

语料库翻译学：研究对象与研究方法

黄立波 西安外国语大学 朱志瑜 香港理工大学

摘要：语料库翻译学近20年的发展历程是一个就研究对象和方法论“提出问题—尝试解决问题—不断消除错误—提出新问题”周而复始的过程。就研究对象而言，主要是对翻译语言/文本整体特征的描写，如翻译共性、译者风格等。研究方法方面，经历了从单一的单语类比或双语平行模式到二者相结合的综合模式。近年来，翻译带来的语言互动与变化成为新的研究对象，根据实际需要建立的多重复合对比模式成为方法论发展的趋势。语料库翻译学的发展历程说明：一种新研究范式要得到认可并取得长足发展，必须不断加深对研究对象的认识，不断革新研究方法，拓展其研究领域。

关键词：语料库翻译学；研究对象；研究方法

中图分类号：H59

文献标识码：A

文章编号：1672-9382(2012)06-0028-09

1 引言

英国科学哲学家Karl Popper在其《客观知识：一种进化方法》(Objective Knowledge: An Evolutionary Approach)一书中给出了一个科学研究方法的图式：

问题1 尝试性理论 错误消除 问题2
(Popper, 1979: 164)

科学研究就是“提出问题—尝试解决问题—同时不断消除错误—再提出新问题”这样一个周而复始的过程。语料库翻译学已有近20年历史，经历了从研究途径到一套方法论，再到一种新研究范式的发展过程，同样遵循了这一模式。本质上，语料库方法就是“假设检验”(hypothesis testing)，具体包括7个步骤：(1)提出假设；(2)设定研究目标；(3)检验假设；(4)分析数据；(5)对发现进行理论阐述；(6)将假设精确化；在此基础上(7)为将来的研究提出新假设(Laviosa, 2002: 2)。Laviosa将语料库翻译学的发展划分为3个时期：第一个时期是语料库翻译学的开

端(1993—1995)；第二个时期是确立在翻译研究中的地位(1996—1999)；第三个时期，是语料库在各个语言文化中的延伸(2000—今)，其中第二个时期既包括描写研究，也包括应用研究(2004: 6; 2011: 14)。从Laviosa的介绍基本上可以看到语料库翻译学的发展脉络及其在各个阶段的研究内容、特点和发展趋势。本文以方法论探索过程为主线，概述语料库翻译学的理论来源及其不同阶段在研究对象和研究方法方面的思考与发展，在此基础上指出此领域的发展带给翻译研究的启示。

2 语料库翻译学的理论支撑

语料库翻译学的发展建立在两块土壤之上：语料库语言学和描写翻译研究。语料库语言学从语言研究角度为语料库翻译学提供理念和方法上的理据，如以真实文本为研究对象、从假设检验入手、研究使用中的语言、关注典型的语言使用模式、注重概率性统计分析等。描写翻译研究则是从翻译研究角度提供具体研究对象和理论方面的支持，如将翻译文本作为独立的研究对象、以目标语文本为导向、关注

作者简介：黄立波，西安外国语大学英文学院副教授，博士，研究方向：语言与翻译、语料库语言学，E-mail: libohuang2003@yahoo.com.cn。
朱志瑜，香港理工大学中文及双语学系教授，研究方向：翻译理论、中国翻译理论发展史，E-mail: ctcychu@inet.polyu.edu.hk。

翻译文本的规律性语言模式或特征等。

语料库语言学注重真实的语言使用,借助语言学理论和计算机工具对真实文本进行统计、分析和描写,遵循“让语料说话”的原则,达到对语言现象的客观认识。但语料库研究方法很长一段时间以来,一直将翻译文本排除在其语料选择之外,原因在于翻译文本被认为是源文本的衍生品,不具备语言代表性。Baker最早提出将翻译文本作为语料库的材料,并坚信大规模的翻译语料库一定能够帮助翻译研究者更好地认识翻译,“以前所未有的规模探究指导翻译行为的原则以及制约翻译运作的因素”(1993:235)。20世纪90年代初期,以单语语料库的构建为模式、以翻译文本为语料的翻译语料库(包括平行和类比两类)开始产生。

描写翻译研究是Baker语料库翻译学思想的另一个重要来源,这一研究学派抛开“对等”,抛开源文本,以“规范”(norms)为核心概念,关注目标语文化中的“翻译”事实,并对翻译事实进行客观描写。描写翻译研究跳出“对等”的局限,将关注点放在翻译文本整体上,尤其是对翻译行为规律性的描写上,而不是关于源文本和目标文本关系的个案研究或对比分析上;强调建立独立的学科分支、完善的方法论和明确的研究步骤,强调翻译行为的概率性,并对此做出合理解释,同时这一研究过程具有可观察性和可重复性(同上:240-241)。无论从理念、对象或方法看,语料库翻译学与描写翻译研究密不可分,甚至可以说,描写翻译研究是语料库翻译学的主要思想来源。

语料库语言学与描写翻译研究的研究对象都是真实使用的文本,本质上都是假设检验,研究对象都放在语言的规律性特征上,方法上以文本对比为基本模式,这些共同点成为二者结合建立语料库翻译学研究分支的基础。

3 翻译文本(或译文)能不能独立作为翻译研究的对象?

