

矿物加工专业英语学术词表的创建

盛露雨

中国矿业大学, 北京

收稿日期: 2022年2月14日; 录用日期: 2022年3月10日; 发布日期: 2022年3月21日

摘要

目前针对各学科的专门用途语料库日渐丰富,但有关矿物加工专业的学术英语语料库却少之又少。因此,本文旨在通过创建矿物加工专业的学术英语语料库,语料库的库容为1391626形符,并建立该专业英语学术词表,对词表进行分析归类,为该学科翻译及学术英语写作教学提供借鉴。本研究将建立的专业英语学术词表按照频次将单词分为通用英语词汇,科普英语词汇以及专业英语词汇三大类,并发现通用英语词汇与日常英语词汇释义基本相同,而科普英语词汇以及专业英语词汇除了具有日常英语词汇释义以外,还具学术性的特殊含义。因此,本研究得到的词表可以帮助该领域学者进行专业英语学习,除一般单词的掌握外,学者还要特别注意对科普英语单词和专业英语单词特殊含义的理解与掌握。

关键词

矿物加工, 英语语料库, 学术词表

Creating the English Academic Word List of Mineral Processing

Bonnie

China University of Mining & Technology, Beijing

Received: Feb. 14th, 2022; accepted: Mar. 10th, 2022; published: Mar. 21st, 2022

Abstract

Corpus is the basic resource of corpus linguistics research and the main resource of empirical language research method. At present, specialized corpora for various disciplines are increasingly abundant, but academic English corpora related to mineral processing are few. Thus, this study intends to build an academic corpus for this major to create the English for specific purposes. This corpus has a capacity of 1,391,626 tokens. The obtained vocabulary is analyzed and classified, and instructive results are obtained. The results can provide reference for relevant scholars in the teach-

ing of translation and academic English writing on the subject. Based on the established English academic word list, this study divides the words into three categories according to frequency: general English words, popular Science English words and professional English words. By comparing the above classified word lists, it is found that the definitions of general English words and everyday English words (the meaning of words which is used frequently in our daily life) are basically the same. It is worth noting that the definitions of some popular Science English words and professional English words are the same as those of everyday English words, and on this basis, these words also have their own special professional meanings. Therefore, it can help scholars to carry out learning and research, and scholars can carry out planned learning and research according to the frequency and classification results. In the process of learning words, scholars should pay attention to not only the general meaning of words, but also the understanding and mastery of the special meaning of popular Science English words and professional English words.

Keywords

Mineral Processing, English Corpus, Academic Word List

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

学术词表一直以来都是热点话题。词表的研制与开发历史悠久,早在计算机出现之前,各国学者就开始手工创建用于编制工具书和指导语言教学的词频表,如德国学者凯定(F. W. Kaeding) 1898年编写的《德语频率词典》、中国教育家陈鹤琴1928年编制的《语体文应用字汇》以及美国的教育心理学家桑代克(Thorndike) 1932年出版的《教师两万词词汇手册》等[1]。当代最成功、影响最广泛的学术英语词表当属Coxhead (2000) [2]开发的学术单词表(Academic Word List, AWL) [3]。目前,学术词表在各个领域被广泛的应用,例如教育、科技、政治、法律等等。而AWL, VL和ASWL等学术英语词表是目前较受欢迎的词表。AWL, VL和ASWL等学术英语词表子学科领域(共28个),合计350万英文词。该词表采用了如下三条标准来提取学术单词。Gardner & Davies (2014) [4]采用全新方法开发了AVL。AVL基于一百二十多万词的当代美国英语语料库(Corpus of Contemporary American English, COCA)学术文本开发而成。先前开发的诸多学术英语词表均为学术笔语词表,而Dang *et al.* (2017) [5]另辟蹊径,开发了迄今为止第一份学术口语词表(ASWL)。ASWL基于1300万词的学术口语语料库开发而成,该语料库包含讲座(lectures)、研讨会(seminars)、实验室讨论(labs)、导师个别指导(tutorials)等学术口语语体,涵盖四大学科门类(纯硬科学、应用硬科学、纯软科学、应用软科学)和24个子学科领域。ASWL采用了如下标准收录单词[3]。然而,目前有关矿物加工(以下简称“矿加”)学科的学术词表的研究却十分稀少,相关文献也比较少,因此,本研究旨在创建矿加学术词表,为该学科的翻译及学术英语写作教学提供借鉴。

