[文章编号] 1003－4684(2025)03-0080-05

基于情感分析的余华《活着》英译本海外读者接受度研究

孙雅茹，刘　洋

（湖北工业大学外国语学院，湖北 武汉 430068）

[摘　要] 为了解余华《活着》英译本海外读者接受度及阅读喜好，通过使用 Python 对美国亚马逊网站该英译本 *To Live*：*A Novel* 的评论进行抓取、降噪、清洗后导入情感分析工具库 TextBlob，获取读者星级评价占比表、读者年度评论量趋势图、词云图、网络语义图、情感评价占比表等多维定量数据，而后通过人工筛选标注的方式对文本内容进行质性研究，结果显示海外读者对该英译本评价高，接受度良好。

[关键词] 情感分析；海外接受度；读者评价；《活着》

[中图分类号] H315　　[文献标识码] A

乡土文学作为近现代中国文学的重要组成部分，既为国内读者提供了解乡村文化与风土人情的途径，同时还肩负着向外国读者传播中国优秀文化的责任。余华小说《活着》是中国乡土文学最具影响力的代表作之一。截至 2022 年《活着》已被译介到英、美、日、韩等多个国家。其英译本 *To Live*：*A Novel* 由汉学家白睿文翻译、Anchor Books 出版社于 2003 年出版发行，一经发行便引起西方媒体的热烈讨论。《时代周刊》《华盛顿邮报》等多家媒体给予了其高度评价[1]。

情感分析(Sentiment Analysis)又名意见挖掘，通过对含有情绪表达的文本进行整合、归纳、处理进而推论得出受众对作品的接受度和评价好坏[2]。目前情感分析的方法主要分为两类，一是基于情感词典，二是基于机器学习。情感词典方法主要利用文本中带有情感倾向词汇对文本进行分类，在识别并提取目标文本中包含的情感词后，根据词典已有数据将单个情感词进行匹配，通过求和或加权平均等算法计算出目标文本整体情感值。机器学习方法则主要依赖于使用机器学习模型，将已标注的文本作为训练集训练出文本情感分类器，并基于此分类器对新的目标文本进行情感分析[3]。情感分析可以用于了解外国读者对文学作品中情感内涵的评价。张璐[4]、李书影和王宏俐[5-6]、石春让和邓林[3]、李洁和魏家海[7]等学者使用情感分析技术对中国文学作品在海外的传播效果和接受程度进行了定量和定性研究，拓展了中国文学作品对外传播研究的方法。《活着》英译本 *To Live*：*A Novel* 由美国汉学家白睿文(Michael Berry)进行翻译，目前针对该译本的海外接受度情况研究主要是定性研究为主，实证研究比较缺乏[8]。本文以该译本 *To Live*：*A Novel* 为例，通过 Python 语言爬虫提取美国亚马逊网站上英文读者对该作品的评价，使用情感分析工具整理海外读者对该作品的评价，使用定量和定性相结合的方法，了解海外读者对该作品的接受度，为其后中国近现代文学对外译介和传播提供有益参考。

1　研究方法

本研究数据来源于美国亚马逊网站。截至 2022 年 11 月 21 日，亚马逊网站数据显示 *To Live*：*A Novel* 在该网站上共有 369 位读者对其进行了评级，其中有 112 位读者发表了在线评论，剩余 257 位读者仅对书籍进行了评级打分。利用 Python 语言爬取共 112 条评论后，通过人工降噪，剔除其中的无效评论，主要涉及仅发布单个表情符号而无文字评价，及与书本内容无关的评论(如图书封面、价格、书本质量、网站配送服务等)，共获得有效评论 105 条。

根据网站自身具备的评分系统，可以统计获得 *To Live*：*A Novel* 历年读者评论人数及各星级评分占比。词云图的绘制需要在 Python 中安装 Jieba

[收稿日期] 2023－09－01

[基金项目] 湖北工业大学基金项目人文社科类(2022SW0068)

[第一作者] 孙雅茹(1997－)，女，安徽蚌埠人，湖北工业大学硕士研究生，研究方向为翻译学。

[通信作者] 刘　洋(1979－)，女，吉林白山人，湖北工业大学副教授，研究方向为翻译学。

和 Wordcloud 两个模块并将筛选后的有效评论导入。Jieba 模块是 Python 中常用的分词模块，可以将导入的句子拆分为词语，Wordcloud 则会根据词语出现的频率绘制词云图。