回溯式翻译研究,即以译文是否忠实于原文或是否与原文对等为标准对译文质量进行评估,很长一段时间以来一直是翻译研究的传统模式。语言学翻译研究以“对等”为核心概念工具,采用形式分析法从词汇、语法、句法、语义、语用、语篇、功能等层面分析单一源文本和目标文本的语言,关注各个层面上的“对等”,期望达到对翻译的科学解释。在这一思想主导下,译文常被视为由原文衍生的“二手”文本,被排斥在其他各类研究的语料选择之外。

但这种以“对等”为核心的研究观念在一定程度上限制了翻译学科的进一步发展。

描写翻译研究前瞻式的研究模式开始将翻译文本作为独立的研究对象,正是在这一背景下,Baker(1993)一反传统,倡导打破“对等”这一思维定势,将关注点放在翻译文本上,提出用语料库的方法来研究翻译文本。Baker在此文中一个突出的理念就是,“大规模语料库(包括原文及译文)将为翻译研究者提供前所未有的机会,让他们能够直接观察自己的研究对象,探索翻译不同于自然生成语言或其他任何文化互动形式的原因”(同上:235),比如,翻译语言与源语言的对比、翻译语言与目标语非翻译语言的类比、翻译语言与二语写作等中介语的对比研究等。在研究方法上,语料库翻译学借鉴了描写翻译研究的方法,创造了“翻译共性”、“译者风格”等概念,从可观察的现象入手,对翻译文本进行描写,逐步达到对各类深层关系的认识,并最终达到对无法直接观察的过程的重构,通过描写达到最终对翻译现象的解释。Baker认为,“翻译文本同样记录了真实的沟通行为,就此而论它既不优于也不亚于其他任何形式的语言”(同上:234),理应作为翻译研究的对象。

4 研究对象的界定和研究方法的初步提出

语料库翻译学发展初期面临的首要问题,就是界定自己的研究对象,建立自己独立的研究方法。研究对象和内容方面,Baker(1993)提出,用语料库研究翻译主要是从翻译文本入手,从以下几方面展开:(1)翻译语言普遍特征(universal features of translation)(后称翻译共性,translation universals),具体包括显化(explicitation)、消歧(disambiguation)和简化(simplification)、语法性(“grammaticality”)、避免重复、过度凸显目标语语言特征,以及某些特征的特定分布;(2)特定社会-文化语境下的翻译规范;(3)翻译中介过程(the intermediate stages of translation)研究、翻译单位的大小与性质,以及对等的类型与层次等(同上:243-248)。可见,语料库用于翻译研究初期的设想,不仅对翻译文本本身,还有对文本背后、翻译过程以及操作方式的关注,不仅包括语言,而且涉及社会、文化、认知等方面。但从后来的发展看,前两类课题得到了较为充分的关注,其中共性研究成果最丰富,但对第三类课题的关注则较少,原因在于:尽管语料库是一个丰富的资源,但由于语料自动标注技术的

限制,数据提取主要侧重于语言形式方面,对语义、语用层面的数据只能通过间接方式获取,限制了此方面的发展。

在方法论方面,语料库翻译学的初衷是要建立一套有别于单一源文本-目标文本对等与否的回溯式研究模式,而以目标语文本为取向,将翻译文本整体视为研究对象,考察翻译文本不同于非翻译文本的特征。但实际研究中,源文本不可能完全被忽略掉。为了有别于传统上以“对等”为核心概念工具的研究,语料库翻译学提出了“翻译共性”这一核心概念,考察翻译语言表现出的普遍规律。尽管如此,语料库翻译学诞生初期在研究对象界定方面仍不能完全抛开对源文本角色的考虑,主要原因在于:这一新途径在具体研究对象界定方面主要是建立在前人传统研究(如Vinay & Darbelnet, 1958/1995; Nida & Taber, 1969; Blum-Kulka & Levenston, 1983; Vanderauwera, 1985; Blum-Kulka, 1986等)的基础之上。例如, Baker对显化的最早界定是“与特定源文本或(目标语中)原创文本总体上相比较,(目标文本)所表现出显性程度的显著提高”(1993: 243)。从这一界定可看出, Baker当时所讨论的各种翻译共性既有语际对比(interlingual),也有语内类比(comparable)。

以显化为例,这一概念最早由Vinay & Darbelnet提出,指“在目标文本中将源文本中的隐含信息加以明示(的过程或结果),此类信息在源文本中可以根据语境或场合获得”(1958/1995: 342)。这样的显化属于语言系统差异造成的强制性(obligatory)语际显化。译文相对于原文的长度也被作为显化的一个标志,如Nida & Taber指出,好的译文往往多少会比原文长,主要因为在翻译过程中译者会通过明示原文中的隐含信息,在一定限度内增加译文的冗余度(1969: 164-165)。这里所关注翻译文本的特点也是相对于源文本而言,但实际上,分属不同语系的源文本和目标文本在长度上

可比性并不强,如英语以词计算长度,汉语通常以字来计算长度,这样的比较其实并不能说明实质性的异同。还有对具体显化策略或技巧的讨论,如Vanderauwera (1985)指出了译者常用的10种显化策略。此外还有对句法显化的考察,如Blum-Kulka从翻译中衔接与连贯的转换入手提出了显化假设(the explicitation hypothesis):“翻译过程会使译文相对于原文更加冗长,其[一种]表现形式为衔接方式的显化程度提高”(1986: 19)。这些研究都主要是语际比较基础上翻译文本相对于源文本表现出的形式方面的特点,都是 Baker基于语料库的翻译共性研究直接或间接的思想来源,研究内容与方法在之后的语料库显化研究中都有不同程度体现,其共同点是关注翻译文本中的规律性语言使用模式。与 Baker方法不同的是,这些研究都是以源文本为标准来衡量译文。

