

《英汉翻译中的汉语译文语料库研究》介评

李德超 香港理工大学

上个世纪90年代初, Mona Baker(1993)首倡借鉴当时已经发展得相当蓬勃的语料库语言学来开展翻译研究。自此以来,语料库翻译研究(亦称语料库翻译学)在国内外译界都取得了长足的进展。

据不完全统计,过去10年用英文出版的基于语料库翻译研究和语言对比的著作和论文集就有20多部(参见肖忠华,2012:4),基于语料库的中英对比及翻译的著作或论文集也有近10部。但总体而言,这些著作有些只是采用对应语料库来关注译文的特定语言特征,或只是关注某一作家或某一作品的译文特点,研究较为单一,较少从多角度、多层面来对汉语译文进行宏观及微观的考察。而浙江大学肖忠华教授的《英汉翻译中的汉语译文语料库研究》(以下简称《研究》)^①可说正是针对上述问题而写的力作,对国内今后基于语料库的翻译研究起到一定的示范及促进作用。

《研究》一书共有九章。第一章绪论回顾了过去半个世纪翻译研究经历的两次范式转变。第二章阐述了语料库翻译学的一些关键术语,指出此新学科在翻译研究中的前景。第三章探讨了语料库用于探讨译文——这种作为客观存在语言变体的总体语言特征(亦称翻译共性)——在国内外的研究现状。第四章在介绍当前一些大型通用汉语语料库的情况的基础上,着重介绍了作者本书研究基于的两大语料库,即兰卡斯特汉语语料库(LCMC)和浙大汉语译文语料库(ZCTC)的设计、研制和加工情况。

在第五章和第六章,作者讨论了基于ZCTC和LCMC这两个类比平衡语料库发现的一些汉语译文宏观和微观的语言特征。

第七章考察了现代汉语的一些典型句式和语法结构在两大语料库中的分布,从词汇接口探索译文的语法特征。作者发现,汉语译文中的短式被动句(即施事没有出现的被动句)的比例大大超过汉语母语,且两者在“被”字句的语用含义及其在各类语体分布上亦有较大差别。

第八章总结前面几章对汉语译文宏观统计特征以及在词汇和语法上有别于汉语母语的语用特征,以此来观照描写翻译学中的翻译共性研究。

第九章总结全书,从发现的汉语译文的词汇及语法特征中重新审视目前流行的各项翻译共性假设。作者的主要观点有:明晰化假设在翻译体汉语中得到了充分的支持,简化假设则有部分语言特征支持,

部分却是反证,这说明翻译语言可能是“简化和复杂化并存的产物”。(肖忠华,2012:158)

《研究》有以下几大特点。一是翻译语料库的设计科学、建设过程严谨,使研究结果的横向可比性大大加强。目前,国内不少汉语翻译语料库的选材比较单一(往往是以文学类语篇为主),且入库的语料往往都是通篇录入,没有任何的抽样。这样做的好处就是对某一文类译文的考察比较深入。由于语料都是整篇呈现,研究者亦易于结合更大的语境来对研究结果作出阐释。但由于这些翻译语料库设计时没有参照国际上一些早已成型且有世界性影响的语料库(如代表美国英语的布朗语料库以及相对应的代表英国英语的FLOB语料库)的建库标准,这就可能令所选的语料代表性不足,造成研究结论的偏颇。同时,由于库与库之间各自为政,这让基于各个库所得出的结论之间无法相互比对和认证,亦无法在世界范围内与其它国家的翻译语料库的研究结果进行横向比较。相较之下,《研究》所基于的汉语语料库和译文语料库均是参照FLOB的抽样方法和标准建库,两库的设计、采样和语料的加工都相当一致,这就确保了两者之间的可比性和对比结果的可靠性。

