

基于语料库对张爱玲自译风格的研究

王海燕

上海海事大学外国语学院, 上海

收稿日期: 2025年9月6日; 录用日期: 2025年9月26日; 发布日期: 2025年9月30日

摘要

张爱玲是中国现代文坛的一颗彗星, 她是一位天生的作家, 她的作品确立了她在文学界的地位。但少有人知, 为了生计, 张爱玲还曾当过翻译, 而她所留下来的翻译作品, 都具有独特的审美情趣。她翻译的作品, 是以读者为导向的, 用词符合目标读者的阅读习惯, 同时还能准确地传达原文的含义。作为一名作家, 张爱玲担得起天才之名, 作为一名翻译家, 她同样算得上传奇。本文以张爱玲自译作品——《金锁记》为研究对象, 自建语料库, 通过语料库软件WordSmith 4.0对译作的类符形符比、平均词长和平均句长等一些基本数据进行统计和量化分析, 剖析张爱玲的自译风格以及在翻译中所使用的翻译策略和翻译方法, 从多个维度客观地挖掘一个张爱玲作为翻译家的自译风格。

关键词

张爱玲, 译者风格, 异化, 直译, 语料

A Corpus-Based Study on Eileen Chang's Self-Translation Style

Haiyan Wang

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: September 6, 2025; accepted: September 26, 2025; published: September 30, 2025

Abstract

Eileen Chang was a comet on the stage of modern Chinese literature, a born writer whose works secured her an unshakable position in the literary world. Yet few people know that, for the sake of livelihood, Chang also worked as a translator. The translations she left behind are marked by distinctive aesthetic qualities. Her works are reader-oriented: the word choices conform to the reading

habits of target readers while faithfully conveying the meaning of the original text. As a writer, Eileen Chang undoubtedly deserves the title of genius; as a translator, she can also be regarded as a legend. This paper takes Chang's self-translation *The Golden Cangue* as its research object. A self-built corpus is constructed, and corpus analysis software WordSmith 4.0 is employed to generate statistical and quantitative data such as type-token ratio, average word length, and average sentence length. On this basis, the paper analyzes Chang's self-translation style, as well as the strategies and methods she employed in translation, aiming to objectively explore her distinctive style as a translator from multiple dimensions.

Keywords

Eileen Chang, Translator's Style, Foreignization, Literal Translation, Corpus

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,国际上对语言文字信息化方面的研究越来越深入,语料库的创建与研究成为世界性学术热点[1]。语料库经过几十年的发展,在理论和实践应用领域都已经趋向成熟,在很多领域中都能见到语料库的应用,其大量的应用经验帮助人们加深对自然语言的认知,使之应用得更加得心应手。

张爱玲作为享誉国内外文坛的女性作家,其文学创作类型包括小说、散文和剧本。其晚年寓居美国研究“红学”,为世人对《红楼梦》的研究留下浓墨重彩的一笔。人们在惊叹着一个绝代才女的文学传奇,谈论着她师承曹雪芹的成败得失。一生大红大紫的张爱玲,至今仍是文艺青年的最爱。人们一开始对于张爱玲的关注大多聚焦在其中文作品创作上,很少有人注意到张爱玲“翻译家”这个身份。她的母亲黄逸梵曾出国留学,并决意用欧美教育模式培养女儿,因此张爱玲的英文功底非常深厚,这也为她日后进行英汉汉英翻译创作活动打下坚实基础。且许多读者不知道的是,“爱玲”这个名字本身就是由其英文名“Eileen”翻译得来的。

作为中国近代最具影响力的女作家之一,张爱玲出生于新文化运动的关键时期,因此她的作品中带有特定时代的烙印和作者特有的思想。她独具魅力的创作风格和写作手法一直以来令人痴迷,这也是文学作品价值的由来。因此张爱玲一直以来文学界研究的热门对象,并且各学者对张爱玲作品研究的切入点也各不相同,比如意向构建、女性主义、东方主义、功能理论和后殖民主义理论等。但是就目前来说,对张爱玲文学作品的分析大都只是传统意义上的文学分析,缺少令人信服的量化分析即多以定性为主即缺乏定量的数据和量化的支持[2]。因此,本文选取张爱玲自译作品——《金锁记》(*The Golden Cangue*)为研究对象,通过 WordSmith 4.0 软件自建语料库,对其自译的文学作品进行量化分析。《金锁记》描写了一个小商人家庭出身的女子曹七巧的心灵变迁历程,揭示了旧中国女性在封建社会中的压抑与苦难,体现了张爱玲对于女性生存问题的深切关注。