在 Baker (1993) 倡导将语料库应用于翻译研究时,语料库还只是被视为一种新工具,而非一种方法论,基于语料库的翻译研究还是属于描写翻译研究的范畴。主要原因在于,当时语料库工具发展程度有限,研究者们对语料库的潜力、研究内容、方法和步骤尚未有一个深入认识。由此可见,语料库翻译学也是对原有翻译研究范式的传承,所不同的是:着眼点不同,方法不同。先前的研究以“源文本”为中心,采用人工的文本对比方法,而语料库翻译学则以“目标文本”为导向,采用计算机数据提取、统计与分析技术,将关注焦点放在翻译语言上。在语料库翻译学的探索时期,传统翻译研究方法的影子依然存在,语料库方法是被作为翻译研究的一条新途径,但系统的、连贯的方法论尚未建立,相关的实证研究也未展开。

5 目标语单语类比语料库翻译研究模式的建立

对比模式方面,为有别于传统的源文本-目标文本语际对比,

语料库翻译学在初期以类比语料库(comparable corpus)为基础建立了目标语单语类比模式,即对比目标语中翻译文本与非翻译文本的文本特征。

Baker提出,语料库翻译学要向前发展,当务之急是着手建立明确、连贯的方法论(1995: 223)。她首先提出了类比语料库的设想, Baker指出,平行语料库的贡献是使翻译研究从规定转向描写;而类比语料库对翻译研究的贡献在于:在不考虑所涉及源语言或目标语言的情况下,识别翻译文本中的特定语言模式,探索翻译本质(同上: 231; 234)。后一种研究模式显然与传统上以探讨源文本与目标文本关系为核心的研究方法不同,标志着目标语单语类比模式的语料库翻译研究方法正式建立。 Baker (1995) 还介绍了从类符-形符比(type-token ratio)和词汇密度(lexical density)统计方面进行翻译共性研究的思路。由于当时语料库技术发展有限,计算机统计手段还不十分完善(例如牛津大学1993年开发的检索工具Microconcord,即WordSmith的前身,当时也只能实现KWIC检索和词频表的生成等功能,对多种语言的兼容性也不够),到1996年之前,语料库翻译学还仅限于理论上的构建,真正意义上的语料库实证研究尚未正式展开。

1996年,第一个基于语料库的翻译研究在曼彻斯特大学展开,第一个英语单语类比语料库建成,并用于“简化”现象的探讨(参见Laviosa-Braithwait, 1996)。具体做法是建立翻译英语语料库(The Translational English Corpus, TEC),将英国国家语料库(BNC)中具有类比关系的子库作为参照库开展语内类比研究。 Baker进一步明确了用翻译语料库考察翻译共性特征的方案,并对显化、简化、范化(或保守化)和整齐化(levelling out) 4类共性特征一一做了界定(1996: 180-184)。但定义本身并没有讲清楚这些特征到底是相对于源文本还是目标语中的原创文

本,或许二者都有。事实上,甚至Baker本人的思想也存在不一致的地方:她一方面指出翻译共性研究主要是借助类符-形符比和词汇密度等计算机统计手段,来探究翻译文本相对于目标语中原创文本而体现出的一些规律性语言特征(1995:235-238);另一方面,在讨论显化、简化和范化时却分别以源文本的文本长度、译者对源文本标点符号的改变,以及源文本及源语言的相对社会地位等语际对比参数作为讨论依据(Baker,1996:180-184)。无论是类比还是语际共性的讨论,对应的实证研究并未展开。

这一时期,在研究对象和研究方法方面还存在许多不统一的地方,语际共性与类比共性的讨论兼而有之。此后,单语类比模式与双语平行模式的翻译共性研究同步发展。共性研究内容还扩展到了传统化、标准化、范化、整齐化、净化、“独特项假设”(unique items hypothesis)、干扰(interference)、非典型搭配、不对称假设(asymmetry hypothesis)等方面,但基本上还是Baker(1993)提出的研究范围。

1998年,Laviosa宣称:“基于语料库的翻译研究方法正在发展成为一种连贯、综合、丰富的范式”(1998:474)。此种论断的依据是,语料库翻译学在短短五年内不仅界定了自己的研究领域,初步建立了自己的方法论,而且拥有了自己的研究团队,并在不同语对中开展了一系列大规模的理论探讨和实证描写。

6 单语类比与双语平行模式的

并存

尽管Baker(1995,1996)的初衷是抛开源文本,倡导单语类比的翻译共性研究模式,但实际上这一时期以类比语料库和平行语料库为基础的研究都有。从这一点可以看出,当时的研究者对研究对象还没有一个统一认识,因此在方法上存在差异,此种各自为政的局面大

约一直持续到新世纪前10年中期。在翻译共性和译者风格研究中都出现了分别使用单语类比与双语平行模式的研究,不同的研究模式产生了不尽相同的研究结果和发现。

6.1 翻译共性研究

以下首先以显化为例,分别概述这一时期的类比与语际共性研究,然后总结两类研究在方法上的特点。

6.1.1 类比模式

类比模式的语料库显化研究主要是对Baker所倡导模式的应用和验证。语料库翻译研究的关键在于如何将研究假设与语言层面的表现形式联系起来,以具体的语言检测项为切入点来考察翻译现象。

类比显化研究中的第一类是对形式词汇或标记的考察,如Olohan & Baker(2000)通过对叙述动词后选择性that在翻译英语语料库(TEC)和英国国家语料库(BNC)中的频次进行考察;Olohan(2001、2002a、2002b)考察的同样是对形式成分的选择性显化,具体切入点包括补语标记词that、关系代词wh-/what、动词补语小句中的to be、that从句中should省略与否等选择性句法成分和词汇成分在翻译文本和目标语非翻译文本中的使用情况。Olohan认为翻译文本中对选择性句法成分的过多使用,目的在于明示各种语法和词汇关系,对选择性词汇,如人称代词的使用频次低,说明译文为了明示语义信息往往会以名词化结构代替原文中的人称代词形式。Olohan(2003)在TEC和BNC基础上,对英语翻译文学和原创文学中各类缩写形式(contracting forms如's, 'll, 'd, 've, 't, 're等)的种类和频次考察发现,翻译文学在两方面均低于非翻译文学,说明翻译文本表现出显化的倾向。这些研究在表面上都支持了Baker的假设,但实际上,是否使用选择性that或省略形式等对于文本的显隐程度并无大的影响,反而在很大程度上与译者的习惯有关。另外还可以看出,尽

