

基于 PDCA 的用户生成内容在线翻译质量管理

高 军, 刘国聪

(上海理工大学 外语学院, 上海 200093)

摘要: 用户生成内容的在线翻译质量好坏直接影响到读者的阅读体验。基于 PDCA 循环理论四个阶段——计划(plan)、执行(do)、检查(check)和改进(action), 结合用户生成内容的在线翻译实例, 通过在上述阶段分别采取预翻译译文检测、术语管理、译前编辑和译后编辑等手段进行在线翻译质量管理, 有效提升译文的翻译质量, 并据此提出可操作的用户生成内容在线翻译质量管理流程图。

关键词: PDCA 循环; 用户生成内容; 在线翻译; 翻译质量管理

中图分类号: H 059 **文献标志码:** A **文章编号:** 1009-895X(2019)03-0213-05

DOI: 10.13256/j.cnki.jusst.sse.2019.03.003

The PDCA-based Online Translation Quality Management of User-generated Content

GAO Jun, LIU Guocong

(College of Foreign Languages, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The online translation quality of user-generated content (UGC) will affect the reading experience of its readers. Based on the PDCA cycle of P (plan), D (do), C (check), A (action) and the online translation of a UGC sample text, this paper discusses the translation quality management of UGC by application of pre-translation quality assessment, term management, pre-editing and post-editing in the process of online translation, which effectively improves the translation quality. This paper also suggests a flow chart of translation quality management of user-generated content.

Keywords: PDCA cycle; user-generated content; online translation; translation quality management

用户生成内容(user-generated content, UGC)是由业余人士以非专业渠道制作的、包含一定的创作性劳动并在网络上公开可用的内容^[1]。用户在脸书、推特、维基和博客等社交媒体网络平台上发表观点、交流意见和分享经验, 由此贡献生成的网站相关内容都属于用户生成内容。

随着经济全球化的推进, 世界知名公司和产品服务商意识到用户生成内容对产品宣传推广的重要

性, 对用户生成内容的翻译服务需求越来越大。值得指出的是, 用户生成内容数量大、更新快, 如此庞大的翻译工作量由人工完成显然耗时低效。因此, 在线机器翻译以其便捷性成为众多语言服务商使用的翻译工具, 在线翻译的质量管理也随之变得愈发重要, 因为翻译质量的好坏将直接影响到读者的阅读体验。本文拟根据 PDCA 循环理论, 借助在线翻译工具——谷歌翻译工具包(Google Trans-

收稿日期: 2018-04-28

作者简介: 高 军(1970-), 女, 教授。研究方向: 英汉对比与翻译、社会语言学、英语教学。

E-mail: gaojun@usst.edu.cn

lator Toolkit), 结合用户生成内容的在线翻译实例, 即某旅游评论网站上海希尔顿酒店客户的英文点评的在线翻译质量管理, 探讨如何设计可操作的在线翻译质量管理流程, 提升用户生成内容的在线翻译质量。

一、PDCA 循环理论

PDCA 循环是能使任何一项活动有效进行的一种合乎逻辑的工作程序, 最早由统计质量管理之父休哈特提出, 后由美国质量管理专家戴明博士于 1950 年再度挖掘并在质量管理中得到了广泛的应用。P, D, C, A 分别是 plan (计划)、do (执行)、check (检查) 和 action (改进) 4 个英文单词的首字母, 即代表质量管理的 4 个阶段。具体而言, 计划阶段包括方针和目标的确定以及活动计划的制定; 执行阶段是根据已知的信息, 设计具体的方法、方案和计划布局, 再根据设计和布局, 进行具体运作, 实现计划中的内容; 检查阶段总结执行计划的结果, 分清哪些对了, 哪些错了, 明确效果, 找出问题; 改进阶段对检查的结果进行处理, 成功的经验加以肯定, 并予以标准化, 失败的教训也要引起重现。对于没有解决的问题, 应进入下一个 PDCA 循环中去解决。

