

基于文献计量可视化方法的学术英语词表分析研究(1996~2022)

宗小妮

上海海事大学外国语学院, 上海

收稿日期: 2023年2月18日; 录用日期: 2023年3月15日; 发布日期: 2023年3月27日

摘要

学术英语词表(AWL), 开始逐渐受到国内外学者的注意。本文通过采用定量与定性相结合的研究方法, 对1996~2022年的学术英语词表进行统计分析。学术词汇指各专业各学科都常用的词汇, 然而由于一方面这些学术词汇在ESP中出现的频率很高, 另一方面这些词汇在不同的学科、不同的领域中有着不同的意义, 学术词汇通常不如研究学者领域的专业词汇为他们所熟悉, 这些也是学术词汇往往给ESP学习者造成困难的原因研究, 为了明确学术词汇的未来研究方向, 同时推动学术词汇研究的快速发展, 很有必要对国内外学术词汇的研究成果进行总结和反思。本文通过探析国内外词表的发表趋势、发文国家、以及研究热点, 呈现出近年来国内外词表的发展历程, 为今后此表的研究发挥积极的作用。

关键词

学术英语词表, 学术研究, 文献计量分析, 研究热点

Analysis of Academic Word List Based on Bibliometric Visualization Method (1996~2022)

Xiaoni Zong

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: Feb. 18th, 2023; accepted: Mar. 15th, 2023; published: Mar. 27th, 2023

Abstract

Academic word list (AWL), gradually attracts the attention of scholars at home and abroad. This

文章引用: 宗小妮. 基于文献计量可视化方法的学术英语词表分析研究(1996~2022) [J]. 现代语言学, 2023, 11(3): 970-977. DOI: 10.12677/ml.2023.113136

paper presents a statistical analysis of the academic English word lists from 1996~2022 through a research approach combining quantitative and qualitative methods. Academic word list (AWL) refers to the vocabularies commonly used in various disciplines, but on the one hand, because the academic vocabulary in ESP is at a high frequency, on the other hand the vocabulary in different disciplines has different meaning of academic vocabulary in different fields, and academic words are usually less familiar to researchers than the specialized words in their fields, these are also the reason why academic vocabulary is often difficult to ESP learners, in order to clarify the future research direction of academic vocabulary, and promote the rapid development of academic vocabulary research, it is necessary to summarize and reflect on the research results of academic vocabulary at home and abroad. By exploring the publication trend of domestic and foreign word lists, publishing countries, and research hotspots, this paper presents the development process of domestic and foreign word lists in recent years, which plays a positive role for the future research of this list.

Keywords

Academic Word List, Academic Research, Bibliometric Analysis, Research Hotspot

Copyright © 2023 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

英语是一种国际交流的工具,近年来,教育国际化和学术国际化的趋势日益明显,来自不同学术背景的学生希望通过阅读英文文献进行学术研究活动。[1]因此,掌握学术词汇对于他们尤为重要:它可以帮助准确理解文章内容、避免错误理解;同时也可以帮助他们更好地应用学术词汇表达和传递信息。由此学术英语词汇更是受到不少关注成为研究热点。在学术词汇(academic vocabulary)的研究中,学者们还在探讨其使用特征、能力发展以及学术词表的创建等方面的问题,这些研究也为我们了解学术词汇提供了重要信息。然而,对于通用核心词汇、通用学术词汇及专业学术词汇的定义和习得顺序,国内外学界尚无定论[2]。

在各种学术性文章中学术词汇虽不能传递主要学术信息,但当作者表达学术观点、阐述研究活动、解释研究结果时经常需要使用这些词。学者们发现,学术词汇涵盖了约 10%的学术文本和约 16%的学校教科书,没有掌握足够的学术词汇对阅读学术文本和撰写学术文章有困难。如果消除这些障碍有助于提高学术水平,人们一致认为教学应该集中于学术词汇[3]。

