

## 【人工智能赋能中医药教育专题】

## 中医药英文论文写作机器翻译质量对比研究\*

邵英俊 张烁 孔祥国# 熊淋宵

(北京中医药大学人文学院 北京 102488)

**摘要:**在机器翻译辅助下,中医药英文论文的翻译效率虽有所提高,但译文仍存在诸多不足。研究通过定性、定量对比百度翻译、DeepL、有道翻译和 ChatGPT 4.0 在处理中医药论文英译中的优劣,并进行全方位质量评估和对比分析,揭示目前应用机器翻译工具辅助中医药英文论文写作时,须重点关注机器翻译工具译文的句式构建、术语英译的准确性和规范性,以及源自中医古籍的文言表达等问题。同时,在译后编辑过程中,应使用多种工具,并参照权威中医药英译国际标准及权威英译本,以确保英文翻译的规范性和准确性。

**关键词:**机器翻译;中医药英文论文写作;译文质量评估

中图分类号:G642.41

doi:10.3969/j.issn.1003-305X.2026.03.366

## A comparative study on the quality of machine translation in English paper writing for traditional Chinese medicine\*

SHAO Yingjun, ZHANG Shuo, KONG Xiangguo#, XIONG Linxiao

(The School of Humanities, Beijing University of Chinese Medicine, Beijing 102488, China)

**Abstract:** The writing of English papers on Traditional Chinese Medicine (TCM) with the assistance of machine translation improves the efficiency of English translation, but there are still many deficiencies in the translated texts. This study qualitatively and quantitatively compares the advantages and disadvantages of Baidu Translate, DeepL, Youdao Translate, and ChatGPT 4.0 in handling the English translation of TCM papers, through comprehensive quality evaluation and comparative analysis, it reveals that attention should be focused on issues such as the sentence structure construction of the translated texts of machine translation tools, the accuracy and standardization of the English translation of terms, and the classical Chinese expressions derived from ancient TCM books. At the same time, multiple tools and authoritative international standards for the English translation of TCM should be used to ensure the standardization and accuracy of the English translation in the post-editing process.

**Keywords:** machine translation; TCM English paper writing; translation quality evaluation

国际期刊是中医药文化传播的重要媒介,中医药优秀学术成果有赖于中医药学者通过发表国际期刊论文向世界展示<sup>[1]</sup>。这些论文在很大程度上提升了中医药的国际知名度,同时,也让越来越多的西方民众看到中医药的科学性和有效性。

朱梦梦等<sup>[2]</sup>针对中医药研究生国际期刊论文写作与发表的现状展开问卷调查与分析,结果表明,

在列出的15个不同的写作困难原因中,选择“英语水平低”的人数占比最多。另一项基于问卷调查的研究发现,美国医学类期刊主编认为,中国作者发表的英文论文中存在的问题远多于优势。65.4%的主编将语言列为最常见的问题,他们普遍的评价是论文语言不符合英语语法的表达方式。这是中国作者论文的最突出的短板。此外,语言很差的论

邵英俊,男,硕士,副教授,硕士生导师,E-mail:15210369502@163.com

#通信作者:孔祥国,男,硕士,教授,硕士生导师,E-mail:13810283596@163.com

\*基金项目:教育部产学合作协同育人项目(No. 231101339071101);北京中医药大学社科培育课题(No. 2024-JYB-900208-011)

文往往很难使审稿人给出应有的评价<sup>[4-6]</sup>。由此可见,中医药论文英文写作中的语言表达和运用一定程度上影响文章的发表,且对很多人而言依然还是一个较大的挑战。

机器翻译的迅猛发展,尤其是人工智能(artificial intelligence, AI)技术在机器翻译领域的应用为翻译领域带来深刻的变革。虽然随着机器翻译工具越来越多地应用于中医药英文论文写作中,中医药英文论文写作效率得到较大地提高,但是由于中医药独特的理论体系、术语表达和语句结构,机器翻译产出的中医药内容的英文译文质量往往不尽如人意。因此,有必要在积极应用不断更新的机器翻译工具和技术的同时,探讨并揭示其存在的不足,以便为机器翻译辅助下的中医药英文论文写作提出建设性的意见。

