



上海理工大学学报(社会科学版)

Journal of University of Shanghai for Science and Technology(Social Sciences Edition)

ISSN 1009-895X,CN 31-1853/C

## 《上海理工大学学报(社会科学版)》网络首发论文

题目: ChatGPT 人机对话式翻译研究  
作者: 李梅, 孔德璐  
DOI: 10.13256/j.cnki.jusst.sse.230918456  
收稿日期: 2023-09-18  
网络首发日期: 2024-05-21  
引用格式: 李梅, 孔德璐. ChatGPT 人机对话式翻译研究[J/OL]. 上海理工大学学报(社会科学版). <https://doi.org/10.13256/j.cnki.jusst.sse.230918456>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# ChatGPT 人机对话式翻译研究

李梅, 孔德璐

(同济大学外国语学院, 上海 200092)

**摘要:** 聚焦 ChatGPT 人机对话式翻译, 分析其译文特征并通过与以谷歌翻译为代表的神经机器翻译质量对比发现二者差距。研究显示, ChatGPT 可提供多种译文、变化译文风格并识别文本标记。但就译文质量方面, ChatGPT 在准确性、完整性及一致性上仍存在提升空间。ChatGPT 的人机对话式翻译将会为翻译从业者、翻译学习者和技术开发带来人工智能时代的新启示。

**关键词:** ChatGPT; 人机对话; 神经机器翻译; 启示

**中图分类号:** H 319.3 **文献标识码:** A

**DOI:** 10.13256/j.cnki.jusst.sse.230918456

## A Study on ChatGPT's Human-Machine Interactive Translation Performance

LI Mei, KONG Delu

(School of Foreign Languages, Tongji University, Shanghai 200092, China)

**Abstract:** The current research explores ChatGPT's human-machine interactive translation performance in machine translation in comparison with neural machine translation (NMT) systems represented by Google Translate in English-Chinese translation. Case analyses show that ChatGPT overrides NMT systems in that it is able to provide multiple translations, produce translations of various styles, and adjust translations based on the meta information of the source text. In contrast, Google Translate does a better job than ChatGPT in accuracy, adequacy, and consistency. Implications for translation practitioners, students, and system developers, when faced with challenges from ChatGPT's human-machine translation performance, are also discussed.

**Keywords:** ChatGPT; human-machine interactive; neural machine translation; implications

近年来, 翻译界对于技术的冲击已不再陌生, 诸多“机器翻译替代人工翻译”的话题已是老生常谈<sup>[1-3]</sup>。自 2014 年神经网络机器翻译(NMT)模型推出以来, 尤其是百度和谷歌等公司相继推出大规模在线 NMT 翻译系统后, 机器翻译质量显著提升, 机器翻译译后编辑已然成为翻译新模式<sup>[4-6]</sup>。Chat-

GPT 于 2022 年底横空出世, 俨然成为“翻译工作流程上的‘游戏规则改变者’”<sup>[7]</sup>。该工具不仅能实现如写诗作曲等即时文本生成, 而且通过人机对话, 在语言翻译方面也表现出不凡实力。在此背景下, 本研究聚焦 ChatGPT 人机对话式翻译实践, 结合英汉翻译案例, 通过将 ChatGPT 生成的译文与以

收稿日期: 2023-09-18

基金项目: 国家社会科学基金年度项目(19BYY128)阶段性成果

作者简介: 李梅, 女, 教授。研究方向: 机器翻译译后编辑、技术写作/传播。E-mail: may@tongji.edu.cn

谷歌翻译为代表的 NMT 译文进行深度对比分析,从多重译文、风格复现等方面揭示 ChatGPT 与以往所有机器翻译引擎大相径庭的翻译实现手段和特征。同时,研究也表明 ChatGPT 在翻译质量上与 NMT 引擎相比仍存在一定差距。本研究能够帮助我们拓宽对机器翻译技术的认识,掌握 ChatGPT 在实际翻译场景中人机对话式翻译模式的特色和亮点,同时认识到 ChatGPT 在翻译过程中可能存在的问题和挑战,从而灵活将之运用于翻译实践和教学之中,改善翻译质量、提高翻译准确度。

## 一、ChatGPT 工作机制及特点

ChatGPT 呈现方式为问答式窗口<sup>①</sup>。用户在对话框内通过语音或者键入指令(Prompt)点击回车后,ChatGPT 将给出相应回答。如图 1 所示,用户键入指令之后,ChatGPT 即刻给出相应回答。用户还可以根据 ChatGPT 回答继续提问,进行交互。最后,用户可将所有指令和回答复制到文档中进行储存或编辑。通过图 1 中的回答可见,ChatGPT 并不是“胡言乱语”,而是具有较高的逻辑性和流畅度。