AWL (英语学术词汇表),是语言学家通过追踪主流媒体和学术文章所发现的在所有学术文章中出现频率最高的 570 个英语单词集合。这个表在许多国家的英语教学中都得到了很高的评价,本文主要借用 citespace 软件分析 AWL 中词的出现频率、发表趋势、发文国家、以及研究热点。AWL 的分析研究具有重要的理论意义和实践意义。主要体现在,在教学理论以及教学实践中可以指导学习者进行独立地学习。对于 EAP 的研究者和教学者来说进行相关的教科研都是有重要意义的。并且国内相关领域的研究尚不多见,选择这一视角旨在推动国内 AWL 的编纂和 ESP 的发展。

2. 研究背景

Coxhead 的学术词汇表是一个处理大规模语料的研究,它旨在帮助学生正确理解和使用学术性的单词。通过使用语料库语言学和各种检索软件,从不可胜数的英语词汇中分辨出哪些是学术性的单词,并

将其列入学术词汇表中,以便在教学中提高学生对学术词汇的掌握,帮助学生正确理解和使用这些单词。语料库语言学的发展和各种检索软件的出现使大规模语料处理成为可能□这也为词表的制定提供了新的途径。近年来这方面比较有影响的研究便是 Coxhead 的学术词汇表。

2.1. Coxhead 的 AWL 研究背景

为了提高初级英语学习者的词汇学习效果,West 在 1953 年编纂了 2000 常用词汇表(General Service List, 简称 GSL)。GSL 包含 90%小说文本、75%非小说文本和 76%学术文本[4],可以为所有初级学习者使用。然而,随着 GSL 的使用和验证,其中单词较普通无法满足学术要求而受到一些批评。因此,为了满足学术需要,后来的研究者开始创立出具有较高出现频率、覆盖面较广、可用于学术目的的新一代学术性单词表。

2.2. Coxhead 的 AWL 开发

Coxhead [5]的 Academic Word List (AWL)是一个包含 3,513,330 词标的语料库,它来自 4 个学科(人文、商业、法律和理科)的 28 个专业以及 414 篇文本(见表 1)。AWL 收集了学术词汇中出现频率较高、覆盖面较广的单词,可以帮助英语学习者提高词汇学习效果在国际化教育和学术交流中取得成功。Coxhead 基于这个语料库,收集出现频率超过 100 次、在 15 个专业中出现 10 次以上的词族,最终完成了一份包含 570 个词族的“学术词汇表”(Academic Word List 简称 AWL) [6]。Coxhead 选择词族的三个标准是: 1) 特定出现(specialized Occurrence): 入选词族必须不在 West 的 2000 常用词汇表中; 2) 范围(range): 每个词族里的单词至少在四个学科中出现 10 次以上、至少在 15 个专业领域中有所体现; 3) 频率(frequency): 每个词族里的单词在整个文本库里至少出现 100 次[3]。

Table 1. Composition of the academic corpus
表 1. 学术语料库的组成

	学科				
	人文	商业	法律	理科	
词标	883,214	879,547	874,723	875,846	351,333
文本	122	107	72	113	414
专业	教育	会计学	宪法	生物学	
	历史—	经济学	刑法	化学	
	语言学	财政学	家庭和法医学	计算机科学	
	哲学	劳动关系	国际法	地理学	
	政治—	管理学	纯商法	地质学	
	心理学	营销学	准商法	数学	
	社会学	公共政策	权利和赔偿	物理学	

在 range 软件的帮助下 Coxhead 的 AWL 收录了 570 个词族,按照词族频率的高低分成 10 个子表。子表 1 包含了在 Coxhead 语料库中出现频率最高的词族子表 2 次之后以此类推。子表 1~9 各有 60 个词族子表 10 有 30 个词族。作为一个刚出现的新词表有必要深入研究它的效度和适用性,以及它在英语教学中的应用。目前, Mingtzu 和 Nation [7]已经从一词多义的角度对 AWL 进行了效度研究;而其他方面的

研究则少之又少。此外, AWL 收录了 570 个词族, 并把它们按频率分成 10 个子表。Coxhead 的 Academic Word List (AWL) 包含了覆盖学术语料库 10% 单词的词族, 其中 94% 至少出现在 20 个专业领域中。因此, 使用 AWL 来进行学术英语教学和进行相关研究都是非常重要的一步。

