

基于翻译教学的评价语料库与译文质量 评估的实证研究

王红玫

(温州大学外国语学院, 浙江温州 325035)

摘要: 在翻译教学中, 学生译文的质量是教学效果的直接体现。学生译文质量的评价机制一直都是翻译教学研究及改革的重点及难题之一。在介绍基于双语平行评价语料库的学生译文评价的方法及其验证效果的基础上, 试图说明评价语料库能够有效的提高评价的客观性, 以便给学生提供建设性的反馈意见, 帮助学生提高译文质量及翻译能力。评价语料库对翻译教学内容及其方法的改善也能够为学生提供有益的参考和帮助。

关键词: 评价语料库; 译文质量评估; 评价机制; 客观性

中图分类号: H059 **文献标志码:** A **文章编号:** 1674-3555(2020)04-0093-09

DOI: 10.3875/j.issn.1674-3555.2020.04.012 本文的 PDF 文件可以从 <http://xuebao.wzu.edu.cn/> 获得

在翻译教学及翻译学科的研究中, 译文质量的评估始终是一个复杂棘手的问题。译文质量评估的难点在于如何将评价标准的主观性降至最低。朱利安·豪斯的译文质量评估模式, 试图使用一个参数体系来评定译文的质量。该模式的确大大提高了评价的客观性, 但朱利安·豪斯自己也曾指出该模式“没有普遍适用性”, “对大多数可用于该模式的文学翻译的评价, 很大程度上是主观的”^[1], 然而在翻译教学中, 对学生译文的评价, 需要设定一些老师和学生都能够接受的相对客观的评价标准, 以便老师能够依据这些标准为学生的译文打分, 给学生提供建设性的反馈, 以帮助学生进行针对性的翻译训练, 提高学生的翻译技巧和能力。

目前, 国内基于语料库的翻译教学研究已取得一定的成果, “但该研究领域起步较晚, 在翻译专业教学和语料库翻译学迅速发展的映衬之下, 语料库应用于翻译教学的研究略显不足”^{[2][173]}。国内基于语料库的翻译教学研究, “主要涉及语料库建设与翻译教学实践、基于语料库构建翻译教学理论框架以及基于语料库的翻译教材建设等”^[3]。如双语对应语料库的建设、单语语料库与翻译教学的结合及自建学习者翻译语料库等方面的研究。这些研究多见于理论层面, 如王克非^[4]、李德超^[5]等学者的研究。也有学者尝试过实证性的研究, 如刘泽权等^[6]对自建学习者译本语料库的应用进行分析, 并尝试建立一套基于语料库的翻译教学模式与译作评析体系。再如王佳健以实证的方法“揭示大学英语教学中, 翻译测试人工评分员关注的反映译文质量的文本特征, 为英汉翻译测试译文的自动量化评价研究提供一些有益的参考”^[7], 然而“国内至今从某一视角出发利

收稿日期: 2019-04-17

基金项目: 浙江省高等教育教学改革研究项目(kg20160380)

作者简介: 王红玫(1974-), 女, 河北沧州人, 讲师, 博士, 研究方向: 翻译理论与实践

用语料库进行翻译教学实证研究的探讨并不多见”^[21]⁷⁵。

本文基于实验测试数据,探讨评价语料库在评定学生译文上的可行性、客观性、高效性以及改善翻译教学方法的功用。本研究所涉及的相关内容评价语料库在学生译文质量评估方面的优势、评价语料库的构成、基于评价语料库的学生译文质量评估的实证测评数据以及对测评结果所反映出的问题的分析讨论。

