

# 翻译风格研究综述

翻译风格既是译者个体选择的体现，也是不同翻译模式差异化的重要标志，长期以来为翻译学研究所关注。随着神经机器翻译与大语言模型的介入，风格问题在机器译文与人工译文的比较中进一步凸显：机器译文能否再现原作的文学风格、不同翻译路径在风格上呈现出何种系统性差异，已成为近年文学翻译与机器翻译交叉领域的核心议题。既有研究一方面从语料库与计量视角切入，考察翻译文本的风格特征，如张继光（2016）借助科学知识图谱对国内语料库翻译学的发展脉络进行了梳理；另一方面则在神经机器翻译的框架下探讨风格的迁移与适应，如宁秋怡等（2022）提出基于风格感知的无监督领域适应算法，以应对电商文本中的风格差异。Yao（2025）更以熵值方法比较了人工翻译、神经机器翻译与大语言模型译文在“简化”这一风格维度上的差异。围绕这一议题，国内外研究在关注侧重与方法路径上均呈现出不同取向。下文将分别梳理国外研究现状与国内研究现状，并在此基础上展开比较与反思。

## 一、国外研究现状

翻译风格研究在国际学界由来已久，但近年来随着神经机器翻译（NMT）与大语言模型的快速发展，这一议题被重新置于人机对比的语境中加以审视。早期以语料库语言学为基础的风格考察，主要关注译者个人偏好与文本特征之间的对应关系；而随着机器翻译在文学与非文学领域的渗透，研究重心逐步转向“机器译本是否具备风格意识”以及“何种风格属性可被量化评估”这两个问题。近期一批实证研究（Pastor, 2024; Yao, 2025; Gao, 2024）通过对比人工翻译、NMT与大语言模型在文学文本上的表现，试图厘清风格在不同翻译主体之间的差异及其成因。与此同时，另一批研究（Araghi, 2024; Leiter, 2024; Zhang, 2025）则致力于从评估指标与质量估计角度切入，探索如何让风格这类较为隐性的维度进入自动评价体系。这两条路径之间尚存明显张力：前者更关注译本的文学性与创造性，后者则受制于可计算性，难以真正把握文本风格的微妙层次。

第一，关于机器翻译在文学文本中再现原文风格的能力，这一议题构成了国际学界近年最为活跃的讨论场域之一。文学翻译长期被视作“人类翻译的最后堡垒”，其根据正在于文体、语气、修辞等风格要素难以被算法充分处理。Wang (2023a) 从社会学视角出发，讨论了 AI 介入文学翻译所引发的职业边界与主体性争议，指出尽管 NMT 在一般文本上进展显著，但在文学领域仍面临根本性的风格再现困境。这一判断得到多项实证研究的呼应：Pastor (2024) 通过考察文学文本中被改造的多词表达 (MWE) 发现，NMT 在创造性表达的处理上与人工译者之间存在明显落差；Ferragud (2023) 对比机器译本与学生译本后也观察到，机器译文在错误类型上呈现出较为规律化的模式，与人类译者出现的非系统性误差形成对照。随着大语言模型的出现，这一图景被部分改写。Gao (2024) 比较了 ChatGPT、Google Translate 与 DeepL 在中国古典诗歌翻译中的表现，发现 ChatGPT 在意象与节奏的处理上优于传统 NMT；Ed-Dali (2025) 对 DeepSeek R1 与 ChatGPT 4.5 在阿英文学翻译中的对比研究也得出类似结论。You (2025) 则以《西游记》的葡语译本为案例，揭示即便是最先进的系统，在面对文化负载与语体层次丰富的古典文本时仍显乏力。这一议题的方法论也从早期以错误分析为主的定性研究，逐步扩展至定性定量混合的设计。理论贡献在于打破了“文学翻译不可机器化”的简单二分，但仍未能回答一个根本问题：当机器译本在表层流畅度上接近人工译文时，其风格差异究竟源于模型架构、训练数据，还是源于“风格”本身作为一种文化实践的不可还原性。