图 1 ChatGPT 交互界面

Fig. 1 ChatGPT interactive interface

ChatGPT 的核心算法是大规模预训练语言模型 (Large Language Model), 通常使用深度学习去处理海量的文本数据, 以学习和模拟自然语言中的模式和知识。最近研究表明, LLM 在一系列自然语言处理任务中取得了巨大突破, 例如文本摘要生成、文本分类、机器翻译以及人机问答等<sup>[8]</sup>。OpenAI 公司开发的 GPT 模型同属 LLM 技术, 于 2018 年推出 GPT-1。当时该模型的语言理解能力已达到较好水平<sup>[9]</sup>。GPT 经多次迭代在 GPT-3.5 模型基础上推出了 ChatGPT。与 NMT 相比, GPT 模型有所不同, 为方便读者快速了解, 我们将其与翻译相关的特点概述如下<sup>②</sup>。

### (一) 语言模型庞大

GPT 模型是迄今为止最大的 LLM 模型之一,

① ChatGPT 官网为 <https://chat.openai.com/chat> (访问日期: 2023 年 10 月 26 日)

② OpenAI 公司提供的模型既用于研究目的, 也用于生产开发, 但 OpenAI 提供的模型和论文中发表的内容并不完全匹配。此处的 GPT-3.5 实际上是采用了 Brown 等人在 2020 年发表论文中介绍的 GPT-3 模型的架构、参数和算法, 并基于截至 2021 年 6 月的数据进行训练, 比 GPT-3 模型 (在截至 2019 年 10 月的数据上进行训练) 更具有时效性。此处讨论的特点均为 GPT-3 模型, 至于该公司在 2023 年 3 月发布的 GPT-4, 由于并未向大众公开使用, 采取付费体验模式, 故不在本文讨论之列。

其内部参数量超过 1 750 亿,采用多样化互联网文本数据集(约 45TB)进行训练,内容包括各类书籍、文章和网站等,涵盖多种文本类型。ChatGPT 从庞大的多语言混合训练数据中学习,与 NMT 翻译引擎相比,能够翻译的语言对只多不少。研究显示,在德语、英语、罗马尼亚语三种语言之间,ChatGPT 可以自由翻译,但腾讯旗下的 TranSmart 囿于数据的原因,暂时无法实现罗马尼亚语的翻译<sup>[10]</sup>。当然,建立如此巨大的模型,花费也是庞大的,该模型训练一次的费用保守估计是 460 万美元,总训练成本估计达到了 1 200 万美元<sup>[11]</sup>。可以说,庞大的预训练数据使得 ChatGPT 能够发挥出高质量的机器翻译能力。

### (二) 实现一专多能

以往自然语言处理(NLP)任务在研发阶段,数据和算法往往具有唯一性,如机器翻译研发所采用的平行数据集专门为机器翻译任务搜集制作,放在其他 NLP 任务当中较难通用。然而 ChatGPT 实现了所谓的“一专多能”。“一专”指的是对话任务,ChatGPT 专门针对会话任务进行微调<sup>[12]</sup>,自然语言理解和生成能力达到质的突破。“多能”表现在能处理多个 NLP 任务,这归功于训练阶段的多任务学习以及应用阶段的微调过程,前者能够让模型同时学到多任务的特征和规律,后者则可根据具体任务选择不同输入和输出层,调整模型参数,达到更好性能。这种一专多能的特点赋予了 ChatGPT 在进行翻译任务时的人机交互能力,能够满足译者不同的需求。

### (三) 人工强化学习

ChatGPT 之所以能说人话,主要是采取了人工反馈强化学习的调优策略。与传统的强化学习方法不同,该技术不依赖于预先设定的奖励函数,而是利用人类专家的经验 and 反馈来指导模型的学习过程,从而显著提高学习效率和性能<sup>[12]</sup>。这种方法在 OpenAI 公司较早推出 InstructGPT 模型中就已经采用,并不断对该模型进行迭代升级。由此可见,ChatGPT 的翻译能力会随着模型的更新迭代得以不断进化,能够更快速地产出准确译文。

由此可见,ChatGPT 在运行机制上和谷歌翻译所采用的 NMT 引擎有所不同。ChatGPT 在具体翻译中表现如何?与 NMT 翻译引擎相比翻译质量如何?下面通过英汉翻译实例进行深入探讨。

## 二、ChatGPT 非比寻常的翻译实现

本节探讨 ChatGPT 英汉翻译实现方式,并通过案例分析展示其完成翻译任务呈现出哪些别具一格的特点。本研究原文选自英文主流报刊媒体的 10 篇文章、每篇约 900 单词<sup>③</sup>。同时,为凸显 ChatGPT 的独特的翻译特性,我们选取全球市场上技术领先且应用广泛的谷歌翻译作为 NMT 代表与之进行对比分析。

### (一) 一种原文、多种译文

ChatGPT 通过人机对话方式可为译者提供多种译文选择。在问答框内点击“重新生成回答(Regenerate response)”,或键入“换一种译法”提示语,ChatGPT 就会给出另一种译文,并对连接词、语序等结构进行微调。译者可以要求 ChatGPT 不断生成新的译文,直到自己满意为止。译者也可以在第一次提问时就要求其“提供 n 种不同翻译”,并在得到不同译文后进行甄选。例 1 展示了 ChatGPT 生成不同译文的过程。

例 1 人工指令(以下简称问)翻译成汉语,提供四种不同译法: The charismatic corporate climber is a common target for resentment in office life.

