

AIGC 赋能下的高校英语写作智能语评实证研究

王姝月

西安欧亚学院，陕西西安，710065；

摘要：AIGC 技术的革新在教育界引发了广泛讨论。本文以 AIGC 在英语写作中的应用为切入点，系统探究 AIGC 写作反馈系统对差异化英语能力学习者的适配状况，并基于智能语评模型的八个维度，对比讯飞星火、豆包及文心一言三大平台的反馈质量及特点。研究发现 AIGC 对初级英语写作水平的反馈重点为准确使用单词和语法，以及完善文章内容；对进阶英语写作水平的反馈重点则为表达优化，基本满足学生使用 AIGC 进行写作自评的需求。同时，不同 AIGC 平台提供的反馈各有千秋，学生要根据自己的批改需求，并结合本文提出的智能语评模型来获取更完善的写作评价。

关键词：AIGC；英语写作；智能语评；写作评价

DOI：10.69979/3029-2735.25.09.057

引言

生成式人工智能（GAI）的突破性发展正重塑教育生态，其核心人工智能生成内容（AIGC）在语言教育领域展现出巨大潜能。AI 写作辅助工具的使用率近年来呈现显著增长趋势，尤其在二语写作中，智能系统通过实时纠错和多维分析，正颠覆传统教学反馈模式。当前针对 AIGC 在教育教学中的研究呈现三重视角：技术赋能层面验证人机协同可提升工作效率；教育伦理领域警示过度依赖导致元认知退化；教学模式方面探索人机协同教学新模式。然而对于高等教育学术写作场景关注不足、跨平台效能对比缺失需要进一步的验证。本研究聚焦高校英语写作，通过实验与文本挖掘，着力解决三大问题：（1）基于国内外写作评价维度构建 AIGC 智能语评模型；（2）解析主流平台（讯飞星火/豆包/文心一言）的反馈特性差异；（3）对比 AIGC 写作辅助工具的分层支持能力。本研究旨在构建“教师-AIGC-学习者”三元协同模式，为智能教育建设提供新路径。

1 AIGC 在教育中的使用现状

AIGC 在教育领域的应用日益广泛，尤其在外语教学和写作教育中展现出巨大潜力。钟秉林等（2023）指出 AIGC 在教育中的优势，包括定制个性化学习计划、自动化处理任务、即时反馈与动态调整学习路径、提升教师工作效率及打破时空限制促进资源共享。在写作教学中，陈代云（2024）认为 AIGC 简化写作过程，影响教学效果与评价标准，且能辅助教师教学设计，激发学生潜能。陈茉和吕明臣（2024）基于活动理论视角，分析 ChatGPT 在英语写作教学中的作用，包括课堂活动设计、目标语输入输出监督及作文评阅，但教师需关注学生使用 AI 情况并培训。然而，AIGC 在教学中的广泛应用也存在

潜在问题与风险。张震宇和洪化清（2023）认为过度依赖 AIGC 可能阻碍批判性思维发展，影响学习公平与诚信，尤其学生直接使用 AIGC 工具生成文本上交作业或论文存在剽窃风险，且 AIGC 平台生成文本内容可能存在事实性错误或捏造信息，误导学生。

综上所述，AIGC 在教育领域展现出巨大潜力，但从智能语评方向的深入分析研究尚少。本研究将从智能语评角度出发，分析 AIGC 应用于学生独立进行英语写作修改的可行性，以及不同 AIGC 平台写作语评反馈的对比。

