

# 基于语料库的《离骚》英译本译者风格研究

张浩清

中国矿业大学外国语言文化学院, 江苏 徐州

收稿日期: 2025年12月22日; 录用日期: 2026年1月16日; 发布日期: 2026年1月28日

## 摘要

《离骚》是我国一部著名的长篇叙事抒情诗歌, 其外译具有很高的研究价值。本文基于自建的《离骚》汉英双语平行语料库, 选取英国汉学家大卫·霍克斯及中国翻译家许渊冲英译本为研究对象, 从语言特征方面和非语言特征方面对二人译者风格进行对比研究。研究发现, 许译本突显对“美”的追求, 用词丰富多变, 有“显化”特征, 倾向于采用“归化”策略, 韵律佳, 可读性好, 霍译着重信息准确传达, 其用词较为简化, 但相比许译更难, 信息负载量高, 语言自然流利, 有“隐化”特征, 译文兼顾“归化”“异化”两大策略。

## 关键词

《离骚》, 译者风格, 语料库

# A Corpus-Based Study of Translator Style in English Translations of *Li Sao*

Haoqing Zhang

School of Foreign Studies, China University of Mining and Technology, Xuzhou Jiangsu

Received: December 22, 2025; accepted: January 16, 2026; published: January 28, 2026

## Abstract

*Li Sao* is a famous epic lyrical poem in China whose translation has high research value. Based on a self-built Chinese-English parallel corpus of *Li Sao*, this paper selects the English translation by David Hawkes, the British sinologist and the Chinese translator Xu Yuanchong for comparative study from the perspectives of linguistic and non-linguistic features of the two translators' styles. The study finds that Xu's translation emphasizes the pursuit of "beauty", with rich and varied vocabulary. His translation shows the features of "explicitation" and "domestication", and has excellent rhythm and good readability. On the other hand, Hawkes' translation focuses on accurately conveying in-

formation, with limited vocabulary richness but more difficult vocabulary compared to Xu's. His translation has a high information load which is narrating in natural and fluent language flow. It shows the feature of "implication" and the balances between "domestication" and "foreignization" strategies.

## Keywords

*Li Sao*, Translators' Styles, Corpus

Copyright © 2026 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

《离骚》战国时期楚国诗人屈原所著的一部长篇叙事抒情诗，其在中国古代文学史上具有十分重要的地位。作为中国古代抒情诗典范之一，《离骚》以屈原自叙离别楚国政治的心路历程为主线，表达出了作者屈原对国家兴衰的担忧、对人生沉浮的感慨以及对理想境界的追求。《离骚》不仅在文学艺术上达到极高水平，在思想和文化上也具有十分深远的影响，具有很高的研究价值。百余年来，《离骚》一作吸引了大批翻译家进行翻译研究，其中包括庄延龄、大卫·霍克斯(David Hawkes)、许渊冲、孙大雨等[1]。本文选取英国汉学家霍克斯翻译的《楚辞——南方之歌》(The Song of the South)中《离骚》英译本(以下称“霍译”)及中国翻译家许渊冲翻译的《楚辞》(Elegies of the South)中《离骚》英译本(以下称“许译”)进行对比，建立平行语料库，借助 Lanksbox6.0 及 BFSU Huge Mind Readability Analyzer 进行译本分析，助力新时代中国文化外译进程。

## 2. 理论概述

Mona Baker 在 1993 年首次提出将语料库与翻译研究相结合，由此展开了语料库翻译学的发展，随后，其团队搭建首个翻译语料库，即翻译英语语料库(TEC, Translational English Corpus)，并基于该语料库展开了系列翻译研究[2]。译者风格研究一直以来都是传统翻译研究领域热点，在语料库翻译学的提出背景下，Baker [3]巧妙结合 Gideon Toury 提出的描述性翻译研究，基于其团队搭建的 TEC 语料库，对译者风格从形符、类符、平均句长、词汇密度等方面进行考察，由此创立了语料库译者风格的研究范式。Baker (2000)提到，对译者风格的研究应该关注其译本中典型的表达方式，她明确提出，风格是通过“语言”和“非语言”特征的组合而得以表现，因此可以通过考察译本的语言特征及非语言特征来探究译者风格。Baker 的研究为语料库译者风格领域奠定了坚实的基础，在 Baker 提出这一研究范式之后，许多学者也相继提出了各自对译者风格的理解。Saldanha [4]认为研究译者风格需要找到同一译者在不同翻译作品中体现出来的易于识别的独特特征，这些特征将该译者与其他译者区分开来，构成一种连贯的选择模式。胡开宝[5]提出译者风格既包括译者运用独特语言展现出来的个人特征，也包括译者对文本的原则、翻译策略的选用、脚注等这些非语言特征，而后，黄立波和朱志瑜[6]提出了 T 型译者风格研究和 S 型译者风格研究，T 型译者风格研究主要关注译者的表达方式，而 S 型译者风格主要关注译者对原文所含有的语言特征的重现。