PDCA 循环的四个阶段不是运行一次就完结, 而是周而复始地进行。一个循环结束了, 解决了一部分问题, 没有解决或新出现的问题, 则进行下一个 PDCA 循环, 依此类推, 产品、服务或工作质量随之呈阶梯上升。PDCA 循环应用了科学的统计观念和处理方法, 是全面质量管理所应遵循的科学程序。PDCA 循环不仅在质量管理体系中运用, 也适用于一切循序渐进的管理工作, 包括翻译质量管理。

二、文本来源及翻译工具介绍

创立于 2000 年的猫途鹰网站 (TripAdvisor.com) 是一个大型的在线旅行数据平台, 收录了世界各地的海量旅行信息, 包括逾 5.35 亿条旅游点评及建议, 平均每分钟就有 290 多条新的用户生成内容。本研究选取该网站上海希尔顿酒店客户的英文点评为文本对象, 时间跨度为 2014—2015 年, 按点评打分星级分别选取 5 星到 1 星点评各 5 条 (每条点评大约 80~90 个单词), 共计 170 个句子, 2 110 个单词。

选择的在线翻译工具——谷歌翻译工具包, 是由谷歌公司 2009 年推出的一个比较成熟的在线翻译辅助系统。该工具包的最大亮点是将谷歌自动翻译融入其中, 为用户提供了实时翻译的便利。将翻译文档上传至谷歌翻译工具包账户后, 谷歌翻译对文档进行预翻译, 用户在此基础上可以逐句修改译文^[2]。

谷歌翻译工具包操作简单, 翻译记忆库和术语库等一目了然。翻译记忆库可自行保存译文, 当遇到相近或相同的文字或句段时, 翻译记忆就会自动搜索、比对和匹配, 选出系统设定值范围内的译文作为参考译文。而译员也可发挥主观能动性, 选择接受或修改译文, 改译的文字随即被存入记忆库中, 以备下次参考之用。与其他计算机辅助翻译工具不同的是, 谷歌翻译工具包的翻译记忆库可以全球共享, 也支持译者对推荐的例句进行评分, 因而高质量的译文被推荐的机率就会加大, 翻译也就愈加精准。此外, 用户可提取术语并上传至术语库, 还能邀请他人共同使用术语库。译员在翻译过程中若发现了新术语, 可通过即时通信功能, 讨论并确定新术语, 添加到已有术语库, 从而更新和充实术语库。

三、翻译质量管理流程

(一) 计划阶段

计划阶段首先应明确翻译质量标准, 从而确定质量目标并制定质量计划; 其次, 与人工翻译不同, 机器翻译可以直接、快速地进行预翻译, 因此可以先对预翻译译文进行抽样检测, 评估预翻译的翻译质量, 为执行阶段作好准备。

1. 确定翻译质量标准

翻译质量标准的选择和制定是翻译质量管理中很重要的一环。国内外行业协会先后颁布了翻译服务译文质量的基本要求、质量评定和检测方法, 为翻译行业服务管理的规范化提供依据。但是, “随着文本类型的增加和变化, 推特和论坛等用户生成内容等文本在翻译生产中扮演着越来越重要的角色, 当前基于企业的翻译质量评估模型已不大适用, 新文本类型的翻译需要新的质量评估方式”^[3]。

O'Brien^[3]认为, 推特和论坛等用户在线生成内容属于“客户对客户”的交流模式, 其翻译质

量要求一般低于“商家对客户”,可采用充分性(Adequacy)和流畅度(Fluency)两个参数对此类在线翻译译文进行动态翻译质量评估。当不存在参考译文时,充分性可用来表示原文内容被传递的情况,根据内容传递的程度可以分为5个等级;流畅度用来表示译文的流畅程度、语法是否正确以及用词是否地道等,也分为5个等级,详见表1。

