewood Drive, Danvers MA 01923, USA.

First published by St. Jerome Publishing Ltd. Manchester, United Kingdom

This edition is authorized for sale in the People's Republic of China only, excluding Hong Kong SAR, Macao SAR and Taiwan, and may not be bought for export therefrom.

版权所有。未经许可不得以任何方式,在世界任何地区,以任何文字,作全部和局部之翻印、仿制或转载。

本书由英国圣哲罗姆出版有限公司授权外语教学与研究出版社出版影印版。此版本仅限在中华人民共和国境内(不包括香港、澳门特别行政区及台湾省)销售,不得出口。

图书在版编目(CIP)数据

语料库与译者培养 = Corpora in Translator Education / (意)扎内廷(Zanettin, F.)等编. — 北京:外语教学与研究出版社, 2007.1

(外研社翻译研究文库)

ISBN 978 - 7 - 5600 - 6185 - 6

I. 语… II. 扎… III. ①外语教学—文集—英文 ②翻译—文集—英文 IV. H3 - 53

中国版本图书馆 CIP 数据核字 (2006) 第 154348 号

出 版 人: 李朋义

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: 10.5

版 次: 2007 年 1 月第 1 版 2007 年 1 月第 1 次印刷

书 号: ISBN 978 - 7 - 5600 - 6185 - 6

定 价: 16.90 元

* * *

如有印刷、装订质量问题出版社负责调换

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

外研社翻译研究文库

专家委员会

(按姓氏笔画排列)

王东风	王克非	王宏印	毛思慧
文 军	方梦之	申 丹	申雨平
冯庆华	冯志伟	吕 俊	朱志瑜
朱纯深	庄绎传	刘士聪	许 钧
孙艺风	孙致礼	李亚舒	李德凤
李德超	杨自俭	连淑能	汪榕培
张柏然	张美芳	陈宏薇	陈国华
陈德彰	范守义	柯 平	郭建中
黄忠廉	黄振定	辜正坤	程镇球
谢天振	廖七一	谭载喜	潘文国
穆 雷			

外研社翻译研究文库*

现有书目:

About Translation P. Newmark

《论翻译》

A Practical Guide for Translators G. Samuelsson-Brown

《译者实用指南》

Can Theory Help Translators? A Dialogue Between the Ivory Tower and the Wordface A. Chesterman & E. Wagner

《理论对译者有用吗? 象牙塔与语言工作面之间的对话》

Corpora in Translator Education F. Zanettin et al.

《语料库与译者培养》

Corpus-based Approaches to Contrastive Linguistics and Translation Studies S. Granger

《基于语料库的语言对比和翻译研究》

Crosscultural Transgressions: Research Models in Translation Studies II, Historical and Ideological Issues T. Hermans

《跨文化侵越——翻译学研究模式 (II): 历史与意识形态问题》

Electronic Tools for Translators F. Austerlühl

《译者的电子工具》

Intercultural Faultlines: Research Models in Translation Studies I, Textual and Cognitive Aspects M. Olohan

《超越文化断裂——翻译学研究模式 (I): 文本与认知的译学研究》

Method in Translation History A. Pym

《翻译史研究方法》

Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis (Second Edition) C. Nord

《翻译的文本分析模式: 理论、方法及教学应用》(第二版)

The Translator's Turn D. Robinson

《译者登场》

Translated! Papers on Literary Translation and Translation Studies J. S. Holmes

《译稿杀青! 文学翻译与翻译研究文集》

* 本文库采用开放式结构, 今后还将陆续添加其他有影响的翻译研究著作。

Translating Literature: Practice and Theory in a Comparative Literature Context
A. Lefevere

《文学翻译: 比较文学背景下的理论与实践》

Translation and Empire: Postcolonial Theories Explained D. Robinson

《翻译与帝国: 后殖民理论解读》

Translation and Language: Linguistic Theories Explained P. Fawcett

《翻译与语言: 语言学理论解读》

Translation and Literary Criticism: Translation as Analysis M. G. Rose

《翻译与文学批评: 翻译作为分析手段》

Translation and Nation: Towards a Cultural Politics of Englishness

R. Ellis & L. Oakley-Brown

《翻译与民族: 英格兰的文化政治》

Translation and Norms C. Schäffner

《翻译与规范》

Translation, Power, Subversion R. Álvarez & M. Vidal

《翻译, 权力, 颠覆》

Translation Today: Trends and Perspectives G. Anderman & M. Rogers

《今日翻译: 趋向与视角》

Unity in Diversity? Current Trends in Translation Studies L. Bowker et al.

