

# 基于语料库的BBC Science Focus科普文章 词汇概况及可读性研究

李文文

上海海事大学外国语学院, 上海

收稿日期: 2025年7月15日; 录用日期: 2025年8月18日; 发布日期: 2025年8月29日

## 摘要

本研究旨在讨论科普文章对于English for Academic Purposes (EAP)教学的应用价值, 深入分析其词汇特征和文本难度。研究选取了BBC Science Focus网站的科普文章作为语料, 利用AntWordProfiler软件和Flesh-Kincaid公式分别从词汇覆盖率、所需词汇量及文本可读性展开分析。结果显示, 科普文本General Service List (GSL)覆盖率为78.69%; Academic Word List (AWL)覆盖率为5.62%; 词汇需求分别为5000词(95%覆盖率)和9000词(98%覆盖率); 文本可读性在10.0~10.4之间。这表明科普文章可作为EAP学习的有效过渡材料, 作为桥梁, 帮助学习者逐步适应学术英语的词汇与表达。基于此, 本研究为EAP教学提供了详细的教学建议。

## 关键词

词汇覆盖率, 可读性, 科普文章, 学术英语

# A Corpus-Based Study of Lexical Profile and Readability of BBC Science Focus Popular Science Articles

Wenwen Li

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: Jul. 15<sup>th</sup>, 2025; accepted: Aug. 18<sup>th</sup>, 2025; published: Aug. 29<sup>th</sup>, 2025

## Abstract

This study aims to explore the application value of popular science articles in English for Academic

**Purposes (EAP) teaching, conducting an in-depth analysis of their lexical characteristics and text difficulty. The study selected popular science articles from the BBC Science Focus website as the corpus, using AntWordProfiler software and the Flesh-Kincaid formula to analyse lexical coverage, required vocabulary size, and text readability. The results showed that the General Service List of English Words (GSL) coverage of popular science texts was 78.69%; the Academic Word List (AWL) covered 5.62%; the vocabulary size were 5000 words (95% coverage) and 9000 words (98% coverage); and the text readability was ranged from 10.0 to 10.4. This indicates that science popularisation articles can serve as effective transitional materials for EAP learning, acting as a bridge to help learners gradually adapt to the vocabulary and expressions of academic English. Based on this, this study provides detailed teaching recommendations for EAP instruction.**

## Keywords

Lexical Coverage, Readability, Science Popularization Articles, Academic English

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

在当今知识经济时代,学术英语(EAP)已成为全球学术交流与合作的重要媒介。对于非英语母语的学者而言,掌握一定的学术英语能力,尤其是阅读和理解学术文本的能力,是其在学术领域取得成功的关键因素之一。

在学术英语的教学与研究中,词汇作为语言学习的核心要素,为学术英语学习者搭建起了通往专业知识的桥梁。Laufer (2010)指出,词汇研究在语言学习中起着至关重要的作用[1]。EAP 教科书虽然适应了 EAP 学习者的需求,但也具有“密集使用技术词汇和实词”的特点(Bondi, 2016) [2],因此第二语言学习者可能仍然难以理解。另一方面,英语作为第二语言(ESL)教科书既没有提供给学生接触学术词汇的范围,也没有提供他们可能需要的学术词汇的接触次数(Miller, 2011) [3]。而科普文章作为桥接大众与科学知识的重要媒介,承担了传播科学、引导大众理解复杂科学问题的重任(Ciapuscio, 2003) [4]。科普类文本中使用的语言通常不如研究文章中使用的语言专业,但它保留了与学术写作惯例相一致的正式和精确程度(Batchelor, 2023) [5]。因此,在各种形式的科普话语中,科普文章对 EAP 学习和教学特别有价值(Batchelor, 2023) [5]。

