

旅游翻译语料库的构建及应用

——以浙江省旅游文化汉英翻译平行语料库为例

余 姿

(中国计量大学 外国语学院, 浙江 杭州 310018)

摘 要: 本文从语料库的设计、语料的选取与采集、语料整理、优化搜索、后期维护等方面, 深入探讨了浙江省旅游文化汉英翻译平行语料库的构建, 并从旅游文本翻译、翻译教学和翻译实践三方面, 分析了浙江省旅游文化翻译平行语料库的应用。

关键词: 旅游翻译; 语料库; 构建; 应用

中图分类号: H315.9

文献标识码: A

文章编号: 1008-9640 (2017) 08-0043-02

21 世纪翻译研究领域的一大进展就是语料库翻译学范式的建立、研究和发展。当前, 语料库以一种前所未有的技术呈现语言运用事实, 成为一种真实可靠的资源及信息检索和处理平台, 为研究者提供了大量真实的语言数据库, 使语言分析研究能够基于自然发生的语言结构和形式, 而不是靠自觉感知的语言现象展开。基于语料库的旅游文本翻译的最大优势在于从语言现象入手, 使用真实的旅游语料描述语言特征, 通过大量语料从纵向找寻重复出现的语言形式, 更加客观准确地揭示本质, 提高对语言系统的洞察力, 从而为翻译实践服务。

一、浙江省旅游文化汉英翻译平行语料库的构建

为直观地研究浙江省旅游文本的翻译现象, 笔者研制了一个小型语料库——浙江省旅游文化汉英翻译平行语料库, 容量约 15~20 万词, 主要为翻译研究、翻译教学和翻译实践。

(一) 语料库设计

此语料库的设计基于书面文本和口语副本两方面。这种设计理念是基于当代语料库研究领域的两种不同观点: 一种认为在某些特定方面, 口语形态的文本更能反映出翻译最基本的组织形式; 另一种更倾向于文本的翻译模式。在构建浙江省旅游文本语料库时, 笔者将两种不同模式的文本融合在一起, 从而增强了语料库的多样性, 更加符合翻译需求。运用语料库专业软件 AntCont 作为语料建设平台, 将所有搜集和整理好的语料原始文件批量转换和输入到 AntCont 中。为更好地管理和搜索源文件, 所有的文件将以“城市+旅游景点+日期”的格式储存在系统中。例如, “西湖十景”是以“杭州+平湖秋月+2017”的格式储存十个景点的翻译语料。

(二) 语料的选取与采集

因语料的收集范围直接影响语料的质量, 因此, 本文采用以下几个标准来确定语料收集的范围, 以保证语料的代表性、平衡性和权威性。

第一, 语料内容为与浙江省相关的旅游文化材料。为保证语料的权威性和代表性, 本文选取了浙江省 4A 和 5A

级旅游景区的资料作为建库语料, 且以拥有有效中英文网页链接的旅游景区内容介绍为主要选择, 包括景点介绍和推广材料、宣传文化传统和习俗的资料等中英文双语语料。此外, 在选取过程中还应侧重语料的可比性和多样性, 如关于三潭印月的翻译文本, 该语料搜集了来自不同时期、不同出版物, 甚至来自不同英语国家共八项翻译素材。

第二, 正式出版或非正式出版的语料分别入库。正式出版物如书籍、杂志、音像资料等, 非正式出版物如介绍旅游文化景点的网页、各类旅游文化宣传资料, 包括酒店介绍、旅行社传单、景点公示牌上的中英文介绍、旅游公示语等。

第三, 语料载体不限。选取的语料可以是书籍、杂志、网页、旅游小册子, 包括旅游景观介绍、旅游广告、景点中英文招牌、民俗风情介绍、饮食文化介绍、人名地名介绍、历史文化介绍、物产资源等。