二是对汉语译文语言特征的探讨全面,针对性强。对于一些常见的考察项目,如译文的词汇密度、文本信息密度等,《研究》均作了详尽及令人信服的探讨。而对于一些较少涉及的课题,如正负关键词、习语和话语重述标记这些词汇特征在汉语译文中的体现,《研究》也作了颇有新意的论述。例如,一般认为习语在译文中出现频率更高,因为译者为求稳,趋向于在译文中采用“安全”而非创新的语言表达式(Baker,2007),但作者却发现习语在汉语译文及其在母语中的分布非常相似,对规范化的假设提出反证。目前基于语料库翻译研究检验的许多研究课题和假设均脱胎于翻译英语或相近欧洲语言对,当中有不少在汉语语境下并不适用(如探讨削尾词),而汉语的某些独特的特征亦未有在以往的基于语料库翻译研究中得到应有的重视^②。因此《研究》一书最令人称道的是能结合汉语的特点,在汉语译文中调查一些汉语所特有的语法特征,如“被”字句、“把”字句、“有”字句、“是”字句、量词结构、体标记“了、着、过”、结构助词“的、地、得、之”,以及语气词等。这些具有译入语特色的语言特征,正如作者所言,“越能用来区别译入语的原创文本和翻译文本”。(肖

忠华, 2012: 159)

三是在研究方法上采取复合研究范式, 从多角度对译文特征作出考察及阐释。国际上基于语料库的翻译研究大多采用单一的方法, 要么是利用类比语料库来“比较对应译入语母语和译文语料”, 要么是利用双语平行语料库“研究原文与译文之间的对应”(同上: 161)。前者是 Mona Baker 提倡的把翻译语言作为一种客观存在的语言变体, 来研究其“出现的非源语具有的典型语言特征”的必然结果 (1993: 243); 后者则是将语料库运用于比较源语对译文的渗透作用的必经之道。长久以来, 这两种研究方法通常都是各行其道, 较少结合于一项研究之中^③。针对这一不足, 《研究》采用了 McEnery & Xiao (2002) 首先提出的复合研究范式, 在考察时综合运用对应语料库和平行语料库, 不仅有从宏观上由对应语料库对翻译汉语的语言特征研究的统计分析, 亦有从微观上对统计发现的特点交由平行语料库中的数据去检验和分析原因。这种多角度、多层次利用不同语料库来充分解释翻译现象的复合研究范式定会对将来的语料库翻译学起到示范作用。

笔者亦想以《研究》一书为契机, 谈谈语料库翻译研究目前存在的问题, 并探索其未来的发展方向。

第一, 如上所述, 采取复合研究范式是《研究》的一大亮点。长期以来, 对翻译体语言的研究, 尤其是翻译汉语特征的研究, 都是长于对表象数据进行分析 and 归纳, 短于对所发现的具有高度可预测性的语言特征进行分析和解释。但正如 Toury (1995) 指出, 在描写翻译学研究中, 描写只是其中的一步骤而已, 解释和预测翻译现象具有同样的重要性。采用复合研究范式由于可以利用平行语料库对类比语料库中发现的语言规律作出解释, 因而有助于我们了解翻译语言特征形成的深层原因。但正如要取得具有可靠性的对翻译语言的描写结果, 首先要有建库科学、抽样均衡的类比语料库一般, 要从平行语料库中分析得到令人信服的解释结果, 充分了解源语渗透对特定翻译语言特征起到的作用, 那就同样要有与类比语料库建库思想一致、取样标准和样本收集相同的类比语料库。但遗憾的是, 《研究》参照的两个平行语料库(兰卡斯特大学 Babel 英汉平行语料库和北外通用汉英平行语料库)并非专门参照类比语料库而建, 双方的可比性不强。这就令人不得不质疑将同质的类比语料库得出的结果置于非同质的两个平行语料库中讨论与分析是否合适(因为这两个语料库对文类的选择、抽样的方法与标准与 ZCTC 和 LCMC 大不相同), 这变相也令作者所采用的复合研究范式的协同效果打了折扣。

第二, 早在六年前, 已有学者指出不少对翻译共

性的研究, 均是“以对共性的验证和描写为主, 解释较少”(吴昂等, 2006: 299)。六年后的今天, 仍然还有学者发现, 对翻译共性的“解释性研究更是尚未起步”(胡显耀等, 2011: 60)。对翻译共性重描写而轻解释的倾向在《研究》一书中得到了一定的纠正。但在解释过程中亦应该防止把现象或问题简单化的做法。例如, 在讨论 Babel 英汉平行语料库中“被”字句的分布情况时, 作者发现 526 例“被”字句译文中 446 例汉语被动句式是从英语不同形式的被动句翻译过来, 并以此作为例证来证明翻译汉语中的“被”字句是受英语源语的影响, 也即是译文中体现了源语的渗透作用(肖忠华, 2012: 119-120)。作为例证, 作者引用了以下的译文:

(1a) If he turns out to be a presentable, coherent but otherwise ordinary young man—reasonably law-abiding, amused by his good fortune and never taking himself too seriously—he will be from time to time **spotted by photographers, snapped with girlfriends, mentioned in gossip columns, invited onto talk shows** and we will know him so well we won't care that, strictly speaking, somebody else was born first.

