

国际写作评估研究主题动态演化趋势分析

汤小艺

北京科技大学外国语学院, 北京

收稿日期: 2025年5月21日; 录用日期: 2025年7月7日; 发布日期: 2025年7月21日

摘要

本研究使用文献可视化工具SciMAT, 对2003~2023年Web of Science核心数据库中写作评估领域的文献, 进行主题动态演化的文献计量学分析。通过关键词演进覆盖图、聚类战略图以及动态演化路径图的剖析, 研究发现写作评估领域呈现出繁荣发展的态势, 形成了“句法复杂度”“学生感知”“效度”“写作自动评估”“教学策略”“二语写作”及“性能评估”七个研究方向的十条演化路径。未来写作评估相关研究需融合基于语言规则的理性主义和基于语言大数据的经验主义的跨学科研究方法, 解构“写作质量”多维度、多层次构念, 拓宽写作评估领域的研究内容与视角。

关键词

写作评估, SciMAT, 主题动态演化, 文献计量分析

A Bibliometric Analysis of Dynamic Evolution Trends in International Writing Assessment Studies

Xiaoyi Tang

School of Foreign Studies, University of Science and Technology Beijing, Beijing

Received: May 21st, 2025; accepted: Jul. 7th, 2025; published: Jul. 21st, 2025

Abstract

The present study harnesses the literature visualization tool SciMAT to conduct a bibliometric analysis of the dynamic thematic evolution within the writing assessment domain, drawing from documents indexed in the Web of Science Core Collection from 2003 to 2023. Through an in-depth examination of keyword evolution coverage maps, clustering strategy maps, and thematic dynamic evolution pathway diagrams across three distinct periods, our investigation unveiled that, over the past

two decades, the domain of writing assessment has experienced a flourishing development, culminating in the establishment of seven research directions: Syntactic Complexity, Students Perceptions, Validity, Automated Writing Evaluation, Instructional Strategies, L2 Writing, and Performance Assessment with ten evolutionary pathways. Future research on writing assessment could integrate interdisciplinary approaches that combine rationalism based on language rules and empiricism based on big language data, thereby deconstructing the multidimensional and multilevel constructs of writing quality and expanding the perspective of research within the field of writing assessment.

Keywords

Writing Assessment, SciMAT, Thematic Dynamic Evolution, Bibliometric Analysis

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

写作在外语教学与测评中占据重要地位，其过程涉及批判性思维、逻辑推理及问题解决等高阶认知活动，为写作评估领域带来诸多挑战。随着语料库语言学、自然语言处理、人工智能等学科的蓬勃发展，国际写作评估领域理论体系和研究方法日趋成熟。国内学者对二语写作评估和学术写作评估的研究已有深入的梳理和分析，涵盖了领域内的基本趋势、研究内容以及方法，并对写作评估与反馈、写作教学策略等热点问题进行了探讨[1]-[3]。然而，对于该领域研究成果的可视化分析和综述较少，且近年来写作评估领域的研究热点和方向亦呈现出新的演化趋势。因此，有必要对其进行全面系统的梳理，有效推动国际写作评估领域的纵深发展。

鉴于此，本研究采用文献计量可视化工具 Science Mapping Analysis Software Tool (SciMAT)，选取 Web of Science (WOS) 核心数据库 2003~2023 年的写作评估领域相关文献作为研究对象，从发文量基本趋势、关键词演进覆盖、主题聚类战略和主题动态演化路径等角度，深入剖析国际写作评估领域的多维动态发展图景，旨在为研究者把握该领域的发展脉络和演化路径提供有益参考，并为国内写作评估领域的相关研究提供新的研究思路和议题方向。

2. 文献来源与研究方法

2.1. 文献来源

本研究选取 WOS 核心数据库中的子库：Science Citation Index Expanded (SCI-E) 和 Social Science Citation Index (SSCI) 作为文献来源，以 2003 至 2023 年间写作评估领域的相关文献作为研究对象，检索关键词为“writing assessment” OR “writing evaluation” OR “essay assessment” OR “essay evaluation” OR “essay scoring”，文献类型为文章(article)和综述(review)，语种为英语。检索工作完成于 2024 年 3 月 29 日，通过严格的手动筛查标题和摘要，剔除无关文献，最终确定 770 篇文献用于后续分析。

2.2. 研究工具

本研究基于 SciMAT 工具，对近二十年国际写作评估领域发展进行脉络梳理，以期对写作评估理论和应用研究提供参考。相较于 HisCite、CiteSpace 和 VOSviewer 等文献计量工具，SciMAT 在预处理方面

如自动去重、智能合并关键词等方面更为先进，并能执行时序分析。这一特性使其能够准确识别研究主题的时序变化，并据此生成主题演进覆盖图、聚类战略图和动态演化路径图，展现研究主题的发展轨迹[4]。

2.3. 数据预处理

预处理包括文献导入、数据清洗与时间分段三个关键步骤：(1) 文献导入：选取 2003 年至 2023 年间与写作评估相关的 770 篇文献，数据格式为 txt 和 RIS 格式，记录内容为全记录与引用的参考文献；(2) 数据清洗：利用 SciMAT 自动清洗功能，统一单复数形式，如将“writing assessment”与“writing assessments”统一为“writing assessment”；对无法整合的近义关键词进行手动合并，如“automated writing assessment”“automated writing evaluation”与“automated essay evaluation”统一为“automated writing evaluation”；同时，手动删除意义较模糊且高频的关键词，如“adolescents”“adults”与“school”等；(3) 时间分段：考虑到 2003~2013 年与后十年发文量的差异，将此期间定为首个研究时段，其后时间以固定时间窗口法分为两段。最终，通过 Periods Manager 工具，将所收集的文献分为三个研究时段：2003~2013 年、2014~2018 年和 2019~2023 年。