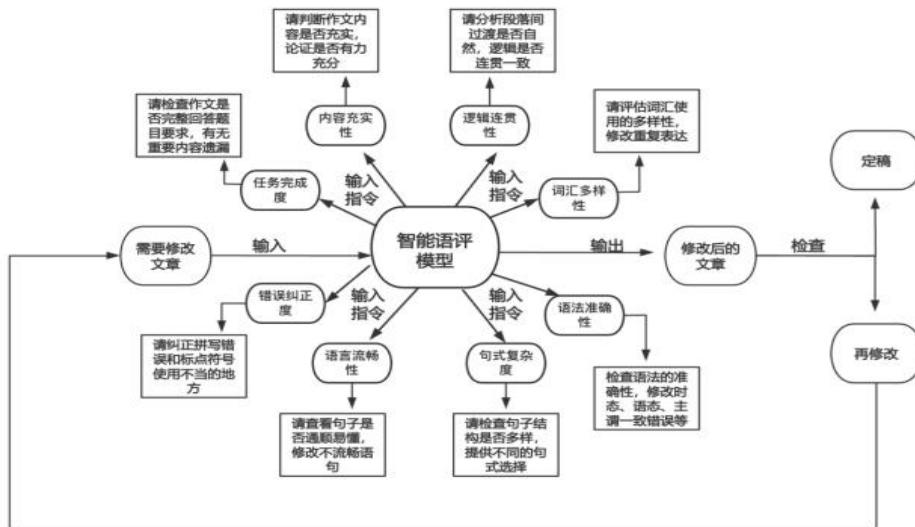
2 AIGC 智能语评模型设计

写作评价研究一直是语言研究的重点。评价过程中，写作任务、评价方法、评价标准、评价策略、等变量均会影响评价结果。写作评价通常采用整体评分法或分项评分法。整体评分法是对文本给出整体评价，分项评价则从不同维度给考生作文进行评价（孙海洋、韩宝成，2013）。目前，最有影响的二语写作分项评分标准是 Jacobs 等在 1981 年开发的二语作文标准（ESL Composition Profile），包括 5 个维度，：内容（content）、组织结构（organization）、词汇（vocabulary）、语言使用（language use）、写作规范（mechanics）。值得注意的是，国际权威英语测评体系如雅思考试（IELTS）通过任务完成度、连贯与衔接、词汇丰富度以及语法多样性及准确性这四个维度进行评价。尽管我国高考、CET 等大规模英语测试仍采用整体评分法，但其等级描述细则已系统涵盖内容质量、论证效度、结构完整性和语言准确性等关键指标。

鉴于此，本研究基于现有写作评估理论框架，构建了一个包含八个核心评价维度的英语写作智能评估模

型,涵盖英语写作评价的各个方面,分别是:写作任务、写作内容及论证、写作逻辑、写作词汇、写作语法、写作句式、写作语言以及拼写错误和标点使用。学习者可参照系统生成的诊断性评语,全面核查文本各要素的评估覆盖度。也可以根据模型输入相应的指令对自己的作

文在AIGC平台进行某一方面的评价和修改,并根据AIGC提供的修改建议进行润色,必要时可以重复检测。该模型支持学习者开展作文自主诊断与多轮次评估。也为智能教育技术的发展提供新的理论视角和实践路径。



3 研究设计

3.1 研究问题

本文主要研究两个问题:第一,使用AIGC平台辅助学生写作是否能满足不同英语能力学生写作修改需求,侧重点是否有所不同。第二,参照智能语评模型的八个核心维度,对比三个主流AIGC平台(文心一言、豆包、讯飞星火),分析不同平台的评价反馈异同。

3.2 研究对象

本研究总共有90名2023级英语专业的大一学生参与。他们是第一次参与大学英语写作课程,英语写作能力有比较明显的区分。

3.3 研究过程

本研究主要使用文本案例收集数据,数据收集过程如下:

作者以教师身份进行学生测试作文的收集。收集了90篇学生在30分钟内完成的命题作文。题目如下: Suppose your university student union is planning to hold a speech contest. You are now to write a proposal for organizing the contest. The proposal may include the topic, aim, procedure and selection of contestants. 根据教师初步评分,选择出两篇具有代表性的文章,分别代表的学生初级英语写作水平和进阶英语写作水平。并用AIGC对学生测试作文进行修改,对比如分析不同AIGC工具对作文的修改情况。