一、评价语料库作为学生译文质量评价参考标准的优势

在翻译教学中,教师大都会尽力多给学生提供不同题材、不同主题、不同文体的翻译材料,对学生进行翻译训练,以便将来学生能够处理不同领域的翻译文本。这就要求教师和学生不仅要掌握原语和译入语的语言规则,更要很好的掌握和理解材料涉及的学科和专业知识,同时也要求作为评判官的老师,不但要掌握各领域所涵盖的语言知识(如词汇、短语、语法等),而且要熟知这些领域文本的特殊文体和特殊的表达方式等,以便公正客观的评价学生的译文。然而在大多数情况下,要求作为译者和评判员的老师精通各种特殊用途的语言和文体是不现实的,也是不合理的。因此,在翻译教学中,对学生译文质量进行客观而有效的评估,并进行建设性的反馈,成为阻碍翻译教学的一大障碍。汉斯·洪宁(Hönig H G)指出在对学生进行翻译训练时,“译文质量评估不应该是目的,而应该是手段”,“学生译后须得到建设性的反馈”^[8]³²,以便使“学生即使离开教室很久之后,仍可以继续独立进行学习”^[9]。这就意味着教师必须能够指出译文的具体问题,告诉学生错在哪里,为什么会错,而不是依靠模糊的印象,凭直觉或感觉得出判断及评价。翻译教师由于自身语言和专业背景知识的局限性,对学生译文进行评价时,需要经常借助外部资源。传统的查询资源包括字典(包括纸质词典和电子词典)、相似译入语文本,自身的语感直觉等。下面就这三方面与语料库评价机制进行比较,以观其优缺点。

(一) 语料库与传统的字典

传统的字典用作评价资源,其主要功能是查询词汇的用法,如在译中文词语的意义、词语的搭配、词汇的语法功能等是否准确。字典的主要缺陷是存储空间有限、词条收录不完全,且不能提供词语使用的具体语境。虽然有些词典会给每个词条提供一个例句,但大部分情况下,这些例句并不是真实语境中的实际用语,而是考虑到词典有限的存储空间而人为制造的没有上下文的孤句、短句。与传统字典相比,语料库在词汇查询上具有一定的优势。由于没有存储空间的限制,语料库在最大数量限度上收录了一个词实际存在的真实文本,为词汇作为概念和术语提供了最大覆盖范畴。语料库不仅可以查询词语的意义、词语的搭配、词汇的语法功能等内容,还可以提供词语的使用频率,词汇所在文本的文体,具体的上下文、文本的来源出处,出版或使用的时间等具体信息,并可以与其它词汇进行平行比对。如使用语料库分析工具 Concorde,可以在语料库中瞬间检索到医学术语“injection(注射)”一词的使用情况,包括该词在上下文语境中的词意、在具体语境中的功能、在所处语境的文本类型、使用频率以及搭配使用情况等。这可以帮助评判教师提高词汇错误识别的准确性。与传统字典相比,语料库收录的词汇时效性更强。对于新闻、科技等题材的内容而言,会不断产生大量新词,传统字典由于编纂印刷的限制,没有办法及时更新,而以电子信息储存方式存在的语料库则可以实时更新,实现即时查询。语料库用作词汇评判标准优点之一是检索方法简单、查询速度快、查询结果准确。相比传统的字典而言,语料库查询可以使用模糊匹配、排序,通配符等方法,快速准确地定位所要查询的内容。

(二) 语料库与平行文本

平行文本指的是由译入语国家相同领域的专家撰写的,与原文内容题材相似的文本。平行文本还应该与原文具有相似的文本类型、文本功能和使用时间段。大部分情况下,教师都会借助平行文本作为评判学生译文的参考文本,但是教师选择传统平行文本作为参考资源有两点不足,其一,教师收集和查阅的相关专业的文本数量有限,不能确保原文中所有相关概念、术语和语言类型都能够出现。其二,教师需花费大量时间和精力在平行文本中人工检索需要比对的词汇和语言类型。与传统人工检索的平行文本相比,评价语料库收录的文本不受数量限制,可以全面覆盖需检索的词汇和语类,同时,评价语料库采用电子数据检索方式,以实现快速精准查询。此外,教师还可以利用语料库分析工具,如 WrodSmith、Concorde、Wordfast 等,快速实现词频语类统计、多个平行文本比对等功能。这不仅能大大提高评判的效率,还可提高评判的客观性。毋庸置疑,使用语料库会大大节约教师查询的时间,并帮助教师提高评判译文词意时的准确性、代表性和客观性。

(三) 语料库与评判教师的语感直觉

虽然靠母语的直觉来判断哪些词汇或表达是恰当的翻译,有时候会得到“神来之笔”的译文,但由于个体对语言掌握的程度差异,所以无法保证评判教师直觉判断的准确性频率。因此,他们通常会求助于专业领域的专家文本或专家本人,来确定一个特定的语句或段落为目的语中应该如何表达才算是正确恰当的译文。但通常情况下,评判教师很难有机会当面请教所有文本涉及各领域专家,而要在各种学科领域专家撰写的大量文本中查询一个特定的段落或语句又极其耗时费力。而以电子数据形式存在的平行语料库不仅可以快速准确地查询出特定语句或段落的多个相似文本,而且还可以有效避免直觉的随意性和参考文本的孤证性。