《多元下的统一? 当代翻译研究潮流》

Western Translation Theory: from Herodotus to Nietzsche D. Robinson

《西方翻译理论: 从希罗多德到尼采》

20 世纪 90 年代翻译界所经历的文化转向给翻译研究带来了崭新的视角和方法。在这场浩大的文化转向运动中,翻译研究逐渐从“规定性研究”的束缚中解脱出来,“描述性研究”逐渐得到译界认可和越来越多的应用。在描述性研究中,翻译文本不再被视作原文文本的衍生,逐渐从翻译与文化的边缘走向中心。多元系统理论认为,翻译文本是目的语文化的一个重要组成部分,是与目的语文化产生互动的一个独立子系统。该观点丰富并扩大了翻译研究的领域,让“离开原文文本独立探讨翻译文本”(study translated text in its own right)的研究在方法论层面上有了支撑和立足点。语料库翻译研究的产生便基于这一背景。语料库翻译研究有两大支撑点:语料库语言学和描述性翻译研究¹,其中后者与翻译的描述转向息息相关。

一般认为,Mona Baker 于 1993 年发表的文章²是带动语料库翻译研究发展的开山之篇,此后一批学者利用语料库探讨翻译规范和翻译普遍性等问题,内容涉及单语可比语料库、双语平行语料库,甚至多语语料库等。从研究类别来看,基于语料库的翻译研究应从属于 Holmes 提出的翻译学研究分支中纯研究类别下的描述性研究,再具体而言,此类研究应归属于翻译产品研究。随着语料库翻译研究的深入,研究者逐渐发现语料库在培训译者领域也有着重要用途,这使得语料库研究得以进入应用翻译研究的类别,充实了原有的研究维度,并扩充了语料库在翻译研究与实践中的应用领域。

早期探讨把语料库运用于翻译教学的研究主要见于 1998 年 *Meta* 杂志所出版的语料库翻译研究特辑,其中有学者探讨翻译的普遍性问题,也有学者开始关注语料库在翻译教学中的应用。如今,这批学者依然活跃于这一领域,有些人在本书中仍然撰文参与讨论。本论文集的研究者都是翻译界的活跃学者,他们从不同角度对使用语料库进行翻译教学提供了自己第

1. Laviosa, Sara. 2002. *Corpus-based Translation Studies*. Amsterdam-New York: Rodopi. p5.

2. Baker, Mona. 1993. "Corpus Linguistics and Translation Studies: Implications and Applications" in Mona Baker, Gill Francis, and Elena Tognini-Bonelli (eds) *Text and Technology: In Honour of John Sinclair*. Amsterdam-Philadelphia: John Benjamins. pp. 233-252.

一手的实践体验和理论上的敏锐见解。

本论文集所收集的论文来自 2000 年在意大利 Bertinoro 召开的 CULT2K³ 会议。论文集共收集论文 10 篇, 由 Federico Zanettin, Silvia Bernardini 与 Dominic Stewart 三位学者编辑而成, 主要探讨使用语料库作为译者培训资源及语料库在翻译教学课堂上的应用。

本书首篇论文由三位编者联手撰写, 主要探讨语料库发展的历程, 对语料库应用于翻译教学作了历史回顾, 还探讨了用语料库研究翻译的普遍性等问题, 随后深入剖析了语料库作为翻译辅助工具与语言教学资源等方面的运用。文章对语料库作为译者培养的辅助工具进行了详细分析, 指出该论文集之所以命名为“译者培养”(translator education)而非“译者培训”(translator training)有以下考虑:“编者确信语料库能帮助未来译者提升个人的自主性和灵活性, 此经验不仅具有教育价值, 而且能使译者更具专业优势。”值得指出的是, 作者认为使用语料库进行翻译教学最大的教育价值在于能够唤起译者思考(thought-provoking), 而非仅提供现成答案(question-answering)。该论文集的开篇之作给随后的论文提供了有效的背景知识, 并简明扼要地对论文集集中的其他论文作了归纳综合, 这对于那些对语料库不甚熟悉的读者不啻为有用的启蒙之作。

