

A Study of the Translator's Style in the Three Chinese Versions of *Sense and Sensibility* Based on Statistics and Analysis of Self-Built Corpus

Tingting Zhao

College of Foreign Languages, Shanghai Maritime University, Shanghai
Email: 759172882@qq.com

Received: Sep. 2nd, 2018; accepted: Sep. 14th, 2018; published: Sep. 21st, 2018

Abstract

This paper adopts a corpus-based approach to compare the three Chinese Versions of *Sense and Sensibility* through corpus retrieval software. The language characteristics of the four translators, Wu Chonghan, Sun Zhili and Wang Jiqing, Tang Yinsun are investigated from lexical, syntactic and textual level through statistics and preliminary quantitative analysis. Thereby the similarities and differences of the three Chinese versions in style are compared and discussed. The findings suggest that influenced by the style and theme of the original novel, the vocabulary richness and lexical density of the three versions are lower than those of the original Chinese texts. And mean length of utterance of the three versions is close to the original Chinese texts. In addition, there are differences in the three versions in terms of vocabulary richness, lexical density, mean length of utterance, translation strategies and so on.

Keywords

Sense and Sensibility, Translator's Style, Corpus

《理智与情感》三个汉译本的译者风格研究——基于自建语料库的统计与分析

赵婷婷

上海海事大学外国语学院, 上海
Email: 759172882@qq.com

收稿日期: 2018年9月2日; 录用日期: 2018年9月14日; 发布日期: 2018年9月21日

文章引用: 赵婷婷. 《理智与情感》三个汉译本的译者风格研究——基于自建语料库的统计与分析[J]. 现代语言学, 2018, 6(4): 581-587. DOI: 10.12677/ml.2018.64068

摘要

本文基于自建多译本语料库,应用语料库检索软件对孙致礼、武崇汉和王纪卿&唐荫荪所译的《理智与情感》三个译本进行了对比分析,从词汇、句法和语篇三个层面对译本的语言特征进行数据统计和初步的量化分析,比较和探讨三个汉译本在风格上的异同。结果发现,受原著小说文体和创作主题的影响,三个译本的词语变化程度和词汇密度低于汉语原创文本,平均句长与汉语原创文本相近。另外,三个译本在词汇丰富度、词汇密度、平均句长,翻译策略等方面都有差异,各有特色。

关键词

《理智与情感》,译者风格,语料库

Copyright © 2018 by author and Hans Publishers Inc.

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

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



Open Access

1. 引言

自 20 世纪 90 年代以来,基于语料库的描写翻译学发展迅速,而且取得了丰硕成果。译者风格的考察是语料库在翻译研究中的一项重要应用。Baker (2000: 245)认为译者风格指“留在文本中的一系列语言和非语言的个性特征”(a kind of thumbprint that is expressed in a range of linguistic-as well as non-linguistic-features) (Baker, 2000: 245) [1]。

译者风格表现在译者对所译文本和翻译策略的选择、前言、后记、脚注、文本中的注释,以及译者个人偏爱的表达形式和重复出现的语言行为方式,即典型语言运用习惯。但是这些习惯特征,尤其是语言表达方面的特征很难用肉眼捕捉,而且主观判断在一定程度上会影响研究的深度和结论的可信度。如果借助语料库技术,就可以对一些难以捕捉和容易被忽视的语言习惯进行多层面的客观对比、全方位的定量分析,从而提出具有可信度和说服力的结论。

《理智与情感》是简·奥斯汀的第一部小说,小说通过睿智的对话和生动的人物形象描绘两姐妹的情感经历和婚姻观,充满讽刺幽默的韵味,引人入胜。埃莉诺善于用理智来控制感情,玛丽安却将自己的情感展露于表,面对爱情时,她们的做法截然不同。多年来,对该小说的研究多从女性文学视角进行,从译者风格,翻译策略进行研究的文章却不多,采用语料库研究的文章也是寥寥无几。通过在中国知网、万方数据库和维普资讯网的搜索,与《理智与情感》翻译相关的文献均不超过二十篇,其中研究中文译本对比的不到十篇,而采用语料库研究相关译本的也不多。因此本文想要通过自建小型语料库来探讨《理智与情感》三个译本的译者风格,并进行对比。

