

译前编辑对机器翻译质量影响的研究^{*}

胡 倩

中南财经政法大学外语学院，武汉

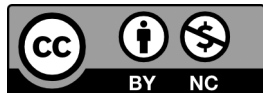
摘 要 | 本研究将探索翻译质量评估的科学方法，并借此研究译前编辑对翻译质量的影响，力图填补译后编辑领域的研究不足。通过客观的评分系统，本研究结果表明，合理的译前编辑和预处理能够显著提升文本对机器的适应度，进而提高机器翻译的精准性和质量。

关键词 | 译前编辑；翻译质量；翻译质量评估体系

Copyright © 2023 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



一、引言

随着科技的快速更新，机器翻译（Machine Translation，MT）在翻译行业中发展迅猛。从最早的基于规则库的语法和词汇知识系统，到统计机器翻译（Statistical Machine Translation，SMT），再到谷歌 2016 年推出的神经机器翻译系统（Neural Machine Translation，NMT）以及 2023 年初出现的 ChatGPT，都标志着机器翻译的不断改进和发展。如今，机器翻译已经成为主流，并且常常与译后编辑相结合。但是，对于某些特殊领域的文本和复杂句子，机器翻译仍然效果不佳，需要大量的译后编辑工作，从而降低了整体翻译效率。相反，译前编辑在机器翻译之前对需要翻译的文本或文档进行有针对性的修改编辑以提高机器翻译产出的质量（冯

全功，2017：63–68）。然而，由于人们翻译实践中的思维习惯或者时间压力，译前编辑常存在于译后编辑的阴影下，其角色常被忽略，其作用常被低估。

本文以湖南省人民政府门户网站发布的《湖南省优化营商环境规定》为材料，以百度翻译和 DeepL 为机器翻译工具，进行汉译英机器翻译实践，主要探析译前编辑对机器翻译质量的影响。本文先回顾了译前译后编辑和翻译质量评估的研究现状，再对机器翻译的试句结果进行质量评估，以检验译前编辑对翻译质量的影响，最后介绍本研究的结果和意义。

二、译后编辑和译前编辑的研究现状

从国内外对译后编辑和译前编辑的研究来看，

^{*}2023 年度中南财经政法大学研究生科研创新平台项目（202310810）。

主要关注点在两个方面,一是对翻译效率的影响,二是对翻译质量的影响。首先,关于对翻译效率的影响,根据 O'Brien (2007) 的研究,译后编辑时包含消极翻译指征的句子会导致整体编辑速度变慢,例如较长的名词短语和动名词会对编辑负荷有较大的影响。Garcia (2011) 进行了英汉双向翻译后编辑的研究,发现在逆向翻译(汉译英)时,译后编辑可以显著提高速度。此次,关于译前译后编辑对翻译质量的影响,Barreiro (2011) 介绍了 SPIDER 系统,这是一种用于文档编辑和修订中的改写系统,适用于机器翻译译前编辑。通过收集实际的译前编辑实例,以提高机器翻译系统的翻译质量。Rei Miyata (2017) 进行了针对日语到英语的翻译实验,共收集了 400 个源句子的 12687 个译前编辑实例,结果表明,超过 85% 的源句子被 MT 系统准确翻译。

而国内主要研究译前和译后编辑的方法和策略。袁亦宁(2003)提出将“受控语言”(controlled language)引入译前编辑,此举意在限制句子的表达,使其能被机器翻译系统理解。胡清平(2005)则更进一步讨论了受控翻译及其前景和设想,认为受控语言是解决自然语言使用中的随意。魏长宏(2008)提出格式转换、使用受控语言、建立翻译记忆库等手段。黄越悦(2017)、曹媛(2018)和陈丽(2020)则重点分析了机器翻译中译前编辑的几种方法,包括句法机构调整、增补、省略、提前翻译术语、补充标点符号等。

总而言之,随着机器翻译和翻译实践的发展,译前和译后编辑的研究和应用也应该不断前进和调整。首先,从知网上的论文数量来看,以“译后编辑”为主题的中英总库文章超过 800 篇,而以“译前编辑”为主题的文章只有不到 70 篇,显示出译前编辑的研究明显少于译后编辑。但事实上,译前编辑在确保翻译质量方面起着至关重要的作用,它可以帮助机器更好地理解原文。其次,如何确定编辑后的译文质量得到提升? 这的确缺乏客观标准,主要是因为翻译本身就是一项艺术,难以将其完全量化。然而,我们可以通过与原始文本的准确度、流畅度、一致性等因素进行比较来评估翻译的质量。

