

旅游文本《找个理由Go无锡》机译英文版偏误分析

王晓蕊 库莹莹

江西水利电力大学, 中国·江西 南昌·330099

摘要: 人机交互的翻译的模式是当下主流的翻译模式, 机器翻译给译员带来了便利, 同时译文也存在一定程度的偏误。本文以旅游文本《找个理由 Go 无锡》为研究对象, 在 MQM (Multidimensional Quality Metrics) 质量评估框架基础上对三个翻译引擎 (DeepL Translator、百度翻译、有道翻译) 的机译译文进行质量测评、错误值统计。分析表明, 机器翻译在习惯用语、文化负载词和成语上均存在偏误, 由此, 本文提出对原文进行译前编辑, 建立语料库以及加强译后编辑等机译偏误的解决办法。

关键词: 机器翻译; 旅游文本; 机器翻译质量测评; 译前编辑; 语料库; 译后编辑

Analysis of the error of the travel text "Find a Reason for Go Wuxi" in the machine translation of the English version

Wang Xiaorui, Ku Yingying

Jiangxi University of Water Resources and Electric Power, China Jiangxi Nanchang 330099

Abstract: The translation model of human-computer interaction is currently the mainstream translation model. Machine translation has brought convenience to translators, but there are also certain degrees of errors in the translated texts. This paper takes the tourism text "Reasons for Staying in Wuxi" as the research object and conducts a quality evaluation and error value statistics of the machine-translated texts from three translation engines (DeepL Translator, Baidu Translate, and Youdao Translate) based on the MQM (Multidimensional Quality Metrics) quality assessment framework. The analysis shows that machine translation has errors in idiomatic expressions, culture-loaded words, and idioms. Therefore, this paper proposes solutions to machine translation errors, such as pre-translation editing of the original text, establishment of a corpus, and strengthening of post-translation editing.

Keywords: Machine translation; Tourism text; Machine translation Quality assessment; Pre-editing; Corpus; Post-editing

0 引言

机器翻译是指利用计算机将一种自然语言转换为另一种自然语言的过程。在全球化竞争日益激烈的今天, 高质量的机器翻译对于提升工作效率、降低沟通成本至关重要。时至今日, 机器翻译在翻译行业中发挥着日益显著的作用, 已然成为专业翻译人员日常工作中不可或缺的一部分 (朱慧芬, 2019)。尽管机器翻译技术在众多领域取得重大进步, 但仍面临诸多挑战。例如, 在翻译文学色彩较强的旅游文本时, 对习惯用语、文化负载词、成语以及语句连贯性和流畅性的处理上, 机器翻译所输出的译文质量还有很大的提高空间。

Reiss 根据 Karl Buhler 的语言功能划分。将文本划分为四种类型: 信息型、表情型、操作型和依靠各类非语言视听技术媒介手段的视听型。旅游文本属于信息型文本、表情文本和召唤型文本的结合体, 在同时具有信息功能和召唤功能的基础上也兼具一定的审美和艺术功能。《找个理由 Go 无锡》为旅游网站新闻翻译, 属新闻翻译类别, 原

文带有明显的文学化倾向, 但属于实用文本翻译。翻译旅游文本相较于其他文本而言难度更大, 需要译者有一定的专业领域知识储备和严谨的翻译态度。

本研究以《找个理由 Go 无锡》旅游文本为研究样本, 通过对比三个机器翻译引擎的机译译本与相对标准的译文, 进行旅游文本机器翻译质量测评, 采用基于 MTI 教学实践优化的 MQM (Multidimensional Quality Metrics) 框架, 从术语 (Terminology)、准确性 (Accuracy)、语言惯例 (Linguistic Conventions)、风格 (Style)、区域惯例 (Locale Conventions)、受众适应性 (Audience appropriateness) 和格式与标记 (Design and Markup) 等 7 大维度对译文进行翻译质量评估, 最终通过数据聚合生成各平台译文的量化评估结果, 并提出相应的解决办法。