近年来,语料库翻译研究在方法论上不断丰富,除了传统的词频和 TTR 外,学界已广泛采用多种词汇多样性指标、搭配/搭配网络分析、关键词(keyness)分析、词块/短语搭配(lexical bundles)检索、语篇层面的主题建模和语用/语篇指标等。本研究以《金锁记》为例开展探索性分析,认为量化证据应与文本细读相结合,以避免对文本产生单一维度的解读。

2. 文献综述

2.1. 自译研究现状

自译(self-translation)是指“将自己的作品翻译成另外一种语言的行为或是这种行为的结果”[3], 自译者就是把自己的作品翻译成另一种语言的译者。自译活动发端于西方中世纪与文艺复兴时期, 而对于自译的研究则直到二十世纪七八十年代才开始陆续出现[4]。四十多年来, 自译研究逐渐兴起。有相当一部分学者认为, 自译不能称为纯粹意义上的翻译, 因为作为原文本创作者, 在将其作品翻译成其他语言时, 不可避免或多或少增加了自己的创作成分。然而, 赵彦春认为自译本身就是翻译, 因为它同样涉及同一内容的两种语言的不同版本, 体现了语言形式和功能的转换[5]。

2.2. “风格”研究现状

风格是指语言使用者在说话或作文中表现出的规律性特征, 并且这些特征在一定时期、一定范围内相对稳定[6]。许钧认为, 风格可以再现, 具有客观可观性。译者风格表现的是翻译者自身的个性, 这种个性具有稳定性和规律性[7]。译者风格这个研究方向是随着描写翻译理论的出现而出现, 通过将译者的翻译文本与非翻译文本或其他译者文本相比较得出相关结果。随着翻译研究方向由规范性走向描写性, 加上计算机科技的进步和发展, 基于语料库的译者风格研究得以全面展开[8]。研究者借助语料库工具, 在大量数据事实的基础上, 描述和阐释翻译文本的特点和译者的翻译倾向, 并以此刻画出译者的整体翻译风格。

2.3. 语料库语言学研究现状

语料库是一个提供语言的描述、研究和分析的资料库。而语料库语言学兴起于二十世纪五十年代后期。研究者通过检索工具对语料库的文本进行统计和分析。基于语料库的翻译研究采用语料库语言学的工具和方法对大量的翻译文本进行描述, 并从文本体现的语言特征中探究翻译的规律性。Sinclair 认为语料库翻译学起源于应用语料库工具分析译者风格或文本的语体特征[9]。

王克非率先以“语料库翻译学”这一术语向国内学界引介该领域研究以来, 经过了多年发展历程, 语料库译学的学科属性逐渐清晰[10]。现今, 语料库检索分析法对研究文学作品的语言特点或主题的表达提供了可靠的量化依据[11]。与其他语言学分支学科相比, 语料库语言学的优势体现在它通过分析 and 统计功能, 对文学著作进行实证性研究, 使得短时间内把握目标文本的核心骨架成为可能, 从而能更全面、理性地对目标文本进行剖析, 得出客观、严谨的结论。

2.4. 《金锁记》(The Golden Cangue)自译本研究现状

国内对张爱玲的自译作品的研究开始于上世纪八十年代, 而以《金锁记》英文版为研究对象的学者也有不少: 杨轶跳出直译和意译的对立结构, 反映出译作所蕴含的复杂时代意义[12]; 叶爱香认为张爱玲对《金锁记》的翻译以异化策略为主, 并研究了这一翻译方法的成因[13]; 在译文中的体现; 杨妮从翻译适应选择论的视角探索张爱玲如何适应翻译生态环境并采用何种翻译方法, 揭示以译者为中心地位的翻译适应选择论对《金锁记》翻译的指导意义[14]。

