

人工智能大模型驱动的翻译软件在跨语言学术论文快速理解中的辅助效能研究

黄梦璐

上海理工大学管理学院, 上海

收稿日期: 2025年5月20日; 录用日期: 2025年6月16日; 发布日期: 2025年6月23日

摘要

在全球学术交流需求激增的背景下, 本研究基于认知负荷理论与信息处理模型, 运用准实验设计和用户问卷, 实证分析人工智能大模型驱动的翻译软件在跨语言学术论文快速理解中的辅助效能。研究表明, 该类软件通过自动化语义处理和逻辑图谱构建, 显著提升文献筛选与表层信息提取效率, 在术语解析和句式转化中有效减轻认知负荷。然而, 其在处理图表、公式、算法伪代码等非文本元素的语义对齐, 以及跨学科知识关联和长段落逻辑连贯性还原方面存在不足, 易导致用户对文献深层方法论和创新逻辑的理解缺失, 还可能引发过度依赖、削弱批判性思维。研究明确其作为“表层信息加速处理工具”的定位, 指出在低阶认知任务中效能显著, 但高阶逻辑整合仍需人工干预。

关键词

人工智能大模型, 翻译软件, 跨语言学术论文, 辅助效能, 认知负荷

Research on the Auxiliary Effectiveness of Translation Software Driven by Artificial Intelligence Large Language Models in the Rapid Comprehension of Cross-Lingual Academic Papers

Menglu Huang

Business School, University of Shanghai for Science and Technology, Shanghai

Received: May 20th, 2025; accepted: Jun. 16th, 2025; published: Jun. 23rd, 2025

文章引用: 黄梦璐. 人工智能大模型驱动的翻译软件在跨语言学术论文快速理解中的辅助效能研究[J]. 运筹与模糊学, 2025, 15(3): 504-510. DOI: 10.12677/orf.2025.153180

Abstract

Against the backdrop of growing global academic exchanges, this study empirically explores the auxiliary effectiveness of AI large model-driven translation software in cross-lingual academic paper comprehension. Using quasi-experimental design and user surveys, grounded in cognitive load theory and information processing models, it finds that the software boosts literature screening and surface-level information extraction efficiency via automated semantic processing and logical graph construction, reducing cognitive load in term parsing and sentence transformation. However, it struggles with semantic alignment of non-text elements (charts, formulas, pseudocode), cross-disciplinary knowledge linking, and long passage logic restoration, causing gaps in understanding deep methodologies and innovation logics. Over-reliance may also undermine critical thinking. The study positions it as a “surface-level information accelerator,” effective for low-level tasks but requiring human input for high-level logical integration.

Keywords

Artificial Intelligence Large Language Model, Translation Software, Cross-Lingual Academic Papers, Auxiliary Effectiveness, Cognitive Load

Copyright © 2025 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

在全球学术创新加速的背景下，跨语言知识共享成为推动科技进步的重要动力。根据《2024 全球研究报告》显示，全球学术论文年发表量已突破 320 万篇，较 2019 年增长 47%，其中非英语论文占比达 38.6%。然而，语言壁垒仍制约着知识传播效率，大多数非英语母语研究者因语言障碍延迟或放弃引用重要外文文献。以英语为主导的学术生态中，学术人员需耗费大量时间克服语言壁垒，传统翻译工具因排版混乱、术语不精准、公式图表还原效果差等问题，难以满足学术文献高效理解的需求。

人工智能技术的突破为跨语言学术交流提供了新路径。早期的机器翻译以规则和统计模型为主，受限于上下文理解能力和专业领域适配性，译文生硬且准确率低。而基于深度学习的大模型技术通过海量学术语料训练，显著提升了语义理解和术语翻译的精准度。例如，百度推出的“AI 论文精翻”依托文心大模型，结合百万级学术数据，可精准翻译生物医药、人工智能等领域的专业术语，并支持 LaTeX 排版还原，使复杂公式和图表以原版格式呈现，极大优化了阅读体验。类似工具如 DeepL、谷歌翻译也在实时翻译、多语言处理方面表现出色，但其在学术场景下的综合效能仍需进一步验证。

当前，集文献整理、翻译、解读于一体的工具逐渐普及，其中“小绿鲸英文文献阅读器”表现突出，但其核心效能是否真正提升科研人员对跨语言论文的理解效率，尚缺乏系统性研究。现有研究多聚焦于翻译准确性(如术语匹配率、语法错误率)，而忽略了对用户认知负荷、信息整合效率等维度的实证分析。AI 翻译工具虽能快速生成译文，但若缺乏辅助阅读功能(如论文概要提炼、观点解析)，研究者仍需反复对照原文与译文，导致效率折损。