2. 文献综述

对于《理智与情感》的翻译研究,根据在中国知网、万方数据库和维普资讯网的搜索,只有二十篇左右。学者们从不同角度出發,对《理智与情感》的译本进行研究对比。有些学者从译者主体性的视角切入分析,例如:金妙(2012)从译者主体性角度对比研究了《理智与情感》的武崇汉译本和孙致礼译本,从翻译策略和文体风格方面进行对比,分析译者主体性在翻译活动中的体现[2]。梁心爱(2012)从德里达

的“Relevant 翻译”思想探讨了《理智与情感》的翻译，突出了意义的不确定性，强调了译者的主体性地位[3]。李小静(2008)从译者主体性角度对孙致礼和王纪卿两译本进行分析，讨论了两译本在语意表达、文化信息传递、文风演绎和人物性格表述等方面的差异，并浅析了其中的译者主体性因素，最后强调译者是具有创造意识的主体[4]。他们的研究中，重点强调了在翻译过程中译者主体性的体现，同时表现处译者在翻译中的创造意识。

另有一些学者站在文体学角度，从词汇、句法和修辞等层面对《理智与情感》的译本进行探讨。例如：牛桂玲，李孜(2017)研究了孙致礼和武崇汉两译本中 AABB 式叠词的使用情况，并从四个方面分析了 AABB 式叠词在译本中应用的因素[5]。江茹(2012)从文学文体学角度研究了《理智与情感》的小说翻译，从词汇、句法、话语引用、讽刺修辞手段四个层面，分析对比各译本在凸显主题方面的不同，又探讨了各翻译方法和方式的合理性[6]。高华婷(2012)分析了《理智与情感》的武崇汉和孙致礼两译本在词汇层面、语篇层面和修辞层面上传递原文风格的异同，并强调忠实传递原作不仅是对中国传统翻译研究成果的继承，也是翻译的本质要求[7]。以上学者利用语料库技术，对《理智与情感》的两译本从文体学角度分析译本的词汇、句法、修辞、语篇等特点，使得文学翻译作品的研究更加科学化。

根据上述相关研究文献，我们可以发现：文学翻译研究不再停留在传统翻译学视角，一部分学者已将语料库应用到文学翻译的研究中，采用定量分析的方法探讨文学翻译的整体语言特征，但不足的地方是，将语料库研究方法应用到具体的译本对比或者译者风格的研究中的文献较少。而对于《理智与情感》的翻译研究，基本上还都是从传统的翻译学视角研究，在语料库的应用方面很少，因此本文就借助语料库研究方法，通过自建小型语料库来探讨《理智与情感》三个译本的译者风格，并进行对比。

3. 语料库与研究方法

3.1. 语料库

本文选用《理智与情感》的三个译本自建小型语料库，大小约 60 万字。选取的三个译本分别是武崇汉(2011 版)、孙致礼(2009 版)和王纪卿，唐荫荪(2005 版)的三个版本，主要考虑到以下两个因素：首先，译者应是国内翻译界享有盛誉的翻译家，或者译本较受欢迎，译文准确生动。另一个因素是各译本之间的用语习惯，存有有一定差异能更加凸显不同的译者风格和翻译策略。本文选择两个语料库作为参照数据库，一个是兰开斯特现代汉语语料库(The Lancaster Corpus of Mandarin Chinese, 下文简称 LCMC)，兰卡斯特现代汉语语料库是第一个完全免费对公众开放的现代汉语书面语语料库，为肖忠华教授与 McEnery 教授主持创建，百万词级，定位为 Freiburg-LOB Corpus of British English (FLOB)的汉语可比语料库。而且 LCMC 选取了 15 个文体类型，语料平衡，能够比较客观地反映现代汉语普通文体书面语的特征。由于《理智与情感》属于小说体裁，为了能够进一步进行比较，本文同时选取兰开斯特现代汉语语料库的小说子集库(下文简称 LCMC-N)作为参照语料库。

