

基于语料库的泰戈尔《飞鸟集》三个汉译本风格研究

徐 旭

上海海事大学外国语学院, 上海

收稿日期: 2022年5月25日; 录用日期: 2022年6月23日; 发布日期: 2022年6月30日

摘 要

近年来, 语料库成为译者风格研究的重要工具, 借助语料库能更好地揭示翻译的本质属性和译者的主体性。本文运用语料库研究方法, 在高频词、类符形符比、词汇密度、平均句长等层面对泰戈尔诗集《飞鸟集》的三个汉译本进行了考察, 比较郑振铎、徐翰林和冯唐的风格差异。依靠语料库软件, 结合源文本和目标文本进行考察。经过定量和定性分析发现: 在三个译本中, 郑译本语言丰富, 译文注重形式感, 对语篇的掌控能力较强, 译文靠近读者, 多采用归化的翻译策略。徐译本词汇丰富, 译文流畅, 多采用归化的翻译策略。而冯译本注重信息传达, 用词简洁, 注重对原文信息的保留, 多采用异化的翻译策略。

关键词

译者风格, 语料库, 《飞鸟集》

A Corpus-Based Study on the Styles of Three Chinese Versions of Tagore's *Stray Birds*

Xu Xu

School of Foreign Languages, Shanghai Maritime University, Shanghai

Received: May 25th, 2022; accepted: Jun. 23rd, 2022; published: Jun. 30th, 2022

Abstract

In recent years, corpus has become an important tool for the study of translators' style, and the corpus can better reveal the essential attributes of translation and the subjectivity of translators. Using corpus research methods, this paper investigates three Chinese translations of Tagore's collection of poems *Stray Birds* in terms of high-frequency words, quasi-symbol ratio, lexical density,

average sentence length, etc., and compares Zheng Zhenduo, Xu Hanlin and Feng Tang's style differences. This thesis combined with source text and target text for investigation and relied on corpus software. Through quantitative and qualitative analysis, it is found that among the three translations, Zheng's translation is rich in language and has a strong ability to control the text. The translation pays attention to the sense of form. Xu's translation is rich in vocabulary and fluent in translation, and mostly adopts the translation strategy of domestication. Feng's translation, on the other hand, pays attention to the communication of information, uses concise words, pays attention to the retention of the original information, and mostly adopts the translation strategy of foreignization.

Keywords

Translator's Style, Corpus, *The Stray Birds*

Copyright © 2022 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

随着语料库语言学和描述性研究的兴起,由蒙娜·贝克为首的许多学者尝试运用语料库来辅助翻译研究。Baker认为译者在翻译过程中会有一定的主观性和目的性。在译文表达上,一方面会考虑读者,一方面会发挥译者目的语言的掌控能力[1]。胡开宝认为译者风格是译者表现的个性化特征,在翻译策略、方法和翻译文本选择上体现出译者的个人风格[2]。自1999年语料库被引入中国以来,近十年基于语料库的翻译研究迅速增加。

译者风格研究是翻译研究的一个相对新的领域。自从1999年开始,中国学者开始基于语料库的翻译研究。在国内的翻译研究中,译者的风格通常是隐形的。人们常常认为好的译文就是对原文风格的再现,然而常常忽视译者的主观性和创造性。借助语料库能更好地解译翻译的本质属性和译者的主体性。近十年来,语料库成为译者风格研究的重要工具。

刘宓庆认为传统的翻译理论注重对原文的忠实性,主张与原文的风格保持一致[3]。换句话说,传统的翻译理论注重以源语言为中心,在语言学层面对源语言和目标语进行转换。译者只是再现源语言的风格,因此对译者风格的研究并不多[4]。

《飞鸟集》是一部蕴涵了丰富的思想、深奥的哲理的小诗集。由于其在文学的巨大影响力,1922年开始有许多译者翻译这部诗集。其中,郑振铎和徐翰林的译本较为出名。郑振铎的译本是最畅销的。诗歌讲究节奏美、形式美和韵律美。因此诗歌的翻译难度不小,译者对诗歌的翻译能够准确地反映出译者的能力与风格。使用语料库方法对译者的风格研究还相对较少。因此,本文笔者不再以译文的“忠实性”和“对等性”进行翻译研究,而是基于语料库进行译者风格对比研究。