在众多科普平台中, BBC Science Focus 凭借其权威性、主题多样性受到广泛关注。此平台涵盖多个学科领域,如健康、太空、地球科学等,且文章更新频率高,内容紧贴最新科研结果与社会热点,为语料库的构建提供了丰富而稳定的数据来源。基于此,本研究选取 BBC Science Focus 中的科普文章作为研究对象,运用语料库语言学的方法,对其词汇概况及文本可读性进行全面分析,旨在揭示科普文章的词汇特点和难度水平,为 EAP 教学提供新的视角和资源。因此本文将主要关注科普文本学术词汇覆盖率(Academic Word List coverage, AWL coverage)和所需词汇量,同时引入文本可读性(Readability)作为研究科普文本的补充视角,综合分析科普文本的学术性及其语言难度。

## 2. 文献综述

研究文本词汇概况(Lexical Profile)的一种重要研究方式是检查文本的词汇覆盖率。词汇覆盖率(Lexical Coverage),即给定文本中的读者或听众已知的单词的比例(Nation & Waring, 1997) [6],可能是影响理解的最重要因素(Schmitt *et al.*, 2011) [7]。此外,文本可读性是衡量文本理解难度的指标(唐玉玲等, 2022)

[8], 也就是说, 文本易于阅读和理解的程度以可读性呈现(李绍山, 2000) [9]。且可读性将文本的难度进行完全量化(孙哲露等, 2017) [10], 可以更加清晰地分析科普文本。本研究选取应用最为广泛的 *Flesh-Kincaid* 公式计算文本的可读性, 多方面评估科普文本的教学价值。

与语法结构和主题知识相比, 学习者的词汇资源对他们的理解更重要。而估计理解文本所需的词汇知识的一种常见方法是确定达到某些词汇覆盖率所需的单词数量(Nation & Waring, 1997) [6]。Laufer 和 Ravenhorst-Kalovski (2010)研究发现, 对于书面文本, 95%和 98%是学术文本中最小和最佳阅读理解的阈值[1]。因此本文基于此结论分别确定学习者达到 95%和 98%覆盖率所需的词汇量。

国内外对于学术类书面文本的词汇概况研究逐渐增多, 研究对象主要包括学术教材(Hsu, 2014) [11], 文章摘要(Le Nguy & Ha, 2022) [12], 专业期刊(Hsu, 2011; Vuković-Stamatović, 2020) [13] [14], 报纸(Ha, 2022) [15]等。例如, Hsu (2014)研究了工程教科书的词汇量, 发现需要 5000 个词族和 10,000 个词族加上专有名词(PN)、复合词(TC)和首字母缩略词(AC)才能分别达到 95%和 98%的覆盖率[11]。Ha (2022)认为, 许多语言能力测试, 如雅思, 长期以来一直将报纸和杂志文章纳入其阅读部分, 因此, 作者分析了网络新闻(NOW)语料库的词汇概况, 以确定理解在线报纸和杂志文章所需的词汇知识, 从而为英语教学的发展提供了有益的思路[15]。李华东, 姜慧慧(2022)对国内四套主流大学英语教材的语言学术性和语言难度进行了实证研究, 发现传统大学教材偏向通用英语, 为大学英语教学改革试验提供了数据支持, 并对教材选用和教材修订提供参考[16]。

科普话语位于学术文本与普通文本之间, 以可理解的语言将学术见解传达给非专业大众的话语(Sterk & van Goch, 2023) [17]。Myers (2003)指出学术话语和科普话语不是截然不同的两种话语, 而是同一话语的连续体[18]。因此, 科普新闻文章对 EAP 学习和教学特别有价值(Batchelor, 2023) [5]。在此类观点的基础上, Muñoz (2015)将农业类准科普文章(semi-popularization articles)作为一种“中间体裁”, 通过语料库词汇描述和高频词分析, 发现农业类准科普文章学术词汇覆盖率在 6%左右, 低于学术论文(10%~12%), 并且建立了农业类词表, 帮助 ESP 教师建立明确的词汇目标, 对本科阶段的 ESP 阅读教学具有重要的启示意义[19]。对此, Yu 和 Wen (2024)做了进一步研究, 扩大科普文章研究范围, 选取生物学, 化学, 数学, 空间与物理和社会科学五大类科普文章作为语料, 研究科普文章对于学术阅读的意义[20]。