完成评论筛选后在 Github 网站获取 TextBlob 文本处理库，粘贴到 Python 代码编辑器 Jupyter Notebook 中。TextBlob 工具是一个基于 NLTK 构建的开源文本处理库，它在 NLTK 库的基础上进行了简化，相对来说更加容易使用。它可以执行诸如拆分词语、词性标注、名词短语提取、词形还原、文本翻译等文本处理任务，其包含了用来执行情感分析的预先训练好的机器学习模型。在 Python 编程器中加载完 TextBlob 工具包后，使用其默认的 Patter-n Analyzer情感分析器进行情感分析[9]。将评论文本导入后可以获得文本情感分值，极性(Polarity)浮动范围为[－1，1]，值越接近－1 则评论情绪越消极，值为 0 则为中立，值越接近 1 则评论情绪越积极。

2　数据分析

2.1　历年读者评论量及星级评分

译作成功与否的衡量标准之一是译介的效果，而来自读者们的评价与反馈是决定译介效果的重要标准[10]。研究采用亚马逊网站读者的星级评价标准，满意度从一星到五星逐渐递增。截至 2022 年 11 月，根据网站统计，《活着》英译本总评分数为 369 人，总体评分为四点五星，与章璐[8]2018 年的研究相比，上升了 0.1 星，证明好评量有所上升。其中五星及四星好评打分人数共有 326 人，占比高达总人数的 87％(表 1)，由此可见绝大部分读者对《活着》英译本内容关注和认可度较高，其在海外读者中广受好评，备受青睐，这一结果与章璐[8]和赵佳薇[11]的研究一致。其中一些人给出“Awesome book!”“I highly recommend it.”“Great novel!”等积极评价，表明该译本引起了海外读者的强烈兴趣。这些积极的评价不仅提高了该译本的国际知名度，也为吸引更多的读者阅读该译本提供了强劲的动力。同时，给出一星评价的 6 人中有 2 人发表了评论。这两位读者针对网站物流及电子书排版给出较低评价，与译文文本本身质量并无关系。

表 1　读者星级评价占比

星级	评分人数	人数占比/％
五星	261	72
四星	65	15
三星	26	8
两星	10	3
一星	6	2

由图 1 所示，自 2003 年到 2016 年亚马逊网站上关于 *To Live*：*A Novel* 评论数量基本呈上升状态，其中 2013 年至 2014 年陡然上升并在 2015 年达到顶峰，2016 年则保持平稳，说明外国读者对该书关注度持续增加，展示出该译本在国际上的影响力。2016 年至 2021 年书本评论量少许下降，读者关注度有所降低，但评论量仍高于 2003 年至 2012 年中的任意一年，表明外国读者近年对该书的关注度相较于十几年前仍有所增加。

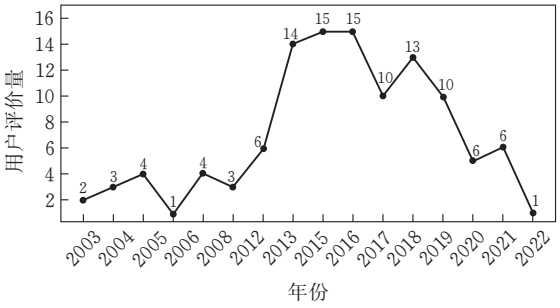

图 1　读者年度评论量趋势

2.2　高频词云图

词云图按关键词出现频率以字体大小进行可视化分布，实现了更直观的情感分析和展示。图 2 的词云图能够直观、详细和全面地呈现海外读者对《活着》英译本所关注的主题和焦点。

由运算得出的词云图可以看出“book”“life”“story”“fugui”“Chinese”等词在词云图中最为突出，出现频次共为 221 次，表明读者在发表评论时多谈论小说故事本身及主人公生活经历，同时也更加着重于对背后的中国元素和中华文化提出相关看法。除上述词语外“recommend”“well”“great”“good”等夸赞性词语也较为显眼，证实读者对该文有积极的情感倾向。作者余华的名字占据空间也较大，表明部分读者对作者本人也十分关注。此外“written”“writing”“translation”等词语的出现频次共为 26 次，展现了读者对写作手法、写作风格以及翻译质量发挥的作用有所意识。在这些因素的共同作用下，《活着》英译本保障了阅读的流畅性，让读者感觉《活着》在英语环境中自己开口说话，极大呈现了中文原作的精神世界[12]。“hard”“hardship”“survive”“tragedy”“poverty”等带有消极色彩的词语出现频次共 35 次，展示全文写作的基调、福贵生活环境的艰辛以及读者对福贵命运多舛的一生的同情。“civil”“war”“famine”“political”等关键词则表明了海外读者对中国当时的社会和文化现状及人民生活状态的关注。综合主要词频，可以发现海外读者通过该译著了解了中国当时的文化和历史背景，深刻地感受到了小说中主要人物的生活场景和生存