管Baker所倡导的单语类比模式是翻译描写研究的一个新视角,但在对所描写现象的解释中,源文本还是一个不容忽视的重要因素,否则就只能说明翻译文本与非翻译文本有差异,仅此而已。

第二类是对句法形式的考察。如Puurtinen(1998,2003a,2003b)对翻译芬兰语的语言特征以及意识形态规范进行考察发现:非翻译芬兰语作品中非定式结构的出现频次低,而译自英语的翻译作品中,非定式结构的使用频次明显高于当代非翻译芬兰语作品。由于非定式结构的使用降低了译文的明晰度,增大了词汇密度和信息负载,大量的非定式结构背离了芬兰语非翻译儿童文学的语言规范,这一发现显然与Baker等所提出的翻译文本的显化相矛盾。Puurtinen(2004)的研究表明,无论翻译或非翻译芬兰语儿童文学中,连接词(包括连词、特定副词、关系代词)频次的增高或降低趋势均不大明显,一部分连接词在翻译文本中表现出较高频次,而另一部分则在非翻译文本中频次较高,因此不能完全支持原有的显化假设。Puurtinen认为造成显化不明显的原因可能与翻译中的直译倾向有关;另一方面,儿童文学子库成分的复杂性,也是导致显化呈现多样性的因素。

以上考察结果说明,在具体语言形式使用方面,翻译文本与目标语非翻译文本相比较表现出一些规律性差异,说明翻译文本与原创文本的确存在系统差异,各类考察有些支持原有假设,有些则与原有假设相反。

6.1.2 平行模式

同一时期,以平行语料库为基础从语言转换入手,可以双向考察语际显化,其结果不同于类比显化的单语模式。

第一类主要是从词汇或句法双向转换入手,如Schmied & Schffler(1997:21)以克姆尼茨英-德翻译语料库(Chemnitz)为基础,将英-德翻译中的显化分为结构显化和非结构显化两类,分别体现在词汇和语法

两个层面。相反，德-英翻译中会出现浓缩（condensation）（即隐化）现象。Schmied & Schffler继而从认知和信息加工的角度对显化和浓缩现象的原因进行了解释。这项研究表明两个方向上均表现出显化，初步说明显化确是翻译过程的共有特征。但需要考虑的是，不同语对、不同文类中是否会表现出相同的情况呢？

Øveras (1998) 以英语-挪威语平行语料库 (ENPC) 为基础关注了词汇衔接现象，结果表明从翻译方向上看，总体上英语-挪威语方向上的显化要比挪威语-英语方向上更为突出。这一发现，与Schmied & Schffler (1997) 的发现不完全一致。Øveras的研究表明尽管两个方向上都表现出显化，但程度有差异，说明具体语对是一个重要变量。Øveras特别指出，语际显化的研究结果可以作为类比显化的必要补充，因为翻译文本相对源文本显性程度的提高并不一定能够说明翻译文本就比目标语中的同类原创文本显性程度高 (1998: 568)。这一研究初步将平行与类比两种模式联系起来。

柯飞 (2003) 以英汉/汉英平行语料库为基础，对汉语特殊句式“把”字句在英汉转换中的分布特点进行考察发现：翻译汉语作品中的“把”字句要比非翻译汉语中的“把”字句频次高，文学类作品比非文学类作品频次高，由于“把”字句适于表达复杂和细微的意思，因此“把”字句的频次说明了翻译文本中的显化现象。“把”字句属于汉语中独有句法结构，在翻译文本中表现出频次的增加。而其他语言中同样是特殊句式，却表现出相反的情况。如Tirkkonen-Condit (2000, 2002) 以翻译芬兰语语料库为基础提出了独特项假说：目标语语言中的某些语言项/元素在源语中缺少语言对应项，相对于原创文本而言，此类语言项/元素在翻译文本中的频次低。柯飞 (2003) 的研究结果与Tirkkonen-Condit形成了鲜明对照，这说明所讨论的语对也是显化研究一个重要的变量。

第二类研究属于纯形式的对

比。如王克非 (2003) 根据汉英/英汉平行语料库的4个子库 (汉译英文学、汉译英非文学、英译汉文学、英译汉非文学) 中句对应抽样检索发现，1:1的语句对应比例在汉英、英汉对译文本中都呈现出高频的特点，英译汉1:1的语句对应比例高于汉译英，而且在英汉互译中都呈现出译语文本扩增的特点，但不同类型的文本在扩增程度上有差别。这里的文本扩增可以作为显化的一个参数，但需要将文体类型这一因素考虑进去。此外，Frankenberg-Gacia (2004) 以葡-英双向平行语料库 (Compara) 为基础，以定量分析的方法证明了翻译文本中词汇数量的整体增加表明译文往往比原文更为明晰，而且这并不受制于两种语言之间的差异。

同样是研究语际显化，其他的切入点还包括：Klaudy & Károlly (2005) 对英译匈和匈译英文学文本两个方向上转述动词 (reporting verbs) 的考察；Englund-Dimitrova (2003) 对俄语-瑞典语翻译中转折连词 (adversative conjunctions) 的考察等，研究结果对原有假设都是部分支持，部分背离。

