

英语专八试题人机翻译句法复杂度对比分析研究

黎雨婷, 黄贻宁

福建理工大学人文学院, 福建 福州

收稿日期: 2026年2月24日; 录用日期: 2026年3月8日; 发布日期: 2026年3月25日

摘要

随着生成式人工智能技术在自然语言处理领域的深度渗透, ChatGPT等大语言模型在翻译任务中的应用备受关注, 但现有评估多聚焦准确性与流畅性, 对反映翻译专业性的句法复杂度关注不足。本研究以1996~2025年(不含2020年)英语专业八级(专八)汉译英试题为对象, 构建ChatGPT-5译文与专八官方参考答案的平行语料库, 采用BFSU系列句法分析工具, 从基础句法指标、复杂句法结构指标、句法多样性与规范性三个维度, 结合独立样本t检验开展量化对比。结果显示: 两类语料在基础句法规模与句法多样性上无显著差异; 人类翻译语料在从属分句、补语从句等复杂结构的使用密度, 以及句法规范性、场景适配稳定性上显著优于ChatGPT-5; ChatGPT-5存在复杂句法结构使用保守、异常值占比偏高的问题。本研究首次系统揭示了二者在句法复杂度上的深层差异, 丰富了AI翻译质量评估的理论体系, 为AI翻译模型优化、英语专业教学及专八翻译备考提供了实证依据与实践指导。

关键词

英语专八翻译, ChatGPT-5翻译, 人类翻译, 句法复杂度

A Comparative Study on Syntactic Complexity between Human and ChatGPT-5 Translations of TEM-8 Translation Task

Yuting Li, Yining Huang

School of Humanities, Fujian University of Technology, Fuzhou Fujian

Received: February 24, 2026; accepted: March 8, 2026; published: March 25, 2026

Abstract

With the deep integration of generative artificial intelligence technologies into the field of natural

language processing, large language models such as ChatGPT have attracted considerable attention in translation tasks. However, existing evaluations mainly focus on accuracy and fluency, while relatively neglecting syntactic complexity, which reflects translation professionalism. This study takes the Chinese-English translation tasks of the Test for English Majors Band 8 (TEM-8) from 1996 to 2025 (excluding 2020) as research data. A parallel corpus consisting of ChatGPT-5 translations and the official reference translations was constructed. The BFSU series of syntactic analysis tools were employed to conduct quantitative comparisons from three dimensions: basic syntactic indices, complex syntactic structure indices, and syntactic diversity and well-formedness, combined with independent-samples t-tests. The results show that there is no significant difference between the two corpora in terms of basic syntactic scale and syntactic diversity. However, the human translation corpus demonstrates significantly higher density in the use of complex structures such as subordinate clauses and complement clauses, as well as stronger syntactic well-formedness and greater stability in contextual adaptability. ChatGPT-5 shows a conservative tendency in the use of complex syntactic structures and exhibits a relatively higher proportion of outliers. This study systematically reveals the underlying differences between the two in syntactic complexity, enriches the theoretical framework of AI translation quality evaluation, and provides empirical evidence and practical guidance for AI translation model optimization, English-major teaching, and TEM-8 translation preparation.

Keywords

TEM-8 Translation, ChatGPT-5 Translation, Human Translation, Syntactic Complexity

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

人工智能技术的飞速发展推动机器翻译成为自然语言处理领域的重要分支[1], 极大地提升了跨语言交流的便捷性, 在专业文本翻译中展现出强劲技术潜力。ChatGPT 等新一代大语言模型基于海量数据预训练并经翻译任务精细调整[2], 虽非专门为翻译设计, 但可通过精确指令(prompt)引导发挥翻译潜能, 其表现受到学界广泛关注。然而, 现有机器翻译评估多侧重准确性和流畅性[3], 对句法复杂度这一反映翻译文本专业性与语言精密性的关键维度关注不足。英语专业八级翻译考试的官方参考答案凝聚专业译者智慧, 在句法组织、语义传达与文本适配性上具有标杆意义, 为衡量 AI 翻译深层语言能力提供了权威参照。当前学界尚未充分聚焦句法复杂度开展 AI 翻译与人类专业翻译的对比研究, 针对 ChatGPT-5 在权威考试场景下的句法生成能力也缺乏系统性量化分析。基于此, 本研究构建量化分析框架, 对 ChatGPT-5 译文与专八官方参考答案进行深入对比, 揭示两者在深层语言组织能力上的异同, 为 AI 翻译模型的评估优化、专业翻译教学提供实证依据与理论参考。

