

文本可读性与知识传播效率——基于林业期刊摘要的文献计量学分析



刘玲雨, 罗桂花*

中南林业科技大学外国语学院, 湖南长沙 410004

摘要: 可读性反映了文本因写作风格而易于理解的程度。摘要是学术论文知识传播的重要窗口, 其可读性不仅关系到读者对学术研究的初步理解, 也可能影响学术成果的检索与传播。现有可读性研究多集中于医学、信息科学、语言学等领域, 较少关注林业学术语篇。林业研究涉及自然资源保护、生态治理和环境政策制定等重要议题, 探讨林业学术摘要的可读性有助于理解该领域知识传播的语言特征。本研究以 2008—2021 年 Web of Science 数据库收录的 198 种林业 SCI 期刊论文摘要为主要研究对象, 共筛选出 5285 篇摘要作为分析样本, 采用四个经典可读性指数探讨其与被引频次的关系。研究发现, 林业期刊摘要整体可读性较低, 读者通常需要较高教育水平和专业知识基础才能充分理解文本。随着时间推移难度呈微弱上升趋势, 这与学科知识积累和专业术语增加相关。可读性与标准化被引频次之间未呈现显著相关性, 但包含符号的复杂表达在林业论文引用中较不受青睐。本研究为学者探讨可读性与科学影响力的关系提供了普适性启示, 在未来的学术研究写作中, 建议研究者在保证摘要信息充分性的基础上, 兼顾专业术语所体现的学科知识积累与文本表达的可理解性。

关键词: 文本可读性; 被引次数; 摘要; 林业

DOI: 10.57237/j.ssr.2026.03.006

The Relationship Between Readability and Citations in Forestry Journals

Liu Lingyu, Luo Guihua*

School of Foreign Languages, Central South University of Forestry and Technology, Changsha 410004, China

Abstract: Readability refers to the ease of understanding or comprehension of a text due to the style of writing. As an important window for knowledge dissemination in academic papers, the readability of abstracts is not only related to readers' initial understanding of academic research, but may also affect the retrieval and dissemination of scholarly findings. Existing readability studies have mainly focused on fields such as medicine, information science, and linguistics, while relatively little attention has been paid to forestry academic discourse. Forestry research involves important issues such as natural resource conservation, ecological governance, and environmental policy-making. Therefore, examining the readability of forestry abstracts can help reveal the linguistic features of knowledge dissemination in this field. The present study examined abstracts published in 198 forestry journals indexed in the Web of Science database from 2008 to 2021. A total of 5,285 abstracts were selected as the analytical sample. Four classic readability indexes were utilized to analyze the

基金项目: 湖南省社科基金项目—林学英语学术词表开发与应用研究 II (项目编号: 23YBA116).

*通信作者: 罗桂花, luowenting8552@126.com

收稿日期: 2026-05-07; 接受日期: 2026-06-05; 在线出版日期: 2026-06-13

<http://www.socscires.com>

readability of forestry abstracts and its relationship with citations. The results showed that forestry abstracts are generally difficult to read, and readers usually need a relatively high level of education and disciplinary knowledge to fully understand the texts. The difficulty of the abstracts increased slightly over time, which may be related to the accumulation of disciplinary knowledge and the increasing use of technical terms. The study found an insignificant relationship between readability and the normalized citations via correlation analysis. However, complex expressions with symbols are less preferred in forestry article citations. This study provides general implications for scholars to further investigate the relationship between readability and scientific impact. For future academic research and writing, it is suggested that researchers should ensure the informativeness of abstracts while also taking into consideration both the disciplinary knowledge reflected by technical terms and the comprehensibility of textual expression.

Keywords: Readability; Citations; Abstracts; Forestry

1 引言

可读性受特定文本中影响读者理解与参与的所有要素制约[1]。克拉尔将其凝练定义为“因写作风格而产生的理解便利性”[2]。文章若与读者能力不匹配，会阻碍信息的传播与接收，而当文本难度略超读者知识基础时，学习效果最佳[3]。因此，基于不同读者群体的可读性研究对知识传播至关重要，相关研究已涵盖教育、医学、军事、政府出版物、经济报告、商业网站及学术写作等领域。

早期可读性测量聚焦于评分标准的开发，主要基于语言风格因素或个体经验判断[4]。其中，词长和句长是预测阅读难度的核心语言特征，对可读性公式影响深远。1923 年，Lively 和 Pressey 基于此词表推出首个可读性公式，用于编写初中科学教材[5]。20 世纪 20 年代至 80 年代，超过 200 个可读性公式相继涌现[6]，最常用的传统特征包括经典的弗莱士易读度（Flesch Reading Ease, FRE）、弗莱士-金凯德（Flesch-Kincaid）、SMOG、自动可读性指数（Automated Readability Index）等。其他基于计算的指标如形符比率（TTR）也在同期得到开发。TTR 是 Templin 纳入测量的词汇多样性指标，最初用于儿童语言研究，计算样本中唯一词类型占总标记的比例[7]。作为词汇多样性和表达能力的衡量标准，TTR 值在语料库语言学中广泛应用于跨文本和语料库的词汇多样性研究。

学术写作可读性包括跨学科平行比较[8, 9]及单一具体学科领域研究。Bauerly 等使用 FRE 公式研究了市场营销期刊的 70 篇文章，结果分数在“困难”到“非常困难”之间波动[10]。Dolnicar 和 Chapple 考察了旅游期刊文章的可读性，证实了旅游文章的高难度本质[11]。雷蕾和严绍典使用 FRE 和 SMOG 指数分析了信息科学四

种期刊文章的可读性，发现该领域摘要和全文均难以阅读，前者整体保持稳定水平，后者随时间推移可读性有所提高[12]。Wang 等采用九个可读性指数分析语言学期刊摘要的可读性，发现摘要可读性低且可读性持续下降[13]，这些发现均证实了学术出版物的高难度阅读特性。