2.4. 参数设置

本研究参照李文婷、祝朝伟的参数设定[5]：分析单元为词，包含作者关键词和来源关键词(Author words & Source words)；为优化绘图效果，三个时段的数据缩减均采用频率缩减法(Frequency reduction)，设定最小频率阈值为 1；矩阵类型为共现矩阵(Co-occurrence matrix)；网络缩减均为边缘值缩减(Edge value reduction)，默认最小值为 1；标准化网络相似度指标选择关联强度(Association strength)；聚类算法选择简单中心算法(Simple center algorithms)，设定最大值为 10，最小值为 2；绘图选择核心绘图(Core mapper)；聚类质量的评估指标选择总引用量(Sum citations)；演化图(Evolution map)和覆盖图(Overlapping map)的标准化系数均选择关联强度。

3. 国际写作评估研究现状分析

3.1. 国际写作评估研究发文量基本趋势

本研究对近 20 年写作评估领域的年度发文量进行统计(见图 1)，旨在宏观上呈现该领域的发展趋势。由图 2 可知，国际写作评估领域的发文量呈现上升趋势，2003~2013 年间，发文量增长较为缓慢。2014~2018 年间，相关研究开始步入加速发展的稳定期。2019 年以后，相关研究的增长呈现显著上升，从 2019 年的 47 篇增至 2023 年的 125 篇。总体而言，国外写作评估领域正逐渐受到学术界的广泛关注，研究成果持续增多。究其原因主要有以下几点：第一，多元化理论丰富写作评估的研究视角。研究者们基于外语教学与测试、语类研究、评估反馈等领域的理论基础，整合跨学科的理论洞见，进一步深化了多元视角下的写作评估研究。第二，多样化语料库工具助力解构写作质量的复杂构念。随着自然语言处理、语料库语言学和计算语言学的发展，研究者们能够提取多维度的细粒度语言特征，从而促进写作评估与反馈的相关研究[6]-[8]。第三，跨学科方法推动写作评估的纵深发展。随着语言智能的快速发展为写作评估与人工智能的融合开辟新的路径，其中写作自动评估与反馈、人机协调的写作评估等研究呈现快速发展的趋势[9]-[11]。

为梳理国际写作评估的发展历程和演进脉络，本研究以出版年份为横轴，参考文献被引次数为纵轴，绘制了写作评估参考文献 - 出版年图谱(图 2)。通过计算各年份被引频次与相邻两年平均值的偏差，绘制中值曲线。由图 2 可知，2003~2023 年写作评估领域的参考文献 - 出版年图谱呈现出先上升后下降的波动，中值曲线的峰值标识出了可能对该学科或领域发展有着重要作用的历史文献。

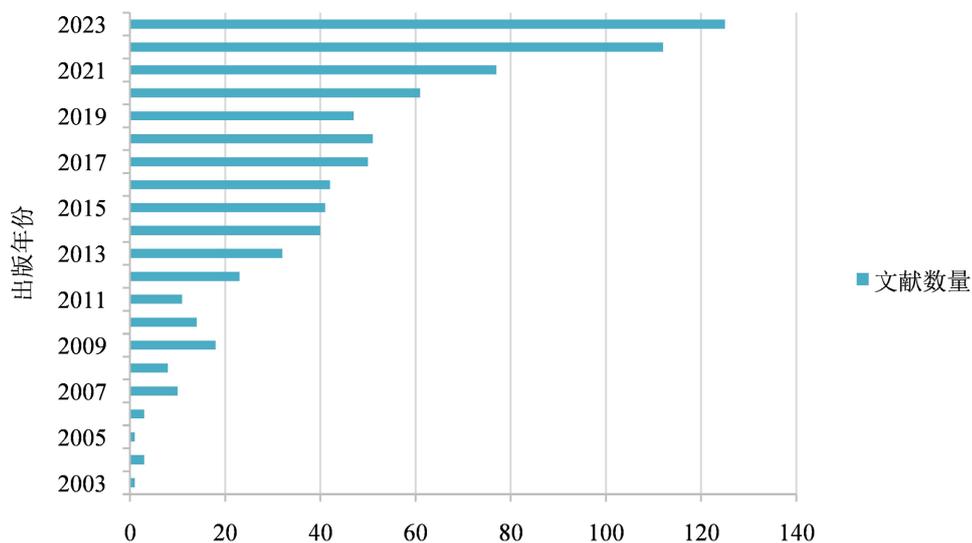
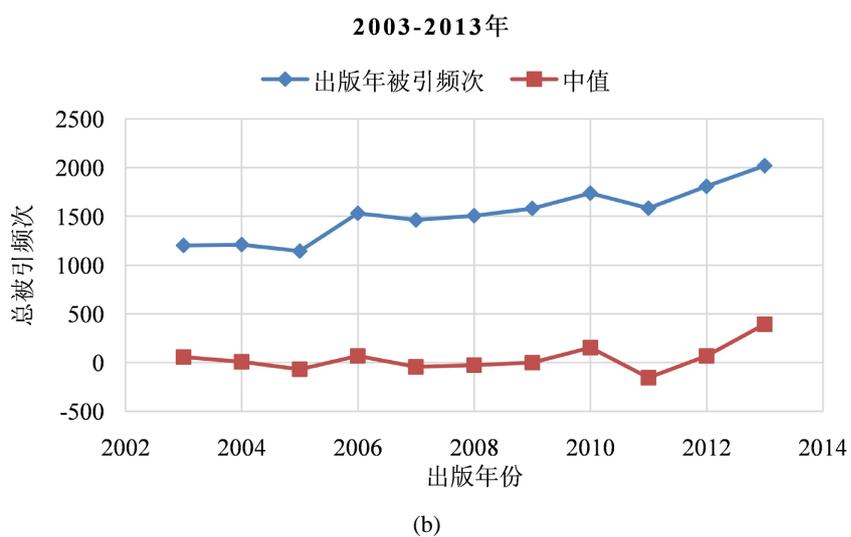
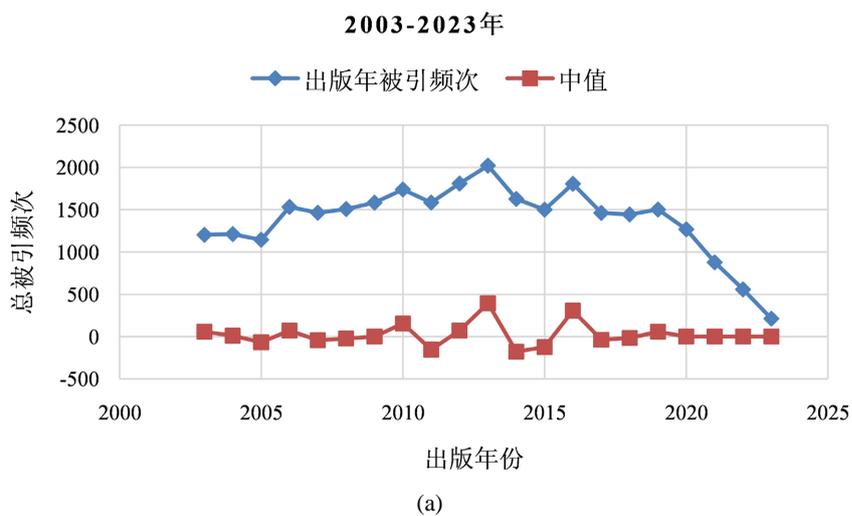