本研究旨在填补这一空白，通过实证方法评估大模型驱动的翻译工具在跨语言学术论文理解中的辅助效能，为工具开发者优化功能设计提供依据，同时帮助学术人员选择适配自身需求的翻译工具，推动学术资源共享与全球科研协作的深化。

2. 文献综述

近年来,人工智能大模型的飞速发展对翻译领域带来了革命性变化,其在跨语言学术论文理解中的应用也成为研究热点。Kocmi 和 Federmann (2023) [1]提出大语言模型已成为翻译质量评估的先进工具,这为翻译软件的质量把控提供了新方向,也暗示着基于大模型的翻译软件在学术论文翻译评估中的潜在价值。

在生成式人工智能的应用研究方面,李艳等(2025) [2]对典型科研场景下生成式人工智能的使用进行差异性分析,揭示了学科背景与人工智能素养对其使用的影响,这为理解不同学科领域学者对翻译软件的需求差异提供了理论依据。卢云峰等(2025) [3]以偏见研究为例,探讨了生成式人工智能对社会科学研究方法的拓展,表明人工智能技术在学术研究中具有重要作用,其驱动的翻译软件对跨语言学术交流意义重大。

赵正平(2025) [4]阐述了人工智能大语言模型和 AI 芯片的新进展,技术的进步为翻译软件的性能提升提供了硬件和算法基础。在翻译领域的具体应用上,李群(2025) [5]分析了 ChatGPT 在翻译领域的机遇、挑战与应对,刘心怡和何高大(2025) [6]通过中国古诗词人机翻译对比,展现了人工智能在翻译实践中的表现,袁筱一(2025) [7]探讨了人工智能文学翻译的“主体性”与“创造性”,这些研究从不同角度揭示了人工智能在翻译中的特点,为翻译软件在学术论文翻译中的功能设计提供了参考。王华树和张成智(2025) [8]展望了 GenAI 时代的翻译实践模式,强调技术迭代带来的业态变革,指明翻译软件在学术翻译领域的发展趋势。

在翻译教学与项目研究方面,隋晓冰等(2025) [9]研究了新文科背景下人工智能辅助翻译教学,梅晓明(2024) [10]探讨了人工智能赋能翻译项目的产出策略,这为翻译软件在学术论文翻译教学与实践项目中的应用提供了思路。此外,李禹佳(2024) [11]对生成式人工智能训练合理使用的研究,为翻译软件的数据训练和应用规范提供了理论支持。

3. 研究设计

3.1. 理论基础

本研究以认知负荷理论与信息处理模型为核心理论框架,探索人工智能大模型驱动的翻译工具在跨语言学术论文理解中的效能机制。

Sweller 于 1988 年在认知负荷理论中指出,个体工作记忆的有限性导致认知资源分配直接影响信息处理效率。在跨语言学术阅读场景中,认知负荷可分解为三类:内在认知负荷源于学术论文的复杂性(如高密度专业术语与逻辑结构),外在认知负荷由信息呈现缺陷(如翻译排版混乱、术语误译)引发,而相关认知负荷则指向对信息的深度整合(如通过辅助功能提炼核心观点)。基于此,人工智能翻译工具可能通过精准翻译降低内在负荷、优化排版与图表还原减少外在负荷,并借助摘要生成、逻辑图谱等辅助功能增强相关认知负荷,以此促进知识关联,从而系统性降低研究者认知负担。

信息处理模型进一步从动态视角解析跨语言阅读的认知路径。该模型将信息处理分为输入、加工、存储与输出四个阶段:在输入阶段,翻译工具的术语误译或排版失真可能导致原始信息偏差;在加工阶段,辅助功能(如观点可视化)影响信息整合效率;在存储与输出阶段,工具提供的笔记管理与知识图谱可能强化长期记忆与知识复用。由此可见,大模型驱动的翻译工具通过优化输入质量(精准翻译)与加工效率(智能辅助),有望缩短“信息接收-知识内化”周期,实现学术论文的快速理解。

3.2. 研究假设

基于上述理论框架,本研究提出以下假设:

假设 H1: 相较于传统翻译工具, 大模型驱动的工具能显著降低外在认知负荷, 表现为阅读时间缩短与关键信息检索准确率提升。

假设 H2: 翻译工具的排版还原度与术语准确性通过减少输入阶段的信息偏差, 正向调节内在认知负荷, 进而影响整体理解效能。

假设 H3: 大模型驱动的辅助功能对相关认知负荷的增强效应, 预期其能提升核心观点复述准确率与知识关联完整性。

3.3. 研究对象

本研究以分层抽样方法选取 60 名学术研究者对象, 涵盖计算机科学、医学、经济学三个学科领域(各 20 人), 旨在通过跨学科对比检验翻译软件的辅助效能。样本学科特征鲜明: 计算机科学文献涉及算法描述与技术术语, 医学文献包含复杂病理机制与诊疗标准, 经济学文献强调逻辑推演与数据模型分析。参与者需满足以下条件: 英语水平达 CET-6 标准, 且具有 10 篇以上英文文献阅读经历并持续从事学术研究半年以上, 以确保其对文献阅读流程的熟悉度及对翻译软件效能评估的准确性。此样本设计参考了李群(2025) [6]的研究方法, 通过控制学科差异与语言能力变量, 有效聚焦翻译软件在学术场景中的核心功能, 为后续效能分析提供了可靠的实证基础。

3.4. 研究方法

本研究采用准实验分组设计, 将 60 名研究对象随机分为实验组与对照组(各 30 人)。实验组使用人工智能大模型驱动的“小绿鲸英文文献阅读器”辅助阅读, 全程启用其 AI 问答、术语解释及逻辑图谱功能; 对照组仅可借助基础机械翻译工具(如谷歌翻译基础版)阅读文献原文, 禁用智能辅助功能。实验刺激材料选自 Web of Science 数据库, 从计算机科学、医学、经济学领域各随机抽取 10 篇近三年高影响力期刊论文(共 30 篇), 要求受试者在限定时间内完成核心信息提取任务, 包括研究目的、方法与结论的识别。实验实施双盲控制, 研究者与受试者均不知晓分组目的, 任务文献通过系统随机分配以避免顺序效应。数据采集涵盖客观效能与主观体验两类指标: 客观效能通过关键信息提取准确率, 由两名独立评审对照人工标注基准进行双盲评分和毫秒级计时的任务完成时间量化。

为了进一步验证实验结果, 本研究通过问卷星小程序平台, 针对正在使用人工智能大模型驱动的翻译软件的计算机科学、医学、经济学领域的学生, 如表 1 所示, 结合研究主题从多维度评估 AI 翻译软件的辅助效能, 问卷采用 Likert 5 级量表的形式, 评估内容准确度、信息全面性及工具易用性。

Table 1. Sample statistics

表 1. 样本统计

接触翻译软件频率	每天	每周 3~5 次	每月几次	极少	总计
经济	24	32	44	13	113
专业					
计算机	32	43	41	15	131
医学	25	37	27	17	106
总计	81	112	112	45	350

4. 研究结果

4.1. 描述性统计

本问卷在多方线上平台以有奖征集的方式发布, 用时 48 小时, 在剔除不合理作答时间人群后, 共收

集有效问卷 350 份。结果如表 2 显示，各项功能评分平均值在 1.93~3.57 间，功能表现存在差异且部分评分离散度较高。其中，一键提问功能、AI 生成摘要和可视化解析报告得分靠前，用户认可度高；而阅读效率提升感知一般，在加深文献理解深度方面表现欠佳，得分最低。

Table 2. Descriptive row statistics

表 2. 描述性统计

	N	平均值	标准差
译文能准确传达原文的专业术语含义	350	3.32	1.246
长难句的翻译符合学术表达规范	350	3.45	1.231
对图表和公式等非文本内容的解读准确	350	3.34	1.296
AI 问答功能能纠正原文可能的语义歧义	350	3.27	1.292
文献解析的准确性是否减少反复查阅	350	3.33	1.273
软件提供的背景知识补充有助于理解文献	350	3.37	1.206
AI 生成的摘要能覆盖文献核心论点	350	3.51	1.239
对比传统翻译工具能识别更多隐含逻辑	350	3.40	1.273
界面交互设计符合学术阅读习惯	350	3.37	1.194
文献解析报告的可视化呈现加速重点抓取	350	3.48	1.196
一键提问功能减少传统翻译工具的切换耗时	350	3.57	1.266
文献关键信息提取的便捷性	350	3.35	1.275
使用后文献阅读效率提升幅度	350	3.10	1.039
对比传统方法理解深度的变化	350	1.93	1.042

4.2. 信度检验

对 14 项文献阅读工具使用效果评价指标进行信度检验，结果如表 3 所示，克隆巴赫 Alpha 系数为 0.825，高于 0.8 的标准值，表明量表内部一致性良好，数据可靠性较高，可用于后续分析。