3.2. 研究方法

第一步，使用 ABBYY 软件将 PDF 格式文件转换成 Word 文档，并使用 TextEditor 文本整理器对文本进行降噪处理。第二步，通过 MyTxtSegTag 分词和词性标注工具将语料库文本切分到词汇级别，根据词性给每个词赋予标记代码，并进行人工校对，以便进行后续分析处理。第三步，使用 AntConc 检索软件对三个译本翻译语言词汇、句法和语篇三个层面的不同特征参数进行检索和计算。第四步，将得出的数据与兰开斯特现代汉语语料库、兰开斯特现代汉语语料库的小说子集库以及当代汉语翻译语料库的相关数据进行对比，分析四位译者在语言表达和翻译策略上的不同。

4. 语料库检索结果统计分析

4.1. 词汇层面

1) 类符/形符

形符(Token)是指文本中所有出现的词,而类符(Type)是指文本中不同的词(Baker, 1995) [8]。类符/形符比(TTR)即语料库文本中所使用不同词的数量与总词数量的比值,在一定程度上可以反映语料的用词变化,是统计文本总体特征的重要参数。类符/形符的比值可以体现出译者在用词上的不同:类符/形符比越高,表明该文本使用的不同词汇量越大;反之,不同词汇量越小。因此,通过类符/形符的比值可以比较不同语料库中译者词汇使用的丰富程度和多样性。但是文本长度的差异会直接影响这一比值,在相关研究中一般使用更为可靠的标准化类符/形符比 STTR 衡量语料库文本的词汇变化。

通过表 1 数据显示,三个译本的形符数王&唐译本最高,武译本次之,孙译本最低;在类符数量方面,王&唐译本和武译本分别是最高和最低,孙译本处于中间;标准类符/形符比从大到小排列顺序为:王&唐译本 > 孙译本 > 武译本。由此我们可以得出,王&唐译本文本较长,词汇使用最丰富。孙译本篇幅最短,词形种类也较少,词汇丰富度排在第二。武译本虽然比孙译本文本长,但其使用的词形种类却比孙译本少,因此两者相比孙译本词汇更加具有多样性的特征,而武在翻译过程中一方面尽可能扩增文本长度,另一方面在有意地减少类符数量,以降低译文的阅读难度。三个译本的类符/形符比都明显低于 LCMC 和 LCMC-N,这一点说明:受原著的影响,《理智与情感》三译本的词语丰富度明显低于现代书面汉语,这在一定程度上验证了翻译小说词语“简化”的假设。

2) 词汇密度

词汇密度是实词与总词数比值的百分比,是文本信息量大小的一个衡量标准(Baker, 1995) [8]。词汇密度指实词在语料库中占的比例,也就是实词数占总词数的比例。实义词指具有稳定词汇意义的词语,包括名词、动词、形容词和数量词四大词类。在文本中实词传达的信息通常比虚词要多,虚词主要起着语法和文本的连接功能。篇章中的实词越多则密度越大,传递的信息也越多,难度也相应增加。反之,文本中词汇密度越低,该句承载的信息量就越少,就越容易被理解。

从图 1 可以看出,武译本和孙译本的名词所占比例相差很小,王&唐译本最低;武译本的动词比例最高,王&唐译本最低;形容词所占的比例也相差不甚明显,武译本和孙译本稍高,王&唐译本偏低;数词所占的比例方面,孙译本最高,武译本次之,王&唐译本最低。

