

语料库在科技期刊编辑工作中的应用

翟 振 王久丽 蔡建伟 郭红丽

中国科学院物理研究所《中国物理 B》编辑部,100084,北京

摘 要 介绍语料库在科技期刊编辑工作中的 2 种应用:送审和语言润色。通过自建 2 个小型语料库,利用语料库检索工具帮助科技期刊编辑准确、快速查找审稿人和修改英文,从而提高期刊论文的学术质量和写作质量。

关键词 科技期刊;语料库;审稿人;语言润色

Application of corpus in the editing work of scientific journals

//ZHAI Zhen, WANG Jiuli, CAI Jianwei, GUO Hongli

Abstract This paper shows how the corpus can help the editors of scientific journals in two manners: selecting peer reviewers and language polishing, with the purpose of improving the scientific quality and writing quality.

Keywords scientific journal; corpus; peer reviewer; language polishing

Authors' address Editorial Office of Chinese Physics B, Institute of Physics, Chinese Academy of Sciences, 100190, Beijing, China

DOI:10.16811/j.cnki.1001-4314.2016.05.012

科技期刊是报道、宣传科研成果的载体,其质量主要取决于发表论文的学术质量和写作质量;因此,期刊编辑在实际工作中,一方面必须加强审稿人队伍建设,制订严格的审稿标准和流程,做到准确、及时送审,另一方面,严格按照科技论文的写作要求,提高论文的可读性。毋庸置疑,审稿人是控制稿件学术质量的关键。目前,国内科技期刊为了提升稿件的学术质量,已经在审稿人来源^[1]、送审制度^[2-4]、审稿制度^[5]、审稿人网络数据库建设^[6]等方面积累了一些经验^[1-3],但仍存在一些不足和困难,降低了送审的效率和准确度。另一方面,英文科技期刊作为国际间学术交流的平台,对论文的撰写也提出了很高的要求。科研工作的学术价值会因为语言不通顺,逻辑混乱而大打折扣;但是,我国大部分英文科技期刊的稿件大多来自国内,英文写作是国内作者的先天劣势。此外,优秀稿件外流严重,投在国内的稿件存在明显的写作经验不足的问题。因此,帮助作者改善稿件写作质量是期刊编辑不可推卸的责任。

为了更好地完成送审和英文修改这 2 项重要任务,本文将介绍如何把语料库运用于编辑的日常工作,希望为科技期刊编辑提高工作效率以及期刊更好地发展提供借鉴。

1 语料库简介

语料库(corpus)一般是指经过科学取样和加工存

放在电子计算机上的大规模文本或语音集合。20 世纪 80 年代以来,科技的发展,尤其是计算机和大容量硬盘的普及,促使语料库的应用领域不断扩宽,比如词典编纂、教材编写、英语教学、语音识别^[7-8],等等。语料库之所以具有广阔的应用前景,主要得益于容易建设和维护及资源高度共享。有些大型语料库都支持免费下载或免费在线使用,比如北京大学中国语言学研究(CCL)的古现代汉语语料库,由 Princeton 大学的心理学家、语言学家和计算机工程师联合设计的 Wordnet。目前,语料库已经不仅仅是一种收集和获取数据的工具,而且是作为语言学的一个分支,在国内外迅速发展。

随着信息时代的来临,人们越来越容易获取自己想要的信息。网络资源就是一个巨大的语料库,通过百度等检索工具,个人完全可以建立属于自己的小型语料库,从而实现特定目标,解决特定问题。语料库在科技期刊编辑日常工作中的潜在作用同样值得我们探索。

2 语料库在稿件送审中的应用

同行评议(peer review)是目前国际上科技期刊通用的稿件评价制度,是决定稿件录用与否的关键,保证学术质量的重要环节^[9]。随着学科的不断细化和交叉,准确、快速查找审稿人的难度也越来越大。科技期刊编辑积极应对这一挑战,在建立专家库、送审、缩短审稿时间等方面提出了有效的改进措施;然而,无论如何改变,送审工作仍然没有脱离传统的基于采编系统的信息录入。这里,笔者首先分析基于采编系统的专家库的特点,然后介绍如何建立专家信息语料库,并应用于送审工作。

2.1 采编系统专家库的特点 目前,国内大多数科技期刊都拥有自己独立的采编系统^[10-11],建立了自己的审稿专家库。专家信息的采集和维护主要依靠编辑长期送审过程中的积累,是一项巨大的工程;因为专家信息越全面,越有利于编辑做出及时准确的判断。除了姓名、性别、单位、职称、研究方向等基本信息外,教育背景、工作经历、发表文章、承担项目、担任职务等,都对送审有重要的参考价值。这就要求编辑从专家的个人网页将以上信息逐条复制到专家库相应位置。由于专家信息格式固定,编辑不可能录入所有信息。此外,还要对研究方向进行提炼和精简,从而节省审稿库容