综上所述，译者风格是译者区别于其他译者的翻译或语言模式，Baker 提出的语料库译者风格研究范式主要可以通过考察译本所展现出的语言特征与非语言特征进行研究。

### 3. 《离骚》英译综述

《离骚》作为在中国文学史上具有极高地位的著作，其译本一直以来都备受关注，具有很高的研究价值。《离骚》中丰富多样的文化意象是翻译的一大难题，不少学者针对各个译本中涉及的文化意象翻译进行了深入研究。李月月[7]以许渊冲的英译本作为研究对象，基于框架理论对其中文化意象词英译进行研究，并得出许渊冲主要采用了非理想化框架操作的结论。俞森林和雷佳豪[8]以庄延龄的《离骚》英译本为研究对象，对《离骚》中涉及到的神话相关术语英译进行研究，得出庄译采用的套译、改译等翻译方法在一定程度上造成相关意象原意有所缺失，而其采用的增译、抽象化译法能有效使原文中神话意象传达。由于《离骚》译本高达十几种，因此，对不同译本的对比研究也十分常见。何丹[9]选取了许渊冲、杨宪益、卓振英及孙大雨四人的《离骚》英译本，对“灵”的英译进行了对比研究，并得出译者英译所表现出的差异受译者主观能动性、翻译背景等因素影响。晏红玲与董春枝[10]基于功能对等理论，对杨宪益和孙大雨二人《离骚》译本进行对比，发现两人译本大致上都能实现功能对等，杨译更注重体现原诗的韵律，重现原文氛围，比起孙译，杨译可读性更好。此外，一些学者也借助了语料库进行《离骚》相关译本对比。曹霞[11]在音系学及韵律理论框架的指导下，选取许渊冲、霍克斯、孙大雨、杨宪益译本作为研究语料，建立相应语音语料库及文本语料库，对骚体诗歌韵律翻译进行探究，并得出杨译本在时长方面与原文最为贴合，许译本在情感表达方面与原文最为贴合。

经前期文献搜集发现，《离骚》英译本对比研究是学界热点，但基于语料库对《离骚》译者风格对比的研究较少，此外，在知网搜索关键词“离骚英译”“语料库”“译者风格对比”并未发现有学者基于语料库以许渊冲、霍克斯二人的《离骚》译本作为研究对象进行译者风格对比，而将中国翻译大家许渊冲与英国汉学家霍克斯译本进行对比具有较好的研究价值，因此，本文将依照 Baker 提出的语料库译者风格研究范式，借用 Lancsbox6.0 与 BFSU Huge Mind Readability Analyzer 两大语料分析工具，与从语言特征与非语言特征两个方面对《离骚》许译本及霍译本进行分析，主要选取语言特征中类形符比、词汇密度、平均词长、平均句长以及文本可读性五个参数，非语言特征中的翻译策略进行对比，对二人翻译风格进行研究。

## 4. 语料库翻译学视角下《离骚》英译案例研究

### 4.1. 语言特征对比

语言特征包括形式类参数，即类形符比、平均句长等；语言类参数，即情态小词、语气助词、文化特有词项、人称代词等；叙事类参数，即及物性、自由间接引语、叙事结构等；综合类参数，即可读性[12]。本文主要选取形式类参数中的类形符比、平均句长、平均词长、词汇密度以及综合类参数中的可读性进行分析。

