

大数据背景下机器翻译与人工翻译的应用对比研究

——以公文语体为例

李兴婕¹ 张萌¹ 李梓壮²

1. 吉林大学外国语学院 吉林长春 130012

2. 吉林大学仪器科学与电气工程学院 吉林长春 130061

摘要: 21世纪以来,人工智能技术高速发展,深度学习技术和大模型技术催生了文心一言、ChatGPT等生成式AI产品,对翻译行业产生了深刻影响,前沿技术正在改变此行业的工作范式,科技赋能翻译行业及国家翻译能力建设已是大势所趋,“人机共译”创新工作模式受到中国超500万译者的广泛关注。在不同的应用环境中,如何选择机器翻译和人工翻译模式,经济高效地产出高质量文本成为亟待解决的难题。本文将通过分析机器翻译与人工翻译的工作特点,以公文语体文本为例,对两者的应用进行对比研究,为该问题的解决提供一些建议和参考,同时对未来的翻译工作模式提出展望。

关键词: 人工翻译; 机器翻译; 公文语体; 应用对比; 人机共译

引言

从1949年机器翻译被提出以来,它的发展经历了规则驱动、统计、神经机器翻译三个阶段。在人工智能迅猛发展的浪潮下,神经机器翻译取得显著的进步,广泛应用于社会各领域。党的二十大报告着重强调创新发展,为机器翻译技术产业化提供了理论支持,刺激了语言服务行业和产业各方对机器翻译技术的研究应用。全球化的深入和“一带一路”倡议的推进深化了中外交流与合作,这为翻译行业带来了新的机遇和挑战。

目前,人工与机器翻译结合的方法正在成为主流,不同类型的文本可根据需求选择适当的翻译方式。翻译工作者如何找到二者的平衡点,以提升翻译质量和效率?面对不同文本,该如何将两者的优势实现最佳结合?本文将公文语体为例,探究机器翻译和人工翻译的具体应用,为以上问题的解决提供思路。

1. 正文

1.1 机器翻译的发展历程

机器翻译(machine translation, MT)作为人工智能和自然语言处理领域重要的研究方向之一,旨在通过计算机将源语言转化为目标语言,使翻译结果更符合人类常识。其探索历程充满曲折,学术界一般将其划分为开创期、受挫期、恢复期及发展新时期四大阶段。

机器翻译的开创期最早可以追溯到20世纪30年代初,机器翻译设想初现,直至1949年,美国科学家W. Weaver发表《翻译备忘录》,正式提出机器翻译的思想。然而,60年代初,由于机器翻译结果的不稳定性,人们开始对此提出质疑,美国科学院成立语言自动处理咨询委员会,在《语言与机器》报告中全面否定了机器翻译的可行性,机器翻译的发展进入受挫期。70年代,计算机硬件技术的飞跃发展和人工智能在自然语言处理上的应用,为机器翻译研究的复苏提供了技术支撑;同时,国家交流频繁带来翻译需求增加,为机器翻译研究带来了新机遇。机器翻译的发展于1990年迎来新浪潮,第13届国际计算语言学大会的召开,标志着基于统计的机器翻译时代的开启,借助庞大语料库和预处理技术机器翻译取得显著成果。进入21世纪,神经网络的机器翻译得到迅猛发展,Transform、Bert、GPT、Elmo等训练模型不断涌现,且功能日益强大。

1.2 人工翻译和机器翻译的工作原理及优劣对比

人工翻译作为传统的翻译方式,一般运用在书籍翻译、政治会面、大型会议现场等,可分为归化翻译和异化翻译:归化翻译注重贴近目标语言的文化习惯,是一种倾向于普通读者的变通转换,使译文读起来地道和生动;异化翻译则以源语言或原作者为翻译归宿,注重保留异域特色。

以目前作为主流的神经机器翻译为例,其工作流程包

括预处理、核心翻译、后处理三个部分。预处理是将原文规整成符合规范的句子；核心翻译阶段是将输入的字符序列翻译成目标语言序列的过程，是机器翻译中最关键最核心的地方。后处理阶段是将翻译结果进行进一步优化，使其更加符合目标语言的阅读习惯。