2. 文献综述

《飞鸟集》深受全球读者喜爱。1922年,郑振铎出版了第一本中译本。随后其他翻译家陆续翻译出版这部诗集的中译本。近些年,许多论文从不同角度对《飞鸟集》进行研究。这些研究大多从译者主体性、翻译策略、意境、修辞、比较文学、语言欧化现象等着手分析。

有些学者从译者主体性视角进行分析,比如:国内学者朱璐和贺桂华基于郑振铎译本和冯唐译本,利

用阐释翻译四步骤研究译者主体性差异,认为译者的主体性需建立在忠实原文的基础上[5]。研究强调了译者在翻译中的创造性理念。刘玉涛和周霞从译者翻译作品的文体风格和译者作品的时代语境特征分析郑振铎和冯唐的译本,研究发现两译本在不同程度上都有不同程度的译者主体性的体现[6]。另有一些学者从文体学出发,从意象、修辞、女性主义、翻译策略和译者风格对比研究等层面对《飞鸟集》的译本进行研究。朱文欣从关联理论视角出发,从用词的恰当性、过分押韵和将自己的风格凌驾于作者之上三方面分析冯译本的不足之处,认为冯译本缺乏翻译经验和理论指导而失去一众读者[7]。陈颖超从副文本理论为立足点,对比分析《飞鸟集》郑译和冯译两译本,研究发现了不同译者的个人特点对译作风格的影响[8]。

在传统的翻译研究中,对译者风格的评判取决于读者的主观印象。然而随着语料库研究的发展,把语料库和译者风格结合起来研究会更加客观和系统。蒙娜·贝克认为原作者和译者都有各自独一无二的写作风格[9]。Dorothy·Kenny在《翻译中的词汇及其创造性:基于语料库的研究》一书中详细介绍了基于语料库的翻译研究理论及其方法论[10]。张继东,朱亚菲(2021)采用定量定性的研究方法,从词频、四次成语和译者的非语言特征、翻译策略层面探究李继宏和李静宜的译者风格差异[11]。杜若凡借助 WordSmith Tools 6.0 软件,对《红楼梦》词汇层面进行对比分析,研究发现不同译者的用词丰富度、人称代词的使用、复杂句和定冠词 the 的使用情况不同,并结合语言、文化背景等因素对其原因做出说明[12]。根据以上相关文献,我们可以发现:一部分学者已经开始采取定性和定量的方式研究文学译作的语言特征和译者风格。然而基于语料库的《飞鸟集》中译本研究数量还相对较少。大多数的翻译研究还是用翻译学理论对文本进行探究和不同译本对比分析。因此本文利用 AntConc 和 WordSmith 8.0 软件,探讨《飞鸟集》的郑振铎译本、徐翰林译本与冯唐译本中的译者风格及诗集中译策略探究。

3. 研究工具与研究方法

3.1. 研究工具

(一) ABBYY 软件能够将 PDF 转化成常见的 Word 等格式的文件。

(二) 对中文语料进行标注决定了语料库的质量。因此我们采取 Corpus Word Parser(肖航汉语分词工具)对《飞鸟集》三个中译本进行分词以及词性标注操纵。

(三) EmEditor 软件可以对文档进行降噪,清理语料,提高文字的匹配度和数据的准确性。

(四) 语料库检索软件: AntConc3.5.8 版本和 WordSmith 8.0。利用 AntConc3.5.8 中的 Concordance, Concordance Plot, Word List 功能和 WordSmith 8.0 中 Word List 功能对中文文本进行词,句层面和代词进行分析。

3.2. 研究方法

首先把三个中译本分别转成 TXT 文本格式,去除译本中前言、序言、文内注释,对语料进行清洗和降噪。其次,借助语料库检索软件 AntConc 对自建《飞鸟集》中译语料库,从高频词、标准类符形符比(STTR)、词汇密度、平均句长等参数和代词等层面分析。最后,通过数据分析,对三位译者的翻译异同点进行比较,归纳三个中译本整体翻译策略的倾向,从而总结三位译者的翻译风格和特点。

4. 语料库检索结果统计分析

本小节借助语料库软件 AntConc 和 WordSmith 8.0,对中译本的语言特征进行分析,从而得出三位译者的风格。

4.1. 高频词

AntConc 的一项基本功能就是对高频词进行检索,整理出前十位的高频词绘制成表,进行横向比较,

如表 1 所示:

Table 1. The top ten high-frequency words in the three Chinese translations
表 1. 《飞鸟集》三译本高频词前十位统计

序列	郑振铎译本		徐翰林译本		冯唐译本	
	字	频率	字	频率	字	频率
1	的	728	的	578	的	505
2	我	254	我	236	我	230
3	在	166	在	151	你	113
4	了	133	你	123	在	107
5	你	130	了	91	了	77
6	着	96	着	77	是	57
7	是	93	是	74	中	37
8	它	76	它	56	让	35
9	里	46	中	51	着	32
10	世界	45	世界	40	世界	31

从表 1 我们可以发现, 在三个中文译本中, 频率排前十位的字区别不大, “的”字和“我”字均是三个译本中出现频率最高的字, “的”字多用在定语后, 用来构成没有中心词的“的”字结构, 从侧面说明诗歌集多用修饰词; 第二高频字是“我”字由于诗歌多以第一人称为叙事视角展开描写, 这种独特的手法与第三人称来描述诗歌内容产生不同的艺术效果。另外, 郑振铎版本中第五高频字是“你”, 出现了 130 次, 而在徐翰林版本和冯唐版本中“你”字分别排第四和第三, 分别出现了 123 次和 107 次, 而冯唐版本中“你”字是第四高频词。出现这一现象的原因可能是他们采取的翻译策略不同。

例(1): Do not say, “It is morning,” and dismiss it with a name of yesterday. See it for the first time as a new-born child that has no name.

郑振铎译本: 不要说: “这是早晨”, 别用一个“昨天”的名词把它打发掉。你第一次看到它, 把它当作还没有名字的新生孩子吧。

徐翰林译本: 不要说“这是清晨”, 然后用一个“昨天”的名词来摈弃它。初次见它, 把它当作还没有名字的新生儿吧。

冯唐译本: 别说“又早晨了”, 别用一个早已用过的名字, 见每个早晨如同初相见。

同一个句子, 在三个中译本中, 郑译本在翻译祈使句“See it for the first time”时添加人称代词“你”作主语让句子更加完整, 符合目的语的语言表达习惯, 主动靠近读者, 使用归化的翻译策略。通常第二人称在诗歌中能够让情感传达更真挚, 容易引起读者的共鸣。而徐译本和冯译本没有译出“你”这一人称代词, 采取的是异化翻译策略。

4.2. 标准类符形符比

类符(type)指文本中, 不重复计算的形符数。形符(token)是指一个文本中出现所有词的个数。类符/形符(TTR)在一定程度上能够体现语料的用词变化。一般来说, 译文中类符和形符的数目越高, 文本表达越清晰。我们可以从中看出译者在翻译过程中词汇的丰富程度以及用词偏好。刘肖港认为译者在表达相近的原文时倾向于使用多样的词汇, 那么 TTR 就越高[13]。当译本容量较大时, 通常采用标准类符形符

比(STTR),引入文本固定值后经过均值处理,可以不被功能词影响,STTR和词汇的变化范围成正相关。《飞鸟集》中三个中文译本的标准类符形符比

Table 2. STTR of the three Chinese versions

表 2. 《飞鸟集》三个中译本的标准类符形符比

	郑振铎版本	徐翰林版本	冯唐版本
类符	1661	1658	1641
形符	6496	5868	5054
标准类符形符比(STTR)	25.57%	28.25%	32.47%

结合表 2,通过库内和跨库对比发现:郑译本类符数为 1661,徐译本类符数为 1658,冯译本类符数为 1641。郑译本、徐译本和冯译本的形符数分别为 6496、5868 和 5054。郑译本和冯译本的形符数相差 1442 个。类符形符比可以反映译者用词的多样性,TTR 比值越高,表达越丰富。郑译本、徐译本和冯译本的标准类符形符比分别为 25.57%, 28.25%, 32.47%,说明冯译本的用词更多样,但用词简单;徐译本和郑译本词汇相对固定,但用词复杂。