方式[10]。

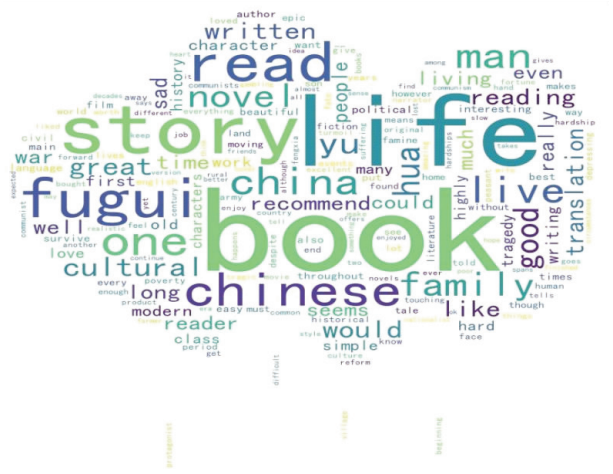

图 2 高频词云图

2.3 语义网络分析

在文本分析中语义网络分析是指从文章中提取关键词，利用词语之间的相似度关系构建词语的关系网，进而探索文本表达的意思。网络语义图以图形表现词语间的结构关系，帮助分析词组间的相关性。相同颜色的点代表同一主题，点的大小表示连线多少，点越大则连线越多，与该词相关联的词语也越多，该词语在网络中的作用也越突出。词语距中心节点（核心节点）距离越近，与中心节点词语的联系就越紧密[13]。而线条的疏密代表共现频率的高低，线条越密，表明共现次数越多[14]。

由图 3 可以看出，“book”一词为网络图的核心节点并向外发散构成核心层，主要包括“live”“yu hua”“novel”“life”“story”“one”“written”“well”“sad”“hard”等词语来构成读者对该书的第一印象，即《活着》是作者余华所写的一部关于一个男人悲惨、艰难的一生的小说。

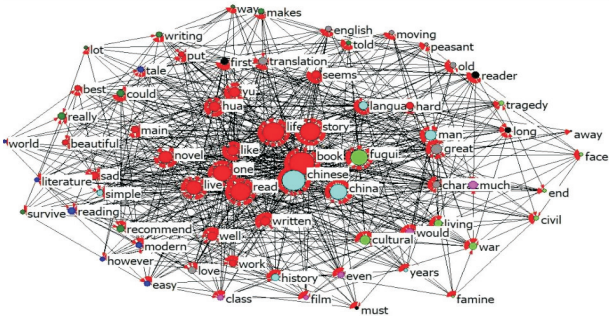

图 3 语义网络

次一级的关键词为“Chinese”“China”“language”“simple”“history”读者紧接着关注的是小说中的语言难度、语言表达、文化和历史。而后是以“fugui”为中心节点，相关主题词为“living”“war”“tragedy”“famine”与书中福贵经历的诸如“战争”的时代洪流相关。除去以上同一主题下的关联词组，还有位置相邻的关联词组的关系可以加以分析。如“class”“film”这对词组展示了读者了解到《活着》这本书的多种渠道，即朋友推荐、班级学习了解抑或是喜欢电影而来了解原著。

3 读者评价情感分析

机器学习情感分析手段效率高但存在一定缺陷。机器学习情感分析的难点在于有时不结合语境的情况下，机器学习进行的情感判断很可能是错误的，而一些程度词也会对文本的情感分值造成影响。当一些情感词语前存在否定词时，考虑程度词和不考虑程度词获得的结果可能是截然相反的。鉴于此次研究文本量不大，且机器在筛选时受程度词影响易产生歧义，因此除了使用机器学习手段进行情感分析外，笔者也根据读者评论所表达的感情色彩对文本进行了人工标注。

由图 4 可知，使用 TextBlob 进行情感分析后可得读者情感占比。其中 83.9%的读者对 *To Live*：*A Novel* 给出了积极评价，4.5%的读者情感较为中性，11.6%的读者则给出了负面评价。以上数据表明 *To Live*：*A Novel* 获得了大部分的读者的认可，这部分读者对小说内容给予了正向肯定。一小部分的读者则对小说表达了不满，不符合他们的期望。