以上两种显化研究模式都主要局限于语言形式方面，对语义、语用层面的关注相对较少，主要原因在于，语料库数据的提取是以对语料的标注程度为前提，标注越细致，提取的信息就越具体。就现有技术而言，对语料的标注还仅限于词类和句法方面，因此主要的的数据提取方式也局限于这些方面。单语类比模式下，主要是对比翻译文本与非翻译文本表现出的规律性模式差异，双语平行模式则侧重于对具体语言转换的关注，即跨语言的变化。但类比模式与平行模式哪一个更有意义？切入点或检索项应如何选择？直到新世纪前10年的中期，研究者对这些问题才有了更深入、更全面的认识。

6.2 译者风格研究

除上述翻译共性研究外，“译者风格”是语料库翻译学的另一个研究课题。就这一课题，同样存在类比与平行两种对比模式：Baker (2000)

提出的译者风格属于类比基础上的研究，关注不同译者各自所有译文表现出的整体翻译风格差异，如在词汇多样性、句子复杂程度、叙事方式等方面的不同，不考虑源文本的影响；平行语料库基础上的译者风格研究则关注不同译者对同源文本中语言现象在各自译文中的规律性处理方式，如不同译者对源文本叙述风格的不同处理等。

Baker的研究方法仅仅关注译者的翻译文本，分析过程通常不考虑对应的源文本，关注译者下意识的语言行为。自Baker (2000) 提出用语料库来考察文学翻译译者风格的方法论以来，相关的研究有很多，有些基本上是按照Baker的方法论 (如Olohan, 2004; Saldanha, 2011a、2011b等)，有些虽然也是考察译者风格，但方法上并不遵照Baker的模式。

6.2.1 类比模式

Baker (2000) 译者风格方法论的第一个前提是目标文本导向，首先不考虑源文本，直接关注翻译文本，采用类比模式，即两位译者所有翻译文本之间的比较。第二个前提是研究切入点，即所考察的语言项不是源文本影响的结果，而应当完全是译者的下意识选择，如用词多样性、语言复杂程度等。根据这一模式，Olohan (2004: 153-156) 考察了Peter Bush和Dorothy S. Blair两位文学译者的风格，切入点放在缩略形式 (如t's=it is/it has) 的使用上，研究发现：Bush总体上使用缩略形式多，Blair则很少使用；与Olohan (2003) 对BNC和翻译文本库中缩略形式考察结果相对照，Bush使用缩略形式的频次接近于BNC中的情况，而Blair则与翻译文本中的频次相似。单从数据上看，两位译者在缩略形式使用方面存在明显差异，但这里要提出的问题是：第一，缩略形式能不能作为译者风格的一个标志？第二，译者是否选择缩略形式受不受源文本或源语言的影响呢？

根据Olohan (2003) 的发现，缩略形式使用频次可以用来区分翻译

文本和原创文本。但以频次区别两位译者的翻译风格说服力并不强。其次, Olohan分析译者使用缩略形式的原因时指出, 源文本的风格和文类对译者的选择有很大的影响(2004: 157)。也就是说, 我们很难将译者自己的风格与原作者或源文本风格在译文中的体现清晰地区分开。

同样根据Baker的方法, Saldanha(2011a)考察了Peter Bush和Margaret Jull Costa两位译者的翻译风格, 切入点放在翻译文本中对外来词使用的差异上, 并使用葡萄牙语-英语双向平行语料库(COMPARA)为参照库。研究表明, 两位译者在外来词的使用方面表现出明显差异, 研究者将这种差异作为译者风格的表现形式, 并指出这种译者风格的差异在很大程度上取决于译者对自身角色认识的差异。Saldanha(2011b)除外来语又增加了两个考察项目, 一个是对比两位译者翻译文本中斜体使用的差异, 另一个则是仿照Baker考察转述动词SAY和TELL后连接词的省略与否, 统计分析发现两位译者在这三个方面存在系统的差异。为了确保这些译者不同的选择不是源文本的影响, Saldanha又检索了对应的源文本, 结果表明源文本并不是影响译者系统选择的因素, 似乎说明这一方法是奏效的。

以上几个个案参照了Baker的方法, 但与Baker的不同之处在于, 切入点不同, Baker的切入点主要放在软件统计的文本整体特征上, 如标准类符-形符比, 平均句长等方面, 而Olohan和Saldanha的考察项主要放在具体的语言项上。但仅仅以纯粹形式上的差异来区分不同译者的风格显然还远远不够。至少Baker(2000)的方法论还存在不完善之处, 其最大的问题在于很难对译者风格进行清楚的厘定, 纯形式的统计缺乏一定的说服力。

6.2.2 平行模式

除以Baker方法论为基础的研究外, 另一些研究则将源文本纳入译者风格考察的范围(如Mikhailov & Villikka, 2001; Bosseaux, 2001,

2004, 2007; Winters, 2004a, 2004b, 2007, 2009等), 从一开始就不同于Baker目标语文本导向的大前提, 讨论的内容包括译者风格存在与否、不同译者对源文本中特定语言现象的规律性处理方式等, 研究切入点也从单纯的形式发展到话语和叙事方式等方面。如Mikhailov & Villikka(2001)以一个俄语-芬兰语小说平行语料库为数据库, 借鉴文体计量学中对作者身份归属确认的方法, 从词汇丰富程度、高频词、译者偏爱词汇几方面对同一作家和不同作家创作的俄文小说、同一译者用芬兰语翻译的不同作品, 以及不同译者用芬兰语翻译的同一作品进行了对比分析, 统计分析显示译者似乎没有明显的属于自己的翻译风格。这样看来, 如果译者风格的确存在, 那么就是考察译者风格的方法不够完善, 至少作者身份归属确认的方法并不完全适用于考察译者的翻译风格。这就从另一个角度说明, 仅依靠类符-形符比、平均句长等软件统计数据不足以区分不同译者的翻译风格。