第二，关于翻译风格的量化刻画与自动评估，近期研究呈现出从表层指标向深层特征转移的趋势。早期以 BLEU 为代表的 n-gram 指标难以捕捉风格层面的差异，这一局限在 (Leiter, 2024) 对可解释性评估指标的梳理中得到确认，该研究指出当前基于大模型的 COMET、BERTScore 等指标虽然表现更优，但其“黑箱”属性使风格判断缺乏透明依据。Mrinalini (2022) 则提出基于句向量相似度的 SBSim 指标，试图在语义层面更贴近人工判断，这一思路后续在 Zhang (2025) 关于中英机器翻译质量估计的极小极大优化机制研究中得到进一步技术化。与这些以指标改进为目标的研究不同，Yao (2025) 采用熵值方法考察了人工译文、NMT 与大语言模型译文在“简化”这一风格特征上的差异，发现 ChatGPT 译文的简化程度介于 NMT 与人工翻译之间，这一发现既挑战了“机

器译文必然简化”的旧有假设，也提示不同翻译主体之间的风格分布呈现连续谱而非截然两分。Araghi (2024) 则从另一角度切入，探讨源文本的翻译难度特征是否可以预测机器翻译质量，为理解风格再现的影响因素提供了新的解释路径。方法上，这一议题的研究普遍结合了语料统计、神经网络嵌入和人工评分，呈现出计算翻译学与传统译学方法的融合态势。但计算视角下的“风格”往往被化约为可测变量，对于风格作为译者审美选择的一面仍难触及，这一张力构成该路径尚未克服的理论局限。

第三，关于翻译工具与工作流程对译者风格的塑造作用，相关讨论将风格研究推向了更具社会学色彩的方向。Vieira (2023) 通过考察 CAT 工具中的分段设置对科幻小说翻译的影响，指出机器辅助环境下的句段切分本身就在重塑译者的风格判断；Mah (2020) 更早地在英韩语对的文学机器翻译中提出面向特定文本类型的译后编辑指南，强调不同语言组合下的风格再现需要差异化的后编辑策略。Wang (2023b) 将视野扩展至视听翻译教学场景，从可读性角度分析 NMT 在中文影视翻译训练中的适用性，触及了媒介形态对译文风格的约束。Ouldelhaj (2025) 从宏观层面勾勒了 AI 与文学翻译的关系图景，强调技术介入下译者角色的重新定位。这一议题的方法论以个案研究与行动研究为主，较少依赖大规模语料。它的贡献在于把风格从“文本属性”重新界定为“工作流程的产物”，提示研究者必须关注生产条件；但现有研究多停留于现象描述，对工具-译者-文本三者如何在长期实践中形成稳定的风格模式缺乏历时追踪。

第四，关于低资源语言与特定语对中的风格再现问题，这一方向虽然规模不大，却揭示了主流风格研究的隐含偏倚。Tafa (2025) 对低资源语言机器翻译的系统综述指出，数据稀缺不仅影响译文的准确性，也使得风格迁移几乎无法实现，因为模型缺乏足够的风格样本进行学习。Asebel (2025) 以阿姆哈拉语-英语为例追踪了该语对机器翻译的演进，发现从规则系统到神经网络的转变虽提升了流畅度，但在风格一致性上仍有明显不足；Lin (2020) 提出的改进版束搜索算法则试图在英汉翻译中通过解码策略的优化来改善译文的表达自然度。Wang (2025) 则从数据增强角度切入，探索如何在 NMT 训练中通过随机复合操作生成更丰富的语料，从而间接影响译文的风格多样性。Karakus