1.1. 研究背景

生成式人工智能的快速发展推动了机器翻译范式的革新。ChatGPT 系列模型凭借较强的自然语言理解与生成能力, 已在科技、文学等专业翻译场景中得到应用。英语专业八级翻译考试(以下简称“专八”)作为我国衡量英语专业高年级学生语言应用能力的权威测评, 其官方参考答案凝聚专业译者智慧, 在句法组织、语义传达与文体适配性方面具有标杆意义。现有研究表明, ChatGPT 在多种语言任务中

展现出良好性能[4], 其优势源于大规模语言模型的支撑[5], 并在译后编辑教学[6]与翻译指令设计等方面引发广泛关注。然而, 在涉及深层语义理解、专业知识与文化敏感性的翻译任务中, ChatGPT 仍存在一定局限。

句法复杂度是翻译质量评估的重要维度, 指句法结构的繁复程度与组织精密性, 主要体现为结构层级嵌套、扩展方式及语义逻辑关联密度。合理运用复杂句法有助于实现语义精准传达与文体正式性。既有研究普遍将句法复杂度界定为语言产出中句法结构的多样性与复杂程度, 常以句子长度与结构复杂性为衡量指标, 句子越长、结构越复杂, 复杂度通常越高[7]-[9]。在翻译过程中, 句法复杂度受源语与目标语的词汇语法特征、文本体裁及译者语言偏好等多重因素影响[10][11], 如汉语话题突出、语序灵活, 而英语句法规则相对严谨[12]。相关研究还表明, 翻译文本在从属结构和整体复杂性上往往呈现简化倾向[8]。在实践层面, 句法过简可能造成信息缺失, 过繁则削弱表达清晰度; 人类译者可灵活调控句法以适应语境, 而机器翻译系统则可能出现过度简化或复杂化。尽管 ChatGPT 在句法处理方面已展现潜力, 其译文句法复杂度与人类翻译之间的差异仍有待系统评估。目前相关对比研究多集中于语义准确性与流畅度, 对 ChatGPT-5 等新一代模型译文的句法复杂度缺乏深入、量化的探讨。

1.2. 研究意义

句法复杂度是衡量翻译文本语言精密性与专业性的关键维度, 直接体现译者对目标语语法规则的掌控程度, 并与语义逻辑传达和文体适配性密切相关, 是区分普通翻译与专业翻译的重要特征。而现有 AI 翻译研究多关注语义忠实度与流畅度等表层指标, 对句法复杂度等深层语言特征重视不足, 尤其缺乏在权威语言测试场景(如专八翻译)中对 AI 翻译与人类专业翻译的系统性量化比较。基于此, 本研究选取 1996~2025 年专八翻译语料(共 29 个年份, 2020 年除外), 借助 BFSU 系列句法分析工具构建量化对比框架, 系统揭示 ChatGPT-5 译文与人类专业译文在句法复杂度上的差异, 拓展 AI 翻译质量评估的理论视角, 并为自然语言处理中句法生成能力的测评提供实证依据。研究结论可为 AI 翻译模型的定向优化提供参考, 促进其复杂句法处理能力的提升, 同时也为英语翻译教学与专八备考提供可操作的实证支持。

1.3. 研究问题

基于上述背景与意义, 本研究提出以下核心研究问题:

- 1) 英语专八试题人机翻译在基础句法指标上存在何种差异?
- 2) 两类语料在复杂句法结构的使用上是否存在显著区别?
- 3) 两类语料的句法多样性与规范性表现如何, 差异成因是什么?

2. 研究设计

本研究以 1996~2025 年(不含 2020 年)专八汉译英试题为基础构建平行语料库, 包括在统一提示词下生成的 29 篇 ChatGPT-5 译文(60~187 词, 均值 99.62 词)及 29 篇华研外语专八官方参考译文(65~193 词, 均值 114.90 词)。为客观比较两类译文的句法复杂度, 研究采用“工具解析、指标量化、统计检验”的分析范式, 借助经国内学者优化的 BFSU Stanford Parser 1.0 与 BFSU Syntactic Complexity Analyzer 1.0: 先进行句法解析并提取原始数据, 再计算各项派生复杂度指标。