2. 研究方法

2.1. 语料库现状

语料库是语料库语言学研究的基础资源,也是经验主义语言研究方法的主要资源,应用于词典编纂、

语言教学、传统语言研究、自然语言处理中基于统计或实例的研究等方面。随着越来越多语料库可供普通学习者免费使用,语料库在外语教学中的应用也日益广泛[6]。首先,本研究选取矿加学科, *International Journal of Coal Preparation and Utilization*, *Minerals Engineering* 这两种权威期刊,收集其在 2020~2021 年间发表的论文,并进行文本清理,只留下文献的文本内容,最后建立矿加学科的语料库。语料库的库容为 1391626 形符(tokens)。

2.2. 语料库分析

在建立了相关语料库之后,本研究使用 Antconc 语料库分析软件,以 Crown corpus (2009)为对比语料库,对我们所建立的语料库进行了分析。首先,本研究在 Antconc 中导入文本整理后的语料,导入英语停用词表,并在关键词表(keyword list)中导入 crown corpus (2009)为对比语料库,得到了相应的词表(word list)。

3. 结果与讨论

3.1. 结果

给予 Antconc 软件得出的词表,本研究选出了词频在 50 次以上共计 1852 词,而其中出现频率最高的为“coal (煤炭)”(9557 次),最低的为“photography (摄影术)”(1 次)等,如表 1 所示。这类词汇中除了一些矿业加工学科的专业词汇外,还有许多其他学科的词汇,学科的覆盖率很高。通过和 crown corpus (2009)对比语料库进行对比,本研究将词表中的高频词分为三类,分别为通用英语词汇(频次在 1000 以上),科普英语词汇(频次在 500~1000)以及专业英语词汇(频次在 50~500),见表 2。其中,通用英语词汇主要包括“coal”,“particle”,“flotation”等,这类词汇大多用于日常英语,较为常见,学术性较低。科普词汇介于通用英语及专业英语之间,具有一定的专业性,除了与日常英语释义相同的部分以外,还包含了自己特殊的含义;专业英语是该专业中的学术性较高的术语,在日常生活中使用频率很低,专业性较高,具有自己特殊的含义。

Table 1. Some high frequency content words in corpus
表 1. 语料库中的部分高频实词

| 高频实词 | 频次 |
|---------------|------|
| coal | 9557 |
| particle | 6471 |
| flotation | 5721 |
| surface | 5510 |
| mineral | 4308 |
| sample | 4293 |
| size | 3798 |
| process | 3677 |
| concentration | 3324 |

Table 2. High frequency academic English content word classification
表 2. 高频学术英语实词分类

| 通用英语 (频次在 1000 以上) | 科普英语 (频次在 500~1000 之间) | 专业英语 (频次在 500~50 之间) |
|-----------------------|---------------------------|-------------------------|
| coal (9557) | lead (1000) | heat (496) |
| particle (6471) | interaction (964) | quality (494) |
| flotation (5721) | gas (960) | dissolution (492) |
| surface (5510) | structure (955) | degree (483) |
| mineral (4308) | measurement (952) | step (483) |
| sample (4293) | carbon (951) | operation (481) |
| process (3677) | reduction (934) | bond (477) |
| concentration (3324) | grade (930) | settle (474) |
| water (3273) | dosage (906) | component (470) |
| recovery (2971) | mass (891) | chloride (469) |
| ash (2730) | spectrum (891) | pore (468) |
| solution (2711) | plant (887) | approach (467) |
| content (2694) | oil (885) | extract (467) |
| model (2618) | mill (880) | industry (467) |
| temperature (2564) | species (876) | gold (459) |
| analysis (2366) | characteristic (864) | power (457) |
| rate (2349) | type (862) | ray (457) |
| separation (2295) | liquid (857) | technique (457) |
| bubble (2161) | slurry (855) | standard (455) |

