

doi:10.3969/j.issn.1672-4348.2015.05.003

# 描写翻译学视阈下林纾译文语料库的创建

戴光荣

(福建工程学院 人文学院, 福建 福州 350118)

**摘要:** 学界对林纾的研究卷帙浩繁。从描写翻译学发展的趋势及语料库在翻译文体、译者风格研究中的运用,探讨林纾译文语料库的创建,并描述林纾译文语料库的创建与标注,指出创建林纾译文语料库的重要意义。

**关键词:** 林纾研究; 林纾译文语料库; 语料库翻译学; 翻译风格

**中图分类号:** H059      **文献标志码:** A      **文章编号:** 1672-4348(2015)05-0419-04

## On corpus construction of Lin Shu's translation works in the vision of descriptive translation study

Dai Guangrong

(School of Humanities, Fujian University of Technology, Fuzhou 350118, China)

**Abstract:** There are diversified studies on Lin Shu in the academic arena. A discussion on constructing Lin Shu's translation corpus is performed with a description of the development tendency of translation science and the application of corpus in the study on translation and translators' style. It elaborates the construction and indication of the corpus from the angle of corpus linguistics and corpus translation study along with an illustration of the significance of the research.

**Keywords:** research on Lin Shu; Lin Shu's translation corpus; corpus translation study; translation style

林纾借助留洋朋友帮助,翻译西洋小说达200多种,成为中国近代文化史、文学史、翻译史上的一大奇观。林纾译文在清末民初的文坛上产生过巨大影响,使国人得以了解西洋文学,直接影响了我国的近代文学发展,而且吸引了包括钱钟书先生在内的一大批读者。<sup>[1-5]</sup>

纵观多年来学界对林纾的研究,特别是译作译品的研究,大多是定性的,带有研究者个人感悟性的。为了更好地对林纾翻译的作品进行科学客观的描述与分析,需要我们借助当代先进的计算机技术与语言学研究的最新成果。随着学科的发

展,尤其是近年来计算语言学、语料库语言学、语料库翻译学、翻译风格学等的发展,为我们更好地开展林纾研究提供了新的视角。本研究试图从描写翻译学发展的趋势,探讨林纾译文语料库的创建可行性及其初步创建,以便在此基础上开展更深入的研究。

### 一、描写翻译学发展的趋势

自1972年Holmes提出翻译研究框架开始,描写翻译学就逐渐引起学界的注意。Holmes认为,翻译研究作为实证科学的一个分支,是一种注

收稿日期: 2015-09-15  
基金项目: 福建省社科规划一般项目(2013B222); 福建工程学院地方文献整理研究中心重点项目(2015DFWX-A01)  
作者简介: 戴光荣(1973-),男,湖南武冈人,副教授,博士,研究方向:语料库语言学、语料库翻译学。

重实证事实描述,较少抽象的理论概括的研究。Holmes 确立了翻译实证研究的两个目标,其一为描述翻译过程与翻译产品有关的各类现象(如语言特征、文化背景、翻译策略等),在此基础上建立普遍原则以解释和预测翻译现象。<sup>[6]</sup>翻译学界特拉维夫学派代表人物之一的 Toury 也认为,一门实证科学如果没有一个描述性分支就不能称其为完整的和相对独立的学科。<sup>[7]</sup>

语料库翻译研究注重从实证的角度对真实的翻译文本进行研究,并以概率性规则作为表述形式,<sup>[8]</sup>使得这类研究成为翻译界的一种新范式。<sup>[9]</sup>该研究范式对翻译研究的助益是多方面的,如帮助研究者提出研究假设,通过语料库收集相关数据与例证,对假设进行验证(或证实或证伪);利用大型数据库,对翻译过程进行探讨研究,从而将语言学、文化与意识形态等方面的研究与翻译研究有机地结合起来,为翻译过程中所面对的复杂现象进行详细的探讨;通过计算机对语料库样本进行自动语言分析(包括词频、词长、句长、型次比、标准型次比、词性统计、词语搭配、语义韵、语法特征等),从而为更好地理解译作提供数据。<sup>[10]16</sup>