Figure 1. Basic trend of publications of Writing Assessment from 2003 to 2023
 图 1. 2003~2023 年国际写作评估领域发文量基本趋势



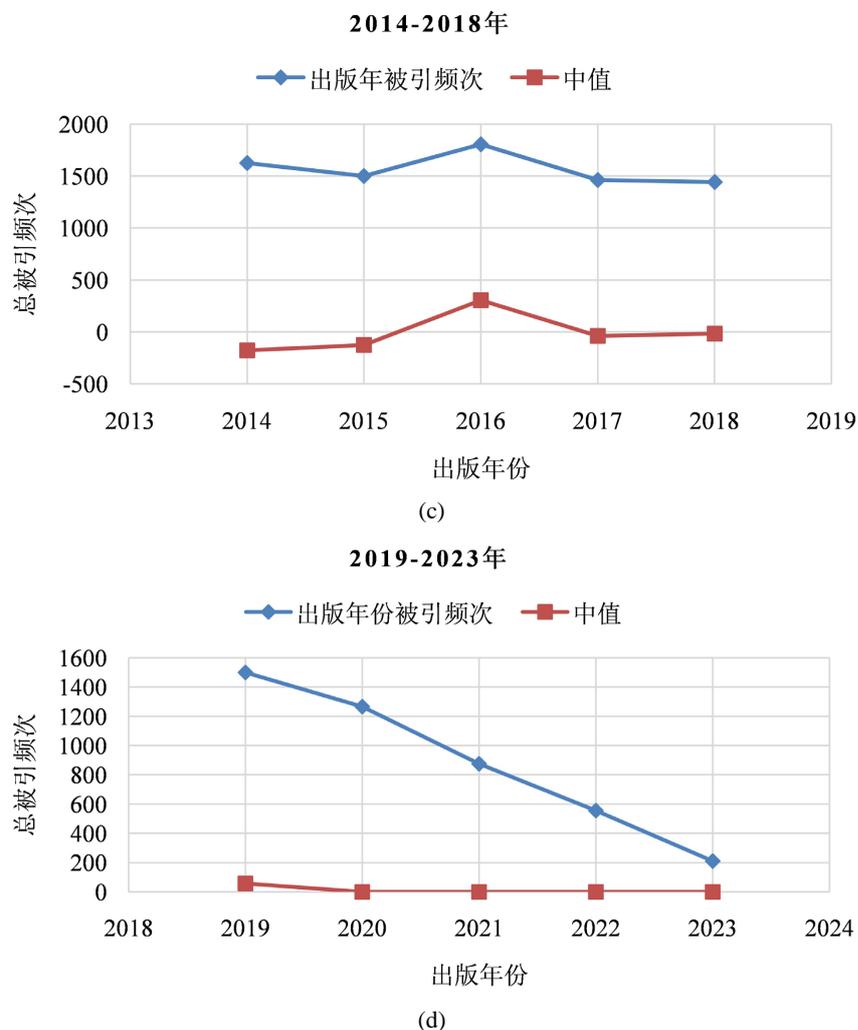


Figure 2. Citation-Year spectrum of Writing Assessment from 2003 to 2023
图 2. 2003~2023 年国际写作评估领域参考文献 - 出版年图谱

根据图 2(b)可知, 2003~2013 年间的中值曲线峰值出现在 2013 年, 对应的关键文献为 Deane 在 *Assessing Writing* 期刊发表的文章: *On the relation between automated essay scoring and modern views of the writing construct* [12]。该文聚焦写作自动评分与写作构念的复杂关系, 通过概述写作评估领域的理论与实践, 凸显从以产出为导向(product-oriented)到以过程为导向(process-oriented)的写作观念的转变, 提出基于社会认知理论框架的写作自动评估方法。

根据图 2(c)可知, 2014~2018 年间的中值曲线峰值出现在 2016 年, 对应的关键文献为 Wilson 和 Czik 发表在 *Computers & Education* 期刊的文章: *Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality* [13]。该文深入探讨了写作自动评估在教育实践中的应用, 尤其是在提供即时反馈、激励学生写作动机和提高写作质量方面所做出的重要贡献。通过与传统教学法的对比, 突出根据教学需求调整自动评估策略的必要性, 为提高写作教学效能和写作评估准确性开辟了跨学科途径。

由图 2(d)可知, 2019~2023 年间的中曲线峰值出现在 2019 年, 是发表在 *English Language Teaching (ELT)* 期刊上的文章: *Automated writing evaluation* [14]。该文聚焦写作自动评估对第二语言学习环境的影

响, 揭示其在促进写作能力、支持个性化学习和及时反馈等方面的重要贡献。由发文量基本趋势和参考文献-出版年图谱分析可知, 近年来, 语言智能技术应运而生并不断促进写作评估的智能化, 为写作评估提供新的跨学科路径。

