

科技期刊英文摘要学术词汇的语步特点研究^{*}

曹 雁 牟爱鹏

(鲁东大学 烟台 264005)

提 要:本研究基于科技期刊英文摘要语料库考察学术词汇的语步分布和使用特点。研究显示,学术词汇在英语摘要各语步的覆盖率均超过 10%,使用频率较高。这说明,学术词汇的学习在摘要写作中不可或缺。同时,这种词汇在不同语步具有不同搭配、类联接和语义韵特点。部分高频词族有明显语步倾向性,具有特殊语用功能。分析学术词汇的语步特点,为从更深层次掌握英文摘要写作规律提供有价值的参考。

关键词:科技论文摘要;语料库;学术词汇;语步

中图分类号:H313.2

文献标识码:A

文章编号:1000-0100(2011)03-0046-4

The Characteristics of Academic Words Across Different Abstract Moves of English Scientific and Technical Journals

Cao Yan Mu Ai-peng

(Ludong University, Yantai 264005, China)

In this study we investigate the coverage and use of academic words across different abstract moves of scientific journals on the basis of English abstract corpus. It is found that academic words account for over 10% of each move as well as the whole corpus. This indicates academic words are indispensable part of abstract writing. Some academic words are characterized by their different collocation, colligation and semantic prosody in different moves. Certain frequently used academic words show a strong tendency to appear only in one or two moves. The study provides valuable information for scientific journals' abstract writing.

Key words: abstract of scientific and technical journal; corpus; academic words; move

1 引言

学术词汇(academic words)指在普通文本中不经常使用,却频繁出现在各类学术文本里的半专业词汇(sub-technical words)。这些词汇虽不能传递主要学术信息,但作者在表达学术观点、阐述研究活动或解释研究结果时经常需要它们(Farrell 1990)。Coxhead(2000)的学术英语词汇表(academic word list, AWL)近年来备受关注,不少研究都以 AWL 为参照,考察和分析医学、应用语言学、农业、理工等不同专业领域学术词汇的频率、分布或适用性(Chen & Ge 2007, Vongpumivitch *et al.* 2008, Martínez *et al.* 2009)。

学术词汇的掌握情况是衡量学习者能否顺利完成学术英语学习的一个重要指标(Kuhen 1996)。研究生的学

术英语阅读约有 24% 的困难源于学生对学术词汇的掌握不充分。Granger(2006)研究发现学生并不是不使用这些高频词,而是在词汇语法形式上,尤其是主动语态和被动语态的转换上与本族语者有很大差距。因此,笔者认为仅仅制定学术词汇表,对于学术词汇的掌握远远不够,有必要深入研究其用法。本研究对学术词汇在科技期刊英文摘要中的覆盖率和语步分布进行统计分析,着重考察摘要中学术词汇的语步使用特点,旨在为摘要写作和相关的编辑工作提供有价值的参考和依据。

2 科技期刊英语摘要语料库

本研究采用的语料库是科技期刊英文摘要语料库,

^{*} 本文系国家社科基金项目“基于语料库的科技期刊英文摘要的研究与应用”(07BSW012)、教育部人文社会科学青年基金项目“中外科技期刊英文摘要语体的多维对比研究”(10YJC740007)和中国外语教育基金项目“中外科技期刊英文摘要语体的多维对比研究”(WH2010005)的阶段性成果。

简称 IESJAC。语料来自国家科技图书文献中心提供的科技类国际核心期刊,具有较强的学术影响力。语料库涵盖数学、力学、物理学、化学等 12 个理科学科,共 5246 篇摘要,全部为英语本族语者所写,共计 91 万词。

我们基于 Swales 的 4 语步模式(引言、方法、结果、结论)(Swales 1990)结合英文摘要的语篇特点,确立包含 4 语步和十几个次语步的英文摘要语步体系和标注代码。在此基础上,对 960 篇摘要约 18 万词的子语料库进行手工语步划分和标注,为纵深挖掘英文摘要语言特点打下基础。