1 旅游文本机器翻译质量测评

1.1 MQM 模型介绍及评估标准

德国人工智能研究中心发布的 MQM 模型是一个系统化的框架, 评估指标 (Metric) 处于核心地位。该模型

在顶层结构上将问题划分为 8 个主要的维度,即准确性 (Accuracy) 流利度 (Fluency)、术语 (Terminology)、区域惯例 (Locale-convention) 风格 (Style)、真实性 (Verity)、格式 (Design) 和国际化 (Internationalization)。本研究以以上 7 个维度进行评估,其中涵盖 21 个错误类型,覆盖面广,适用于全方位测评机器翻译的译文质量。

1.2 旅游文本翻译质量测评结果

本研究对比分析了百度翻译、有道翻译和 DeepL 三款机器翻译引擎的译文质量。研究邀请 3 位专业评审对 3 份样本在三个平台生成的译文 (共 9 份) 进行测评。由于评审人员对翻译错误的判断标准、主观经验及对机器翻译的接受度存在个体差异,可能导致其对错误数量和严重程度评分不一致。

基于前期构建的 MQM 多维度质量评估表格,通过仔细分析最终得出各平台罚分总计 (Penalty Totals) 为 690 分,百度翻译因在术语、准确性以及语言习惯方面失分较多,累计罚分 690 分;有道翻译官在准确性和语言习惯上表现较好,罚分降至 365 分;而 DeepL Translator 在处理文本时展现出显著优势,在术语和语言习惯上表现最好,以 300 分的最低罚分脱颖而出。

分析表明,当前机器翻译样本存在较多误译现象,整体翻译质量评分未能达到预设阈值,后续需投入大量精力进行人工校对。从七个评估维度分析,各类错误分布于零错误、轻微错误、较大错误以及严重错误这四个等级之中。研究结果显示,“术语 / 专有名词”“准确性”与“风格”这三个维度所对应的罚分值显著偏高。

2 偏误案例分析

2.1 偏误类型

虽然机器翻译研究已取得显著进展,但必须承认的是,当前机器翻译的产出仍存在诸多不规范和不准的缺陷。崔启亮 (2014) 认为,整体来说机器翻译的质量仍然无法与人工翻译相提并论。本文选取的研究对象为旅游翻译文本《找个理由 Go 无锡》,文中涉及旅游观光、景点介绍、风俗与文化、交通、餐饮、购物及教育、文化、休闲等诸多方面。本文通过机译版本和官方译文对比分析,笔者发现,机器翻译在习惯语、文化负载词和成语层面均存在一定的问题。

2.2 案例分析

2.2.1 习惯语的偏误

《找个理由 Go 无锡》的机译版本在词汇层面的问题,主要体现在文化负载词的转换上。文化负载词涉及生态文

化、物质文化、社会文化、宗教文化、语言文化等多个文化维度。机器翻译在处理具有中国特色的表达时,往往难以准确传达原文的文化意蕴和信息内容,从而影响了译文的阅读体验。

例 1: Tips: 0510-85689010 无锡市滨湖区马山镇灵山路 1 号 88 路、89 路。

DeepL 翻译: Tips: 0510-85689010 No.1 Lingshan Road, Mashan Town, Binhu District, Wuxi City 88 Road, 89 Road.

百度翻译: Tips: 0510-85689010, No. 1 Lingshan Road, Ma Shan Town, Binhu District, Wuxi City, Route 88 and 89.

有道翻译: Metro Line 1, No. 88, No. 2, No. 158, No. 207, No. 13, Bian Jiawan, Liangxi West Road, Binhu District, Wuxi.

译文: Tips: 0510-85689010 No. 1 Lingshan Road, Mashan Town, Binhu District, Wuxi City Bus No. 88, Bus No. 89.