尽管现有研究已广泛应用语料库语言学和词汇概况分析方法研究学术英语与通用英语的词汇覆盖率, 但是针对科普文本的研究仍然较少。而 BBC Science Focus 作为一个内容权威、主题多样的科普平台, 其文章不仅具有广泛的读者覆盖面, 也具有较强的学术规范性。因此, 本文选取此网站科普文本作为语料, 研究科普文本的词汇覆盖率, 文本难度及其词汇理解要求, 不仅能够填补现有研究空白, 还能为 EAP 教学提供新的视角。

### 3. 研究设计

#### 3.1. 研究问题

本研究选取 BBC Science Focus 网站中的科普文章作为语料, 借助 Coxhead (2000)的学术词汇表(AWL), West (1953)的通用英语词汇表(GSL)以及 Nation (2018)的英国国家语料库(BNC)/当代美国英语语料库(COCA)词族列表, 同时通过 *Flesh-Kincaid* 公式计算文本的可读性, 综合分析科普文本的词汇概况及其难度水平。因此本文主要探讨三个研究问题:

- 1) 总语料库和三个子语料库中 GSL 和 AWL 的覆盖率分别是多少?
- 2) 要达到总语料库及三大主题的子语料库的 95%和 98%覆盖率所需的词汇量分别是多少?
- 3) 语料的难度水平如何?

### 3.2. 语料

BBC Science Focus 是由英国广播公司(BBC)推出的科学杂志,以其权威性、前沿性和深度分析而闻名。其内容由科学记者,科学家等研究人员撰写,由专业编辑团队和科学顾问审核,内容具有权威性,且涵盖多个主题,更新频率较快,这些内容为语料库的构建提供了稳定的数据来源,能够满足大规模语料的需求。本文从 BBC Science Focus 网站中收集约 900 篇科普文章,包括三大板块: health、space、planet earth, 分别对应生命科学、物理科学、地球与环境科学三大基础学科方向,能够系统反映当代科普话语的学科分布特征。具体来说,即健康,涉及人体健康、心理学、医药、健康生活方式等方面的科学知识;太空,涵盖与宇宙、行星、星系、航天探索、天文学等航空相关的科普内容;地球,聚焦地球科学,包括环境保护、生态系统、气候变化、生物多样性等内容。各个板块选取约 300 篇文章,共 900 篇,建立一个约 100 万词的语料库。

### 3.3. 工具

在词汇分析方面,主要运用的软件是 AntWordProfiler (version 2.2.1),此软件包含了 AWL、GSL 词表以及 BNC/COCA 词族列表,能够满足测试文本的词汇覆盖率以及词汇量分析方面的需求。其分析原理基于将文本语料库中的词汇与预设词表进行比对,进而生成详尽的统计数据和频率信息。软件中包含默认词表 GSL 通用服务列表(West, 1953) [21]和 AWL 学术单词列表(Coxhead, 2000) [22],需手动添加到软件中 BNC/COCA 词族列表(Nation, 2018) [23],列表由二十五个频带组成,每个频带包含 1000 词族,此外,此词族列表还包括四个附加单词列表,主要包括专有名词列表(PN)(例如, Greece, Grecian, Greenberry),边缘词列表(MV)(例如, hm, huh, er, ah),缩写词列表(AC)(例如, PHD, UFO, and UDA),复合词列表(TC)(例如, afterword, airbag, and powerboat)。在 BNC/COCA 词表, AWL 和 GSL 词表中缺席的词被标记为“不在列表中”。BNC/COCA 词族列表可以在 Paul Nation 的网站中下载:

<https://www.victoria.ac.nz/lals/about/staff/paul-nation>。

在文本难度测量方面,本研究采用了 Flesch-Kincaid Grade Level (FKGL)公式来评估文本的可读性,这是微软 Word 中内嵌的功能。FKGL 公式通过计算文本的平均句长(以单词数计算)和平均词长(以音节数计算),将文本的可读性可视化对应的美国教育年级,广泛应用于教育和语言学习领域。同时,本研究结合文本的学术词汇覆盖率,可以更全面地衡量文本的整体难度。

### 3.4. 研究过程

为了提高科普文本语料库词汇概况研究的准确性和可靠度,本研究对原始语料进行了一定的预处理。首先,修改文章中的缩写词和连字符,缩写词比如 here's 改为完整形式 here is;复合词中的连字符需由空格代替,例如 salt-feedback 改为 salt feedback,确保这些词汇在分析中不被误归类为“不在列表中”的词汇,从而避免对研究结果产生影响。其次,对于软件可能误判为“不在列表中”的 PN、AC 和 TC 等词类,进行了人工复核,并将其重新分类至相应的词表中,比如地点(例如, Charlottesville, Busoler),种族名(例如, Erdogan, Huma),通过再次检查将其归类并添加到相应列表。完成这些修正后,使用更新后的 BNC/COCA 词表对语料进行了新一轮的深入分析。

在此基础上,利用 Microsoft Word 内嵌的 Flesch-Kincaid 公式对语料进行可读性统计,计算出各子语料库及总语料库的 FKGL 值,从而对文本的可读性进行量化评估,然后对各个主题进行比较和分析。

## 4. 研究结果

### 4.1. 研究问题一

表 1 具体内容包括总语料库和三个子语料库中 GSL 和 AWL 的覆盖率, BBC Science Focus (以下简

称 SF)总语料库中 GSL 覆盖率占 78.69%，AWL 占 5.62%。

**Table 1.** The coverage of the GSL and AWL in the SF corpus and its sub-corpora

**表 1.** 子语料库和总语料库 GSL 和 AWL 的覆盖率

主题	GSL 1 <sup>st</sup> 1000	GSL 2 <sup>nd</sup> 1000	AWL	PN + AC + TC + MW	Not in the lists
Health	73.50	6.47	6.24	2.11	0.73
Planet earth	71.44	5.96	5.63	4.21	1.09
Space	73.56	5.00	5.05	3.95	0.79
Total (SF)	72.91	5.78	5.62	3.41	0.86

不同主题的科普文章在 GSL 和 AWL 的覆盖率上呈现出一定的差异性。Health 语料库 GSL 覆盖率为 79.97%，AWL 覆盖率为 6.24%，Planet earth 语料库 77.40%，AWL 覆盖率为 5.63%，Space 语料库 GSL 覆盖率为 78.56%，AWL 覆盖率为 5.05%。

Coxhead (2000)依据专业性、广泛分布性及高频性三项标准，编制了涵盖人文、商业、法律、理工四大学科领域共 570 个词族的英语学术词汇表(Academic Word List) [22]。相关研究显示，该学术英语语料库中学术词汇的覆盖率约为 10%，而在通用英语语料库中的学术词汇覆盖率则约为 1.4%。而对于本研究的科普文本语料库中，AWL 覆盖率约为 5.62%，介于英语学术文本和通用文本之间，表明科普文章在词汇难度上处于中间位置，兼具通用英语和学术英语的特点。对于 EAP 学习者而言，这种特性使科普文章成为理想的过渡材料。在 EAP 教学中，教师可以充分利用科普文章的这一优势，将其作为教学资源，帮助学习者逐步适应学术英语的词汇难度。

不同主题科普文章在 GSL 和 AWL 覆盖率上的差异，以及科普文本语料库在词汇难度上的中间位置，都能体现科普文章在 EAP 学习中的独特价值。这一研究结果不仅为 EAP 教学提供了新的视角和资源，也为教材编写者在选取和编排教学材料时提供了有益参考。