Table 3. Reliability test

表 3. 信度检验

克隆巴赫 Alpha	项数
0.825	14

4.3. 结果分析

1) 信息提取速度与准确性分化显著

本研究基于 350 份有效问卷及对照实验，剖析人工智能大模型驱动的翻译工具在跨语言学术阅读中的效能与矛盾。实验显示，使用“小绿鲸”工具的实验组信息提取速度显著优于对照组，问卷中 48.29% 的用户认为工具使阅读效率提升超 80%，其自动化功能可快速定位核心信息。

然而，效率优势伴随着细节把控不足的风险，实验发现实验组在方法论描述的还原度上表现较差。问卷表明，39.7% 的研究方法关键图表被 AI 忽略，27.3% 的公式翻译因排版失真致逻辑误读。用户易依赖摘要而忽视原文细节，反映出工具虽擅长表层信息提取，却难以完整还原学术深层逻辑，可能影响系统性知识框架构建。

2) 工具功能呈现“强弱分野”

工具效能受术语库覆盖、上下文语义建模及多模态解析技术制约。其在术语解析与逻辑辅助上表现优异，大语言模型驱动的翻译工具可精准转化复杂表述，降低认知负担，84%的用户认可 AI 生成摘要能覆盖核心论点，逻辑图谱助于快速定位重点。

但工具在跨模态与跨学科支持上短板明显，经济学文献模型变量缺失率达 34.5%，计算机科学伪代码缩进丢失致逻辑误读，医学影像难以语义化，跨学科文献知识关联能力弱。尽管大模型缓解了复杂句式(56.29%用户认可)和方法论描述(64.29%用户反馈障碍缓解)的理解困难，长段落逻辑处理仍存断裂风险(2.7 次/篇)，嵌套结构中易误读。多模态解析能力缺失是核心瓶颈，仅 22.86%用户对图表/公式翻译满意，51.14%用户急需“数据可视化解读”，文本与非文本元素语义对齐困难，阻碍学术文献完整解析。

3) 用户反馈揭示矛盾效应

研究揭示人工智能大模型翻译工具存在技术依赖风险与效率深度悖论。低语言水平用户虽获 31.2%准确率提升，但 22.86%用户因过度依赖 AI 产生误解，依赖摘要功能而忽略原文细节，印证技术依赖削弱批判性思考。48.29%用户认为阅读效率提升超 80%，却有 47.71%表示理解深度下降，工具擅于处理表层信息，却难以解析深层逻辑，导致创新性方法论被简化，高阶逻辑整合仍需人工介入。

未来优化应聚焦用户需求，提升多模态解析能力，完善动态术语库，并通过“认知留白”设计，在保障自动化效率的同时，激发用户主动性，推动工具向“认知增强系统”升级。

5. 讨论

5.1. 技术优化层面

1) 工具定位需从“信息提取器”转向“认知脚手架”

研究发现，工具在信息提取速度上的优势(48.29%用户效率提升超 80%)与深层逻辑解析不足(47.71%用户理解深度下降)形成鲜明对比，本质上反映了技术设计对“效率”与“认知价值”的优先级取舍。当前工具过度聚焦表层信息的自动化处理，却在支撑学术研究核心需求上存在结构性缺失，如方法论细节校验、跨模态知识整合、批判性思考训练。

2) 多模态解析能力是突破效能瓶颈的关键枢纽

实验中 39.7%的关键图表被忽略、27.3%的公式因排版失真引发误读，暴露出工具对非文本元素的“语义盲区”。这提示技术开发需突破单一文本处理框架，构建“文本 + 图表 + 公式 + 代码”的多模态语义融合系统。例如，通过光学字符识别(OCR)与结构解析技术，实现图表数据与文本论述的自动关联标注，而非简单忽略或机械翻译，使工具成为连接不同知识载体的认知桥梁。

3) 动态化、场景化的领域适配是技术落地的核心逻辑

跨学科文献语义网络密度低于单学科(反映知识关联构建能力不足)、经济学模型变量缺失率达 34.5%等现象，表明通用模型难以满足细分学科的深度需求。技术开发应转向“基础模型 + 领域插件”的架构，允许用户根据学科特性(如计算机科学的代码逻辑、医学的影像数据)自定义解析规则，形成“千人千面”的智能辅助系统，而非依赖标准化处理流程，平衡自动化与用户控制，避免过度替代引发的认知惰性。

5.2. 用户实践层面

1) 警惕“效率陷阱”，重建“主动校验”的学术阅读伦理

22.86%用户承认“过度依赖 AI 导致误解”，揭示技术使用中“认知惰性”的滋生风险。研究者需建立“工具使用 - 人工校验”的双轨机制：在利用工具快速定位核心信息(如通过逻辑图谱锁定文献结构)后，必须对方法论描述、数据图表、公式推导等关键学术要素进行原文精读，形成“机器初筛 - 人工复