翻译研究中采用的语料库包括平衡与专门语料库、历时与共时语料库、平行与可比语料库等。所谓平行语料库,即由源语文本及其译文文本构成的双语语料库。按照 Baker 的定义,它是用 A 语写成的源语文本和用 B 语翻译的译文”组成的一类语料库。<sup>[11]</sup>根据双语对齐层次,可以分为词汇层面对齐平行语料库、句级层面对齐平行语料库以及段级层面对齐平行语料库等。根据翻译方向来划分,可以分为单向平行语料库(uni-directional parallel corpus)、双向平行语料库(bi-directional parallel corpus)以及多向平行语料库。<sup>[12]</sup>

可比语料库是指由同一语言不同变体的文本所构成的两个或两个以上的语料库。用 Baker 采用的术语定义,即“同一语言里两组互为独立的文本集合,其中一个语料库由相关语言中的源语文本组成,另一个语料库则由该语言中译自特定源语的译本组成。”<sup>[11]234</sup>而专门语料库则是指所收集的语料样本来自于同一个具体的话题(如小说、戏剧等)。不同类型的语料库,在翻译研究中承担着不同的功能。

近年来,语料库开始运用到翻译文体、译者翻

译风格的研究中。<sup>[13]</sup>翻译文体是指翻译文本在各种限制性因素的作用和影响下所表现出的规律性语言和非语言特征。这些特征的出现,都取决于译者所做出的各种选择。

Baker 认为,正如手拿一个物体一定会留下手模一样,译者生产出一段译文时,也一定会留下个人的痕迹。<sup>[14-15]</sup>她认为译者风格指的是某个译者的语言特色,属于译者个人的语言习惯。在这个层面上,风格不是有无创造性的问题,而是一种个人倾向性选择的问题。在西方传统上,译者是隐形的,<sup>[16]</sup>因此,传统翻译研究中很少关注译者文体。

近年来这一格局发生了变化。Hermans 指出,翻译文体清楚地表明“在译本中有另一个声音(other voice),即译者的声音”。其实,译者的声音体现在译文的字里行间,只是有时可能完全隐藏在叙述者的背后,使读者无法觉察到它的存在罢了。但译者有时会“冲出语篇层面为自己说话,甚至用自己的名字,例如在译文后的注释中用第一人称解释所述的问题”<sup>[17-18]</sup>。这种情况在林纾译文中体现得更为明显,学界尚未从语料库角度探讨林纾译文的特征。

## 二、林纾译文语料库的创建

林纾一生翻译了大量作品,尤其在文学作品翻译方面,数量众多,有许多译作堪称精品。然而世人对其翻译的作品很少进行穷尽性挖掘,很多研究也只是针对其某些或某一部译作进行探讨,没能从整体上对林纾的翻译风格进行分析与把握,从而做出更为科学、客观而详尽的分析。进行语料收集与语料库创建,对于普通个人研究者来说,是一个浩大的研究工程,会遇到很大的瓶颈,受到很多制肘。这就向我们提出了创建林纾译文语料库的必要性问题。

当前计算语言学、语料库语言学、语料库翻译学的快速发展,可以为林纾翻译研究的瓶颈提供新的解决方法及科学有效的数据。

福建工程学院图书馆拥有一个特色数据库“林纾文化研究专题库”(以下简称“专题库”),已经收集了林译小说 475 种(其中“说部丛书”34 种,“林译小说丛书”27 种,“其他”193 种,“连载小说”221 种),<sup>[19]</sup>如图 1 所示。这些馆藏资料以及当前快捷便利的互联网,为创建林纾译文语料



图1 林纾文化研究专题库截图

Fig.1 Screenshot of special studies on Lin Shu

库提供了可行性和便利性。福建工程学院人文学院着手进行了林纾译文语料库创建的研究工作。

这一拟定的创建林纾译文语料库研究计划的主要任务是语料收集与语料库创建,包括以下阶段:(1)语料选取;(2)语料样本除噪;(3)样本标注;(4)样本入库。

目前初步建立的专题库收藏的都是超星电子文本(pdf格式),且电子书大都是采用古派排版,如图2所示:

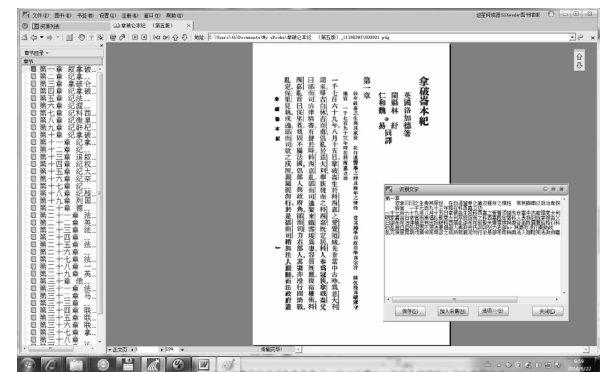


图2 专题库所收录的超星文本截图

Fig.2 Screenshot of superlib text

由于当前语料库处理工具对语料样本要求比较高,必须是纯文本格式,对文本除噪也有较高的要求,因此我们利用批量转换工具尝试对超星文本进行初步转换并尝试在超星阅读器的文字转换功能的基础上,利用人工选取有效区域进行文字识别、校对和保存。如图3所示。

专题库收录的林纾翻译小说都是文言及繁体格式,部分文字识别与输入存在很多困难,需要进行人工校对。为此,研究过程中不可避免地要大

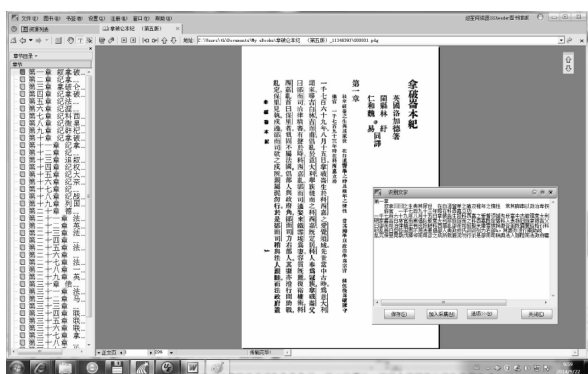


图3 文本识别截图

Fig.3 Screenshot of text recognizing

量投入人力与物力。目前林纾译文的总量基本保持不变(不排除其未曾刊发的译作被发现的可能性),因此,该库的建成将是国内外学界首个专门用途的林纾翻译语料库。

### 三、林纾译文语料库的标注

一个语料库的价值除了所收集的样本之外,还体现在对所收集的语料样本进行标注。这是对语料库样本进行新的增值。语料库标注层次的细化与采用科学的标注方法,将对今后在此平台上展开的林纾翻译研究提供非常重要的帮助。

林纾翻译林纾译文语料库与一般的翻译语料库相比,具有其特殊性。原因在于林纾不通外语,其“翻译”出来的译文是依赖口译者对源语的翻译,在翻译过程中留下的“译者”痕迹是很明显的。另外,林纾本身所具有的桐城派文笔,对其翻译也产生了重要的影响。

我们对林纾译文语料库的信息标注,包含以下几个方面:

1. 元信息的标注:包括源语标题,源语作者(含国别),翻译合作者名字,翻译时间,出版时间(含出版社等相关信息);
2. 译者按语的标注;
3. 语法信息标注;
4. 翻译策略标注;
5. 其他信息如特殊时代各种语言信息标注。