本论文集的其他论文主要来自学者们的第一手研究体验, 是语料库运用于翻译教学等相关领域的成果汇报。由于双语文本的缺乏, 对齐句子耗时, 以及软件不足等诸多缺陷, 过去语料库在翻译教学领域内的研究主要局限于单语语料库(monolingual corpus)。一些学者对利用平行文本进行翻译教学的可能性多持保留态度, 主要也是基于以上考虑。在这方面, Jennifer Pearson 的论文“在译者培训环境中运用平行文本”一文可以说是这一领域的有益探索。作者利用 Scientific American 杂志及其法文译文建成一小型平行语料库, 用法文创作的原文文本建成一可比语料库, 并用该平行语料库对文化词(主要是大学译名)的翻译进行了探究。作者发现, 不译、漏译甚至改译等现象在该语料库中表现得极为普遍。通过研究, 作者认为平行语料库可以帮助学生考察专业译者所采用的翻译策略, 从而形成自己的翻译方针(translation guideline)。本文从某种程度上证明了平行语料库相对于可比语料库的潜在优势, 然而由于研究涉及的语料较少, 且考察的文化词范围相对狭窄, 所得出的结论仍有待进一步研究。但是, 作为一项尝试性研究, 仍不失其探索的价值。

3. CULT2K 全称为 Corpus Use and Learning to Translate (2000)。本次在意大利 Bertinoro 举行的研讨会是 CULT 的第二次会议。

接下来的两篇论文主要讨论了利用语料库培训译者在翻译专门用途语言 (LSP) 与术语方面的研究。Natalie Kübler 的研究主要汇报作者在本科生与研究生中所进行的一个实验项目。该实验考察在训练 LSP 译员中使用电子语料库和语料库处理工具, 探讨训练学生利用可比语料库、平行语料库以及互联网进行术语抽取、编辑和其他自然语言处理的技巧。众所周知, 术语对于 LSP 翻译极为重要, 译者往往必须自己编辑术语条目。传统上译者主要利用相关杂志或词典等工具编辑术语, 而新术语往往鲜见于传统词典, 语料库以及网络的出现提供了另一有效途径, 使术语编辑更为灵活。研究者通过对学生进行此类训练, 发现业界对此类培训的正面反馈: 语料库和语料库索引软件在很大程度上帮助译者提高了工作效率。因此作者认为, 对学生进行此类训练显得极为必要和迫切, 尤其是在技术翻译领域。

另一篇探讨语料库与 LSP 的论文作者是 Belinda Maia。她在论文“某些语言比另一些更对等”中主要讲述了自己指导学生进行专业术语收集过程中所遇到的情况和难题。Maia 在文中谈到, 相对于其他常规文本来说, LSP 平行文本往往比较稀少, 且翻译质量不高。在实际的培训过程中, 可比单语语料库显得更加实用。另一方面, 作者在指导学生进行葡萄牙语与英语的术语编辑过程中发现, 小语种 (指葡萄牙语) 文本往往得来不易, 大大加重了术语收集和编辑的难度。最后作者指出, 随着翻译越来越具专业化特征, 技术翻译的需求日益要求译者进行术语编辑以提高翻译质量, 因此对译者进行“术语与翻译”方面的培训极为必要, 事实上, 此类培训课程在欧美其他国家的翻译教学中也的确日益普遍。