近年来,越来越多的学者将目光转向中医药文本的机器辅助翻译。王南鑫和云红<sup>[7]</sup>探讨中医药说明书中功效主治部分机器辅助翻译的新模式,即译前编辑+机器翻译+机辅工具+译后编辑,并探索该机器翻译模式应用于中医药功效、主治类文本的可行性。宋熹月等<sup>[8]</sup>聚焦中医古籍智能机器翻译模型的构建,运用大规模预训练语言模型,构建基于Seq2Seq的机器翻译模型Pre-training+Seq2Seq,并分析该模型应用于中医古籍智能翻译的可行性。曹媛和马永良<sup>[9]</sup>以北京中医药信息网新闻稿英译为例,探讨通过译前编辑优化谷歌翻译译文质量的方法和途径。

关于具体机器翻译工具在中医药文本翻译中的应用及错误分析也逐渐成为研究热点。蒙洁琼<sup>[10]</sup>在机译错误分析的基础上分析使用谷歌翻译和百度翻译处理中医药论文结构性摘要英译的策略。温海煜<sup>[11]</sup>分析谷歌翻译和百度翻译处理中医药文本英译中的错误类型和特点,并进行数据分析,进而提出相应的机器翻译译后编辑建议。以上研究多涉及机器翻译辅助中医药文本翻译中的特定文本翻译模式研究、翻译模型构建及应用、翻译工具的对比,错误归类分析等。目前,关于中医药英文论文写作中机器翻译译文质量评价的相关研究仍较为缺乏。

本研究选用中医药论文相关语料,分别使用百度翻译、有道翻译、DeepL和ChatGPT 4.0将语料译成英文,将不同工具生成的译文进行定量、定性对比分析。运用Python对译文的BLEU值进行计算,并辅以人工评分,旨在从主客观,多角度评估不同

工具在处理中医药论文中文语句英译中的优势和局限,针对不同翻译工具在中医药英文论文写作中产出的译文质量进行对比评估,有望为机器翻译辅助中医药学术文本翻译研究提供借鉴和启示,进而助力中医药国际传播。

## 1 资料和方法

### 1.1 研究问题

本文将聚焦中医药英文论文写作中应用常用机器翻译工具翻译中医药特色语句过程中存在的问题,探讨不同翻译工具辅助中医药英文论文写作的译文质量评估方法,在翻译或者译文生成过程中存在的具体问题,以及未来机器翻译辅助中医药英文论文写作的应用前景。

### 1.2 文本选取

研究在语料选取上,主要聚焦于中医领域的T1级科技核心期刊,同时涵盖中英文类别的期刊。经过文献检索及阅读,明确SCI期刊中医药相关英文学术论文主要在引言、讨论、结论等部分较多地使用中医药特色表述,语句内容以介绍中医相关的基本概念、病症、草药、方剂、特色疗法等为主。因此,根据中医药学科的特色精心筛选语料,包括:中医经典古籍、中医基础理论(如辨证分型、病因病机、中医体质分类、中医证候等)、中医学(含中药学、方剂配伍等)及针灸学(涵盖各类针灸技术、穴位选择及其作用机理等)。这些内容在英文文献的理论阐述中占据重要地位,有助于确保语料选择的全面性和代表性,进而更深入地探究机器翻译工具在处理各类语料时的性能表现。最终,本研究从10篇中医药论文中选取10个代表性汉语语句,所选取语句从内容上涵盖中医经典、中医辨证分型、中医病因病机、中医体质分型、草药性味功效、针刺穴位选取等中医药特色表达,基本涵盖中医药英文论文写作中较频繁使用的术语类型和句式。值得说明的是,本文未选取已公开发表的中医药英文论文作为研究语料来源,理由是此类已发表的英文论文的中文原文无法获取,亦或并无对应中文原文。本研究纳入10个语句,共458个中文字符,所选例句原文具体见表1。

### 1.3 研究方法

BLEU(Bilingual Evaluation Understudy)是一种广泛使用的机器翻译质量自动评价指标,最早由IBM公司在2002年提出。该方法的核心思想是利用N-gram匹配和惩罚因子对机器翻译和人工译文相似度进行计算,二者的相似度越高,译文质量就