笔者发现,在通用英语词汇中,词汇虽然在学术英语语料库中出现频次高,但其使用的意思大都以日常释义为主,例如,“coal”在语料库原索引行中释义以其最常见义“煤”为主(频次为 9557 次),如“the oxidized coal or the low-rank coal”(氧化煤或低阶煤),“In the current work, the screening of moist coal was carried out for various experimental conditions.”(目前的工作中,在各种实验条件下对湿煤进行了筛选。)都以名词的形式出现。在日常用语中,大部分的“coal”都表示名词“煤;煤炭”的含义。只有极少一部分会被用作动词“加煤;把……烧成炭”。而这种使用方式在日常生活中以及矿加领域使用的例子都十分稀少。同样的,“particle”在语料库中出现的释义以“颗粒;粒子”为主(频次为 6471 次),如“the influence of particle bed and size was considered”(颗粒床的影响以及尺寸都被考虑在内)以及“Understanding the effects of these factors on bubble-particle attachment is thus crucial as they may in turn affect the mineral recoveries attained.”(因此,了解这些因素对气泡颗粒附着的影响至关重要,因为它们可能

反过来影响获得的矿物回收率)。而在日常生活, 诸如语言学等其他专业领域中, “particle”除了以上常见的释义, 还有“小品词; 质点”的意思, 例如“particle kinematics”(质点运动学)中, “particle”就表示“质点”的意思。但金国分析, 本研究发现, particle无论在通用词汇还是日常释义中, 都以名词的形式出现, 且含义较为相近, 都作为“颗粒; 粒子”的含义理解。因此通过总结, 笔者发现, 通用英语词汇的含义基本与日常使用词汇的释义接近, 在词性上也没有太大变化, 使用方式几乎相同。因此, 对通用词汇的学习也是对日常英语的学习, 这方便了相关学者的学习研究工作。

而在科普英语中“lead”的原句有“galena was used instead of lead-activated sphalerite”(使用方铅矿来替代铅活化的闪锌矿)以及“The dust that is produced from the EAF often contains hazardous metals such as lead, cadmium, chromium or zinc and these waste products need to be immobilized before being considered safe or benign.”(电弧炉产生的粉尘通常含有铅、镉、铬或锌等有害金属, 这些废物在被视为安全或无害之前需要固定化)。此处“lead”的含义为“铅”, 属于名词(频次为499次)。而在本研究范围内的语料库里, 极大一部分的“lead”被翻译成“铅”。这与日常英语中较常使用的“引导”(动词), 例如“I will lead them to their destiny.”(我将引导他们走向他们的命运。)以及“领先”(名词), 例如“She took the lead in the second lap.”(她在第二圈时领先。)含义相差较大。在日常英语释义中, “铅”的使用频率低, 但在专业语料中, “铅”的使用频率很高, 因此需要特别注意。在专业英语词汇中的“current”, 包含该单词的原句有“it resulted in dendrite growth and partial loss of current”(它导致枝晶生长和部分电流损失), 以及“No limiting current was seen in the range studied.”(在所研究的范围内未发现限制电流。)此处的“current”含义为“电流”, 为名词, 与较常使用的日常释义“如今的, 现在的”相差较大, 词性也不同。例如“a budget for the current year”(今年的预算)。但是, “lead”和“current”在专业语料中的含义除了以上几个特殊的意思意外, 同样也具有一般的含义。拿“lead”作为例子, “lead”也表示动词“导致”(动词)。例如在语料库中的语句“The high dose of water glass in the cleaning flotation circuit will also lead to a difficulty in the settling of the tailings.”(清洗浮选回路中的高剂量水玻璃也会导致尾矿沉降困难)。而语料中的“current”也表示形容词“现在的, 如今的”(形容词), (频次超过500次), 例如语句“Particles in the swarm will be updated based on their own fitnesses (current and historical) and the fitness of the swarm.”(粒子群中的粒子将根据其自身的适合度(当前和历史)以及粒子群的适合度进行更新)。通过对比归类能够发现, 有关以上两个词汇的中文释义, 除了单词具有的特殊的专业的释义外, 其他的释义都与日常释义相同, 在词性上也没有变化。因此, 学者在进行专业词汇和通用词汇的学习时, 不但要掌握此类词汇的日常释义, 还要特别注意每个词汇特殊的含义以及词性的变化, 这样才能更好的进行学习与研究。

