

国际语料库翻译学研究的研究 动态与发展趋势(1993~2022) ——基于CiteSpace可视化分析

姜 琰, 李 芳

上海海事大学外国语学院, 上海

收稿日期: 2023年10月12日; 录用日期: 2023年11月15日; 发布日期: 2023年11月24日

摘 要

本文利用CiteSpace文献计量工具对Web of Science核心数据库中收录的以“语料库翻译”(corpus translation)为主题的文献进行了分析, 绘制出“语料库翻译学”的知识图谱。通过对高影响力作者、高频关键词和突变专业术语的可视化图谱, 对当前国际语料库翻译的热点和前沿问题进行了图谱解读, 客观展现了三十年来国际语料库翻译的基本态势, 并展望未来的趋势。研究发现: 1) 语料库的翻译研究在世界范围内呈现出不均衡的趋势, 欧洲地区已成为研究领域的主要领域, 比利时根特大学是其中的佼佼者。2) 在过去的三十年里, 语料库翻译学的研究主要集中在翻译学的语料库建设、语言特性、视听翻译和非文学翻译四个领域。3) 语料库的研究正逐步趋向于多学科的发展。这些研究结果对于推动语料库翻译学的发展具有重要意义, 为相关领域的学者提供了参考和启示。

关键词

语料库翻译学, 文献计量统计, CiteSpace

The Research Status and Trends of International Corpus Translation (1993~2022): Visual Analysis Based on CiteSpace

Yan Jiang, Fang Li

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: Oct. 12th, 2023; accepted: Nov. 15th, 2023; published: Nov. 24th, 2023

文章引用: 姜琰, 李芳. 国际语料库翻译学研究的研究动态与发展趋势(1993~2022)——基于 CiteSpace 可视化分析[J]. 现代语言学, 2023, 11(11): 5396-5403. DOI: 10.12677/ml.2023.1111726

Abstract

This paper utilizes the CiteSpace bibliometric tool to analyze the literature on “corpus translation” in the Web of Science core database. It constructs a knowledge map of “corpus translation studies” based on high-impact authors, high-frequency keywords, and mutated technical terms. The map provides an objective interpretation of the current hot and frontier issues in international corpus translation, presenting the fundamental trends of international corpus translation over the past thirty years and predicting future directions. The findings reveal that: 1) Corpus translation research exhibits an uneven global distribution, with Europe being the main research hub and Ghent University in Belgium emerging as a prominent institution in this field. 2) Over the past three decades, research in corpus translation studies has primarily focused on four domains: corpus construction in translation studies, language features, audiovisual translation, and non-literary translation. 3) Corpus research is gradually moving towards multidisciplinary development. These findings are of great significance in promoting the development of corpus translation studies and providing insights for scholars in related fields.

Keywords

Corpus Translation, Bibliometric Analysis, CiteSpace

Copyright © 2023 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

语料库翻译学的研究最早可以追溯到 20 世纪 80 年代。Mona Baker 她的文章《Corpus Linguistics and Translation Studies: Implications and Applications》中首次将语料库应用于翻译研究[1]。语料库翻译学是翻译学和语料库语言学的结合,是翻译学和语言学研究的最新发展趋势,也是一个最新的研究方向。在过去的三十年里,各国学者在语料库翻译学领域取得了丰硕的研究成果。对语料库翻译学知识结构的梳理对于了解国内外该领域的发展动态具有极高的价值。张继光和何春艳利用 CiteSpace 信息可视化软件对国内这一新的研究模式进行了全方位的梳理,为我国的翻译学界勾勒出了一个新的学术格局[2] [3]。然而,由于缺乏从国外角度的全方位分析,本文作者利用 CiteSpace 信息可视化软件从文献计量学的角度对与“语料库翻译”相关的文献进行了系统的整理和分析,准确、形象、直观地呈现了刊发载体、主要作者、研究机构、研究热点和前沿动态等重要信息,以期更好地认识该领域的发展。