二、评价语料库的构成

评价语料库指的是用来作为评价参考标准的语料库,它应该包括以下所述子库。

(一) 双语平行术语库

它应由某一专业领域一些常用的术语、概念词汇、缩写词及其目的语中对应的词汇构成,用来评判学生译文中的专业术语、概念翻译是否标准、规范和准确。

(二) 双语专业语料库

专业语料库里的平行文本由某一个专业领域内的专家编写和翻译的真实文本组成。它能够涵盖所要检测的文本主题,提供必要的概念信息,同时它应与原文本具有相同的文本类型,以便评判译文的文体和词汇用法是否准确。双语专业语料库可以是(自建)小型语料库,语料库中的文本能涵盖所检测主题,为评判教师提供该领域的基本概念和一些关键术语。专业语料库的主要功能为评判教师提供了一个基准,教师可以根据它来判断学生是否正确理解原文本中的概念和主要内容。

(三) 一般双语语料库

它需由学科范畴更广泛的一般性文本组成,语料应达到一定的规模,可以体现双语语言的使用规范,文本的使用时间或出版日期应明显比原文本的时间跨度大。它主要用来帮助教师评判译文的文体和表达方式是否准确、语篇是否通畅连贯、译文是否符合译入语规范。

需要指出的一般语料库并不仅仅是文本的随机组合,而需确保语料库的内容契合预期用户的特定需要。本文的目标用户是作为评判者的教师,他们的需求是评估一个陌生领域里的一篇专业文本的汉语译文。在以下测试中,被测试人员使用了评价语料库作为评判和反馈的参考资源。

三、基于评价语料库的学生译文评估测试

(一) 实验设计

1. 研究的问题

该实验旨在测试评价语料库作为评估学生译文的标准,参考资源的可行性、客观性、有效性,对教师和学生的帮助度以及对改善翻译教学方法的功用。

2. 评价语料库的设计

本次测试使用的双语专业术语库为双语平行医学术语库,该术语库为本地双语平行术语库,内含医学术语词条 74 893 条,涵盖基础医学和临床医学,包括解剖、病理、生化、内科、外科、妇儿科、诊断学、神经精神病学等学科的术语。使用的双语专业语料库为医学英汉双语平行语料库,该语料库为在线语料库^[10],其中的英文语料均为英美医学专家撰写,汉语翻译则由医学专家和英语专家共同完成,语句符合业内表达规范,内容涵盖基础医学和临床医学两个专业领域,涉及解剖、病理、生化、内科、外科、妇儿科、诊断学、神经精神病学等多个学科的内容,英汉语总词次共约 300 万;使用的一般双语语料库为 NiuTrans 中英平行语料库和英汉双语平行语料库,前者为本地语料库,含中英文对照语句共 10 000 句,后者为在线语料库^[11],由厦门大学开发,共收集英汉双语语料 215 713 句,其中英文约 3 290 670 词次,中文约 5 370 429 字次。

3. 实证研究的设计

参加本次测试的共有 23 人。其中学生 18 人,包括 8 名翻译人员(均为来自翻译专业大三的学生),10 名学生评判人员(均为来自英语专业大三的学生),他们均没有任何医学领域的专业背景。测试专家有 5 人,包括 1 名医学领域的双语专家,1 名没有汉语背景的英语语言专家,1 名双语背景非医学专业的香港留学生,1 名从业多年的中国职业译员,1 名从事翻译教学多年的中国高校教师。本测试要求 8 名翻译专业的学生独立翻译一篇与医学相关的新闻报道,该报道原文 993 个汉字,题为《2016“微笑联盟”温州站昨落幕》,学生被要求以职业译员的角色为客户将该中文报道翻译成英文。参加测试的 10 名学生评判员被分为两组,其中 5 名在控制组中,5 名在对照组。测试要求对照组成员仅使用常规资源评判译文并提供反馈意见,同时控制组成员仅使用评价语料库评判译文并提供反馈意见。评判结束后,负责翻译的学生仅根据反馈意见对译文进行修改。然后,由 5 名专家以译文用户的身份评判学生修订后的译文,并根据自己的满意度对译文进行排序。