表1 例句原文及出处

序号	例句原文	文章标题	期刊名称
1	肾为先天,脾为后天,精气得脾运化之精微以滋养,方能“精气溢泻”而繁衍后代	基于铁死亡探讨参苓白术散对高尿酸血症小鼠睾丸生精功能的改善作用	《北京中医药大学学报》
2	痰湿体质是由于津液运化失常导致痰湿凝聚,以重浊黏滞为主要特征的体质状态,其形成与先天遗传或后天饮食偏嗜、过食肥甘等有关	骨不连与中医体质的相关性研究及“治未病”理念的应用探讨	《国际中医中药杂志》
3	MIRI在中医学可归属于“胸痹”范畴,病机多为气血不足、阴阳亏虚、气虚血瘀、痰瘀阻络	基于血清代谢组学探究蛭龙活血通瘀胶囊对心肌缺血再灌注损伤小鼠的保护作用及机制	《北京中医药大学学报》
4	太冲为肝经原穴,有研究表明,针刺太冲能够影响包括小脑在内的不同区域脑功能,保持身体平衡,调节肌张力及运动功能,发挥平肝潜阳、息风止痉的作用	电针对帕金森病模型小鼠运动功能的影响及相关分子机制探讨	《北京中医药大学学报》
5	方中葛花味甘、微辛,性微凉,归脾、胃、肝经,具有醒酒解毒、生津止渴之功,故立为君药	葛花解酒涤脂汤对酒精性肝病小鼠肝脏组织能量代谢的影响	《北京中医药大学学报》
6	三阴交属足太阴脾经,可健脾疏肝、行气补血	灸法改善化疗所致骨髓抑制的研究进展	《针刺研究》
7	痰热互结证:形体肥胖,腹部胀大,口干口渴,喜冷饮,饮水量多,脘腹胀满,易饥多食,心烦口苦,大便干结,小便色黄,舌质淡红,苔黄腻,口干脉弦滑	2型糖尿病患者中医证型与血清代谢学指标相关性研究	《中医杂志》
8	此方君臣佐使,九味为伍,共建健脾渗湿、清热解毒、疏肝解郁、消食理气、化痰通络、和胃降浊、生津止渴之功	葛花解酒涤脂汤对酒精性肝病小鼠肝脏组织能量代谢的影响	《北京中医药大学学报》
9	高尿酸未出现临床症状,如痛风性关节炎、痛风性肾病时,多归属于中医学“湿浊”范畴	基于铁死亡探讨参苓白术散对高尿酸血症小鼠睾丸生精功能的改善作用	《北京中医药大学学报》
10	对于寒湿型RA患者,多选择温经散寒通络的穴位,如足三里、阳陵泉等	针灸治疗类风湿关节炎临床及机制研究进展	《国际中医中药杂志》

越好<sup>[12]</sup>。其得分通常为0~1分之间,数值越高,表明译文质量越好。

首先,课题组邀请中医药英译领域的专家给出以上10个例句的参考译文,经团队讨论、复核后确定译文终稿,最后,由笔者校对整理后作为本研究中例句的对应译文,作为计算机器翻译译文各自的BLEU值计算的依据,所得数值作为评估不同机器翻译工具译文的指标之一。

鉴于BLEU评价指标的局限性,同时设计译文质量人工评分,邀请2名中医药翻译领域资深教师对机器翻译生成10个例句的译文分别进行打分,打分的全过程,教师不知道译文的来源。根据BLEU计算得出的译文评分和人工评分,经过深入的分

析,综合评价不同工具产出译文中存在的问题。

## 2 结果

### 2.1 BLEU值评分

分别将10个例句输入“百度翻译”“有道翻译”“DeepL”即时生成英文译文,后经Python计算得出每一个译文句子的BLEU值。在使用ChatGPT 4.0处理例句的过程中,输入统一的指令(prompt)——“Please provide the English translation for the sentence”,而后同样经过Python计算得出对应译文的BLEU值。以上不同版本的译文各自对应的BLEU值,课题组拟从几个维度结合不同工具生成译文的BLEU值分析不同机器翻译工具得分情况。具体见表2。

表2 译文质量BLEU评分(%)

翻译工具	例1	例2	例3	例4	例5	例6	例7	例8	例9	例10
百度翻译	11.97	37.55	36.85	53.24	39.26	55.44	31.79	28.77	49.44	40.89
有道翻译	15.08	50.84	41.86	58.40	33.74	37.99	56.22	28.37	56.70	19.42
DeepL	13.05	42.21	30.28	50.49	38.86	33.79	46.12	43.48	30.18	47.71
ChatGPT 4.0	35.22	89.43	100.00	88.84	87.68	94.88	89.29	96.08	95.67	95.81

周成彬和刘忠宝<sup>[13]</sup>认为,一个翻译工具生成译文的 BLEU 值达到 31.4%,就表明译文质量较好,可达到机器翻译的基本要求。基于表 2 中不同工具产出译文的 BLEU 值,计算不同工具处理后译文 BLEU 值的平均值。经计算,ChatGPT 4.0 产出译文的 BLEU 平均值最高,达到 87.29%。有道翻译产出译文的 BLEU 平均值为 41.80%,百度翻译产出译文的 BLEU 平均值为 38.53%,DeepL 产出译文 BLEU 平均值最低,为 37.62%。应该说这 4 种不同的翻译工具在所选 10 个例句的翻译中都达到基本要求,均完整地将中文原文翻译成英文。其中,ChatGPT 4.0 在本研究所选 10 个例句的翻译中表现优异,在很大程度上较充分地满足翻译的需求。然而,除 ChatGPT 4.0 外,另外 3 种翻译工具在流畅性、达意度等方面都还有较大提升空间。