2. 研究方法和文献收集

本文使用的分析工具是 CiteSpace,这是由美国德雷赛尔大学的陈超美教授开发的一款多元、分时、动态的信息可视化软件。该软件目前被广泛应用于自然科学领域的前沿和发展趋势的分析和预测。为了进行本文的研究,作者于 2023 年 1 月 12 日在 WoS 核心数据库中使用“corpus OR corpora AND translation”作为检索词,将研究文献类型限定为“Article”,删除了会议等其他类型的文献,并将研究领域限定为“language and linguistics”和“linguistics”两类期刊,同时剔除了与计算机科学相关的文献。最终,作者获得了 610 篇有效文献,并将其以纯文本的 TXT 格式导出,用于本文的元数据分析。

3. 统计结果分析与讨论

根据研究目的, 本研究在利用 Citespace 对数据进行分析时, 对有关参数进行了设置。例如, 时间窗口 Time Slicing 被设置为 1993 至 2022, 时间切片 Year per Slice 设置 1, 节点类型 Node Type, 根据研究目的先后勾选 Author 和 Keyword, 将 Top N 的阈值设定为 40, 在此过程中始终选择寻径网络算法 (pathfinder)。视觉可视化效果选择为静态(Cluster View Static)以显示最终的网络图谱效果。

3.1. 研究力量分布

3.1.1. 发文量年份分布



Figure 1. Annual publication volume (1993~2022)

图 1. 年度发文量(1993 年~2022 年)

图 1 清晰地显示, 第一篇有关语料库研究的论文出现在 1993 年。然而, 直到 2000 年, 语料库翻译学研究才逐渐开始兴起, 在随后的十年中发展缓慢。2012 年是语料库翻译学研究的转折点, 发表的论文数量开始显著上升。在此后的几年中, 论文数量呈现出爆发式增长, 其中最为明显的增长年份为 2012 年、2016 年、2019 年, 而 2021 年则达到了顶峰, 共发表了 77 篇论文。在 2017 年至 2022 年期间, 国际学者对语料库翻译研究的关注度最高, 平均每年发表论文数量在 55 篇至 60 篇之间。

3.1.2. 发文国家机构及作者

根据表 1, 全球范围内的语料库翻译学研究呈现不均衡的态势。尽管英国是语料库翻译学的发源地, 但其在这一领域的研究成果并不多。相比之下, 西班牙的论文产量最高, 达到了 129 篇, 中国排名第二, 共有 86 篇论文发表, 美国位列第三, 发表了 57 篇论文, 其余比利时、德国、法国、意大利、波兰等欧洲国家也成为了语料库翻译学研究的重要阵地。同时, 中国也是语料库翻译学研究的中坚力量之一。

图 2 是共现图, 其中每个圆形节点代表一个关键词, 节点的大小表示出现频次的高低, 而年轮的厚度则代表在该时间分区里论文产出的频率。节点由不同颜色的年轮构成, 每一个年轮对应出现时间, 由内到外对应的时间分区由远及近, 最早的时期为淡灰色至深灰色, 向冷色转变, 近期为温暖的颜色。从图中可以看出, “西班牙”、“德国”、“中国”、“美国”和“英国”这几个节点从紫色到红色均有覆盖, 说明该主题研究在各年份都有广泛的分布。同时, 从节点之间的连线也可以得出一些结论, 例如

法国、芬兰和荷兰与美国早期合作发表了较多的论文。而英国和美国、德国和法国、荷兰和瑞士之间的连线呈现红色，说明近三年来这些国家之间的合作发文较为频繁。

Table 1. Top 10 countries and top 10 authors with the highest productivity

表 1. 语料库翻译学最高产的 10 个国家和前 10 位作者

序号	国家	频次	作者	频次
1	Spain	129	Gert De sutter	7
2	People R China	86	Bart Defrancq	5
3	USA	57	Mario Bisiada	5
4	England	55	Lucja Biel	4
5	Belgium	42	Hermann Ney	4
6	Germany	33	Haidee Kruger	4
7	France	27	Andrew KF Cheung,	4
8	Italy	26	Feng Pan	3
9	Poland	25	Rosario Caballero,	3
10	Netherlands	17	Bonnie J Dorr	3

