

采矿工程专业学术英语词表的创建

王奕雯

中国矿业大学(北京), 北京

收稿日期: 2022年3月6日; 录用日期: 2022年4月5日; 发布日期: 2022年4月13日

摘要

具体学科学术词表的建立对学习具有重大意义。Coxhead在2000年推出了Academic Word List学术词汇表(简称AWL), 该词表语料来自于文、商、法及理四个学科的英文材料, 同时每门学科下还有七个子学科, 共计350万字符, 涵盖28个学科。尽管其中包含的学术词族总体上适用于理工科, 但无法充分覆盖具体学科特别是采矿工程专业的需要。采矿业作为我国经济发展的组成要素之一, 其专业学术词表的建立能够满足专业学习的需求。因此本文旨在通过语料库的方法下建立采矿工程英语学术词表。该词表通过提取汇集语料库中高频出现的词汇形成的学术词表, 且收录的词汇基于大规模真实语料, 能够助力研究者日后的论文写作与翻译实践。

关键词

学术英语词表, 词频, 语料库应用

Establishment of Academic English Vocabulary for Mining Engineering

Yiwen Wang

China University of Mining and Technology (Beijing), Beijing

Received: Mar. 6th, 2022; accepted: Apr. 5th, 2022; published: Apr. 13th, 2022

Abstract

The establishment of academic word list of specific subjects is of great significance to learners. In 2000, Coxhead launched the Academic Word List (AWL for short), which is based on English materials from four disciplines: liberal arts, business, law and science. At the same time, there are seven sub-disciplines under each discipline, totaling 3.5 million characters, covering 28 disciplines. Although the academic word families included are generally applicable to science and engineering, they cannot fully cover the needs of specific disciplines, especially mining engineering. Mining

industry is one of the components of China's economic development, the establishment of its professional academic thesaurus can meet the needs of professional learning. Therefore, this paper aims to establish a mining engineering English academic lexicon based on the method of corpus. The academic word list formed by extracting and collecting frequently occurring words from the corpus, and the included words are based on large-scale real corpus, which can help researchers in future paper writing and translation practice.

Keywords

Academic English Word List, Word Frequency, Corpus Application

Copyright © 2022 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

科学研究和学术交流的国际化趋势对专业人员,尤其是大学生的学术交际能力提出了较高要求。黄宽宏(2020) [1] 调查结果显示,82.65%的受访者表示由于专业词汇量的不足导致英文文献阅读较为困难。同样在英文学术论文写作方面,最突出的问题就是专业术语使用不当。(朱海清 2020) [2] 表明由于专业所需的术语不同,非英语专业的研究生们对研究领域内的英语专业词汇掌握有限而导致用词不当的现象时常发生。Cohen (1998) [3] 表明,由于学术词汇储备的匮乏,导致二语学习者在阅读学术类文本时感到困难的现象是普遍的。由此可见,学术词表的建立对专业学习作用重大。

学术词表是对学术文本中除高频使用的实义词和功能词以外使用频率较高的词汇汇总成表。目前,对于各专业的学术词表建设已取得显著进展,如 Academic Word List 学术词汇表 AWL (Coxhead 2000) [4] 中的语料来自于文、商、法及理四个学科的英文材料,每门学科下还有七个子学科,共计 350 万字符,涵盖 28 个学科。AWL 词表发布后引起了学者们的关注,研究者们对其进行了许多研究。吴瑾、王同顺(2007) [5] 针对 Coxhead 建立的通用学术词汇表在自然与理工学科的适用性进行研究,研究表明 AWL 包含的学术词族虽然总体上也适用于理工科,但在所选专业的侧重上有显著差异。国内不同学者也曾尝试创建具体学科的专业词表,如能源英语(景漱玉, 2020) [6] 等,大大推动了学术词汇表在具体学科应用的发展。

与此同时,何玲(2016) [7] 指出现在虽存在着针对矿业专业的纸质词典,如何长姬教授(中国矿业大学)编写的《露天采矿词汇》,当中列出了部分词汇的中英对照,但没有涉及到实际应用当中,并且还有着少量翻译错误的现象。同时市面上常见的翻译软件如有道词典,谷歌翻译等只能查询到词汇常见的意思,并不能够有效解决学习者在专业上的困惑。

尽管在国内已建立许多具有专业用途的学术词表,但有关采矿专业英语学术词表领域研究较少。目前,还存在着语料库库容较小,语料来源主要从教材,英文原著,书籍和期刊论文等,缺乏针对性和代表性。语料的年份较为久远,无法满足当下的需求。矿业作为国民经济增长的产业之一,具有十分重要的地位。因此,本研究旨在创建采矿学术词表,为该学科的翻译及学术英语教学提供借鉴。

2. 研究方法

语料库的创建及数据处理

语料的采集对于建设语料库有着至关重要的作用,其选择的语料来源要具有针对性,代表性以及时效性(丁自华, 2011) [8]。本文通过 Web of Science 数据库进行期刊检索,并咨询行业内学者后,最终确