除以上不同翻译工具处理 10 个例句的平均得分外,也很有必要关注整体得分分布中的最高分和最低分。从研究中不难发现,百度翻译处理的例句 1 的 BLEU 值为 11.97%,为所有得分中的最低分。ChatGPT 4.0 处理的例句 3 的得分达到 100%,与人工翻译的译文完全一致,达到相应例句翻译需求,甚至无需译后编辑,这也基本符合本研究中的整体得分情况,ChatGPT 4.0 处理的所有例句的翻译质量平均 BLEU 值达 87.29%。

文本的整体特点一定程度上决定翻译的难度,其中就包含句式特点、术语的数量、文化负载词及概念隐喻等。课题组尝试从表 1、表 2 中挑选得分总分最高的例句和得分总分最低的例句,进而分析不同例句的语言特点。经计算,例句 1 的 BLEU 值得分总分最低,为 75.32%,就连 ChatGPT 4.0 的得分只有 35.22%,勉强达到机器翻译的基本要求。而得

分总分最高的例句为例 4,达到 250.97%。进一步分析发现,例句 1 中的相关表述如“肾为先天,脾为后天”源于李中梓《医宗必读·肾为先天本脾为后天本论》,而“精气溢泻”源于《素问·上古天真论篇》。虽然后半句在古籍原文的基础上有所改动,但整体语言风格保留了中医典籍中的古汉语表达,语言高度凝练、概括。由此可以看出,本研究中包括 ChatGPT4.0 在内的 4 种机器翻译工具在处理古汉语表达的语句时,整体可靠性和准确性较低。例句 4 的 BLEU 值总得分最高,其语言内容为相关穴位针刺的治疗作用,其中多数表述为西医学内容。除句尾 2 个四字格结构以外,语言风格上亦为现代汉语,因而该例句 BLEU 值得分总分最高。

以上分析表明,ChatGPT4.0 辅助中医药英文论文写作中的文本翻译整体上优势明显。而对于中医药英文论文写作中引自中医典籍的古汉语表达,所有工具都存在明显的局限性。

值得注意的是,虽然 BLEU 指标作为自动评价工具经常用来评估机器翻译译文质量,该指标主要是通过与参考译文比对来评估整体系统性能,而非关注单个翻译案例,语言表达中的复杂性和细微差别往往在评估中得不到体现,且翻译本无定论,不同的参考译文也一定会导致 BLEU 值出现变化,因此,一个较低的 BLEU 值不一定意味着完全无效或者完全不准确的翻译。为了更加全面准确地评估机器翻译生成的译文,课题组邀请从事中医英语翻译的一线教师针对每一个例句不同工具生成的译文进行人工评分,以便更好地弥补 BLEU 自动评价工具在文化适应性,语言风格等方面可能的不足。2 位专家给出的人工评分,具体见表 3。

表 3 译文质量人工评分(分)

项目	有道翻译		DeepL		百度翻译		ChatGPT4.0	
	评分一	评分二	评分一	评分二	评分一	评分二	评分一	评分二
例句 1	84.0	80.0	56.0	61.0	58.0	63.0	92.0	90.0
例句 2	71.0	73.0	76.0	80.0	70.0	68.0	95.0	93.0
例句 3	78.0	76.0	76.0	79.0	73.0	75.0	100.0	100.0
例句 4	89.0	90.0	82.0	84.0	90.0	91.0	93.0	94.0
例句 5	89.0	87.0	81.0	80.0	80.0	82.0	93.0	95.0
例句 6	81.0	83.0	85.0	87.0	83.0	85.0	97.0	95.0
例句 7	86.0	83.0	87.0	84.0	86.0	88.0	96.0	98.0
例句 8	84.0	81.0	78.0	77.0	81.0	79.0	97.0	95.0
例句 9	87.0	85.0	86.0	82.0	86.0	88.0	94.0	96.0
例句 10	90.0	92.0	88.0	90.0	91.0	92.0	94.0	96.0
平均得分	83.9	83.0	79.5	80.4	79.8	81.1	95.1	95.2