ChatGPT 回答(以下简称答)

(1) 有号召力的公司职场攀登者是办公室生活中的常见怨恨目标。

(2) 个性魅力的企业晋升者在职场生活中往往是惹人嫉妒的对象。

(3) 集人气的企业攀升者在办公室中常常成为被怨恨的对象。

(4) 魅力独具的企业扶梯者在工作场所中是常见的被人们怨恨的目标。

谷歌翻译 富有魅力的企业攀登者是办公室生活中常见的怨恨目标。

在上例中,ChatGPT 应人工要求为“corporate climber”提供了四种不同译文。其中,“职场攀登者、企业晋升者”有参考价值,而“企业扶梯者”不正确。译者可对 ChatGPT 的多重译文中或直接采

<sup>③</sup> 本研究所引 ChatGPT 案例均于 2023 年 2 月进行,ChatGPT 版本号是 GPT-3.5-turbo。由于 ChatGPT 是基于无监督模型的问答系统,且模型迭代速度极快,故复现性不强。



用,或译后编辑,或获取灵感直至得到满意的译文。而谷歌翻译的译文仅给出一种译法,且“企业攀登者”较为生硬,译者也无法从其他译法中获得参考,很可能先入为主地保留 NMT 译文,从而让读者困惑。

**例 2 问** 翻译成汉语,提供三种不同译法,进行适当的词汇、语序变换: Many totemic bosses of the age, from bankers to tech founders, come with big egos, showy antics and volatile tempers.

**答** (1) 银行家和科技创新者等时代的权威人物,个个自负、喜欢炫耀、情绪不稳,成为了图腾老板。

(2) 时代的许多重要人物,从银行家到科技创始人,个个都是自信心爆棚、行为张扬、情绪多变的图腾老板。

(3) 无论是银行家还是科技创新者,当今时代的许多重量级领导人都表现出了高度的自我意识、喜欢表演的倾向和情绪反复无常的特点,成为了时代的重要人物。

**谷歌翻译** 这个时代的许多图腾老板,从银行家到科技创始人,都非常自负、爱炫耀和脾气暴躁。

例 2 中的英文原文的难点首先在于对“totemic boss”的理解,以及句子最后一系列的修饰成分。在 ChatGPT 的回答中,所有译句都针对“totemic boss”做出了额外阐释,如“权威人物、企业大佬”等,让读者一目了然,明白“图腾老板”的引申意义。同时对于句末一系列的修饰成分,ChatGPT 回答既有译句 1 中的连续三个四字词语的使用,符合汉语的行文特征,也有译句 3 中对每个修饰成分深入细致的阐释。译者可以权衡 ChatGPT 给出的不同译句,考量原文中的疑点难点,给出合适的译文。更为重要的是,ChatGPT 得益于强大的交互能力,能够根据译者个性化的指令调整输出的译文,如例 2 指明需进行词汇、语序变换。如此一来,ChatGPT 可以根据译者导向提供准确的译文,给译者留出抉择的余地。而谷歌翻译仅给出了一种译法,一方面没有对“图腾老板”的概念进一步阐释,另一方面在修饰成分的处理上也略显平庸、不出彩,译后编辑可能会花费更多时间,译者有时不知如何下手修改。

NMT 较难提供多种译文主要因为其受制于较为固定的交互体系和运作机制。神经网络是一种基

于序列到序列模型(Sequence-to-sequence)的方法,完全采用神经网络完成源语言到目标语言的翻译过程<sup>[5,13]</sup>。除非针对原文进行微调,否则 NMT 给出的译文往往较为相似。其次,虽然 NMT 较以往而言在质量和效率上有了显著提升,但 NMT 产生的译文是基于训练数据的“二度摹仿”,同时受制于固定已有的算法架构,译文具有一定的“机械性”<sup>[1]</sup>。如例 1 中仅较机械地译出“企业攀登者”,没能译出“晋升”的含义,以及例 2 中没能针对“图腾老板”做出更深的阐释。而 ChatGPT 的译文多样性既避免了 NMT 机器译文的“机械性”,也为译者提供了选择的便捷。

## (二) 多种译文、不同风格

文本风格通常被认为产生于对常规的变异,翻译文本风格则是在对原语的风格意义进行分析的基础上获得译文风格对原文风格的“适应性”<sup>[14]</sup>。也就是说,译文风格的理想状态便是符合原文风格,人工翻译固然可以实现,但机器翻译处理风格仍旧较为困难。从宏观来说,文本风格在不同类型或领域上的差异很大,比如新闻文本简明直接,学术文本正式严谨;从微观上讲,不同写作主体在创作时会突显出较为一贯的文本风格,比如文学中的鲁迅文体、海明威风格等。与 NMT 相比,ChatGPT 在文本类型风格层面表现较为出色,见例 3。

**例 3 问** 翻译成汉语,分别以新闻风格、口语风格进行翻译: “Despite the company’s repeated assurances that its products are safe, recent investigations have uncovered numerous instances of potentially harmful chemicals and undisclosed ingredients in their manufacturing process, raising serious concerns among regulators and consumers alike.”