通过阅读大量文献可以发现, 目前对《金锁记》的自译研究主要是结合各类理论视角进行研究, 缺少定量的数据和量化的支持。

3. 研究问题和方法

本文通过语料库的研究方法, 采取定性和定量相结合的研究方法, 选取张爱玲中英自译作品——《金锁记》(The Golden Cangue), 进行量化研究和数据分析, 结合张爱玲自译中采取的翻译方法, 探究张爱玲

在中英自译小说中的风格。本文首先通过对比类符形符比、平均词长和平均句长 3 个方面的统计数据，对张爱玲的这部汉英自译作品进行分析，从而得出张爱玲在作为翻译家，自译作品时的风格以及作品的文体特征；其次，探索张爱玲在自译时多采用的翻译策略与翻译方法。

本文选取的语料库工具为 WordSmith 4.0。WordSmith 4.0 是一个综合性的语料库分析工具套件，其主要功能为检索、词频列表和主题词统计。该软件界面简洁，操作难度适中，是语料库研究中最常用的工具之一。确定语料库工具之后即可开始建库工作，该小型自建语料库的建构过程比较简单快捷。第一步，将英译版小说进行断句处理，存储为 txt 文档，注意要把格式转为 UTF-8。第二步，用 WordSmith 4.0 打开 txt 文档即可进行检索研究。由于以单部《金锁记》建立的语料库规模较小，且为了使数据更具说服力，本文选取 BROWN 语料库(普通英语语料库)中小说部分作为参考语料库。

4. 研究结果与动因分析

4.1. 类符形符比

类符(type)指“不重复出现的形符数”；形符(token)指“一个语言单位”，一个单词代表一个形符。Baker 认为，类符/形符比值的高低与写作者词汇使用的丰富程度和多样性成正比因此，类符/形符比(TTR)可以用来判断译文词汇的丰富程度[15]。TTR 越高，则译文词汇使用重复率越低，词汇越丰富。利用 WordSmith 4.0 中的 Word List 功能可以检索出语料库中的类符和形符的数量。检索得，中文版《金锁记》类符为 3100，形符为 3299；英文版类符为 3466，形符为 22,776；BROWN 语料库类符为 32,238，形符为 351,549 (表 1)。

Table 1. Comparison of TTR between the English Version of *The Golden Cangue* and the fiction section of the BROWN Corpus

表 1. 《金锁记》英文版和 BROWN 语料库(小说部分)的 TTR 比较

	《金锁记》英文版	BROWN 语料库(小说部分)
类符	3466	32,238
形符	22,776	351,549
类符/形符	15.21	9.17

数据表明，英译版《金锁记》TTR 为 15.21%，BROWN 语料库(小说部分)TTR 为 9.17%，由此可得张爱玲译文用词更丰富、更多样，词语变化程度高于相关原创英文文本。TTR 较高，不仅与张爱玲深厚的英文语言功底密切相关，而且很大程度上依赖于张爱玲原作的艺术风格——对意象的扑捉与精当的描写、用大量比喻通感修辞手法来推进情节并烘托人物心理。

4.2. 平均词长

平均词长指特定语料库中形符的平均长度，反映词汇复杂程度。平均词长越长，译文用词更为复杂、高级，阅读难度也更大，而单词长度一般可以直接反映译者的用词习惯和风格。根据 WordSmith 4.0 的 Word List 功能可知，英译版《金锁记》平均词长为 4.34，BROWN 语料库(小说部分)的平均词长为 4，两个语料库平均词长如表 2。

Table 2. Comparison of average word length between the English version of *The Golden Cangue* and the fiction section of the BROWN corpus