3.2. 写作评估领域研究主题关键词演进覆盖分析

SciMAT 生成的关键词演进覆盖图(见图 3)呈现了领域内研究焦点的变迁。图中的圆圈代表时段, 圆圈内的数字为对应时段的主题数量。上斜箭头代表该时段消亡的关键词数目, 下斜箭头代表新生的关键词数目。圆圈间箭头上的数目代表前一研究时段保留至后一研究时段的关键词数目。

2014~2018 年与 2003~2013 年相比, 保留主题词 43 个, 新生主题词 28 个, 消亡主题词 5 个, 可见 2014~2018 年的国际写作评估研究发展快速。2019~2023 年与 2014~2018 年相比, 保留主题词 66 个, 新生主题词 10 个, 消亡主题词 5 个。笔者发现相比于前一时段, 后一时段的保留主题词数量有所增加。这一现象表明, 写作评估领域的研究主题呈现出较高的成熟度和研究连续性。从新生与消亡主题词的更迭角度来看, 新生主题词的数量始终高于消亡主题词, 这一趋势说明国际写作评估领域的研究主题丰富度在不断增加, 研究活跃度较高, 具有广阔的发展前景。

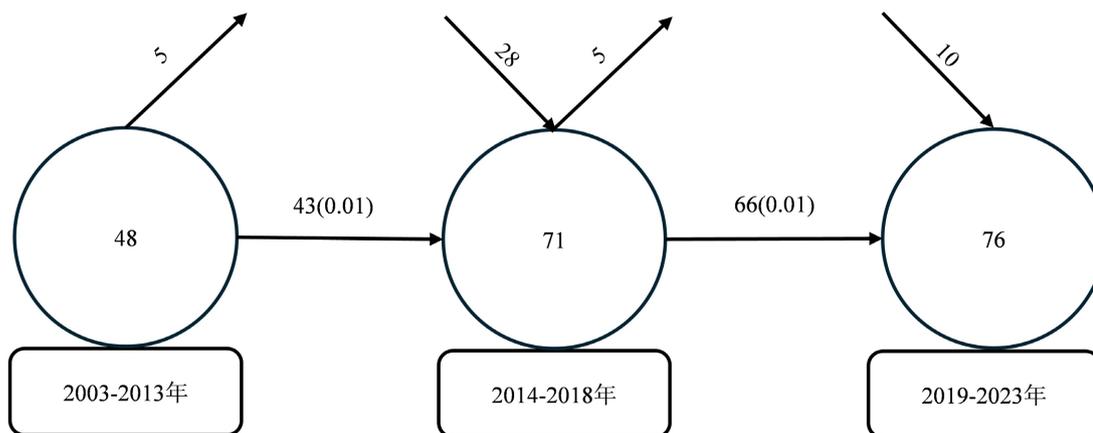


Figure 3. Keyword evolution coverage maps of research themes in Writing Assessment from 2003 to 2023

图 3. 2003~2023 年国际写作评估领域研究主题关键词演进覆盖图

3.3. 写作评估领域研究主题聚类战略分析

SciMAT 绘制的聚类战略图(见图 4)从文献计量、中心度(centrality)、密度(density)和象限位置等维度将研究主题的发展趋势进行可视化。图中的圆圈代表通过关键词聚类形成的研究主题。圆圈内的数字代表该主题在特定时段的文献数量, 数字越大, 说明该主题的关注度和活跃度越高。横轴代表中心度, 反映了该主题与同一时段内其他主题的关联强度; 中心度高的主题在研究领域中通常占据核心地位, 具有较大影响力。纵轴代表密度, 反映了主题内部关键词的紧密度和凝聚力; 密度高的主题表明其内部结构紧凑, 关键词间关联性强。象限位置描绘了研究主题的核心性、成熟度、边缘性及成长趋势。第一象限(右上): 高中心度、高密度的特点, 代表领域中核心且发展成熟的研究主题; 第二象限(左上): 低中心度、高密度, 代表成熟但相对孤立的研究主题; 第三象限(左下): 低中心度、低密度主题, 新兴或衰退的边缘研究; 第四象限(右下): 高中心度、低密度, 代表研究主题的发展潜力大。

以聚类质量评估指标和聚类网络图, 对三个时段聚类主题的演化状态进行分析, 结果如下:

(1) 2003~2013 年有六个聚类主题: 语言特征(Linguistic Features)、写作任务(Writing Task)、自动反馈

(Automated Feedback)、性别差异(Gender Differences)、效度(Validity)、共同核心标准(Common Core Standards)。该时段的聚类主题集中在第一、第二和第四象限,表明该阶段写作领域研究主题的发展势头相对稳定。第一象限的“写作任务”主题包含综合写作任务、学术写作、语言复杂度等关键词,属于中心度高、成熟度高、影响力大的核心主题。“性别差异”和“共同核心标准”主题位于第二象限,处于中心度较低,成熟度较高的位置,可知其属于研究热点较低、发展比较稳定的外围主题。前者包含教学策略、写作策略、语言学习障碍、写作动机、自我效能等关键词,后者聚焦共同核心标准视域下的写作语言和教师信念的相关研究。“自动反馈”位于第四象限(中心度较高、成熟度较低),说明其在写作评估领域的发展潜力大。该主题包含评分标准、二语写作、学生感知、纠正性反馈等关键词。位于中心轴的“语言特征”主题包括写作素养、智能教学系统、自动写作评估工具、写作过程等关键词,其研究热度和影响力高,将持续成长为热点主题。位于密度轴上的“效度”主题包含性能评估和大规模写作评估等关键词,是具有一定演化能力的研究方向。