被引频次通常被视为科学影响力的重要指标[14]，可能影响学术成果的传播。研究发现，被引频次低的学术出版物可能被认为质量较低[15]，但学术可信度可能更多与沟通的复杂性而非信息的本质相关[16, 17]。因此，关于可读性与被引频次关系的研究对学者改进语言以提升学术效果具有深远意义。Dowling 等调查了《经济通讯》的文章，发现可读性分数与被引频次呈正相关[18]。然而，这种关系并不具有普遍性，部分研究发现更复杂的文本反而会受到更多引用。例如，Ante 分析了新兴技术领域 135000 篇摘要，发现可读性较低的文本更可能被引用，尤其是在不成熟的研究领域，表明语言复杂性可能被视为可信度或专业性的标志[19]。在语言与语言学领域，Wang 等观察到过去三十年摘要可读性下降，可读性与被引频次之间存在弱但统计上显著的负相关[13]。其他研究发现可读性与被引频次无显著关联，Hartley 等通过四个实验发现，按被引频次划分时无显著差异，表明可读性与被引频次无直接关联[20]。Didegah 和 Thelwall 在考察高被引论文影响因素时，将摘要可读性与作者、期刊影响因子及文献长度等元变量并列分析，说明可读性应置于多重文献属性中理解[21]。雷蕾和严绍典分析了 2003—2012 年四种信息科学期刊的摘要和全文，发现信息科学摘要整体阅读难度较高，摘要和全文可读性变化甚微，与被引频次几乎无关联[12]。

然而, 学术写作可读性研究尚未扩展到林业领域。林业研究的传播与自然资源的再生和保护及环境政策的制定密切相关, 因此, 调查林业期刊文章的可读性对林学知识传播至关重要。本研究基于林业期刊摘要语料库, 使用四个可读性指标, 旨在确定林业学科是否存在特定学科特征和变化趋势。研究目的有三: 一是考察林业期刊摘要的可读性难度水平; 二是探究可读性随时间的变化趋势; 三是识别可读性与以被引频次为目标的科学影响力之间的关系。

2 研究方法

2.1 数据收集

本研究从 198 种林业期刊中提取了 5285 篇摘要 (不含标点约 1242486 词), 总被引频次为 23187 次。由于 Web of Science (WOS) 是权威的文献数据库[22],

故选择其作为语料库来源。数据通过高级检索表达式“TS=Forestry AND DT=Article AND LA=English AND PY=2008-2021”在 WOS 中筛选, 并在更精细层面进行选择, 共检索到 5285 篇林业期刊摘要, 以考察林业出版物的可读性及其与被引频次的关系。这些期刊均收录于科学引文索引扩展版 (SCIE), 部分收录于社会科学引文索引 (SSCI), 涵盖林业植物学、森林生态学、造林学、测树学、森林管理学、森林经济学、遥感、土壤科学等主要子学科。选择 2008—2021 年的时间跨度有两层原因, 一是过去 10 年的数据能有效反映当前研究趋势[12], 且所选时间段数据完整可靠; 二是在文献计量学研究中, 最新三年的文章被引频次统计时间过近, 可能导致数据集出现偏差。

研究采用四个指标考察摘要可读性: FRE [23]、SMOG [24]、Flesch-Kincaid [25] 和自动可读性指数[26]。这些指标均将句长和词汇密度等因素纳入计算, 具体公式如下所示:

$$\text{FRE} = 206.835 - 1.015 \times (\text{平均句长}) - 84.6 \times (\text{平均词音节数}) \quad (1)$$

$$\text{Flesch-Kincaid grade level} = 0.39 \times (\text{平均句长}) + 11.8 \times (\text{平均词音节数}) - 15.59 \quad (2)$$

$$\text{SMOG} = 1.043 \sqrt{\text{多音节词数量} \times \frac{30}{\text{句子数量}}} + 3.1291 \quad (3)$$

$$\text{Automated Readability} = 4.17 \left(\frac{\text{字符数}}{\text{词数}} \right) + 0.5 \left(\frac{\text{词数}}{\text{句子数}} \right) - 21.43 \quad (4)$$

FRE 通过计算平均句长和每词音节数来评估文本可读性, 其得分范围在 0—100 之间, 分数越高, 文本的可读性越强。Flesch-Kincaid 年级水平基于 FRE 衍生而成, 得分范围为 0—18, 每三年为一个周期, 分别对应基础、平均和高级的可读性水平。SMOG 和自动可读性指数的得分则代表了读者理解文本所需的受教育年限。其中, SMOG 公式的主要特点是通过计算多音节词的数量来反映段落的词汇复杂性; 而自动可读性指数则不计算音节, 而是通过计算字符数来评估词汇难度, 字符数越多, 词汇难度越大。Flesch-Kincaid、SMOG 和自动可读性指数均以教育标准为基础来划分难度等级, 其得分与阅读难度呈正相关关系。FRE 作为经典公式, 在可读性研究中得到广泛应用, 便于与早期研究对比[21, 12]。SMOG 通过计算段落中多音节词的固有数量, 以 30 个句子为可读性判断标准[24], 长度标准适配摘要测量。Flesch-Kincaid 和自动可读性指数

基于教育标准划分难度等级, 可提供补充解释[25, 26]。综上所述, 这四个指标是经过广泛认可且在前期研究中被大量应用的可读性测量工具。

2.2 数据处理与分析

首先, 使用 Alpha Readability Calculator (ARC)¹ 计算可读性分数。ARC 可稳健地计算九种可读性指数, 以及词汇和句法层面的多种指标。其次, 计算 2008—2021 年每篇林业出版物的标准化被引率, 标准化被引频次指每本期刊平均一篇文章的被引率, 用于衡量每篇文章的被引次数, 以文章被引频次之和除以该期刊当年的平均被引次数[27]。当一篇文章的 RCR 等于 1 时, 代表该期刊理想文章的平均被引频次。最后, 使用斯皮尔曼等级相关分析 (Spearman's rho) 探究可读性分数与标准化被引频次的关系, 斯皮尔曼可通过等级转换减少干扰因素, 适用于初步相关性探索。

1 <https://github.com/leileibama/AlphaReadabilityCalculator>

3 研究结果

3.1 林业期刊摘要可读性水平

句法复杂性和词汇难度是阅读过程的重要组成部分

表 1 林学摘要的词汇特征

指标	样本量	均值	中位数	标准差	最小值	最大值
每词音节数	9733	1.84	1.84	0.17	0.99	2.45
每词句数	133326	25.23	24.50	6.24	8.90	209
形符比	3180	0.60	0.60	0.06	0.29	0.89

2008—2021 年林业期刊摘要中，每词平均音节数为 1.84，最小值为 0.99，最大值为 2.45，表明多音节词较少；平均每句 25 词，最小 8.90，最大 209，句长差异大，整体呈现为中等长度。TTR 用于量化文本在给定语料库中的词汇多样性，通过词类型数除以词标记总数计算。作为词汇丰富度指标，TTR 值越高词汇多样性越大，越接近 1 表明文章用词复杂度越高，反