## 2.2 人工评分

在评分过程中,2位评审专家并不了解译文的来源,以此保证评分的客观性。从第一组人工评分的情况看,ChatGPT 4.0平均分最高,为95.1分,然后是有道翻译和百度翻译,得分分别为83.9分和79.8分,DeepL的得分最低为79.5分。第二组人工评分从高到低分别为ChatGPT 4.0(95.2分),有道翻译(83分),百度翻译(81.1分)和DeepL(80.4分)。虽然2组评分存在差异,但整体上不同翻译工具得分的排序是一致的,即2位评审专家均认为ChatGPT 4.0在处理所选例句的翻译时表现最佳,而DeepL表现最差。

## 2.3 ChatGPT 4.0在翻译本文所选例句中的亮点

### 2.3.1 句式构建

例句1阐释中医学对于脾肾2个脏器重要性的认识,ChatGPT 4.0的译文为:The kidneys are considered the foundation of congenital constitution, while the spleen is the foundation of postnatal constitution. The essence chamber is nourished by the refined substances produced by the spleen's transportation and transformation, allowing "essence and qi to overflow", which enables reproduction and the continuation of life. 此译文运用1个简单句和1个从句,很好地阐明了源语所表达的内容,第一句译文中通过while很好地阐述两相对照的关系,第二句则巧妙地运用"allowing……"及which引导的从句层次分明,简洁、易懂地阐明脾胃运化及生育繁殖之间的关系。

例句2是关于中医体质学中寒湿体质特征和形成原因的简要描述,ChatGPT 4.0的译文中相应的分割成3个独立的句子:① Phlegm-dampness constitution is a physical state characterized by the accumulation of phlegm and dampness due to abnormal fluid metabolism. ② It is primarily marked by heaviness, turbidity, and stickiness. ③ Its formation is related to congenital factors or acquired dietary habits, such as an excessive intake of rich, greasy foods. 虽然译文中仍略有不足,但整体上语义层次清晰。尤其后面2个句子添加代词it作为主语,断句恰当,简单易懂。

### 2.3.2 术语英译准确性及规范性

例句1中的“先天”和“后天”,有道翻译分别误译成“original”和“acquired”,百度翻译分别误译成“congenital”和“acquired”,以上2组译文虽然在一定程度上接近原文语义,但明显有失准确,且过于抽

象。ChatGPT4.0则分别译作“foundation of congenital constitution”和“foundation of postnatal constitution”,虽然略显冗长,且并非标准译文,但可以较为准确地阐明中医学对这2个脏器重要性的认知及二者之间在这个语境中的关联性。

例句4中提到“太冲”这一穴位,有道翻译、百度翻译和DeepL均直接译为拼音,而ChatGPT 4.0则在译为Taichong(LR3),在拼音后补充国际通行的数字+字母组合,这在论文写作中术语使用的规范性上非常重要。同样地,在例句5中,“葛花”一词ChatGPT在翻译的时候译做Kudzu flower(葛花, Puerariae Flos),保留这一草药的法定药材名,中文及植物学拉丁文表达,而其他3个工具均只给出一种表达。虽然不同学术刊物对中草药的名称标注要求不同,但大多都要求提供包括法定药材名,植物学拉丁文及汉语拼音在内的至少2种表述,以确保准确性和规范性。

例句5中“归脾、胃、肝经”为中医药文化专有术语,有道翻译和DeepL均错译成“belongs to……”,而百度则错译成“returns to……”,只有ChatGPT 4.0准确地译成“enters to……”,例句7中句尾的“脉弦滑”,有道翻译的译文为“smooth slippery pulse”,DeepL的译文为“slippery veins”,百度翻译的译文为“smooth pulse”,只有ChatGPT 4.0准确地译为“wiry and slippery pulse”。例句8中的“君臣佐使”作为另一个中医药文化专有术语,在有道翻译和DeepL的译文中直接被省去未译,百度翻译则只译出部分,且该术语所在语句的译文出现严重误译,将该句“此方君臣佐使,九味为伍,共建……之功”错译成“The monarchs, ministers, and envoys of this side are combined with nine flavors to jointly……”,而ChatGPT 4.0则给出准确的译文“This formula follows the principle of Jun-Chen-Zuo-Shi( soverign, minister, assistant, and courier herbs) with a total of nine ingredients”,除courier一词并非标准译文,整体上原文语句的含义得以完整、准确地表达,“君臣佐使”这一术语的内涵也较为准确地译出,且同时提供4个字的汉语拼音。