三、翻译质量评估

本文采用翻译多维质量标准 Multidimensional

Quality Metrics (MQM) 对译前编辑前后的译文进行评估,来比较译文质量的变化。MQM 翻译质量评估系统可以对人工翻译和机器翻译进行翻译质量评测,它提供了一种标准化的方法来评估翻译质量,它由本地化行业标准协会(LISA)开发,现在由全球化和本地化协会(GALA)维护,并以两个国际标准为体系基础:(1) ISO DIS 5060: 2022, 翻译服务—翻译产出的评估—通用指南;(2) ASTM WK46396: 分析性翻译质量评估的标准做法。

MQM 系统基于一组指标,涵盖翻译质量的各个方面,这些指标包括若干个维度,比如准确性、流畅度、术语、风格、一致性和文化适应性。每个指标都被分配 0 至 4 的分数,其中 0 代表完全不符合该指标,也就是说没有出现该错误;4 代表最符合的程度,表示该错误很明显。评估者可以使用这些分数计算翻译的整体 MQM 得分,该得分可用于比较不同翻译的质量或跟踪随时间的改进(总分 100 分)。需要注意的是,由于错误类型众多、文本类型多样,评估者应该只选择他们需要的那些错误类型,并根据特定内容类型的质量要求和利益相关者的要求来评分。

本文将根据 MQM 翻译质量评估体系对所有被试句子进行评分,包括没有进行译前编辑的一组 and 进行了译前编辑的一组,以比较译前编辑前后的评估分数,探究译前编辑对翻译质量的影响。

四、译前编辑方法探究

译前编辑即通过对文本的改写和调整,使其更容易被机器翻译系统所理解,最终帮助我们提高翻译实践的质量与效率。上述提到的译前编辑方法可以分为多种,但总体来说可以分为大三类:切分组合、转换扩充和适当删减。

(一) 切分和组合

第一种是需要切分的情况,因为中文中很多长句不见句号,还有很多并没有体现逻辑和结构,所以在汉译英时,译前编辑可以尝试多进行适当的断句。而另一种需要组合的情况,汉语重意合,句子以动词为中心,有很多流水句,一般不需要介词、连接词和分句,所以译前编辑可以选择组合词句的方法。

句子 1: 依法必须进行招标的项目以特定行政

区域、特定行业的业绩、奖项作为投标、加分和中标条件。

百度翻译: Projects that must be tendered in accordance with the law are subject to bidding, bonus points, and winning conditions based on the performance and awards of specific administrative regions and industries. (MQM 质量评估: 64.5 分)

DeepL: Projects that must be tendered in accordance with the law to specific administrative regions, specific industry performance, awards as bids, bonus

points and winning conditions. (MQM 质量评估: 56 分)

第一,从术语维度来看,“投标、加分和中标”的翻译需要更贴合语境,错误较为明显;第二,准确性不高,原文重点是“项目以若干指标作为投标条件”,但是机器翻译译文把重点放在了“项目进行投标”;第三,语法方面基本符合英文语法规则;第四,译文风格与原文相差不大,基本符合要求;第五,表达方式不够自然,需要进一步改进。

以句子 1 的百度翻译结果为例,基于以上分析,各维度罚分和最后总分如表 1 所示。

表 1 句子 1 的 MQM 质量评分

Error Severity Levels	Neutral	Minor	Major	Critical	Error Type Penalty Total	
Severity Penalty Multiplier	0	1	5	25		
Error types	Error Counts				ET Weights	ETPTs
Terminology			1		1	5
Accuracy			1		1	5
Linguistic Convention	1				1	0
Style	1				1	0
Custom		1			1	1
Absolute Penalty Total						11
Evaluation Word Count						31
Penalty per 100 word						35.5
Overall Score						64.5

原句第一个顿号之前的“招标的项目”是后面一连串介词短语“以特定行政区域、特定行业的业绩、奖项作为投标、加分和中标条件”的修饰对象。由于后面介词短语部分过长和复杂,可以选择把这两部分分开的方法,使机器更明晰句子结构,以免误解整句的修饰对象。

译前编辑后: 依法必须进行招标的项目,以特定行政区域、特定行业的业绩和奖项作为投标、加分和中标条件。

百度翻译: Projects that must be tendered in accordance with the law shall have performance and awards from specific administrative regions and industries as the conditions for bidding, bonus points, and winning the bid. (MQM 质量评估: 81.25 分)