定了采矿专业影响因子排名靠前的期刊，主要收录研究性论文的期刊 *International Journal of Rock Mechanics and Mining Sciences* 以及主要收录综述性论文的期刊 *Ore Geology Reviews* 作为语料来源。同时为了兼顾时效性与准确性，语料的发表年份来源于 2018 年至 2020 年。

每篇论文的文本整理本着遵循原文这一主要原则，重点整理语料中的图表公式、多余空格、换行符及其他不标准符号等，形成 UTF-8 编码格式的 txt 文本，并进行人工校对逐一核查。

本文采用语料库分析软件 AntConc 对语料进行统计分析。AntConc 是语料库检索工具，适用于语言学、翻译学、外语教学等领域的研究者。将采矿工程专业学术英语语料库导入 Antconc 后，采用 Crown corpus (2009) 作为对比语料库，并相应导入英语停用词表与 lemmatization (词性还原) 作为参照，所有带有不同形式的词汇都将还原成基础词性，词性的统计都以原形为单位。

Antconc 统计出语料库中出现频次较高的词汇，形成了初步的词表。在手动删除了其中留有的介词、代词等功能词后，选取其中频次大于等于 50 的实词收录入采矿工程学术英语词表，词表内共收录了 2070 词。

3. 结果与讨论

在开发词表时，首先对词汇进行了词性还原，这样能减少学习者的负担，有利于词汇的学习与记忆 (Lei & Liu, 2016) [9]。该词表共收录 2070 词，筛选标准为选取频次大于等于 50 的高频实词。通过观察词表可得出出现频率最高的词是“rock 岩石”，频率为 9051 次。其中“fracture 岩层的裂纹”(6256 次)“specimen 样本”(4139 次)“coal 采煤”(3736 次)出现频率较高，分别位于该词表的第 3，第 7，第 10 位(见表 1)。

Table 1. Mining engineering professional part of high-frequency content words
表 1. 采矿工程专业部分高频实词

序号	频次	词元
1	9501	rock
3	6256	fracture
7	4139	specimen
10	3736	coal
11	3647	process
12	3565	sample
13	3409	mineral
14	3376	test
15	3329	strain
17	3017	deposit
18	2958	mine
19	2935	ore
20	2901	surface
21	2753	joint
25	2607	energy
27	2535	strength
28	2527	temperature
33	2345	parameter
34	2335	datum
35	2247	material

在观察词表中收录的词所在的索引行时发现,词汇在文中的时态具有规律性。辛闻,王会娟(2016) [10]指出,矿业类英语文献主要以说明性质为主,句法的表达一般采用陈述式的语气,以达到表达客观严谨的目的。如 ME2_RA_16 中提到的“The surface asperity is one of the most important characteristics affecting both the mechanical and hydraulic behaviors of rock fractures during shearing.”(地表微凸性是影响岩石在剪切过程中断裂力学和水力特性的重要特征之一。)就使用了陈述式表达。其次就是矿业英语句法的显著特征之一就是频繁使用被动语态,目的是减少人为的感情色彩与主观判断,重要的表达也可以放置于句首,使读者能够理解文本。如 ME2_RA_32 中提到的“he specimen was covered with a heat shrink tape”(样品上覆盖了一层热收缩胶带)以及 ME2_RA_16 中提到的“In this method, a fractal surface is represented by a continuous and single-valued function $Z(x, y)$.”(该方法用连续的单值函数 $Z(x, y)$ 表示一个分形曲面。)

采矿工程学术英语词汇专业性极强,收录的词汇大多由普通名词或半专业词汇转化形成的,在日常生活中很少使用且词义的内涵也比较广泛。部分词汇应用于采矿专业时,会出现与其普通用法截然不同的含义。如何正确精准的识别与应用这些词汇对于学习者十分重要。例如“burden”原意为“负担,责任”,但在采矿英语中的意思为“抵抗线”。“fault”在普通用法下是“错误”,但在采矿英语中可以翻译为“断层”。

该词表中收录的词汇可与其他词汇组成固定搭配,构成采矿工程专业的专有名词,辅助学习者掌握记忆专业词汇。如 strain 原意为“压力,应变”(频次 3329)在 ME3_RA_128 中有 Both the matrix strain and the bulk strain also increase accordingly. 该词组合形成 matrix strain (基体应变)和 bulk strain (体应变)。还有 deposit 原意为“放置,定金”,在 ME_RV_07 中提到“The ore deposit is hosted within the Taft and A bkuh formations.”(矿床产于塔夫脱和 A-bkuh 岩层中。)中 deposit 与单词 ore 组合后形成的固定搭配 ore deposit 应该翻译为矿床。像这样的固定搭配还有很多,如 energy spectrum (能谱), flow parameters (气流参数), fluid displacement (流体驱替), shear cracks (剪切裂隙)等等。

同时,固定搭配形成的专有名词在文中的出现频次在一定程度上可以反映出该领域是最近学术界的热点。据统计,rock damage “岩石损伤”一词共出现 66 次。学界对其的讨论主要围绕着 1) 假设其是研究中影响结果的变量之一进行讨论。如“For blasting excavation of rock slopes, the size of the EDZ and the degree of rock damage are influenced by a large number of factors in addition to excavation methods and in-situ stress.”(对于岩质边坡的爆破开挖,除开挖方法和地应力外,还受多种因素的影响,如爆破带的大小和岩石损伤程度。)2) 强调岩石损伤的研究对于其他研究具有重大意义。原文中提到“Furthermore, much attention has been paid to the law of rock damage accumulation which is helpful for evaluation of fatigue life and establishment of damage model under cyclic loading conditions.”(此外,对岩石损伤累积规律的研究有助于疲劳寿命评估和循环加载损伤模型的建立。)

原文 1: W represents the hydraulic aperture, which is calculated based on the stretch of the bond.

