

● 乌蒙论坛

昭通旅游翻译双语平行语料库的构建

闵西鸿

(昭通学院 外国语学院, 云南 昭通 657000)

摘要: 旅游文本具有明显的地域性和专门性, 通用的旅游翻译语料库不能完全适用于昭通地区的旅游翻译, 根据地区现状, 通过总体设计、语料收集整理、标注对齐、语料整合几个步骤建设昭通旅游翻译双语平行语料库, 有利于昭通旅游业的长远发展。

关键词: 昭通; 旅游翻译; 双语平行语料库

中图分类号: H315.9

文献标志码: A

文章编号: 2095-7408(2015)04-0019-04

一、引言

云南是全国旅游资源最丰富的省, 而素有“小昆明”之称的昭通市地处川滇黔三省结合处, 是云南文化三大发源地之一, 是中国著名的“南丝绸之路”的要冲, 素有“锁钥南滇, 咽喉西蜀”之称。这片古老而神秘的净土, 有着丰富的旅游资源、源远流长的历史文化、古朴浓郁的民族风情, 是云南旅游一块待开发的宝地, 具有强劲的后发优势。2012年, 这片“养在深闺”的土地以“磅礴乌蒙 神奇昭通”的形象在全国旅游推介会亮相, 并在中央电视台宣传。2013年, 央视栏目《远方的家》百集系列特别节目《百山百川行》之《磅礴乌蒙在昭通》更为昭通旅游业的快速发展埋下伏笔。而此时, 着手建立一个基于昭通旅游翻译的电子双语平行语料库对昭通旅游的长远发展是十分必要的。

二、建设昭通旅游翻译双语平行语料库的必要性

(一) 旅游文本分类和特点

旅游文本按其功能不同可划分为八类, 分别是: 宾馆饭店、交通指南、城市概况、地方文化、购物美食、景点介绍、旅游服务和景区公示语, 它们是旅游对外宣传材料的基本组成部分, 为游客提供各种类型的信息, 并能起到吸引游客的作用。由于各地的地理、人文环境迥异, 旅游资源也各不

相同, 旅游业具有较强的地域特色, 旅游文本也因此具有较强的地域性和专业性, 翻译时应注重其语言功能。

英国翻译理论家纽马克(Peter Newmark)将文本的语言功能划分为三大类型: 表达型(expressive)、信息型(informative)、呼唤型(vocative)。表达型文本注重作者情感的主观表达, 如日记、小说等; 信息型文本注重信息的客观传递, 如新闻、纪录片等; 呼唤型文本注重对受众产生的反应和影响, 如广告、宣传材料等。很多情况下, 一个文本都同时具有多种功能。以旅游文本为例, 其根本目的是要吸引游客到景区游览, 具有较强的呼唤功能, 其次还要向游客说明如何到景区, 到景区后的吃、住、行、玩等具体细节, 起到传递信息作用即信息功能, 所以旅游文本同时具有信息和呼唤双重功能。翻译旅游文本时, 应遵循交际翻译原则, 注重文本的呼唤功能, 同时兼顾其信息功能。

(二) 昭通旅游翻译现状

目前昭通旅游相关的文字材料主要为中文资料, 一部分包含在云南省旅游宣传资料内, 如云南省旅游局编写的《云南景区景点导游》、《走遍彩云南》等, 一部分是专门针对昭通景区的宣传材料, 如《昭通龙氏家祠》、《明月万里照关山: 大关的历史文化解读》等, 还有部分未公开出版的内部资

收稿日期: 2015-03-09

基金项目: 2013年度全国高校外语教学科研项目(云-0001-A)。

作者简介: 闵西鸿(1979—), 女, 重庆永川人, 讲师, 硕士, 主要从事翻译理论与实践研究。

料。关于昭通旅游的英文材料却不太多,云南省旅游宣传的全英文资料“YUNNAN FORGING AHEAD”中只用了 3 页的篇幅对昭通作了一个简略的介绍。2007 年出版的中英对照《黑颈鹤:孙德超黑颈鹤生态摄影》专门介绍了昭通大山包黑颈鹤自然保护区,而对昭通景区景点比较全面的英文介绍主要集中于 2009 年出版的中英双语对照的《昭通旅游导游词》。

