

双一流建设背景下 *NATURE* 摘要语篇内部差异的多维度分析

高江珊

中北大学, 山西 太原 030051

摘 要 : 在当前我国全面迈向高质量发展的新阶段, 面对着建设社会主义现代化强国的宏伟目标和全球百年难遇的重大变革, 对高素质人才的渴求在国家和社会各界达到了前所未有的高度。优秀学术论文的发表能引领、示范、激励广大研究生开展原创性研究, 发现和培育高层次创新人才。本研究将在双一流培养高层次人才的背景下, 基于自建 *NATURE* 学术论文摘要语料库, 探索 *NATURE* 论文摘要语篇的内部差异, 明确 *NATURE* 摘要写作特点及语篇特征, 为英语学习者今后的学术论文摘要写作提供一定的参考和指导。

关 键 词 : 多维度分析; *NATURE*; 摘要; 双一流建设

Multidimensional Analysis of Internal Differences in *NATURE* Abstract Discourse under the Background of Double First Class Construction

Gao Jiangshan

North China University, Taiyuan, Shanxi 030051

Abstract : In the current new stage of comprehensive high-quality development in China, facing the grand goal of building a socialist modernized strong country and the major changes that are rare in the world for a hundred years, the thirst for high-quality talents has reached an unprecedented height in the country and all sectors of society. The publication of excellent academic papers can lead, demonstrate, and motivate graduate students to conduct original research, discover and cultivate high-level innovative talents. This study will explore the internal differences in the discourse of *NATURE* academic paper abstracts based on the self built *NATURE* academic paper abstract corpus, in the context of cultivating high-level talents through the Double First Class initiative. It will clarify the writing characteristics and discourse features of *NATURE* abstracts, and provide some reference and guidance for English learners' future academic paper abstract writing.

Keywords : multidimensional analysis; *NATURE*; abstract; double first class construction

一、研究背景

在双一流建设的推动下, 评价高层次人才学术成就的标准既包括他们发表的论文数量和质量, 也包括他们的研究成果对同行产生的影响。简而言之, 学者们需要在权威期刊上发表更多论文, 并期待自己的论文能够获得广泛的阅读和引用。为了实现这一目标, 他们不可避免地要面对用英语发表论文的挑战。论文摘要是学术论文中不可或缺的重要组成部分, 能够起到吸引读者并为读者提供精确信息的作用, 在摘要写作中, 分析出语篇内部差异, 有助于了解期刊的写作风格。对理工科写作者来说, 掌握 *NATURE* 摘要写作特征, 可以完善提高摘要写作能力, 提高发表论文的机率。高质量的英文摘要有助于提高期刊的总体质量, 从而提升期刊的国际影响力。综上所述, 本研究将基于自建2018年-2023年 *NATURE* 学术论文摘要语料库, 探索 *NATURE* 论文摘要语篇的内部差异, 明确 *NATURE* 摘要写作特点及语篇特征, 为英语学习者今后的学术论文摘要写作提供一定的参考和指导。^[1]

二、文献综述

摘要又称概要、内容提要。基本要素有研究目的、方法、结果和结论。摘要, 作为论文的“门面”, 其质量是国外知名检索数据库决定收录与否的关键因素; 同样, 读者也会依据中文摘要提供的内容来决定是否深入阅读或引用该论文^[2]。近年来, 学术论文摘要作为一篇独立的文本, 已经引起了国内外众多学者的极大关注。在过去几十年间对学术摘要部分研究较为丰富, 语类分析的学术摘要研究占大多数, 主要包含对引言、方法、结果和讨论这四大语步的分析研究^[3]。在语篇结构领域, Swales 提出的语步分析框架、英文研究论文的 IMRD 结构 (引言-方法-结果-讨论) 以及引言部分的 CARS 模型 (Create a Research Space) 是最具影响力的理论^[3-4]。多维度分析法 (MDA), 亦称作多维度多特征分析法 (MD/MF Analysis), 是由美国语料库语言学家 Douglas Biber 所倡导的一种量化研究手段。这种分析技术涵盖了七个不同的分析维度, 每个维度都旨在揭示文本的功能特性:

第一维度聚焦于交互性和信息性表达，第二维度探讨叙述性与非叙述性关切，第三维度分析指称明晰性与情景依赖型指称，第四维度关注显性劝说型表述，第五维度涉及信息的抽象程度与具体性，第六维度考察即席信息组织精细度，第七维度则探讨学术模糊性表达。

Conrad & Biber (2001) 指出，在多维分析方法中，维度6与维度7出现的语言特点较为稀缺，多数研究都不会研究这两个维度。因此，本研究仅考察 Biber (1988) 多维分析方法中的前5个维度。已有多项研究采用了多维分析方法，分别对学科之间的变异规律^[6]、二语学习者^[6-7]、跨语言及跨文化差异^[8]、以及非学术语域^[9]进行了广泛的考察，为研究不同学科和不同语言的语域变异提供了科学的参考标准。