分[28]，可作为语料库语言学视角下文本可读性的补充指标。林业期刊的词汇和句法分布见表 1。

映词汇多样性增强和可读性难度提升，TTR 随文本长度增加而降低，适用于林业科学期刊摘要等较短文本，其平均 TTR 值为 0.6，表明词汇多样性中等偏高。

同时，研究计算了 FRE、Kincaid、SMOG 和自动可读性分数考察林业摘要可读性难度，描述性统计见附录（附表 1—附表 4），各指标的分数解释依据见附录（附表 5—附表 8），可读性分数解释见表 2。

表 2 林业摘要可读性分数

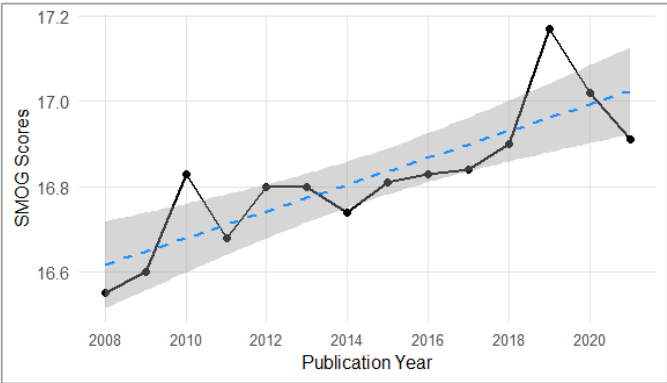
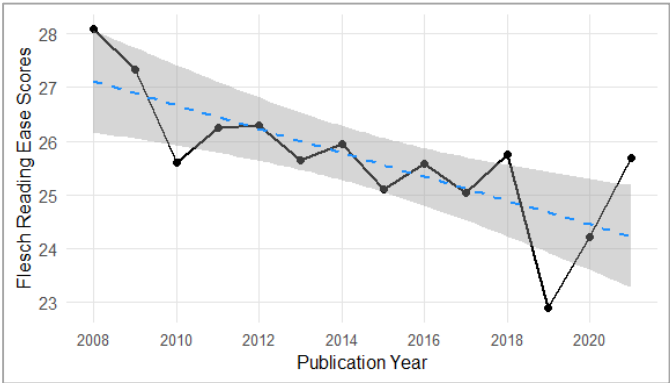
指标	平均值	中位数	分数范围	教育水平	阅读难度
FRE	25.67	25.51	-163.97-92.23	大学本科	非常困难
Kincaid	15.94	15.86	8.13-32.96	研究生	高级
SMOG	16.82	16.76	11.42-25.53	大学本科	难以阅读
Automated Readability	17.87	17.65	9.89-38.69	教授	难以阅读

FRE 得分越高，表明文本可读性越强。林学摘要的 FRE 最高分为 92.23，最低分为-163.98，表明林业期刊可读性差异较大。所选时间段内均值和中位数分别为 25.67 和 25.51，均属于“非常困难”范围。Kincaid 分数范围从 8.31 至 32.69，均值 15.94，中位数 15.86，平均得分属于大学研究生的“高级”阅读水平。SMOG 衡量显示，摘要平均分数 16.82，最低 11.42，最高 25.53，中位数 16.76 接近均值，表明读者至少需接受 16 年教育才能阅读林学期刊摘要，对应本科后期至研究生教

育阶段。自动可读性指数的均值和中位数分别为 17.87 和 17.65，超过美国教授的年级水平，即林业期刊读者应具备高学术背景，才能够无障碍阅读该领域专业文章。四个指标结果均证实林业学术摘要阅读较困难。

3.2 可读性随时间变化趋势

林业学术摘要的 FRE、Kincaid、SMOG 和自动可读性分数的可读性趋势见图 1。



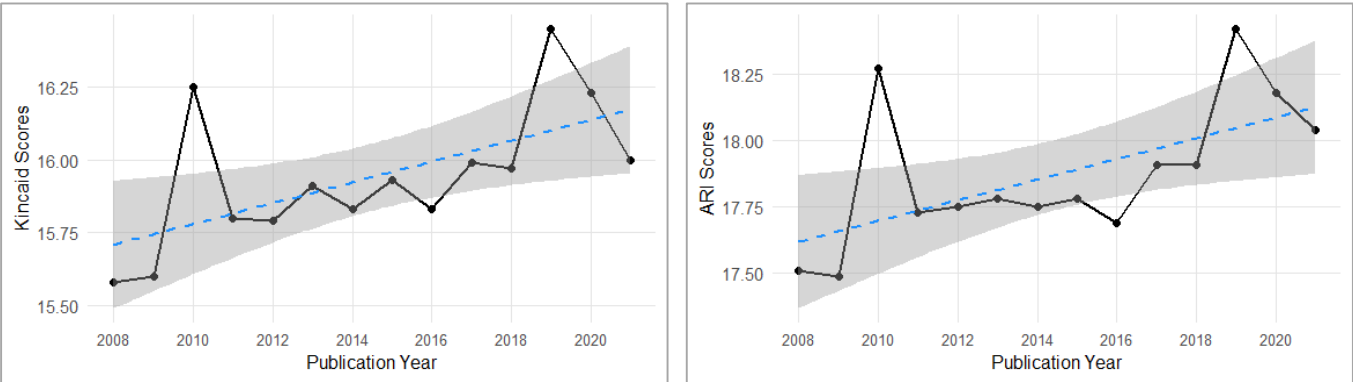


图 1 可读性随时间变化趋势

在 2008—2021 年间，林业论文摘要的 FRE 分数呈下降趋势，而 SMOG、Kincaid 和自动可读性分数则呈上升趋势。由于可读性与 FRE 正相关，而与后三者负相关，这表明四种指标均显示林业期刊的可读性在下降，整体仍处于困难水平。具体而言，摘要的 FRE 可读性随时间小幅波动下降，图 1 显示 FRE 分数呈下降趋势，简单线性回归结果表明这一下降趋势统计显著，尽管效应量较小 ($F(1,5283)=12.83, p=0.00034, \text{Cohen's } d=-0.0986$)。SMOG 分数值越高，文本可读性越低，其分数在研究期间小幅波动变化，2008—2020 年 SMOG 可读性分数显著上升，2017 年后尤其急剧上升，这表明文本阅读难度逐年增加，出版年份与可读性分数的线性关系显著，尽管效应量较小 ($F(1,5283)=18.86, p=0.000014, \text{Cohen's } d=0.1195$)。按年级水平分层，Kincaid 和自动可读性分数均呈相似上升趋势，简单线性回归分析显示显著，由于这两个分数与阅读易度负相