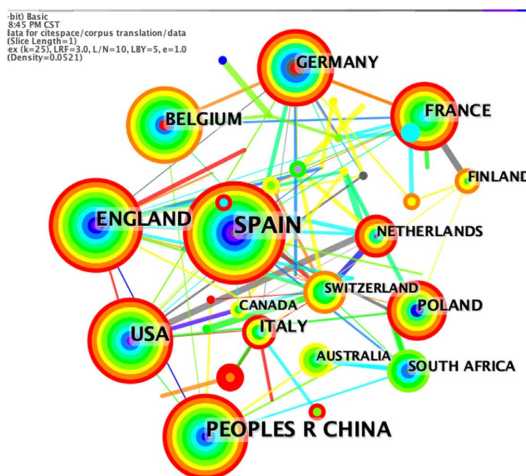


Figure 2. Co-occurrence network map of highly productive countries

图 2. 论文高产国家共现图谱

根据表 1 所示，这些作者在 1993 年至 2022 年期间在 WOS 核心期刊上发表了三篇以上的研究性语料库翻译学论文，可视为该领域的高产作者。为了获得更为详尽的语料库翻译学研究高影响力作者网络共现知识全景图，笔者将高影响力作者信息导入 CiteSpace，并得到了图 3。在语料库语言学研究领域，比利时根特大学的 Gert De sutter [4] [5] 和 Bart Defrancq [6] [7] 是最高产的作者。根据节点外围颜色，我们可以看出近两年来，荷兰乌得勒支大学的 Martijn van der Klis 和华中科技大学的潘峰 [8] 等新进入该领域的研究者在该领域的研究中非常活跃。

机构和作者之间有着广泛的合作和联系。笔者通过对比语料库翻译学研究发文量排名前 10 位的科研机构，发现西班牙有三所机构上榜，这与其在世界发文量排名中位居第一的地位相符。发文量最大的机构是比利时的根特大学，共发表 19 篇论文。该机构最突出的学者为 Gert De Sutter 和 Bart Defrancq，与高产作者前两位的地位相符。西班牙的格拉纳达大学 (13 篇)、巴塞罗那自治大学 (9 篇)、比利时的鲁汶大

学(9 篇)以及中国的香港理工大学(8 篇)紧随其后。中国在该领域的研究机构主要有广东外语外贸大学、北京外国语大学等较为领先。从图 4 中各个机构的节点外围颜色可以看出, 乌特勒支大学、巴塞罗那自治大学、都柏林城市大学和香港理工大学在语料库翻译学领域的研究一直持续至今。

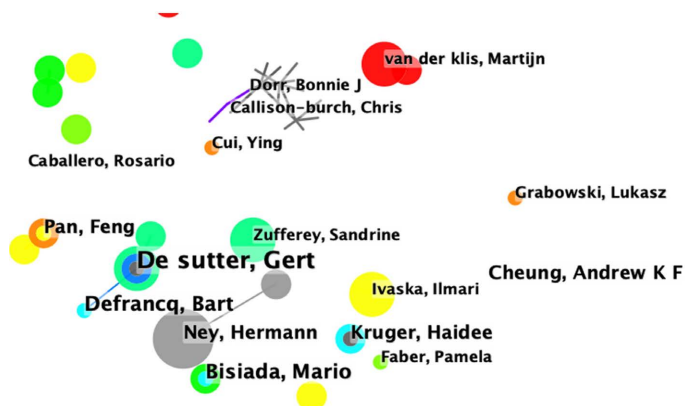


Figure 3. Co-occurrence network map of highly influential authors
图 3. 高影响力作者网络共现图谱