人工翻译在情感、修辞和文学价值方面独具优势，可与读者之间建立深层联系，使句段之间连接更为流畅自然。但成本高、速度慢，且受限于人的精力，长时间工作后可能影响翻译质量的稳定。相比之下，机器翻译具有高度自动化、成本低、速度快等优点，但局限于直接翻译，对于翻译要求较高的文本，如复杂句式和不同语言中的特定文化元素，很难达到“信”“达”“雅”的翻译效果。

由以上简单分析看来，机器翻译与人工翻译各有其优点与局限性，机器翻译的译后编辑和译文润色需依靠人工翻译，二者之间并非绝对对立，而是在互动中相辅相成的关系。

2. 机器翻译和人工翻译在公文语体中的适用性

2.1 语体特征

公文语体：书面语体的一种，是适应公务交际需要而运用全民语言所形成的言语体式。在公务活动中，按照特定体式，经过一定的处理程序形成和使用，是有一定格式的书面材料，主要用于办理公务，其作用是协调安排各种关系及事务，保障工作完成的优质优量。

公文语体具有以下特征：

①简洁性：在使用公文语体时，要求简明扼要，一针见血，即用最少的文字表达出最理想的效果。

②庄重性：通常使用规范的书面用语。不能带文学色彩，不能出现诙谐语，语言在总体上体现出公文发布者和公文的决断性、权威性。

③准确性：公文语体要求行文准确，在选择同义词上要求精确，在语气上应与公文类别保持一致。词语含义单一，不能产生歧义。

2.2 机器在公文语体下的翻译特点

①注重原意，无法适应语体风格。

宏观看，公文语体具有简洁性、庄重性、准确性三大特征。在翻译该类文本时，译文风格应遵循该语体色彩。微观看，在不同类型的公文语体中，译文语气应与公文类别保持一致。如批评错误的通报要用否定和警示语气，而表扬先进的通报，要用肯定和赞许语气，机器翻译现难满足该需求。

②通过词意达到对句子的理解，译文部分理解大于整体理解。

机器翻译运用大数据算法分析和自然语言学习实现语言转换，具有模式化和程序化的特点。但翻译并非简单的字词对应，需考虑不同国家语言习惯和思维逻辑的差异。对于公文语体，不同国家的公文都有大量专门用语和专业术语，行文逻辑和语法层面均有差异。机器翻译的运作模式，有以偏概全，词不达意的风险。如：优化财政支出结构，……把钱花在刀刃上。（2010年《政府工作报告》），百度翻译为：Optimize the structure of fiscal expenditure,... spend money on the cutting edge. (2010 Report) 在这里机器翻译把“刀刃”直译为了“the cutting edge”，但此处的意义为“关键地方”，造成了语义的偏差。

2.3 人工在公文语体下的翻译特点

①用语准确、精当。例如将 I hereby declare...译为带有公文色彩的“本人谨声明如下……”，而非“我在此声明”；选用“令”“兹”“拟”等传统公文语体的文言词语，使行文庄重典雅，增强概括性；使用礼节用语“I have the honour to...”“众望所归，领导有方”等。

②选用缩略语。缩略语的运用在使行文简洁流畅的同时，能保证意义的准确传达。如“人大”（全国人民代表大会和各级地方人民代表大会）。

③使用模糊语言。将模糊词语依据语义性质进行分类，主要分为：时间、程度、趋向、范围、频率、条件、数量。在公文语体中使用模糊语言可以简代繁，减少信息的冗杂。如运用“有些”“不少”这样表示范围的模糊词语，在不影响文章主题的基础上对数据信息进行了概括；也能避免绝对，留有余地，使得行文灵活。如：会议报告里常有 A view was expressed that…。在了解本意（“只有，一个代表（团）认为……”）的基础上，译员出于多种因素考虑，译成“有一种意见认为……”。在这类情况下往往采用“朦胧”的翻译方法。

④采用程式化表达，这也是公文语体的一大特点。如会议记录中“It was so decided.”应为“就这样决定”；“close consultation”应译为“密切磋商”而非“紧密协商”。在国际组织的公文体文件中，由于这些表达长期沿用，要求译者在翻译时也使用固定说法。