(二) 数据分析

虽然该实验规模相对较小,但却产生了大量的数据,现就一些最突出的问题给予讨论。

1. 错误的识别和纠正

测试的实验数据包括学生评判员(以下简称评判员)在译文中标注的错误数量以及学生译员(以下简称译员)根据评判员的反馈意见改正的错误数量,如表 1 所示。相同词汇错误识别的统一性,即评判员是否能在英文译文中识别出某一汉语医学术语并判别其翻译的正确性,如表 2 所示。为了保证测试的客观性,选取了由翻译软件 Trados(塔多思)萃取出的术语,表 2 栏目中所

列举的 8 个医学术语只是源文诸多术语中的一部分, 为随机筛选的结果, 非特别圈定的测试内容, 也并未提前告知被测试人员, 这样做的目的是测试评判员和译员是否能真正识别术语并给出正确译文。

表 1 错误识别量及纠错量统计表

组名	评判员	学生	学生	学生	学生	学生	学生	学生	学生	合计	纠错及改正比 (%)
		译文	译文	译文	译文	译文	译文	译文	译文		
		1	2	3	4	5	6	7	8		
对照组	学生评判员 1	12/8	10/6	11/8	12/6	6/5	12/5	13/5	11/7	87/50	57.5
	学生评判员 2	7/3	13/2	9/5	9/3	9/6	11/3	11/4	10/6	79/32	40.5
	学生评判员 3	1/0	6/1	5/1	12/4	4/2	7/3	9/3	8/5	52/19	36.5
	学生评判员 4	6/2	7/5	10/5	11/6	8/5	9/6	7/4	6/2	64/35	54.6
	学生评判员 5	4/3	10/5	5/3	9/3	5/3	6/5	8/6	8/5	55/33	60.0
	合计										337/169
控制组	学生评判员 6	17/15	15/14	18/15	21/20	17/16	19/17	16/15	15/14	138/126	91.3
	学生评判员 7	14/12	18/16	12/12	14/13	13/12	11/10	12/11	13/12	107/98	91.6
	学生评判员 8	28/23	23/22	20/20	25/23	19/19	24/23	20/19	21/19	180/168	93.3
	学生评判员 9	19/18	21/21	24/22	20/19	17/16	15/14	17/14	21/20	154/144	93.5
	学生评判员 10	16/14	16/15	15/15	17/14	19/15	16/15	14/13	13/13	126/114	90.4
	合计										705/650

注: 学生译文栏目中的第 1 个数字表示评判员识别出的错误数量, 第 2 个数字表示译员根据反馈意见改正的错误数量, 如学生译文 1 的数字为 12/8, 这表示评判员的识别错误数量 12, 译员的改正数量为 8。

表 2 相同词汇翻译错误识别统计表

组别	评判员	患者	附属医院	唇腭裂	口腔	主治医生	麻药	缝合	筛查
对照组	学生评判员 1	√	√	√	√	×	√	×	√
	学生评判员 2	√	√	√	×	√	√	√	×
	学生评判员 3	√	√	√	√	√	√	√	√
	学生评判员 4	√	×	√	√	√	√	×	×
	学生评判员 5	√	√	√	√	×	√	√	×
控制组	学生评判员 6	√	√	√	√	×	√	√	×
	学生评判员 7	√	√	√	√	×	√	√	×
	学生评判员 8	√	√	√	√	×	√	√	×
	学生评判员 9	√	√	√	√	×	√	√	×
	学生评判员 10	√	√	√	√	×	√	√	×

注: “√”代表评判员判定该词翻译正确, “×”代表评判员判定该词翻译有误。

表 1 和表 2 的统计数据, 反映出如下问题:

其一, 控制组评判员识别的错误量远高于对照组。就错误总量来看, 控制组识别的错误量是对照组的一倍多 (705 : 337)。评判员 7 是控制组识别错误最少的人, 评判员 3 是对照组识别错误数量最少的人, 而评判员 7 识别的错误量是评判员 3 的两倍多 (107 : 52)。需要说明的是, 虽然任务要求每个学生需在三小时内完成任务, 但对照组的成员完成任务总共只用了 5 小时, 而

控制组的学生总共用了8小时。表1的统计数据显示,对照组在5小时内总共发现了337个错误,大约每小时可发现67个错误。假设对照组学生能够继续以这种速度发现更多错误,那么在8小时内,他们总共可以发现539个错误,这仍然只是控制组发现错误总量的76%。此可以说明比起传统资源,评判语料库作为评判标准具有更好的可行性和有效性。