关，这表明 Kincaid ($F(1,5283)=10.439, p=0.001241, \text{Cohen's } d=0.0889$) 和自动可读性 ($F(1,5283)=10.485, p=0.001211, \text{Cohen's } d=0.0891$) 的摘要可读性随时间增加，尽管变化幅度有限。

3.3 可读性与标准化被引次数的关系

相对被引率 (RCR) 用于评估林业论文的被引情况，实现了不同学术出版物期刊影响力的直接比较和线性排名。2008—2021 年林业文章平均 RCR 为 1，作为期刊特定标准化后的基准值，消除了时间对被引频次的影响。尽管平均 RCR 保持 1 不变，但中位数范围为 0—1.16，2020 和 2021 年中位数为 0，表明这两年出版的一半林业文章被引率接近 0，标准化被引频次 0.00—79.20，表明林业论文被引频次差异大。

表 3 标准化引用的描述性统计分析

Year	N	Mean	Median	Std. deviation	Minimum	Maximum
2008	194	1.00	1.16	6.39	0.00	57.00
2009	241	1.00	1.21	3.95	0.00	41.00
2010	251	1.00	1.20	7.93	0.00	79.20
2011	311	1.00	0.87	3.03	0.00	34.46
2012	328	1.00	0.98	2.17	0.00	13.70
2013	302	1.00	0.75	2.60	0.00	31.63
2014	379	1.00	0.59	2.51	0.00	34.15
2015	372	1.00	0.49	1.65	0.00	20.20
2016	403	1.00	0.37	1.45	0.00	22.04
2017	438	1.00	0.32	0.83	0.00	6.54
2018	493	1.00	0.18	0.58	0.00	6.43
2019	480	1.00	0.09	0.29	0.00	2.67
2020	540	1.00	0.00	0.14	0.00	1.26
2021	553	1.00	0.00	0.01	0.00	0.15
Total	5285	1.00	0.59	2.40	0.00	79.20

通过斯皮尔曼等级相关 (rho) 分析可读性分数与 RCR 的关系，四个分数的散点图见图 2。

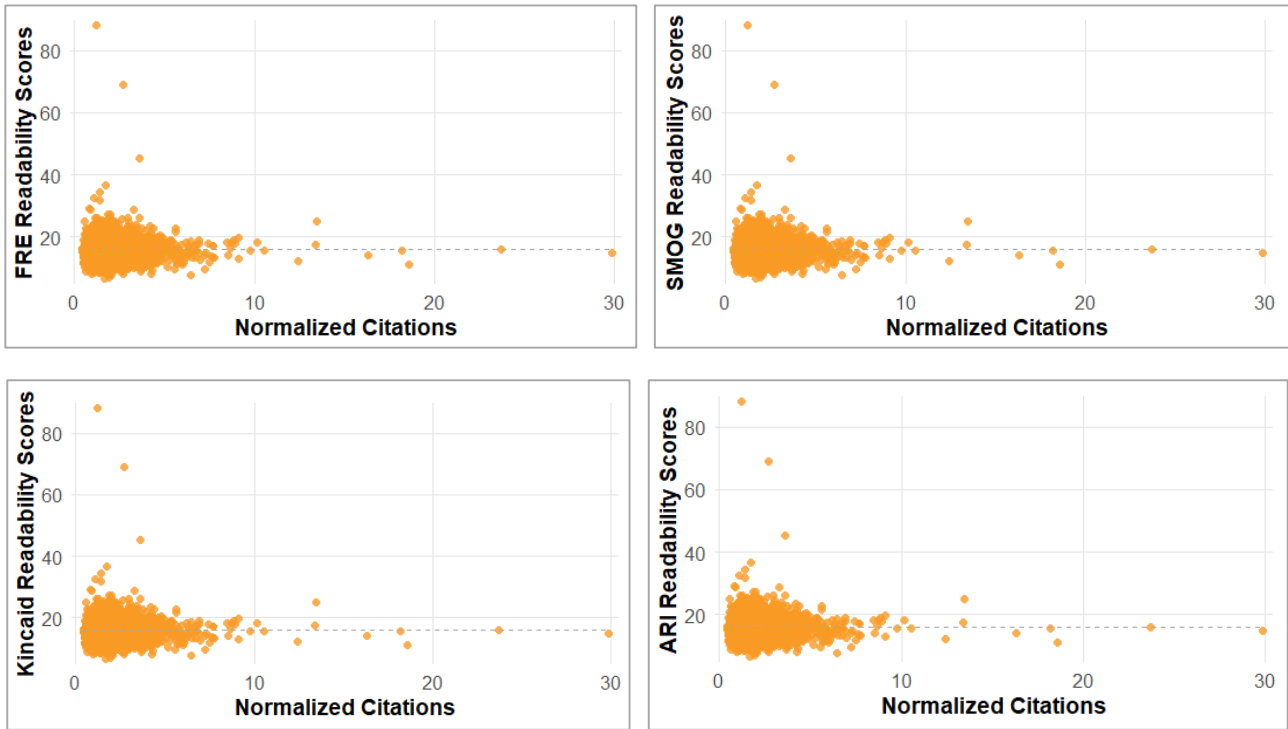


图2 可读性分数与标准化引用散点图

斯皮尔曼秩相关分析结果显示，三个变量与标准化被引频次无显著相关，即 FRE 指数 ($r=0.017$, $p=0.223$)、SMOG 指数 ($r=-0.01$, $p=0.453$) 和 Kincaid 指数 ($r=-0.02$, $p=0.149$)，表明可读性影响被引次数，但未达到统计学显著。自动可读性分数与被引频次呈现极弱的负相关 ($r=-0.027$)，在 5% 显著性水平下略达统计显著 ($p=0.0473$)，由于相关强度接近于零，其对实际引用次数的解释力可忽略不计，无足够影响力说明可读性得分能用于预测或解释被引量的变化。总体而言，四种经典指标显示林业期刊文章可读性与被引频次关系甚微。