4.3. 词汇密度

词汇密度是另一个衡量译者风格的重要标准。实词数和文本总字数的比值是词汇密度,比值越高,在某种程度上说明文本的难度大。Olohan 认为词汇密度可以用来衡量不包括语法词的词汇词的比例,也可以通过计算词汇词的百分比得到,可以与词的总数进行比较[14]。在语料库研究中,词汇密度在文本分析中用作描述性参数。一般认为词汇密度大于 60%~70%是密度较高,低于 40%~50%为密度较低。汉语主要由实词和虚词组成,实词分别为六类,即:名、动、形、副、数与量;虚词包括助词、叹词、介词、连词和语气词这类承担语法功能的词。胡显耀认为形合度是指文本中虚词在全文的占比,比值越高,文本形式化程度越强[15]。由于代词本身属于虚词,而它们所代替的却是实词,那么分两次计算形合度。通过中文分词软件 Corpus Word Parser 进行标注,统计结果见图 1 和表 3:

结合图 1 和表 3 分析,有以下两点发现:在实词方面,冯译本的词汇密度最高,郑译本的词汇密度最低,徐译本的词汇密度居中。冯唐译本和郑振铎译本的词汇密度相差了 5.5 个百分点。冯译本的词密度最高,信息负载量大,文本相对难读。可见,郑译本词密度小,信息量小,在词汇的选择上更加通俗

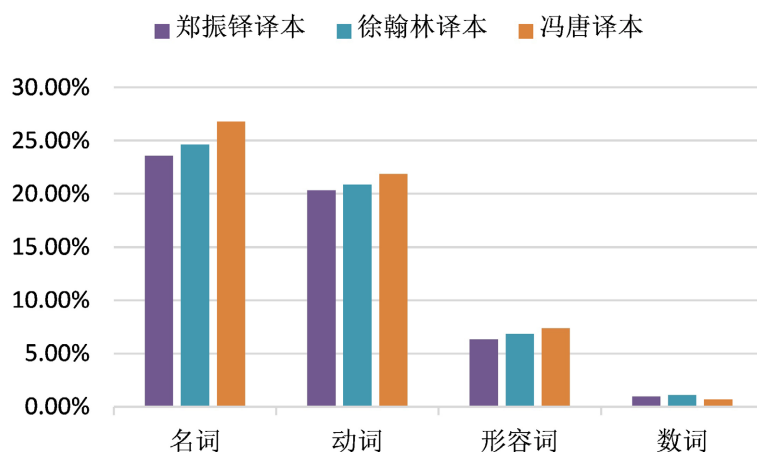


Figure 1. Frequency of use of four types of words

图 1. 四类词的使用频率

Table 3. Lexical density of three Chinese versions
表 3. 《飞鸟集》三个中译本的词汇密度

	郑振铎译本	徐翰林译本	冯唐译本
实词合计	3539	3354	3044
词汇密度	55.46%	58.4%	61.02%
虚词 1	27.34%	24.61%	21.05%
代词	12.98%	12.30%	11.67%
虚词 2	40.32%	36.91%	32.72%

易懂。这一现象是由于郑振铎在对原文充分理解的基础上, 尽量增补信息向译者传递泰戈尔诗歌的内涵。经过两次对虚词的统计, 郑振铎译本的虚词频率皆高于徐翰林译本和冯唐译本, 这说明郑译本的形式化更强。而徐译本和冯译本相对更加注重信息传达。

例(2): This is a dream in which things are all loose and they oppress.

I shall find them gathered in thee when I awake and shall be free.

郑振铎译本:

在梦中, 一切事都散漫着, 都压着我, 但这不过是一个梦呀。

但我醒来时, 我便将觉得这些事都已聚集在你那里, 我也便将自由了。

徐翰林译本:

这只是一个梦, 万物放荡不羁, 压迫着我。

但我醒来时, 我将觉得这些都已聚集在你那里, 于是我自由了。

冯唐译本:

梦里,

所有事物松软,

所有事物侵染。

醒来,

所有事物聚集于你,

于是我自由于天地。

由三个中译本有两点发现: 从翻译策略来看, 郑振铎和徐翰林往往采用复杂的长句, 郑译本的句法成熟度比徐译本略高。郑译本添加了原文字面没有的意思和转折词, 例如: “但这不过是一个梦呀”。郑译本利用连接词来增强诗句之间的衔接, 传达的信息更全面, 更多采用意译法, 并且偏重于语篇的互文性; 徐译本多采用直译法, 长短句结合。冯译本高类符形符比, 高词汇密度更能体现原诗短句特色, 用简练的语言传达诗歌中的丰富思想。从虚词的使用来看, 译文中的“着”和“了”都是虚词, 在三个译本中, 郑译本的虚词使用最多, 文本显化程度最高, 形式化强于徐译本和冯译本。