3. 研究指标体系

本研究选取以下三类核心指标, 涵盖基础句法特征、复杂结构使用及句法多样性与规范性三个维度, 各指标定义如表 1 所示:

Table 1. Core indicators of syntactic analysis

表 1. 句法分析核心指标

指标类型	指标缩写	指标名称	指标定义
基础句法指标	W	总词数	文本中所有实词与虚词的总数量
	S	句子数	以句号、问号、感叹号结尾的语法单位数量
	C	分句数	包含主谓结构的最小语法单位数量
	T	T单位数	最小的不能再分割的主谓结构单位数量
	MLS	平均句子长度	总词数/句子数(W/S)
	MLT	平均T单位长度	总词数/T单位数(W/T)
	MLC	平均分句长度	总词数/分句数(W/C)
	C/S	分句 - 句子比	分句数/句子数, 反映单句包含的分句密度
	C/T	分句-T单位比	分句数/T单位数, 反映T单位的句法复杂度
	复杂句法结构指标	DC	从属分句数
CT		补语从句数	充当句子补语的从句数量
CP		并列短语数	由并列连词连接的短语数量
CN		名词短语数	以名词为中心词的短语数量
DC/C		从属分句 - 分句比	从属分句数/分句数, 反映分句的从属化程度
DC/T		从属分句-T单位比	从属分句数/T单位数, 反映T单位的从属结构密度
CT/T		补语从句-T单位比	补语从句数/T单位数, 反映补语从句的使用频率
句法多样性与规范性指标	VP/T	动词短语-T单位比	动词短语数/T单位数, 反映动词短语多样性
	CP/C	并列短语 - 分句比	并列短语数/分句数, 反映并列结构使用密度
	CN/T	名词短语-T单位比	名词短语数/T单位数, 反映名词短语丰富度
	异常值占比	无指标缩写	某指标为0或显著偏离均值(± 2 标准差)的文本占比

4. 研究对比分析

为检验 ChatGPT-5 译文与人类译文在各句法指标上的差异可靠性, 本研究采用独立样本 t 检验, 显著性水平设为 $\alpha = 0.05$ 。表格中“t”对应 t 统计量(t-statistic), “p”对应概率值(p-value), 二者共同用于判断差异是否显著; 标注“*”表示差异达到统计学显著性($p < 0.05$)。指标数值后的“ $M \pm SD$ ”代表均值 \pm 标准差(Mean \pm Standard Deviation), 其中 M (均值)反映两类语料在该指标上的平均水平, SD (标准差)反映数据的离散程度; 样本量均为 $N = 29$, 涵盖 1996~2025 年专八译文(不含 2020 年), 为统计分析提供可靠基础。t 值绝对值越大, 说明组间差异相对随机波动越明显; p 值越小, 差异由偶然因素导致的可能性越低。基于上述统计结果, 本文将从基础句法指标、复杂句法结构以及句法多样性与规范性三个维度, 对两类译文的句法复杂度进行系统比较。

4.1. 基础句法指标对比

基础句法指标直接反映文本的语言组织规模与基本结构特征, 两类语料的描述性统计结果如表 2 所示:

Table 2. Comparison of basic syntactic measures between two corpora (M ± SD)**表 2.** 两类语料基础句法指标对比(M ± SD)

指标	ChatGPT-5翻译语料 (N = 29)	人类翻译语料 (N = 29)	t值	p值
W (总词数)	99.62 ± 24.31	114.90 ± 27.58	-2.376	0.021*
S (句子数)	5.55 ± 1.82	5.83 ± 2.01	-0.634	0.528
C (分句数)	9.72 ± 3.15	11.07 ± 3.52	-1.689	0.097
T (T单位数)	6.86 ± 2.24	7.59 ± 2.46	-1.215	0.230
MLS (平均句子长度)	19.22 ± 4.87	21.04 ± 5.32	-1.457	0.151
MLT (平均T单位长度)	15.36 ± 3.28	16.78 ± 3.65	-1.793	0.079
MLC (平均分句长度)	10.83 ± 2.15	11.56 ± 2.37	-1.284	0.204
C/S (分句 - 句子比)	1.78 ± 0.42	1.90 ± 0.45	-1.052	0.300
C/T (分句-T单位比)	1.42 ± 0.28	1.46 ± 0.31	-0.567	0.573

由表 2 可知, 在基础句法指标上, 两类语料呈现以下特征:

1) 总词数差异显著: 人类翻译语料的平均总词数(114.90)显著高于 ChatGPT-5 翻译语料(99.62) ($p = 0.021 < 0.05$)。

2) 句子数、分句数等指标无显著差异: 两类语料在句子数($p = 0.528$)、分句数($p = 0.097$)、平均句子长度($p = 0.151$)等指标上虽存在一定数值差异, 但均未达到统计显著水平。

3) 句法密度指标表现相当: C/S (分句 - 句子比)与 C/T (分句-T 单位比)的均值差异极小(分别为 0.12 和 0.04), 且 p 值均大于 0.05。

4.2. 复杂句法结构指标对比

复杂句法结构是体现文本专业性的核心特征, 两类语料的对比结果如表 3 所示:

Table 3. Comparison of complex syntactic structure measures between two corpora (M ± SD)**表 3.** 两类语料复杂句法结构指标对比(M ± SD)

指标	ChatGPT-5翻译语料 (N = 29)	人类翻译语料 (N = 29)	t值	p值
DC (从属分句数)	2.69 ± 2.03	3.55 ± 2.27	-2.013	0.049*
CT (补语从句数)	2.93 ± 2.15	3.86 ± 2.41	-2.175	0.034*
CP (并列短语数)	3.72 ± 2.38	4.05 ± 2.53	-0.542	0.590
CN (名词短语数)	25.34 ± 14.62	29.86 ± 16.37	-1.328	0.191
DC/C (从属分句 - 分句比)	0.28 ± 0.16	0.35 ± 0.18	-2.247	0.029*
DC/T (从属分句-T单位比)	0.41 ± 0.23	0.50 ± 0.25	-2.089	0.042*
CT/T (补语从句-T单位比)	0.43 ± 0.25	0.52 ± 0.27	-2.136	0.038*

表 3 数据显示, 两类语料在复杂句法结构指标上的差异更为显著:

1) 从属分句与补语从句使用差异显著: 人类翻译语料的从属分句数(3.55)、补语从句数(3.86)均显著

高于 ChatGPT-5 翻译语料(分别为 2.69 和 2.93), 且 p 值均小于 0.05。

2) 比率指标差异显著: DC/C(从句分句 - 分句比)、DC/T(从句分句-T 单位比)、CT/T(补语从句-T 单位比)三个核心比率指标中, 人类翻译语料的均值均显著高于 ChatGPT-5 翻译语料($p < 0.05$)。

3) 并列短语与名词短语无显著差异: CP(并列短语数)与 CN(名词短语数)的均值虽均低于人类翻译语料, 但差异未达到统计显著水平($p > 0.05$)。

4.3. 句法多样性与规范性分析

句法多样性反映语言表达的丰富程度, 规范性则体现句法使用的准确性与合理性, 两类语料的对比结果如表 4:

Table 4. Comparison of syntactic diversity and accuracy measures between two corpora ($M \pm SD$)

表 4. 两类语料句法多样性与规范性指标对比($M \pm SD$)

指标	ChatGPT-5翻译语料 (N = 29)	人类翻译语料 (N = 29)	t值	p值	异常值占比 (GPT/人类)
VP/T(动词短语-T单位比)	2.03 ± 0.65	2.18 ± 0.71	-0.924	0.360	6.90%/3.45%
CP/C(并列短语 - 分句比)	0.42 ± 0.21	0.45 ± 0.23	-0.618	0.540	3.45%/0%
CN/T(名词短语-T单位比)	3.82 ± 1.57	4.15 ± 1.72	-0.897	0.374	6.90%/3.45%
整体异常值占比	-	-	-	-	17.24%/6.89%

由表 4 可知:

1) 句法多样性指标无显著差异: VP/T(动词短语-T 单位比)、CP/C(并列短语 - 分句比)、CN/T(名词短语-T 单位比)三个多样性指标的均值差异均未达到统计显著水平($p > 0.05$)。

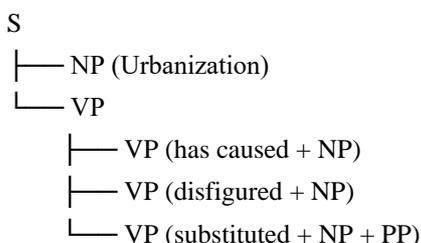
2) 规范性存在显著差距: ChatGPT-5 翻译语料的整体异常值占比(17.24%)显著高于人类翻译语料(6.89%), 主要表现为部分年份的 DC、CT 等指标为 0, 或 CN/T 指标显著偏离均值(如 2024 年 GPT 的 CN/T 为 12.00, 远超均值 + 2 标准差)。而人类翻译语料的异常值占比极低, 仅在 2009 年出现 CP = 0 的情况。