(2025) 在医疗场景中考察 NMT 的实际部署，虽非严格意义上的风格研究，但其对译文“语域恰当性”的讨论与风格再现紧密相关。这一议题的理论意义在

于提醒学界：关于翻译风格的主流话语长期建立在英语及其他高资源语言之上，低资源语对的风格问题既受制于技术条件，也受制于研究关注度的不平衡。

国际研究已形成一幅围绕“翻译风格的可计算性”展开的知识图景：一端是针对文学文本的定性考察，强调风格作为创造性实践的不可化约性（（Pastor, 2024; Wang, 2023a; You, 2025））；另一端是以指标改进与模型优化为目标的计算研究，试图让风格进入可评价的范畴（（Yao, 2025; Leiter, 2024; Zhang, 2025））。两端之间由工具研究与特定语对研究加以连接。主要理论贡献，在于把“翻译风格”从传统文体学议题扩展为涉及人机协作、评估体系与生产条件的综合问题。存在的盲点也较为明显：一是现有研究多以英语为中心或以高资源语对为主，非西方语言及跨文化语境下的风格差异仍未获得充分探讨；二是“风格”一词在不同研究中所指并不统一，有时指词汇-句法特征，有时指审美倾向，概念界定的模糊削弱了跨研究的对话性；三是对大语言模型介入后译者主体性与风格形成机制的变化，尚缺乏系统的社会学与伦理学审视（（Ouldelhaj, 2025; Ed-Dali, 2025））。对国内研究的启示主要有两点：其一，在方法论上需要把语料库路径与人机对比实验相结合，避免把风格研究简化为指标竞赛；其二，在理论上应当重新厘清“译者风格”“机器风格”与“译文风格”之间的概念层级，使中文语境下的风格研究具备与国际对话的可比性基础。

## 二、国内研究现状

翻译风格研究在国内学界的兴起，与语料库翻译学的本土化进程及近年来人工智能技术对翻译领域的冲击密切相关。从研究路径看，国内相关讨论大致沿两条线索展开：一条依托语料库方法对译者风格、文本风格进行描写性考察，另一条则随着神经机器翻译与大语言模型的介入，转向对机器译文风格可控性的技术探索。张继光（2016）通过科学知识图谱的方式对1993至2014年间的语料库翻译学研究进行了整体梳理，显示出这一领域在方法论层面已积累起相当规模的成果，但也指出论文质量良莠不齐、研究层次有待深化的问题。进入2022年之后，随着领域适应、风格迁移、智能体协同等技术议题的涌入（杨晨等，2025；刘淼，2024；宁秋怡等，2022），翻译风格不再只是文学译

本比较的对象，而逐渐成为机器翻译系统在垂直领域落地时无法回避的工程问题。当前国内研究的分歧在于：翻译风格究竟是译者主体性的语言学外显，还是可以被建模、迁移与评估的可计算对象？两种取向之间的对话尚未充分展开。

第一，关于语料库翻译学路径下翻译风格研究的总体格局。此议题关涉国内翻译风格研究的方法论底色。张继光（2016）借助 CiteSpace 工具对二十余年间的文献进行计量分析，描绘了国内语料库翻译学的发展脉络，指出该领域在高频关键词、高产作者与机构分布上已形成相对稳定的学术共同体。这一工作的价值在于，它用可视化手段确认了“风格”作为语料库翻译学核心议题之一的位置，也为后续研究提供了知识基础。然而该研究同时指出，国内语料库翻译学研究存在论文质量参差不齐、研究层次偏浅的问题，许多成果停留在词频统计与类符形符比的描述层面，缺乏对风格背后认知机制与社会语境的深入追问。这一判断放在今天看仍具参考价值：国内翻译风格的语料库研究多以单一文本或单一译者为对象，宏观的比较性研究与跨语种、跨时代的系统考察相对稀少。更为根本的是，该路径所积累的描写性成果如何与新近兴起的神经翻译、大模型研究形成对话，目前尚缺乏清晰的议程。这使得语料库路径虽然在方法上成熟，却在面对新的技术语境时显得较为被动。

第二，关于神经机器翻译中风格差异处理的研究。这一议题直接回应了机器译文在垂直领域应用时暴露出的风格失配问题。宁秋怡等（2022）针对电子商务领域平行语料稀缺及中外产品信息表达风格差异的双重困境，提出了一种基于风格感知的无监督领域适应算法，通过互训练方法利用单语数据，并引入拟知识蒸馏机制处理风格差异。这一工作的贡献在于，它把“风格”从一个模糊的修辞学概念转化为可被神经网络感知和调节的技术变量，使得机器翻译系统不仅追求语义等价，也开始关注语域与表达方式的适配。顺着这一思路，杨晨等（2025）把视线投向航空手册这一高度规范化的技术文档领域，提出了一种简化技术英语引导的偏向性检索增强集成翻译模型，使译文能够符合受控自然语言在语法和词汇层面的严格规则。两项研究虽然针对不同领域，却共同指向一个方法论转向：翻译风格研究开始从对既有译本的事后描写，转向对译文生成过程的事前干预与约束。这一转向具有方法论意义，但其局限也较为明显：风格在模型中往往被简化为词汇分布或句法规则的统计特征，对于文学翻译中