## 4.2. 研究问题二

表 2 展示了 SF 总语料库以及各个子语料库要达到 95%和 98%的覆盖率分别需要的词汇量。结果

**Table 2.** Vocabulary size

**表 2.** 词汇量需求

Word list	SF corpus	Health	Planet earth	Space
1000	88.61	89.38	88.23	88.21
2000	90.37	90.83	90.08	90.17
3000	92.63	93.23	92.47	92.2
4000	94.57	94.79	94.42	94.49
5000	95.67 <sup>1</sup>	95.76 <sup>1</sup>	95.79 <sup>1</sup>	95.47 <sup>1</sup>
6000	96.62	96.67	96.57	96.58
7000	97.26	97.24	97.15	97.35
8000	97.72	97.74	97.57	97.8
9000	98 <sup>2</sup>	98.05 <sup>2</sup>	97.83	98.07 <sup>2</sup>
10,000	98.18	98.22	98.03 <sup>2</sup>	98.24

上标 1 表示达到 95%覆盖率所需词汇量；上标 2 表示达到 98%覆盖率所需词汇量。

显示, 对于 SF 总语料库, 要达到 95% 的覆盖率需要 5000 词, 而达到 98% 的覆盖率需要 9000 词汇量。

各个子语料库词汇量需求基本一致, 三个主题要达到 95% 覆盖率需要 5000 词汇量, 而要实现对文章较好的理解, 达到 98% 覆盖率, 学习者需要 9000~10,000 词汇量。科普文章旨在向非专业读者普及科学知识, 因此其语言特点是在保证内容准确性的前提下, 尽可能使用通俗易懂的词汇和表达方式。然而, 由于科学知识本身的复杂性和专业性, 科普文章中不可避免地会涉及到一定数量的学术词汇和专业术语。这些词汇虽然在整体文本中的比例相对较小, 但对于理解文章的核心内容和概念至关重要。Yu 和 Wen (2025) 通过研究 *Scientific American* 杂志中的科普文本(包括生物学, 化学, 数学, 空间物理, 社会科学五类主题), 发现总语料库要达到 95% 和 98% 覆盖率的词汇量分别是 5000 和 10,000 [20]。本研究与前人研究结果相类似, 选取不同主题的科普文本, 证明了不同主题的科普文本均可作为 EAP 过渡学习材料的适用性。因此, 根据此结论, EAP 教师可以根据学生词汇量和阅读水平, 优先选择与之词汇量相匹配的科普文章, 避免因难度过大影响学习的积极性。

### 4.3. 研究问题三

FKGL 值共分 12 个级别, 分别与美国教育年级相对应, 用于衡量文章的可读性。FKGL 值越高, 表示可读性较低, 文本越难理解; FKGL 值越低则表示文本可读性高, 越容易理解。

**Table 3.** The coverage of the AWL and the FKGL in the SF corpus and its sub-corpora

**表 3.** 子语料库和总语料库 AWL 覆盖率与 FKGL 值

主题	AWL	FKGL
Health	6.24	10.4
Planet earth	5.63	10.4
Space	5.05	10.0
Total (SF)	5.62	10.2

表 3 呈现了各个子语料库与总语料库的 AWL 覆盖率与 FKGL 值, 可以看出, 各个主题的 FKGL 值分布在 10.0~10.4 之间, 总语料库值为 10.2, 表明其整体难度大致相当于美国十年级学生的阅读水平, 因此, 从文本整体难度水平上看, 科普文本的语言难度适中, 这也侧面证明此类文本适合作为 EAP 学习的过渡性材料, 可帮助学习者逐步适应学术英语的学习要求。