3 研究工具和方法

本研究提取学术词汇的频率和分布信息所使用的工具是 Range 词汇分析软件。Range 的词汇分析基于其内置的三个基础词表。词表₁和词表₂包含英语最常用的前 2000 个词族(general service list,简称 GSL),词表₃为 AWL 词表,含 570 个词族(word family)。利用 Range 可以统计出 AWL 词汇在科技摘要语料库中的覆盖率和分布。我们参考 Coxhead 学术词汇的选词标准(Coxhead 2000),把本研究的学术词汇规定为在整个语料库中出现至少 20 次,并且在每一学科至少出现一次。

4 结果与讨论

4.1 IESJAC 学术词汇的覆盖率

统计显示,IESJAC 语料库中共出现 564 个学术词族,占整个语料库的 11.8%,这一结果高于 Coxhead 人文子语料库的 9.3%(Coxhead 2000)。其中,前 198 个学术词族覆盖整个语料库的 8.8%。较高的覆盖率说明高频学术词汇对于摘要写作具有重要意义。同时,学术词汇在各个语步(引言、方法、结果、结论)中的覆盖率比较接近,均超过 10%(见表₁)。可见,学术词汇在摘要各语步的写作中都不可或缺。

表₁ 学术词族在整个 IESJAC 语料库和各语步的覆盖率

	IESJAC	引言	方法	结果	结论
Academic words	108078	12048	3783	5270	1128
Total	916020	98034	30693	46361	9516
Coverage	11.80%	12.29%	12.33%	11.37%	11.85%

4.2 IESJAC 学术词汇的语步分布特点

表₂分别列出整个语料库以及每个语步中最常用的前 20 个词族,共记 37 个词族。进一步分析这些词族发现,尽管学术词族在各语步所占比例相当,但其分布和排序有较大差异:有些词族只在其中一个或两个语步中高频出现;有些词族虽然在各语步均有分布,但使用频率有较大差异。

表₂ 各语步中的高频学术词族(前 20 个)

排序	IESJAC	引言	方法	结果	结论
1	analyse	data	analyse	analyse	indicate
2	data	method	method	region	vary
3	method	analyse	data	vary	significant
4	structure	structure	vary	significant	structure
5	vary	vary	estimate	data	predict
6	function	image	technique	indicate	method
7	significant	process	structure	structure	analyse
8	region	theory	function	estimate	data
9	process	function	approach	consist	similar
10	indicate	estimate	equate	range	error
11	range	investigate	distribute	distribute	conclude
12	distribute	region	image	phase	occur
13	respond	parameter	simulate	similar	process
14	theory	distribute	phase	predict	site
15	investigate	approach	investigate	occur	contribute
16	energy	equate	obtain	stable	complex
17	predict	source	region	obtain	phase
18	estimate	range	theory	error	demonstrate
19	react	derive	energy	source	region
20	occur	technique	derive	method	consist

在这些词族中,analyse、data、method、structure、vary 和 region 等词族出现在每个语步前 20 位之列。有 28 个词族仅在其中的一个或两个语步中出现次数较多,表现出明显的语步倾向性。

表₃ 37 个核心学术词族分布的语步数目

词族	语步数目	词族	语步数目
analyse	4	occur	2
data	4	predict	2
method	4	process	2
region	4	range	2
structure	4	significant	2
vary	4	similar	2
distribute	3	source	2
estimate	3	technique	2
phase	3	theory	2
approach	2	complex	1
consist	2	conclude	1
derive	2	contribute	1
equate	2	demonstrate	1
error	2	energy	1
function	2	parameter	1
image	2	simulate	1
indicate	2	site	1
investigate	2	stable	1
obtain	2		

4.21 各语步共享的高频学术词汇

统计显示,有 9 个词族在三个以上的语步中均高频出现。这些词族的广泛分布揭示科技论文和摘要写作的共性特征,但同时部分词族的用法存在语步差异性。

(1) 高分布词族的名词化特征:高分布词族具有较明显的名词化特征。Analyse 词族是科技论文摘要使用频率最高的词族,其词形变化包括名词、动词、形容词和副词。统计显示,名词 analysis 的使用频率最高,尤其是在结果语步,名词的使用频率高达 83.3%。Vary 和 distribute 词族中的名词比例也分别占词族的 78% 和 89%,远超其动词形式。名词化不仅在语篇的信息结构中举足轻重,而且也和其他语法或词汇衔接手段一样,对语篇的连贯、流畅起着重要的衔接作用。恰当使用名词化结构体现科技论文的客观性和严肃性。Data 和 method 两个词族的高频使用和广泛分布揭示基于数据和具体方法的定量研究在科技论文中的中心地位,并作为主轴贯穿整个摘要写作。Structure, region, phase 等名词的高频使用在一定程度上表明当今科技论文的研究主题和研究方向。