**答** 新闻风格: 尽管该公司多次保证其产品安全,但最近的调查发现其生产过程中存在大量潜在有害化学物质和未披露成分,这引起了监管机构和消费者的严重关注。

口语风格: 尽管这家公司一再保证他们的产品是安全的,但是最近的调查发现他们的制造过程中存在许多潜在有害的化学物质和未披露的成分,这让监管机构和消费者都感到非常担心。

例 3 显示 ChatGPT 人机对话式翻译能够为译者提供针对同一句话的不同文本风格的译文。新闻风格用词简洁正式,对关键信息翻译到位;口语风格则在词汇上有所改动,添加了日常用语中多见的

“的”,同时降低词汇难度,将“raising serious concerns”译为“非常担心”,而不是“严重关切”。

例4问 翻译成汉语,分别用鲁迅的散文和李白的古诗风格进行翻译:“Their eyes met in a sudden, electric moment that left them both breathless.”

答 鲁迅散文式:电光火石间,他们的目光交汇,顿时令他们俩骇然不已。

李白古诗式:两眼相触间,电火飞舞忙。惊魂乍现处,气馁俱踟躇。

例4中,虽不能完全理解所谓的“鲁迅风格”,但 ChatGPT 给出的回答还算差强人意,至少在词汇上进行了一定的模仿,如运用多个四字词语,体现出些许鲁迅白话文的色彩。相比之下,在李白古诗风格上,ChatGPT 处理得较好。译文形式为五言诗,在正确传达原意的同时,读起来也颇具意境和韵味,几可以假乱真。早前,国内研发团队已经发布了 AI 作诗工具,引发广泛关注<sup>④</sup>。但在机器翻译任务中,将原文以如此正确且仿真的形式翻译成为古诗,ChatGPT 已然成为行业翘楚。

针对例3和例4,谷歌翻译均只能提供一种译文(囿于文本篇幅,在此不再展示),因此也就不存在所谓的风格差异。从宏观上来说,针对不同文本类型的原文,NMT 引擎需要从训练数据和算法调优两方面入手,对文本类型进行风格上的调整,仅有少部分厂商提供细分文本领域的机器翻译服务;同时,NMT 引擎的算法和数据往往针对通用翻译任务,很难实现某位作者风格的定制化,因此目前也无法在微观上再现不同作家的风格。对于译者而言,掌握原文风格并在译文中充分复现是一项较为重要且困难的能力。借助 ChatGPT,译者能够对原文风格进行快速响应;而 NMT 给出的风格信息极为有限,很难帮助译者处理原文的风格要素。

### (三) 文本标记、轻松识别

元文本标记通常是指用于描述或解释文本内容的元数据或信息,往往是附加在文本之外的文字或标记,可以提供关于文本的背景、结构、作者等方面的信息。在网站开发中,元文本标记通常需要保留原文,如果被翻译成汉语,则无法发挥标记功能。因此,翻译含有元文本标记句子,对 NMT 和 ChatGPT 提出了挑战,见例5。

例5问 “进行英译汉翻译任务,第一遍识别元文本标记,说明是哪种标记类型,并按照元文本标记的风格进行翻译,第二遍保留标记并翻译:

<title> America's jobs boom rolls on, fuelling fears of sustained inflation. </title>”

答 第一遍识别标记并翻译:这是 HTML 标记语言中的<title>标记,用于定义标题。按照标题风格翻译为:美国就业激增引发持续通货膨胀担忧。

第二遍保留标记并翻译:<title> 美国就业激增引发持续通货膨胀担忧。 </title>

谷歌翻译 <title>美国的就业热潮不断,加剧了对持续通胀的担忧。 </标题>

通过例5可见,ChatGPT 能够识别元文本信息“<title> </title>”,并针对不同的元文本信息功能对原文的译法进行调整。例5指令中明确要在第一遍翻译中识别元文本标记,得到的译文首先指明这是超文本语言中的文章标题标记,在译法中就对应了新闻标题那种简短精悍、突出要点、一目了然的特点。其次,译者也可以提出保留元信息的要求,ChatGPT 一样能够识别,如例5中第二遍翻译。这样一来,网站开发者可以直接用 ChatGPT 得到的句子进行网页设计,不再需要转写。

NMT 在翻译词汇时能够做到较为正确,但却缺乏专业背景知识,不能理解和处理元文本标记这个概念,可能会将元文本标记中的词汇翻译出来,反而破坏其功能。同样是例5中的英文,谷歌翻译无法识别“</title>”的元文本标记功用,而是简单将其对待为一个单词去翻译,这样的句子就无法使用,必须将“</标题>”再还原成“</title>”才能继续用于网站开发。

总而言之,ChatGPT 作为大语言模型训练和人工强化学习的产物,在翻译任务中确有让人眼前一亮的表现。与 NMT 相比,ChatGPT 发挥了其自身人机对话的特点:可以针对同一原文产生多种译文,为译者提供了选择的便利;能够实现多种风格的处理,甚至可以模仿不同作者的写作风格;产生的译文也可以识别原文中的元文本标记,并根据译者的需求对译文进行灵活调整。然而,研究表明,ChatGPT 在对 NMT 构成严峻挑战的同时,仍存在一定的不足之处。