例 1 关于“路”的翻译,三个机译软件翻译的各不相同,分别译为“Road”“Route”和“No.”。实际上本句中“路”的含义是公交车线路,DeepL 和百度的处理方式相似,直译成 Road 和 Route,读者会误将此处理解为马路或者路线,这两种翻译虽然语法上没有问题,但在实际语境中可能会引发误解,因为 Road 通常指的是道路,而 Route 虽然可以表示路线,但在没有上下文的情况下也可能被理解为航空或铁路等路线的泛指。而有道翻译的处理方式出现了明显的错误,它将“88 路、89 路”与一系列地铁线路和地名混淆在一起,这显然与原文不符。译文中“Bus No. 88, Bus No. 89”则准确地传达了原文的含义,清晰地指出了这是两条公交车线路。这种翻译方式既符合英语中表达公交车线路的常用方式,又能够确保读者正确理解原文信息。

2.2.2 文化负载词的偏误

例 2: 这里的山水很江南

DeepL 翻译: The landscape here is very Jiangnan

百度翻译: The scenery here is very Jiangnan

有道翻译: The landscape here is very southern

译文: Views are gentle

在探讨“江南”这一地理文化概念在翻译中的处理时,我们需要首先明确“江南”在中文语境下的具体含义和它所承载的文化内涵。在汉语地理文化语境中,“江南”一词通常指代长江中下游南岸区域,尤以苏南、浙北地区

为核心范围。该地域凭借其典型的水乡泽国生态景观、绵延千年的文化积淀以及独具特色的地域人文品格,在中国区域文化研究中占据重要地位。

例 2 中,“这里的山水很江南”是一个简洁而富有文化意涵的表达,旨在传达无锡地区山水风光的细致温润特点,同时也暗示了其地理位置处于“江南”这一文化区域内。而关于“江南”的翻译,DeepL 和百度翻译两者都将“江南”直译为 Jiangnan。这种翻译方式在形式上保留了原词的音韵和拼写,但在缺乏上下文和解释的情况下,对于不熟悉中国地理文化的读者以及国际读者来说,可能难以准确理解其含义。有道翻译将“江南”译为“southern”。虽然这一翻译在地理上大致正确,指出了长江以南的地理位置,但它忽略了“江南”作为一个具有丰富文化内涵的地理名词的特殊性。“southern”一词过于宽泛,无法准确传达“江南”所承载的历史、文化和情感价值。译文中“Views are gentle”则保留了原文的文化内涵,由于上下文内容已经表明了无锡位于长江以南的地理位置,同时又通过“gentle”一词强调了山水风光的美丽和温润,标题的简短表述也更准确地传达原文的意图。

2.2.3 成语的偏误

例 3: 有湖的地方,就有烟波浩渺。

DeepL 翻译: Where there's a lake, there's smoke and mirrors.

百度翻译: Where there is a lake, there is a vast expanse of smoke and mist.

有道翻译: Where there is a lake, there is a vast expanse.

译文: Where there is lake, there is a vast expanse of misty waters.

例 1 关于成语“烟波浩渺”的翻译,三个机译软件翻译的各不相同,分别译为“smoke and mirrors”和“a vast expanse of smoke and mist”和“a vast expanse”。实际上本句中“烟波浩渺”是形容烟雾笼罩的江湖水面极其辽阔。DeepL 的译文存在显著的误译。它将“烟波”错误地翻译成了“smoke and mirrors”,这在英语中通常用于比喻虚假的表象或幻象,与原句中的“烟波浩渺”所表达的雾气笼罩的湖面景象大相径庭。百度翻译试图表达“烟波”的意象,但将“烟波”直译为“smoke and mist”并不准确。“烟波”在此并非指真实的烟雾和雾气,而是形容湖面因雾气笼罩而显得辽阔无边的景象。虽然部分捕捉到了“浩渺”的意境,但对“烟波”的理解不够准确。有道翻译漏译了“烟波”一词,只保留了“浩渺”的意境,即湖面的

辽阔。这种做法虽然简洁,但丢失了原文中雾气缭绕的视觉效果,未能完整传达原文的意境。而译文准确且生动地传达了“烟波浩渺”的意境。其中,“misty waters”恰当地描述了雾气笼罩的湖面,而“vast expanse”则准确地传达了湖面的辽阔无边。