3. 研究设计

3.1. 学术英语词汇研究

本研究将采用定量和定性相结合的方法, 通过深入分析学术词表的变化趋势, 分析 1996~2022 年间 AWL 相关文献。重点梳理研究趋势、研究热点和研究突显。本文借助 WOS 检索 1996~2022 年期间的研究, 在高级检索中以 academic word list 或者 academic word lists 为主题词, 与 AWL 为主题词, 期刊类别限定为 linguistics 和 language linguistic。语种限定为英语。两者分别搜出 303 篇和 19 篇文献, 在经过 CiteSpace 分析时共收集有效论文 302 篇。

3.2. 学术英语词汇研究理论框架

英语学术词汇表简称: AWL。它是由新西兰惠灵顿维多利亚大学语言学和应用语言研究学院的 Averil Coxhead 开发的, 是一种通过语料库研究得出的, 包含了所有英语世界主流媒体和学术文章中出现频率最高的 570 个英语单词。Averil Coxhead 的研究旨在帮助学习者快速提高英语水平, 他将最常用的 570 个单词归纳出来, 作为一个参考清单, 帮助学习者先从这些单词开始学习, 然后再扩展到更多的英语单词, 帮助学习者快速提高英语水平, 从而更好地理解和使用学术文章。

该列表包含 570 个单词家族, 之所以被选中是因为它们在广泛的学术文本中出现频率很高。该列表不包括在英语中最常见的 2000 个单词(一般服务列表)中的单词, 因此使其特定于学术语境。该词表主要是为了让它可以被教师使用, 作为计划的一部分, 为学习者准备高等教育水平的学习, 或用于学生在学院和大学独自学习最需要学习的单词。词表可以广泛用于学术论文写作、各种国际英语测试如雅思 (IELTS)、托福 (TOFEL) 等。570 个单词被分成 10 组。这些组的顺序是这样的, 第一组的单词是最频繁的单词, 而在最后一组是最不频繁的单词。

4. 研究与讨论

4.1. 发文量总体趋势与特点

本文根据年度发表情况, 运用 Excel 软件分别以年份和发文量为横、纵坐标绘制出论文发表量分布图(见图 1)。随着年代的推移, 英语学术文章的发表量也在不断变化。这些数据使得 AWL 发文量更加明了。在 AWL 诞生之际, 发文量比较稳定, 少则 9 篇, 多达 15 篇, 这一阶段学术英语词表相关研究很有起色。然而此后很长一段时间年经历了大萧条, 甚至在 03 年发文量为 0, 多年发文量为 1 篇, 说明在此期间词表研究不为研究者所吸引, AWL 经过一段时间的瓶颈期后最终在 2010 年开始发文量直线上升, 逐渐开始进入成熟发展, 且始终保持一定规模的产出, 在 2021 年达到峰值, 也表现出 AWL 逐渐成为学者的研究热点。

4.2. 关键词

CiteSpace 软件的关键词聚类功能可以清晰地展现出某研究领域的热点词汇[8], 这有助于我们判断出论文内容的精华, 并更好地了解其中的研究焦点和发展趋势[4]。AWL 的另一个用处是, 它可以帮助研究者快速搜集学术文章中使用频度最高的 570 个单词, 从而更好地分析出文章的核心内容和研究焦点。