更为细腻的语气、节奏与审美意蕴，现有技术路径仍难以处理。此外，这类研究多发表于计算机学科期刊，与翻译学界内部的风格理论讨论缺乏实质对话。

第三，关于大语言模型介入下翻译风格一致性的研究。这是近两年国内研究中新兴的方向。刘淼（2024）以《习近平谈治国理政》（第四卷）及其俄译本为语料，提出了多智能体互动翻译框架 MAGIC-PTF，实验显示该框架在术语准确性、风格一致性和整体质量等指标上均优于主流的机器翻译系统。这项研究之所以值得关注，在于它把“风格一致性”作为与术语准确性并列的评价维度，正式纳入大模型翻译质量评估体系。这一做法既延续了语料库翻译学对风格的量化关切（张继光，2016），又借助多智能体架构实现了对生成过程的分工控制，在一定程度上回应了前述神经翻译研究中风格处理粗糙的问题（宁秋怡等，2022）。然而该研究也反映出国内此类工作的共性局限：首先，风格一致性的评价仍主要依赖与官方译本的对齐程度，这种“以官方译本为金标准”的做法，潜在地预设了某一译本的风格合法性，回避了翻译风格的多元可能。其次，时政文本与文学文本在风格的约束性上差异悬殊，将前者的经验推广到后者需要更审慎的论证。再次，此类研究多由技术驱动，翻译学界对其评估框架的学理介入尚显不足。

从整体看，国内翻译风格研究已初步形成语料库描写与技术建模两条并行路径，前者在方法层面相对成熟（张继光，2016），后者在应用层面推进较快（杨晨等，2025；刘淼，2024；宁秋怡等，2022）。但若与国际学界近年关于译者风格、文本性与风格迁移的讨论相对照，国内研究至少存在三方面不足。其一，宏观论述与技术报告多，真正扎根于文本细读与译者主体性考察的微观实证研究偏少，风格研究容易被化约为数据指标。其二，学科之间壁垒明显，翻译学、语料库语言学与计算机科学各自为政，像（刘淼，2024）这类技术性成果在翻译学界的理论回响有限，而翻译学内部的风格理论也较少被纳入模型设计的前端考量。其三，研究对象集中于政治文献、商务文本与技术手册等规范性较强的语料，文学翻译风格尤其是诗歌、戏剧等文体在新技术语境下的处理问题讨论不足。未来研究有必要在以下方向继续推进：建立兼顾人文阐释与计算建模的风格分析框架，开展跨语种、跨文体的系统比较，以及在大模型时代重新审视“译者风格”这一概念本身的理论边界。

### 三、国内外研究的比较与反思

翻译风格研究在国内外学界的演进轨迹并不一致。国外研究大体沿着机器翻译与译者研究两条平行线索推进，近年逐渐在神经机器翻译、计算机辅助翻译与文学翻译的交叉地带形成议题（Ferragud, 2023; Vieira, 2023; Wang, 2023a）；国内相关成果则更多嵌入具体的应用场景，服务于特定文本类型的翻译质量提升（杨晨等, 2025; 刘淼, 2024; 宁秋怡等, 2022）。这种分化不只是选题的偶然差异，背后折射出两种不同的问题意识：国外研究更倾向于把翻译风格作为一个需要被理论化、被质疑、被重新界定的概念对象，而国内研究更倾向于把风格视为一个需要被达成、被控制、被优化的技术指标。理解这种立场差异，是判断两者能否真正形成对话的前提。下文从关注点、方法论以及共同盲点三个层面展开。

第一，研究关注点的深层差异与互补。国外研究的关注点集中在风格在机器翻译中的流失、扭曲与再生产问题上。文学机器翻译中译文风格同质化、译后编辑规范的制定、人机风格差异的错误分析，构成了稳定的议题群