表1 机器翻译质量评估参数

Tab.1 Assessment parameters of machine translation quality

等级参数	5	4	3	2	1
充分性	全部	大部分	较多	较少	几乎没有
流畅度	完美	良好	不地道	不通顺	完全无法理解

2. 预翻译及其译文质量检测

首先,使用谷歌翻译工具包对选取的文本进行在线预翻译;其次,从预翻译译文中随机抽取5星到1星点评各一条,共计38个句子和420个单词,邀请5名专业翻译人士根据充分性和流畅度两个参数对预翻译质量进行评分,结果取5个评分的平均值。结果显示,充分性为3.10分,介于“较多”和“大部分”之间;流畅度为2.67分,介于“不地道”和“不通顺”之间。由此可见,在线翻译译文基本能够传递原文较多内容,但在句子流畅度方面表现不够理想,大部分句子不够通顺。酒店的顾客点评旨在传达信息,帮助其他用户了解酒店设施和顾客体验。而预翻译评测结果表明,译文没有完全传达原文的信息,在流畅度方面也存在问题,因此需要对用户生成内容的在线翻译进行有效的质量管理。

3. 分析问题并制定计划

翻译过程一般分为译前、译中和译后三个阶段。而在机器翻译过程中,人机交互主要体现在译前和译后,通过对这两个阶段进行分析,找出影响译文质量的因素。

在译前阶段,不仅要检查原文的词汇以及语法结构,还要检查原文的语言错误和语言歧义现象^[4]。经分析,发现原文存在一些机器翻译经常误译的人名、地名和酒店领域相关术语,而且有较多的语法和拼写错误,部分句子过于口语化,且存在一些长句和复杂句,因此有必要进行术语管理和译前编辑。

Vasconcello & Bostad^[5]认为,在译后编辑阶段,应尽量保留机器翻译输出的译文,有针对性地修改,而不是重新翻译。因此,为确保译文准确流畅,在译后编辑中需要修改明显的语言错误、词汇和句子的漏译及误译,并调整语序,保证句子通顺。

此外,为了检测和把控翻译质量改进情况,还需要再次运用充分性和流畅度两个评估参数对译前、译中和译后三个阶段的译文进行进一步的语言质量检测,从中发现问题,总结经验,进入下一个质量管理循环。

(二) 执行阶段

执行阶段主要涉及术语管理、译前编辑和译后编辑三个过程。

1. 术语管理

术语漏译和误译是机器翻译的常见问题。例如,在预翻译译文检测中,笔者发现“Wifi”漏译,人名“Crystal”误译为“水晶”,因此译前的术语整理是翻译顺利执行的关键。首先查找原文,完成术语编辑,将术语保存成符合规范的UTF-8编码的CSV格式文件,然后点击谷歌翻译工具包的术语库,再点击上传,即可进入上传术语库的界面,选择文件,为文件命名,最后点击上传至术语库。当翻译至带有术语的句段时,工具包栏就会提示相应的术语,并且会在句段中圈出该术语。整理的部分术语如表2所示。

表2 术语
Tab.2 Terms

序号	原文	译文
1	Crystal	克莱斯特
2	CNY	人民币
3	Jing'an Temple	静安寺
4	Wifi	无线网络
5	Hilton Hotel	希尔顿酒店
6	advanced deposit	定金

2. 译前编辑

译前编辑过程中,首先要纠正原文语言错误。只有提高原文语言质量,才能改进机器翻译译文质量。此外,由于谷歌翻译工具包对长句和复杂句的

处理能力较差,需要根据相关改写规则对原文进行译前编辑。表3列出了在译前编辑中的主要修改类型。

表3 主要修改类型
Tab.3 Major correction types

修改类型	修改前	修改后
拼写错误	frendly	friendly
词义歧义	The staff is exceptional.	The staff is nice.
句式修改	What an experience!	It is an amazing experience!
长句切分和改写	My booking indicated that I had internet access but the front desk people didn't honor that arguing that the internet access here meant the internet access is available only in the lobby not in the room.	My booking indicated that I had internet access but the front desk people didn't honor that. He argued that the internet access is available only in the lobby, not in the room.