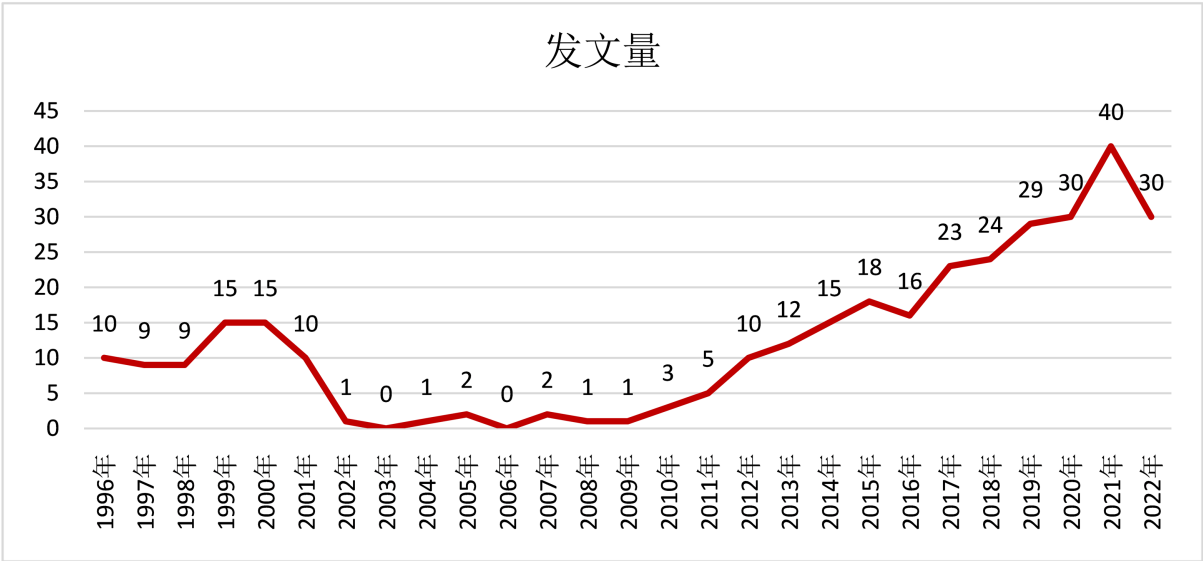


Figure 1. Number of AWL research papers published at home and abroad in 1996~2022

图 1. 1996~2022 年国内外 AWL 研究论文发文量

从词频由高到低的排序中可以看出(见表 2)，国际研究者对 AWL 的 English、language、academic vocabulary 的研究论文发表数量排在前三位；表明学者们对词汇研究本身的一个重视。frequency, wordlist, knowledge, list 和 vocabulary 等方面的研究论文发表数量次之。而学生的 comprehension、performance 相对来说比较薄弱。

Table 2. Key words of AWL research papers from 1996 to 2022

表 2. 1996~2022 年国内外 AWL 研究论文关键词

count	Keywords	count	Keywords
35	English	11	student
31	language	11	retrieval
23	academic vocabulary	10	academic writing
23	frequency	9	2rd language
21	word list	9	awl
15	knowledge	9	acquisition
15	list	9	comprehension
15	vocabulary	8	memory
14	research article	8	performance
14	academic word list	7	word
13	lexical bundle	7	episodic memory
13	children	6	collocation list
13	learner	6	lexical coverage

与此同时，我们可以看到研究的焦点已经从单词本身转向记忆单词，比如“learner”，“retrieval”，因此通过 CiteSpace 数据分析，我们可以清楚地知道单词记忆过程已经引起了人们的注意。

4.3. 主要发文国家分布

由表 3 可以看出，以 AWL 为题的文章主要在美国发表的为 116 篇，占总数的 38.2%，美国与英国占据了 51.1%，而中国仅占据了 7.5%。说明英美国家的外语学界对于词表分析研究非常重视，Coxhead 的 Academic Word List (AWL) 就是一个很好的例子。因此，使用 AWL 来进行学术英语教学和进行相关研究都是非常重要的一步。这些发现展示出英美国家外语学界对于词表分析有深入的关注和重视，以便能够开展多领域的有效词表分析研究。而同时也说明我国对词表研究并不太重视。

Table 3. Research of AWL published nations in1996~2022
表 3. 1996~2022 年 AWL 分析研究发文国家统计

citation counts	References
116	USA, 1996
39	ENGLAND, 1997
23	PEOPLES R CHINA, 2015
22	CANADA, 1998
11	AUSTRALIA, 1998
11	JAPAN, 2000
10	TAIWAN, 2014
10	SOUTH AFRICA, 1999
10	GERMANY, 1993
9	NEW ZALAND, 2014