从表 2 可以看出,孙译本词汇密度最大,武译本次之,王&唐译本最小。篇章中的实词比例越大,信息量越大,难度也随之增加。据此可以认为,三个译本中孙译本难度最大,武译本难度居中,王&唐译本最小,最易于理解。一方面,三个译本的词汇密度均低于 LCMC 与 LCMC-N,这可基本印证关于翻译小说词汇密度通常低于译入语原创文本或译入语原创小说文本词汇密度的假设。这说明从词语总体特征来看,翻译小说存在用词简化特征,即试图通过降低实词所提供的信息量来降低译文的难度,从而提高其可接受性(胡显耀 2006: 139) [9]。另一方面,三译本的词汇密度虽与 LCMC-N 的词汇密度相差不是

Table 1. STTR
表 1. 标准类符/形符比

	武崇汉	孙致礼	王&唐	LCMC	LCMC-N
形符	107,614	105,495	109,422	760,324	171,670
类符	10,371	11,206	14,136	27,231	12,945
STTR (%)	23.58	27.49	29.87	43.35	44.02

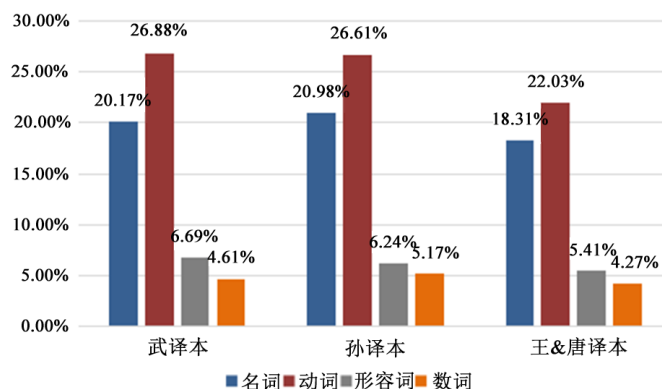


Figure 1. Four word classes use frequency
图 1. 四种词类使用频次对比图

Table 2. Lexical density
表 2. 词汇密度

	武崇汉	孙致礼	王&唐	LCMC	LCMC-N
词汇密度	58.35%	59.00%	50.02%	63.33%	59.13%

很大，原因在于《理智与情感》是一部长篇小说，阅读人群有一定的知识储存，更易于理解，接收小说的内容。

3) 高频词

词频指各词形在文本中出现的频次，词频统计是语料库分析中一个基本的统计手段，在翻译研究中使用词频统计对原文的用词风格、译文的用词风格乃至文体进行研究也有一定的帮助。

从表 3 中我们可以看出，尽管排序和频率有些不同，除了“说”这个特殊词外，这三个译本和两个参照库的高频词语非常相似，这些高频词主要以代词和虚词居多，如“我”“她”“你”等代词，和“的”“了”等虚词，这应和汉语语言的基本特点有莫大的关系，符合汉语语言的习惯。这三个译本和两个参照库的高频词中，“的”字都位列榜首，这主要是因为“的”是汉语中最为常见的虚词之一，汉语中形容词在修饰名词时，都须借助于“的”来进行表达。这说明在词汇层面，译者的翻译一方面会受到原文风格的影响，在某些层面反映出了原作的风格；另一方面，也必然会反映出了译入语本身的语言特点。此外，在三个译本的高频词汇中，“她”字都排在第二位，比两个参照库要靠前，这事实上是和原作的文体特点有很大的关联。因为《理智与情感》是一部描写女性的小说，小说中的人物多以两姐妹展开，因此，“她”字出现的几率必然就会非常大。

4.2. 句法层面

4) 平均句长

与类符/形符比一样，平均句长也是译者风格的一般标记(Olohan, 2004) [10]。平均句长是指文本中的句子的平均长度。虽然句子的长度与句子的复杂程度并不一定成正比，但就整个语料库而言，句子的长短在一定程度上反映了句子的复杂程度。本文通过 word 文档查询功能，分别查询句号、问号、感叹号和省略号的频次，相加后得出句子总数，平均句长的计算方法是形符数/句子数。