3.2. 讨论

3.2.1. 学术词汇与高频词汇

目前针对学术词汇与高频词汇之间关系的讨论有很多, 不能仅凭词汇出现的频率判断该词汇是否为学术词汇。高频词汇可能包含学术词汇, 而学术词汇也可能是高频词汇(Gardner & Davies 2014) [4]。高频词与学术词汇之间并没有明晰的界限, 高频词可能是学术词汇, 学术词汇也可能是高频词, 因此, 没有必要在开发学术词表时排除高频词(Dang *et al.* 2017 [5]; Gardner & Davies 2014 [4]; Lei & Liu 2016 [7]) [3]。

针对以上问题, 本研究在整理有关矿加领域的专业词汇时, 并不将高频词汇也就是通用词汇剔除。相反的, 包括通用词汇, 科普词汇以及专业词汇在内的所有词汇在剔除掉虚词以后, 都被归纳在学术词表内。本研究通过剔除出现频次较低(50次以下)的词汇, 获得矿加英语学术词表, 并附上专业性的含义, 该词表可以极大的提高矿加英语学习者的学习效率。

3.2.2. 词表的使用方式

本词表将按照频次对得到的词汇进行分类：通用词汇，科普词汇以及专业词汇，并附上含义。学者在使用词表时，要特别注意以下几点：

首先，通过对高频实词的归类统计，我们发现大多数通用英语类词汇与日常英语词汇含义相同，并且大多被包含在四级词汇和高考词汇表内，含义较为统一，词性没有很大的变化，因此不是学者学习的重点。

其次，通用英语类词汇和专业英语类词汇则包含了日常生活中较常使用释义外，还拥有特殊的含义，词性的变化也较大。这是学习者在写作及翻译中的难点。因此，学者在使用词表的时候，要有轻重，

最后，相比通用英语词汇，科普词汇及专业英语这两类词汇是学术英语教学中的重中之重，需要学术英语学习者特殊注意，着重掌握。

4. 结论

对矿物加工专业英语学术词的分析证明了矿物加工英语专业词表的建立具有可行性，并且具有很大的发展空间。同时，该词表中高频实词大部分都被包含在我们平时应试需要的词表中，含义也较为统一，通常没有太大的变化，但要特别注意科普性词汇和专业类词汇的特殊含义，进行着重的学习与掌握。学者可以根据本研究得到的分类词表，对词表中出现频率较低的单词进行掌握，依次列举出该类单词的特殊含义和基本含义。其次，我们还要认识到，矿业加工专业英语学术词汇也涵盖了各种其他学科所需要的词汇，这就需要各个学科互相促进发展，使该词表更加完善。这可以帮助学习者提高学习效率，也帮助相关研究者提高学术英语写作效率。

致 谢

本文的顺利完成，离不开我的导师张汝莹的指导，感谢她不辞辛苦，多次询问研究进度，帮我开拓研究思路，悉心校对文章内容。同时，我也要感谢我们大创项目的小组成员们。他们一丝不苟的共同努力建立的语料库。最后，向为本项目付出的所有人献上真挚的感谢。

基金项目

中国矿业大学(北京)大学生创新训练项目资助和中央高校基本科研业务费专项资金资助，大创项目“矿业工程学术英语语料库的创建及应用”项目编号：202108009。

参考文献

- [1] 赵志刚. 专门用途英语学术词表创建研究——以航海英语为例[J]. 重庆交通大学学报(社会科学版), 2015, 15(6): 140-144.
- [2] Coxhead, A. (2000) A New Academic Word List. *TESOL Quarterly*, **34**, 213-238. <https://doi.org/10.2307/3587951>
- [3] 刘迪麟, 雷蕾. 学术词表研究综述[J]. 外语教学, 2020, 41(2): 34-38+50.
- [4] Gardner, D. and Davies, M. (2014) A New Academic Vocabulary List. *Applied Linguistics*, **35**, 305-327. <https://doi.org/10.1093/applin/amt015>
- [5] Dang, T.N.Y., Coxhead, A. and Webb, S. (2017) The Academic Spoken Word List. *Language Learning*, **67**, 959-997. <https://doi.org/10.1111/lang.12253>
- [6] 徐锦芬, 刘文波. 信息技术背景下的外语创新教学与研究[J]. 外语与外语教学, 2019(5): 1-9+147. <https://doi.org/10.13458/j.cnki.flatt.004614>
- [7] Lei, L. and Liu, D.L. (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>