## 三、ChatGPT 有待提高的翻译质量

与传统的机器翻译工作机制形成鲜明对比的

<sup>④</sup> 国内人工智能作诗详见:<http://jiuge.thunlp.org> (访问日期:2023年10月26日)

ChatGPT 虽然在多重译文、风格复现、文本标记的处理上呈现出不同凡响之特征,但在翻译质量上目前仍有较大提升空间。下文将从完整性、准确性、一致性等方面展示 ChatGPT 相较于以谷歌翻译为代表的 NMT 所存在的局限。

### (一) 信息漏译、文不完整

翻译首先要忠实于原文的语义表达,信息漏译会导致重要的信息被省略或遗漏,从而使译文的完整性大打折扣,严重影响译文质量,也会导致译文读者对原文中的信息产生误解。在本研究中,ChatGPT 出现了几次关键信息漏译现象,见例 6、例 7。

**例 6 原文** Given the importance of *solar* to India's ambitions, this was ominous.

ChatGPT 对印度雄心勃勃的野心而言,这是不祥之兆。

**谷歌翻译** 鉴于太阳能对印度雄心壮志的重要性,这是不祥之兆。

例 6 中,我们发现 ChatGPT 漏译了原句中较为关键的“solar”一词,直接导致译句意义的转变。原句意思是太阳能资源对于印度的宏伟蓝图是非常重要的,但如果漏掉了 solar 这个关键词,就变成了印度本身野心勃勃,这显然会造成理解错误,这种翻译一旦发生在外交等重要场合,就会酿成大错。而谷歌翻译对原句的处理较好,甚至都不需要译后编辑。

**例 7 原文** But the main impact of the visit was symbolic, coming as it did just four days before the first anniversary of the full-scale invasion on *February 24th*.

ChatGPT 但是,此次访问的主要影响是象征性的,因为它是在全面入侵第一周年之前的四天。

**谷歌翻译** 但这次访问的主要影响是象征性的,就在 2 月 24 日全面入侵一周年的前四天。

例 7 中,同样的问题发生在日期的漏译上,ChatGPT 给出的译文漏掉了原文中的日期,这将会给读者带来不便,不清楚原文中所谓的一周年的起止是什么时候,而谷歌翻译不存在漏译的问题。

### (二) 词不达意、佶屈聱牙

评测机器翻译性能的重要指标是机器译文的质量,翻译错误会大大降低机器翻译的准确性,增加译后编辑人员的工作负荷<sup>[15]</sup>。请看以下例句中 ChatGPT 与 NMT 译文的对比。

**例 8 原文** That statistical difference contributed to an unusually steep fall in British national income compared with its *peers* during the pandemic; other economies may still be flattered by the comparison as a result.

ChatGPT 这种统计差异导致英国国民收入在大流行期间与其同行相比出现异常的急剧下降;其他经济体可能会通过这种比较而受益

**谷歌翻译** 在大流行期间,与其他国家相比,这种统计差异导致英国国民收入异常急剧下降;因此,其他经济体可能仍因这种比较而受宠若惊。

例 8 在词汇处理上,ChatGPT 并没能超越谷歌翻译为代表的 NMT。ChatGPT 将“peers”译为“同行”乍看似正确,但却忽略了上下文,误置了对比对象。这里对比的是国家之间的经济状况,因此谷歌翻译的“其他国家”更为准确。

**例 9 原文** *Far from* admitting their error, they doubled down and declared him fit for service.

ChatGPT 远离承认他们的错误,他们强化声明,宣称他适合服务。

**谷歌翻译** 他们非但没有承认错误,反而加倍努力,宣布他适合服役。

例 9 中的“far from”是一个经常使用的短语,放在句首表示“非但没有……还……”的意思,此处 ChatGPT 就单纯翻译出了字面含义,翻译成“远离”,这与原文想要表达的意思相去甚远,造成了意义翻译错误。

在句式上,由于中文和英文在句式表达上存在差异,译者需要根据具体情况对译文句式进行调整。以往的统计机器翻译多基于平行语料库和概率算法进行翻译,常会在英译汉过程中造成译文句式上的成分错位或过于西化等问题<sup>[16]</sup>。而在目前的 NMT 系统中,这种现象得到了明显的改善。

**例 10 原文** *Not every one* of these trends will last.

ChatGPT 不是每一个这些趋势都会持久。

**谷歌翻译** 并非所有这些趋势都会持续下去。

上面例子中,原文存在一个较长的主语,且 not every one 表示部分否定的意思,应当译为“并非所有”,因此 ChatGPT 的译文直接分别翻译了 not 和 every one 的意思,没有经过句式上的调整,读起来就会显得佶屈聱牙,过于西化,相反谷歌翻译表现较好。



### (三) 专有名词、前后不一

译文前后一致性关乎整体译文的质量, 人名或术语等专有名词的翻译前后不一致是较为明显的错

误之一, 会直接导致读者错误领会上下译文的意思。研究发现, ChatGPT 在执行机器翻译时, 对于人名或术语翻译的前后一致性表现较为逊色<sup>⑤</sup>。举例见表 1。