另外一些研究则是对比不同译者对同一源文本的翻译, 这样源文本就成了一个控制变量, 从而以译者作为变量考察不同译者的规律性语言习惯。如Bosseaux(2001)从翻译转换的视角考察了Virginia Woolf的小说《海浪》(The Waves)的两个法语译本在翻译文体方面的差异, 主要关注了平均句长、类符-形符比、文化特有词项等方面, 研究表明两个法语译者的翻译策略存在明显差异, 以此说明两位译者风格的不同。但严格来讲, 尽管前两个考察项与Baker相同, 但以两个译者对同一源文本翻译做比较, 单从两个译本来说明译者风格, 显然不符合Baker对译者风格的界定。Bosseaux(2004, 2007)又分别考察了Virginia Woolf的小说《海浪》和《到灯塔去》(To the Lighthouse)的法语译本中各类视点体的语言表现形式的翻译, 探究译者的选择对叙事结构转换会产生什么样的影响。与Baker不同的是, Bosseaux从源文本到目

标文本的语言转换入手, 考察项已经从纯粹的形式发展到叙事方式方面, 具体做法是找出叙事风格的语言标志, 然后借助语料库比较不同译者使用这些语言标志的特征。

另一个类似的系列研究是Winters(2004a, 2004b, 2007, 2009)以Fitzgerald的小说《美女和被诅咒的人》(The Beautiful and Damned)的两个德文译本为语料, 分别考察情态小词(model particles)、外来词、语码转换和言语行为转述动词, 讨论了两位译者的风格差异。Winters与Bosseaux的研究方法相似, 但不同于Baker的单语类比模式, 使用的语料库为平行语料库, 考察项均与源文本相关联。这些研究最大的问题在于, 以同一个源文本的不同译本来讨论译者风格显然违背了Baker从一个译者所有翻译文本中发现译者风格的基本原则。这些讨论的已经不是Baker严格意义上的译者风格。

实际上, 无论是类比模式还是平行模式, 我们都很难将纯粹的译者风格与原作者或源文本风格在译文中的表现区分开, 因为二者的表现形式都是对语言的规律性使用。Olohan提出了一个设想, 就是在上述研究的基础上将译者创作文本与其翻译文本加以比较(2004: 150), 但问题是, 即便可以发现两类文本的规律性的异同, 我们还是很难判断哪些是译者翻译风格, 哪些是其创作风格。Saldanha(2011b)也尝试对译者风格和翻译文体做区分, 但在实际研究中, 译者的选择本身就受到多种因素的影响或限制, 要清楚地区分这两者十分困难。实际上, “译者风格”像上文讨论的“翻译共性”一样, 是语料库翻译学的另一个基本概念工具, 其功能主要在于帮助我们从事者的角度更好地认识翻译。无论前一种还是后一种模式, 这些研究从不同角度补充了Baker的方法论。既如此, 我们不妨将严格的区分放在一边, 将研究范围扩大一些, 既关注不同译者自己的风格, 也关注他们对源文本或者源语言中特殊现象在其翻译文本中的不同处理方式, 因为这两类文体/风格的表

现形式都是规律性的语言模式，都是译者选择的结果，无论是有意还无意识的选择，都体现出译者在翻译方式上的规律性差别。相比较而言，平行模式的研究对翻译研究的意义更大一些，因为纯粹译者自己的风格在翻译文本中的表现形式非常有限。

7 基于语料库的语言变化与多重复合对比模式的应用

近年来，语料库翻译研究的对比模式不再是单纯的平行或类比，或是简单的平行加类比，大多都是复合式对比模式。具有代表性的研究如Laviosa (2007)以“词汇启动”(lexical priming)理论为基础考察翻译所导致目标语的变化；Laviosa (2010)考察了意大利语经济类文本中英语外来语business的词汇启动模式；Hoey (2011)对比了一篇英语游记散文开篇句子原文与三种葡萄牙语译文中的词汇启动现象，并以葡萄牙语语料库为参照库对相关词汇进行了检索；Munday (2011)以英国国家语料库(BNC)和西班牙语语料库CREA (Spanish Real Academia Corpus)为参照库，考察了英语loom (large)与其西班牙语对应词cernerse的语义韵律差异；House (2011)讨论了如何运用语料库方法，来考察作为世界通用语(lingua franca)的英语怎样通过翻译和多语文本生成等语言接触方式来引发并推动语言变化(关于此部分的讨论可参见黄立波、王克非，2011: 917-918)。这些研究表现出的一个共同特点就是新研究对象的确立和新研究方法的应用，即基于语料库的语言变化与多重复合对比的研究模式。

以上这几个研究已经不同于原有的翻译共性或译者风格研究，但在细节的研究方法上还有类比或平行对比模式的影子，总体上让我们意识到，翻译语料库不仅仅可以研究翻译现象，还可以透过翻译来探索翻译活动所导致的语言变化。由此可见，一种新研究范式要得到认可并取得长足发展，必须不断加深对研究对象的认识，不断革新研究方法，拓展其研究领域，翻译研究也不例外，语料库翻译学近20年所取得的突出成绩就充分证明了这一点。□

注释

本文是国家社科基金项目“基于汉英平行语料库的翻译文体学研究”(编号: 09XYY002)的阶段性成果，同时得到香港理工大学科研项目(编号为G-YX4G)的资助，特此鸣谢!