（Ferragud, 2023; Wang, 2023a; Mah, 2020）。国内研究的关注点则更多落在“如何让译文风格符合某种既定规范”上，例如航空手册要求符合简化技术英语规范（杨晨等, 2025）、时政话语要求术语准确与风格一致（刘淼, 2024）、电商文本要求跨文化的风格适配（宁秋怡等, 2022）。表层看是议题之别，深层则是知识生产逻辑的不同。国外研究承接了描写翻译学与文学翻译社会学的传统，倾向于把风格当作一个需要被观察和解释的现象，关心其生成机制；国内研究更多承接应用语言学与工程导向的研究范式，把风格当作一个可操作、可评测的目标变量。两种立场各有盲区：脱离国内应用场景的国外研究容易停留在对“风格危机”的忧虑，而难以回应实际生产中的风格调控需求；缺少国外批判视角的国内研究则容易把“风格一致性”简化为术语与句式的对齐，忽略风格背后的文学性、主体性与文化差异问题（Wang, 2023a）。当前两者之间的对话并不充分，国内研究较少援引国外对机器翻译风格同质化的批判讨论，国外研究也较少关注中文语境下受控语言与领域适应带来的风格新议题。

第二，研究方法的差异与可相互借鉴之处。国外研究方法呈现出较为多元的分布：既有基于错误类型学的对比分析（Ferragud, 2023），也有基于可读性指标的实证评估（Wang, 2023b），还有引入社会学视角对译者劳动与行业结构的考察（Wang, 2023a），以及语义相似度等自动评价指标的开发（Mrinalini, 2022）。国内研究则相对集中在模型构建与算法改进上，通过检索增强、多智能体协作、无监督领域适应等技术路径来改善风格表现（杨晨等, 2025；刘淼, 2024；宁秋怡等, 2022）。这种方法差异折射出不同的认识论取向。国外研究倾向经验优先，从个案深描或小规模实验中提炼对风格的理解，承认风格的复杂性与不可完全量化性；国内研究倾向工程优先，预设风格可以被分解为若干可测量的特征，并通过数据与模型加以逼近。两种路径可以互相补足。国内研究若引入错误类型学和译者访谈等质性方法，可以更好地回答“模型究竟在什么意义上改善了风格”这一问题，而不只是依赖自动评价分数的提升（Ferragud, 2023；Mrinalini, 2022）。国外研究若关注中文语境下针对特定领域的受控生成实验，也有助于把抽象的风格讨论落到可验证的操作层面（杨晨等, 2025；宁秋怡等, 2022）。早期国内语料库翻译学的知识图谱分析已经提示，方法层面的整合仍有较大空间（张继光, 2016）。

第三，共同盲点与研究前沿。两类研究共有若干薄弱环节。其一，对“风格”本身的概念界定仍显粗糙。无论是机器翻译中的文学风格（Ferragud, 2023；Wang, 2023a），还是领域翻译中的规范风格（杨晨等, 2025；刘淼, 2024），多数研究默认风格可由词汇、句法、可读性等表层特征近似刻画，较少追问这些特征与译者主体性、读者感知之间的关系（Mah, 2020）。其二，译者作为风格承担者的角色在两类研究中被弱化。国外研究虽有社会学转向的尝试（Wang, 2023a），但对译者与机器协作过程中的风格协商过程仍缺少细致刻画；国内研究则更多把译者置于“后编辑者”的位置，对其在风格形成中的能动性探讨不多。其三，评价指标与风格判断之间的落差问题尚未获得足够关注，自动指标能否真正捕捉风格差异，仍是一个开放问题（Mrinalini, 2022）。这些盲点的存在，与学科壁垒、数据获取的不对称以及理论工具的相对单一都有关系。未来的研究或可在风格特征的可计算定义、译者与模型的交互日志分析、以及跨语对的风格对比实验上寻找突破口，把社会学视角、语料库方法与大语言模型评测结合起来。