## 5. 科普文本对 EAP 教学的意义

本研究结果证实, EAP 学习者需要掌握大量的词汇, 才能进行有效阅读。根据科普文本的学术词汇覆盖率和文本可读性值, 不同主题的科普文本在词汇量需求和文本难度上存在一定差异, 因此, 在文本类型的选取层面, 建议 EAP 教师优先选取 BBC Science Focus 科普文章中的 Space 主题作为初步学习资料, 帮助学习者适应学术英语的词汇难度, 其次选取 Planet earth, Health 主题作为拔高, 逐步提升学习者的阅读理解能力和分析能力。确定主题后, 教师可以对具体文章筛选进行细化, 利用 Microsoft Word 中的 Flesch-Kincaid 公式测量文章难度, 优先选择 FKGL 值较低的文章作为初始阅读材料, 同时教师密切关注学生的进步和反馈, 根据实际情况灵活调整教学安排, 以达到不断提升 EAP 学习者的语言运用能力和学术表达水平。

在教学实践层面, 建议教师使用 AntWordProfiler 工具检索并提取出教学科普文本中的学术词汇, 根据词频整理成词汇表, 用于课堂教学, 例如, 可以提前讲解, 帮助学生克服阅读障碍, 提高阅读自信心;

教学结束后可以运用词表设计词汇练习, 测试等等, 加深学生的学术词汇记忆。此外, 本研究的科普文本中除 GSL 和 AWL 覆盖词汇, 还有四个附加词表(专有名词列表 PN、缩写词列表 AC、复合词列表 TC、边缘词列表 MV)中包含的词汇占少量比例, 难度较大, 出现在文章中可能会影响学生理解, 建议教师提前标注, 作为拓展学习资源, 供学生自主探究与学习。

另外, EAP 学习者也可充分运用 BBC Science Focus 网站(<https://www.sciencefocus.com/>), 进行自主阅读, 此网站更新迅速, 且免费提供, 学习者可以根据自己的兴趣和学习目标, 选择合适的科普文章进行阅读学习。

## 6. 结语

本研究通过分析 BBC Science Focus 科普文章的词汇概况和文本难度, 得出以下结论: 首先, 从词汇角度看, 科普文本 AWL 覆盖率约为 5.62%, 介于通用英语文本和学术英语文本之间, 从文本难度上看, 科普文本 FKGL 值在 10.0~10.4 之间, 难度中等偏上, 综合表明科普文章可以作为 EAP 学习者有效的过渡材料, 帮助 EAP 初学者逐步适应学术英语的词汇难度, 同时为 EAP 教学提供了新的视角和资源。其次, 学习者需要掌握 5000 个词族能达到 95% 的词汇覆盖率, 可基本理解文章内容; 掌握 9000 个词族可达到 98% 的词汇覆盖率, 可实现对文章的深入理解。本研究为 EAP 教学实践提供了明确指导, 同时为 EAP 学习者提供了有效的学习材料, 进而逐步提升相关专业英语学习的效率。

然而, 本文也存在一些局限性。首先, 本研究的语料选取相对单一, 仅选取了同一科普网站中的文本作为研究对象, 未来研究可以扩大语料库范围, 涵盖来自不同平台的科普文本, 以进一步验证和丰富本研究的发现。其次, 本研究并未考虑学习者学科背景的不同, 来自不同学科背景的学习者的词汇基础和需求可能存在差异, 未来研究可以针对不同学科背景的 EAP 学习者进行分类研究, 以更精准地了解科普文章在不同学科教学中的适用性和效果。