核-深度加工”的闭环，避免将工具输出视为“知识终点”而非“认知起点”。

2) 差异化使用工具，匹配研究阶段的认知需求

工具在信息检索阶段(低阶认知任务)的效能优势与深度分析阶段(高阶认知任务)的局限性，要求用户根据研究目标动态调整使用策略。例如，在文献调研初期可依赖工具快速筛选文献，但在确定研究方向、构建理论框架时，需回归原始文献进行批判性阅读，尤其关注工具可能简化或误读的创新点(如新型统计模型的方法论细节)，防止因技术处理的“标准化过滤”错失学术洞见。

3) 将工具缺陷转化为认知训练的“元思考素材”

实验中工具对复杂句式的处理能力(56.29%用户认可)与长段落逻辑断裂风险并存，提示用户可将工具输出作为“对照样本”而非“标准答案”。用户应理性看待工具的效能边界，在利用其提升信息筛选效率时，保持对关键细节的主动校验，尤其在构建学科知识框架或开展深度分析时，需警惕技术依赖导致的逻辑断层风险。

6. 结论

本研究表明，人工智能大模型驱动的翻译软件在跨语言学术论文快速理解中展现出显著的效率优势与深层逻辑解析局限的双重特征：工具通过自动化语义处理、逻辑图谱构建等功能，能够快速定位核心论点并简化复杂表述，有效提升文献筛选与表层信息提取效率，尤其在术语解析和句式转换中降低认知负荷；但在图表、公式、算法伪代码等非文本元素的语义对齐，以及跨学科知识关联构建、长段落逻辑连贯性还原等方面存在关键短板，导致用户对文献深层方法论与创新逻辑的理解出现断层，且存在过度依赖工具削弱批判性思考的风险。研究揭示此类工具作为“表层信息加速处理工具”的核心定位，其效能边界明确——可高效辅助低阶认知任务(如信息检索与语言转换)，但在支撑高阶学术逻辑整合、多模态内容解析等深层认知需求时，仍需研究者深度介入原文校验与批判性分析。未来需聚焦多模态解析技术突破、领域动态适配能力强化及“人机协同”设计优化，推动工具向兼顾效率与深度的“认知增强系统”演进，为跨语言学术阅读提供更均衡的智能辅助。

参考文献

- [1] Kocmi, T. and Federmann, C. (2023) Large Language Models Are State-of-the-Art Evaluators of Translation Quality.
- [2] 李艳, 朱雨萌, 孙丹, 等. 典型科研场景下生成式人工智能使用的差异性分析——学科背景与人工智能素养的影响[J]. 现代远程教育研究, 2025, 37(2): 92-101+112.
- [3] 卢云峰, 黄星尧, 吴语菲, 等. 生成式人工智能的兴起对社会科学研究方法的拓展——以偏见研究为例[J/OL]. 贵州民族大学学报(哲学社会科学版), 1-22. <http://kns.cnki.net/kcms/detail/52.1155.c.20250327.1042.002.html>, 2025-04-01.
- [4] 赵正平. 人工智能大语言模型和 AI 芯片的新进展[J]. 微纳电子技术, 2025, 62(3): 7-37.
- [5] 李群. ChatGPT 在翻译领域中的应用: 机遇、挑战与应对[J]. 湖南科技学院学报, 2025, 46(1): 77-82.
- [6] 刘心怡, 何高大. 人工智能背景下中国古诗词名篇翻译欣赏——以几首经典词作人机翻译对比为例[J]. 名作欣赏, 2025(6): 23-26.
- [7] 袁筱一. 人工智能文学翻译的“主体性”与“创造性”[J]. 上海交通大学学报(哲学社会科学版), 2025, 33(1): 1-10.
- [8] 王华树, 张成智. GenAI 时代的翻译实践模式: 技术迭代、业态变革与趋势展望[J]. 外语教学, 2025, 46(1): 53-58.
- [9] 隋晓冰, 周宇薇, 谢洪. 新文科背景下人工智能辅助翻译教学研究[J]. 上海翻译, 2025(1): 52-56.
- [10] 梅晓明. 人工智能赋能翻译项目的产出策略研究[J]. 英语广场, 2024(33): 23-26.
- [11] 李禹佳. 生成式人工智能训练合理使用研究[D]: [硕士学位论文]. 北京: 中国政法大学, 2024.