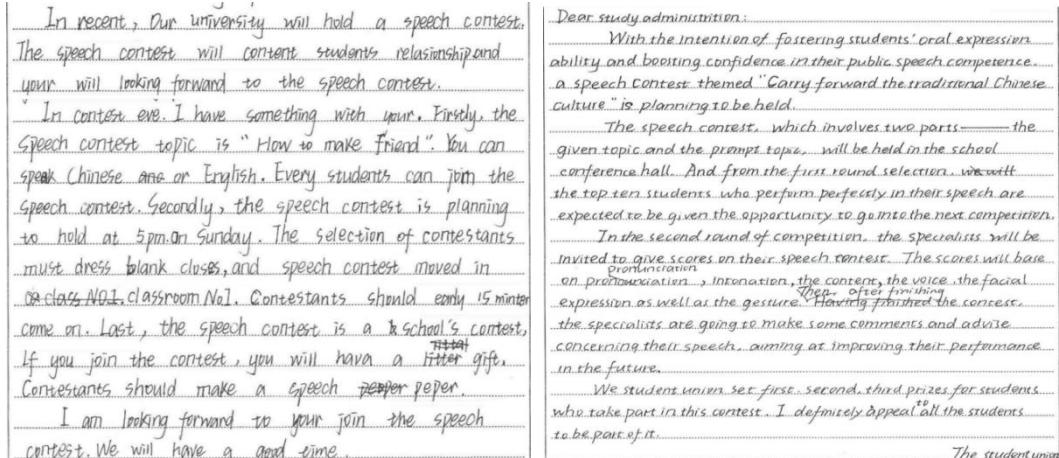


图1: 初级英语写作水平

图2: 进阶英语写作水平

图 1 为抽选的初级英语写作水平的作文。老师在初步评分中发现, 该作文出现较多低级语言错误, 如单词拼写错误、词义混乱、主谓不一致、词性使用错误、语句不通、对题目的回应不完整。这些问题反映出该生在写作中语言运用能力较弱, 对于写作任务的完成度不足的问题。图 2 为抽选的进阶英语写作水平的作文。老师在初步评分中发现该作文整体任务完成度较好, 逻辑较清楚, 语言流畅, 语法整体准确性较好。无单词拼写错误, 句型相对多样, 单词使用比较地道, 仅有个别语态, 句子表达啰嗦的问题, 能从侧面反映出该生英语写作能力较好。

3.4 研究结果与分析

(1) 多维评价体系对比

如表 1、表 2 所示, 本研究参照智能语评模型的八个核心维度, 对讯飞星火、豆包及文心一言三个平台进行系统化评估。数据显示, 三者在词汇语法层面的反馈覆盖率最高, 但在高阶写作能力指导方面存在显著差异。

表 1 写作评价维度对比表 (初级英语写作水平)

评估维度	讯飞星火	豆包	文心一言
任务完成	√	▲	—
内容论证	√ 4	√	—
词汇表达	√ 2	√ 2	√ 3
拼写标点	√	√ 3	√
写作逻辑	▲	√	▲
句式结构	—	—	▲
语言流畅	▲	—	—
语法准确	√ 10	√ 6	▲

注: √=系统建议(数字为具体建议数), ▲=间接建议, —=无建议

表 2 写作评价维度对比表 (进阶英语写作水平)

评估维度	讯飞星火	豆包	文心一言
任务完成	▲	—	—
内容论证	√	√	▲
词汇表达	—	▲	√
拼写标点	—	—	—
写作逻辑	▲	—	▲
句式结构	▲	—	—
语言流畅	√	—	—
语法准确	▲	√ 2	√ 1

注: √=系统建议(数字为具体建议数), ▲=间接建议, —=无建议

(2) 平台特性分析

低阶学习者适配度:

讯飞星火: 基础写作修改比较完善, 在词汇语法维度提出 12 条具体建议, 包含拼写校正、词性修正和主谓一致。其修改范文有效补充了原文缺失的赛事目的,

但在逻辑衔接建议方面仅通过范文对比呈现, 缺乏显性指导。

豆包: 在错误检测方面存在漏检率, 如 relationship 拼写错误未检出, 其内容拓展建议集中于表层信息补充, 对选拔标准等要素未作延伸。虽建议使用连接词提升逻辑性, 但范文未提供具体范例, 导致建议可操作性降低。

文心一言: 呈现明显的“范文导向”特征, 提供字数达原文两倍的修改文本, 包含 6 个高级句式结构。然而其直接建议仅涉及 3 处词汇修正, 且范文采用小标题结构, 可能影响学习者对自然段落过渡的理解。

高阶学习者适配度:

讯飞星火: 在语言表达和信息明确方面提出了一些建议, 提升了文章的内容深度。对比修改后的文章, 语言方面提供了更为地道的表达。例如原文中一些使用不恰当的从句, 或者被动语态, 在修改后表达更为通顺流畅。

豆包: 语法表达上给出了具体的改进建议, 会在文本基础上进行句子的修改, 创新性表达较少。

文心一言: 语法和用词上给出具体的改进建议, 对结构提出的建议更偏重句法方面。修改后的文章, 用词更学术。

4 讨论

对比分析了讯飞星火、豆包、文心一言这三款 AIGC 写作辅助工具的修改建议, 本文将探讨三个平台对不同英语能力的修改反馈优势及平台差异化表现。

4.1 AIGC 写作辅助工具的分层支持能力

AIGC 在提供写作改进意见时, 存在以下优势: (1)能提供词汇使用和语法表达上详细且具体的反馈, 包括包括单词拼写错误、词性使用错误及主谓不一致等问题的具体指出与修正。(2) 内容与结构的补充。或是通过具体建议, 或是通过修改后的范文, 有效补充原文中缺失的内容, 提升文章的完成度。(3) 提供地道表达。在提升文章地道表达方面给出了有效的建议。

AIGC 通过差异化策略满足不同层次英语学习者的需求。对于写作基础薄弱的学生, AIGC 聚焦于词汇与语法的规范性修正。词汇作为语言构建的基本单元, 其丰富程度直接影响到写作表达的细腻度与精准性。例如, 当检测到频繁使用 “good” 等简单词汇来描述美好事物时, AIGC 会主动推荐 “gorgeous”、“charming”、“attractive” 等替换词, 并标注具体语境中的适用差异, 帮助学生理解并掌握其细微差别, 从而在写作中更