## 3 讨论

虽然本研究中,ChatGPT 4.0在处理中医药论文语句英译中具有明显优势,然而其文化适应性、术语翻译的准确性及规范性并非完美无缺,因此,也须警惕其在这一应用场景中的短板。

### 3.1 文化差异挑战

中医药文化中包含大量文化负载词,这对于相关内容的翻译提出很大的挑战,研究中所涉及4个不同的工具无一例外在处理此类术语和表达时或多或少出现不足和问题。例句1中的“精气溢泄”,有道翻译的译文为“the essence is abundant can it ‘overflow’”,DeepL的译文为“overflow of the essence of the diarrhea”,而百度翻译的译文为“overflow and diarrhea of the essence”,只有ChatGPT译作“essence and qi to overflow”,只不过遗憾的是此处的“精气”正确的译文应为“semen and qi”,另外该例句中的“精室”,也误译成“essence chamber”,有道翻译直接省去该词的翻译,而DeepL译成 essence room,百度翻译的译文为 seminal chamber,较为接近准确译文。虽然该例句较短,且“精”在部分中医药英文词典中也确有“essence”的译本,但上下文已然直观明确主题,即脾、肾2个脏器在生殖繁衍方面的重要性,显然这几个工具均未能依据语境完全识别出此处“精”和“泄”字的真实所指,此处“精室”仍应取“semen chamber”。这些问题说明所涉4个翻译工具或语言模型均不能完整、全面、准确地识别文化差异,尤其是对于中医药文化负载词的解读和翻译。

### 3.2 术语英译准确性、标准化挑战

例句9原文为“高尿酸未出现临床症状,如痛风性关节炎、痛风性肾病时,多归属于中医学‘湿浊’范畴”,该例句中的“出现”一词,在有道翻译、DeepL和百度翻译中分别译做 present (clinical symptoms), (high uric acid) without (clinical symptoms) 和 appear (clinical symptoms),只有ChatGPT4.0准确地译做 when hyperuricemia has not yet manifested clinical symptoms。该例句内容多为西医学表述。就医学翻译而言,一般来说,目前可用的翻译工具或语言模型在开发过程中多以西医内容的语料库作为训练素材,这也是为什么相较于中医药内容而言,西医文本的机器翻译译文准确率会更高。然而从这一例句的翻译来看,本研究中所涉及的工具除ChatGPT4.0外,针对西医学的翻译仍然存在不足。

例句7开头的“痰热互结”,有道翻译、DeepL和百度翻译给出的译文分别为 phlegm-heat interweaving、phlegm-heat inter-conjugation 和 phlegm-heat,而ChatGPT4.0的译文为 phlegm-heat accumulation,虽然前3个工具译出痰热之间相互影响,但未能译出规范的标准译文,即 combined phlegm and heat,而ChatGPT4.0也只译出痰热的积聚,未能译出二者之

间的相互影响。例句4中的“原穴”,其余3个工具均译做 original point/acupoint, ChatGPT4.0译做 Yuan-Primary point,虽然一定程度上均译出穴位甚至初始穴位之意,但是都未能译出规范的标准译文,即 Yuan-source point。

### 3.3 译文评分分析

通过对比分析4种工具产出的译文质量,发现ChatGPT4.0在翻译中医药文本时,译文质量明显优于另外3种工具,这一结论从人工评分中也得到印证,其BLEU平均值最高,达到87.29%,人工评分的平均分为95.15分。而DeepL产出译文的BLEU平均值最低,为37.62%,人工评分平均分为79.45分,亦为得分最低。说明ChatGPT4.0可作为中医药英文论文写作中的首选翻译辅助工具。王和私和马柯昕<sup>[14]</sup>针对ChatGPT应用于生物医学文本的翻译研究得出同样的结论。同时,整体上机器翻译自动评分工具BLEU的评分和2名人工评分的得分保持一致,说明BLEU可以作为中医药文本机器翻译质量评估的工具。通过分析也注意到,例句1的BLEU值得分总分最低,为75.32%。而得分总分最高的例句为例4,达到250.97%。究其原因,例句1的文本源于中医古籍,且其语言风格为古汉语文言文,提示在使用包括ChatGPT4.0在内的翻译工具处理源于中医古籍的古文表达时须谨慎、认真校对,必要时查阅相关权威英译本或请教行业专家。而得分最高的例句4其内容为西医学表述,故而其得分最高,这和翻译工具或语言模型训练过程中大量使用现代医学素材直接相关。