#### 4.1.1. 类形符比

形符(token)类似于日常所提及的“词汇”，其属于一种语言单位。在一段语料中，相同形符重复出现十分常见，若只将语料中的相同形符计算一次，则得到的是类符(type)。类符与形符的比值在一定程度上可以体现出语料用词的多样性，这一比值被称为类形符比，即 TTR (type-token ratio)。

笔者利用 Lancsbox6.0 对两个自建平行语料库进行分析，得到如表 1 的数据：许译本的形符数为 2608，类符数为 927，其类形符比为 35.54%；而霍译本的形符数为 3464，类符数为 1222，类形符比为 35.27%。由此可看出，霍译本篇幅相较于许译本更长，由于语料规模的限制，二者类形符比差距较小，许译本略高于霍译本，但仍可以得出许译本用词相较于霍译本来说更为丰富，善于变化，而霍译本更为简化。细读二人译本也可以找到相应的例子，例如二人对原文植物的翻译。《离骚》中涉及各类植物名称，以“杂

杜衡与芳芷”中的“芳芷”及“朝搴阰之木兰兮”中的“木兰”两种植物为例。对比发现，许译与霍译对“芳芷”及“木兰”两种植物的翻译各有不同。“芳芷”为“辟芷”的别名，为多年生草本植物，因常年生长在幽静偏僻之地，故得此名。“木兰”为常绿大乔木，与“白芷”有所区别。在翻译“白芷”时，霍译为“fragrant angelica”，许译为“fragrant grass”，而在翻译“木兰”时，霍克斯仍然选用“angelica”，而许渊冲则换为“magnolia”，相较于霍译，许译更贴近原文，同时也能在一定程度上反映出许译词汇选择更富于变换。

**Table 1.** Type-token ratio  
**表 1.** 类形符比

	许译本	霍译本
形符	2608	3464
类符	927	1222
形类符比	35.54%	35.27%

#### 4.1.2. 词汇密度

词汇密度可以在一定程度反映文本的信息负载量。通常来说，词汇密度越大，原文的信息负载量则越高。词汇密度无法直接通过 Lancsbox6.0 获得，需要在词类检索的基础上通过人工进一步简单计算得出。词汇密度即实义词总数在原文本中的占比，实义词包括名词、形容词、实义动词和副词，而原文中单词数即形符，通过词类检索和简单计算，笔者得出数据如表 2 所示。

**Table 2.** Lexical density  
**表 2.** 词汇密度

	许译本	霍译本
名词	573	831
形容词	206	221
副词	144	269
实义动词	570	719
词汇密度	57.24%	58.89%

由表 2 可知，许译本词汇密度为 57.24%，霍译本词汇密度为 58.89%，二者相差较小，霍译本略高于许译本，因此，将二者进行对比可得出霍译本实义词使用比例更高，其译本的信息负载量更高，文本难度相对较大。除此之外，经分析可知，许译本中虚词使用的占比略高于霍译本，在翻译时相比于霍译本而言使用“显化”策略，通过增添连词等虚词来增强译本的逻辑性，使文本变得更为清晰明白，而相比而言，霍译本则使用“隐化”策略，比起许译本的表达，霍译本倾向于概括信息或对源文本某些部分进行省略。

#### 4.1.3. 平均词长与平均句长

平均词长即指一篇语料中所选用单词的平均长度，平均句长即指语料中各句所含单词的平均数。表 3 列出了许译本与霍译本句子总数、平均词长及平均句长的数据。

**Table 3.** Mean word length and mean sentence length  
**表 3.** 平均词长与平均句长

	许译本	霍译本
句子总数	374	201
平均词长	4.14	4.37
平均句长	6.97	17.25

通过分析研究得到, 许译本平均词长 4.14, 霍译本平均词长 4.37, 两者相比较, 霍译本平均词长值略高。将二者对比, 许译本虽然整体篇幅上相比于霍译本而言较短, 但其句子总数明显高于霍译本, 因此许译本中善用短句, 而霍译本善用长句。此外, 平均词长在一定程度上与单词难度相关, 平均词长越长, 语料所用长难词越多, 文本阅读难度越高。通过对比两译本平均词长数值可初步判断出霍译本相比于许译本选用了更多长难词, 阅读难度相比许译本略高。除此之外, 在英文中, 通常将 6 字母及其以上的单词认定为长难词。因此, 笔者还利用 Lanksbox6.0 计算了两译本中 6 字母及以上的单词数, 通过检索得出数据如表 4 所示。