3) 体裁适配性差异显著: ChatGPT-5 的异常值主要集中在 2002、2007、2019、2024 等年份的文本中, 这些文本多为科普说明类体裁, ChatGPT-5 倾向于使用简单句实现语义传达, 而人类翻译即使在这类体裁中仍会保留适当的复杂结构以保证表达的专业性。

为避免统计结果停留于指标层面, 有必要结合具体语例对结构差异进行验证。以 2022 年 TEM-8 翻译材料为例, 原句“城市驱散了旷野原有的住民, 破坏了旷野古老的风景, 越来越多地以井然有序的繁华, 取代我行我素的自然风光。”在人类译文中被处理为:

“Urbanization has caused the displacement of original inhabitants in the pristine wilderness, disfigured its age-old scenery, and increasingly substituted its unspoiled natural landscape with the orderly hustle and bustle.”

该句在句法层面呈现典型的多重谓语并列结构, 其核心树形结构可概括为:

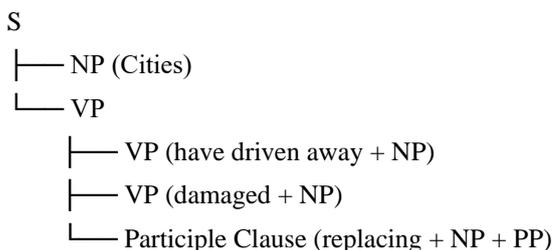


三个谓语在同一主语支配下并列展开, 形成对等递进关系; 同时, 第一分句中的名词化结构 “the displacement of...” 进一步增加了内部层级嵌套, 使整体句法密度较高。

相比之下, GPT 译文为:

“Cities have driven away the dwellers of the wild and damaged its ancient landscapes, replacing nature’s unrestrained beauty with ever more orderly prosperity.”

其核心结构可表示为:

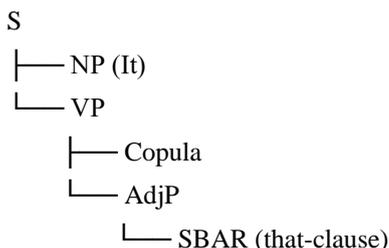


在该结构中, 第三个动作未与前两项形成完全对等的并列谓语, 而是降格为现在分词结构, 成为附加性补充信息。从句法层级看, 原本三重并列的对等结构被改写为 “主干并列 + 附属成分” 的形式, 嵌套深度相对降低。语义命题并未缺失, 但原句所呈现的递进强化效果有所减弱。

再如结尾句 “人类永远不可能以城市战胜旷野”。人类译文为:

“It is noteworthy that humankind will never be able to prevail over it through urbanization.”

其结构包含形式主语与 that 引导的名词性从句, 句法层级如下:

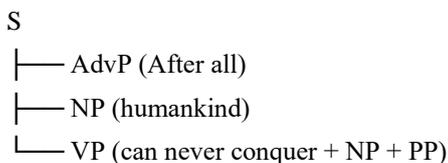


该结构通过评价框架嵌套, 使判断语气在结构层面得到显性强化。

GPT 译文为:

“After all, humankind can never conquer the wilderness with cities.”

该句采用简单主谓结构, 仅以句首副词 “After all” 实现语气提示, 未使用名词性从句或评价性嵌套结构。句法结构可概括为:



相较人类译文, 句法嵌套层级明显减少。评价语气由结构层面转移至词汇层面, 体现出复杂结构使用上的保守倾向。

由此可见, 在散文类文本中, GPT 译文在并列递进结构与评价性结构的处理上, 倾向于通过分词结构或简单主谓结构完成表达, 而较少采用名词性从句或评价框架嵌套。这种复杂结构的相对回避并未造成核心命题意义的缺失, 但削弱了语篇层级与修辞张力, 与量化统计中复杂从句比例与嵌套深度下降的结果形成呼应。