3 机译英文版偏误的解决方案

针对《找个理由 Go 无锡》机译版本存在的问题,笔者建议可通过以下方法改善机器翻译与计算机辅助翻译的局限性。

3.1 译前编辑

在机器翻译流程中,译前编辑作为一种提升译文质量的前置处理手段,指在开展机器翻译前针对原文实施优化调整。通过科学有效的译前编辑操作,能够大幅降低后续译后编辑环节的工作负荷,故而在专业领域常被纳入广义译后编辑体系的范畴。具体实践过程中,编辑策略需依据文本类型进行差异化设计。编辑者应着重对原文进行符合目标语言表达范式的适应性修改,通过调整句法结构、优化词汇选择等方式,强化文本的语篇连贯性与语义可译性。这种前置性处理不仅有助于提升机器翻译系统的处理效率,更能从源头上保障译文质量的可控性。

以《找个理由 Go 无锡》为例,这类旅游文本语言风格书面化,善用修辞手法(如四字格)增强抒情效果。其显著特点是大量使用无主句(即省略主语的句子),而英语除祈使句等特殊情况下,通常要求主谓结构完整。若忽视汉英句法差异,译文可能机械保留无主句形式,违背英语表达规范。故对该文本的译前编辑十分必要,第一、全面分析文本,识别出需要编辑的部分,包括无主句、长句和复杂意群。第二、根据上下文和文本的语境,为无主句确定合适的逻辑主语。第三、对长句进行结构上的调整,将其拆分为较短的句子,同时保持原意不变,以提高可读性和翻译的准确性。第四、对长意群进行重组,确保翻译后的文本在逻辑上清晰,同时保持原文的意境和风格。故通过对《找个理由 Go 无锡》等特殊旅游文本进行细致的译前编辑,可以显著提高机器翻译的质量和效率。这不仅减少了译后编辑的工作量,也确保了翻译文本的流畅性和准确性,从而更好地传达原文的意境和情感。

3.2 建立语料库

为了提升人机协作效率并优化计算机辅助翻译技术,可通过扩充术语库和翻译记忆库来构建高质量的数据库,从而显著提高翻译速度与质量。需要注意的是,机器翻译存在固有缺陷,因此将机器翻译与人工审校相结合尤为关

键,这样才能确保译文的准确度。专业译员若能熟悉各类机器翻译系统的特性,就能精准识别其输出的常见错误,特别是对旅游文本等特定文体特征的把握。在审校过程中,译员可通过查阅专业资料、词典及语料库及时修正机器翻译的偏差。同时,充分利用翻译记忆库和术语库资源,能有效减少机器翻译的差错率。以《找个理由 Go 无锡》的语料库建设为例,该项目由一位具有五年以上从业经验的资深译员主导,并配备 15 名 MTI 研究生(均持有 CATTI 二级笔译证书且累计翻译量超过 25 万字)协助完成。团队主要负责术语的提取、统计、查询及校对等建库工作。

目前,语料库的机器翻译的应用极为广泛,让译者不再局限于传统的翻译法。建立《找个理由 Go 无锡》语料库是翻译过程中不可或缺的资源,它包含了特定领域内的专有名词和术语。通过不断扩充和更新术语库,可以显著提高机器翻译的准确性和可靠性。翻译记忆库记录了以往翻译的文本和对应的翻译结果,与术语库相结合,可以为机器翻译提供更为丰富的上下文信息,从而提高翻译的一致性和准确性。笔者认为建立和维护一个高质量的术语库是促进人机合作、提高计算机辅助翻译技术的关键。术语库不仅能够提升机器翻译的准确性,还能够作为人工翻译的重要参考,确保翻译工作的质量和效率。通过人机协作,我们可以充分利用机器翻译的速度和人工翻译的深度,以追求更高质量的译文。

3.3 译后编辑

译后编辑(post-editing)指根据一定目的和要求对原始机器翻译产出进行加工和修改,既是提高机器翻译质量的重要途径,也是人机交互翻译的重要表现形式。”(冯志伟、张灯柯 2022)。《找个理由 Go 无锡》包含大量历史文化词汇,机译后需要译者查阅资料、词典或语料库进行校对修正。人工翻译可借助记忆库与术语库,有效规避机译错误。在必须使用机辅翻译的情况下,建议采用“机译+译后编辑”的混合模式,既能保留机译的高效优势,又能确保人工翻译的精准度。本文基于真实项目经验,将译后编辑分为三个层级:轻度译后编辑(Light Post-editing, LPE)、中度译后编辑(Moderate Post-editing, MPE)和重度译后编辑(Full Post-editing, FPE)。