量。除了编辑日常添加和维护专家库之外,投稿时以作者身份注册的专家和他们推荐的审稿人也会进入专家库。有时已经入库的审稿人投稿时,忘记用户名和密码时,以及单位变动时都会再次注册;因此会导致同一审稿人多个用户名和邮箱的情况。如果编辑不及时和经常维护,势必会造成审稿库混乱,影响送审效率。综上所述,基于采编系统的专家库信息采集和维护都有一定的局限性。

2.2 专家信息语料库的建立 简而言之,专家语料库的建立是将专家个人网页以文本形式分类保存在个人电脑的过程。根据个人需要,分类方式可以是专家单位或学科方向,主要用来避免同名专家的混淆。以专家单位为例,我们先在个人电脑新建以专家单位为名的文件夹,然后逐次登录单位主页,将相关科研人员的个人网页以文本格式保存到新建的文件夹内,最后删掉文本文件中多余的内容。

不难看出,与采编系统专家库相比,专家语料库具有独特的优势:第一,不需要逐条录入专家信息,更简单高效;第二,没有固定格式的限制,信息保存更完整;第三,信息全部由编辑采集,不容易造成混乱,因此更容易维护。

2.3 利用 DirFind 软件查找审稿人 语料库检索软件可选择性比较大,DirFind 是一款非常适合用来检索文件名和文件内容的工具。它不仅支持 txt 纯文本文件,还支持 htm 文件,同时支持中英文检索,还可以定义文件名,也可以定义文件内容。如果一篇稿件的作者单位是华中科技大学,其中一个关键词为“超导”,检索方式如图 1 所示。我们打开窗口下方显示的检索结果,可以看到的内容和专家个人网页一样详细。接下来,我们就可以甄选专家,然后通过采编系统进行送审。最后,如果选出的专家还没有在采编系统注册,我们就打开语料库中的文本文件,将必要信息录入采编系统专家库,以备后用。



图 1 DirFind 软件检索审稿人界面

3 语料库在语言润色中的应用

3.1 提高英文写作质量的方法 英文写作质量是评判科技期刊整体水平的一把重要标尺。国内英文期刊要提高国际化程度,必须对写作质量进行严格把关,让论文的学术思想在国际科研领域无障碍传播。为此,编辑同人已经做了大量工作来提高论文写作质量,大致包括以下 3 个方面。

一是提高对作者的要求^[12-13]。初审严格控制,对语言不合格的稿件予以退修,甚至退稿;语言润色过程中与作者良好沟通。

二是提高编辑自身的写作水平^[14-15]。加强编辑培训和自我学习,丰富科技论文写作技巧和英语语法知识。

三是聘请英语为母语的人员帮忙。

笔者认为,以上 3 个方面相辅相成,应以编辑自身建设为主,作者和“外援”为辅。首先,将所有稿件让作者退修不切实际,而且作者短期内修改回来的稿件难保质量有很大提升。其次,聘请英语为母语的人员费用太高。

期刊编辑的自我学习可以是日常工作中一点一滴的积累,比如研读科技论文写作教材和英语语法书。俗话说,近朱者赤,近墨者黑。提高英文写作水平还有一个非常重要的方法,那就是大量阅读国际知名期刊发表的科研论文。然而,期刊编辑忙于日常工作,很难抽出时间学习写作教材和阅读文献。为了更充分利用这些资源,笔者提出建立英文文献语料库,利用软件检索的方法,帮助编辑语言润色。

3.2 英文文献语料库的建立 根据不同期刊的定位和需要,语料的选取来源有所不同,一般遵循 3 项准则:一是来自国际著名期刊;二是稿件第一作者的母语是英语,写作质量较高;三是文献研究方向尽量宽,但不超出期刊报道的范畴。有鉴于此,符合要求的语料可以从期刊发表文章的参考文献中选取。国际期刊文献可获取的格式一般是 PDF 和 HTML。为了保证语料库可检索,需要将 HTML 文件中的文字部分复制到 txt 文本文件中。如果选取的文献只有 PDF 格式,仍可采用类似的办法,但比 HTML 麻烦得多。最后,所有的文件都保存在个人电脑的同文件夹内。

3.3 利用 AdTAT 软件检索 笔者用来检索上述语料库的软件是 AdTAT,是一款由阿德莱德大学开发的非常实用的语料库检索软件(concordancing software)。首先,它可以将多个文本文件合并为一个 cps 文件,因此每次检索只需调用 cps 文件;其次,它有基本检索、单词关联检索、词频检索 3 种模式。以论文中出现频次较

高的单词 increase 为例,它的基本用法可以通过图 2 的检索结果一目了然。increase 有名词和动词双重词性,作为名词的固定搭配有 3 种,即“with (tenfold) increase in”“increase of more than one order of magnitude”“increase by two orders of magnitude”。不难看出,这种检索方式的优势在于快速,专业,全面,容易记忆。