整体而言，翻译风格研究目前处于方法快速更新但概念基础相对滞后的阶段。国外研究在理论反思上走得较远，国内研究在技术落地上推进较快，两者之间尚未形成稳定的相互引用关系（Wang, 2023a; 刘淼, 2024; 张继光, 2016）。一个值得进一步探讨的空白，是如何在大语言模型介入翻译生产的新条件下，重新界定“风格”这一概念，并建立起兼顾可计算性与文学解释力的分析框架（Mrinalini, 2022; Mah, 2020）。这一空白既涉及基础概念层面的澄清，也涉及方法工具的重构，构成后续研究可以切入的具体问题。

## 选题建议

基于本综述的研究空白与争议点，以下是 10 个可供参考的研究选题：

### 【1】基于熵值谱系的汉译风格连续体重构

Yao (2025) 提出人工、NMT 与大语言模型译文在“简化”维度上呈连续谱而非二分对立，这一发现挑战了传统人机对立的研究框架，但其研究以英语语料为主，汉语语境下是否成立仍待验证。建议以鲁迅、张爱玲等现代汉语文学作品的英译为语料，构建涵盖人工翻译、NMT (Google/DeepL) 与多款大模型 (GPT-4、DeepSeek、文心) 的平行语料库，采用熵值、类符形符比、平均句长等多维指标刻画“简化”“显化”“规范化”三大翻译共性在不同译者主体间的分布谱系。预期贡献在于把翻译共性假设从二元对立推进到连续体模型，并为汉英方向提供本土化的计量证据。

### 【2】大语言模型古典诗歌翻译的意象-节奏双维评估

Gao (2024) 发现 ChatGPT 在中国古典诗歌意象与节奏处理上优于传统 NMT，但其评估框架较为粗糙，缺乏对意象等价、格律还原的系统化指标。建议选取《唐诗三百首》中具有代表性的律诗、绝句各 50 首，构建意象密度指数与节奏对齐度两个可操作化指标，对比 ChatGPT-4o、Claude、DeepSeek 与 Google/DeepL 的译文表现，并引入母语读者的审美评分作为外部效度检验。研究可填补大模型诗歌翻译评估的方法论空白，厘清模型架构与训练数据对诗学风格再现的差异化影响，推动计量诗学与机器翻译评估的交叉对话。

### 【3】低资源语对的汉语方言文学机器翻译风格再现

Tafa (2025) 与 Asebel (2025) 揭示低资源语言的风格迁移因样本稀缺而几近失效, 但主流研究仍集中于非洲或南亚语言, 汉语内部的方言文学 (如粤语、吴语、闽南语小说) 同样面临低资源困境, 却未获关注。建议以《繁花》(沪语)、香港粤语小说等方言文学为语料, 考察 NMT 与大模型在方言标记性词汇、语气助词、句法倒装等风格特征上的再现能力, 结合数据增强 (参考 Wang, 2025) 与少样本提示工程进行干预实验。预期贡献在于把低资源风格研究从跨语言维度扩展到语言内部的方言维度, 丰富翻译风格研究的社会语言学视野。

#### **【4】"译者风格""机器风格""译文风格"的概念层级辨析**

综述明确指出"风格"一词在不同研究中所指不一, 有时指词汇-句法特征, 有时指审美倾向, 概念界定的模糊削弱了跨研究的对话性。建议以概念史与元研究方法, 系统梳理 1993 年以来中英文翻译学核心期刊中"风格"相关术语的使用语境与定义差异, 构建"译者风格, 机器风格, 译文风格, 文本风格"的四层概念框架, 并以若干典型研究为案例检验该框架的区分力。研究的理论贡献在于为大模型时代的风格研究奠定术语基础, 推动语料库翻译学、计算翻译学与传统文体学之间的概念对接。

#### **【5】多智能体翻译框架在文学文本中的风格一致性检验**

刘焱 (2024) 提出的 MAGIC-PTF 框架在时政文本风格一致性上表现优异, 但综述指出时政与文学文本在风格约束性上差异悬殊, 其经验能否迁移至文学领域存疑。建议复现并改造 MAGIC-PTF 架构, 设计专门面向文学翻译的多智能体分工 (语气代理、修辞代理、文化负载代理、一致性审校代理), 以莫言或余华小说英译为测试集, 与单一大模型译文及人工译文进行对照, 采用风格指纹分析与人工审美评分双轨评估。预期揭示多智能体架构在文学风格再现上的潜力与边界, 推动技术与文学翻译理论的对话。

