

教育部人文社会科学研究“典籍英译译者文体分析与文本的译者识别”(编号: 10YJAZH034)项目资助

“辽宁省高等学校优秀人才支持计划”(编号: WR2014010)阶段性成果

大连外国语大学科研基金“译者研究的多维视角”(编号: 2013XJYB01)项目资助

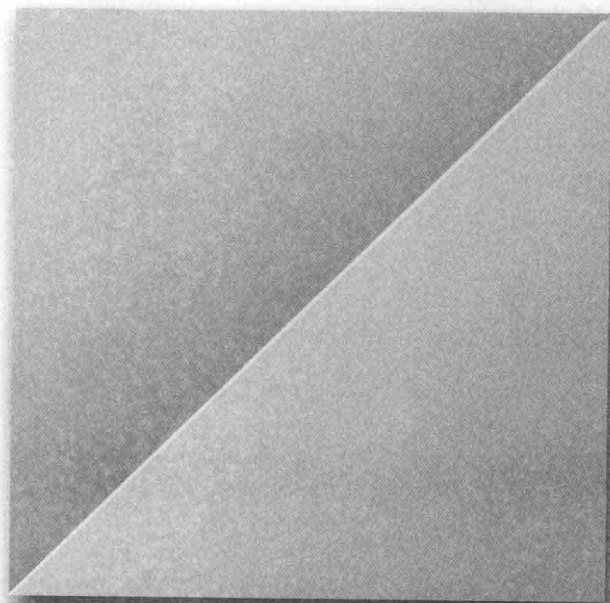


霍跃红 著

译者研究

——典籍英译译者的文体分析与文本的译者识别

中西書局



霍跃红 著

译者研究

——典籍英译译者的文体分析与文本的译者识别

中西書局

图书在版编目(CIP)数据

译者研究：典籍英译译者的文体分析与文本的译者识别 / 霍跃红著. —上海：中西书局，2014. 10
ISBN 978-7-5475-0735-3

I. ①译… II. ①霍… III. ①英语-翻译-研究
IV. ①H315.9

中国版本图书馆 CIP 数据核字(2014)第 223642 号

译者研究

——典籍英译译者的文体分析与文本的译者识别
霍跃红 著

责任编辑 贺寅

装帧设计 贺寅

上海世纪出版集团

出版 中西书局(www.zxpress.com.cn)

地址 上海市打浦路 443 号荣科大厦 17F(200023)

发行 上海世纪出版股份有限公司发行中心

经销 各地新华书店

印刷 常熟市兴达印刷有限公司

开本 700×1000 毫米 1/16

印张 21.5

版次 2014 年 10 月第 1 版 2014 年 10 月第 1 次印刷

书号 ISBN 978-7-5475-0735-3 / H·037

定价 55.00 元

教育部人文社会科学研究“典籍英译译者文体分析与文本的译者识别”(编号: 10YJAZH034)项目资助

“辽宁省高等学校优秀人才支持计划”(编号: WR2014010)阶段性成果

大连外国语大学科研基金“译者研究的多维视角”(编号: 2013XJYB01)项目资助

序

霍跃红教授的专著《译者研究——典籍英译译者文体分析与文本的译者识别》即将付梓,嘱我为她的呕心沥血之著作序。这部专著是在她博士论文的基础上,经过几年反复加工而成的,作为她的博士生导师,写一篇序言是当仁不让的。不过,我总觉得在她的专著即将问世的时候,对这部专著再评头论足的话,似乎已经有点不合时宜,因为读者可以自己来衡量作品的成败得失、判断它的价值。我在这里谈自己的几点感想。

典籍英译研究在 21 世纪以来受到的关注越来越多,出现的成果也越来越多,这跟中华文化加速走向世界是密切相关的。不过,从总的发展趋势来看,正如作者在前言中提到的那样,“国内的典籍英译研究有两个薄弱之处:一是综述性研究多,定性研究为主,有宏观把握,却缺少数据支持;二是译本研究、译者研究、翻译策略研究等个别现象研究多,普遍联系比较少;而且多以已故译者和他们的译作为研究对象,对那些仍然健在、对典籍英译发展贡献重大的翻译家则缺少深入研究。”