### 3.4 译文质量分析

通过对比分析,发现相较于其他3种翻译工具,ChatGPT4.0在翻译中医药文本中的优势体现在句式构建及术语翻译的准确性和规范性等多个方面。首先,在句式构建上,ChatGPT4.0在处理较长句式时,所产出的译文断句层次清晰,逻辑连贯且简单易懂。其次,针对中医药术语翻译,研究纳入的文本中包含中医基础理论、针刺、穴位、草药及其性味归经、方剂等,除少量错误外,整体上ChatGPT4.0基本可以给出准确英译,且同时参照相关中医药术语英译国际标准和规范,如穴位和草药在英译中的规范表达等。因此,在术语翻译的准确性和规范性方面,ChatGPT4.0也展现出绝对的优势。相比之下,其余3个工具产出译文在句式构建和术语翻译的准确性等方面均存在较多问题,使用过程中会为译后编辑增加较大工作量。

然而在使用以上机器翻译工具辅助中医药英文论文写作时,也不能忽视翻译工具的局限性,以上分析表明,由于中西方的文化差异和中医药独有的表达,包括 ChatGPT4.0 在内的4种翻译工具均存在一定的问题,如文化负载词的英译,尤其是诸如上文提到的“精气溢泻”“九味为伍”等语言高度凝练且文化内涵丰富的表达等。当然也包括部分术语英译中的准确性和规范性不足等问题。中医药国际传播过程中,术语翻译的规范化和标准化一直被认为是至关重要的一环。而中医药英文论文的写作和发表更加强调术语的统一,然而本研究中所涉所有翻译工具在中医药论文术语英译中的规范化和标准化均存在不足。

### 3.5 研究局限性

第一,本研究从中医药权威中文期刊选取语句,通过机器翻译自动评分工具 BLEU 及人工评分从主客观、多角度针对机器翻译译文进行全面的评价并展开对比分析,但所选语句数量较少,且未能包含相对完整的中医药文本语段及语篇,未来的研究中,应当针对中医药文本的较长语篇展开相关研究,以更加全面、深入、多维度评价中医药文本机器翻译译文质量。第二,研究方法亦存在一定的不足,研究中所选语句均源于中文期刊,均无相应的权威英译,虽然语句参考译文经资深中医英译专家翻译定稿,但难免存在一定的主观性。未来,针对中医药文本机器翻译质量的研究可选取公开出版或发表的权威中医药中英对照文本作为研究对象,以期增加整体研究的信度和效度。第三,未来的相关研究中,可考虑纳入更多的机器翻译自动评价工具,以便拓展中医药文本机器翻译译文质量的评价维度。第四,在过去的短短数月内,国内 AI 工具或大语言模型的发展日新月异,本研究未能纳入更多 AI 工具,未来研究应聚焦国内 AI 工具及大语言模型翻译中医药文本的整体适用性评测。

## 4 结语

从机器翻译译后编辑(Machine Translanton Post-Editing, MTPE)的应用场景来说,中医药论文英文写作中可以选择 ChatGPT 4.0 作为论文翻译的预处理工具,提供较高质量的初始译文,便于译后编辑和校对。但是也应该看到,包括 ChatGPT 4.0 在内的本研究涉及4种翻译工具依然存在不足。首先,关于长句的翻译,应当格外留意译文的断句、衔接及连接词的使用,确保译文层次分明、表意清晰,如怀疑存在语法错误,则可考虑使用语法纠错软件

辅助识别、纠错。同时也可继续使用 ChatGPT4.0 进行单句细化的纠错,但是要注意给出准确的指令。其次,关于术语的翻译,应该格外留意术语英译的准确性和规范性问题,尤其是涉及文化内涵丰富的表达、草药名称、穴位名称及方剂的英文名称等,必要时查阅相关国际标准予以确认。笔者在翻译实践中也发现,有些翻译工具在分段处理整篇论文英译过程中,不止一次出现过前后术语英译不一致的情况,这也提示在使用翻译工具英译论文的过程中,要把握核心术语及关键词的前后一致。再次,关于出自中医古籍的古汉语表达,在使用机器翻译过程中同样要格外谨慎,尽量查阅权威英译本中准确、规范的英译表达。

虽然所有的机器翻译工具都在不断地更新迭代,相信多数的工具都会变得越来越完善。但是,目前在中医药论文的英译中,机器翻译依然不能完全替代人工翻译或者人工校对。总体而言,ChatGPT 还未对传统人工翻译或现行机器翻译工具造成根本性变革<sup>[15]</sup>,认真、细致的译后编辑仍然必不可少。本研究通过实证研究对不同工具产出的译文质量进行全面的评估,从定量的角度,量化地对比本研究中各个工具在翻译中医药论文语句中的得分,也从定性的角度细致对比、分析量化指标中未包含的译文质量细节,以期为中医药英文论文写作中机器翻译的高效应用提供借鉴。