#### **【6】CAT 工具分段设置对译者风格的历时塑造研究**

Vieira (2023) 指出 CAT 工具的分段设置在重塑译者风格判断, 但现有研究多为横截面的现象描述, 缺乏对工具-译者-文本长期互动的历时追踪。建议选取 3-5 位从业 10 年以上的文学译者, 收集其使用不同 CAT 工具 (Trados、MemoQ、OmegaT) 前后的译作语料, 结合深度访谈与语料计量方法, 考察句

段长度、衔接手段、句法复杂度等风格指标的历时变迁。研究的社会学贡献在于把翻译风格从文本属性重新界定为"工作流程的产物", 回应综述中提出的生产条件视角缺口。

### 【7】大模型介入下译者主体性的伦理与身份重构

综述指出对大语言模型介入后译者主体性与风格形成机制变化的社会学与伦理学审视尚显不足 (Ouldelhaj, 2025; Ed-Dali, 2025)。建议采用质性研究方法, 通过对国内 30 位文学译者与技术译者的深度访谈, 结合他们的译后编辑过程数据 (击键记录、修改痕迹), 考察译者在与大模型协同时的主体性让渡、风格妥协与身份焦虑。研究可引入 Pym 的译者伦理框架与 Bourdieu 的场域理论, 构建"人机协同译者"的主体性谱系模型。学术贡献在于补充技术主导研究中缺失的人文关怀与伦理维度。

### 【8】可解释性风格评估指标的构建与验证

Leiter (2024) 批评 COMET、BERTScore 等指标的"黑箱"属性使风格判断缺乏透明依据, Mrinalini (2022) 的 SBSim 仅在语义层面改进。建议在现有句向量相似度的基础上, 设计分层可解释的风格评估指标 StyleX, 将风格分解为词汇丰富度、句法复杂度、修辞密度、语域一致性四个子维度, 每个维度配以可视化解释模块, 并在汉英文学翻译测试集上与 COMET、BLEURT 对比其与人工评分的相关性。预期贡献在于提供兼具解释力与判别力的风格评估工具, 推动自动评价体系向"风格透明"方向演进。

### 【9】受控自然语言引导下的技术文本风格可控翻译扩展

杨晨等 (2025) 在航空手册领域提出简化技术英语引导的检索增强翻译模型, 但其经验尚未扩展到医疗、法律、专利等其他强规范性领域。建议选取医疗知情同意书与法律合同两类文本, 借鉴 Karakus (2025) 对医疗 NMT 语域恰当性的讨论, 构建领域专属的受控自然语言规则库, 设计检索增强+规则约束的混合翻译框架, 并以"风格合规率"替代传统 BLEU 作为核心评估指标。研究可推动技术文本翻译的风格可控性研究从单一领域走向跨领域方法论提炼, 兼具学术与产业价值。

### 【10】译文难度特征对风格再现质量的预测模型

Araghi (2024) 探讨源文本翻译难度是否可预测机器翻译质量, 但其研究聚焦一般质量而非风格维度, 难度特征与风格再现之间的关系尚未明确。建议构建一个涵盖词汇罕见度、句法嵌套深度、文化负载密度、修辞标记数四类难度特征的源文本刻画体系, 以 1000 段汉英文学翻译样本为数据, 训练可预测"风格再现偏离度"的回归模型, 并检验不同翻译主体 (NMT、大模型、人工) 在面对高难度文本时的风格退化模式。预期贡献在于把翻译难度研究从"质量预测"推进到"风格预测", 为译前难度评估与译后风格修正提供理论依据。