DeepL: Projects that must be tendered in accordance with the law, take the specific administrative regions, specific industry performance and awards as bids, extra points and winning conditions. (MQM 质量评估: 74.1 分)

句子 2: 企业取得建设用地后,已经实行区域评估,且符合适用条件的审批事项可以实行告知承诺制办理。

百度翻译: After obtaining construction land, the enterprise has implemented regional assessment, and the approval items that meet the applicable conditions can be handled through a notification commitment system. (MQM 质量评估: 77.8 分)

DeepL: Enterprises obtain land for construction, has been the implementation of regional assessment, and meet the applicable conditions of approval matters can be implemented to inform the commitment system for. (MQM 质量评估: 58.6 分)

原句是典型的中文流水句,逻辑体现不强,所以可以通过加连接词体现其层次,帮助机器理解。添加“则”体现行为的结果,并用“且”紧密连接前后两个平行条件,防止翻译机器系统分开理解。

译前编辑后: 企业取得建设用地后,则可以对已经实行区域评估且符合适用条件的审批事项,实

行告知承诺制办理。

百度翻译: After obtaining construction land, enterprises can implement a notification and commitment system for approval items that have already undergone regional assessment and meet applicable conditions. (MQM 质量评估: 92 分)

DeepL: After enterprises obtain land for construction, then they can implement the notification and commitment system for the approval matters that have been implemented for regional assessment and meet the applicable conditions. (MQM 质量评估: 80.6 分)

句子 3: 行政机关以及法律法规授权的具有管理公共事务职能的组织办理依申请的行政事项要求提供证明材料时, 应当按照便民利民原则, 推行证明事项告知承诺制。

百度翻译: Administrative agencies and organizations authorized by laws and regulations to manage public affairs shall, when providing proof materials in accordance with the requirements of the applied administrative matters, implement a system of notification and commitment of proof matters in accordance with the principle of convenience and benefit to the people. (MQM 质量评估: 76 分)

DeepL: Administrative organs and laws and regulations authorized by the management of public affairs functions of the organization for the application of administrative matters required to provide supporting documents, should be in accordance with the principle of convenience and benefit to the people, the implementation of the commitment to inform the certification system. (MQM 质量评估: 77 分)

句子太长时, 机器不好做出判断, 在句子 3 中有多个动词, 如“授权”“办理”“提供”“告知”“承诺”等, 可以预判机器容易弄错其中的逻辑关系或主动和被动关系, 所以译前编辑可以做切分处理。

译前编辑后: 行政机关以及法律法规授权的具有管理公共事务职能的组织, 办理依申请的行政事项要求提供证明材料时, 应当按照便民利民原则, 推行证明事项告知承诺制。

百度翻译: Administrative agencies and organizations authorized by laws and regulations with the function of managing public affairs, when handling administrative

matters according to the application and requesting proof materials, should follow the principle of convenience and benefit to the people and implement a system of notification and commitment for proof matters. (MQM 质量评估: 92 分)

DeepL: Administrative organs and laws and regulations authorized by the organization with the management of public affairs functions, for administrative matters on application required to provide supporting documents, should be in accordance with the principle of convenience, the implementation of the commitment to inform the certificate system. (MQM 质量评估: 76.1 分)

句子 4: 鼓励中央垂直管理部门将其实施的政务服务事项进入所在地综合性政务服务大厅集中办理。

百度翻译: Encourage central vertical management departments to centrally handle their government service matters in the comprehensive government service hall where they are located. (MQM 质量评估: 63.6 分)

DeepL: Encourage the central vertical management departments to implement its government services into the location of comprehensive government services hall centralized processing. (MQM 质量评估: 61.9 分)

同样地, 句子 4 比较长且句子中动词多, 译前编辑可以选择在适当处断句。

译前编辑后: 鼓励中央垂直管理部门, 将其实施的政务服务事项纳入当地综合性政务服务大厅集中办理。

百度翻译: Encourage central vertical management departments to include their implemented government service matters in the centralized processing of local comprehensive government service halls. (MQM 质量评估: 86.4 分)

DeepL: Encourage the central vertical management of its implementation of government services matters into the local comprehensive government services hall centralized processing. (MQM 质量评估: 81 分)

(二) 转换和扩充

语言的表达方式因文化、时代、地域等因素而异, 因此直接将一种语言中的句子直译成另一种语言中的句子可能会导致信息的丢失或误解。在进行

翻译时,适当的扩充和转换可以让翻译更符合目标语言的语言习惯和表达方式,同时使翻译更加流畅易懂。此外,有时原文中含有语言的一些特有的表达方式,而这些方式在目标语言中并不存在,因此在翻译时需要进行适当的转换和替换,以便更好地传递原文的意思。