**Table 4.** Words of six or more letters in the translations  
**表 4.** 译本中 6 字母及以上单词

	许译本	霍译本
6 字母以上单词数	454	841
占比	17.40%	24.27%

由表 4 可知, 在许译本中含有 6 字母及以上单词数共计 454 个, 占全文 17.4%, 在霍译本中 6 字母以上单词数共计 841 个, 占全文 24.27%。由此可得出, 霍译本所含有的长难词明显高于许译本, 由于长难词与文章阅读难度呈正相关, 因此可以进一步说明霍译本的阅读难度比许译本高。

**4.1.4. 译本可读性**

本文笔者借用 BFSU Huge Mind Readability Analyzer 2.0 中的 Flesch Recording Ease Score 功能对许译本和霍译本进行译本可读性衡量, 该功能便于分析译本阅读难度。文本可读性以评分形式呈现, 分数范围在 0~100 之间, 根据分数可将文本分为 7 个不同等级: 评分在 90 分到 100 分之间可归为 Very Easy; 评分在 80 分到 89 分之间可归为 Easy; 评分在 70 分到 79 分之间可归为 Fairly Easy; 评分在 60 分到 69 分之间则归为 Standard; 评分在 50 分到 59 分之间则为 Fairly Difficult; 如果评分在 30 分到 49 分之间则该文本等级为 Difficult; 若低于 29 分, 则该文本为 Very Confusing。将霍译本与许译本语料分别导入 Flesch Recording Ease Score 功能得出分数如表 5 所示。

由表 5 可知, 许译本分数为 95.47, 所属等级为 Very Easy, 霍译本为 75.35, 属于 Fairly Easy 等级, 两者可读性均较容易, 但相比之下许译本可读性明显更强, 更易于读者理解。如身世开篇:

原文: 帝高阳之苗裔兮, 朕皇考曰伯庸。  
霍克斯: “Scion of the high lord Gao Yang,/Bo Yong was my father’s name.”  
其中专有名词完全音译(Gao Yang、Bo Yong), 读者必须先查背景才能建立“谁是谁”的概念。同时倒装 + 同

位语结构(Scion of ..., Bo Yong...), 对非母语者不算最直观。

许渊冲: “Descendant of High Sunny King, oh!/My father’s name shed sunny ray.”

他把“高阳”意译为 High Sunny King, 瞬间给出“上古帝王”这一语义, 无需脚注。同时两行都是主谓宾正常顺序, 外加句尾“oh!”形成固定节奏, 朗读顺口。因此许用常用词 + 正常语序, 霍用音译 + 倒装, 阅读负荷明显更高。

Table 5. Readability measures  
表 5. 可读性评估分数

	许译本	霍译本
Flesch Recording Ease Score	95.47	75.35

4.2. 非语言特征对比

非语言特征即包括译本的选择、译者翻译方法、翻译策略的选用及译本副文本信息等。本文在非语言特征对比方面主要选取译者采用的翻译策略进行研究分析, 而要对这一特征进行对比可以借助语言类参数。黄立波(2018)曾提到, 语言类参数包括文化负载词、外来词等。《离骚》中有大量与神话相关的文化负载词, 数据如表 6 所示。

Table 6. Mythological terms in *Li Sao*  
表 6. 《离骚》中的神话术语

神话术语类别	数量
神人	40
神地	16
神物	12

本文笔者主要对比霍译本与许译本在对“神人”类别的神话术语进行翻译时所采取的翻译策略。在离骚中, “神人”类神话术语主要包括“神”以及被“神化的历史人物”两类, 本文依次选取了两类中具有代表性例子进行对比分析。

Table 7. English translation of “武丁” and “说”  
表 7. 关于“武丁”“说”的英译

	许译本	霍译本
武丁	his sovereign	Wu Ding
说	a convict	Yue

表 7 中展示的是“武丁”、“传说”两个被“神化的历史人物”的翻译。许渊冲在翻译时将“武丁”译为“his sovereign”, “传说”译为“a convict”, 是采用的二者身份来进行翻译, 采用了归化的翻译策略, 使用了突出精要的含糊释义法, 这样的处理方法增加了原文的可读性和欣赏性, 减少文化隔阂, 使译文流畅通顺。而霍克斯在翻译时直接将“武丁”译为“Wu Ding”, “传说”译为“Yue”, 并在后面