表 2. 《金锁记》英文版和 BROWN 语料库(小说部分)平均词长比较

	《金锁记》英文版	BROWN 语料库(小说部分)
平均词长	4.34	4

由表 2 可知,《金锁记》英文版的平均词长比 BROWN 长 0.34。这个结果说明,相对于 BROWN 语料库(小说部分)中的原创作品,张爱玲自译作品中更倾向于使用比较复杂的词汇,阅读难度更高。比如,在翻译“凤求凰”(Phoenix Seeking the Phoenix)时,中文原句:“她的婚姻就像是凤求凰一样的安排。”张爱玲将英文版处理成:“Her marriage was arranged like the phoenix seeking the phoenix。”“凤求凰”这个成语来源于中国古代文学,用来形容一种理想的恋爱或婚姻。它的字面意思是“凤凰寻求凰鸟”,寓意着一种完美无瑕的匹配或爱情。这里便是采用异化策略,保留了“凤求凰”,富有中国特色,这一传统文化背景的比喻,未转化为西方读者更容易理解的表达。但是作为中国读者,我们可以肯定的是以张爱玲的英文功底,她完全可以用地道的英文来表达,帮助目标语读者理解,但她选择了更接近原文的译法。张爱玲在翻译时选择了字字对译的手法,为的是尽量保持习语原有的结构和文化特征,从而有助于目标读者了解中国文化,体会文化差异。亦或是在处理“闺房”这个中文特色词汇时,中文原句为“她的闺房布置得非常讲究。”,英文版翻译为“Her boudoir was meticulously arranged.”这里采用异化策略:将“闺房”翻译为“boudoir”,而不是对标目标语中的“卧室——bedroom”,这样处理保留了原有的文化色彩和隐私空间的特性,读起来更有中国韵味。

4.3. 平均句长

平均句长指的是文本语料中平均每个句子的长度,可以用文本总词数除以文本总句数计算出一个语料的平均句长,可以用每句平均多少词来表示。根据 WordSmith 4.0 的 Word List 功能可知,英译版《金锁记》共有 1507 个句子,因此可得这部作品平均句长为 15.11 个单词(表 3)。

Table 3. Comparison of average sentence length between the English version of *The Golden Cangue* and the fiction section of the BROWN corpus

表 3. 《金锁记》英文版和 BROWN 语料库(小说部分)平均句长比较

	《金锁记》英文版	BROWN 语料库(小说部分)
平均句长	15.11	15.02

两语料库的句子都属于中等长度的句子,张爱玲汉英自译本的句子稍长一些。除却别的影响因素,文本的平均句长越长,语句所含的信息量越大,句法复杂度会增加,可读性也会相对降低[16]。因此,平均句长可以在一定程度上反映句法复杂度和文本可读性。译者在翻译过程中如果选择尽量保持原文的形式,就要求以直译为主,句子则较短,而平均句长直接影响作品意思的传达,信息量大,句长也会随之增加。数据表明,张爱玲自译作品平均句长与 BROWN 语料库(小说部分)相差甚微,说明张爱玲在尽可能传递中文文本信息时,也保证了英文版《金锁记》的可读性。

平均句长也可以反映张爱玲所采取的翻译方法,作为译者,张爱玲在翻译过程中,坚持以直译为主。多使用直译的翻译方法,可以减少译者干预,译者不进行过多的解释或注释,让读者自己去理解和感受原文的意义;同时直译常常保留原文中的文化特征和背景信息,张爱玲有更强烈的愿望将原文的内容和韵味传达给读者,因此采用直译翻译会保留原文中的特定文化术语、习语或风俗习惯。例如翻译称呼与称谓时:七巧(Qiqiao), Sister Yun(云妹妹),直译法可传达出主要人物的年龄、地位、人物关系、辈分等信息;翻译成语与谚语时:早起的鸟儿有虫吃(The early bird has worms to eat);三寸金莲(Three-Inch Golden Lotus,直译中国传统的束脚习俗。)

通过研究数据可得,张爱玲自译作品《金锁记》(The Golden Cangue)中,类符形符比相对于目标语小说类型来说要高,这直接表明张爱玲在自译作品时使用更丰富的词汇。此外,就平均词长来看,英译版《金锁记》比 BROWN 语料库(小说部分)要高,可看出张爱玲自译作品中更倾向于使用比较复杂的词汇,

具有一定的阅读难度。她多采取异化的翻译策略,从而保证原文特色,这对更好传播中国文化起到积极作用,同时利于目标语读者更好感受中国文化的韵味。最后,通过平均句长的数据可知,张爱玲自译作品与 BROWN 语料库(小说部分)相差甚微,张爱玲多采用直译的翻译方法,减少译者对目标语读者的干预。不管是在语言层面还是在文化层面,张爱玲的翻译策略和翻译方法都在不断地进行探索和调整,为的是让译文保持平衡,力求能让目的语读者理解和接受,更重要的是让目标语读者感受到中国文化的特色。