(1b) 如果他长成了一个体面的、思路清晰的、而在其他方面又同于一般的年轻人——相当地遵纪守法, 为他的好运而高兴, 从来 not 把自己太当回事——他会时常被摄影师捕捉, 与女朋友在一起的时候被拍快照, 在闲聊栏目中被提及, 被邀请参加电视访谈节目, 而我们将如此地了解他, 严格说来, 我们不会在乎第一个出生的其实另有其人。(同上: 120)

作者认为, 中文例句中的“被”字的运用都是受到英语由 be 引导的被动式的影响, 并觉得“这些例句中的‘被’字大多数可以省略”(同上)。但如果认真细读原文, 就会发现其被动用法另有深意。在原文中, 原作者通过大量地使用被动式的排比句 (be) spotted, snapped, mentioned, invited 表现所描述的一切均非出于人物自己的意愿。这样, 一位成长之后就只得被人摆布, 在现实面前一切都无能为力的人物形象就呼之欲出。从这个小小的片断中亦可以折射出现代人生活的无奈。在英文文学作品中, 采用被动句方式来折射出人类在自然力量面前毫无反抗能力的例子是相当多的:

(2) I'm finding out that a lot of what I thought had been bonfired, Oxfam-ed, used for land-fill, has in fact been tidied away in sound archives, stills libraries, image banks, memorabilia mausoleums, tat troves, mug morgues.

It seems that as long as you're in print or on film or a name on a buff envelope in an archive somewhere, you're never truly dead now. You can be electronically colored, emulsified, embellished, enhanced, coaxed

towards some state of virtual reality. You can be reactivated or reembodyed; simulated and hologrammed. In just the last two years my voice has been artificially reprocessed for stereo effect and reincarnated in half-speed remasterings and on digital compact-disk.

(Alma Cogan, by Gordon Burn)

可以看到,在以上由英国作家 Gordon Burn 所写的小说 *Alma Cogan* 中的片断中,大量地采用了划线部分的被动式。它们用来表示在现代科技面前,人类只能任其摆布的文体效果,读之令人惆怅和无奈之感油然而生。

很明显,(1b)译文的译者领悟到了原文被动用法的文体特征,因而有意识地在译文中复制这种被动用法。换言之,译者在这段文字中对被动句的选择其实是有其在复制原文风格上的考量和深意,不能简单地把“被”字视为是受到源文影响的翻译腔。或许我们可以对平行语料库中发现的规律性译文特征,区分强制性保留 (obligatory retention) 和选择性保留 (optional retention),以准确考察源语对译文的渗透作用。前者是指译文由于源语和目标语在语言结构上的本质差别,或是出于文体上的考虑而不得不要保留原文语言特征的情况(比如前面所引的《研究》中的例子);后者则是译者个人主观和具有倾向性决定的结果,且不保留这些原文语言特征对翻译质量并无明显影响。在此后者才是研究源文渗透作用的对象。

第三, Kenny(2006:46)曾对语料库给翻译研究带来的贡献作了两点总结:一是语料库可以用于分析“表现形式高度可测的语言特征”;二是语料库可用于研究单凭个人无法处理的大量的常见语言特征。但若从另一个角度看,语料库翻译研究以上的长处却暗含一个预设,即只有反复、大量出现的语言特征才有研究的价值和意义。但 Kenny 亦指出,有不少语言现象“较不易预测,或是难以用语料库工具来发现”(ibid)。这些较少出现的语言现象不能因为不具备统计学上的显著意义而被认为在语言研究中无意义或不重要。恰恰相反,正如《研究》一书在对汉语译文中呈现不足的独特项目研究所示,往往是越具有译入语特色的语言特征,就越少在译文中出现。换言之,在译文中不常见、甚至不出现的语言特征亦可能是构成译文“第三语码”语言特色的重要原因。相对于后者而言,由于前者较显著、较容易统计而得,长期以来占了国内基于语料库翻译研究的大部分版图,而后者往往由于在语料库中呈现不足而不受重视。希望《研究》一书的出版能扩大将来国内的语料库翻译研究内容的光谱,为翻译共性的研究提供新的角度。