其二,从整体上看,对照组纠错率约有50%,而控制组纠错率约为92%,控制组评判员的纠错百分比比对照组的高出42%。其原因在于,参与这个实验的10个学生评判员都没有医学专业背景,而他们所掌握的传统资源不能为他们提供足够的信息,这使他们难以快速了解掌握医学领域的专业知识,因此对照组学生评判员常常会忽略(或忽略未纠正的)学生译文中可能出现的错误,而控制组成员使用评价语料库可以快速并准确的找到译文所涉及的专业术语和段落语句,因此可以更准确地发现和纠正更多译文中的错误。

其三,控制组学生完成任务的总时长虽然比对照组多了3个小时,但比较单位时间内两组识别和纠正错误的数量发现,对照组每小时的平均速度为67.4/33.8个,控制组每小时平均速度为88.125/81.25个,可见,控制组的纠错速度是对照组的两倍多。这是因为语料库可以实现快速精准检索,可以提高查询和比较的速度,从而进一步提高控制组识别和纠正错误的效率。从以上数据分析可以看出,利用评价语料库对译文质量所作的评估,在错误识别率、纠错率以及评估效率上都优于使用传统资源所作的评估。

其四,在相同错误识别率上,控制组远高于对照组。在随机抽取的8个词汇的译文错误识别统计中,如表2所示,对照组成员对相同词汇的翻译错误缺乏识别的一致性。对照组的评判员识别出有些栏目的一些错误,其他评判员并不能识别出来,如表中的“筛查”一栏,这也是评判主观性的表现,即一个评判员认为不正确的,另一个评判员则认为是可以接受的。相比之下,控制组评判员在词汇错误的识别上更加一致,这说明控制组的判断标准更加统一客观。

2. 评判反馈

对评判员反馈意见的数据统计分析结果如下。

其一,反馈意见的主客观性。从反馈意见的话语分析来看,在对照组评判员的评论和反馈意见中,使用代表个人主观意见的词更多,如“我认为、我觉得、感觉、似乎、印象中”等,因此他们的反馈意见显得比较主观。在控制组的学生评判员的评论和反馈意见中,只少量出现了“我”,更多出现的是“根据语料库中的例句”“查询术语库得知”“从未用于语料库”“搜索词分类语料库数据显示”等语句,因此他们的反馈表现的更为客观。在控制组的反馈意见中,包含较少的模糊词语和提示,这一方面表明他们的反馈意见更加客观,另一方面也显示出他们对自己提出的纠错意见更有信心,对其正确性也更有把握。

其二,反馈意见的接受度和改正错误的有效性。测试要求译员将评判员给出的反馈意见,按帮助提高翻译技巧的有效程度从低到高进行排序,1分代表最低档,10分代表最高档,其统计结果如表3所示。根据表3的统计数据显示,得分排在前三名的都是控制组提供的反馈,而倒数三名是对照组提供的反馈。这表明学生认为控制组的反馈意见比对照组的在提高翻译质量和帮助翻译学习方面更有效,更容易为学生所接受^①。因此,利用评价语料库提供的反馈意见在客观度、可接受度以及有效性方面优于传统资源。

^① 学生的书面反馈意见也反映了评价语料库在帮助学生自主对译文进行纠错方面的作用,如译员3认为“学生评判员2的反馈意见只标注了翻译错误,但是学生评判员6的反馈意见还解释了错误的原因,并提供了5个语料库中的相似例句,这给我改正错误提供了参考依据,还帮助我提高了学习和翻译能力”。

其三，译文质量控制。对照组译员被要求只能按照以传统资源为参照标准的评判员提供的反馈意见修改译文，控制组译员被要求只能按照以评价语料库为参照资源的评判员提供的反馈意见修订译文，然后要求参与实验的医学专家根据修改后的文本从技术的准确性和译文的可读性，语言专家根据文章的连贯性、衔接性和总体写作质量，将文本从低到高进行排序，其中 1 分代表最差，8 分代表最好。该项实验旨在确定哪种类型的反馈能真正帮助学生提高翻译质量，其统计结果如表 4 所示：