表 1 信息一致性案例分析

Tab. 1 Case analysis of information consistency

人名一致性			术语一致性		
原文	ChatGPT	谷歌翻译	原文	ChatGPT	谷歌翻译
Jochen Buck and Lonny Levin, who work at ...	Buck	Buck	These are called silent synapses.	沉默突触	沉默突触
Dr Buck and Dr Levin say this is not a cause for concern.	布克博士	Buck	Silent synapses —which, as their name suggests ...	无声突触	沉默突触
Whether Dr Buck's and Dr Levin's discovery will ...	桶博士	Buck	Just over a quarter of the connections ... were silent synapses.	静默突触	沉默突触

人名的处理需要遵循一致性原则。虽然人名翻译并无确定唯一的译法, 如“Mary”可被译为“玛丽”或者“玛莉”, 但为避免混淆, 应在上下文统一使用选定的译法。在本研究中, 如表 1 左侧所示, 人名如“Dr. Buck”在原文上下文中出现了三次, ChatGPT 译文对应出现了三个不同的译法, “Buck”

“布克博士”“桶博士”, 如不进行译后编辑, 读者很容易误认为这是三个不同的人, 更不用说第三种译法直接将人名当作一个普通名词来译, 这是错误的; 而谷歌翻译的译文统一保留了原文, 译成“Buck”, 虽然没音译成汉语读者多见的“巴克”, 但至少做到了上下文统一, 不会给读者带来歧义。

术语翻译同样需要保证前后一致, 确保译文连贯和准确。尤其在翻译大型文档或专业文件时, 保持术语一致应贯穿整个项目始终。本研究发现, ChatGPT 在同一专有名词的翻译上前后一致略逊于 NMT (见表 1)。比如术语“silent synapses”, 指未激活的神经突触, 专业医学术语称之为“沉默突触”。ChatGPT 在上下文中出现了三种不同的译法:

“沉默突触”“无声突触”“静默突触”。虽然“沉默、无声、静默”都是“silent”的正确意思, 但术语需保持前后一致性才不会让读者产生困惑。这一点谷歌翻译做到了。ChatGPT 对同一原文给出不同的译文表达, 的确增加了译文灵活性, 但反之也降低了译文一致性, 增加译后编辑工作量, 造成额外时间消耗。在这一点上, NMT 引擎做得较好, 如若额外添加了术语表的功能, 就能在确保术语的前后一致性的同时, 做到更为地道的目标语转换。

综上, 我们发现与 NMT 引擎相比, ChatG-

PT 确实存在一定的局限性, 如: 在译文准确性上, 会出现误译和句式直译问题, 造成事实性错误; 在信息完整性上, 译文中存在漏译现象, 影响原文意义表达; 在术语一致性上, 专有名词的前后一致性较弱, 造成读者误解。值得一提的是, 以上提及的各类翻译错误在 ChatGPT 中并未频繁出现, 我们只是选取了较为典型的案例。通过具体翻译实践, 译者能够对 ChatGPT 和 NMT 的实际表现有较为深入的了解, 也能够后续使用 ChatGPT 进行翻译任务时采取相应的翻译或译后编辑策略。

## 四、人机对话翻译带来的启示

人工智能技术就如同一个快速迭代、学科融合、注重实践的大浪潮, 避而驱之绝不可行, 偃旗息鼓亦不可为。在翻译实践、翻译学习、技术研发中, ChatGPT 和未来的诸多人机对话式 AI 产品给我们提供了一定的启示。

### (一) 对于翻译从业者而言: 正视挑战、为我所用

ChatGPT 虽在执行翻译任务上表现优异, 但并不是完美无瑕、毫无破绽的超级机器。通过本研究的案例, 能够看出 ChatGPT 在翻译中仍会产生较多错误。这说明对机器译文的人工介入仍是必不可少

<sup>⑤</sup> 目前学界对于人名是否应纳入术语的范畴仍存在争议, 有研究将人名纳入术语的范畴<sup>[30-31]</sup>, 另有研究称术语不包括人名<sup>[32]</sup>。本文参考现行国家标准《术语工作 原则与方法》(标准号 GB/T 10112—2019) 中规定的术语应具有“单名单义性”, 采纳第二种观点, 将人名和术语区别开来进行研究。



的,“人工智能取代译者”的论调还言之尚早。不管是对于常规语言还是陌生化语言,机器译文目前还很难做到一步到位,必须有译后编辑参与<sup>[17]</sup>,译者仍需不断提高翻译和译后编辑技能,及时发现机器译文当中的错误。

研究发现,ChatGPT和NMT在翻译实践中各有利弊,译者应该知晓两种技术之间存在的异同,在翻译过程中灵活运用不同的机器翻译工具,充分发挥其自身的特点,达到良好的翻译效果。例如:在翻译文学、散文等风格变化较大的文本时,译者可选择ChatGPT并参考本文例1至例4中的指令,利用模型能够适应不同文体风格译文的特点,助力特定文本类型的翻译;而在翻译精度要求高、数据繁多、术语数量大的文本时,为避免ChatGPT灵活性过大而导致的误译、漏译或一致性差等弊端,应选取结合术语表或垂直领域的NMT引擎,并进行适当的译后编辑。