注 释

① 《研究》属王克非与胡开宝主编的“语料库翻译学文库”

系列中的一册。

- ② 目前可见的文献中只有柯飞(2003,2005)在此方面作过比较深入的探索。
- ③ Chesterman(2004)对此还采取了一种折中的方法,把前者称之为是对目标型共性(T-universals)的研究,后者称之为是对源语型共性(S-universals)的考察,试图把它们均统一于“翻译共性”这个概括性的术语之下。但共同采用“共性”这个标签反而模糊了这两种不同语料库研究方法本质的区别。

参 考 文 献

- [1] 董娜.基于语料库的“译者痕迹”研究:林语堂翻译文本解读[M].北京:中国社会科学出版社,2010.
- [2] 胡开宝.语料库翻译学概论[M].上海:上海交通大学出版社,2011.
- [3] 胡显耀.现代汉语语料库翻译研究[M].北京:外文出版社,2008.
- [4] 胡显耀、曾佳.基于语料库的翻译共性研究新趋势[J].解放军外国语学院学报,2011(1):56-62.
- [5] 黄立波.基于汉英/英汉平行语料库的翻译共性[M].上海:复旦大学出版社,2007.
- [6] 柯飞.汉语“把”字句特点、分布及英译[J].外语与外语教学,2003(12):1-5.
- [7] 柯飞.翻译中的隐和显[J].外语教学与研究,2005(4):303-307.
- [8] 刘泽权.红楼梦中英文语料库的创建及应用研究[M].北京:光明日报出版社,2010.
- [9] 王克非.双语对应语料库研制与应用[M].北京:外语教学与研究出版社,2004.
- [10] 王克非.语料库翻译学探索[M].上海:上海交通大学出版社,2011.
- [11] 吴昂、黄立波.关于翻译共性的研究[J].外语教学与研究,2006(5):296-302.
- [12] 肖忠华.英汉翻译中汉语译文语料库研究[M].上海:上海交通大学出版社,2012.
- [13] Baker, M. Corpus linguistics and Translation Studies: Implications and applications [A]. M. Baker, G. Francis and E. Tognini-Bonelli (eds.) *Text and Technology. In Honor of John Sinclair* [C]. Amsterdam: John Benjamins, 1993: 233-250.
- [14] Baker, M. Patterns of idiomaticity in translated vs. non-translated text [J]. *Belgian Journal of Linguistics*, 2007 (21): 11-21.
- [15] Chesterman, A. Beyond the particular [A]. A. Mauranen and P. Kuyamaki (eds.) *Translation Universals: Do They Exist?* [C]. Amsterdam: John Benjamins, 2004. 33-49.
- [16] Frawley, W. Prolegomenon to a theory of translation [A]. W. Frawley (ed.) *Translation: Literary, Linguistic and Philosophical Perspectives* [C]. London: Associated University Press, 1984: 159-175.
- [17] Kenny, D. Corpus-based Translation Studies: A quantitative or qualitative development? [J]. *Journal of Translation Studies*, 2006 (1): 43-58.
- [18] Laviosa, S. Corpus-based Translation Studies: Where does it come from? Where is it going? [J]. *TradTerm*, 2004 (10): 29-57.
- [19] McEnery, A. & Xiao, R. Domains, text types, aspect marking and English-Chinese translation [J]. *Languages in Contrast*, 2002 (2): 211-231.
- [20] Toury, G. *Descriptive Translation Studies and Beyond* [M]. Amsterdam: John Benjamins, 1995.

[作者简介] 李德超,博士,香港理工大学中文及双语学系副教授,研究方向:翻译史、实证翻译研究、口译研究。
[作者电子信箱] ctdechao@polyu.edu.hk