表 3 学生译员对评判员的反馈意见的接受度与帮助度统计表

组名	评价员	排名得	排名得	排名得	排名得	排名得	排名得	排名得	排名得
		分译员	分译员	分译员	分译员	分译员	分译员	分译员	分译员
		1	2	3	4	5	6	7	8
对照	学生评价员 1	5	4	3	4	3	5	5	5
	学生评价员 2	4	5	5	5	5	4	4	4
	学生评价员 3	1	1	1	2	1	3	3	2
	学生评价员 4	3	2	4	3	4	1	1	1
	学生评价员 5	2	3	2	1	2	2	2	3
控制	学生评价员 6	9	8	9	10	8	8	8	9
	学生评价员 7	6	6	6	6	6	6	7	6
	学生评价员 8	10	7	10	9	10	9	10	10
	学生评价员 9	7	10	7	7	9	10	9	8
	学生评价员 10	8	9	8	8	7	7	6	7

表 4 的统计数据显示，在专家给出的排名中，排在前两名的译文，都是根据控制组提供的反馈意见修订过的译文。排在倒数两名的，都是根据对照组提供的反馈意见修订的译文。此外，根据控制组反馈意见修订的译文都排在前三位，根据对照组的反馈意见修订的译文，只有一位专家给出了第 3 的排名。专家提供的排名表明，不管从译文的专业性上看，还是从语言的内容和形式上看，采用控制组的反馈意见修订的译文比采用对照组的反馈意见修订的译文更符合目标客户的要求，译文质量更高。

3. 其它实验结果

从纠错的准确性上看，在对照组中出现的错判和提供错误修改意见的情况，在控制组中并未出现。从纠错类型上看，对照组发现并纠正的错误大部分都属于词汇选择和语法错误，控制组除了在识别并纠正了同样的错误外，还处理了术语的语域错位、搭配错误、语义错误、文体错位、标点不当等其它类型的错误。对学生评判员后续访谈中发现，对照组评价员都认为，若仅使用现有的传统资源去熟悉一个完全陌生领域的概念，用语非常困难，需要花相当长的时间。控制组的评判员认为，评价语料库提供了大量可比较的目标语言文本，使得他们可以迅速了解该领域及其相关领域的概念和语言特征，从而能快速熟悉一个新的领域。因此评价语料库比传统资源更能够帮助译者和评判员快速地进入不熟悉的主题或领域，但其缺点是，它需要用户提前掌握语料库和语料库分析工具的使用方法。只有在接受过培训的情况下，语料库作为评价资源才能发挥作用。

四、实验结果讨论

表 1 和表 2 的数据分析表明，评价语料库作为错误识别、纠正参考资源和评价标准具有较强的可行性。在单位时间内，错误识别和纠正的数量、效率、准确率远远高于传统资源。对于同一

错误的识别和纠正以及术语的全文统一性上,有着不可比拟的优势,对于应用文本的翻译评价,如科技文本、法律文本、商务文本等,评价语料库显示出更强的客观性和准确性。表3、表4的数据分析表明,评价语料库能够切实有效地帮助学生改进译文质量,帮助学生提高翻译能力。相较于传统评价标准和资源,评价语料库的优势显而易见。评价语料库包含大量的文本,既可以提供翻译的文本材料,又可以提供评价标准,还可以改进译文质量、提高学生自主学习的效率。然而目前其却没能在校翻译教学中广泛应用,究其原因,也许在于:其一,成品专业术语库较少。相较于翻译的内容和题材类型,目前现存的术语库要么词条庞杂,要么词条数量太少,很难满足各种专业翻译的需求,而术语库需要译者和教师长期积累整理,很难在短时期建立各种专业性强、体量大的专业术语库;其二,现存的能够满足各种专业翻译的语料库太少,而自建和使用语料库都需要很强的专业技术,因此,只有少量的译者或教师能够将语料库应用到翻译教学中。大部分高校应用于翻译教学中的现代科技工具,还不是当今最先进的科技工具,这一方面是由于师资的原因,另一方面也是由于对话料库和现代翻译技术认识的不足,缺乏相关的课程设置,因此在高校课程中设立翻译技术课程,培养学生应用语料库和各种计算辅助翻译技术的能力已经刻不容缓。另外,值得指出的是,在翻译教学 and 实践中,传统的评价资源和标准也不可弃之不顾。相比于客观性较强的应用文本,文学文本的翻译更多体现的是情感、美感和文化的异质性,对译文的主观评价大于客观,所谓千人千译,不求趋同,因此评价语料库标准在文学文本翻译的评价中会稍显生硬,所以传统的评价标准和资源仍是不可或缺的。