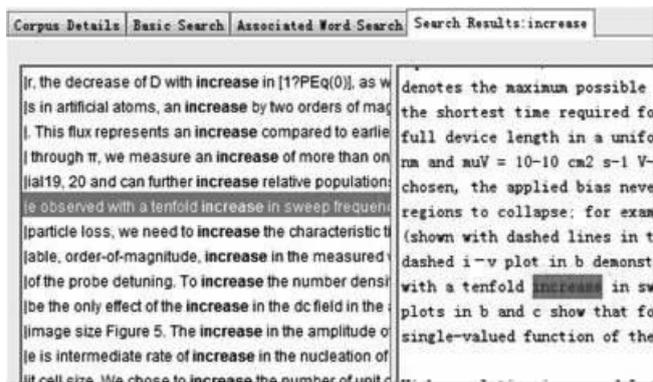


图 2 AdTAT 软件的单词检索界面

4 结束语

在计算机和网络技术高速发展的今天,语料库在诸多领域发挥着不可忽视的作用。本文初步探索了语料库在科技期刊编辑工作中的 2 种应用。一是运用于专家库的建立,帮助编辑更快、更准确实现稿件送审;二是运用于稿件语言润色。两者并驾齐驱,对编辑提高期刊的学术质量和写作质量发挥一定的辅助作用。由于语料库存放在个人电脑里,具有方便移动和共享的优势,其价值会得到更大的利用。

5 参考文献

[1] 李金丽,徐永吉,林清华. 利用中国教育网选择合适的审

稿专家[J]. 编辑学报,2006,18(2):147

- [2] 王俊丽,郭焕芳,郑爱莲. 英文科技期刊遴选审稿专家的途径与原则:以《中国化学快报》为例[J]. 中国科技期刊研究,2015,26(4):351
- [3] 黄劲松,彭超群,杨兵. 审稿专家的选择与管理[J]. 编辑学报,2013,25(1):55
- [4] 任胜利,王久丽. 同行评议中审稿人遴选方式对审稿结果的影响:以《自然科学进展》为例[J]. 中国科技期刊研究,2006,17(5):722
- [5] 汪挺,阮星星,陈小宇,等. 开放审稿的必要性及可行性探讨[J]. 编辑学报,2014,26(1):19
- [6] 王霞,贾贤,蒋伟,等. 科技期刊如何完善审稿专家库[J]. 编辑学报,2011,23(增刊1):43
- [7] 许葵花,张卫平. 论语料库语言学在外语教学中的应用[J]. 外语与外语教学,2003(4):21
- [8] 桂诗春. 语料库语言学与外语教学:语料库语言学的发展前景与资源共享[J]. 现代外语,2010(4):419
- [9] 吴锦雅. 同行评议面临的问题与可行性措施[J]. 编辑学报,2011,23(3):238
- [10] 杨琦,刘晓莉. 国内外期刊稿件采编系统功能特点[J]. 长江大学学报(社会科学版),2015,17(2):99
- [11] 黄莘,漆蓉,税红,等. 稿件采编系统的操作分析[J]. 编辑学报,2008,20(2):153
- [12] 郭红明. 论稿件退修过程中编辑与作者的互动[J]. 编辑学报,2015,27(1):21
- [13] 王桂珍,颜志森. 作者自校清样有利于提高期刊编校质量[J]. 中国科技期刊研究,2005,16(5):716
- [14] 葛建平,范真真,李明敏,等. 关于国内英文科技期刊语言国际化的思考[J]. 编辑学报,2013,25(1):48
- [15] 吕盈盈,张欣. 英文版科技期刊编辑人员素质培养[J]. 中国科技期刊研究,2010,21(2):230

(2016-04-01 收稿;2016-06-06 修回)

敬告《编辑学报》的投稿者

1)《编辑学报》(以下称本刊)是定位于科技期刊学术研究的学术期刊,发表有关科技期刊的编辑出版理论与实践问题研究的文章。凡跟科技期刊的编辑出版、经营管理、改革发展等关联性不强的论文,请改投他刊。

2)本刊暂未设立网站,稿件请投至 bjxb_bj@163.com,目前存在的多个本刊官网均为假冒,注意切勿上当受骗。

3)投稿前认真阅读《〈编辑学报〉论文格式简介》(见本刊 2016 年第 4 期第 408 页),务请按其规范撰

写论文。

4)本刊不收取稿件审理费;对经严格审稿后决定利用的稿件,按 500 元/面收取发表费。

5)为杜绝作者署名不正之风,本刊不赞成滥标通信作者,尤其反对标注并列第一作者,也不支持在校样上添加新作者、通信作者,改变署名顺序等行为。

6)本刊反对一稿多投、一稿多发,坚决抵制形形色色的学术不端、学术不当,吁请投稿者与我们一起维护科技期刊界弊绝风清的学术氛围。