参考文献

- [1] Baker, M. Corpus linguistics and translation studies: Implications and applications [A]. In M. Baker, G. Francis & E. Tognini-Bonelli (eds). *Text and Technology: In Honour of John Sinclair* [C]. Amsterdam: John Benjamins, 1993: 233-250.
- [2] Baker, M. Corpora in Translation Studies: An overview and some suggestions for future research [J]. *Target*, 1995 (7/2): 223-243.
- [3] Baker, M. Corpus-based Translation Studies: The challenges that lie ahead [A]. In H. Somers (ed). *Terminology, LSP and Translation* [C]. Amsterdam: John Benjamins, 1996: 175-186.
- [4] Baker, M. Towards a Methodology for Investigating the Style of a Literary Translator [J]. *Target*, 2000 (12/2): 241-266.
- [5] Blum-Kulka, S. & E. A. Levenston. Universals of lexical simplification [A]. In C. Faerch & G. Kasper (eds). *Strategies in Interlanguage Communication* [C]. London: Longman, 1983: 119-139.
- [6] Blum-Kulka, S. Shifts of cohesion and coherence in translation [A]. In J. House & S. Blum-Kulka (eds). *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies* [C]. Tübingen: Gunter Narr Verlag, 1986, 17-35.
- [7] Bosseaux, C. A Study of the Translator's Voice and Style in the French Translation of Virginia Woolf's *The Waves* [A]. In M. Olohan (ed). *CTIS Occasional Papers 1* [C]. Manchester, Centre for Translation and Intercultural Studies, UMIST, 2001: 55-75.
- [8] Bosseaux, C. Point of view in translation: A corpus-based study of French translations of Virginia Woolf's *To the Lighthouse* [J]. *Across Languages and Cultures*, 2004 (5/1): 107-123.
- [9] Bosseaux, C. *How Does it Feel? Point of View in Translation: The Case of Virginia Woolf into French* [M]. Amsterdam: Rodopi, 2007.
- [10] Englund-Dimitrova, B. Explicitation in Russian-Swedish translation: Sociolinguistic and pragmatic aspects [A]. In B. Englund-Dimitrova & A. Pereswetoff-Morath (eds). *Swedish Contributions to the Thirteenth International Congress of Slavists, Ljubljana, 15-21 August 2003* [C]. Lund: Lund University, 2003: 21-31.
- [11] Frankenberg-Gacia, A. Are translations longer

- than source texts? A corpus-based study of explicitation [A]. Paper presented at the Third International CULT (Corpus Use and Learning to Translate) Conference, Barcelona, 22-24 January 2004. <http://www.linguateca.pt/Repositorio/FrankenberGarcia2004.doc> (accessed 04/07/2006).
- [12] Hoey, M. Lexical priming and translation [A]. In A. Kruger, K. Wallmarch & J. Munday (eds). *Corpus-based Translation Studies: Research and Applications* [C]. London: Continuum, 2011: 153-168.
- [13] House, J. Using translation and parallel text corpora to investigate the influence of global English on textual norms in other languages [A]. In A. Kruger, K. Wallmarch & J. Munday (eds). *Corpus-based Translation Studies: Research and Applications* [C]. London: Continuum, 2011: 187-208.
- [14] Klaudy, K. & K. Károly. Implication in translation: Empirical evidence for operational asymmetry in translation [J]. *Across Languages and Cultures*, 2005 (6/1): 13-28.
- [15] Laviosa-Braithwait, S. *The English Comparable Corpora (ECC): A Resource and a Methodology for the Empirical Study of Translation* [M]. Unpublished PhD Thesis, Center for Translation and Intercultural Studies (CTIS), University of Manchester, 1996.
- [16] Laviosa, S. Corpus-based translation studies: Where does it come from? where is it going? [A]. In A. Kruger, K. Wallmarch & J. Munday (eds). *Corpus-based Translation Studies: Research and Applications* [C]. London: Continuum, 2011: 13-32.
- [17] Laviosa, S. The corpus-based approach: A new paradigm in translation studies [J]. In Sara Laviosa (ed). *The Corpus-Based Approach, Special Issue of Meta*, 1998 (43/4): 557-570.
- [18] Laviosa, S. *Corpus-based translation studies: Theory, Findings and Applications* [M]. Amsterdam: Rodopi, 2002.
- [19] Laviosa, S. Corpus-based translation studies: Where does it come from? where is it going [J]. *Language Matters*, 2004 (35/1): 6-27.
- [20] Laviosa, S. Similarity and differences in corpus-based translation studies [J]. *Journal of Foreign Languages*, 2007 (5): 56-63.
- [21] Laviosa, S. Corpus-based translation studies 15 years on: Theory, findings, applications [J]. *SYNAPS*, 2010 (24): 3-12.
- [22] Laviosa, S. Corpus-based translation studies: Where does it come from? where is it going? [A]. In A. Kruger, K. Wallmarch & J. Munday (eds). *Corpus-based Translation Studies: Research and Applications* [C]. London: Continuum, 2011: 13-32.
- [23] Mikhailov, M. & M. Villikka. Is there such a thing as a translator's style? [A]. In *Proceedings of Corpus Linguistics*. Birmingham, 2001: 378-385.
- [24] Munday, J. Looming Large: A cross-linguistic analysis of semantic prosodies in comparable reference corpora [A]. In A. Kruger, K. Wallmarch & J. Munday (eds). *Corpus-based Translation Studies: Research and Applications* [C]. London: Continuum, 2011: 169-186.
- [25] Nida, E. A. & C. Taber. *The Theory and Practice of Translation* [M]. Leiden: E. J. Brill, 1969.
- [26] Olohan, M. & M. Baker. Reporting *that* in translated English: Evidence for subconscious processes of explicitation [J]. *Across Languages and Cultures*, 2000 (1/2): 141-158.
- [27] Olohan, M. Spelling out the optionals in translation: a corpus study [J]. *UCREL Technical Papers*, 2001 (13): 423-432.
- [28] Olohan, M. Comparable corpora in translation research: Overview of recent analyses using the Translational English Corpus [A]. *LREC Language Resources in Translation Work and Research Workshop Proceedings*, 2002a: 5-9.
- [29] Olohan, M. Leave it out! Using a comparable corpus to investigate aspects of explicitation in translation [J]. *Cadernos de Tradução IX*, 2002b: 153-169.
- [30] Olohan, M. How frequent are the contractions? A study of contracted forms in the Translational English Corpus [J]. *Target*, 2003 (15/1): 59-89.
- [31] Olohan, M. *Introducing Corpora in Translation Studies* [M]. London: Routledge, 2004.
- [32] Øverås, L. In search of the third code: An investigation of norms in literary translation [J]. *Meta*, 1998 (43/4): 557-570.
- [33] Popper, K. R. *Objective Knowledge An Evolutionary Approach (revised edition)* [M]. Oxford: Oxford University Press, 1979.
- [34] Puurtinen, T. Syntax, readability and ideology in children's literature [J]. *Meta*, 1998 (43/4): 524-533.
- [35] Puurtinen, T. Genre-specific features of