《昭通旅游导游词》由昭通市委、市政府组织全市旅游系统编写,该书全面介绍了昭通辖区内的景点,并由云南大学外国语学院教师完成了对应的英译工作,目前昭通各景区景点的英文介绍主要来源于此。该书的出版不仅为当地涉外旅游提供了参考,也为翻译研究者提供了研究素材,围绕该书已有两个科研项目:昭通市社科联课题《功能翻译与语用学视角下的〈昭通旅游导游词〉》、昭通学院校级科研课题《昭通市辖主要旅游景区介绍英语文本的语用考查与研究》。笔者有幸参与了两个项目的研究工作,在此期间,项目组成员发现了该书的一些不足之处。首先,英译文中出现了一些错误或称误译,如将“威信县”的拼音误拼为“Wenxin County”,“会理”误拼为“Hu-li”^{[1](P. 204)},如果说此类错误是由于编辑的疏忽造成的,那么将“眉山”误译为“Mei Mountain”^{[1](P. 209)}，“岔河”误译为“Cha River”^{[1](P. 213)}就是译者不可推卸的责任了,因为原文语境中所称“眉山”非山,“岔河”非河,而是两个地名,根据国家颁布的《关于改用汉语拼音方案作为我国人名地名罗马字母拼写法的统一规范的报告》,正确的译法应为“Meishan”、“Chahe”。其次,因为该书篇幅较长,由多人分组翻译,后期审译工作量大,疏漏在所难免,出现了译文前后不统一的现象,普通名词的译文不统一通常不会造成理解上的困难,但专有名词的译文不统一却容易产生误解。以“五尺道”为例,出现的译文有“Five - chi Path”^{[1](P. 212)} 和 “Wuchidao” (Five Chi Path)^{[1](P. 386)},如果读者是对中文一无所知的外国游客,他可能会感觉有些困惑:这是同一个地方还是两个不同的地方? 该书中有大量的专有名词,译文不统一的现象多次出现,这些专有名词译文的不统一不仅影响到翻译的权威性,也容易引起读者误解。

该书的另一不足之处是检索不便。在项目研究的过程中,项目组成员为此深感苦恼,由于笔者并非昭通本地人,对昭通的地理环境比较陌生,对于一些小景点,如果不知道它们所在的大地名,便无从检索。例如在整理独立的景点名称译文时遇到“六井仙人洞”译为“Xianren Cave in Liujin”,对“六井”译为拼音有所质疑,打算在书中查证,但苦于不知道其出现章节,还得花时间向他人请教出处,颇费周折。在信息技术高速发展的今天,信息技术广泛应用于各行各业,将信息技术应用于旅游业能够大大提高资源利用率,电子语料库也因此应运而生。

(三) 昭通旅游翻译语料库的针对性

语料库是为一个或多个应用目标而专门收集的、有一定结构的、有代表性的、可被计算机程序检索的、具有一定规模的语料的集合。^[2]基于计算机的电子语料库兴于 20 世纪 60 年代,经过近五十年的发展,语料库种类越来越多,分类也越来越细。语料库有多种分类方法,按语料库中语种数量分类,可分为单语语料库、双语语料库和多语种语料库;按其语种对应关系,可分为类比语料库和平行语料库;按其用途,可分为通用语料库和专门语料库。近十几年来,不少学者把目光投向双语平行语料库的研究和开发,双语平行语料库采用原文与译文本对齐的方法,实现词汇、短句、句子、段落、篇章等层面的双语自动对齐,用户可以根据自己的需要精确地检索各层面的对应翻译,具有很强的实用性和针对性,对翻译、翻译教学和翻译研究都非常实用。到目前为此,国内针对旅游业的语料库以香港理工大学李德超教授研发的新型中英双语旅游语料库容量最大,该语料库分为双语旅游翻译对应语料库和双语旅游翻译类比语料库,容量各为 100 万字/词。^[3]但由于该语料库主要收录香港地区的英译汉和汉译英的旅游翻译文本,对其它地区的旅游翻译并不完全适用,所以建设一个针对昭通地区的旅游翻译双语平行语料库是非常必要的。

三、昭通旅游翻译双语平行语料库的构建

(一) 语料库总体设计

语料库的设计包括语料模块的设计和系统功能设计。整个设计应遵循三个原则:针对性、实用

性和扩展性。针对性主要指语料模块分类应尽量针对不同用户的需要,昭通旅游翻译语料库主要针对的用户是政府旅游管理部门、旅游翻译从业人员、旅游翻译教学人员和境外游客,在设计语料模块时,应充分考虑这几类人员对旅游翻译语料的分类需求,并使分类具有一定的灵活性和交互性。实用性和扩展性主要针对语料库的系统功能。实用性主要体现在为用户提供便捷、精确的检索功能,用户根据自己的需要可通过词语、句子等关键词对语料库进行分类检索或者全库检索。扩展性主要指语料库中语料的扩展和系统功能的扩展。语料的扩展是指管理员可对语料进行增加、删除、修改等操作,并能实时更新到语料库中。系统功能的扩展主要指系统设计时应预留接口,以便在长远发展过程中,需要添加新的功能模块。