## 参考文献

- [1] Laufer, B. and Ravenhorst-Kalovski, G.C. (2010) Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size. *Reading in a Foreign Language*, **22**, 15-30.
- [2] Bondi, M. (2016). Textbooks. In: Hyland, K. and Shaw, P., Eds., *The Routledge Handbook of English for Academic Purposes*, Routledge, 323-334.
- [3] Miller, D. (2011) ESL Reading Textbooks vs. University Textbooks: Are We Giving Our Students the Input They May Need? *Journal of English for Academic Purposes*, **10**, 32-46. <https://doi.org/10.1016/j.jeap.2010.12.002>
- [4] Ciapuscio, G.E. (2003) Formulation and Reformulation Procedures in Verbal Interactions between Experts and (Semi-)laypersons. *Discourse Studies*, **5**, 207-233. <https://doi.org/10.1177/1461445603005002004>
- [5] Batchelor, J. (2023) Writing Science for Diverse Audiences: A Corpus-Based Discourse Analysis of the Language of Science News and Research Articles. Georgia State University.
- [6] Nation, P. and Waring, R. (1997) Vocabulary Size, Text Coverage and Word Lists. *Vocabulary: Description, Acquisition and Pedagogy*, **14**, 6-19.
- [7] Schmitt, N., Jiang, X. and Grabe, W. (2011) The Percentage of Words Known in a Text and Reading Comprehension. *The Modern Language Journal*, **95**, 26-43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- [8] 唐玉玲, 张宇飞, 于东. 结合深度学习和语言难度特征的句子可读性计算方法[J]. 中文信息学报, 2022, 36(2): 29-39.
- [9] 李绍山. 易读性研究概述[J]. 解放军外国语学院学报, 2000, 23(4): 1-5.
- [10] 孙哲露, 李华东, 杜馥炎. 人教版高中英语课文可读性级别研究[J]. 英语教师, 2017, 17(20): 24-29.
- [11] Hsu, W. (2014) Measuring the Vocabulary Load of Engineering Textbooks for EFL Undergraduates. *English for Specific Purposes*, **33**, 54-65. <https://doi.org/10.1016/j.esp.2013.07.001>
- [12] Le Nguy, N.Q. and Ha, H.T. (2022) Lexical Profile of Academic Written English Revisited: What Does It Take to Understand Scholarly Abstracts? *Sage Open*, **12**, 1-12. <https://doi.org/10.1177/21582440221126342>

- 
- [13] Hsu, W. (2011) The Vocabulary Thresholds of Business Textbooks and Business Research Articles for EFL Learners. *English for Specific Purposes*, **30**, 247-257. <https://doi.org/10.1016/j.esp.2011.04.005>
- [14] Vukovic Stamatovic, M. (2020) Vocabulary Complexity and Reading and Listening Comprehension of Various Physics Genres. *Corpus Linguistics and Linguistic Theory*, **16**, 487-514. <https://doi.org/10.1515/cllt-2019-0022>
- [15] Ha, H.T. (2022) Lexical Profile of Newspapers Revisited: A Corpus-Based Analysis. *Frontiers in Psychology*, **13**, Article ID: 800983. <https://doi.org/10.3389/fpsyg.2022.800983>
- [16] 李华东, 姜慧慧. 大学英语教材语言学术性和语言难度研究[J]. 杭州电子科技大学学报(社会科学版), 2022, 18(6): 59-65.
- [17] Sterk, F.M. and Van Goch, M.M. (2023) Re-Presenting Research: A Guide to Analyzing Popularization Strategies in Science Journalism and Science Communication. Springer, 141.
- [18] Myers, G. (2003) Discourse Studies of Scientific Popularization: Questioning the Boundaries. *Discourse Studies*, **5**, 265-279. <https://doi.org/10.1177/1461445603005002006>
- [19] Muñoz, V.L. (2015) The Vocabulary of Agriculture Semi-Popularization Articles in English: A Corpus-Based Study. *English for Specific Purposes*, **39**, 26-44. <https://doi.org/10.1016/j.esp.2015.04.001>
- [20] Yu, H. and Wen, J. (2025) Lexical Coverage in Science Popularization Discourse: The Case of Popular Science News from Scientific American. *English for Specific Purposes*, **77**, 45-55. <https://doi.org/10.1016/j.esp.2024.10.001>
- [21] West, M. (1953) A General Service List of English Words. Longman.
- [22] Coxhead, A. (2000) A New Academic Word List. *TESOL Quarterly*, **34**, 213-238. <https://doi.org/10.2307/3587951>
- [23] Nation, I.S.P. (2018) The BNC/COCA Word Family Lists. <http://www.victoria.ac.nz/lals/about/staff/paul-nation>