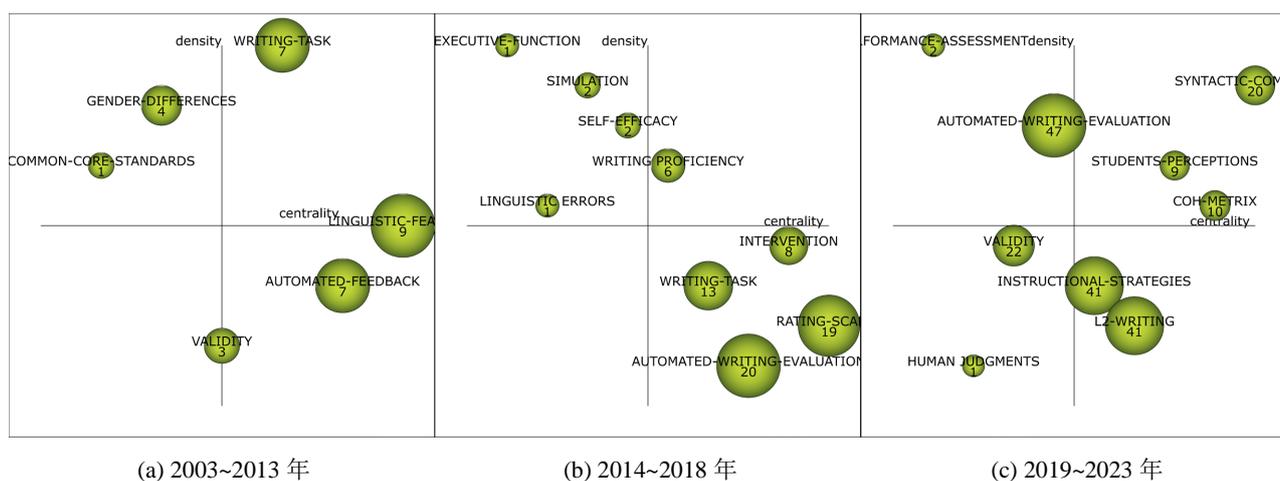


Figure 4. Clustering strategy maps of research themes in Writing Assessment across three periods

图 4. 三个时段国际写作评估研究主题的聚类战略图

(2) 2014~2018 年有九个聚类主题: 写作自动评估(Automated Writing Evaluation)、评分标准(Rating Scales)、写作任务(Writing Task)、干预(Intervention)、写作熟练度(Writing Proficiency)、自我效能(Self Efficacy)、模拟(Simulation)、语言错误(Linguistic Errors)、和执行功能(Executive Function)。该时段的聚类主题主要集中在第二和第四象限,文献数量较上一时段有明显增加。“写作熟练度”主题位于第一象限,包含写作发展、词汇复杂度、课堂写作评估、N-gram、词汇密度和外语学习者等关键词,具有高成熟度和中心度,显现出良好的发展态势。第二象限包含“自我效能”“模拟”“执行功能”和“语言错误”等主题,处于成熟度较高、中心度较低的位置,且文献量较少,表明说明这些主题在写作评估领域受到少数研究者的关注。第四象限包含“写作自动评估”“评分标准”“写作任务”和“干预”这四个研究主题。“写作自动评估”包含自动反馈、写作素养、准确度、纠正性反馈、学术写作和人工智能等关键词,在这一时段有很高的研究热度和影响力,是写作评估领域中有发展潜力的研究主题。此外,“写作任务”这一主题由上一阶段的第一象限发展至第四象限,且文献数量增加,表明该主题有待发展的潜力。“干预”主题包含写作策略、语言学习障碍、形成性评价、年级发展、记叙文、写作流利度等关键词,处于中心度高、成熟度低的位置,是具有发展潜力的研究主题。

(3) 2019~2023 年有九个聚类主题: 写作自动评估(Automated Writing Evaluation)、教学策略(Instructional Strategies)、二语写作(L2 Writing)、效度(Validity)、句法复杂度(Syntactic Complexity)、Coh-Metrix

(Coh-Metrix)、学生感知(Students Perceptions)、性能评估(Performance Assessment)、人工评分(Human Judgments)。第一象限包含“句法复杂度”、“Coh-Metrix”和“学生感知”这三个研究主题，处于高中心度和成熟度的位置，属于成长状态的热点主题。“句法复杂度”包含写作熟练度、写作流利度、写作发展、搭配等关键词。“Coh-Metrix”包含语言特征、自我效能、评分员效应、语类/体裁、写作能力等关键词。这两个主题的中心度和中心度都较高且相关文献数量较多，属于关注度较高的研究主题，具有较强的演化能力。“写作自动评估”和“模型性能评估”位于第二象限，其中“写作自动评估”是从上个时期的第四象限上升至第二象限且文献数量增加，表明该主题内部联系紧密，成熟度高，且有持续成长的趋势。第四象限包含“教学策略”和“二语写作”两大主题，处于中心度高、成熟度低的位置，发展潜力巨大。“教学策略”主题包含写作策略、写作素养、写作动机、评分标准、形成性评价、语言学习障碍和干预等关键词，是影响力较大的研究主题。“二语写作”主题包含纠正性反馈、语言复杂度、课堂写作评估、有声思维等关键词。这两个主题的研究热度和影响力较高，文献数量多，是有持续发展潜力的研究主题。“人工评分”和“效度”主题均位于第三象限，处于中心度低、密度低的位置，属于边缘化主题。

3.4. 写作评估领域研究主题动态演化路径分析

主题动态演化路径图(见图 5)描绘了写作评估领域的发展脉络和兴衰变化。图中的圆圈代表聚类得到的研究主题，其大小与对应主题的文数量成正比；实/虚线分别代表主题的继承和分化；连线的颜色、粗细与主题相似度成正比，连线越粗颜色越深，表明两个研究主题的关联强度高，演化能力强。