4.4. 平均句长

平均句长指的是文本的平均句子长度, 对于译本语句的研究也是译者风格的一项重要指标, 语句特征在一定程度上能够反应译者主体风格。句子是能够表达语言中思想的最小单位, 句子长度是一个句子中包含的所有单词总和。按照词数句子可以分成三大类: 短句(1~9 个词); 中等句子(10~25 个词); 长句(超过 25 个词)。可以通过观察译者多用长句还是短句可以看出译者的语篇处理习惯。在某种程度上, 平均句长越长, 句子越复杂。句长标准差反映了句子幅度的变化, 差值越小, 变化程度越小。因此, 句长标

准差是译者的写作布局和文本的操纵能力的体现。本小节采取 WordSmith 8.0 分析《飞鸟集》的三个中文译本的语言特征，在词表选框中勾选统计，便可得到相关数值。结果如表 4 所示：

Table 4. Sentence length
表 4. 三个译本中句长的统计情况

	郑振铎译本	徐翰林译本	冯唐译本
形符数	6496	5868	5054
句子个数	424	414	363
平均句长	14.55	13.39	13.03
句长标准差	6.58	5.65	6.93

从表 4 可知，从平均句长来看，郑译本句长(14.55 个字)长于徐翰林(13.39 个字)和冯唐(13.03 个字)的译本。这说明相对来说郑译本在原文的基础上增译较多，向译作受众传递泰戈尔诗歌的内涵。另外，从表中句长标准差可知，冯唐译本为 6.93，略高于郑振铎译本的 6.58 和徐翰林译本的 5.65。这说明冯唐译本句子句式结构多样化。而徐翰林译本的句长标准差最小，可知徐译本的句子变化幅度最小。英文版《飞鸟集》的句子个数是 348 个，由表可知英文句子数少于三个中译本的句子个数，这说明三个中译本比英文文本更容易理解。通过在句子维度上的两个指标的分析，能够看出郑振铎译本对语篇的掌控力略强于李翰林和冯唐译本。

4.5. 指示代词的使用

指示代词的用法包括用来指示或标识人或事物。译者对中文译本中的指示代词的分析可以帮助读者理解文本。因此，在 AntConc 软件的 Concordance 功能框下检索栏输入“这/那”，设置为“case”，手动剔除不符合规范的检索结果。结果如表 5 所示：

Table 5. Demonstrative
表 5. 指示代词

	郑振铎译本	徐翰林译本	冯唐译本
指示代词	114	92	41
频次	1.06%	0.96%	0.54%

例(3): The sadness of my soul is her bride's veil. It waits to be lifted in the night.

郑振铎译本：我灵魂里的忧郁就是她的新婚的面纱。这面纱等候着在夜间卸去。

徐翰林译本：我灵魂的忧伤是新娘的面纱，等候着在午夜被掀开。

冯唐译本：我消魂的黯然是新娘的面纱等夜晚到来之后再摘下。

例句选自《飞鸟集》中第 98 首，原文中徐译本和冯译本没有使用指示代词，徐译本通过把后半句翻译成了无主语的被动句，让读者明白了原文所指，而郑译本将“it”处理为指示代词“这”，同样体现出读者意识。

例(4): If you shut your door to all errors truth will be shut out.

郑振铎译本：如果你把所有的错误都关在门外时，真理也要被关在门外了。

徐翰林译本：如果你把所有的错误拒之门外，那么真理也会被关在外面。

冯唐译本：如关上门，摒弃所有错误，真题也被堵住。

例句选自《飞鸟集》第 130 首。在徐译本中，原文中“if + 主句”句式翻译为了“如果……那么”。

If you shut your door to all errors 这一条件句十分冗长, 因此, 徐译本使用指示代词“那么”进行说明, 语篇衔接更加流畅。

从表 5 和例句分析我们可以发现, 郑译本和徐译本都尽可能地让译文更好理解, 符合翻译共性中的“显化”现象, 让译文更加靠近读者。姜菲和董洪学认为译者经过加工, 将隐形表达显化, 挖掘出译出语文本中字面下的意义, 使得其在译入语中显化[16]。