(二)语料收集和整理

语料收集包括中文语料收集和英译文语料收集,来源以互联网和电子文库语料最为便捷,但在互联网上搜索昭通地区的旅游语料,结果基本都是中文语料,英文语料非常少,更不要说对应的英译文语料了。就互联网上的中文语料,收集过程中要注意其准确性和权威性,昭通市旅游局政府网站昭通旅游网是比较可靠的来源,但该网站所列景区景点只有11个,还有很多景点都没有列出,所以有很大部分的语料还需要从纸质文本转换为电子文本。纸质文本以昭通市委、市政府组织编写的《昭通旅游导游词》最为全面、最具权威性。该书对昭通市的自然和人文旅游资源进行了全面翔实的收集整理,挖掘了各地悠远的人文历史和多彩的民风民俗,而且是中英文双语对照,是收集旅游翻译语料的重要来源,但仍然需要重新梳理,特别是英译文部分,我们应在充分理解原文的基础上,找准原文的文化信息和交际意图,本着交际翻译的原则,对其中不恰当的译文进行修改。由于旅游文本中人名、地名、景点名等专有名词很多,翻译时要特别注意其统一性。对于其它没有英译文的纯中文的语料,翻译时也要充分理解原文的交际意图,实现其呼唤功能和信息功能。同时,在语料的收集整理过程中,所有电子文本格式需要统一为TXT或WORD文档,以便后继工作。

(三)语料标注和对齐

语料的标注在这里指词性标注,又称词性赋码,是指对语料中的每一个字词按其在句子中的语法功能加注词性标记,如单数普通名词、形容词的比较级、动词的过去分词等。^[4]语料的标注和对齐是建好语料库的关键,它直接影响到语料库建成之后检索功能的精确性和实用性。标注和对齐可通过人工和软件两种方式完成,人工方式过于耗时耗力,所以可以采用人工和软件相结合的方式,先用软件进行标注和对齐,然后辅以人工方式校对。从成本角度考虑,有两款免费的词性标注软件比较经济适用,一是德国斯图加特大学计算机语言学研究所研发的Tree Tagger赋码器,另一款是英国兰卡斯特大学计算机语料库研究中心UCREL研发的CLAWS赋码器免费在线试用服务。语料对齐以句子为单位实现汉英对照。由于中英文句法、表达方式等不同,有时候一句汉语需要分译为几句英文,而有时候几个汉语句子又会合译为一句英文,对于这种情况,就不能按照句子一对一的方式对齐,而应采用句意对应的方式对齐,也就是说会出现一个句子对应几个句子的情况,这样可以保证翻译的完整性,也能体现出翻译技巧,对翻译教学和研究有一定的借鉴意义。另外,旅游文本中专有名词较多,所以对人名、地名、景点名、展品名、特产名等专有名词可按词组方式对齐,便于检索,检索结果可附带显示该词出现过的句子,以便用户进一步了解。

(四)语料库的整合

语料标注对齐后,就需要用语料库软件将所有语料按一定方式整合起来,实现检索功能。由于软件涉及计算机网络、数据库技术层面的问题,而语料库软件又属于小众型应用软件,所以现成的软件并不多。现成的软件可以使用香港科技大学开发的Compulang WordPilot,由于具备强大的检索功能和软件自身开放性的特点,可以将它作为小型语料库的建库工具使用。^[5]也可以使用ParaConc等语料库检索软件,但这类软件的免费版功能都有所受限,在资金允许的条件下,可以请专业的计算机软件人员针对语料库开发新的软件。语料整合完成后还需经过网络测试,测试成功后便可放在相关网站上对用户开放。

四、结语

昭通的旅游资源是昭通经济的重要组成部分

分,建设昭通旅游翻译双语平行语料库不仅能向旅游管理部门和旅游从业人员提供实用的参考,也能够为昭通旅游业营造良好的文字环境和文化氛围,提升昭通旅游的对外形象,为吸引更多境外

游客起到积极促进作用。虽然语料库的建设需要大量的人力、物力、财力,但从昭通旅游业的长远发展来看,投资建设昭通旅游翻译双语平行语料库具有巨大的应用价值。

参考文献:

- [1]饶祥碧. 昭通旅游导游词[M]. 昆明:云南大学出版社,2009.
- [2]KARIN A, BENGT A. 语料库语言学的进展[M]. 北京:世界图书出版公司,2009:序.
- [3]李德超,王克非. 新型双语旅游语料库的研制和应用[J]. 现代外语,2010,(1):46—54.
- [4]余国良. 语料库语言学的研究与应用[M]. 成都:四川大学出版社,2009:44.
- [5]梁茂成. 利用 WordPilot 在外语教学中自建小型语料库[J]. 外语电化教学,2003,(6):42—45.

On Establishment of Bilingual Parallel Corpus of Zhaotong Tourism

MIN Xi-hong

(School of Foreign Languages, Zhaotong University, Zhaotong 657000, China)

Abstract: As Tourism texts are regionalized and specialized, general tourism translation corpora do not apply to zhaotong tourism. Based on regional situation, a bilingual parallel corpus can be established by following the steps of designing, text collecting, marking, integrating, which is conducive to the long-term development of zhaotong tourism.

Key words: Zhaotong; tourism translation; bilingual parallel corpus