另外,ChatGPT最大的特点是人机对话、交互性强,其回答质量与译者的指令息息相关,给出的译文也不具有复现性,这就需要译者不断训练模型,在提高“搜商”的同时,还要提高“问商”,提高人机交互效率,使其提供更加准确的译文。译者同样要探索NMT和ChatGPT二者潜在的交互性,后者可为NMT译文提供修改建议并改善整体翻译的流畅性和连贯性,从而帮助译员加快译后编辑过程、提高译文整体质量。在此过程中,译者也要强化在数据采集、处理、分析、共享、保护和利用等方面的能力,提高个人数据权益和数据伦理意识<sup>[18]</sup>。

最后,ChatGPT绝对不是最终的划时代人工智能产品,未来随着科技发展,还会出现更多、更新、更强的产品,译者需要做到在态度上拥抱技术、在学习中享受技术、在头脑中思考技术。一味的抵触和“摆烂”实际上都无济于事,技术的车轮滚滚而来,“顺之者昌、逆之者亡”。“优秀的译者应借助机器寻找意义”<sup>[19]</sup>,要争取利用技术不断去挖掘新的资源,而不是被技术裹挟进而变成技术的附庸。

## (二) 对于翻译学习者而言:学而用之、勿入歧途

如今,随着学界和产业界对翻译技术关注的不断升温,各大高校开始充分利用高校特色跟进和更新针对MTI的翻译技术课程教育<sup>[20-21]</sup>,建立翻译技术人才能力培养框架和模型<sup>[22-23]</sup>,这一系列努力都有助于培养翻译学习者正确对待和使用人工智能的意识和能力。因此有学者称,在人工智能时代,

翻译技术教学大有可为<sup>[24]</sup>。

对于翻译学习者来说,ChatGPT能够让其跳出课堂教育的框架,可以随时随地进行“泛在学习”<sup>[25]</sup>。同时,ChatGPT能够提供翻译练习、译文打分、意见反馈,还能提供学习资源和工具等。可见,人工智能已经从感知智能进入了认知智能,我们进入了一个人机合作的新时代<sup>[26]</sup>。在此背景下,翻译学习者的自学能力将是成功的关键,要时刻考虑这几个问题:用不用、何时用、怎么用。在专业技术技能上打好基础,充分利用好人工智能技术和资源,实现所谓的“弯道超车”。

值得一提的是,翻译学习者在享受ChatGPT带来的快捷和便利的同时,要树立正确的“技术观”,适度借助技术进行翻译学习和实践,把握技术和人文之间的平衡,这在学习阶段尤为重要。一方面,我们不能过分强调技术,完全依赖技术,这不但会造成“人工翻译无用论”思想,还会导致“译者身份异化”,带来思维模式化、单一化、机械化的不良影响<sup>[27]</sup>,要充分认清ChatGPT在翻译中表现出的劣势,提高给机器“挑错”的能力,既能够改正较为明显的译文不准确现象,也能够及时发现较为隐蔽的信息不完整、前后不一致等现象;另一方面,过度排斥技术同样不可取,翻译不再是过去依赖纸笔的时代,语言服务企业追求效率和利益,需要翻译技术的加持。对于学习者而言,从计算机辅助翻译,到NMT引擎,再到如今的ChatGPT,技术的更新迭代极其迅速,“一步跟不上,步步跟不上”,只有积极拥抱技术、学习技术,树立人机耦合的意识,构建人机交互、人机协同、人机共生的学习理念<sup>[28]</sup>,才能不落后于时代。

## (三) 对于技术开发者而言:迭代更新、以人为本

通过实践,我们发现将ChatGPT应用于机器翻译任务,仍在以下两个方面有待提升。首先,虽然在人工反馈强化学习模式下,ChatGPT会自动规避生成一些有悖道德、常识、意识形态方面的错误或有害信息,但由于“语言污染”或数据偏差,无法完全避免偏见和道德伦理不正确的信息,这对机器翻译伦理的忠实和忠诚原则提出了挑战<sup>[29]</sup>。如果在译文中出现此类差错,不仅增加了译者进行译后编辑的工作负荷,一旦漏改,也会对译文质量带来致命的影响。

另一方面,ChatGPT在集成翻译平台(如CAT工具)上的效果较差,这对当前CAT的开发

提出了挑战,即如何在编辑器界面融入 ChatGPT 的交互式功能。现如今的 CAT 工具(Trados、Memoq 等)多将机器翻译作为插件集成入编辑器,机器翻译提供“答案”,用户进行“批改”,NMT 的“答案”通常是唯一、固定的,而 ChatGPT 具有交互性,可根据原文特征提供多样、灵活的“答案”,并接受用户反馈和再交互。这要求开发者针对 ChatGPT 的特点进行改进,从人机工效学的角度调整编辑器交互界面,充分结合 ChatGPT 的交互功能和多语言处理能力,真正做到“以人为本”,让技术更好地服务于译员。