5. 结论

本文运用语料库的方法, 对比考察诗歌《飞鸟集》的三个中文译者的风格特征, 对不同中译本进行了客观的定量分析。研究发现三个译本具有一些共性: 三个译本中的高频词大多是代词和语气助词, 符合诗歌体裁的特点。诗歌多以第一人称为叙事视角展开描写, 所以三译本中“我”字使用频率都较高。三个译本的句子数目都高于英文版句子数目, 更易于读者理解。译者风格指的是译者在语言表达中的使用偏好。

郑译本词密度较低, 中文文本句子数目最多, 译本注重形式感, 用词典雅, 说明他考虑到英文诗歌及其文化背景可能对中文读者带来的影响, 使得读者容易理解, 对语篇的掌控力略强, 采用归化的翻译策略。徐翰林译本句长标准差最小, 说明句式变化不大, 译文流畅, 注重信息传达, 采用归化的翻译策略。冯译本类符/形符比和词汇密度最高, 注重对原文信息的保留, 使用词汇较丰富, 平均句长最短, 说明其句式简短, 难度较高, 句长标准差最大, 句子的长短变化大, 采用异化的翻译策略, 部分翻译没有遵守忠实原文的准则, 因此译本的接受度不高。

笔者认为, 尽管语料库方法对译者风格的研究有些许优势, 但同时也有其局限性。与此同时, 语料库研究为译者风格的研究提供了新的视角, 通过定性和定量分析, 可以使得读者对《飞鸟集》中译本三位译者有更进一步的了解。

参考文献

- [1] Baker, M. (1993) Corpus Linguistics and Translation Studies: Applications and Implications. In: Baker, M., Francis, G. and Tognini-Bonelli, E., Eds., *Text and Technology: In Honour of John Sinclair*, John Benjamins, Amsterdam, 233-250. <https://doi.org/10.1075/z.64.15bak>
- [2] 胡开宝, 朱一凡, 李晓倩. 语料库翻译学[M]. 上海: 上海交通大学出版社, 2018: 104.
- [3] 刘宓庆. 翻译的风格论(上) [J]. 外国语, 1990(1): 3-7+41.
- [4] 胡开宝. 语料库翻译学概论[M]. 上海: 上海交通大学出版社, 2011.
- [5] 朱璐, 贺桂华. 阐释翻译观下《飞鸟集》中译本的译者主体性对比研究[J]. 黑龙江工业学院学报(综合版), 2022, 22(1): 151-156.
- [6] 刘玉涛, 周霞. 论译者主体性在冯唐译《飞鸟集》中的体现[J]. 作家天地, 2020(4): 37-38.
- [7] 朱文欣. 从关联理论看冯唐翻译《飞鸟集》的不足[J]. 海外英语(下), 2019(3): 15-16.
- [8] 陈颖超. 从副文本角度比较《飞鸟集》郑振铎译本与冯唐译本[J]. 智库时代, 2017(10): 236+238.
- [9] Baker, M. (2000) Towards a Methodology for Investigating the Style of a Literary Translator. *Target*, 2, 241-266. <https://doi.org/10.1075/target.12.2.04bak>
- [10] Dorothy, K. (2001) *Lexis and Creativity in Translation: A Corpus-based Study*. St. Jerome Pub.
- [11] 张继东, 朱亚菲. 胡塞尼小说译者风格对比——平行语料库视角[J]. 外国语(上海外国语大学学报), 2021, 44(5): 102-114.
- [12] 杜若凡. 基于语料库的《红楼梦》诗词英译本译者风格对比研究[J]. 安徽理工大学学报(社会科学版), 2021, 23(6): 72-79.
- [13] 刘肖港. 基于语料库的《琵琶行》译者风格研究[D]: [硕士学位论文]. 保定: 河北大学, 2017.
- [14] Olohan, M. (2004) *Introducing Corpora in Translation Studies*. Routledge, London.

<https://doi.org/10.4324/9780203640005>

- [15] 胡显耀, 曾佳. 用语料库考察汉语翻译小说定语的能力和结构[J]. 解放军外国语学院学报, 2009, 32(3): 61-66.
- [16] 姜菲, 董洪学. 翻译中的显化思维和方法[J]. 外语学刊, 2009(4): 106-109.