(2) 学术词汇搭配、类联接和功能的语步差异性:我们比较 method, data, analysis 等高频词在几个不同语步的搭配、类联接和功能差异。在方法语步, method 一词有许多较显著的动词搭配词,如 have, is, propose, use, develop, require, provide 等。常用的类联接主要是“名词 + 动词 + method”或“method + 动词”,这些句式通常出现在句子的前半部分,作为描述研究方法的起点。对比之下,结论语步的类联接更加多样化,主要可以归纳为三种:(1)“名词 + 介词 + the method”,其中名词多是表示事物属性的抽象名词,如 validity, accuracy, improvement 等;(2)“形容词 + method”;(3)“method + 动词(或系表结构)”。这些表达方式在功能上主要用于评价和界定研究方法,有时与 can, will, would 等情态动词连用,预示研究方法对未来产生的影响。在语义上, method 常与 effective, feasible, promising, rational, new, powerful, attractive, reliable, popular, useful 等具有积极语义韵的词连用,通常对研究方法实施正面、肯定评价。

Data 在方法语步较显著的动词搭配词有 use, obtain, derive, collect, compare, base, test 等,主要出现在“动词 + data”类联接中,描述数据的来源或收集数据的方法。而在结论部分,具有显著意义的动词搭配词有 suggest, indicate, demonstrate, show, obtain, support 等,主要类联接是“data + 动词”,描述由研究数据得出的结论或总结研究结果;此外, data 常出现在“data + be + 形容词(介词短语)”类联接中,常见搭配词如 consistent, suitable, capable, ambiguous, available, in good agreement 等,对数据作出结论或与先前的数据加以对比。

Analysis 在方法语步常与 use, carry out, perform 等词的

被动语态搭配使用,典型类联接是“analysis + be + 过去分词”;有时也使用“动词 + analysis”结构。这两个结构主要陈述研究中的分析方法。Analysis 在结果语步的主要结构为“analysis + 动词”,常用搭配词有 reveal, show, demonstrate, indicate 等,主要目的是揭示分析和讨论研究结果。

我们进一步考察 analysis 在不同语步典型类联接的检索行。Analysis 在引言语步的常用句式主要用于明确研究目的和研究对象, analysis 的前置定语或介词后置定语界定研究范围和对象。

引言语步: this paper(we) + 动词 + a/an + 形容词 + 名词

1. this paper present an elasto-plastic analysis
2. this paper present a detailed analysis
3. this work extend a fractal analysis
4. this paper present an elasto-plastic analysis
5. we present a critical analysis
6. we present a detailed analysis

方法语步的 analysis 常接被动语态,后面的不定式动词有 use, base, extend, obtain, examine, investigate 等,说明具体研究方法。

方法语步: analysis + be + 过去分词 + to do

1. analysis be conducted to examine
2. analysis be shown to improve
3. analysis be used to demonstrate
4. analysis be used to extract
5. analysis be performed to find

结果语步的 analysis 主要和 indicate, show 等引述动词一起引出研究结果。

1. analysis indicate
2. analysis show
3. analysis suggest
4. analysis reveal

由于这些高频词在不同语步承载不同功能,其用法存在明显差异。因此,学习高频学术词汇,仅仅是记住远远不够,还须要从单个词汇的学习转变为在具体语境中的短语和惯用法的学习,掌握摘要各个语步的典型类联接和搭配,才能使表达更地道、更准确。

4.22 带有语步倾向性的高频学术词汇

与上述高分布的学术词族相反,有些词族仅在其中的一个或两个语步中高频出现。这说明这些词应是某一个或几个语步专用的表达方式,具有特殊语用功能。我们结合摘要的 4 个基本语步,对它们的功能进行归类。