笔者再次对经译前编辑后的机器翻译译文进行抽样评估,并与之前的预翻译译文进行比较,详见表4。

表4 译文质量检测结果
Tab.4 Results of translation quality assessment

译文参数	预翻译译文	经译前编辑后的译文
充分性	3.10	3.78
流畅度	2.67	3.50

表5 译后编辑修改类型及其修改方法
Tab.5 Types and methods of post-editing

修改类型	原句	修改前	修改后
词汇误译	We were in Shanghai adopting our newest daughter.	我们在上海采用我们最新的女儿。	我们在上海领养了我们第一个女儿。
词性误译	Crystal met us at the desk.	克莱斯特迎接我们的办公桌。	克莱斯特在办公桌旁迎接我们。
词序混乱	One issue is that there is often a petroleum cleaning product smell throughout the hotel which is very pungent.	一个问题是有经常整个酒店,这是一个问题,经常有非常刺鼻的石油清洁产品气味。	一个问题,经常有非常刺鼻的石油清洁产品气味充斥着整个酒店。
冗余	This is why this hotel is so comfortable.	这就是这家酒店为什么如此舒适	这就是这家酒店如此舒适的原因。
漏译	Unless you have to stay here because the company puts you on business here.	除非你留在这里,因为该公司把你的业务在这里。	除非你留在这里,因为该公司把你的业务安排在这里。

(三) 检查阶段

为了检查执行阶段的效果,分别从每个评分星级中各抽取1条评论(样本字数约400字,占总字数的20%)作为样本,对样本的前期、中期和后期译文,即预翻译译文、经译前编辑后的译文和经

表4显示,经译前编辑后的译文质量相比之前有了较大提升,基本传递了原文大部分内容,流畅度也较理想。

3. 译后编辑

鉴于机器翻译译文质量的不确定性,译后编辑成为提高译文质量的必要环节。谷歌翻译工具包的译后编辑功能十分方便,用户可在预翻译的基础上逐句修改译文。

译后编辑分为轻度译后编辑(Light Post-Editing)和完全译后编辑(Full Post-Editing)。轻度译后编辑倾向于调整自动翻译输出的结果,尽量忠实于机器产出的译文,修改明显的语义误译、增添漏译和删除多译等。该方式效率高、成本低,但译文质量提升效果有限。完全译后编辑不仅要确保词义和句义准确无误,对译文的流畅度和风格也有一定要求。该方式能大大提升译文质量,使其尽可能地贴近人工译文质量,但成本高,效率低。根据翻译自动化用户协会(TAUS),译后编辑方式的选择要基于机器译文质量状况和最终译文质量标准的要求。当机器译文质量较高而且最终译文质量标准较低(不需要达到出版要求)时,可选择轻度译后编辑。由于用户生成内容对翻译质量的要求不是很高,译文只要达到信息准确、句子通顺即可,因此,笔者选择了轻度译后编辑,部分修改信息见表5。

轻度译后编辑后的译文进行了翻译质量抽样检测,三种译文的充分性和流畅度结果如图1所示(上方为充分性,下方为流畅度)。

图1显示,译前编辑、术语管理和译后编辑等措施能有效提升译文的充分性和流畅度(分别提升了1.70分和2.07分),其中流畅度整整提高了

2 个分数档。此外笔者还发现译前编辑和术语管理对在线机器翻译的质量提升有着不错的效果,充分性和流畅度分别提升了 0.68 分和 0.83 分;译后编辑对译文流畅度的改善更为明显,提升了 1.24 分。