4 讨论

本研究基于自建林学学术论文摘要语料库，深入考察了可读性水平、其随时间的变化趋势以及与被引频次的潜在关系。研究发现，林业期刊摘要的可读性整体处于困难水平。具体来看，FRE 分数在 30—50 之间表示“困难”，低于 30 则为“非常困难”。本研究中，林业摘要的平均 FRE 分数为 25.67，这表明文本的可读性较低。类似地，旅游期刊摘要的平均 FRE 分数为 18.0 [11]，也属于“非常困难”的范畴。在 SMOG 指数中，分数 11 及以上意味着对阅读技能有一定要求，相当于 17

岁及以上人群的阅读水平。林业文本的平均 SMOG 分数为 16.82，这意味着读者需要具备研究生阶段的知识基础才能完全理解文本内容。信息科学领域的期刊摘要同样要求读者完成 16 年的教育[12]。根据 Kincaid 公式，12—15 分代表大学生的高级阅读水平，15—18 分则代表研究生的高级阅读水平；而自动可读性指数中，分数 13 对应大学生的阅读能力，14 及以上则对应 24 岁以上教授的识字水平。Wang 对语言学期刊的研究表明，自动可读性指数为 18.46，Kincaid 指数为 16.87，对于非熟练读者来说难度较高[13]。本研究中，林业文章摘要的 Kincaid 平均分数达到“高级”水平，而自动可读性指数则暗示需要教授级别的阅读理解实力。

尽管不同学科之间存在差异，但研究结果证实了学术出版物普遍具有高难度的特性，这一发现与先前的结果一致[8, 12, 13, 19, 29, 30]。除了使用传统可读性公式进行宏观测量外，TTR 值还表明，林业文章摘要的词汇多样性高于平均水平。低可读性与高词汇多样性的特征可能源于林业语言的语义特性，尤其是技术术语的广泛使用。Sawyer 指出，复杂主题可能需要更复杂的写作技巧，例如使用行业术语来准确传达信息[31]。学术写作通常面向专业群体，因此需要使用较长的技术术语[11]。

工程类文章通常包含大量特定领域的术语。在林

业出版物中,林业文本包含大量中低频词和技术术语,理解这些文本需要掌握 8000—11000 词族的知识,这远高于一般学术文本的要求[32]。Hsu 通过分析一个包含 457 万字的工程教材语料库,发现英语作为外语的工程专业本科生需要掌握 5000 常见词词族,结合专有名词、复合词和缩写,才能实现 95% 的词汇覆盖率[33]。研究者开发了工程英语词汇表(EEWL),以帮助学生填补基础词之后的词汇缺口,具有一定的跨学科适用性。由此推断出,促进学术词汇的标准化使用可能对提高学术论文的可读性具有积极影响。

此外,还存在不透明词汇现象,即某些词汇在特定学术语境中的含义与其常用含义不同[34]。Todd 创建了一个包含 186 个工程关键词的列表,其中识别出 41 个高频不透明词。研究发现,这些词汇在工程教材中的含义与普通英语或在线词典的定义有显著差异。其中 64% 的不透明词是常见高频词,但在工程语境中有特定含义。这些术语多为通用词和学科特定术语,只有一小部分是学术术语。不透明词汇的使用可能会影响论文的可读性,导致目标读者对知识产生误解,尤其是对于学术语言能力较低的读者,如学生或非英语母语者,他们可能难以自主处理这些词汇。这些研究为环境和教育政策制定者以及信息检索研究者提供了重要启示,只有当读者能够理解研究内容时,才能在此基础上进行正确的引用。

在考察学术摘要的可读性随时间的变化趋势时,FRE、SMOG、Kincaid 和自动可读性分数均显示出下降趋势。尽管变化幅度较小,但可能反映出触底效应,即这些指标在基准年已处于高难度水平[35]。FRE 分数的显著下降趋势与旅游、市场营销和新兴技术领域的先前研究结果一致[10, 11, 19]。Plavén-Sigraý 等使用 FRE 指数检查科学期刊文章后发现,科学摘要的可读性下降,并提出通用科学行话的增加可能是导致可读性下降的原因[36]。SMOG 分数的上升趋势与信息科学的结果一致[12]。通过计算三个或更多音节的词来衡量可读性,SMOG 分数的上升表明林业学术文章摘要中使用了越来越多的多音节词,这意味着可读性的下降和复杂性的增加。摘要通常是研究文章最常被阅读的部分,有着吸引读者的重要任务[37]。因其不可或缺的特性,摘要应包含足够的信息。信息性摘要作为研究人员快速定位符合其兴趣研究的重要导航工具,需要在通常仅数百字的有限文本空间内概括全文并呈现核心信息。作为论文全文关键论点和发现的简洁反映[12, 38],若摘要不够充分,读者可能不会选择继续阅读文

章的其余部分[39]。

同时,学术知识的稳步积累也增加了这种复杂性。随着学术研究的不断进展,越来越多的概念、主题、理论、范式和方法等知识得以积累[11, 36]。由于这种累积性,近期研究可能比早期研究承载更多的信息内容,从而导致编码难度的增加。因此,林业摘要的写作风格可能由最新术语和概念的使用所决定,以确保可信度和学术一致性。

通过斯皮尔曼秩相关分析检验可读性与引用次数之间的关系,发现林业摘要的可读性与被引频次之间总体并不显著相关。FRE、SMOG 和 Kincaid 的可读性与被引频次的相关性未达到统计显著水平,这一发现与心理学[20]、纳米科学与纳米技术[21]和信息科学[12]领域的研究结果一致。这一结果可以根据“福克斯博士现象”来解释,即复杂的结构或流畅的表达可能比实质内容更影响感知的可信度[16]。当说话者使用过于复杂的技术语言时,即使内容琐碎,也常被视为更有知识和声望。此外,目标受众的引用行为受多种因素影响,如研究相关性、学术影响力和作者声誉。还存在被引频次统计错误的微小可能性。一项研究发现,论文标题中连字符的使用会扭曲被引频次,在高被引或较旧的论文中这种效应更为明显。不同数据库处理连字符的方式不同,导致聚合结果存在差异[40]。先前研究表明,错误引用是学术领域的一个主要问题,测算估计学术界四分之一的文章存在不恰当引用[41]。