(1) 表明目的:引言语步主要说明研究对象和范围,指出研究目的和研究在该学科中的位置,必然涉及相关理论基础、研究方法、研究过程,因此这一部分词汇以与

研究相关的概念类名词居多,如 image, process, theory, function, parameter, approach, equation, source, range, technique 等;还有少量构建陈述句的动词,如 investigate, derive 等,用于说明研究目的。

(2) 说明方法:方法语步的主要作用是说明具体方法和实验过程、使用的材料、设备及统计方法等。因此,其典型的词族与引言语步相似:以名词为主,如 technique, function, approach, equation, image, theory, energy 等。同时,部分行为动词在方法描述中起着不可或缺的作用,如 derive, simulate, investigate, obtain 等。方法语步的前 20 个高频词中没有形容词,这与该部分侧重研究方法的客观表述目标一致。

(3) 描述结果:结果语步重在描述研究结果,所以这部分出现不少描述性形容词,如 stable, significant, consistent, similar 等,用以比较研究结果或描述结果的变化、程度等;同时,还有一些构建结果类陈述句必备的动词,如 indicate, predict, occur, obtain 等。

(4) 陈述结论:结论语步主要概括性地总结从该项研究结果中取得的正确观点、理论意义或实用价值等,因此构建总结性语句的引述动词在这一部分扮演重要角色,如 conclude, demonstrate, indicate, predict, occur, contribute 等;同时,描述性形容词出现在这一部分,如 significant, similar, complex, consistent 等。

5 结论

本研究着重考察 IESJAC 语料库中学术词汇的语步分布和使用特点。研究发现,学术词汇在科技论文摘要中使用频率较高,在各语步的分布均超过 10%。这说明学术词汇对摘要各语步的写作具有不可或缺的作用。其中,有近 200 个词族出现频率较高,学科分布面广,应当列为英文摘要写作的重要词汇。同时,这些学术词族都有各自特定的搭配特点、典型的类联接形式,它们选择性地和某一类词搭配使用。部分高频词族的语步倾向性明显,具有特殊的语用功能,这些功能与各语步的写作目标基本对应。因此,有必要把搭配、类联接和语义韵概念引入英文摘要写作,收集来自语料库的语言证据,学会分析、概括所学词项的深层特点,增强语法和语义选择趋向

意识,从而提高灵活运用词汇的能力。学术词汇的特点分析为从更深层次掌握摘要写作规律提供有益参考。

当然,本研究仅局限于学术词汇的研究,在剩余的 24% 词汇中,仍有部分专业词汇在英文摘要中频率高、分布广,也应当成为词汇学习的重点。

参考文献

- Chen, Q. & G. Ge. A Corpus-based Lexical Study on Frequency and Distribution of Coxhead's AWL Word Families in Medical Research Articles [J]. *English for Specific Purposes*, 2007 (26).
- Coxhead, A. A New Academic Word List [J]. *TESOL Quarterly*, 2000(2).
- Farrell, P. *A Lexical Analysis of the English of Electronics and a Study of Semi-technical Vocabulary* [M]. Dublin: Trinity College, 1990.
- Granger, S. *Lexico-grammatical Patterns of EAP Verbs: How Do Learners Cope?* [C]. Paper Presented at the Exploring the Lexis-Grammar Interface Conference, Hanover (Germany), 2006.
- Kuhen, P. *Assessment of Academic Literacy Skills: Preparing Minority and Limited English Proficient (LEP) Students for Post-secondary Education* [M]. Fresno, CA: California State University, Fresno, 1996.
- Martínez, I. A., S. C. Beck. & B. Carolina. Academic Vocabulary in Agriculture Research Articles: A Corpus-based Study [J]. *English for Specific Purposes*, 2009 (3).
- Swales, J. M. *Genre Analysis: English in Academic and Research Settings* [M]. Cambridge: Cambridge University Press, 1990.
- Vongpumivitch, V., J. Huang. & Y. Chang. Frequency Analysis of the Words in the Academic Word List (AWL) and Non-AWL Content Words in Applied Linguistics Research Papers [J]. *English for Specific Purposes*, 2009 (1).

收稿日期:2010-05-12

【责任编辑 郑丹】