加注，对“武丁”与“傅说”二人的事迹进行了详细的阐述，采用的音译加注的翻译策略。霍克斯的译本整体没有保留“骚体”的风格，而是用自然畅快的英语进行表述，而在这些流畅的英语中出现汉语拼音，会给读者带来一定的陌生感，更易于让读者感受异国文化，使文化得到有效传播。

**Table 8.** English translation of “飞廉” and “雷师”  
**表 8.** 关于“飞廉”“雷师”的英译

	许译本	霍译本
飞廉	The Curtain-Rolling Wind	Wind God
雷师	The Lord of Thunder	The Thunder God

表 8 展现的是许译本和霍译本中对“神”这类神话术语的翻译。许渊冲在翻译“飞廉”这一意象时，没有将其翻译为一个具体的名称，而是将“飞廉”这一人物物化。“飞廉”是“风神”，而许渊冲将其翻译为“The curtain rolling wind”，即汉语中的“卷帘风”，将其力量十分生动传神地展现在读者眼前，而在翻译“雷师”这一意象时简要地翻译为“The Lord of Thunder”，采用归化的策略，易于读者理解。而霍译本也采用了归化的手法，其中，对“飞廉”和“雷师”都用到了“God”，即西方读者熟知的“神”，“上帝”，这也是体现了其采用“套译法”，使用英语中已有的为人熟知的词汇意象来进行翻译，使得文章更为自然。利用 Lanksbox 6.0 中的 KWIC 对“god”进行检索发现，霍译本中共有 4 处将与神相关的神话术语翻译为“god”，而许译本中没有。

5. 研究结论与讨论

经研究发现，许译本用词丰富，富于变化，而霍译本用词相对简化。两者相比，霍译本阅读难度更大：霍译本中实词占比更高，虚词占比更低，其译本相比于许译本所含信息负载量更大，有“隐化”倾向，而许译本虚词占比更高，有“显化”倾向；此外，霍译本选词更难，并偏向用长难句，而许译本用词相对简单，且善用短句，可读性更高。在翻译文化负载词时，许偏向采用归化，增加原文可读性，减少文化隔阂，而霍译本归化异化更为平衡，在采用音译加注的方式异化翻译的同时，对于“神”这类意象也会采用套译法，在翻译时替换为“god”这类西方熟知的意象。笔者将从翻译背景与译者素养、时代背景两方面进行分析。

5.1. 译者素养

霍克斯对《离骚》进行翻译很大程度源自于其对中国文化的浓厚兴趣。霍克斯在牛津大学学习汉语文学期间，主要是研究一些中国典籍，例如《论语》《诗经》等。在当时，《诗经》是诸多西方汉学家的研究热点，但针对汉学领域的研究人少之又少，为了弥补这一空缺，霍克斯选取了《楚辞》中的《离骚》作为自己的本科阶段的研究课题。后来，霍克斯在研究生阶段更进一步研究了楚文化，并将其研究对象扩大到整本《楚辞》。他深度学习了唐诗及白话文，后又进入北大中文系学习，在中国学习的阶段中，翻译是他在课余时间做得最多的事情，在这期间，他对《红楼梦》和《楚辞》进行了部分翻译。由于霍克斯的研究经历，他认为成功的译本应该兼顾准确性及文学性。霍克斯作为一个对中国传统文化有着浓郁兴趣的西方汉学家，他有着西式思维，因此在翻译时，他追求“真实”，着重强调知识的准确性，他曾提到，很多时候，正是因为译者的过度阐释或过度解读才会导致一个译本丢失了原本的美，而翻译的难点正是在于译者需做到将原文的信息准确传达，同时避免冗杂且保留原文原本的美(Hawkes, 1953)。因此，从霍克斯的译文中我们可以看到他十分灵活地运用归化异化的策略，而不是一味的忠实亦或是完全的创