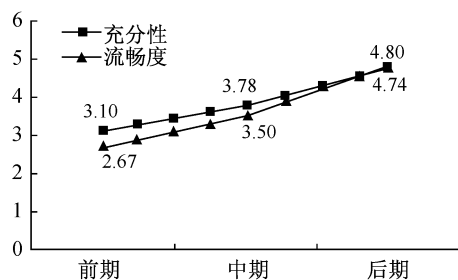


图1 翻译质量抽样评测

Fig.1 Translation quality assessment of the sample

(四) 改进阶段

根据 PDCA 循环法,需要对检查结果进行分析,成功的经验加以肯定,并予以标准化,失败的教训也要引起重视。没解决的问题进入下一个 PDCA 循环中予以解决。

1. 总结经验并标准化

上述分析表明,运用全面质量管理 PDCA 循环法对在线机器翻译出现的几大问题采取相应改进措施之后,译文质量有了较大提升。为了对成功经验进行总结和标准化,笔者归纳了用户生成内容的在线翻译质量管理的流程图,见图 2。

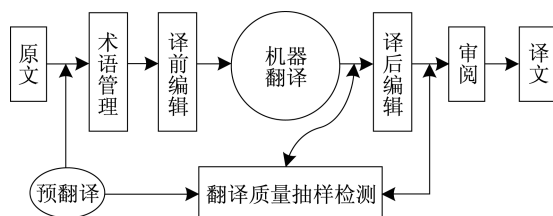


图2 用户生成内容的在线翻译质量管理流程

Fig.2 Flow chart of online translation quality management of UGC

翻译质量管理流程大致可分为两个环节:译文生产和质量监控^[6]。根据本研究的文本特征、采取的质量改进措施和相应的译文提升情况,我们认为运用谷歌翻译工具包的用户生成内容的译文生产环节应包括机器预翻译、译前编辑、术语管理、机器翻译、译后编辑和审阅等若干个阶段;质量监控环节则采取机器翻译质量检测参数对各个生产环节进行翻译质量抽样检测,为译前编辑、译后编辑和

审阅提供指导。两个环节处于并行状态,协同保证翻译质量达到预期目标。

2. 待解决问题

在质量全面管理中,成本控制,尤其是时间成本控制十分重要。而上述翻译质量控制流程中未设置时间控制环节,没有对译前和译后编辑的使用时间进行监测。如果译前和译后编辑所花费的时间与人工翻译所用时间相差无几甚至更多,那么机器翻译就得不偿失了。因此在下一个 PDCA 循环中,将对译前和译后编辑的时间进行监测,并设置人工翻译对照组,进一步探讨译前和译后编辑效率问题。

四、结束语

本研究运用 PDCA 循环理论,以“猫途鹰”上海希尔顿酒店客户的英文点评的在线机器翻译为例,从翻译质量管理的计划、执行、检查和改进四个阶段分别针对用户生成内容的在线翻译执行了包括译前编辑、术语管理、译后编辑和翻译质量检测等质量管理措施,从而使译文质量在内容传递和流畅度上都有显著提升。此外,从译文生产和翻译质量监控两个环节探讨和制定了用户生成内容的在线翻译质量管理流程图。最后,提出该流程仍未解决的时间成本问题,并投入下一个循环进行解决,不断完善翻译质量管理流程,提升译文质量。

参考文献:

- [1] Vickery G, Wunsch-Vincent S. Participative web and user-created content: web 2.0, wikis and social networking [R]. Paris: OCDE Publishing, 2007.
- [2] 徐彬. 翻译新视野: 计算机辅助翻译研究 [M]. 济南: 山东教育出版社, 2010.
- [3] O'Brien S. Towards a dynamic quality evaluation model for translation [J]. Journal of Specialized Translation, 2012(17): 55-775.
- [4] Gross A. Limitations of computers as translation tools [J]. Translation Directory on Line, 1992: 96-130.
- [5] Vasconcellos M, Bostad D A. Machine translation in high-volume translation environment [J]. Computers in Translation: A Practical Appraisal, 1992: 58-77.
- [6] 王华伟, 王华树. 翻译项目管理实务 [M]. 北京: 中国对外翻译出版公司, 2013.

(编辑: 朱渭波)