

基于 DeepSeek 的 MTI 研究生科技笔译语用能力评价研究

马颖¹ 邓鹂鸣²

1. 沈阳理工大学 沈阳 110168

2. 武汉大学 武汉 430072

摘 要: 本研究基于《中国英语能力等级量表》、借助 DeepSeek 构建了“科技笔译语用能力等级评价体系”，聚类到专业性、准确性、功能性及文化适应性四个一级维度，并采用非线性赋分原则，以更精准地反映科技翻译能力的非线性发展特征。实证研究表明，MTI 研究生的科技翻译能力集中在基础水平，即 5—6 级，高阶语用能力明显不足。研究建议翻译教育应转向“语用能力驱动”，并建立量化的认证标准，以提升科技翻译的质量与可靠性，为未来的个性化培养提供理论依据。

关键词: 科技笔译；语用能力；MTI 研究生；《中国英语能力等级量表》；DeepSeek

随着全球科技交流的深化，科技笔译需求激增，但是传统翻译评价多聚焦语言准确性，而忽视了语用能力。有研究者发现，非英语专业学生的科技翻译译文中存在词汇、语法、语用失误问题^[1]。建立科技笔译语用能力等级评价体系，可以规范行业标准，为人才培养提供量化依据，强化译者“专业+语用”的双重能力，避免因表述上的歧义而引发技术或法律问题。该体系的核心挑战在于量化语用能力、适配不同领域特征及跨文化差异，这一探索对提升科技翻译的专业性和可靠性具有重要意义。

1. 研究方法

DeepSeek 通过对相关技术算法、体系架构与创新思维的极致优化，实现了大规模知识资源的低成本训练和高效输出^[2]，为本研究提供了研究工具。本研究采用定性与定量相结合的混合研究方法，以《中国英语能力等级量表》为基础框架，采用 DeepSeek 文本分析技术对科技翻译语料进行深度挖掘，提取核心评价维度；运用层次分析法（AHP）确定各维度权重并构建五级评价量度，再通过实证检验确保评价体系的科学性与适用性，最终形成一套兼具客观量化标准和专业语用特征的科技笔译能力等级评价体系，为翻译人才培养和质量评估提供可靠依据。

2. “科技笔译语用能力等级评价量表”维度解析

《中国英语能力等级量表》对笔译能力和语用能力进行了描述，并提供了“笔译能力总表”和“语用能力总表”^[3]。本研究对上述两个总表进行文本分析，从中提取了科技笔译的语用需求，初步获取六个维度，利用 DeepSeek 进行维

度聚类后获得四个一级维度和八个二级指标，即：专业性（术语准确性、专业文本适应性）、准确性（逻辑与结构严谨性、技术规范符合度）、功能性（信息完整性、功能明确性）、文化适应性（信息适应性、传播性）。这也符合科技英文客观、准确、逻辑性特点^[4]。“科技笔译语用能力等级量表”对科技翻译语用能力的四个维度进行了九个等级的划分，旨在全面评估翻译者在不同情境下的表现。

在专业性维度上，九级能精准翻译高度专业化术语，处理前沿科技文献；八级能准确翻译多领域专业术语，翻译专业应用文本时简洁无歧义；七级能翻译熟悉领域的科技术语，但翻译科普类文章时需要校对；六级能识别常见专业术语，但翻译大众科技文章时需要简化内容；五级依赖词典翻译日常科技词汇，能处理非专业科技文本，但需要大量简化。

在准确性维度上，九级译文符合国际/行业技术规范，且内容无遗漏、信息无歧义；八级遵循常见技术规范，能够准确翻译关键信息，需要进行少量校对；七级能翻译常规技术文档，但识别基本技术规范能力有限；六级了解基本技术写作规范，但是细节需进一步修正；五级仅能遵循最基础的格式，整体翻译的技术规范意识和准确性不足。

在功能性维度上，九级确保功能指示明确和有效传达信息，逻辑清晰且层次分明；八级功能指示较为明确，适应性良好，能应对技术文档；七级功能指示一般，逻辑基本清晰，翻译时的功能表现有待提升；六级功能指示较模糊，内容传达不够明确，难度稍高的内容显得不足；五级功能

指示不清,常常导致理解上的困难,翻译效果较差。

在文化适应性维度上,九级能够调整文化敏感内容,确保目理解无歧义,传播效果极佳;八级跨文化技术交流能力强,复杂文化隐喻处理需要优化;七级能进行基础技术交流,但对文化背景理解有限;六级能处理基础的跨文化技术文本,但需要母语者辅助;五级跨文化适应能力较弱,直译易导致误解,理解效果差。