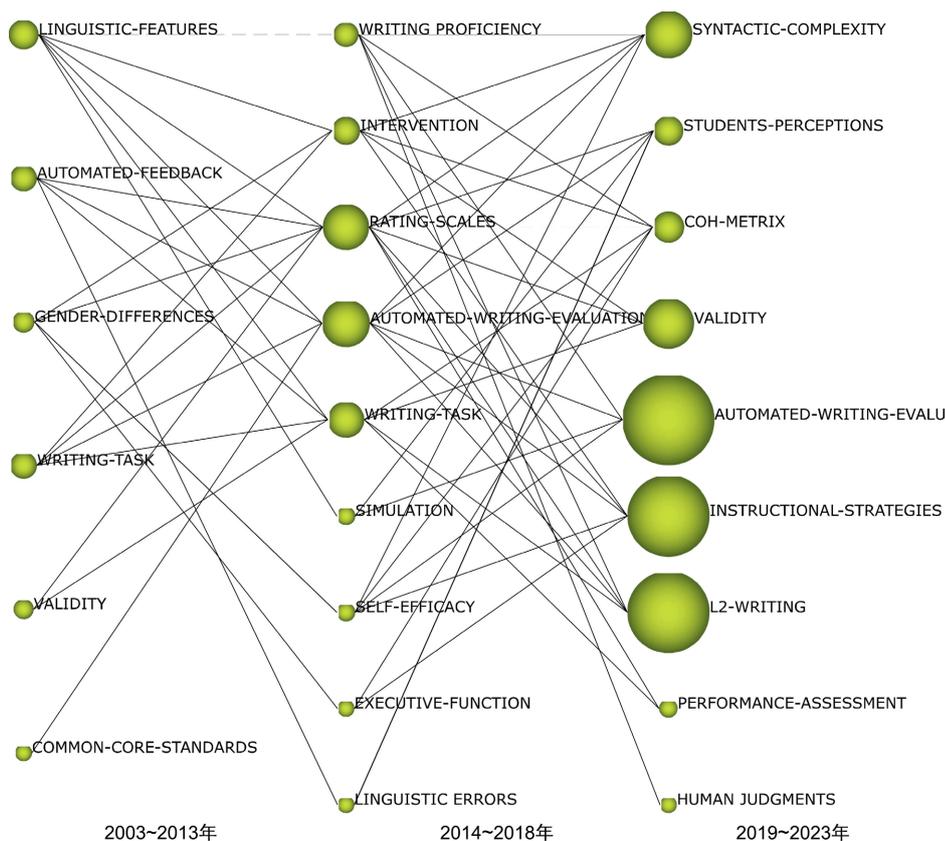


Figure 5. Dynamic evolution pathway diagrams of research themes in Writing Assessment across three periods
图 5. 三个时段国际写作评估领域研究主题动态演化路径图

总体而言, 写作评估领域的研究主题及文献数量持续增加, 主题词呈现多样化趋势, 研究内容日益丰富。该领域主题演化关系复杂, 新的研究主题不断涌现, 且主题关联紧密, 未见与前后时段无关联的孤立点出现。纵向来看, 最后一个时段的主题节点大小相比前两个研究时段有明显差异, 说明近五年来, 写作自动评估、教学策略和二语写作的研究规模大幅度增加, 尤其是写作自动评估的研究规模急剧增长, 成为该时段的核心研究主题。横向来看, 国际写作评估领域研究主题具有动态性和多元性的特点, 呈现出七大研究方向的十条演化路径:

(1) 句法复杂度方向: 路径① Automated Feedback (自动反馈)→Rating Scales (评分标准)→Syntactic Complexity (句法复杂度)。句法复杂度作为衡量句法层面的关键指标, 对于学习者的语言能力及写作水平评估具有关键作用[15][16]。随着语料库语言学的发展, 研究者们对于句法复杂度的探讨已由对浅层句法特征(句子长度、T 单位等)的分析[17][18], 逐步转向至深层细粒度句法特征(动词论元构式、从属成分、主动词词目等)的考察[19][20]。尽管目前对句法复杂度与写作质量的研究已取得丰富成果, 但研究者们大多聚焦于浅层句法特征对议论文单一主题写作质量的预测分析, 不同文体、不同主题的细粒度句法复杂度特征与写作质量的关系亟待深化。

(2) 学生感知方向: 路径② Automated Feedback (自动反馈)→Linguistic Errors (语言错误)→Students Perceptions (学生感知)。学生对写作自动反馈的感知对提升写作能力具有重要作用, 而这一感知受到个体差异、教学环境、互动模式、写作信念、修改行为及师生技术素养等多重因素的共同作用[21][22]。尽管写作自动评估系统在辅助写作教学与评估方面潜力巨大, 但现有研究在师生对于修正性反馈感知对比、提升教师反馈素养、以及多语言教育背景下的人机协同写作评估与反馈等方面尚显薄弱。此外, 目前研究多聚焦教师、同伴或自动反馈等单一反馈形式对写作质量的影响, 多维写作评估反馈体系有待健全。

(3) 效度方向: 路径③ Writing Task (写作任务)→Writing Task (写作任务)→Validity (效度); 路径④ Validity (效度)→Writing Task (写作任务)→Validity (效度)。评估的信度与效度是写作评估领域关注的重点。随着评估范式的演变, 写作测试的设计越来越注重对考生实际运用语言能力的考察, 其中包括考生如何将不同的语言资源整合应用在特定的写作任务中。研究者们通过多层面 Rasch 模型对不同写作任务类型的评分过程进行剖析, 从受试、评分员、受试与评分员的偏性交相互作用分析全面验证评分的效度, 旨在提高写作评分的准确性和公平性[23][24]。

(4) 写作自动评估方向: 路径⑤ Automated Feedback (自动反馈)→Automated Writing Evaluation (写作自动评估)→Automated Writing Evaluation (写作自动评估); 路径⑥ Common Core Standards (共同核心标准)→Automated Writing Evaluation (写作自动评估)→Automated Writing Evaluation (写作自动评估)。随着深度学习和大语言模型(Large Language Models, LLMs)的蓬勃发展, 语言智能应运而生并不断促进写作评估的智能化, 写作自动评估在计算机辅助写作教、学与评等领域呈现广阔的应用前景[25]-[27]。随着大语言模型与可解释人工智能的发展, 未来应进一步关注特征可视化, 通过描绘特征贡献率程度、特征间的交互作用以及特征依赖关系对属性的联合影响, 提升模型的可解释能力和泛化能力。