5. 研究结果讨论

前文基于基础句法指标、复杂句法结构及句法多样性与规范性三个维度的量化分析表明, ChatGPT-5 译文与专八人类译文在句法规模和多样性上整体接近, 但在人类译文中, 从属分句、补语从句等复杂结构的使用密度、规范性与稳定性显著更高, 而 ChatGPT-5 表现出复杂结构使用偏保守、异常值比例较高的特征。下文将从生成机制差异出发, 分析两类译文句法分化的成因, 并讨论 ChatGPT-5 在复杂句式处理中的优势与局限, 为模型优化与翻译教学提供理论依据。

进一步结合 1996~2025 年(2020 年除外) TEM-8 翻译材料的文体分布可以发现, 句法复杂度差异具有明显的体裁依存性。专八汉译英文本主要可分为抒情散文类与议论政论类两种类型。散文类文本强调意象递进与情感铺陈, 原文多采用意合表达, 语义呈层层推进式展开; 在英译过程中, 人类译者往往通过增加定语从句、非限定性从句或分词结构, 将隐含逻辑显性化, 从而形成多层嵌套结构。相较之下, ChatGPT-5 的译文更倾向于采用并列句或独立简单句进行线性展开, 减少从属层级。在议论类文本中, 人类译文常使用 that 引导的补语从句、名词化结构或倒装结构以强化逻辑强调, 而 AI 译文则倾向于改写为条件句或简单陈述句, 从而降低句法嵌套密度。由此可见, 复杂结构差异不仅源于生成机制差异, 也受到专八特定文体规范的制约。

5.1. 语言模型生成机制与人类翻译的本质差异

两类语料在句法复杂度上的差异, 本质上源于 ChatGPT-5 的生成机制与人类翻译认知过程的根本不同。

5.2. 生成逻辑的差异

ChatGPT-5 的翻译生成基于大规模语料库的统计学习与概率预测, 其核心目标是生成“符合语境且概率最高”的文本。因此在句法选择上更偏向高频、结构相对简单的句式, 以降低生成风险。这一机制使其在从属分句、补语从句等复杂结构的使用上趋于保守, 个别译文甚至出现复杂结构回避现象(如 $DC = 0$)。从历年文本具体表现来看, 这种保守性呈现出“句法降阶”特征。即当原句可以通过嵌套从句实现语义限定时, ChatGPT-5 更倾向于拆分为两个或多个并列结构。例如散文类文本中常见的“意象 + 限定性修饰”结构, 在人类译文中往往以关系从句实现, 而 ChatGPT-5 的译文则将修饰成分外置为独立句, 弱化主从依存关系。这种结构选择在命题层面保持等值, 但在信息层级与语篇逻辑上呈现扁平化倾向, 直接导致 DC/C 与 CT/T 指标的显著差异。

人类翻译是一个“语义解构、重组、优化”的主动认知过程, 专业译者在准确传达语义的基础上, 会根据文本体裁与表达需求, 主动运用复杂句法结构构建逻辑严密的语义网络。例如, 在文学散文类文本中, 人类译者会通过定语从句、状语从句增强表达的形象性; 在科普文本中, 会通过补语从句明确概念定义, 这一主动优化过程是当前 AI 模型难以模拟的。

5.3. 语境适配性的差异

人类译者能够精准把握专八翻译的文体规范, 理解该考试对句法复杂度的隐性要求——既需避免过度简单化导致的表达幼稚, 也需避免过度复杂导致的理解障碍。因此, 人类翻译语料的句法指标呈现出“适度复杂、均衡稳定”的特征, 异常值占比极低。

ChatGPT-5 对特定考试场景的文体规范适配性不足, 其句法选择更多依赖通用语料库的统计规律, 而非对考试要求的深度理解。这导致其翻译文本的句法复杂度波动较大, 部分文本过于简洁(如 2009 年 GPT 的 MLS 仅 10.78), 部分文本又出现结构冗余(如 2024 年 GPT 的 CN/T 达 12.00), 难以实现像人类翻

译那样的精准适配。

6. 研究启示与应用价值

本研究通过量化分析揭示了 ChatGPT-5 与人类专业译文在句法复杂度上的核心差异, 为 AI 翻译优化与专八翻译教学提供了可操作的实践启示, 主要体现为“基础能力达标、复杂结构不足、规范适配欠缺”三方面特征。在技术层面, 研究结果为大语言模型翻译能力的定向优化提供了实证依据。模型改进应聚焦复杂从属结构使用不足的问题, 增强专八翻译、学术翻译等专业语料的训练权重, 明确标注复杂句式的语义逻辑与句法功能, 并引入场景适配与句法规范校验机制, 以降低异常值比例, 避免过度简化或结构冗余, 提升译文的句法稳定性与专业性。在教学层面, 研究结论可直接服务于英语教学与专八翻译备考。教师可将 ChatGPT-5 译文与人类参考译文作为平行语料, 引导学生对比分析复杂句式的使用差异, 明确句法提升路径; 学习者亦可借助 ChatGPT-5 生成翻译初稿, 并结合人类译文进行针对性句法优化, 重点强化从属结构与补语从句的运用。教学中应强调“适度复杂”的句法原则, 帮助学生在语义准确与句法规范之间实现平衡。