图4所示的是对样本进行的元信息标注格式。

其他各类标注,将随着语料库样本收集的扩展而逐步展开。语料库信息标注能为研究者提供丰富的信息。这类信息通过语料库检索软件,能

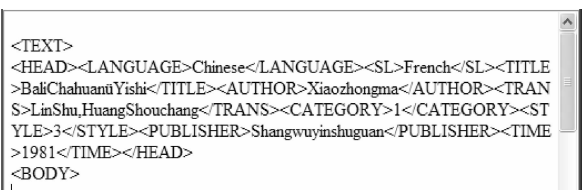


图 4 文本头元信息标注样本

Fig. 4 Head annotation sample

够迅速进行提取,展开相应的统计分析,这是基于语料库基础上进行研究的便捷之处。

## 四、创建林纾译文语料库的意义

林纾译文语料库的创建,将为翻译学界提供一种新的“解读林纾”的手段与方法,也将为翻译

学界提供全新的研究成果,为全方位认识与评价林纾提供客观、科学的数据。这将在描写翻译学研究领域内提供一个可供参照的研究范式,为翻译名家研究、翻译风格研究、翻译语言特征分析、翻译策略研究等多方面提供新的视角。

林纾作为福建工程学院的前身——“苍霞精舍”的创办者,为该大学留下了丰富的遗产。创建林纾译文语料库,是开展相关学科建设、大学校园特色文化建设的组成部分。该语料库的创建,将为人文社科与自然科学(如自然语言处理、信息技术等)研究团队的培养提供良好的合作平台。林纾译文语料库的创建,将在相关学术交流、文化品牌创建等方面发挥应有的作用,也将为进一步发掘林纾文化宝库资源做出一定的贡献。

## 参考文献:

- [1] 吴仁华. 传承林纾文化思想 打造林纾研究重镇——《林纾研究专刊》代序[J]. 福建工程学院学报, 2012, 10(5).
- [2] 张俊才. 林纾评传[M]. 北京: 中华书局, 2007: 249 - 250.
- [3] 苏建新. 真诚勤勇的楷模林纾完美人格论略[J]. 福建工程学院学报, 2008, 6(2): 200 - 204.
- [4] 林农. 林纾翻译思想[J]. 福建工程学院学报, 2012, 10(5): 446 - 450.
- [5] 钱钟书. 林纾的翻译[C]//罗新璋. 翻译论集. 北京: 商务印书馆, 1984: 696 - 725.
- [6] Holmes J S. The Name and Nature of Translation Studies[C]// L Venuti (Ed). The Translation Studies Reader. London & New York: Routledge, 1972/2000: 172 - 185.
- [7] Toury G. Descriptive Translation Studies and Beyond[M]. Amsterdam & Philadelphia: John Benjamins Publishing Company, 1995.
- [8] 王克非. 语料库翻译学探索[M]. 上海: 上海交通大学出版社, 2012.
- [9] 肖忠华, 戴光荣. 翻译教学与研究的新框架: 语料库翻译学综述[J]. 外语教学理论与实践, 2011(1): 8 - 15.
- [10] 戴光荣. 译文源语透过效应研究[M]. 上海: 上海交通大学出版社, 2013.
- [11] Baker M. Corpora in Translation Studies: An Overview and Some Suggestions for Future Research[J]. Target, 1995, 7(2): 223 - 243.
- [12] 王克非. 双语对应语料库研制与应用[M]. 北京: 外语教学与研究出版社, 2004.
- [13] 黄立波. 基于双语平行语料库的翻译文体学探讨[J]. 中国外语, 2011, 8(6): 100 - 106.
- [14] Baker M. Towards A Methodology for Investigating the Style of A Literary Translator[J]. Target, 2000, 12(2): 241 - 266.
- [15] Baker M. The treatment of variation in corpus-based translation studies[J]. Language Matters, 2004, 35(1): 28 - 38.
- [16] Venuti L. The Translator's Invisibility: A History of Translation[M]. London and New York: Routledge, 1995.
- [17] Hermans T. The Translators Voice in Translated Narrative[J]. Target, 1996, 8(1): 23 - 48.
- [18] 左尚君, 戴光荣. 基于 Wmatrix 的《红楼梦》两英译本文体对比研究[J]. 嘉兴学院学报, 2014, 26(1): 60 - 66.
- [19] 陈虹. 高校图书馆林纾文化特藏资源概况与展望——以福建工程学院为例[J]. 长春师范学院学报: 自然科学版, 2012, 31(3): 182 - 185.

(责任编辑: 许秀清)