在2014年，中国学者肖忠华和曹雁在对比中外作者科技论文英文摘要的研究中，运用了163项语言特征进行分析，这些特征覆盖了语法、语义以及类联接形式等方面^[10]。闫鹏飞^[11]在其博士论文中研究了材料科学论文语篇内部的差异。多维分析模型最初是为了探究口语和书面语在语体上的差异而开发的，现在这一模型已经被广泛地应用于各种语域的对比研究中。然而，在中国，这种模型的应用大多还仅限于研究口语与书面语体、专业领域语体以及学习者英语等方面，鲜少有关于具体到一个期刊的写作风格的研究，所以探索 *NATURE* 摘要语篇内部差异的研究就显得尤为重要。本研究将自建 *NATURE* 摘要语料库，多维分析其摘要语篇内部差异，从而使得中国的外语学习者可以了解并掌握 *NATURE* 摘要的写作特征。

三、研究设计

本研究的语料来源于 NATURE 2018年-2023年每年的前30篇摘要。语料首先经过语料清洁软件 Text Editor 的处理，之后统一人工赋码将清洁语料分为引言 (INTRO)、方法 (METHODS)、结果 (RESULTS) 和讨论 (DISCUSSION) 四语步的语步语料。四语步词数分别为 11565, 3731, 16924, 4633, 共计 36853 词。在本研究中，我们采用了 Nini 于 2015 年开发的多维标注与分析工具 MAT 1.3 (Multidimensional Analysis Tagger1.3)，以便对文本进行自动化的标注、特征提取和数据统计。该工具基于 Biber 在 1988 年提出的 67 个语言特征和六个主要的功能维度进行文本分析，并集成了斯坦福词性赋码器 (Stanford POS Tagger) 来自动进行词性标注。此外，本研究还使用了 SPSS19.0 统计分析软件，对 MAT 工具生成的报告中的维度和 Z 值进行了统计显著性检验。

四、研究结果

在本节中，在 MAT 报告了一些基本统计数据后，使用 SPSS 对五个维度上的一般特征数据进行处理，并讨论四个语料库之间的一般特征差异。

(一) 四个语步语料库的显著差异

MAT 处理后的数据形成标准化分数后，SPSS 将进行单因素方差分析。SPSS 计算的 p 值显示了四个语料库在每个维度上的关系。

表1 四个语料库之间的五个维度 p 值分数表

维度 \ p 值	引言. vs. 方法.	引言. vs. 结果	引言. vs. 讨论.	方法. vs. 结果.	方法. vs. 讨论.	结果. vs. 讨论.
维度1	0.007 < 0.05	0.442 > 0.05	0.982 > 0.05	0.141 > 0.05	0.069 > 0.05	0.853 > 0.05
维度2	0.004 < 0.05	0.777 > 0.05	0.000 < 0.05	0.029 < 0.05	0.998 > 0.05	0.002 < 0.05
维度3	0.003 < 0.05	1.000 > 0.05	0.340 > 0.05	0.003 < 0.05	0.465 > 0.05	0.307 > 0.05
维度4	0.553 > 0.05	0.007 < 0.05	0.911 > 0.05	0.626 > 0.05	0.914 > 0.05	0.139 > 0.05
维度5	0.000 < 0.05	0.263 > 0.05	0.000 < 0.05	0.000 < 0.05	0.814 > 0.05	0.000 < 0.05

由表1可知，*NATURE* 摘要语料库中4个语步在五个维度上均存在显著性差异 ($p < 0.05$)，且维度2和5存在较多的显著性差异。维度一仅有引言语料库与方法语料库之间存在显著性差异 ($p=0.007 < 0.05$)，维度二除引言语料库与结果语料库，方法语料库与讨论语料库无显著性差异外其余语料库间均存在显著性差异 ($p=0.004 < 0.05$, $p=0.000 < 0.05$, $0.029 < 0.05$, $0.002 < 0.05$)，维度三存在两组显著性差异，分别为引言语料库与方法语料库 ($p=0.003 < 0.05$)、方法语料库与结果语料库 ($p=0.003 < 0.05$)。维度四中的引言语料库与结果语料库存在显著性差异 ($p=0.007 < 0.05$)。维度五存在四组显著性差异，分别为引言语料库与方法语料库 ($p=0.001 < 0.05$)、引言语料库与讨论语料库 ($p=0.025 < 0.05$)、方法语料库与结果语料库 ($p=0.000 < 0.05$)、结果语料库与讨论语料库 ($p=0.002 < 0.05$)。