Krista Varantola 的论文主要探讨了一次性语料库 (disposable corpora 或 ad hoc corpora) 在译者培训中的作用。本文主要基于作者指导学生在翻译工作坊中所进行的实验。实验中, 学生利用互联网作为语料库资源进行搜索、编辑、查询等技巧训练。作者指出, 一次性语料库结构上虽然简单, 但实际运用中可以帮助学生提高翻译绩效 (translation performance)。同时, 作者也强调掌握如何编辑并且有效使用语料库是现代翻译能力的重要组成部分。最后作者列出了现代翻译能力所必须的各项技巧, 由“语料库编辑”与“使用语料库知识”两大板块构成。

Ana Frankenberg-Garcia 与 Diana Santos 合著的论文为“介绍 Compara 这一葡萄牙语-英语平行语料库”, 正如其标题所示, 论文主要向读者介绍 Compara 这个葡萄牙语和英语的双向语料库, 该语料库主要参照“英语-挪威语平行语料库”的结构建设而成。作者对该语料库的结构、文本选

择标准、对齐和标注等问题进行了较为详尽的阐述,并对该语料库运用于翻译课堂教学、编写翻译教学材料的可能性作了展望。

Tony McEnergy 与 Paul Baker 合著的文章“语料库:翻译和多语处理”与翻译教学似乎相关性不大。该文主要基于作者对英国翻译社以及译者所进行的一项调查研究。研究者主要考察当前翻译研究和实践领域使用语料库资源的状况。调查发现,在英国,虽然业界有翻译小语种的需求,如中文、阿拉伯文、印地语和其他东南亚语言,但专业译者往往对小语种语料的获得感到困难。在这方面,欧洲本土语言在语料方面比起非欧洲本土语言显得更具优势。因此,研究者认为给小语种提供适当的语料显得极为迫切。作者最后介绍了由英国 EPSRC⁴ 所资助的为小语种提供单语口语、书面语及双语平行语料的项目。希望借助该项目的努力,小语种的研究境况能够得到改善。

Lynne Bowker 与 Peter Bennison 的“学生翻译档案库和学生翻译跟踪系统——设计、发展与应用”一文主要讨论学生翻译作品库的建立,包括收集、管理与研究学生的翻译作品。作者从翻译培训的角度给研究学生翻译作品提供了初步的方法。文章介绍了该系统的组成部分和建立过程,并对学生的作品分析作了实例说明。作者认为,语料库很早之前便被用于研究中介语,学生的翻译作品从某种程度上也具有自身特色,值得探讨。作者在题目中没有使用“语料库”而使用“档案库”这个词,主要是认为“语料库是根据某些研究标准严格编辑的可机读文件,应与随机收集的文本区别开来”。本文从一个崭新的角度提供了如何研究学生翻译作品的方法,开创了新的研究方向。

Kirsten Malmkjær 的文章“论译者培训中语料库的伪颠覆性应用”主要对使用语料库进行译者训练进行了反思。她认为,语料库主要“反映过去语言的使用模式,不能代表未来语言发展的趋势”。翻译是一种创造性活动,而语料库所收集的例证可能会对译者进行误导并且扼杀这种创造力。作者通过两个案例分析说明了语料库在这方面的局限性,认为有必要探索其他颠覆性的方法来使用语料库,让译者知道何时该打破某些翻译规范,进行创造性翻译。笔者认为,这种逆向思维可以为语料库在翻译教学中的应用与研究发展提供宝贵的思路和启发,成为新的研究方向。

最后一篇文章的作者是 Stig Johansson,其文章“对语料库的反思及其在跨语言研究中的使用”主要对上述学者的研究进行简要总结,并对语料

4. EPSRC 全称为 Engineering and Physical Sciences Research Council, 是英国政府资助工程学和物理学研究的主要机构。

库的发展和应用进行反思和展望。文章探讨了语料库建设的两个模型——钻石模型与星形模型。作者认为,语料库为学习和研究提供了机遇,使学习带上了研究的特征,使两者的联系更加紧密。文章最后指出,当今运用语料库培训译者是美好时机,在使用过程中只要译者富于想象力和洞察力,语料库将是一个功能全面、有益且令人振奋的资源。