(5) 教学策略方向: 路径⑦ Automated Feedback (自动反馈)→Rating Scales (评分标准)→Instructional Strategies (教学策略); 路径⑧ Gender Differences (性别差异)→Executive Function (执行功能)→Instructional Strategies (教学策略)。现有研究侧重个性化写作教学、在线写作教学对二语/外语写作能力提升的影响, 对于教师自我效能和信念与写作教学的探讨尚需考察[28][29]。此外, 多维度(教师-学生同伴-人机协同)、全过程(课前-课中-课后)、多文体的“线上+线下”写作教学策略与反馈研究均有待系统地探讨。

(6) 二语写作方向: 路径⑨ Automated Feedback (自动反馈)→Rating Scales (评分标准)→L2 Writing

(二语写作)。二语写作评估研究一直以来备受学界关注。早期研究者们依赖手工提取的语言特征,结合线性回归、贝叶斯等传统机器学习方法进行写作评估。预测写作质量的语言特征大致归为三大类:词汇[30][31]、句法[7][16]以及衔接[32][33]。尽管基于语言规则的特征提取方法对写作教学与评估具有重要的理论价值,但研究者们大多聚焦单一文体总分维度的语言特征,不同文体、不同评分维度的细粒度语言特征与写作质量的关系亟待深化。

(7) 性能评估方向: 路径⑩ Automated Feedback (自动反馈)→rating Scales (评分标准)→Performance Assessment (性能评估)。写作自动评估模型的性能评估提升是写作评估研究热点之一。近年来,基于深度学习的写作自动评估依赖大规模文本语料库,以端到端的方式构建具备高性能和强泛化能力的写作自动评估模型。虽然模型的准确度和泛化能力得到提升,但无法描绘特征贡献率程度及特征间的交互作用,而这对写作教学与评估至关重要。其次,缺乏从可解释人工智能角度揭示“黑箱”中的细粒度语言特征及原理机制。

4. 思考和展望

结合发文量趋势、主题演进覆盖图、聚类战略图以及动态演化路径图的分析,可以看出“语言特征”“自动反馈”“评分标准”“作文自动评估”“教学策略”和“二语写作”等主题在国际写作评估领域占据着核心地位。此外,这些研究主题并非孤立存在,而是相互关联、交织在一起,推动着写作评估领域的持续发展与进步。

第一,解构“写作质量”多维度、多层次构念,拓宽写作评估领域的研究视角。现有写作评估研究多聚焦单一维度,忽视英语学习者在不同维度发展的差异性。依托多维度解构“写作质量”构念,实现多维度、全信息的动态剖析并提供个性化反馈,赋能个性化写作的教、学、评,是未来发展的潜力方向。

第二,揭示细粒度语言特征在不同文体中的互动效应,为跨文体写作评估与教学的理论研究提供支撑。已有研究多局限于单一文体,忽视细粒度语言特征在不同文体中的互动效应以及各语言子系统内部的交互关系。未来研究应着眼于可解释人工智能的潜力,描绘各细粒度语言特征依赖关系对属性的联合影响,解码跨文体写作中的文体依赖特征与文体独立特征,为跨文体写作的相关理论提供依据。

第三,融合基于语言规则的理性主义和基于语言大数据的经验主义方法,为写作评估提供“语言智能+”融合的跨学科研究路径。针对基于语言大数据的深度学习方法存在“黑箱”难题,未来研究应将语言学、自然语言处理、人工智能等跨学科研究视角有机结合,融合多任务迁移学习、可解释人工智能以及大语言模型在多维度写作评估中的应用,提升多维度写作自动评估模型的可解释能力和泛化能力,反哺写作个性化教、学、评,为英语写作评估提供智能化、全方位、全信息的研究路径。

5. 结语

本研究通过可视化工具 SciMAT 对近 20 年来国际写作评估领域的文献进行分析,系统梳理了 2003~2013 年,2014~2018 年和 2019~2023 年三个时段主题关键词演进覆盖、聚类战略和动态演化路径分析。三个研究时段的主题保留词不断增加,且新生词均多于消亡词,反映出国际写作评估领域的持续繁荣。研究主题在“句法复杂度”“学生感知”“效度”“写作自动评估”“教学策略”“二语写作”和“性能评估”七个研究方向呈现出十条演化路径。未来写作评估相关研究可融合语料库语言学、计算语言学、人工智能和自然语言处理等跨学科的研究视角,将基于语言规则的理性主义和基于语言大数据的经验主义有机结合起来[34],利用大语言模型和多任务迁移学习方法,实现不同文体、不同主题、不同评分维度的多源和异构特征关联,达到知识迁移共享和智能互通,为多维度写作评估提供“语言智能+”融合的跨学科研究路径。