从图 2 可以看出，孙译本平均句长最短，也就是说武译本中句间停顿最多，表明他更倾向于把新的信息放在独立的句子中表达，使句子的信息和结构简单易懂，更加符合中国人的阅读习惯。武译本的平均句长最长，表明句子结构复杂，在其翻译的过程中，更加侧重逻辑关系的表达。另外，三个译本的平均句长

句长除武译本外，基本和 LCMC 无明显差别。孙译本和王&唐译本与 LCMC 的数据吻合，符合汉语书面使用习惯，翻译的译本与汉语语言相近。武译本的平均句长高于 LCMC，可能是受译者翻译策略的影响，译者在翻译时可能考虑到作品的阅读群众的水平，从而使译文句子偏长。

4.3. 语篇层面

5) 连接词数量

汉语属于意合语言，句子间或段落间的语义关系不是通过连接词作为纽带直接体现，而是通过句子与句子或段落与段落之间的逻辑关系以及语句的先后序列间接地表现出来；篇章连接往往是借助于语义的内部联系；相对而言，法语和英语在连接词句篇章时则频繁使用连接词，是形式化或形合度较高的语言。在翻译中，句法处理倾向于异化翻译策略的译文往往表现为形合度较高，倾向归化则形合度较低。本文通过 AntConc 检索出连接词的频次，以连接词/形符数比来代表译本的形合度。

通过图 3 可以得出，孙译本连接词使用比例最高，武译本次之，王&唐译本最低。由此可以在一定程度上说明，孙译本相较其它译本而言，更加“异化”，而王&唐译本的形合度最低，也更加符合汉语语言风格，因此最为“归化”。

Table 3. The rate of the top ten characters

表 3. 高频词前十位

	武崇汉	%	孙致礼	%	王&唐	%	LCMC	%	LCMC-N	%
1	的	5.42	的	5.51	的	5.33	的	6.75	的	5.0
2	她	3.33	她	3.14	她	3.41	了	1.73	了	2.41
3	了	2.45	我	2.35	我	2.06	是	1.66	是	1.46
4	他	2.38	他	2.10	了	1.95	在	1.38	一	1.43
5	我	2.32	了	1.98	他	1.91	一	1.33	我	1.36
6	是	1.42	你	1.36	你	1.36	不	1.02	他	1.33
7	你	1.30	是	1.19	說	1.17	和	0.97	在	1.19
8	说	1.10	在	1.03	是	1.13	他	0.78	不	0.94
9	就	0.95	不	0.85	在	1.08	我	0.75	她	0.89
10	在	0.89	说	0.79	不	0.78	有	0.74	你	0.89