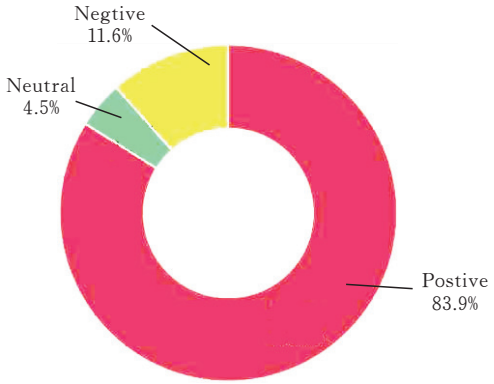

图 4 整体情感分析占比

对书的整体评价进行了情感分析后，笔者针对读者对人物、情节、语言三方面的评价进行了二次筛选并人工标注，并绘制统计图（图 5）以便进一步研究。多数读者评价都集中在情节的设置上，大多数都给予了正面评价，认为其情节抓人，让人手不释卷。其中不乏“love”“surprisingly interesting”“moving”“epic”等字眼，表明读者被福贵飘摇的一生所触动。关于情节的负面评价则认为人生苦难充斥在故事的每一个角落重复而琐碎（The misery seems repetitive and trivial），行文缓慢故事线重复（The story was slow and the story line just repeated over and over），福贵由赌徒浪子向家庭责任承担

者形象转变太快(His turnaround into a responsible family man seems very rapid)。人物方面，给予了正向评价的读者认为福贵虽然经历了如此多的苦难但依然坚强坦然的生活着(The work provides insight into the mind and life of a man who, despite several great tragedies, keeps moving forward)。负面评价多集中于对青年福贵的描述，读者用“Who fritters away his wealth in a drunken night of gambling”描述了该读者眼中醉生梦死、沉迷赌博的纨绔子弟的青年福贵。语言方面，一些读者认为小说译文语言优美，简单易读(Masterfully written, very easy to read, yet goes to the core of life)。另一部分则觉得译者使用的语言太过简单，没有什么文学价值(There is not much literal value)，怀疑译文是否存在扭曲原文的意思的现象，认为译文折损了一部分原文的幽默(Though the book had some witty and funny parts, nothing really struck me. I think many parts of the books were probably lost in translation)。但是整体来说负面评价远远少于积极评价。

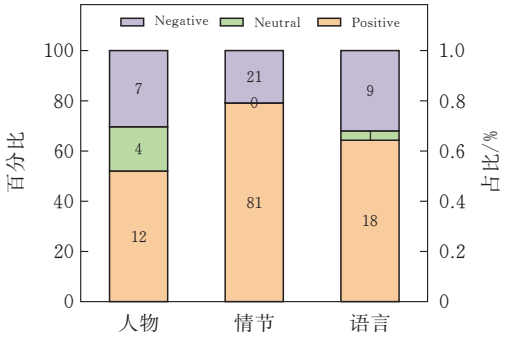

图 5　人物、情节、语言情感分析占比

4　译本质量评价

海外读者对于译本质量的评价主要围绕该译本的词汇、句法、文体、语用等维度。余华是我国著名乡土文学作家，他文字精炼、言语简单流畅而富有幽默感。小说主题往往围绕死亡、苦难、鲜血和希望展开。白睿文作为一名汉学家同时又是英语母语者，其译文表达符合英语读者使用习惯，大多数读者对其接受度较高。一部分读者认为白睿文的译文简单易读，流畅自然，延续了余华的语言叙述风格。面对余华原文中一些字句简单但结构工整的句子时，白睿文在翻译中也用简单的句子将其完美重现出来。如：做牛耕田，做狗看家，做和尚化缘，做鸡报晓，做女人织布[15]。其译文为：Oxen plough the fields, dogs watch over the house, monks beg for alms, chickens call at the break of day and women do the weaving[16]。白睿文在原有排比结构的基础上还将前三句进行了押韵，使得译文产生了音韵美，声音和情感相互交融，读起来更加朗朗上口，重现原句中文顺口溜似的行文格式和巧思。

白睿文的译本也存在一些不足，正如有些读者所言，其译文相较于原文存在一些偏差。如原句为：那个时候空气里满是焦煳和硝烟味，吸到嗓子眼里觉得有一颗一颗小石子似的东西[15]。白睿文将其译为：The air was dense, filled with the burning scent of gunpowder. The smoke-filled air made my throat itch, and small grains of charred ash got caught in my eyes[16]. 在中文语境中“嗓子眼”仅仅是指喉咙，而白睿文的译文却将其划分开来译为“throat”和“eyes”，转换了概念，与原文表达不相匹配，造成了部分文化认知上的障碍。