7. 研究局限与未来研究方向

本研究虽通过量化分析揭示了 ChatGPT-5 与人类专业翻译在句法复杂度上的差异, 但仍受语料范围、研究维度与变量控制等因素限制。语料仅涵盖 29 个年份的专八翻译文本, 体裁集中于文学、文化与科普, 结论的跨领域适用性仍有待验证; 分析维度侧重句法复杂度, 未结合语义准确性与文体适配性等维度, 可能影响对翻译质量的整体判断; 此外, 研究仅采用统一提示词, 未考察提示词差异对模型句法生成的影响。针对上述局限, 未来研究可从三方面拓展: 一是扩大语料范围, 考察不同文本类型对 AI 翻译句法复杂度的影响; 二是构建融合句法、语义准确性与流畅度的多维评估框架, 实现对 AI 翻译质量的综合衡量; 三是系统比较不同 ChatGPT 模型与提示词, 分析其对句法生成能力的作用机制与演变趋势, 从而推动 AI 翻译句法能力研究向更全面、深入的方向发展。

参考文献

- [1] Jiang, K. and Lu, X. (2020) Natural Language Processing and Its Applications in Machine Translation: A Diachronic Review. 2020 *IEEE 3rd International Conference of Safe Production and Informatization (IICSPI)*, Chongqing, 28-30 November 2020, 210-214. <https://doi.org/10.1109/iicspi51290.2020.9332458>
- [2] Ray, P.P. (2023) ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope. *Internet of Things and Cyber-Physical Systems*, **3**, 121-154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- [3] Ayala, B. and Chen, J. (2017) A Machine Learning Approach to Evaluating Translation Quality. 2017 *ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, Toronto, 19-23 June 2017, 1-2.
- [4] Helberger, N. and Diakopoulos, N. (2022) The European AI Act and How It Matters for Research into AI in Media and Journalism. *Digital Journalism*, **11**, 1751-1760. <https://doi.org/10.1080/21670811.2022.2082505>
- [5] 朱光辉, 王喜文. ChatGPT 的运行模式、关键技术及未来图景[J]. 新疆师范大学学报(哲学社会科学版), 2023, 44(4): 113-122.
- [6] 王律, 王湘玲. ChatGPT 时代机器翻译后编辑能力培养模式研究[J]. 外语电化教学, 2023(4): 16-23.
- [7] Ortega, L. (2003) Syntactic Complexity Measures and Their Relationship to L2 Proficiency: A Research Synthesis of College-Level L2 Writing. *Applied Linguistics*, **24**, 492-518. <https://doi.org/10.1093/applin/24.4.492>
- [8] Liu, K. and Afzaal, M. (2021) Syntactic Complexity in Translated and Non-Translated Texts: A Corpus-Based Study of Simplification. *PLOS ONE*, **16**, e0253454. <https://doi.org/10.1371/journal.pone.0253454>
- [9] 朱周晔, 王金铨. 源语依存距离对中国英语学习者汉英译文复杂度和质量的影响[J]. 外语教学与研究, 2023, 55(6): 890-900.
- [10] Al-Jabr, A. (2006) Effect of Syntactic Complexity on Translating From/into English/Arabic. *Babel. Revue Internationale*

de la Traduction/International Journal of Translation, **52**, 203-221. <https://doi.org/10.1075/babel.52.3.01aj>

- [11] Xu, J. and Li, J. (2021) A Syntactic Complexity Analysis of Translational English across Genres. *Across Languages and Cultures*, **22**, 214-232. <https://doi.org/10.1556/084.2021.00015>
- [12] Ma, X. (2021) Coping with Syntactic Complexity in English-Chinese Sight Translation by Translation and Interpreting Students: An Eye-Tracking Investigation. *Across Languages and Cultures*, **22**, 192-213. <https://doi.org/10.1556/084.2021.00014>