4.4. 国家合作图谱

合作研究一直是科学计量学较关注的问题，绘制作者合作图谱能够更加清晰直观地反映出某一学术研究领域不同国家专家学者的合作关系[9]。科学合作可以提高学者科研成果的质量和影响力，大型合作群体能够产生更有说服力的成果[10]。通过绘制作者合作图谱，可以更加清晰地反映出不同国家专家学者之间的合作关系，从而有助于增强学术研究领域的影响力。本部分通过绘制国家共现图谱以揭示国内外词表研究现状。

从图 2 中可以看出国家之间分布较稀疏，说明合作较少。说明各国科研从事 AWL 研究的科研人员跨国合作的较少，大多是孤军奋战。每位作者对应一个节点，连线代表了他们之间的合作关系。总体来看，大多数国家之间存在独立发文的情况,超过半数的国家之间没有节点连线。除了发表文章数量较多的国家之间存在几个较大的聚类合作关系之外，其他聚类都呈小群体分散在四周，这说明国家之间相互联系较少。因此，在近年国内外词表分析研究中，大型科学团队尚未形成，而且各个国家之间的合作也有待加强。只有通过不断扩大合作面、加强协作、促进技术创新，才能使词表分析取得更好的发展。

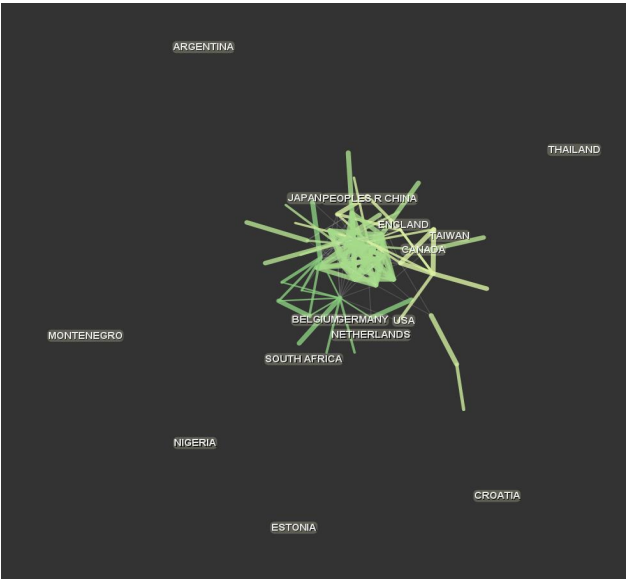


Figure 2. Co-occurrence map of country cooperation in the field of AWL from 1996 to 2022
图 2. 1996~2022 年 AWL 分析研究国家合作统计

4.5. 研究前沿分析

在 CiteSpace 中，前沿研究主要是指在本阶段起始年份即 1996 年开始突现的节点文献所反映的研究主题，共 11 个突发词。我们可以将这 11 个突发词分为三大类。第一类在 1996~1999 三年间，以词汇记忆为主题和内容，包括 memory, retrieval 和 activation 三个突发词。第二类是包括 research article, frequency, academic vocabulary, frequency 和 list，其贯穿在 2014~2017，在此阶段突现集中在词汇本身。最后一类是 English, lexical bundle 以及 knowledge 自 2018 年到现在的 2022 年仍是研究热点，词汇是英语学者及研究者都是至关重要的，词汇知识对英语学习者又是一个很大的挑战。突变率最高的是 English、research article 和 academic vocabulary，对应年份是 2018~2022、2014~2017、2017~2018，表明在此年份中为其的研究热点。通过图 3 分析可知，English、lexical bundle 和 knowledge 在近几年研究比较火热。