这部专著正是针对这两个薄弱之处作出了努力。第一,既在描写理论框架下对典籍英译译者进行了定性研究,又运用了语料库语言学、文献计量分析等科学的手段,对中国典籍英译的译者和作品进行了定量研究。第二,基于语料库对中国典籍英译核心译者进行的研究,尤其是对译者文体的比较研究,得出了一些具有普遍性意义的结论,对今后的相关研究具有启示的意义。

在我的印象中,霍跃红教授是一个头脑聪颖、心地善良、爱家敬业、举重若轻的人,在不经意之间就能成就大事。她不是一个只会自己闭门读书的人,在师生之间人缘极佳,课堂上深受学生的欢迎。她在完成博士论文以后,获得了教育部和省、市等诸项课题,与同事和学生分享自己的科研成果、共同进行深入的研究。

年届不惑的霍跃红正处在大展宏图的黄金年华,我衷心地预祝她在教学和科研中再接再厉,取得一个又一个新的成果,尤其在中国典籍英译的译者研究领域中脱颖而出,成为又一名领军人物。

汪榕培

2014年3月15日于苏州幽兰斋

前 言

本书系教育部人文社科规划基金项目“典籍英译译者的文体分析与文本的译者识别”(项目编号: 10YJAZH034)的结题成果。

本研究在描述翻译学的框架下,结合语料库语言学、认知语言学、文献计量学和计算机科学等学科知识和方法,关注了典籍英译这个领域;运用文献计量分析的方法发现典籍英译领域里的核心译者并计算了译者的影响力,提出了本色译者的概念,并从认知视角研究了翻译过程,然后运用语料库和机器学习的方法研究译者和译者文体。语料库为机器学习提供分类指标特征,机器学习为语料库提供崭新的计算机处理和研究方法,语料库技术和机器学习方法是本书研究译者文体并进行文本译者识别的技术路线。本研究力图在以下方面有所创新:

一、关注典籍英译译者群体,首创了本色译者的概念。

二、拓展译者研究的维度和方法,在描述性研究的基础上,引入其他学科的理论和方法,如语料库语言学、文献计量分析等,定性、定量研究相结合,弥补以往研究主观性强、价值判断成分多的不足;为翻译史的书写增加新的研究视角和研究方法。

三、完善了典籍英译核心译者发现系统,首次把文献计量学中的引文分析法和 h 指数运用到翻译研究中,以测定文学译者的影响力。人为地确定核心译者主观性强,容易以偏概全,甚至会引起学术界无谓的争端。

四、尝试了基于语料库的翻译研究由描述性——解释性——应用性

的转变。提出了把翻译共性作为辨伪标准,用来识别伪译本;把译者文体的特征集作为文本表示的特征集,用来推定不明译本的译者。

从2010年获批教育部人文社科规划基金项目之前就开始孕育,经过数年的思考、整理,这本书终于要呱呱坠地。回首、环顾、展望,带着一颗沉甸甸的感恩的心:真心感谢我的导师汪榕培教授,他的勤恳、包容,不仅促使愚钝的我终于完成学业,而且对我现在的学术研究始终鼓励、帮助,不仅作为我的研究对象,还在百忙之中为我撰序;大连外国语大学英语学院的领导和同仁,大连外国语大学科研处及其他职能部门的领导和同事,还有我指导的几届研究生们,彭丽、刘冀、孙颖轶、刘丽萍……他们都为这本书的付梓提供了大量的帮助。大恩无法言谢,我只有力争做好,并一如既往地走下去。

作者虽年已不惑,仍是学界新人,经验及水平都甚为有限,种种错谬,敬请前辈专家指正。

目 录

序	1
前言	1
1 导论	1
1.1 研究背景	1
1.2 主要概念界定	2
1.3 问题提出	2
1.4 研究意义	3
1.5 技术路线	4
1.6 研究内容	4
2 国内外研究现状	7
2.1 译者研究综述	7
2.2 译者研究的新趋势	9
3 典籍英译核心译者发现及译者影响力研究	23
3.1 典籍英译概述	23
3.2 基于文献计量分析的典籍英译核心译者发现	35
3.3 文学译者的影响力研究	54
4 描写理论框架下典籍英译译者的定性研究	75
4.1 译者研究的理论视角	75