维度一中引言语料库与方法语料库的 Z 值均为负数，表示两个语料库都有信息生产的特征，而方法语料库分值高于引言语料库，说明方法语料库的信息密度高于引言语料库。维度二中四个语料库均值都为负数，表明四个语步均为非叙事性文本。引言语料库与讨论语料库的显著性最高 ($p=0.000 < 0.05$)，讨论语料库的 Z 值最大，引言语料库的 Z 值最小，表明讨论语料库的非叙事性最强。维度三中两个显著性差异都体现在方法语料库上，四个语料库的 Z 值都为正值，表明四个语步的文本均为指称明晰性，而方法语料库 Z 值最低，即方法语步相对于其他语步指称不够明晰。维度四中四个语料库 Z 值均为负数，均表现为非劝说型表述，引言语料库与结果语料库存在显著性差异，引言语料库 Z 值最低，结果语料库 Z 值最高，即结果语步比引言语步展现出了更强的非劝说型表述。维度五中引言语料库与结果语料库的 Z 值为正值，

表现为信息抽象性，而方法语料库与讨论语料库的 Z 值为负值，表明信息抽象性较弱。即引言语步、结果语步比方法语步、讨论语步存在更强的抽象性。

（二）各个维度差异存在的不同语言特征

通过对 67 个特征进行单因素方差分析得出，有 27 个特征值存在显著性差异，占全部特征的 40.29%。维度一中有 11 个特征存在显著性差异，正值特征包含 9 个，负值特征包含 2 个。其中四个语步的类符形符比差异较大，方法语步的均值最大（0.42），表明方法语料库中的词类较为丰富，更突出了方法语步的信息性较强，其次为讨论语步（0.39），也展现出了讨论语步传递信息密度较大，最低的为结果语步（0.25），可见其使用的词汇重复率较高且形式较为单一。现在时动词、分析否定、强调语、Be 作主动词、可能意义情态动词也与类符形符比是同样的情况。方法语步（0.38）与结果语步（0.37）的 Z 值显著大于引言语步（-0.92）与讨论语步（-0.44），这意味着方法语步与结果语步的交互性远远高于另外两个语步。维度二中的 6 个特征值仅有句法否定与过去时动词存在显著性差异，表明维度二非叙事性较强，主要为静态叙述。而引言语步（-1.20）与讨论语步（-1.17）的过去时动词 Z 值大于方法语步（-0.80）与结果语步（-0.92），表明引言语步与讨论语步静态叙述更为注重对事物的评判。维度三中名词化和

副词存在显著性差异，具体表现为方法语步均值远低于其他三个语步，这表明方法语步相比于其他语步加强文本信息指称明晰性不够全面。维度四中不定式、预测模态动词和分裂助动词存在显著性差异，其均为负值，这表明这些显著性差异特征在文本中出现频率并不高，然而预测模态动词与分裂主动词的使用能够增强作者观点的说服力。维度五中连词和无主被动存在显著性差异，引言语步在两者的 Z 值上都远高于其他三个语步的 Z 值，这表明引言语步更加强调句子或段落之间的衔接及联系。

五、结束语

本研究采用多维度分析法从五个维度对比了 NATURE 摘要中四个语步的差异。研究结果表明，引言语步有着较强的信息性，非叙事性与信息抽象，方法语步有着较弱的指称明晰性和较强的信息密度，结果语步有着更强的非劝说型表述与信息抽象性，而讨论语步展现了较强的非叙事性。67 项语言特征中有 27 项语言特征的使用频率存在显著性差异。学者在论文摘要写作中可以使用较多的非叙述性语句如动词过去式与公动词等，加大类符形符比，使用丰富多样的词汇且降低词汇重复率，注重时态转换与语篇连接。

参考文献

[1]Anthony, L. (1999). Writing research article introductions in software engineering: How accurate is a standard model?.*IEEE transactions on Professional Communication*,42(1), 38-46.

[2]周汉香, 江勇, 李朝前. 科技期刊及《钢铁研究》栏目设置的思考 [J]. 2017.

[3]Swales, J. M., & Swales, J. (1990).*Genre analysis: English in academic and research settings*. Cambridge university press.

[4]Swales, J. M., & Feak, C. B. (1994). *Academic Writing for Graduate Students: Commentary: A Course for Nonnative Speakers of English*. Ann Arbor: University of Michigan Press.

[5]Conrad S, Biber, D. 2001. *Variation in English: Multi-Dimensional Studies*. Harlow : Longman.

[6]Friginal, E., & Weigle, S. (2014). Exploring multiple profiles of L2 writing using multi-dimensional analysis. *Journal of Second Language Writing*, 26, 80-95.

[7]刘永厚, 张颖. 中外学者国际期刊英语学术论文摘要写作的对比研究 [J]. 外语界, 2016(05):20-27.

[8]Friginal, E., & Mustafa, S. S. (2017). A comparison of US-based and Iraqi English research article abstracts using corpora. *Journal of English for Academic Purposes*, 25, 45-57.

[9]Hardy, J. A., & Friginal, E. (2012). Filipino and American online communication and linguistic variation. *World Englishes*, 31(2), 143-161.

[10]肖忠华, 曹雁. 中外作者科技论文英文摘要多维度语步对比研究 [J]. 外语教学与研究, 2014, 46(02):260-272+321.

[11]闫鹏飞. 硬学科研究论文语篇内部变异的多维度分析 [D]. 北京外国语大学, 2021.