### 参考文献:

- [1] 廖星,谢雁鸣,罗辉,等. 中医药 SCI 论文投稿资源介绍[J]. 中国中西医结合杂志,2017,37(10):1271-1275.
- [2] 朱梦梦,李逸雯,罗斌玉,等. 中医药研究生国际期刊论文写作与发表现状调查[J]. 中国循证医学杂志,2022,22(6):634-640.
- [3] 范晓晖. 我国英文医学论文写作及发表现状[J]. 中国 ESP 研究,2015,6(1):47-52.
- [4] BORDAGE G. Reasons reviewers reject and accept manuscripts: the strengths and weaknesses in medical education reports[J]. Acad Med,2001,76(9):889-896.
- [5] Byrne D W. Common reasons for rejecting manuscripts at medical journals: A survey of editors and peer reviewers[J]. Science Editor,23(2):39-44.
- [6] TURCOTTE C, DROLET P, GIRARD M. Study design, originality and overall consistency influence acceptance or rejection of manuscripts submitted to the Journal[J]. Can J Anaesth,2004,51(6):549-556.
- [7] 王楠鑫,云红. 中医药说明书功效主治部分的机辅翻译

- [J]. 绵阳师范学院学报, 2019, 38(4): 93-99.
- [8] 宋熹玥, 周净, 刘伟. 中医古籍智能机器翻译模型构建研究[J]. 中国中医药图书情报杂志, 2024, 48(6): 130-135.
- [9] 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.
- [10] 蒙洁琼. 基于神经机器翻译的中医论文结构性摘要翻译策略[J]. 中国中医药现代远程教育, 2019, 17(9): 53-56.
- [11] 温海煜. 中医文本机译错误分析及译后编辑研究[D]. 北京: 北京中医药大学, 2023.
- [12] PAPINENI K, ROUKOS S, WARD T, et al. Bleu: a method for automatic evaluation of machine translation [C]// Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, USA: Association for Computational Linguistics, 2002: 311-318.
- [13] 周成彬, 刘忠宝. 基于语义信息共享Transformer的古文机器翻译方法[J]. 情报工程, 2022, 8(6): 114-127.
- [14] 王和私, 马柯昕. 人工智能翻译应用的对比研究: 以生物医学文本为例[J]. 中国科技翻译[J], 2023(3): 26.
- [15] 文旭, 田亚灵. ChatGPT应用于中国特色话语翻译的有效性研究[J]. 上海翻译, 2024(2): 27-34, 94-95.
- (收稿日期: 2025-05-26)

## 《中医教育》2026年征稿方向

### 一、数字化背景下的中医药教育

1. 人工智能在中医药教育领域的应用与实践
2. 中医药教育的数字化转型的路径及模式

### 二、中医药人才培养模式的创新

1. 科教融汇、产教融合、医教协同背景下的中医药人才培养
2. 新医科背景下中医教育的改革与创新
3. 中医药教育制度变革
4. 新质生产力赋能中医教育改革的路径与模式
5. 中医药拔尖创新人才培养与改革
6. 当代教育体系下中医临床思维的培养
7. 中医药创新创业教育的模式与路径
8. 中医药教育“出海”的路径与策略
9. 中医基础类/经典类学科师资培训研究与探索
10. 中医药与新兴学科交叉融合人才培养案例研究
11. 中医住院医师规范化培训的新要求与实践路径
12. 中医药整合课程与创新课程的构建与实践研究

### 三、师承教育与中医人才培养

1. 师承教育在中医人才培养中的价值
2. 现代师承教育模式的探索与实践
3. 师承教育与现代院校教育相融合的教育体系构建

### 四、教育评价改革与中医药人才培养

1. 新时代中医药教育综合改革的模式与路径
2. 基于多元能力的中医药学生综合评价指标体系构建
3. 中医药教师教学、科研、临床实践一体化评价模式探索

### 五、中医药人才素质教育研究

1. 新时代中医药院校的思想政治教育体系建设
2. “五育并举”在中医药人才培养中的实践
3. 中医药课程思政建设的模式与路径

### 六、“双一流”建设与评价研究

### 七、中医药教育管理模式创新

1. AI赋能中医药教育管理的理论与实践
2. 中医药特色育人管理体系创新实践
3. 新医科背景下中医药教育资源优化配置