4.2 译者研究的认知视角	87
5 基于语料库的翻译普遍性研究	99
5.1 语料库翻译研究	99
5.2 翻译普遍性特征研究	105
5.3 翻译普遍性的应用性研究——伪译本识别	167
6 基于语料库的译者文体比较研究	203
6.1 国外研究现状	204
6.2 国内研究现状	206
6.3 影响译者文体形成因素的定性分析	208
6.4 文言志怪小说译者文体研究	217
6.5 古诗《孔雀东南飞》译者文体比较研究	234
6.6 现代散文《荷塘月色》译者文体分析	262
7 文本的译者识别	272
7.1 文本识别的理论依据	272
7.2 文本译者识别的概述	275
7.3 译者文体的机器表示	277
7.4 文本译者机器识别的实验	291
7.5 小结	311
8 结语	312
参考文献	316
附录 A CLAWS7 (C7) POS Tageset 词性赋码集	327

1 导 论

1.1 研究背景

我国政府大力提倡中华文化走出去,典籍英译作为弘扬中华民族传统文化的重要途径,已经受到社会各界越来越多的关注。把中国优秀的典籍翻译成英语,对于弘扬民族文化、促进东西方文化融合、保持中国固有的文化身份,都有十分重大的现实意义。对典籍英译领域的成果及时进行总结、梳理,以便发现问题和提出对策就尤为重要。

明末清初以来,西方译者英译中国典籍已有 400 多年的历史。中国典籍的翻译与研究是西方传统汉学的中心内容。传统的西方汉学以译介中国文、史、哲典籍为核心,以选译、摘译、编译、个人或零星组织的翻译为主要形式;现、当代的西方汉学则转向现当代中国现实社会的研究,如费正清、安乐哲等。

中国学者自觉向西方译介中国典籍也有 100 多年的历史。改革开放以后,中国典籍英译快速发展,经历了从选译到全译,从零星翻译到结集成套翻译的过程。从 2002 年起已连续召开了八届全国典籍英译研讨会,关于典籍英译的研究也越来越深入。

国内的典籍英译研究有两个薄弱之处:一是综述性研究多,定性研究为主,有宏观把握,却缺少数据支持;二是译本研究、译者研究、翻译策略研究等个别现象研究多,普遍联系比较少;而且多以已故译者和他们的译作为研究对象,对那些仍然健在、对典籍英译发展贡献重大的翻译家则缺少深入研究。

本书的宏观研究范围是典籍英译,具体为从事典籍英译理论与实践的译者及译者文体。

随着语料库语言学与当代计算机应用技术的发展,以及各种学科知识的交叉、融合,关于典籍英译领域的译者研究朝着科学化、客观化、系统化的方向发展。语料库与描述翻译学的结合,为研究译者提供了理论框架,而自然语言处理又为译者研究提供了技术支持。本书旨在探讨译者文体、译者文体的量化、以及基于机器学习的文本译者识别问题。

1.2 主要概念界定

译者,对应英语中的 translator,也称翻译家,泛指任何从事把一种语言转换成另一种语言的人;狭义的概念是指从事文学翻译的人。本书中应用的是狭义的定义。

文体,即风格,是作者在词汇、语法、修辞等方面体现出来的写作特点,是在后天经过不断学习和反复实践形成并巩固下来的写作习惯,具有一定规律性,既有个体的特殊性,又有相对的稳定性。

译者的文体或风格,是译者表达在语言和非语言中的个性特征(thumb-print features),是译者的声音,而不是原文或原作者的风格或文体的再现;译者的文体是译者运用语言的特点及其与他人不同的语言习惯,是译者个人偏爱的表达形式和重复出现的语言行为方式,而不是个别的或一次性的偶然例子。

文本的译者识别,是借用了文本作者识别的概念,通过机器学习代表译者风格的特征,让计算机自动完成对文本的译者识别和认定。

1.3 问题提出

随着计算机科学的发展,人文学科包括语言学、翻译学开始越来越受到计算机技术的影响和支持。研究译者必须依托于译者所翻译的产品即译作。无论多先进的或前端的研究,也不能脱离开文本去寻找生产者即译者的特征或风格。译者文体表现于客观的语言层面上,译者文体分析

就是对这些客观的语言现象的确认。从对文本的分析上升到对人的分析,即译者的分析。有关译者及译者风格,本书对下列问题进行了探讨:

(1) 在众多的典籍英译译者中如何发现核心译者,确定研究对象?