表4 专家对修订译文的排序统计表

组名	学生译文	排名得分	排名得分	排名得分	排名得分	排名得分
		专家1	专家2	专家3	专家4	专家5
对照组	学生译文1	3	3	2	1	3
	学生译文2	1	2	3	3	2
	学生译文3	4	4	4	6	4
	学生译文4	2	1	1	2	1
控制组	学生译文5	7	6	7	8	7
	学生译文6	6	7	7	4	6
	学生译文7	5	5	5	5	5
	学生译文8	8	8	8	7	8

五、结 语

毋庸置疑,“在译文质量评估中,始终会存在主观推测”^{[8]24}。然而,上述实验测试表明,使用评价语料库可以明显降低翻译评价中的主观因素,给评判员和译员提供客观有效的建设性反馈意见,从而帮助他们更准确、更快速、更高效地识别和改正译文中的大量错误,并且使用语料库作为评价资源,学生可以更容易地获得大量真实语境中的例句,能更好地帮助他们学习总结翻译技巧、提高翻译能力,然而评价语料库并不是万能的。它目前还不能完全排除或取代人的主观评判。另外,因本次测试的规模较小,参与测试的人员来源还比较单一,因此本次实验测试的结果还有待完善,但作为一个面向未来的研究领域,有必要对评价语料库应用的潜在可能性做更宽泛的研究。与其简单地把它作为一种资源来帮助教师纠正学生所犯的错误,不如利用评价语料库

开发一种更积极的教学方式来帮助学生自己识别和避免错误。使用评价语料库作为教学和学习资源, 显然能为学习者提供更多自主学习的可能性。

参考文献

- [1] House J. A Model for Translation Quality Assessment [M]. 2nd ed. Tübingen: Gunter Narr Verlag, 1981: 5.
- [2] 单宇, 郭雪梅. 国内基于语料库的翻译教学研究综述: 2003-2014[J]. 湖南科技学院学报, 2015 (12): 171-175.
- [3] 任敬辉. 国内基于语料库的翻译研究综述: 2007-2013[J]. 长春师范大学学报, 2015 (6): 181-185.
- [4] 王克非, 熊文新. 用于翻译教学与研究的英汉对应语料库加工处理[J]. 外语电化教学, 2009 (6): 3-9.
- [5] 李德超, 王克非. 基于双语旅游语料库的 DDL 翻译教学[J]. 外语电化教学, 2011 (1): 20-26.
- [6] 刘泽权, 刘鼎甲. 多媒体计算机技术与语料库方法运用于翻译教学改革的尝试[J]. 外语与外语教学, 2009 (8): 27-30.
- [7] 王健佳. 基于语料库的大学英语英汉翻译量化评价实证研究[J]. 外语教学理论与实践, 2013 (4): 53-57.
- [8] Hönl G. Positions, Power and Practice: Functionalist Approaches and Translation Quality Assessment [C] // Schäffner C. Translation and Quality. Clevedon: Multilingual Matters, 1998.
- [9] Manning A. A Classroom-bound Approach to the Meaning of Translation Quality Assessment [C] // Lewandowska-Tomaszczyk B, Thelen M. Translation and Meaning: Part3. Maastricht: Universitaire Pers Maastricht, 1996: 546.
- [10] 梁平. 医学英汉双语语料库[DB/OL]. [2018-09-05]. <http://www.e-charm.com.cn/yjk.asp>.
- [11] 卢伟. 厦门大学双语语料库[DB/OL]. [2017-02-07]. <http://www.luweixmu.com/ec-corpus/query.asp>.

(编辑: 邢娟妮)

An Empirical Study of Evaluation Corpus and Translation Quality Evaluation Based on Translation Teaching

WANG Hongmei

(School of Foreign Languages, Wenzhou University, Wenzhou, China 325035)

Abstract: In translation teaching, the quality of students' translations is a direct reflection of the teaching effect. The evaluation mechanism of the quality of students' translation has always been one of the key and difficult problems in translation teaching research and reform. This paper introduces the evaluation method for the students' translation, which is based on bilingual parallel evaluation corpus, and its verification effect. It attempts to explain that the evaluation corpus can effectively improve the objectivity of the evaluation, provide constructive feedback to the students, and help improve the quality of their translations and their translation skills. The evaluation corpus can provide useful reference and helpful inspiration for the improvement of translation teaching content and methods.

Key words: Evaluation Corpus; Translation Quality Evaluation; Evaluation Mechanism; Objectivity

(英文审校: 黄璐)