⑤会对原文句子进行拆分和合并。为了符合语言习惯，

在针对表达和思维方式不同的语言时，译者需要重新对句子进行组织，进行整合或拆分。

2.4 机器翻译与人工翻译译文对比（以《联合国宪章》作为典型文本）

例 1:

МЫ, НАРОДЫ ОБЪЕДИНЕННЫХ НАЦИЙ,
ПРЕИСПОЛНЕННЫЕ РЕШИМОСТИ

人工翻译：我联合国人民同兹决心

机器翻译（以 deep l 翻译为例）：我们，联合国人民，决心

人工翻译具有思维性，通过文言词语“兹”字，言简意赅表达了原文思想，简洁有力。而机器翻译相较于整体含义更注重各个词的意义，词与词之间缺乏关联，因此对句子的理解部分大于整体，无法精当翻译文本。

例 2:

И В ЭТИХ ЦЕЛЯХ

проявлять терпимость и жить вместе, в мире друг с другом, как добрые соседи, и

объединить наши силы для поддержания международного мира и безопасности, и

обеспечить принятием принципов и установлением методов, чтобы вооруженные силы применялись не иначе, как в общих интересах, и

...

人工翻译：

并为达此目的

力行容恕，彼此以善邻之道，和睦相处，

集中力量，以维持国际和平及安全，

接受原则，确立方法，以保证非为公共利益，不得使用武力，

.....

机器翻译（以 deep l 翻译为例）：

并本着这些宗旨

以善邻之道，和睦相处，相互容忍，共同生活，并

团结我们的力量，维护国际和平与安全，以及

通过制定原则和方法，确保武装力量只用于共同利益，而不用于其他目的，

.....

可以看到，机器翻译现具有基本的文本翻译能力，如“терпимость”本意为“忍让，宽容”，但在此处选择了“善邻之道”的译法。在句子提炼上，机器翻译注重原意却较为冗杂，不符合公文语体简洁性的特点。同时在加强语势和调动情感上，相较于人工译本而言，机器翻译的译文略显平淡。

例 3:

Члены Организации также соглашаются, что их политика в отношении территорий, на которые распространяется действие настоящей Главы, должна быть основана не менее, чем в отношении их метрополий, на общем принципе добрососедства, с надлежащим учетом интересов и благополучия остального мира в делах социальных, экономических и торговли.

人工翻译：联合国各会员国共同承诺对于本章规定之领土，一如对于本国区域，其政策必须以善邻之道奉为圭臬；并于社会、经济及商业上，对世界各国之利益及幸福，予以充分之注意。

机器翻译（以 deep l 翻译为例）：各成员还同意，它们对本章所适用领土的政策，应与对其大都市的政策一样，以睦邻友好的一般原则为基础，并在社会、经济和商业问题上适当顾及世界其他地区的利益和福利。

通过对比，可以看出，机器翻译的译文缺乏庄重性，译本风格不适应公文语体色彩。同时，对于特定词组把握不够精准，如“политика.....в отношении их метрополий”，机器翻译译为“对其大都市的政策”，而人工翻译的译文为“本国区域”。比较可知，人工翻译的严谨明确远高于机器翻译，对于需要极强规范性的公文语体而言，机器翻译的不精准性会降低文件的严谨性与可信度。

2.5 结论

公文语体特征明显，其严格的规范性显著区别于科技、政治、文艺语体。它作为具有一定法律效力的公务文书，要求简洁明了，严谨周密，精当准确。公文语体语言需避免受到译者主观意识的影响，力求客观还原文本。公文作为一种书面语言运用体系，具有使用的普遍性，因此各国公文的形式虽各有不同，但其简洁性，庄重性，准确性的特点不应有异。

针对公文语体，机器翻译对于现实情况、历史背景了解甚微，对于文本无法准确把握。且机器翻译在注重原意的基础上，难以达到公文语体庄重性、简洁性的要求，无法适

应语体的风格色彩。人工翻译的存在恰恰弥补了这一缺失,遵循公文语体的翻译理念下,人工译者经过分析与思考能够合理选择恰当的翻译方法,如一些程式化表达,同时注重译文的准确性,根据各国语言差异,文化背景解决由表达习惯不同造成的通顺性问题,做到一目了然,言简意赅。