当致力于译者研究的学者面对众多译著丰厚、对文化交流贡献巨大的译者时,选择自己的研究对象,就成为—个值得思考的问题。目前以译者为研究对象的研究,大都是基于研究者主观判断或个人喜好、个人对译者情况的了解。人为地去确定核心译者主观性过强,容易以偏概全,甚至会引起学术界无谓的争端。那么如何在有限的时间内、有限的精力内,确定在某个领域影响力最大的译者来作为研究对象?

(2) 什么是译者文体?译者文体表现在哪些方面?

以往的翻译研究,以原文为主要参照,探讨的是作者风格在译文中的再现问题。那么译者是否具有游离于原文及作者之外的特征?如果肯定译者的文体,该怎样进行验证?本书分析了受原文影响比较小、译者下意识流露出来的一些用词特点、句子结构、语篇衔接手段等特征,界定了译者的风格。

(3) 如何提取译本的文体特征?应用这些特征能否实现计算机对译者的自动识别,从而为未知译者身份的文本识别提供有效的方法?

基于语料库统计出来的译者风格特征除了进行描述和分析外,是否可以得到验证和应用?也就是说,如果能从词、句子、语篇层面统计分析出译者的风格,这些风格被量化后,输入给机器。机器是否能够根据这些特征作为指标识别出某段或某些译文的译者?如果能,说明了什么?如果识别率很低,需要怎样改进?

1.4 研究意义

针对1.3中所提出的问题,考虑到译者研究是翻译史研究中的重要环节,又是个薄弱环节,而翻译界关于译者风格研究常采用的举例法、内省法具有不可避免的局限性,本书尝试构建译者风格辨析和文本的译者

识别模式。该模式建立在描述翻译学和语料库语言学的理论框架之上,以自然语言处理为技术支持,分析译者的个性化文体特征。对这些个性化文体特征整理后上升为一种文本译者识别的特征集。

因此,本书具有一定的理论意义和较强的现实意义。

1.5 技术路线

本书所遵循的技术路线如图 1.1 所示。自然语言处理是总的技术背景。发现核心译者;之后对核心译者进行定性分析;在建立核心译者译作语料库的基础上,利用数量统计方法进行基于语料库的核心译者文体比较研究;最后利用译者文体在语音、词汇、语篇等特征集上的差异性,进行文本译者的识别实验,提供未知译者身份识别的方法。

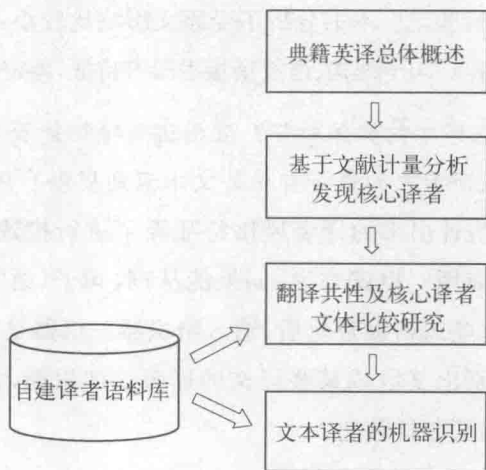


图 1.1 技术路线

1.6 研究内容

首先在概述典籍英译研究现状的基础上,确定典籍英译译者为研究

对象;具体化到译者的文本研究,分析他们的翻译风格,即文体;通过语料库发现其风格特征,并进行量化;把这些特征作为机器学习的资料或特征来识别文本的译者。

在第一章“导论”中,首先从我国典籍英译研究的现状、译者风格科学分析的需求出发,提出了本书所要研究的科学问题,界定了本书的研究范围和研究内容,明确了本书的研究意义。最后,对本书所采用的技术路线和研究内容做了说明。