5. 研究局限与不足

本项研究还存在一些局限与不足,比如本研究采用的 WordSmith 4.0 为经典工具,但并非近年功能最全或对新指标支持最直接的软件;以单部自译小说为研究对象,样本规模有限,易受篇章结构与事件分布影响;Brown 语料库为经典参照,但其构成与语料年代及体裁可能与本研究不完全匹配;TTR 对文本长度高度敏感,宜采用 MATTR、MTLD 等更稳定指标并进行归一化处理;此外需要统计显著性检验来判定差异是否具备统计学意义。

6. 结语

《金锁记》(*The Golden Cangue*)是张爱玲翻译生涯中的代表作,鲜明而突出地体现出她的翻译特色:一方面尽量保持原著的风格,另一方面注重目标语读者的文化和理解,融合中西方文化的同时最大程度保持中文的韵味,尽显她的翻译特色。本文运用语料库工具 WordSmith 4.0 可以剖析作家的写作风格,例如探索作家的用词习惯,推测作家的语言功底以及采取的翻译策略和翻译方法等,可见通过语料库工具能较完整地挖掘一个作家的作品特性,有助于文学研究。本文运用语料库对张爱玲自译的《金锁记》(*The Golden Cangue*)进行分析,为研究张爱玲作为翻译家时提供了较为客观的数据支撑。

利用语料库检索软件对译文进行实证性的研究,将定量和定性研究方法有效结合起来,这样可以客观、直接地反映出译文的特征,从而归纳总结出译者自身的翻译风格,为译本文体研究提供了科学的方法。当然语料库工具只能作为一个辅助工具,为认识文学作品提供直观的数据佐证,并不能代替文学鉴赏[17]。因此,将语料库语言学与从文本出发的文学鉴赏相结合,进行主客观诠释分析,能更好地理解文学作品。

参考文献

- [1] 许阳. 基于语料库的外文字幕转码质量自动评价系统设计[J]. 自动化与仪器仪表, 2020(2): 81-84.
- [2] 陈吉荣. 谈本位论观照之下的《金锁记》自译[J]. 北京第二外国语学院学报, 2007(10): 1-6.
- [3] Grutman, R. (2019) Self-Translation. In: Baker, M. and Saldanha, G., Eds., *Routledge Encyclopedia of Translation Studies*, Routledge, 514-518. <https://doi.org/10.4324/9781315678627-109>
- [4] 张倩. 四十年自译研究: 现状与不足[J]. 外国语(上海外国语大学学报), 2020, 43(3): 110-119.
- [5] 赵彦春. 翻译学归结论[M]. 上海: 上海外语教育出版社, 2005.
- [6] 黄立波. 语料库译者风格研究反思[J]. 外语教学, 2018, 39(1): 77-81.
- [7] 许钧. 文学翻译批评研究[M]. 南京: 译林出版社, 2012.
- [8] 蔡强, 虞琦. 基于语料库的张培基散文翻译风格研究[J]. 宜春学院学报, 2024, 46(5): 83-89.
- [9] Sinclair, J. (1999) *Corpus, Concordance, Collocation*. Shanghai Foreign Language Education Press.
- [10] 王克非, 黄立波. 语料库翻译学十五年[J]. 中国外语, 2008(6): 9-14.
- [11] 李晓培, 高东军. 《冷山》的语料库检索分析[J]. 四川教育学院学报, 2010, 26(5): 83-85.
- [12] 杨轶. 《金锁记》自译的后殖民主义视角[D]: [硕士学位论文]. 北京: 北京外国语大学, 2014.
- [13] 叶爱香. 张爱玲自译作品《金锁记》翻译策略探析[J]. 宁波广播电视大学学报, 2011, 9(3): 56-59.

- [14] 杨妮. 从翻译适应选择论看张爱玲的自译作品《金锁记》[D]: [硕士学位论文]. 长沙: 湖南师范大学, 2017.
- [15] Baker, M. (2000) Towards a Methodology for Investigating the Style of a Literary Translator. *Target. International Journal of Translation Studies*, **12**, 241-266. <https://doi.org/10.1075/target.12.2.04bak>
- [16] 杨晨. 统计方法在计量风格学中的应用研究[D]: [硕士学位论文]. 南京: 南京师范大学, 2021.
- [17] 丁晓. 语料库 AntConc 软件对《推拿》的分析研究[J]. 现代语文(语言研究版), 2014(8): 122-124.