句子 5: 帮办代办

百度翻译: Assistant agent (MQM 质量评估: 0 分)

DeepL: Help Desk (MQM 质量评估: 0 分)

译前编辑后: 帮助办理和第三人代办

百度翻译: Assist in handling and third-party handling (MQM 质量评估: 66.7 分)

DeepL: Help with processing and third-party representation (MQM 质量评估: 50 分)

句子 6: 错时延时服务

百度翻译: Wrong time delay service (MQM 质量评估: 0 分)

DeepL: Staggered time delay service (MQM 质量评估: 0 分)

译前编辑后: 错开时间和推迟时间的服务

百度翻译: Staggered and delayed services (MQM 质量评估: 75 分)

DeepL: Staggered and delayed services (MQM 质量评估: 75 分)

句子 7: 违规增设或者变相增设报装业务办理环节、申请资料、前置条件。

百度翻译: Violation of regulations or disguised addition of installation application procedures, application materials, and prerequisites. (MQM 质量评估: 78.6 分)

DeepL: Illegal additional or disguised additional business processing links, application information, pre-conditions. (MQM 质量评估: 75 分)

句子 7 的难点在于“变相增设”,而“变相增设”与前面的“违规增设”是并列的,内涵一种相反的关系,所以实际上可以理解为“直接与间接”的关系,可以作为译前编辑的切入点。

译后编辑后: 违规增设或者间接增设报装业务办理环节、申请资料、前置条件。

百度翻译: Violating regulations by adding or indirectly adding installation procedures, application materials, and prerequisites. (MQM 质量评估: 84.6 分)

DeepL: Illegal additional or indirectly additional

reporting business processing links, application information, pre-conditions. (MQM 质量评估: 75 分)

句子 8: 有条件的地区对涉及不动产转移登记的供水、供电、供气、维修资金等过户业务实现联动办理。

百度翻译: Regions with conditions can achieve linkage processing for transfer businesses involving real estate transfer registration, such as water supply, power supply, gas supply, and maintenance funds. (MQM 质量评估: 88.5 分)

DeepL: Conditional areas involving real estate transfer registration of water supply, power supply, gas supply, maintenance funds and other transfer business to achieve linkage. (MQM 质量评估: 69.6 分)

“有条件”在中文里实际是指“达到条件的”,此处原句表述容易存在混淆,可以预判机器也许会理解错误。

译前编辑后: 符合条件的地区对涉及不动产转移登记的供水、供电、供气、维修资金等过户业务实现联动办理。

百度翻译: Regions that meet the conditions will achieve linkage processing for transfer businesses involving real estate transfer registration, such as water supply, power supply, gas supply, and maintenance funds. (MQM 质量评估: 92 分)

DeepL: Qualified areas involved in real estate transfer registration of water supply, power supply, gas supply, maintenance funds and other transfer business to achieve linkage. (MQM 质量评估: 87.5 分)

(三) 适当删减

句子 9: 健全政务服务绩效由企业和群众评判的“好差评”制度。

百度翻译: Improve the “good and bad” evaluation system for government service performance, which is evaluated by enterprises and the public. (MQM 质量评估: 78.9 分)

DeepL: Sound performance of government services judged by enterprises and the public “good and bad” system. (MQM 质量评估: 73.3 分)

句子中意思的重复会误导机器。每种语言都有冗余现象,中文也存在一些语言的冗余,比如一些

范畴词和修饰词（赵刚，2004：57–62）。因此在句子 9 中我们在译前编辑时可以删除“好差评”，这不仅不影响原文意思而且能够帮助机器更好地理解文本。

译前编辑后：健全政务服务绩效由企业和群众评判的制度。

百度翻译：Improve the system of evaluating the performance of government services by enterprises and the public.（MQM 质量评估：93.3 分）

DeepL：Sound system of government service

performance judged by enterprises and the public.（MQM 质量评估：83.3 分）

五、MQM 翻译质量评估分数

评估结果分为两大组：未进行译前编辑的译文和进行了译前编辑的译文，并使用了两个主流机器翻译网站进行机器翻译处理，从表 2 可以看出，经过译前编辑后，句子质量得到显著提升，人工译前编辑对机器翻译质量有积极作用。