加灵活自如地运用,使文章内容丰富多彩,表达更加细腻精准。同时,语法规则掌握的不熟练也极大地制约了英语写作的表现。学生在写作中常见主谓不一致(如“People is...”→“People are...”)、时态混乱(如“Students go...”→“Students went...”)等常见语法错误,不仅破坏了语句的通顺与逻辑性,也大大降低了文章的可读性与专业性。针对这部分同学,AIGC能够能够即时标注错误类型并提供修正范例,帮助学生建立基础语言规范。

对于具备一定写作能力的英语学习者,AIGC的反馈重心转向内容逻辑与表达质量提升。三个平台提供了范文都能很好的为学生提供高质量的语言范例。以讯飞星火为例,其修改建议不仅补充原文缺失的比赛目的(如增加“to cultivate cross-cultural communication skills”的论述),还会重构冗余句式——将臃肿的定语从句“Students who are interested in this activity can...”优化为精炼的分词结构“Interested students can...”。这种突破性改写使文本流畅度显著提高。在句子结构优化方面,AIGC也通过补充句子连接词和从句运用,使修改后的文章,句子之间衔接自然,流畅。同时,三个平台都展示了高级词汇、复杂句式的使用,帮助学生通过模仿逐渐提升语言层次,使文章表达更丰富有力。

4.2 主流 AIGC 平台的差异化表现

对比分析三款 AIGC 写作辅助工具(讯飞星火、豆包、文心一言)在提供写作改进建议方面的特点,可以总结出:三个平台都能对文章进行基本的修改和润色,基本能满足同学们进行自评和自测的需求。但三款工具在反馈机制上呈现显著特征差异。

讯飞星火展现出强大的基础支持能力,对于低水平作文批改的反馈更为详细且具体,尤其是在单词,语法方面可以非常详细指出错误,拼写错误也能及时反馈;内容和结构的补充也比较完整,提升文章的完整性。对于作文较好的反馈则更注重语言的质量,因此给出的范文也跳出原文的语言限制,提供了更为地道的语言表达。

豆包则表现出独特的渐进式优化特征。对低水平作文的基础反馈也比较详细。在修改进阶作文时,平台会优先保留原文核心结构,通过微调语态、优化词汇搭配实现表达精准化。但该工具存在拼写漏检率,且句式创新建议不足,限制了高阶提升效果。

文心一言的强项在于学术性提升,其修改后的范文常扩展至原文两倍篇幅,但是其反馈受限于指令,常出现指令接收不稳定的问题。对于低水平作文,文心一言的基础单词语法反馈并不具体,学生需要自己对比范文来发现问题。对于水平较好的英语作文,文心一言的范文学术性较强,可以专注学习写作能力提升。

5 结语

本文通过实证分析验证了 AIGC 工具在高校英语写作中的分层支持效能。首先,三平台均展现出基础语言要素的修正能力,能够满足不同水平学习者的基础批改需求。其次,三平台在反馈机制方面呈现差异化特征:讯飞星火在内容完善度、逻辑衔接性及语言多样化方面表现突出;豆包擅长渐进式优化,但存在拼写漏检率;文心一言虽能提升学术性,但其模式化修改与指令响应不稳定制约了提升效果。因此,建议学生结合使用这些工具,在写作基础建构阶段,优先使用讯飞星火进行系统性语言规范;在质量提升阶段,结合豆包优化表达准确性与逻辑显性化,最后在学术提升阶段,可以借助文心一言实现学术范式升级。未来研究可着重开发智能诊断模块,通过自动识别学习者水平并匹配适配工具,同时加强跨平台协同机制设计,以突破当前单一工具的功能局限。

参考文献

- [1] 陈代云. 从实用到创新: 对人工智能背景下大学写作教学的思考[J]. 河池学院学报, 2024, 44 (1) : 94-100.
- [2] 陈茉, 吕明臣. ChatGPT 环境下的大学英语写作教学[J]. 当代外语研究, 2024, (01) : 161-168.
- [3] 张震宇, 洪化清. ChatGPT 支持的外语教学: 赋能、问题与策略[J]. 外语界, 2023, (02) : 38-44.
- [4] 孙海洋, 韩宝成. 英语写作分项评分和整体评分比较研究[J]. 解放军外国语学院学报, 2013, 36 (06) : 48-54+125.
- [5] 钟秉林, 尚俊杰, 王建华, 等. ChatGPT 对教育的挑战(笔谈)[J]. 重庆高教研究, 2023, 11 (03) : 3-25..

基金项目: 西安欧亚学院 2024 年度校级科研项目“AI GC 赋能下的高校英语写作智能语评与同伴互评的实证研究”(课题编号: 2024XJSK17)