对于自动可读性指数而言,研究发现可读性与被引频次之间存在弱负相关。结果表明,更高的可读性可能伴随被引频次的轻微下降,这一发现与神经影像学[30]、新兴技术[19]和语言学[13]领域的先前研究一致。尽管自动可读性分数与被引频次呈统计学显著负相关($r = -0.027$, $p = 0.0473$),其解释率仅为 $R^2 = 0.0007$ (95% CI [0.0001, 0.0013]),表明可读性变化对被引差异几乎无实际影响。相较于生态学和环境科学领域通常观察到的中等效果量($|r| \approx 0.10-0.30$),本研究中的关联微乎其微。同时,该指数通过计算字符数来评估,计算给定文本中的字母、数字和标点符号,更多字符表示更复杂的单词。更高的自动可读性分数表示文本更难读,由此表明包含符号的复杂表达在林业文章的引用中较不受青睐。因此,作者在撰写摘要时应更注重语言的清晰与读者体验,而非仅仅追求可读性指标的微小优化以期提升引用。

不同结果可能源于四个可读性指数强调不同的语言特征,导致分数差异,从而难以与被引频次建立一致

的关系。FRE 基于句长和音节数进行计算，Kincaid 同样测量句长和音节数，SMOG 公式依赖句子数和多音节词，而自动可读性指数则关注字符数和句长。这些差异使得各指标与被引频次的关联强度不同。

这一结果可能表明可读性并不一定是知识传播的关键因素。主流的学术影响力测量主要基于被引频次，如赫希指数、g 指数、影响因子、特征因子及各种人工排名指标。被引频次作为发现重要性或科学家在研究界地位的定量指标[13]，但除了这些文献计量指标外，研究成果的实际应用也有助于评估学术影响力。能够解决或推进科学问题并转化为现实应用，如政策制定、经济增长或环境保护等，其研究产品也能确认成果学术价值。此外，期刊的开放获取和同行评价所反映的学者影响力等定性因素，也可能进一步影响被引频次[9]。

作为文章最常阅读的部分，摘要的可读性与被引频率之间是否存在相关性值得关注。在生物学、生物化学、社会科学、化学和新兴技术等领域，摘要较长的文章往往被引次数更多[19, 21, 42]。然而，对高被引文章的调查却得出了相反的结论，即较长的摘要被引次数较少[43]。这表明摘要长度并不一定能推动被引频次。值得注意的是，学术写作中摘要的复杂性会影响其被引频次。清晰的写作有助于缩小研究人员、从业者和公

众之间的差距[20]。因此，在传递足够信息的同时保持可读风格的文章够充分发挥其传播潜力。

5 结论

本研究收集 2008—2021 年林业期刊 5285 篇摘要及其被引次数，分析可读性及其与标准化被引频次的相关性。通过四个经典可读性指数计算摘要的易读性，发现林业期刊摘要可读性整体处于较高难度水平，证实了学术论文的难读本质，简单线性回归显示林业摘要随时间推移可读性越来越低，尽管效应量较小。相关性分析发现可读性与被引频次总体无显著关联，但包含符号的复杂表达在林业领域较不受青睐。研究结果为基于摘要探究可读性与科学影响力的学者提供了启示，在学术研究和写作中，应考虑摘要的信息性特征及技术术语所体现的学术知识积累。

本研究将学术写作可读性研究范围扩展到未探索的学科，在林业领域基于经典可读性公式扩展了研究，由于数据规模和可读性指标的限制，未来研究可纳入更长时间段和更大数据集，在充分探索词汇和句法特征后，使用多样化定量指标计算学术写作的语篇和语义特征可读性测量。

附录

附表 1 FRE 分数描述性统计分析

年份	样本量	平均值	中位数	标准差	最小值	最大值
2008	194	28.09	28.50	16.85	-16.13	68.56
2009	241	27.34	25.66	15.24	-11.13	78.12
2010	251	25.60	27.42	21.90	-163.98	68.68
2011	311	26.24	25.31	15.22	-11.55	71.56
2012	328	26.29	25.96	13.99	-16.83	65.56
2013	302	25.64	26.31	16.87	-57.33	80.37
2014	379	25.95	24.72	15.39	-19.46	92.23
2015	372	25.11	24.66	14.79	-16.10	75.64
2016	403	25.57	25.19	14.99	-25.68	74.94
2017	438	25.04	24.46	15.23	-21.40	68.42
2018	493	25.75	26.25	15.39	-22.21	79.53
2019	480	22.89	22.59	14.32	-12.81	64.71
2020	540	24.22	24.28	15.26	-30.22	70.03
2021	553	25.68	25.80	16.63	-25.65	79.42
Total	5285	25.67	25.51	15.86	-163.97	92.23

附表 2 SMOG 分数描述性统计分析

年份	样本量	平均值	中位数	标准差	最小值	最大值
2008	194	16.55	16.42	2.07	11.30	23.25
2009	241	16.60	16.60	1.90	11.37	22.29
2010	251	16.83	16.52	3.18	11.06	44.35
2011	311	16.68	16.75	1.92	11.54	25.08

年份	样本量	平均值	中位数	标准差	最小值	最大值
2012	328	16.80	16.85	1.84	11.90	22.36
2013	302	16.80	16.74	2.18	11.66	27.80
2014	379	16.74	16.78	2.02	11.44	24.21
2015	372	16.81	16.96	1.85	10.75	22.94
2016	403	16.83	16.78	1.85	12.17	24.21
2017	438	16.84	16.72	1.91	12.14	24.04
2018	493	16.90	16.82	1.87	11.60	24.10
2019	480	17.17	17.10	1.79	10.17	22.49
2020	540	17.02	16.80	1.90	11.87	24.68
2021	553	16.91	16.78	2.03	10.91	25.58
Total	5285	16.82	16.76	2.02	11.42	25.53

附表 3 自动可读性分数描述性统计分析

年份	样本量	平均值	中位数	标准差	最小值	最大值
2008	194	17.51	17.20	3.25	9.91	28.45
2009	241	17.49	17.24	3.09	10.79	27.88
2010	251	18.27	17.36	8.18	10.01	110.25
2011	311	17.73	17.70	3.17	9.68	33.50
2012	328	17.75	17.77	2.89	10.65	25.54
2013	302	17.78	17.33	4.04	9.64	55.91
2014	379	17.75	17.59	3.12	9.45	31.21
2015	372	17.78	17.81	2.69	11.30	26.97
2016	403	17.69	17.71	2.91	10.79	29.10
2017	438	17.91	17.69	3.15	11.11	41.50
2018	493	17.91	17.68	3.01	8.68	30.03
2019	480	18.42	18.27	2.84	9.25	29.95
2020	540	18.18	18.01	3.16	9.45	32.34
2021	553	18.04	17.68	3.37	7.73	39.05
Total	5285	17.87	17.65	3.49	9.89	38.69