Top 11 Keywords with the Strongest Citation Bursts

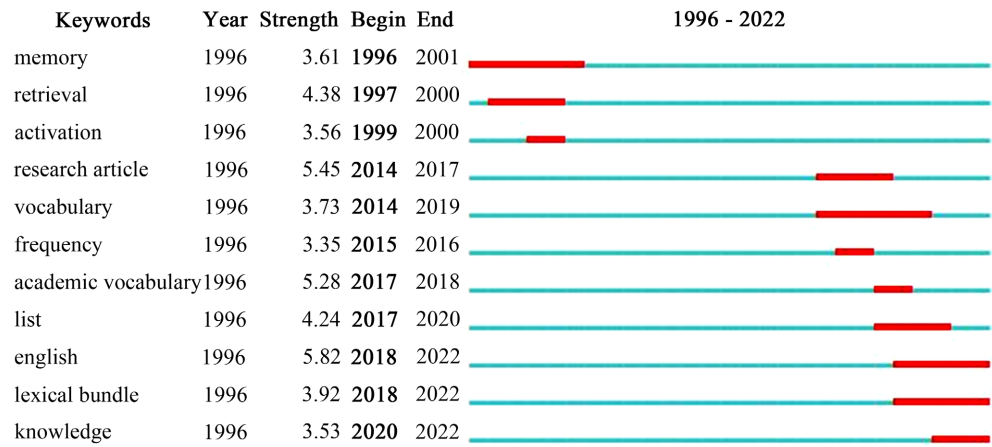


Figure 3. Burstiness of AWL 1996~2022
图 3. 1996~2022 年 AWL 突现表

5. 结论

本研究借助文献计量学知识与可视化分析软件,进行了定性与定量相结合的研究方法。对国内外近十几年学术英语词表相关文献进行了检索、统计、分析和归纳,主要从发展趋势、核心作者以及研究热点等进行总结分析。研究发现:1) 1996~2022年,国内外词表研究文献数量持续增加,而且随之而来的是不断壮大的学术队伍。总体上来看,这些发展趋势表明词表相关的学术领域正在取得可观的成果。2) 国家间连接点较少,国家机构之间合作不紧密。在近年国内外词表分析研究尚未形成大规模的科研团队,只有通过多方面的合作,才会得到更好的发展,也说明国家合作有待加强。3) 通过国家发文量的对比可以看出我国对词表的研究不足。4) 学者们对词表研究的关注已从最初的词汇记忆发展到了更多元化,包括词汇本身、词义知识、专业背景知识等。学术词表在不同国家之间有交流合作,对于英语学习者来说,要掌握学术性的英文单词是非常重要的,这也是取得高等教育成功所必备的一部分。而通过数据分析可知词汇在AWL中的充当着至关重要的作用且为广大研究者所注视。研究者今后可以深入探讨学术词汇的教学方法,以及用于加强学术词汇的练习和测验AWL教学效果的方法,相信这样的研究会有助于加深对学术词汇的认识,并且进一步完善AWL这一新的学术词汇表,以促进学术英语教学和开展相关研究。

参考文献

- [1] Swale, J.M. (1983) Vocabulary Work in LSP—A Case of Neglect. Aston University, Birmingham.
- [2] Paquot, M. (2010) Academic Vocabulary in Learner Writing: From Extraction to Analysis. Continuum, London.
- [3] 吴璟. 学术词汇表(AWL)的研究进展与启示[J]. 云南农业大学学报., 2010, 4(1): 84-89.
- [4] 段春雨, 蔡建东. 国际泛在学习领域知识图谱研究[J]. 现代远程教育研究, 2016(1): 85-95.
- [5] Coxhead, A. (2000) A New Academic Word List. *TESOL Quarterly*, 34, 213-238. <https://doi.org/10.2307/3587951>
- [6] 吴瑾, 王同顺. Coxhead “学术词汇表”的适用性研究[J]. 国外外语教学, 2007(2): 28-33.
- [7] Ming-Tzu, K.W. and Nation, P. (2004) Word Meaning in Academic English: Homography in the Academic Word List. *Applied Linguistics*, 25, 291-314. <https://doi.org/10.1093/applin/25.3.291>
- [8] 姜春林, 杜维滨, 李江波. 经济学研究热点领域知识图谱: 共词分析视角[J]. 情报杂志, 2008, 27(9): 78-80+157.
- [9] 马妍. 基于文献计量可视化方法的国内批评话语分析现状研究[J]. 考试与评价, 2020(4): 9-13.
- [10] 岳洪江, 刘思峰. 管理科学期刊引证指标的灰色关联研究[J]. 科学学研究, 2008, 26(1): 61-65.