第二章的文献综述部分,主要介绍描述翻译学框架下基于语料库的译者研究情况以及文本译者识别的尝试等。

第三章是典籍英译核心译者发现及译者影响力研究。首先以典籍英译为例,基于文献计量学对译者群体进行研究,确定了核心译者群。运用引文分析法和h指数测定了文学译者的影响力。这不仅为译者研究提供了客观发现核心译者的方法,也提供了了解译者影响力和活跃度的参考。

第四章探讨了典籍英译译者定性研究的框架。把目标译者放入其教育背景、认知环境下去考察译者的翻译情况。从译者对原文本的选择和翻译策略出发,提出了本色译者的概念;从翻译过程考虑,探讨了译者研究的认知视角。

第五章“基于语料库的翻译普遍性研究”,首先以《木兰诗》四个时期的英译本为语料,从历时的角度,通过语际对比和语内类比两种方式,证实了显化、简化、范化作为翻译普遍性特征存在于《木兰诗》各个时期的译本中,并且证实了不同时期译本的翻译共性表现形式不同,进而揭示了社会、文化等超语言因素对翻译共性的影响。然后,以伪译本《奥特朗托城堡》为例,验证了翻译共性中简化和显化特征作为辨伪标准的适用性,开创性地进行了基于语料库翻译研究的应用性尝试。

在第六章“基于语料库的译者文体比较研究”中,建立了核心译者平行语料库和对比语料库,并在所建立的平行语料库和对比原料库的基础上对不同译者的文体风格进行了词汇、句子、语篇、修辞、情感流动等语言层面的观测,并从译者自身素质、翻译文本、翻译策略的选择以及译者所

处的社会历史文化背景等非语言层面的因素考察了译者文体。研究的语料包括除古典戏剧外的各种文学典籍体裁,如长篇叙事诗《孔雀东南飞》、《木兰诗》,文言志怪小说,古诗词《诗经》等,最后还对现代散文《荷塘月色》的不同译本进行了计算文体学研究,比较研究了不同译者的文体。

在第七章“文本的译者识别”中,基于第六章语料库分析的结果,提取了代表译者文体特征的三个层面的 31 维特征集,进行文本表示。建立向量空间模型,运用五种文本分类方法进行了三个阶段的实验,并对实验结果进行了分析。三个阶段的实验分别是:运用基于单词频率的文本识别方法进行文本的译者识别,获得了一定的准确率;扩大译者文体特征的维数,进行了基于多层面特征的文本译者识别实验,识别的正确率有显著提高;基于信息增益特征选择进行实验,使实验复杂度降低,时间缩短,识别率提高。

第八章是结论部分,总结已完成的研究,并对以后的工作设想进行了说明。

2 国内外研究现状

2.1 译者研究综述

本书属于翻译史研究中的译者研究。

Jenny Williams & Andrew Chesterman 在《路线图——翻译研究方法入门》一书的第一章简要介绍了翻译研究的 12 个领域,每一领域的研究现状,可供研究的主要课题,以及可以大体了解该领域研究状况的主要参考书籍。翻译史研究被明确列举为第六个研究领域,并指出翻译史研究的主要问题是四个 wh-问题,也就是“谁在翻译? 翻译了什么? 翻译的原因? 翻译的策略?”^①

穆雷概括了翻译研究的类型:

- (1) 国别研究: 研究某一国家或地区的翻译现象。
- (2) 时期研究: 专门研究某一特定时期的翻译现象。
- (3) 译者研究: 专门研究某位译者的翻译活动及其翻译理论与实践,或者研究具有某种特定关系的翻译家群体,如时空关系、地域关系等。
- (4) 理论研究: 如不同时期的理论著作选编,从中可看到翻译理论发展的系统化趋势。
- (5) 翻译机构和出版社翻译贡献研究: 比如中国同文馆译书局翻译作品研究,江南制造局译书研究,联合国翻译机构研究等等。
- (6) 专题分类个案研究: 比如佛经翻译史研究、《圣经》翻译史研究。^②

^① Williams J, Chesterman A. The Map: A Beginner's Guide to Doing Research in Translation Studies[M]. Shanghai: Shanghai Foreign Language Education Press, 2004.

^② 穆雷. 重视译史研究 推动译学发展——中国翻译史研究述评[J]. 中国翻译, 2000(1).