平均句长

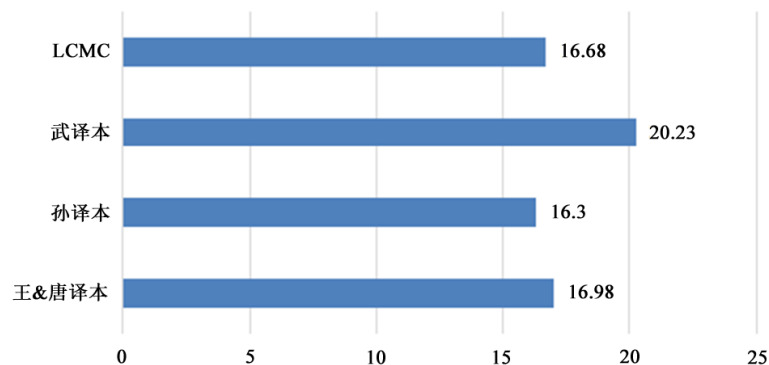


Figure 2. The average sentence length

图 2. 平均句长

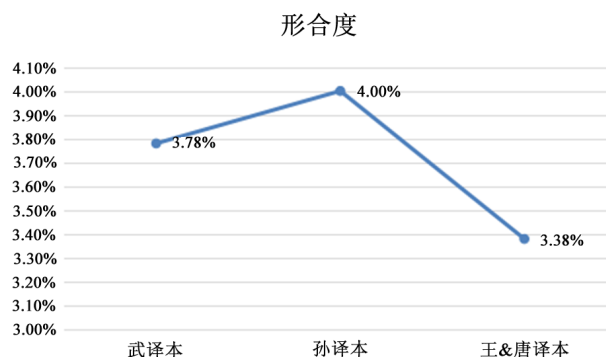


Figure 3. Hypotaxis density

图 3. 形合度

5. 结论

本文采用语料库与翻译研究相结合的手段,从词汇、句法和语篇三个层面对译者风格标记进行了初步的量化研究。首先是三个译本的共性:受原著小说文体和创作主题的影响,三个译本的词语变化程度和词汇密度低于汉语原创文本,平均句长与汉语原创文本相近。高频词主要以代词和虚词居多,符合汉语语言的基本特点和使用习惯。因为《理智与情感》是一部以女性视角的小说,所以“她”字都使用频率很高,比两个参照库要靠前。

由于各译者的职业、所处社会背景、个人进行翻译的时间、目的不尽相同,在不同的研究层面,各译本呈现出各自特有的风格(刘泽权,刘超朋,2011)[11]。其中,王&唐译本文本最长,类符/形符比最高,词汇密度最低,平均句长与汉语原创文本最相近,篇章形合度较低,说明王&唐译本词汇丰富多变,信息量最少,最简单易懂,而且句长和句子间的形合度都比较符合汉语语言风格和使用习惯,翻译策略偏向归化。孙译本文本长度最短,词汇变化程度较低,词汇密度最高,平均句长最短,篇章形合度最高,说明孙译本语言较为精炼,以最短的句子传达最丰富的信息,翻译策略偏向异化。武译本词汇密度最高,词汇丰富度较高,平均句长最长,篇章形合度偏低,说明没有受到英语原文的影响,使用了较少的连接词,易于理解,更易被普通读者接受,译文更贴近汉语语言习惯,翻译策略倾向于归化。

参考文献

- [1] Baker, M. (2000) Towards a Methodology for Investigating the Style of a Literary Translator. *Target*, 2, 245.
- [2] 金妙. 从译者主体性角度对比研究《理智与情感》两中译本[J]. 语文学刊·外语教育教学, 2012(5): 77-78.
- [3] 梁心爱. 从德里达的“Relevant 翻译”思想谈《理智与情感》的翻译[J]. 长春理工大学学报, 2012(7): 80-81, 122.
- [4] 李小静. 译者主体性分析及其在《理智与情感》译本中的体现[D]: [硕士学位论文]. 上海: 东华大学, 2012.
- [5] 牛桂玲, 李孜. 基于《理智与情感》汉译本语料库中 AABB 式叠词的应用研究[J]. 郑州师范教育, 2017(5): 29-32.
- [6] 江茹. 从文学文体学角度研究《理智与情感》的小说翻译[D]: [硕士学位论文]. 上海: 上海外国语大学, 2012.
- [7] 高华婷. 风格传递——《理智与情感》两译本对比研究[D]: [硕士学位论文]. 福州: 福建师范大学, 2012.
- [8] Baker, M. (1995) Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target*, 2, 223-243.
- [9] 胡显耀. 当代汉语翻译小说规范的语料库研究[D]: [博士学位论文]. 上海: 华东师范大学, 2006.
- [10] Olohan, M. (2004) *Introducing Corpora in Translation Studies*. Routledge, New York. <https://doi.org/10.4324/9780203640005>
- [11] 刘泽权, 刘超朋, 朱虹. 《红楼梦》四个英译本的译者风格初探——基于语料库的统计与分析[J]. 中国翻译, 2011(1): 60-64.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2330-1708，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：ml@hanspub.org