## 参考文献

- [1] Pastor, GC. Human versus Neural Machine Translation Creativity: A Study on Manipulated MWEs in Literature. INFORMATION, 2024.
- [2] Yao, GY. An entropy-based study of Simplification in ChatGPT translations compared to neural machine translation and human translation across genres. PLOS ONE, 2025.
- [3] Gao, RY. Machine translation of Chinese classical poetry: a comparison among ChatGPT, Google Translate, and DeepL Translator. HUMANITIES & SOCIAL SCIENCES COMMUNICATIONS, 2024.
- [4] Araghi, S. The link between translation difficulty and the quality of machine translation: a literature review and empirical investigation. LANGUAGE RESOURCES AND EVALUATION, 2024.
- [5] Leiter, C. Towards Explainable Evaluation Metrics for Machine Translation. JOURNAL OF MACHINE LEARNING RESEARCH, 2024.
- [6] Zhang, XM. Application of Minimax Optimization Mechanism in Chinese-English Machine Translation Quality Estimation. IEEE ACCESS, 2025.
- [7] Wang, HT. Defending the last bastion A sociological approach to the challenged literary translation. BABEL-REVUE INTERNATIONALE DE LA TRADUCTION-INTERNATIONAL JOURNAL OF TRANSLATION, 2023a.
- [8] Ferragud, MF. La traducción automática literaria: análisis de errores de la traducción automática y las traducciones de estudiantes del mismo texto original. TRADUMATICA-TRADUCCIO I TECNOLOGIES DE LA INFORMACIO I LA COMUNICACIO, 2023.
- [9] Ed-Dali, R. Assessing DeepSeek R1 and ChatGPT 4.5 in Arabic-English literary translation: performance, challenges, and implications. COGENT ARTS & HUMANITIES, 2025.
- [10] You, M. How well can state-of-the-art machine translation systems render a 16th-century Chinese novel?. CADERNOS DE TRADUCAO, 2025.

- [11] Mrinalini, K. SBSim: A Sentence-BERT Similarity-Based Evaluation Metric for Indian Language Neural Machine Translation Systems. *IEEE-ACM TRANSACTIONS ON AUDIO SPEECH AND LANGUAGE PROCESSING*, 2022.
- [12] Vieira, LN. Translating science fiction in a CAT tool: machine translation and segmentation settings. *TRANSLATION & INTERPRETING-THE INTERNATIONAL JOURNAL OF TRANSLATION AND INTERPRETING*, 2023.
- [13] Mah, SH. Defining language dependent post-editing guidelines for specific content The case of the English-Korean pair to improve literature machine translation styles. *BABEL-REVUE INTERNATIONALE DE LA TRADUCTION-INTERNATIONAL JOURNAL OF TRANSLATION*, 2020.
- [14] Wang, QR. Neural machine translation in AVT teaching in China: An in-depth analysis from the readability perspective. *LINGUISTICA ANTVERPIENSIA NEW SERIES-THEMES IN TRANSLATION STUDIES*, 2023b.
- [15] Ouldelhaj, D. ARTIFICIAL INTELLIGENCE AND LITERARY TRANSLATION: REALITY AND PROSPECTS. *PERSPECTIVAS DE LA COMUNICACION*, 2025.
- [16] Tafa, TO. Machine Translation Performance for Low-Resource Languages: A Systematic Literature Review. *IEEE ACCESS*, 2025.
- [17] Asebel, MH. Exploring the evolution and future prospects of Amharic to English machine translation: a systematic review. *FRONTIERS IN ARTIFICIAL INTELLIGENCE*, 2025.
- [18] Lin, XY. A Novel Beam Search to Improve Neural Machine Translation for English-Chinese. *CMC-COMPUTERS MATERIALS & CONTINUA*, 2020.
- [19] Wang, HJ. RMixDA: A Random Compound Data Augmentation Operation Framework for Neural Machine Translation. *IEEE ACCESS*, 2025.
- [20] Karakus, IS. Bridging language gaps in healthcare: a systematic review of the practical implementation of neural machine translation technologies in clinical settings. *JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION*, 2025.
- [21] 张继光. 国内语料库翻译学研究状况的科学知识图谱分析(1993—2014)[J]. 上海翻译, 2016(34-40+61+93).
- [22] 杨晨, 叶娜, 张桂平. 面向航空手册的偏向性检索增强集成翻译模型[J]. 计算机科学, 2025(112-121).
- [23] 刘淼. 大语言模型赋能中国时政话语俄译研究——多智能体互动翻译框架的创建与应用[J]. 外语电化教学, 2024(15-22+107).
- [24] 宁秋怡, 史小静, 段湘煜, 等. 基于风格感知的无监督领域适应算法[J]. 计算机科学, 2022(271-278).