综观该论文集我们不难发现,语料库与翻译教学的结合已经日益密切。该文集所探讨的主题已经超出了传统翻译教学的研究领域,覆盖诸多方面,包括专门用途语言、术语研究、学生译文评价等。研究者所探讨的语料库概念也有所突破,不仅包括传统上所讨论的单语或双语语料库,也包括一次性语料库、学生翻译作业档案库等。从方法论层面上而言,研究也突破了利用单语语料库进行译者培训的做法,开始探讨双语语料库、学生自建语料库、小语种语料库等方法的其他潜在用途,如 Malmkjær 的文章所谈到的对语料库的颠覆性用法。Malmkjær 一文对语料库所进行的反思,提出了语料库的局限性和不足,指出语言发展不应受到语料库的约束并主张对语料库进行颠覆性使用以打破翻译规范,令我们耳目一新。作为一本会议论文集,本书亦有其不足之处,如某些论文主要为介绍性文章,显得泛泛而谈,研究性较为薄弱。有些论文主要基于小范围的实验和个人见解,缺乏研究的严密性和应用的广泛性。个别研究项目在论文发表时仍在进行当中,尚无定论。在翻译教学领域,语料库使用覆盖面相当广阔,例如利用双语语料编写翻译教材,进行教材革新;探讨语料库在翻译课程设置方面的作用,以更好地适应翻译人才培养目标;如何利用语料库加强翻译课堂互动,提高课堂有效性等等。而这些问题在该论文集中尚未涉及,令人感到遗憾。然而,瑕不掩瑜,该论文集提供了国外学者利用语料库进行翻译教学的最新视角,值得我们好好借鉴。

翻译学科要发展就必须借鉴其他各学科的优势,翻译学跨学科的性质要求我们不可以偏于一隅,必须从其他学科的发展中吸取养分。语料库方法作为语言学的一个分支,也借此契机被引入翻译研究,探究翻译现象,应用于翻译教学等方面。从该论文集我们可以清晰地发现,语料库研究已经超出了传统的翻译纯研究类别,不再局限于探究翻译普遍性等纯理论问题,开始步入应用研究的类别。作为应用翻译研究的一大子类别,翻译教学长期以来都处于边缘化的地位,缺乏有效的方法论指导,缺乏大量构思严密的实验手段,缺乏得心应手的研究工具,等等,这种现象在国内尤为突出。在国内,传统翻译教学多基于教师的经验式、感悟式、师徒式

的传授方法，在翻译教材、课堂教授和学生自主学习等方面都存在不足之处。语料库的出现在某种程度上恰好能弥补这一不足。在语料库教学环境中，学生通过发现式学习和数据驱动的学习方式训练掌握翻译技巧、提高翻译技能、熟悉翻译职业特征。在这方面，有些学者倾向于带领学生自建语料库，主要基于以下考虑：学生通过参与建设语料库，能够在拟真的情景下认识专业译者是如何进行翻译并解决翻译所碰到的难题的。通过在语料库真实材料中发现译者如何对某些术语和文化词进行翻译，学生能够逐渐形成自己的翻译策略。

语料库与翻译教学的结合是翻译教学领域的一次革新，计算机的日益普及和互联网的不断发展使资讯日益丰富。在这个信息爆炸的时代，翻译教学有必要走下神坛，认真考察业界的最新发展动态和需求，并以此调整翻译人才培养的目标与方式。随着技术的发展和社会分工的日益专业化，社会对翻译人才的要求也越来越高，越来越专业化。该论文集的研究正是基于这一社会大背景对社会需求的回应，这对国内翻译人才的培养是一大挑战，同时也是机遇。为了培养合格的翻译人才，我们不可以墨守成规，有必要借鉴国内外在这一领域的有效探索，形成自己的培养特色和方针。该论文集给我们提供了西方同类教育机构在翻译教学人才培养方面的有效探索和最新发展。它山之石，可以攻玉，值得我们认真借鉴。