基金项目

北京科技大学本科教育教学改革面上项目“AI 赋能英语写作教、学、测、评、馈一体化”(项目编号: JG2024M49)。

参考文献

- [1] 赵秀凤. 英语写作评估与语体量化体系[J]. 外语教学, 2003(5): 76-79.
- [2] 高海英. 国内外二语写作评估研究现状与思考[J]. 外语界, 2010(2): 77-82.
- [3] 齐曦. 国外学术英语写作评估研究述评[J]. 外国语文, 2018(5): 94-101.
- [4] Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E. and Herrera, F. (2012) SciMAT: A New Science Mapping Analysis Software Tool. *Journal of the American Society for Information Science and Technology*, **63**, 1609-1630. <https://doi.org/10.1002/asi.22688>
- [5] 李文婷, 祝朝伟. 基于 SciMAT 的国际译学动态演化(2010-2019)分析研究[J]. 外国语文, 2022(1): 123-132.
- [6] McNamara, D.S., Graesser, A.C., McCarthy, P.M. and Cai, Z. (2014) *Automated Evaluation of Text and Discourse with Cohmetrix*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511894664>
- [7] Kyle, K., Crossley, S. and Berger, C. (2017) The Tool for the Automatic Analysis of Lexical Sophistication (TAALES): Version 2.0. *Behavior Research Methods*, **50**, 1030-1046. <https://doi.org/10.3758/s13428-017-0924-4>
- [8] Crossley, S.A., Kyle, K. and Dascalu, M. (2018) The Tool for the Automatic Analysis of Cohesion 2.0: Integrating Semantic Similarity and Text Overlap. *Behavior Research Methods*, **51**, 14-27. <https://doi.org/10.3758/s13428-018-1142-4>
- [9] Shermis, M.D. and Burstein, J. (2013) *Handbook of Automated Essay Evaluation: Current Applications and New Directions*. Routledge.
- [10] 刘应亮, 刘胜蓝, 杨进才. 社会文化活动理论视域下人机协同教学及应用探索——以 iWrite 协同英语写作教学为例[J]. 中国电化教育, 2022(11): 108-116.
- [11] 秦文娟, 周季鸣. 基于写作自动评估的教师反馈素养研究[J]. 外语界, 2023(3): 47-54.
- [12] Deane, P. (2013) On the Relation between Automated Essay Scoring and Modern Views of the Writing Construct. *Assessing Writing*, **18**, 7-24. <https://doi.org/10.1016/j.asw.2012.10.002>
- [13] Wilson, J. and Czik, A. (2016) Automated Essay Evaluation Software in English Language Arts Classrooms: Effects on Teacher Feedback, Student Motivation, and Writing Quality. *Computers & Education*, **100**, 94-109. <https://doi.org/10.1016/j.compedu.2016.05.004>
- [14] Hockly, N. (2018) Automated Writing Evaluation. *ELT Journal*, **73**, 82-88. <https://doi.org/10.1093/elt/ccy044>
- [15] 李慧娟, 郑咏滢, 秦文娟. 精细化句法复杂度对写作质量预测效果的研究[J]. 解放军外国语学院学报, 2022, 45(4): 61-69.
- [16] 张晓鹏, 李雯雯. 句法复杂度对中国大学生英语说明文写作质量的预测效应[J]. 现代外语, 2022(3): 331-343.
- [17] Lu, X. (2010) Automatic Analysis of Syntactic Complexity in Second Language Writing. *International Journal of Corpus Linguistics*, **15**, 474-496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- [18] Lu, X. (2011) A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development. *TESOL Quarterly*, **45**, 36-62. <https://doi.org/10.5054/tq.2011.240859>
- [19] Kyle, K. and Crossley, S.A. (2018) Measuring Syntactic Complexity in L2 Writing Using Fine-grained Clausal and Phrasal Indices. *The Modern Language Journal*, **102**, 333-349. <https://doi.org/10.1111/modl.12468>
- [20] Li, Y. and Yang, R. (2023) Assessing the Writing Quality of English Research Articles Based on Absolute and Relative Measures of Syntactic Complexity. *Assessing Writing*, **55**, Article ID: 100692. <https://doi.org/10.1016/j.asw.2022.100692>
- [21] Zhang, Z. (2020) Engaging with Automated Writing Evaluation (AWE) Feedback on L2 Writing: Student Perceptions and Revisions. *Assessing Writing*, **43**, Article ID: 100439. <https://doi.org/10.1016/j.asw.2019.100439>
- [22] Chen, W. and Zhang, M. (2023) Understanding an Assessment Approach in Computer-Mediated Collaborative Writing: Learner Perceptions and Interactions. *Language Awareness*, **33**, 135-162. <https://doi.org/10.1080/09658416.2023.2180513>
- [23] Evans, N.W., Hartshorn, K.J., Cox, T.L. and Martin de Jel, T. (2014) Measuring Written Linguistic Accuracy with Weighted Clause Ratios: A Question of Validity. *Journal of Second Language Writing*, **24**, 33-50.

<https://doi.org/10.1016/j.jslw.2014.02.005>

- [24] 吴雪峰, 柳焯琛, 殷缘. 英语写作评分标准模型的建构及其效度研究[J]. 外国语, 2018(5): 137-146.
- [25] 郭茜, 冯瑞玲, 华远方. ChatGPT 在英语学术论文写作与教学中的应用及潜在问题[J]. 外语电化教学, 2023(2): 18-23+107.
- [26] 魏爽, 李璐遥. 人工智能辅助二语写作反馈研究——以 ChatGPT 为例[J]. 中国外语, 2023, 20(3): 33-40.
- [27] 陈莱, 吕明臣. ChatGPT 环境下的大学英语写作教学[J]. 当代外语研究, 2024(1): 161-168.
- [28] Li, J., Link, S. and Hegelheimer, V. (2015) Rethinking the Role of Automated Writing Evaluation (AWE) Feedback in ESL Writing Instruction. *Journal of Second Language Writing*, **27**, 1-18. <https://doi.org/10.1016/j.jslw.2014.10.004>
- [29] Abdelhalim, S.M. (2023) Online Writing Instruction: Understanding University EFL Instructors' Beliefs and Practices as a Complex System. *Language Teaching Research*. <https://doi.org/10.1177/13621688221150855>
- [30] Kyle, K., Crossley, S.A. and Jarvis, S. (2020) Assessing the Validity of Lexical Diversity Indices Using Direct Judgements. *Language Assessment Quarterly*, **18**, 154-170. <https://doi.org/10.1080/15434303.2020.1844205>
- [31] 张洁, 王敏, 陈康. 体裁对读后续写中协同效应和写作质量的影响[J]. 现代外语, 2023, 46(2): 259-269.
- [32] Crossley, A.S. and McNamara, S.D. (2016) Say More and Be More Coherent: How Text Elaboration and Cohesion Can Increase Writing Quality. *Journal of Writing Research*, **7**, 351-370. <https://doi.org/10.17239/jowr-2016.07.03.02>
- [33] 贾文峰, 张培欣. TEM8 写作测试文本的语篇特征与写作质量的关系研究[J]. 外语测试与教学, 2020(1): 1-9.
- [34] 冯志伟, 李颖. 自然语言处理中的预训练范式[J]. 外语研究, 2021(1): 1-14+112.