## 五、结语

本文通过案例分析展示,与谷歌翻译相比 ChatGPT 在译文多样性、风格变化和标记识别方面表现出潜力,但在准确性、完整性和一致性方面还有待提高。ChatGPT 已经打开了“潘多拉魔盒”,不论是翻译从业者、学习者,还是技术开发者,都要适应变化、及时调整、随机应变,充分利用好人工智能发展红利,进一步为翻译实践和翻译教学增添新的动力。

### 参考文献:

- [1] 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(5):10-14.
- [2] 严辰松.机器翻译取代人工翻译还只是愿景——机译英译汉实例点评[J].翻译教学与研究,2021(2):1-12.
- [3] 冯志伟,张灯柯.机器翻译与人工翻译相辅相成[J].外国语,2022,45(6):77-87.
- [4] 刘洋.神经机器翻译前沿进展[J].计算机研究与发展,2017,54(6):1144-1149.
- [5] 宗成庆.人类语言技术展望[J].中国人工智能学会通讯,2020,10(1):2-3.
- [6] 李梅.机器翻译译后编辑过程中原文对译员影响研究[J].外语教学,2021,42(4):93-99.
- [7] PEDERSEN F. EasyTranslate raises Eur 3m to scale ChatGPT content generation tool [EB/OL]. (2023-02-01)[2023-10-26].<https://slator.com/easytranslate-raises-3m-build-chatgpt-content-generation-tool>.
- [8] MIN B, ROSS H, SULEM E, et al. Recent advances in natural language processing via large pre-trained language models: a survey[J]. ACM Computing Surveys, 2023, 56(1): 1-12.
- [9] RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding by generative pre-training[J/OL]. [2023-10-26] <https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf>.
- [10] JIAO W X, WANG W X, HUANG J T, et al. Is ChatGPT a good translator? Yes with GPT-4 as the engine [EB/OL]. (2023-11-02) [2023-10-26]. <https://arxiv.org/abs/2301.08745>.
- [11] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2020.
- [12] OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback[C]//Proceedings of the 36th Conference on Neural Information Processing Systems. New Orleans: NeurIPS, 2022.
- [13] 李亚超,熊德意,张民.神经机器翻译综述[J].计算机学报,2018,41(12):2734-2755.
- [14] 刘宓庆.翻译的风格论(上)[J].外国语,1990(1):3-7;41.
- [15] JIA Y F, ZHENG B H. The interaction effect between source text complexity and machine translation quality on the task difficulty of NMT post-editing from English to Chinese: a multi-method study[J]. Across Languages and Cultures, 2022, 23(1): 36-55.
- [16] 李梅,朱锡明.英汉机译错误分类及数据统计分析[J].上海理工大学学报(社会科学版),2013,35(3):201-207.
- [17] 杨文地,范梓锐.科技语篇机器翻译的译后编辑例析[J].上海翻译,2021(6):54-59.
- [18] 王华树,刘世界.新文科背景下翻译技术核心课程教学模式探索[J].外语教学,2022,43(6):57-62.
- [19] 杜磊,刘和平.“优秀的译者应借助机器寻找意义”——刘和平教授访谈[J].东方翻译,2020(6):48-57.
- [20] 李梅.信息时代的“项目+技术”二轮驱动特色翻译教学[J].外国语,2017,40(5):99-101.
- [21] 王华树,刘世界.大数据时代翻译数据伦理研究:概念、问题与建议[J].上海翻译,2022(2):12-17.

- [22] 冯全功, 刘明. 译后编辑能力三维模型构建 [J]. 外语界, 2018 (3): 55-61.
- [23] 崔启亮. AI时代的译者技术应用能力研究 [J]. 外国语言与文化, 2020, 4 (1): 104-110.
- [24] 陶友兰. 中国翻译技术教学研究三十年 (1990—2020)[J]. 上海翻译, 2023 (2): 49-54.
- [25] 刘芹. 泛在学习视域下的理工科学术英语动态评估模型构建 [J]. 上海理工大学学报(社会科学版), 2021, 43 (1): 8-15.
- [26] 冯志伟, 张灯柯, 饶高琦. 从图灵测试到 ChatGPT——人机对话的里程碑及启示 [J]. 语言战略研究, 2023, 8 (2): 20-24.
- [27] 王贇, 张政. 数字化时代机器翻译的风险审视及控制研究 [J]. 中国翻译, 2022, 43 (2): 109-115.
- [28] 秦颖. 人机共生场景下的外语教学方法探索——以 ChatGPT 为例 [J]. 外语电化教学, 2023 (2): 24-29.
- [29] 任文. 机器翻译伦理的挑战与导向 [J]. 上海翻译, 2019 (5): 46-52.
- [30] 顾春辉, 温昌斌. 联合国术语库建设及其对中国术语库建设的启示 [J]. 中国科技术语, 2017, 19(3): 5-9; 34.
- [31] 刘力力, 朱建平, 高新颜. 论中医药英文术语规范的一致性原则 [J]. 中国科技术语, 2013, 15(1): 38-40; 45.
- [32] 马莲. “术语”“名词”“名词术语”辨析 [J]. 新闻与传播研究, 2016 (9): 119-125.

(责编: 朱渭波)