各维度下五个等级的对比,清晰显示出翻译者在专业知识、准确性、功能指示及文化适应能力上的逐级特征。该体系不仅为翻译者的能力提升提供了清晰的定位,也为用人单位和教育机构提供了参考,以便在资源配置和人才培养上作出科学决策。

3. “科技笔译语用能力等级评价量表”等级赋值

量表指标权重及等级赋值明确了科技笔译能力的评估框架。该量表结合《中国英语能力等级量表》中的具体行为描述,通过对译文进行等级定位与达成度评分。第一步,依据评价维度各等级描述语进行达成度判定,基本达成赋

值 1 分,未达成赋 0 分;第二步,根据科技笔译语用能力各等级之间的差异评估,确定等级与其得分区间的对应关系,核算各等级每个指标得分后计算总分,确定各等级对应的得分区间。

DeepSeek 先后输出两种等级与得分区间的对应关系,分别命名为等距得分区间和非线性得分区间。其中等距划分的等级得分区间便于操作和管理,适合初步搭建量表框架,40 分总分完整分配至 5 个等级,符合一般测评逻辑,但是科技笔译语用能力的提升通常是非线性的。科技笔译的语用能力发展存在“门槛效应”和“瓶颈效应”,需要根据实际能力提升难度调整区间跨度。针对科技笔译语用能力的非线性发展特征,优化方案从等距划分调整为“低分宽区间、高分窄区间”的非线性结构,通过细化等级描述突出术语准确性、逻辑连贯性与语境适应性等核心维度,构建更贴合实际能力差异、兼具科学性与实用性的分级体系,兼顾基础能力敏感区分与高阶水平精准评价(表 1)。

表 1 科技笔译语用能力等级评分调整表

等级	等距区间	非线性得分区间	核心能力描述
5 级	1—8 分	1—10 分	术语混淆频繁,技术逻辑断裂,需要系统学习基础术语与句法
6 级	9—16 分	11—20 分	能处理简单技术句子,术语准确率约 70%,长难句翻译困难
7 级	17—24 分	21—28 分	中等文本翻译基本达标,术语准确率超过 85%,初步适应读者认知
8 级	26—32 分	29—36 分	能流畅翻译高复杂度文本,术语准确率超过 95%,风格适配目标用途
9 级	33—40 分	37—40 分	达到出版级质量,文化负载术语处理等效,修辞准确且自然

该评分表依据五个等级划分,逐步细化翻译者在科技翻译中的核心能力。五级译者在术语和逻辑的运用上极为薄弱,需要从基础开始学习;六级译者在术语准确性和复杂句翻译上存在明显不足;七级译者具备中等文本翻译能力,能适应目标读者的认知水平;八级译者能流畅处理复杂文本,术语使用上表现良好;九级译者达到高水平出版标准,具备准确的文化负载术语处理能力,能进行自然且富有修辞的表达经过星云企业级 AI 平台和专家评价,一致认为从等距到非线性划分的调整思路充分考虑了能力提升的非线性特征。这种调整能够更精准地反映出低分段译者在基础能力上的细微进步,同时避免在高分段出现得分饱和和带来的评价失真。

4. “科技笔译语用能力等级评价量表”实证应用

4.1 数据集概览

本次实践应用研究以 S 高校的 32 名 MTI 研究生为测试

对象,以化工类科技文献中文摘要的英译文为评价文本,将 DeepSeek 作为评价工具,对其科技翻译能力进行评分。评分维度涵盖术语准确性、专业文本适应性、逻辑与结构严谨性、技术规范符合度、信息完整性、功能性、信息适应性和传播效果,这些维度共同反映了参加者在科技笔译中的能力表现。DeepSeek 评分的总分范围为 9—17 分,表明参与者在科技笔译的各个评估维度上表现出一定的局限性,为后续的教育和培训提供了改进方向。

4.2 数据分析结果

评分维度的统计数据反映了被试者在科技笔译中的总体表现。从表 2 可知,每个维度最高分仅为 3 分,最低分均为 1 分,平均分数在 1.4—1.7 之间,显示出被试者在术语准确性、专业文本适应性及信息完整性等多个方面普遍存在较大改进空间,后续培训和教育需要聚焦术语使用、逻辑结构、技术规范遵循等关键要素,以提升整体的翻译