(三) 语料整理

收集相应的语料后, 对采集的原始语料进行了初步整理和归类, 纸制材料需经过扫描统一转为 Word 格式。语料的切分采用人工切分对齐和软件切分对齐的方法, 软件采用英文自动分句工具 Multiconcord 进行自动分句对齐处理。中英文由于句式、结构、标点符号及语篇的差异, 在实际翻译中无法以句子为单位进行一对一翻译, 目前软件所能完成的自动句子对齐处理只占 50%~60%, 其余仍须手工对齐和切分。为统一标准, 该语料库按照以句子为单位进行切分和对齐的原则, 即无论是一句原文对应一句译文, 还是一句原文对应数句译文, 都按照一句处理, 具体来说, 简短的公示语按照句子切分和对齐, 人名、地名、旅游景点名称、旅游设施、旅游活动等的翻译按照词组对齐的方法处理。所有资料都需经过清洁标注后方可入库, 且对入库的中英文语料分别存放, 以避免中英文混排而出现乱码现象。

(四) 标注

对采集的全部翻译语料和素材进行标注和细分是语料库设计和开发中必不可少的一步。在语料库中, 主要对文本头和词性进行标注, 如译者国籍、发表时间、性

收稿日期: 2017-08-03

基金项目: 2015 年浙江省旅游科学研究项目(项目编号: 2015YB04)。

作者简介: 余姿(1963—), 女, 浙江义乌人, 中国计量大学外国语学院副教授。研究方向: 语言学及应用语言学、翻译学等。

别、语料出处、关键词等。另外,制定标注规范,抽取部分典型汉英对应语料进行错误标注和分类,并标注赋码与流程,实施严格的标注质量监控。

(五) 优化搜索

设计和运用一些必要的通配符可以一定程度上提升语料库的运用效果。比如,英文大小写和“*”符号的运用。在英文翻译实际练习中,介词的运用十分频繁。例如,在疑问 but 是否可以放句首时,将第一个字母 B 大写,从而使用设计好的语料库来搜索文本资料,并基于搜索到的文本资料分析和判断 but 的用法。符号“*”在语料库软件中是一个通配符,如“We * or”或者“We * * or”,符号“*”代表可能出现在 We 和 or 之中的一个或者两个词语。运用通配符在语料库中进行搜索,可以将词语出现的频繁程度进行排列,从而使不同形态的语言特征可以很好地分布在搜索结果中,有助于旅游文本的翻译。

(六) 后期维护

初步建成的语料库会随着新语料的添加和补充而不断更新和完善,容量会日益增大,语料库规模也会逐渐扩大。基于旅游文本材料的更新速度,本语料库的建立基于静态与动态并存的模式。静态语料库的模式适用于一些较成熟的景点且拥有大量的翻译文本,然而,对于一些新建景点和现存较少的翻译文本资料景点,更加适用动态的语料库模式,并且需要在一定时间内进行文本搜集和语料库更新。

二、浙江省旅游文化翻译平行语料库的应用

(一) 旅游文本翻译研究

翻译语料库能够为研究者提供大量丰富、真实的素材,通过对大量实例的观摩分析,用定量和定性分析相结合的方法,探索浙江旅游文本翻译的特点、规律。通过检索语料库,研究者还可以分析旅游文本翻译在词句搭配、修辞、句型和语篇结构等方面的特色和语言特征,考察原文中某些语言现象的翻译规律,比较不同的翻译形式、语言风格和模式,总结译者对不同译文所采取的翻译策略,从而提高语际转换能力。例如,借助 Wordsmith 5.0 和 Antconc,笔者统计了语料库中公示语翻译策略的分布情况(见图1)。由图1可知,基本上各景点公示语英译时首先考虑的是直译;较长句子的翻译最常用的策略是先拆分原文,再翻译成多个英文句子,且采用合译策略的公示语也不少;景点名称翻译多采用音译加直译的策略;历史典故或者历史事件的翻译基本都选择了漏译或者不译,有些采用了省译;在涉及中国文化特色时,部分景点采用了直译加注的增译策略,有些则采用音译的策略。