原文 2: The hydraulic response of the COx is strongly influenced by stresses and temperature.

译文 1: W 为水力孔径,根据粘结的拉伸程度计算。

译文 2: COx 的液压响应受到应力和温度的强烈影响。

同为 hydraulic 一词,在不同语境下的意思分别为形容词“水力的”和名词“液压”。同样,还有:

原文 3: In top-blown copper smelting furnaces such as the TSL furnace, a single lance is installed from the top of the furnace, extending below the melt surface.

原文 4: It is likely that there will be progress in mitigating the effects of surface alteration.

译文 3: 在顶部吹制的铜熔炼炉中,如 TSL 炉,从炉顶安装一个喷枪,延伸到熔体表面以下。

译文 4: 在减轻地表变化的影响方面很可能会取得进展。

Surface 在原文 3 中的意思是“表面”，而在原文 4 中则为“地表上的”。这要求学习者需要灵活掌握词汇的多种意思的同时，还要在阅读中仔细研读上下文，准确把握词汇的正确意思。

4. 结论与应用展望

目前，有关通用词表与学术词表的建立已形成了一定的规模，同时许多具体学科也建成了专业英语词表。然而目前有关采矿专业学术词表方面仍存在空白。该学术词表的建立能够补充我国在专用语料库上的不足，助力研究者日后的论文写作与翻译实践。矿业作为我国国民经济增长的产业之一，矿业为基础的原材料加工业以及相关产业占全国工业生产总产值约 60% 以上。在国际之间商业贸易与科技交流日益增长的趋势下，我国作为矿业大国需要频繁的使用英语与他国进行对话沟通。专业学习者需要不间断的了解学习国际上新出现的矿业知识，或者在国际期刊上发表相关领域的个人研究并与国外同行进行学术交流。

综上所述，本文基于采矿专业英语语料库，建立了采矿工程学术词表。该词表通过提取汇集语料库中高频出现的词汇形成的学术词表，能够帮助学习者提高综合语言水平。其中收录的词汇基于大规模真实语料，适用性高。其中的应用包括教师及教材编写者可以根据词表中的高频词汇，结合实际情况进行有针对性的教学活动与教材编写，应用于学术英语课堂核心高频词的教学工作以及翻译教学中。与此同时，学习者可以结合自身需求学习采矿专业英语学术词表，以便日后期论文的阅读及写作，提高用词的准确性。但仍存在后续语料库的更新后词表需要更新迭代的问题，词表中的高频词汇收录年份为 2018 年~2020 年，随着学界热点研究话题的变化，词表中收录的词汇也需更新。因此在之后的研究中，在完善语料库实时性的同时，应兼顾学术词表的发展。

基金项目

中国矿业大学(北京)大学生创新训练项目资助与“中央高校基本科研业务费专项资金资助”，项目编号：202108009。

参考文献

- [1] 黄宽宏, 王安妮. 同等学力医学硕士研究生专业英文文献阅读现状调查研究[J]. 卫生职业教育, 2020(38): 121-123.
- [2] 朱海青. 研究生学术英语论文写作中的常见问题及对策[J]. 科技风, 2020(18): 57-58.
- [3] Cohen, A.D. (1988) Reading English for Specialized Purposes: Discourse Analysis and the Use of Student Informants. In: Carrell, P., Devine, J. and Eskey, D.E., Eds., *Interactive Approaches to Second Language Reading*, Cambridge University Press, Cambridge, 152-167. <https://doi.org/10.1017/CBO9781139524513.017>
- [4] Coxhead A. (2000) A New Academic Word List. *TESOL Quarterly*, 34, 218-230. <https://doi.org/10.2307/3587951>
- [5] 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.
- [6] 景漱玉, 吕佳忱. 语料库方法下能源英语学术词表的建立[J]. 中外企业家, 2020(5): 226.
- [7] 何玲. 矿业双语平行语料库的建设理论研究[J]. 海外英语, 2016(1): 187-189.
- [8] 丁自华. 专业英语语料库的开发与应用[J]. 航海教育研究, 2011, 28(3): 101-104.
- [9] Lei, L. and Liu, D. (2016) A New Medical Academic Word List: A Corpus-Based Study with Enhanced Methodology. *Journal of English for Academic Purposes*, 22, 42-53. <https://doi.org/10.1016/j.jeap.2016.01.008>
- [10] 辛闻, 王会娟. 矿业英语的文本特征及翻译技巧[J]. 太原城市职业技术学院学报, 2016(9): 201-202.