穆 雷

广东外语外贸大学高级翻译学院教授，香港浸会大学翻译学哲学博士

刘康龙

广东外语外贸大学高级翻译学院 2004 级硕士研究生，香港中文大学文学院翻译系在读博士生

Contents

Corpora in Translator Education: An Introduction Silvia Bernardini, Dominic Stewart, Federico Zanettin	1
Using Parallel Texts in the Translator Training Environment Jennifer Pearson	15
Corpora and LSP Translation Natalie Kübler	25
Training Translators in Terminology and Information Retrieval Using Comparable and Parallel Corpora Belinda Maia	43
Translators and Disposable Corpora Krista Varantola	55
Introducing <i>Compara</i> the Portuguese-English Parallel Corpus Ana Frankenberg-Garcia and Diana Santos	71
Corpora, Translation and Multilingual Computing Tony McEnery and Paul Baker	89
Student Translation Archive and Student Translation Tracking System Design, Development and Application Lynne Bowker and Peter Bennison	103
On a Pseudo-subversive Use of Corpora in Translator Training Kirsten Malmkjær	119
Reflections on Corpora and their Uses in Cross-linguistic Research Stig Johansson	135
Notes on Contributors	145
Index	149

Corpora in Translator Education

An Introduction

SILVIA BERNARDINI

DOMINIC STEWART

FEDERICO ZANETTIN

1. Applied corpus-based translation studies

Language corpora, i.e. principled collections of texts in electronic format, are increasingly being used as a resource in linguistics and language-related disciplines, and corpus linguistics is now firmly established as a research area and a methodology. One of the many fields where corpora are having a growing impact is translation, both at a descriptive and a practical level. This book is principally concerned with the use of corpora as resources for the translator and as teaching/learning aids in the context of the translation classroom – an area which could be termed “applied corpus-based translation studies” in terms of the ‘map’ of translation studies first proposed by Holmes in ((1972) 1988) and later elaborated by a number of scholars, notably Toury (1995).

The last decade of the twentieth century saw an unprecedented growth of public and private institutions devoted to the training, or, better, *education*, of translators and interpreters throughout Europe and the world (Caminade and Pym 1995). The papers in this volume, which grew out of presentations at the CULT2K Conference¹, record work in progress in some of the liveliest among these institutions. Though the case studies reported and activities proposed reflect curricula that vary according to local educational characteristics, the learning contexts they address are in many ways similar. Such institutions rarely know for certain whether their students will work for large governmental translation agencies, as free-lance literary translators, as interpreters, or as language service providers in some other capacity, and teaching them to translate usually goes hand-in-hand with perfecting their working languages. The use of corpora in these contexts stands at the interface between translation and language

¹The second *Conference on Corpus Use and Learning to Translate* took place in Bertinoro, Italy, in November 2000. See Bernardini and Zanettin (2000) for a selection of papers from the previous CULT Conference (1997). Several papers are also available online <http://www.sslmit.unibo.it/cultpaps/paps.htm>.

teaching/learning, the expectation being that competent use of corpora and corpus analysis tools will enable students to become better language professionals in a working environment where computational facilities for processing text have become the rule rather than the exception.

From this perspective the papers in this volume make, we believe, an important collective statement at a moment when applied corpus-based translation studies are finally becoming established as a framework for research and pedagogy. Like corpus linguistics a few years ago, they can now be said to have come of age (Svartvik 1992), and signs of this newly-attained maturity abound in this collection. Here we shall draw readers' attention to those features we consider particularly central to the way this field of study is progressing, focussing on how they converge or diverge from such related areas as descriptive corpus-based translation studies, computer-aided translation, and corpus-based language teaching.

2. Translator education at the interface

2.1 Corpus-based translation studies

In descriptive translation studies, corpus-based approaches have been used to investigate whether and how translations differ from their source texts, or from original texts written in the target language, and how specific languages and genres as well as translators' stylistic preferences affect translations. Beginning in the early 1990s (following Baker 1993), work in this area has drawn on and brought together aims and methods from descriptive translation studies (cf. Toury 1995) and descriptive corpus linguistics (cf. Sinclair 1991). Translations are seen as subject to socio-cultural norms which are manifested intertextually: the latter can thus be identified through an empirical approach which focuses on the detection and analysis of systematic features of translations in the context of the receiving cultures. Corpus linguistics, as a methodology which focuses on the identification of recurrent patterns of linguistic behaviour in actual performance data, provides the appropriate tool to test hypotheses about norms and regularities in translated texts. Given its emphasis on the target side of translation, work within this paradigm has concentrated on comparable target language corpora, i.e. collections of translated texts and of original texts in the same language (e.g. Laviosa 1997, Olohan and Baker 2000), as well as on parallel or translational corpora, i.e. collections of translated texts and of their source texts in another language (e.g.