3. 机器翻译和人工翻译的相互作用和影响

纵观机器翻译的发展历程,机器翻译至今取得了飞速的发展,在某些方面上,机器翻译确实对人工翻译造成了一定冲击和影响。一方面,机器翻译的低价特色给从业人员造成了极大的竞争,导致译者的生存空间受到挤压,价值地位有所降低。但另一方面,机器翻译可以提高译者翻译工作的效率和质量,使译者更专注于创造性的翻译工作,而非重复的机械劳动。在特定领域的程式化或信息性文本翻译时,利用机器翻译的优势,可以在短时间内解决词汇翻译的疑难。人工翻译也能弥补机器翻译的不足,发挥人工翻译主体作用的同时,还能对机器翻译不够成熟的领域进行补充。

发挥机器翻译和人工翻译的最大优势,更好服务翻译行业,是所有译者的期望。在实际应用中,两者可以进行合理分工,承担不同层次或性质的翻译任务。人工翻译也能以机器翻译为工具,有效结合两者优势,形成新型的“人机共译”模式。机器翻译和人工翻译在不同领域都有着广阔的应用前景,在不同的应用环境中,如何选择合适的翻译模式,快速、经济且高质量地完成翻译任务是关键。因此,探索“人机共译”的创新工作模式,是当前人工智能技术高速发展背景下,翻译的必然趋势。

4. 结语

全球化进程加快,世界各国和地区之间的联系愈发紧密,交流愈发频繁,这催生了大量跨语言翻译需求,传统的人工翻译越来越难以满足快速增长的翻译需求。受益于机器翻译技术的快速发展,机器翻译质量显著提升,应用场景更为广泛。

本文以公文语体为例,发现两者不同的特点,机器翻译虽具备稳定性这一显著优势,但仍存在机械性和随机性的明显短板;人工翻译相较而言更机动统一,但有一定不确定性。因此,“机器翻译完全取代人工翻译”这一说法不具有现实性,要提高编辑效率,提升译文质量,机器翻译还有很长的路要走。在某些方面利用机器翻译,可以帮助译者提高

效率和准确性,但翻译仍然是一项艰巨的工作,机器翻译可以帮助译者做基础工作,但高难度或情感色彩浓郁的文本还是需要人工润色。将机器翻译和人工翻译相结合,实现优势互补,这种“人机共译”的创新模式为翻译领域未来的发展和翻译质量的提升奠定了基础。

参考文献

- [1] 曹菡艾 编著,赵兴民 编著. 联合国文件翻译 [M]. 北京: 中国对外翻译出版公司, 2006.
- [2] 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.
- [3] 李利. 浅析机器翻译的发展历程与前景 [J]. 中国教育技术装备, 2020, (19): 48-49.
- [4] 李爽. 国际组织公文文体文件的翻译难点及翻译策略——以联合国纽约总部笔译项目为例 [D]. 上海外国语大学, 2021.
- [5] 李小睿. 基于语料库的报告类公文语言特征研究——以近十年《政府工作报告》为例 [D]. 华中师范大学, 2019.
- [6] 王菡钰, 郭凤鸣. 论人工翻译的不可替代性 [J]. 现代语言学, 2024, 12(1):15-20.
- [7] 魏光. 人工翻译与机器翻译译文编辑比较研究 [J]. 海外英语, 2021, (19): 18-19, 21.
- [8] 杨卓, 张煜琳, 魏玮, 王慧妍, 王天施. 人工智能影响下人工翻译的现状与未来 [J]. 海外英语, 2022, (13): 45-48.
- [9] 张颖倚. 汉俄公文语体对比和翻译策略 [J]. 山西青年, 2019, (23): 133-134.
- [10] 赵娟廷. 汉韩公文语体对比研究 [D]. 复旦大学, 2004.

作者简介:

李兴婕 (2003.12) 女, 汉族, 贵州遵义人, 吉林大学外国语学院, 本科, 研究方向: 俄语。

张萌 (2003.5) 女, 汉族, 河北保定人, 吉林大学外国语学院, 本科, 研究方向: 俄语。

李梓壮 (2003.1) 男, 汉族, 湖北孝感人, 吉林大学仪器科学与电气工程学院, 本科, 研究方向: 电气工程及其自动化。

基金项目:

2023年吉林省大学生创新创业训练计划 (X202310183033)