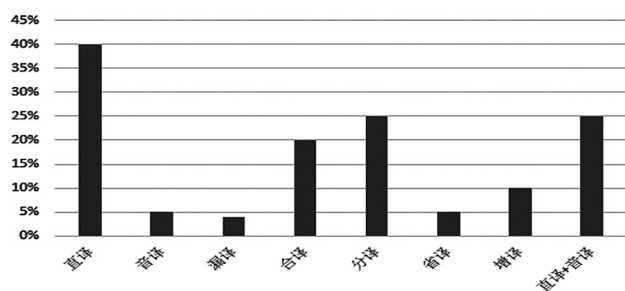


图1 旅游公示语翻译策略分布统计

(二) 翻译教学研究

语料库在翻译教学应用中最突出的特点是其与 Paraconc 和 Wordsmith 定位搜索软件结合之后所具备的功能,如词语互译、结构互译或语句语篇互译的丰富例证

等,这种语料的呈现能客观而有效地提升翻译课堂参与者的积极性。例如,当翻译包含“费用”的汉语句子时,学生使用最多的英语对应词是 cost,但利用平行语料库检索软件 Paracon,以“费用”作为检索项检索自建平行语料库,得到如下结果(随机抽取几例)。

例1:我要仔细核算一下旅行费用。

英译: I need to cost my trip more carefully.

例2:他必须打零工补贴部分大学费用。

英译: He has to do part-time job to pay for part of his college expenses.

例3:请告诉我们你们将收取多少服务费用。

英译: please tell us what your service charges will be.

此外,该词对应的英译还有 rate、fare、cost、payment、fee、charge、bill、expense 等。相比普通汉英词典来说,语料库能提供更为丰富和翔实的语料、例证、参考等,从而帮助译者做出恰当抉择。在翻译教学中使用语料库既能方便教师备课和学习者查询资料,又能消除传统翻译教学中译例单一、缺乏真实性的弊端,其能让学习者比较同一结构或类似结构在不同体裁文本中的翻译技巧,有助于提升翻译教学的时效性。

(三) 翻译实践研究

本语料库及其配套的检索翻译平台能方便研究者全面检索旅游类术语及常用的词语和语篇,并支持译本中的术语自动识别及翻译提示,极大提高旅游文本翻译效率。对译者翻译实践而言,语料库提供的信息量大、例证丰富,译者有较大的选择性和较强的自主性,且可以方便提取和编辑语料,能极大提高译者的翻译质量。例如,借助 Antconc 软件,笔者统计分析了语料库中的主要实词,发现浙江旅游文本中的名词密度较大、动词丰富、形容词单调,而其外置结构、被动句型和“there be”句型比重大,说明文本的客观性较强;具体旅游文化文本均具有一定的特点和行业性,其中的字、词、句都具有其旅游文本的指称意义和文化涵义,相应的检索结果能为中英文自然语言识别和理解提供清晰的诠释和强有力的支撑,也为理解两种文本、确立概念对应和翻译提供可靠的概率数据,为进一步翻译研究和文本分析奠定基础。

三、结语

基于语料库的翻译研究,最大优势在于其能为翻译提供真实的语言事实和参考依据。从语言现象入手,可客观准确地揭示旅游文本的本质。语料库为旅游文本翻译提供了坚实的数据分析和理论基础,也为翻译研究提供了广泛的语料对比标准,语料库的使用可以有效且直观地解析作者的翻译策略和语言规范。此外,翻译者也可以从语料库的素材中学习不同语言的应用策略和不同文化的特性,从而加强语言的丰富性,最终提高文本翻译质量。

参考文献

- [1] 黎士旺. 语料库与翻译教学[J]. 中国科技翻译, 2007, (3): 46-49.
- [2] 梁茂成等. 大型语料库的合理使用与共建共享[J]. 现代外语, 2010, (4): 425-426.
- [3] 王克非. 对应语料库翻译教学平台的应用初探[J]. 外语电话教学, 2007, (12): 3-8.
- [4] 帅莹. 东文化旅游资源汉英平行语料库的建设与应用[J]. 文化月刊, 2017, (4): 90-92.
- [5] 赵佰儒. 旅游资源翻译语料库的构建及其应用研究[J]. 现代语文, 2017, (3): 154-155.

(责任编辑:李汴红)