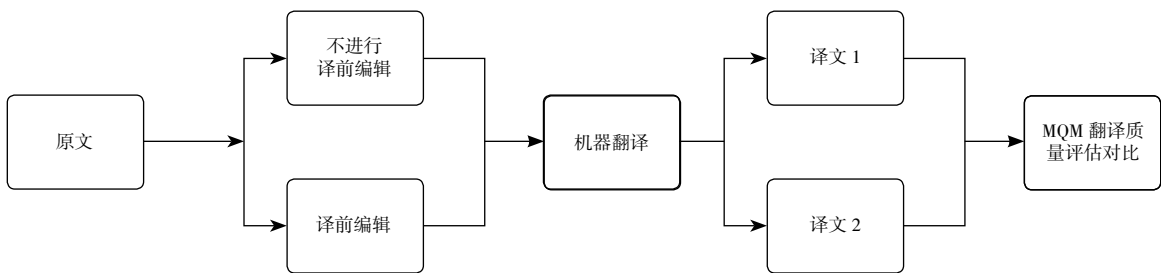


图 1 翻译质量评估流程

表 2 翻译质量评估分数对比

MQM 翻译质量评估分数	未译前编辑的百度翻译译文	未译前编辑的 DeepL 译文	译前编辑后的百度翻译译文	译前编辑后的 DeepL 译文
句子 1	64.5	56	81.25	74.1
句子 2	77.8	58.6	92	80.6
句子 3	76	77	92	76.1
句子 4	63.6	61.9	86.4	81
句子 5	0	0	66.7	50
句子 6	0	0	75	75
句子 7	78.6	75	84.6	75
句子 8	88.5	69.6	92	87.5
句子 9	78.9	73.3	93.3	83.3

六、结语

本研究首先探讨了科学的翻译质量评估方法，以 MQM 质量评估体系为基础，提供更科学的评分指导，减少了主观因素的影响。以此为基础，我们研究了译前编辑对翻译质量的影响，试图填补译前编辑领域的研究不足。研究结果表明，译前编辑普遍对译文质量产生积极影响。未来的研究需要更多关注译前编辑对于提升译文质量的潜力，探索合理的译前编辑和预处理方法，以提高文本对机器的适应度，以满足不同领域和场景的翻译需求。

参考文献

[1] Barreiro A. SPIDER: A System for Paraphrasing in Document Editing and Revision—Applicability in Machine Translation Pre-editing [J] . Lecture Notes in Computer Science, 2011.

[2] Miyata R, Fujita A. Dissecting human pre-editing toward better use of off-the-shelf machine translation systems [C] //Proceedings of the 20th Annual Conference of the European Association for Machine Translation, 2017.

[3] O'Brien S. An empirical investigation of temporal

- and technical post-editing effort [J]. *Translation and Interpreting Studies*, 2007, 2 (1): 83-136.
- [4] Garcia. Translating by post-editing: is it the way forward [J]. *Machine Translation*, 2011, 25 (3): 217-237.
- [5] 陈丽. 机器翻译中的译前编辑翻译技巧——以Google汉译英《边城》为例 [J]. *海外英语*, 2020 (8): 10-11.
- [6] 曹媛, 马永良. 谷歌翻译汉译英的译前编辑——以北京中医药信息网新闻稿的英译为例 [J]. *海外英语*, 2018 (15): 153-155.
- [7] 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 (2): 63-68, 87, 110.
- [8] 胡清平. 机器翻译中的受控语言 [J]. *中国科技翻译*, 2005 (3): 24-27.
- [9] 黄越悦. 浅谈机器翻译中译前编辑的应用 [J]. *湖北函授大学学报*, 2017, 30 (3): 175-176, 179.
- [10] 魏长宏. 机器翻译的译前处理 [J]. *科协论坛* (下半月), 2008 (9): 93-94.
- [11] 袁亦宁. 受控语言: 机器翻译的新途径 [J]. *上海科技翻译*, 2003 (3): 77-80.
- [12] 赵刚. 汉语中的冗余信息及其翻译 [J]. *国外外语教学*, 2004 (4): 57-62.

The Impact of Pre-Editing on the Quality of Machine Translation

Hu Qian

Zhongnan University of Economics and Law, School of Foreign Languages, Wuhan

Abstract: This study seeks to explore scientific methods for evaluating translation quality and investigate the potential impact of pre-editing on translation quality, aiming to fill gaps in post-editing research. Utilizing an objective scoring system, our results indicate that rational pre-editing and preprocessing significantly enhance text adaptability for machines, ultimately improving machine translation accuracy and quality.

Key words: Pre-editing; Translation quality; Translation quality assessment system