质量和能力水平。

表 2 科技笔译语用能力等级评价实证结果

维度	平均分	最低分	最高分	主要问题集中领域
术语准确性	1.4	1	2.5	军事 / 科技术语误译率高
专业文本适应性	1.5	1	2	句式结构混乱，格式不规范
逻辑与结构严谨性	1.6	1	3	长难句逻辑断裂，技术步骤不连贯
技术规范符合度	1.4	1	3	未遵循国际标准
信息完整性	1.7	1	3	关键数据 / 步骤缺失
功能明确性	1.5	1	2	文本用途匹配度低
信息适应性	1.7	1	3	读者认知水平适应不足
传播性	1.4	1	3	文化敏感性缺失，政治表达错误

汇总后可见，1—10 和 11—20 得分区间分别为 9 人、23 人，占比分别为 28% 和 72%，对比非线性得分区间分别达到五级和六级能力水平。利用 SPSS27 根据科技笔译语用能力各维度进行配对检验，结果表明：根据配对样本 t 检验结果，科技笔译语用能力的四个核心维度均值在 2.95—3.15 分，专业性、准确性、功能性、传播性之间均无显著差异（所有 p 值 > 0.05）。具体数值见表 3。

表 3 科技笔译语用能力各维度差异检验结果

	配对差值					t	自由度	显著性（双尾）
	平均值	标准差	标准误差平均值	差值 95% 置信区间				
				下限	上限			
专业性 – 准确性	-.0303	1.1106	.1933	-.4241	.3635	-.157	32	.876
专业性 – 功能性	-.1970	.9515	.1656	-.5343	.1404	-1.189	32	.243
专业性 – 传播性	-.1515	.9802	.1706	-.4991	.1960	-.888	32	.381
准确性 – 功能性	-.1667	1.0801	.1880	-.5497	.2163	-.886	32	.382
准确性 – 传播性	-.1212	1.1861	.2065	-.5418	.2993	-.587	32	.561
功能性 – 传播性	.0455	.8959	.1559	-.2722	.3631	.291	32	.773

尽管统计上无显著差异，但反映出以下问题：能力扁平化，各维度均处于基础水平，未形成高阶能力分异。在保持术语准确性与专业性的基础上，需要通过专项训练打破能力均衡状态，推动高阶语用能力的差异化发展。

5. 结论与启示

本研究创新性地采用非线性赋分原则（低分宽区间、高分窄区间），精准地反映出科技笔译能力的非线性发展特征，如术语习得的“门槛效应”与高阶语用“瓶颈效应”。实证数据表明，MTI 研究生的科技笔译能力普遍集中于基础水平，72% 的被试者处于 6 级（得分 11—20 分），而高阶语用能力明显不足。配对检验显示，各维度之间无显著差异（p > 0.05），这表明被试者在术语准确性和功能性等

关键领域均处于“低水平均衡”状态，未能形成高阶能力的差异化优势。这一现象指向当前翻译教育的普遍短板——过度关注基础技能训练，忽略了语境适配和跨文化传播等关键能力的培养。

在实践层面，翻译教育应从“单一语言转换”转向“语用能力驱动”，通过专项训练强化功能适配与文化敏感性。此外，行业应依托该评价体系建立量化认证标准，推动译者职业能力的分级认证，以确保技术传播的准确性与可靠性。鉴于本次样本量的局限，未来研究可以进一步细化维度权重，探索能力分化的临界点，为个性化培训提供理论支持。

参考文献：

[1] 戈玲玲. 科技论文摘要英译的语用失误：对非英

语专业本科生的一项调查 [J]. 中国科技翻译, 2002(2): 44-46+36.

[2] 张兴旺, 李洁, 李思凡, 等. DeepSeek 赋能图书馆知识服务的理论模型、模式创新与重要启示 [J/OL]. 农业图书情报学报, 1-13.

[3] 教育部. 中国英语能力等级量表 [M]. 北京: 高等教育出版社. 2018.12+14.

[4] 潘冀春, 刘洁. 科技英语翻译 [M]. 北京: 北京理工大学出版社. 2020.7-8.

作者简介: 马颖, 女, 沈阳理工大学外国语学院副教授,

博士, MTI 研究生导师, 研究方向为英语笔译。

邓鹂鸣, 女, 武汉大学英文系教授, 博士, 硕士导师, 中国英语写作教学与研究专业委员会常务理事。研究方向为应用语言学、语体分析、语篇分析、二语习得、二语写作教学研究、学术写作研究、专门用途英语研究等。

基金项目: 教育部产学合作育人项目成果, 项目名称为“基于《中国英语能力等级量表》的 MTI 翻译硕士生翻译能力测评研究”(编号 230828272907294); 2024—2025 年武汉大学访学研究成果。