5　结束语

本研究通过 Python 语言爬取了美国亚马逊网站关于余华《活着》英译本 *To Live*：*A Novel* 一书共 105 条有效评论。整理年度评论数后不难发现海外读者对该书的关注度呈先上升后缓慢下降的趋势。经过数据预处理后，将文本数据导入 TextBlob 文本处理库进行情感分析，获取读者评论情感值，通过绘制读者星级评价占比表、读者年度评论量趋势图、词云图、语义图等手段，将文本数据可视化，得出海外读者对该书评价较好、接受度较高的结论。

由此我们可以获得中国优秀文学作品在译介、出版、推广以及对外传播的一些启示。中国优秀文学作品在外译过程中逐渐获得了认可和关注，但有些作品存在叫好不叫座的现象，如余华《活着》英译本 *To Live*：*A Novel* 虽然进入了海外市场，但海外读者评价的数量相较于其他类型的中国文学作品如科幻小说《三体》和哲学著作《道德经》并不高，这类文学作品不应只是“走近”海外读者，更要“走进”海外读者。除此之外，在翻译层面上，译者要满足目标读者对中国文学作品的期盼，考虑受众水平，并以此为着力点采用合适的翻译方法，精准传递文本信息，再现文本风格，让海外读者易于理解的同时注重可读性。

[参 考 文 献]

[1] 李晗佶.余华作品在英语国家的译介与传播[J].辽宁工业大学学报(社会科学版)，2014，16(05)：61-63.

[2] 赵妍妍，秦兵，刘挺.文本情感分析[J].软件学报，2010，21(08)：1834-1848.

[3] 石春让,邓林.基于情感分析技术的莫言小说英译本在西方的接受程度研究[J].外国语文,2020,36(03):91-96.
[4] 张璐.从 Python 情感分析看海外读者对中国译介文学的接受和评价:以《三体》英译本为例[J].外语研究,2019,36(04):80-86.
[5] 李书影,王宏俐.《道德经》英译本的海外读者接受研究:基于 Python 数据分析技术[J].外语电化教学,2020(02):35-41.
[6] 李书影,王宏俐.基于读者接受的中华文化典籍外译研究:以《孙子兵法》为例[J].出版发行研究,2022(06):92-97.
[7] 李洁,魏家海.基于在线书评的《生死疲劳》英译本海外接受研究[J].解放军外国语学院学报,2022,45(03):144-151.
[8] 章露.《活着》英译本的外国读者接受研究及译介启示[J]. 文艺生活,2018(09):77-78.
[9] 保罗·戴特尔. Python 程序设计:人工智能实例实践[M].王恺,王刚,译.北京:机械工业出版社, 2021.
[10] 石春让,赵慧芳.《活着》的英译策略与译介效果的关系研究[J].语言教育,2020,8(04):63-67.
[11] 赵佳薇.《活着》的英译与传播[J].作家天地,2022(32):117-119.
[12] 吴赟.中国当代文学的翻译、传播与接受:白睿文访谈录[J].南方文坛,2014(06):48-53.
[13] 刘萌玥,陈效萱,吴建伟,等.旅游景区网络舆情指标体系构建:基于蚂蜂窝网全国百家 5A 级景区的游客评论[J].资源开发与市场,2017,33(01):80-84.
[14] 郑华伟. 红色旅游价值观内化的网络文本研究:兼论国民幸福感的生成机制[J].旅游学刊, 2016, 31(05):111-118.
[15] 余华.活着[M].北京:十月文艺出版社, 2013.
[16] YU HUA. To live: a novel[M].Michael Berry, trans. New York:Anchor Books, 2003.

A Study on the Reception of Overseas Readers of *To Live*: *A Novel* from the Perspective of Sentiment Analysis

SUN Yaru, LIU Yang

(*School of Foreign Languages, Hubei Univ. of Tech., Wuhan* 430068, *China*)

Abstract: This paper used Python to collect, denoise, and clean opinions posted on Amazon in the United States regarding the English translation of Yu Hua's *To Live*: A Novel in order to understand the reception and preferences of overseas readers. The TextBlob sentiment analysis library was used to generatereaders ' star rating, the number of reader comments, word clouds, semantic networks, and sentiment scores. Review content was qualitatively researched through manual filtering and annotation. As a result of the study, overseas readers had high opinions of the English translation and it was generally well received by them.

Keywords: sentiment analysis; overseas reception; evaluation of readers; *To Live*

[责任编校:闫 品]