附表 4 Kincaid 分数描述性统计分析

年份	样本量	平均值	中位数	标准差	最小值	最大值
2008	194	15.58	15.37	2.99	7.68	24.95
2009	241	15.60	15.62	2.81	8.68	25.49
2010	251	16.25	15.57	6.55	8.44	88.05
2011	311	15.80	15.78	2.82	8.32	28.60
2012	328	15.79	15.86	2.61	8.98	22.85
2013	302	15.91	15.78	3.51	8.01	45.35
2014	379	15.83	15.74	2.81	7.91	27.12
2015	372	15.93	15.95	2.56	9.01	23.87
2016	403	15.83	15.83	2.66	9.21	26.18
2017	438	15.99	16.05	2.84	9.43	34.42
2018	493	15.97	15.90	2.74	6.95	25.89
2019	480	16.45	16.48	2.53	9.24	26.08
2020	540	16.23	16.19	2.84	7.95	26.10
2021	553	16.00	15.87	3.07	6.49	32.66
Total	5285	15.94	15.86	3.10	8.31	32.69

附表 5 FRE 分数标准

FRE Score	Reading Level
Above 90	Very easy
80-90	Easy
70-80	Fairly easy
60-70	Standard
50-60	Fairly difficult
30-50	Difficult
Below 30	Very difficult

附表 6 SMOG 分数标准

SMOG Score	School Level	Age Range
1	1st grade	3-7
1-5	1st grafe-5th grade	7-11
5-8	5th grade-8th grade	11-14
8-11	8th grade-11th grade	14-17
11 and above	11th grade-college	17 and above

附表 7 Kincaid 分数标准

Kincaid Score	Reading Level	School Level	Age Range
0-3	Basic	Kindergarten/ Elementary	5-8
3-6	Basic	Elementary	8-11
6-9	Average	Middle School	11-14
9-12	Average	High School	14-17
12-15	Advanced	College	17-20
15-18	Advanced	Post-grad	20+

附表 8 自动可读性分数标准

Automated Readability Score	School Level	Age Range
1	Kindergarten	5-6
2	First/second grade	6-7
3	Third grade	7-9
4	Fourth grade	9-10
5	Fifth grade	10-11
6	Sixth grade	11-12
7	Seventh grade	12-13
8	Eighth grade	13-14
9	Ninth grade	14-15
10	Tenth grade	15-16
11	Eleventh grade	16-17
12	Twelfth grade	17-18
13	College student	18-24
14	Professor	24+

参考文献

[1] Dale, E., Chall, J. S. The concept of readability [J]. Elementary English, 1949, 26(1): 19-26.

[2] Klare, G. R. The measurement of readability [M]. Ames: Iowa State University Press, 1963.

[3] Wolfe, M. B. W., Schreiner, M. E., Rehder, B., et al. Learning from text: Matching readers and texts by latent semantic analysis [J]. Discourse Processes, 1998, 25(2-3): 309-336.
<https://doi.org/10.1080/01638539809545030>

[4] Gunning, R. The technique of clear writing [M]. New York: McGraw-Hill, 1952.

[5] Lively, B. A., Pressey, S. L. A method for measuring the “vocabulary burden” of textbooks [J]. Educational Administration and Supervision, 1923, 9: 389-398.

[6] Benjamin, R. G. Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty [J]. Educational Psychology Review, 2012, 24(1): 63-88.
<https://doi.org/10.1007/s10648-011-9181-8>

[7] Templin, M. C. Certain language skills in children: Their development and interrelationships [M]. Minneapolis: University of Minnesota Press, 1957.

[8] Gazni, A. Are the abstracts of high impact articles more readable? Investigating the evidence from top research institutions in the world [J]. Journal of Information Science, 2011, 37(3): 273-281.
<https://doi.org/10.1177/0165551511401658>

[9] Kousha, K., Thelwall, M. Factors associating with or predicting more cited or higher quality journal articles: An Annual Review of Information Science and Technology (ARIST) paper [J]. Journal of the Association for Information Science and Technology, 2024, 75(3): 215-244.
<https://doi.org/10.1002/asi.24810>

[10] Bauerly, R. J., Johnson, D. T., Singh, M. Readability and writing well [J]. Marketing Management Journal, 2006, 16(1): 216-227.