Øverås 1998, Kenny 2001).

Descriptive corpus-based translation studies have led to a better understanding of translation phenomena, and helped raise awareness of what is involved in translating. These insights can benefit not only literary translators and translation theorists, but also scholars in related fields (cf. Malmkjær this volume). For reasons that we hope are illustrated by this volume, translator education too can profit from empirically-based discussions of these theoretical and descriptive issues.

2.2 Computer-aided translation tools

The influence of corpora has been felt not only in the field of translation description, but also in the more practice-oriented area of professional translator support. Virtually all translators in the Western world now use computers in their everyday work, and process texts electronically. At the same time, research in natural language processing, language engineering and machine translation has been spurred by technological developments, in particular the advent of personal computers and the Internet. As a consequence, translators and other language service providers have access to a wide range of computational resources. They can exploit a vast array of electronic reference tools, such as on and off line mono- and bilingual dictionaries and term banks, as well as the largest multimedia encyclopaedia in existence, i.e. the Web. They can obtain real-time help and support from fellow translators through mailing lists and newsgroups. And they can exploit dedicated software suites known as Computer Assisted Translation (CAT) Tools or Translator Workbenches, that combine a text processor with a system for managing 'translation memories' (effectively, a parallel corpus of source texts and their translations) and, optionally, other tools such as a terminology management system, a 'proper' machine translation system, etc. In this last regard, we can note the growing interest in example-based MT systems, partly as a result of the influence of corpus linguistics, which offer promising alternatives to 'traditional' rule-based MT systems (Somers 1998).

Such resources are clearly relevant to the education of prospective translators. It will be an advantage for any future language professional to be skilled in managing state-of-the-art reference tools and translation software, and in integrating different types of computational and managerial resources in the workflow. The emphasis here is generally on non-literary, technical translation, particularly localization and translation in the social

services sector – an area in which McEnery and Baker's contribution to this volume highlights the importance of multilingual corpus tools. But we should not be deceived into thinking that technology is the exclusive province of the 'technical' translator. A working knowledge of CAT tools is nowadays in high demand, and is likely to become necessary for all language service providers in the near future.

2.3 Corpora in language teaching and learning

The uses of corpora in translation teaching contexts are not limited to translation classes proper, where corpus work can be relevant prior, during or after a translation task (Aston 2000). Corpora can also be used in other courses forming part of the curriculum of translation students, such as second language learning (Bernardini 2000) and terminology (Pearson 1998), and such work can complement translation activities in the narrow sense, developing capacities and competences that are far from marginal to translator education.

The use of corpora in language learning contexts was pioneered by Tim Johns, who introduced concordancing into the foreign language classroom in the 1980s. Besides enabling language professionals such as lexicographers and material writers to produce better reference and learning materials, and allowing language teachers to create classroom activities based on real examples, he showed how corpora could provide learners with direct access to virtually unlimited language data. His methodology of *data-driven learning* (Johns 1991) is a corpus-inspired approach to language pedagogy where learners are encouraged to develop their own hypotheses about textual data and to devise their own strategies for extracting information from corpora. This view of *learning as discovery* seems equally applicable to translator education, and can be found in a number of the papers in this volume. Anna Frankenberg-Garcia and Diane Santos, for instance, present exercises in which learners query parallel texts in the *Compara* corpus in order to highlight the variety of strategies adopted by professional translators and the socio-cultural norms guiding their behaviour (Toury 1995). This approach to parallel concordancing seems to have much in common with Johns' experiment in 'reciprocal learning' (Johns 1997). Here, pairs of British and French students discuss parallel concordances derived from an English-French (non-directional) corpus, and help each other to infer generalisations about specific lexicogrammatical differences between the two languages.