新。

许渊冲作为杰出的翻译家，对中国古典诗词的翻译更是得心应手。许渊冲出生于一个审美氛围极高的家庭，因此，在他的翻译视野中，他也一直在追求“美”。人们熟知他提出的“三美论”、“三化论”、“三之论”，这些理论基础也奠定了他译本的成功。对诗词的翻译并不容易，首先对其理解就有一定困难，再加之其对音韵的讲究，需要译者反复斟酌。而细品许渊冲《离骚》译本，随处可见其对“美”的追求。许译本语言十分凝练，选词巧妙，以八个音节为一行，按照“abab”的形式进行翻译，同时他还巧妙地将原文中的叹词“兮”用英文中的“oh”代替，让无数读者赞不绝口。

## 5.2. 时代背景差异

霍克斯在翻译《离骚》过程中并未受到政治因素的影响。在第二次世界大战后，牛津大学开设了相关汉学科。而中英两国在 1950 年之后交流便逐渐密切，而在这过程中，两国文化也得以传播。在这样的时代背景下，霍克斯对中国经典文学研究十分顺利，并且有机会前往中国学习，这使得他能较好掌握中文应用，因此在翻译过程中他的用词准确，作为英语母语者，他的表达也十分自然流利。

许译本富有创造性，用词更为简单，可读性较高，这在一定程度上受到二十一世纪以来人们对翻译的重视的影响。在二十一世纪中，“以人为本”十分关键，人民的需求得到了很高的关注。许译本最初发表于 2009 年，这时的翻译环境较为自由，在翻译自由的背景下，许渊冲在译本中展现出许多独创亮点，他译文中展现出的简单凝练的语言，令读者赞叹。

## 6. 结语

本文基于自建的《离骚》平行语料库，对霍克斯和许渊冲二人译本进行系统研究，将定量研究与定性研究相结合，依照 Baker 提出的研究范式，从语言特征和非语言特征两个层面对译本进行分析，并探讨结果。将两译本对比研究发现，许译本用词赋予变化，善用短句，文本可读性好，富有美感，保留骚体诗歌风格，并表现出“显化”特征，善用归化。而霍译本用词相比许译本变化较小，但长难词更多，更倾向于用长句表达，可读性不如许译本，但其表达自然流利，归化异化策略采用平衡，表现出“隐化”特征。该研究为探索性尝试，在当今呼吁中国文化“走出去”的背景下，典籍英译备受关注，笔者希望在未来可以更加深入学习，进一步加强语料库翻译学研究，更好助力中国文化外译工作。

## 参考文献

- [1] 冯俊, 谢智辉. 十九世纪西方《离骚》英译研究[J]. 南方文学评论, 2022(3): 277-290.
- [2] 康政. 葛浩文英译莫言长篇小说的译者风格考察[D]: [硕士学位论文]. 北京: 北京外国语大学, 2024.
- [3] Baker, M. (2000) Towards a Methodology for Investigating the Style of a Literary Translator. *Target*, 2, 241-266.
- [4] Saldanha, G. (2011) Translator's Style: Methodological Considerations. *The Translator*, 1, 25-50.
- [5] 胡开宝. 语料库翻译学概论[M]. 上海: 上海交通大学出版社, 2011.
- [6] 黄立波, 朱志瑜. 译者风格的语料库考察: 以葛浩文英译现当代中国小说为例[J]. 外语研究, 2012(5): 64-71.
- [7] 李月月. 基于框架理论的《离骚》中文化意象英译研究[D]: [硕士学位论文]. 重庆: 四川外国语大学, 2022.
- [8] 俞森林, 雷佳豪. 庄延龄《离骚》英译本中的神话术语翻译策略探析[J]. 西南交通大学学报(社会科学版), 2023, 24(2): 109-118.
- [9] 何丹. 国内四种《离骚》英译本中核心概念“灵”的英译对比分析[J]. 语言与翻译, 2022(2): 47-53.
- [10] 晏红玲, 董春枝. 功能对等视角下《离骚》英译本对比研究[J]. 英语广场, 2018(10): 22-24.
- [11] 曹霞. 基于语音语料库的《离骚》韵律翻译研究[D]: [硕士学位论文]. 镇江: 江苏大学, 2022.
- [12] 黄立波. 语料库译者风格研究反思[J]. 外语教学, 2018, 39(1): 77-81.