轻度译后编辑(LPE)指对机翻译文进行轻微修正或优化,人工工作量较低,修改幅度小,涉及轻度术语查证(Term Confirmation)、名词统一(Noun Consistance)、调整字符(Character Adjustment)、标记标签(Markup and Tag)、标点/拼写(Punctuation/Spelling)和漏译(Omission)等基

础维度。中度译后编辑(MPE)主要修正较大错误,如术语错误(Wrong Term)、数字/单位/地址格式(Number/Unit/Address Formatting)、误译(Mistranslation)、句式结构(Sentence Structure)、风格不一致(Inconsistance of Style)、文化传播(Cultural Transmission)、视频文本同步(Video-subtitle Synchronization)需较高工作量。重度编辑(FPE)针对机翻的高错误率内容,尤其是语言模型无法处理的逻辑和层级关系(Logical and Hierarchical Relationship)、语域风格(Register and Style)、文化接受(Culture Acceptance)及无法自动生成的习惯表达(Idiomatc Expression)等,此类错误可能导致译文不可用,需深度修改或重译。

4 结语

本文以旅游文本《找个理由 Go 无锡》机译英文版的语料为基础,进行旅游文本机器翻译测试。测试结果表明国内三大主流机器翻译引擎(DeepL Translator、百度翻译和有道翻译)旅游文本机译译文的质量达不到行业标准,错误率较高,其错误主要集中在习惯语言、文化负载词和成语的翻译上,本文从三个层面探讨了翻译偏误问题,提出了针对性的改进措施,指出译前编辑、语料库建立和译后编辑在人机协作翻译中的关键作用。通过优化机器翻译的预处理环节,能够提升译文质量,为后续译后编辑奠定良好基础。未来应顺应翻译技术发展潮流,深化人机协同翻译模式的应用,从而提升旅游外宣资料的翻译水平,更好地推动中华文化走向世界。

参考文献:

- [1] 范梓锐,杨文地. 人机耦合时代机器翻译译后编辑原则与策略例析[J]. 上海翻译,2024,(04):29-34.
- [2] 刘桂兰,邹斯或. 视听文本机器翻译质量测评和译后编辑研究——以《畅游赣鄱——水文化英文之旅》慕课字幕翻译为例[J]. 南昌工程学院学报,2024,43(02):73-79.
- [3] 王华树,刘世界. 人工智能时代机器翻译译后编辑能力探究[J]. 中国科技翻译,2022,35(04):21-24.
- [4] 艾天宇. 基于 DQF-MQM 框架的机器翻译与译后编辑模式[D]. 浙江理工大学,2022.
- [5] 周颖,双文庭. 《2020 年政府工作报告》机译英文版偏误分析[J]. 考试与评价(大学英语教研版),2021,(02):58-62.
- [6] 田朋. 翻译多维质量标准 MQM 模型介评与启示[J]. 东方翻译,2020,(03):23-30.
- [7] 吴昊. 翻译术语库的构建与术语构成[J]. 中国科技翻译,2018,31(02):21-23+8.
- [8] 冯全功,高琳. 基于受控语言的译前编辑对机器翻

译的影响[J]. 当代外语研究, 2017, (02): 63-68+87+110.

[9] 赵涛. 机器翻译译后编辑的现状与问题[J]. 外语教学, 2021, 42(04): 100-104.

[10] 胡开宝, 李翼. 机器翻译特征及其与人工翻译关系的研究[J]. 中国翻译, 2016, 37(05): 10-14.

基金项目: 2023KXJYS163。

作者简介: 王晓蕊(1999-), 女, 汉族, 中国江西上饶人, 硕士研究生, 现就读于江西水利电力大学, 研究方向: 翻译实践与研究。