- [11] Dolnicar, S., Chapple, A. The readability of articles in tourism journals [J]. *Annals of Tourism Research*, 2015, 52: 161-166. <https://doi.org/10.1016/j.annals.2015.03.003>
- [12] Lei, L., Yan, S. Readability and citations in information science: Evidence from abstracts and articles of four journals (2003–2012) [J]. *Scientometrics*, 2016, 108(3): 1155-1169. <https://doi.org/10.1007/s11192-016-2036-9>
- [13] Wang, S., Liu, X., Zhou, J. Readability is decreasing in language and linguistics [J]. *Scientometrics*, 2022, 127(8): 4697-4729. <https://doi.org/10.1007/s11192-022-04427-1>
- [14] Podlubny, I. Comparison of scientific impact expressed by the number of citations in different fields of science [J]. *Scientometrics*, 2005, 64(1): 95-99. <https://doi.org/10.1007/s11192-005-0240-0>
- [15] Padial, J. M., Miralles, A., De la Riva, I., et al. The integrative future of taxonomy [J]. *Frontiers in Zoology*, 2010, 7(1): 16. <https://doi.org/10.1186/1742-9994-7-16>
- [16] Naftulin, D. H., Ware, J. E. Jr., Donnelly, F. A. The Doctor Fox lecture: A paradigm of educational seduction [J]. *Journal of Medical Education*, 1973, 48(7): 630-635. <https://doi.org/10.1097/00001888-197307000-00003>
- [17] Armstrong, J. S. Unintelligible management research and academic prestige [J]. *Interfaces*, 1980, 10(2): 80-86. <https://doi.org/10.1287/inte.10.2.80>
- [18] Dowling, M., Hammami, H., Zreik, O. Easy to read, easy to cite? [J]. *Economics Letters*, 2018, 173: 100-103. <https://doi.org/10.1016/j.econlet.2018.09.023>
- [19] Ante, L. The relationship between readability and scientific impact: Evidence from emerging technology discourses [J]. *Journal of Informetrics*, 2022, 16(1): 101252. <https://doi.org/10.1016/j.joi.2022.101252>
- [20] Hartley, J., Sotto, E., Pennebaker, J. Style and substance in psychology: Are influential articles more readable than less influential ones? [J]. *Social Studies of Science*, 2002, 32(2): 321-334. <https://doi.org/10.1177/0306312702032002004>
- [21] Didegah, F., Thelwall, M. Which factors help authors produce the highest impact research? Collaboration, journal and document properties [J]. *Journal of Informetrics*, 2013, 7(4): 861-873. <https://doi.org/10.1016/j.joi.2013.08.006>
- [22] Prancutè, R. Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world [J]. *Publications*, 2021, 9(1): 12. <https://doi.org/10.3390/publications9010012>
- [23] Flesch, R. A new readability yardstick [J]. *Journal of Applied Psychology*, 1948, 32(3): 221-233. <https://doi.org/10.1037/h0057532>
- [24] McLaughlin, G. H. SMOG grading: A new readability formula [J]. *Journal of Reading*, 1969, 12(8): 639-646.
- [25] Kincaid, J. P., Fishburne, R. P. Jr., Rogers, R. L., et al. Derivation of new readability formulas: Automated readability index, Fog count and Flesch reading ease formula for Navy enlisted personnel: Research Branch Report 8-75 [R]. Millington: Naval Technical Training Command, 1975.
- [26] Senter, R. J., Smith, E. A. Automated readability index: AMRL-TR-66-220 [R]. Wright-Patterson Air Force Base: Aerospace Medical Research Laboratories, 1967.
- [27] Schubert, A., Braun, T. Relative indicators and relational charts for comparative assessment of publication output and citation impact [J]. *Scientometrics*, 1986, 9(5-6): 281-291. <https://doi.org/10.1007/BF02017249>
- [28] Just, M. A., Carpenter, P. A. A theory of reading: From eye fixations to comprehension [J]. *Psychological Review*, 1980, 87(4): 329-354. <https://doi.org/10.1037/0033-295X.87.4.329>
- [29] Vergoulis, T., Kanellos, I., Tzerefos, A., et al. A study on the readability of scientific publications [C]//Doucet, A., Isaac, A., Golub, K., et al., eds. *Digital Libraries for Open Knowledge*. Cham: Springer International Publishing, 2019: 136-144. https://doi.org/10.1007/978-3-030-30760-8_12
- [30] Yeung, A. W. K., Goto, T. K., Leung, W. K. Readability of the 100 most-cited neuroimaging papers assessed by common readability formulae [J]. *Frontiers in Human Neuroscience*, 2018, 12: 308. <https://doi.org/10.3389/fnhum.2018.00308>
- [31] Sawyer, A. G., Laran, J., Xu, J. The readability of marketing journals: Are award-winning articles better written? [J]. *Journal of Marketing*, 2008, 72(1): 108-117. <https://doi.org/10.1509/jmkg.72.1.108>
- [32] Luo, G., Song, J., Wu, J. The lexical profile of forestry academic texts: What does it take to understand a specialized discipline? [J]. *PLoS ONE*, 2024, 19(12): e0315975. <https://doi.org/10.1371/journal.pone.0315975>
- [33] Hsu, W. Measuring the vocabulary load of engineering textbooks for EFL undergraduates [J]. *English for Specific Purposes*, 2014, 33: 54-65. <https://doi.org/10.1016/j.esp.2013.07.001>
- [34] Todd, R. W. An opaque engineering word list: Which words should a teacher focus on? [J]. *English for Specific Purposes*, 2017, 45: 31-39. <https://doi.org/10.1016/j.esp.2016.08.003>
- [35] Bottle, R. T., Rennie, J. S., Russ, S., et al. Changes in the communication of chemical information I: Some effects of growth [J]. *Journal of Information Science*, 1983, 6(4): 103-108. <https://doi.org/10.1177/016555158300600404>

- [36] Plavén-Sigra, P., Matheson, G. J., Schiffler, B. C., et al. The readability of scientific texts is decreasing over time [J]. *eLife*, 2017, 6: e27725. <https://doi.org/10.7554/eLife.27725>
- [37] Nicholas, D., Dobrowolski, T., Withey, R., et al. Digital information consumers, players and purchasers: Information seeking behaviour in the new digital interactive environment [J]. *Aslib Proceedings*, 2003, 55(1/2): 23-31. <https://doi.org/10.1108/00012530310462689>
- [38] Zhao, X., Li, L., Xiao, W. The diachronic change of research article abstract difficulty across disciplines: A cognitive information-theoretic approach [J]. *Humanities and Social Sciences Communications*, 2023, 10(1): 194. <https://doi.org/10.1057/s41599-023-01710-1>
- [39] Lambert, V. A., Lambert, C. E. Writing an appropriate title and informative abstract [J]. *Pacific Rim International Journal of Nursing Research*, 2011, 15(3): 171-172.
- [40] Zhou, Z. Q., Tse, T. H., Witheridge, M. Metamorphic robustness testing: Exposing hidden defects in citation statistics and journal impact factors [J]. *IEEE Transactions on Software Engineering*, 2021, 47(6): 1164-1183. <https://doi.org/10.1109/TSE.2019.2915065>
- [41] Todd, P. A., Guest, J. R., Lu, J., et al. One in four citations in marine biology papers is inappropriate [J]. *Marine Ecology Progress Series*, 2010, 408: 299-303. <https://doi.org/10.3354/meps08610>
- [42] Weinberger, C. J., Evans, J. A., Allesina, S. Ten simple (empirical) rules for writing science [J]. *PLoS Computational Biology*, 2015, 11(4): e1004205. <https://doi.org/10.1371/journal.pcbi.1004205>
- [43] Letchford, A., Preis, T., Moat, H. S. The advantage of simple paper abstracts [J]. *Journal of Informetrics*, 2016, 10(1): 1-8. <https://doi.org/10.1016/j.joi.2015.11.001>

作者简介

罗桂花

1982 年生, 博士, 副教授, 硕士研究生导师. 主要研究方向为学术话语分析、语用学.

E-mail: luowenting8552@126.com

刘玲雨

2002 年生, 中南林业科技大学外国语学院硕士研究生, 主要研究方向为学术话语分析、统计语言学.

E-mail: liulingyu1025@163.com