translationese? Linguistic differences between translated and non-translated Finnish children's literature [J]. *Literary and Linguistic Computing*, 2003a (18/4): 389-406.

[36] Puurtinen, T. Nonfinite constructions in Finnish children's literature: Features of translationese contradicting translation universals? [A]. In S. Granger, J. Lerot & S. Petch-Tyson (eds). *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* [C]. Amsterdam: Rodopi, 2003b: 141-154.

[37] Puurtinen, T. Explicitation of clausal relations: A corpus-based analysis of clause connectives in translated and non-translated Finnish children literature [A]. In A. Mauranen & P. Kujamäki (eds). *Translation Universals: Do They Exist?* [C]. Amsterdam: John Benjamins, 2004: 165-176.

[38] Saldanha, G. Style of translation: The use of foreign words in translations by Margaret Jull Costa and Peter Bush [A]. In A. Kruger, K. Wallmarch, & J. Munday (eds). *Corpus-based Translation Studies: Research and Applications* [C]. London: Continuum, 2011a: 20-35.

[39] Saldanha, G. Translator style: Methodological considerations [J]. *Translator*, 2011b (17/1): 25-50.

[40] Schmied, J. & H. Schäffler. Explicitness as a universal feature of translation [A]. In M. Ljung (ed). *Corpus-based Studies in English. Papers from the 17th International Conference on English Language Research on Computerized Corpora (ICAME 17). Stockholm, May 15-19, 1996*. Amsterdam: Rodopi, 1997: 21-34.

[41] Tirkkonen-Condit, S. In search of translation universals: non-equivalence or "unique" items in a corpus test [A]. Paper presented at UMIST/UCL Conference: Research Models in Translation Studies, 28-30 April 2000, Manchester, UK, 2000: 207-220.

[42] Tirkkonen-Condit, S. Translationese — a myth or an empirical fact: A study into the linguistic identifiability of translated language [J]. *Target*, 2002 (14 /2): 207-220.

[43] Vinay, J. & J. Darbelnet. *Comparative Stylistics of French and English. A Methodology for Translation* [M]. Translated and edited by J. C. Sager & M. -J. Hamel. Amsterdam: John Benjamins, 1958/1995.

[44] Vanderauwera, R. *Dutch Novels Translated Into English: The Transformation of a "Minority"*

Literature [M]. Amsterdam: Rodopi, 1985.

[45] Winters, M. German translations of F. Scott Fitzgerald's *The Beautiful and Damned*: A corpus-based study of modal particles as features of translators' style [A]. In Ian Kemble (ed). *Using Corpora and Databases in Translation* [C]. Portsmouth: University of Portsmouth, 2004a: 71-78.

[46] Winters, M. F. Scott Fitzgerald's *Die Schönen und Verdammten*: A corpus-based study of loan words and code switches as features of translators' style [J]. *Language Matters, Studies in the Languages of Africa*, 2004b (35/1): 248-258.

[47] Winters, M. F. Scott Fitzgerald's *Die Schönen und Verdammten*: A corpus-based study of speech-act report verbs as a feature of translators' style [J]. *Meta*, 2007 (52/3): 412-425.

[48] Winters, M. Modal particles explained: How modal particles creep into translations and reveal translators' styles [J]. *Target*, 2009 (21/1): 74-97.

[49] 王克非. 英汉/汉英语句对应的语料库考察 [J]. *外语教学与研究*, 2003 (35/6): 410-416.

[50] 黄立波, 王克非. 语料库翻译学: 课题与进展 [J]. *外语教学与研究*, 2011 (43/6): 911-923.

[51] 柯飞. 汉语“把”字句特点、分布及英译研究 [J]. *外语与外语教学*, 2003 (12): 1-5.

Corpus-based Translation Studies: Object of Study and Research Method

Abstract: The nearly 20 years' development of Corpus-based Translation Studies (CTS) follows the general model of scientific development — “Problem 1 → Tentative Theory → Error Elimination → Problem 2” proposed by Karl Popper (1979). The overview of the development shows that apart from the remarkable success achieved by studies of translation universals and translator's style, language change brought about by translation has become a new focus of attention. The simple mono-lingual comparable or interlingual parallel research model or an integrated one of both has been replaced by the multiple-complex models in relation to research needs. The development of CTS illustrates a fact that to achieve recognition and significant development, a new paradigm has to deepen its understanding of the object of study and to innovate its research method so as to broaden its territory.

Key Words: Corpus-based Translation Studies; object of study; research method