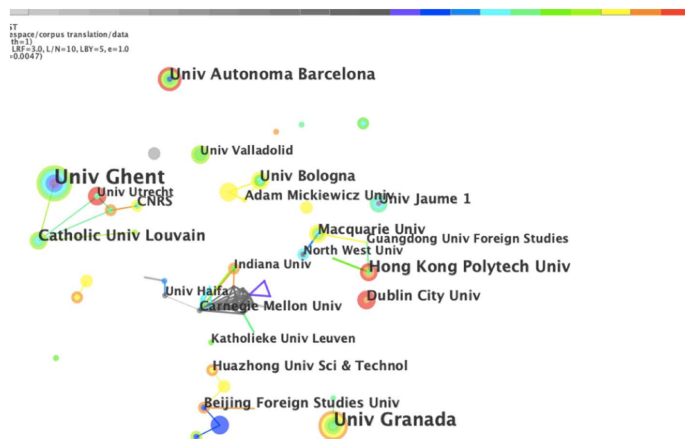


Figure 4. Co-occurrence network map of highly productive institutions
图 4. 高产发文机构共现图谱

3.1.3. 聚类分析

聚类分析能够帮助我们理解世界范围内该领域的研究热点, 笔者将 603 条文献进行关键词的聚类, 得出下图(图 5)。西班牙发文量排名第一, 研究聚焦于 audio description (视听翻译)、media accessibility (无障碍媒体)、terminology (术语管理), 侧重于非文学翻译。比利时的研究内容比较广泛, 包括 modifier attachment (修饰成分)、corpus frequency (语料库频率)、lexical presentation (词汇表达)、late closure (后封闭原则)、disambiguation preference (词义消歧偏好); 中国的研究侧重于机辅翻译(CAT)、film title (字幕翻译); 美国集中在研究 eye tracking (眼动实验)、medical translation (医学翻译); 德国的研究聚焦于 communicative dynamism (交际性动力); 南非主要研究 journalist translation (新闻翻译), 可以看出, 每个国家都有自己的擅长领域。虽然有些国家的论文数量很少, 但其研究内容也值得关注, 例如澳大利亚研究侧重 appraisal theory (评价理论), 伊朗的研究更侧重 bilingual compound verb (双语复合动词)、academic motivation scale (学习动机量表)。

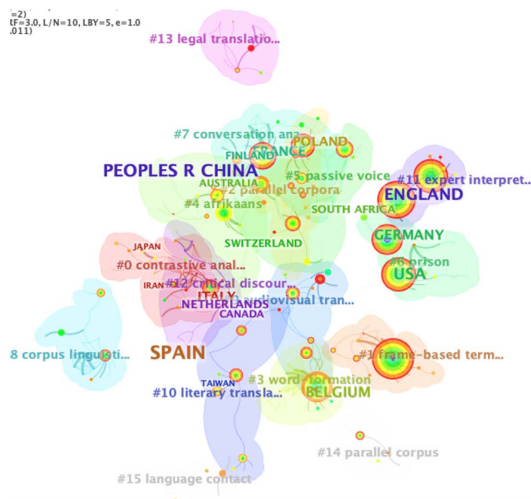


Figure 5. Research strength clustering
图 5. 研究力量聚类共现知识图谱

3.2. 语料库翻译学研究热点分析

通过对关键词变化的分析, 可以了解近十年语料库翻译学领域的研究热点和重要主题。本文使用了 CiteSpace 信息可视化软件来统计关键词的频次和共现情况, 并通过可视化图谱展示关键词的频次和聚类关系, 形成了一个知识图谱, 以分析国际语料库翻译学研究领域的研究热点和趋势(见表 2 和图 6)。根据 CiteSpace 的计算结果, 频次最高的五个关键词分别是 language (42 次)、English (39 次)、translation (36 次)、corpus-based translation study (31 次)和 translation universal (17 次)。通过中介中心强度的大小(强度有效取值为 0.01), 可以探测一个领域的变化程度。在 2006 年, 中介中心强度最大的词汇是西班牙语和机器翻译; 2012 年的关键词是英语、基于语料库的翻译研究、视听翻译、类比语料库、文学翻译和交流; 2013 年的关键词是电影字幕翻译、话语分析和法律翻译; 2014 年的关键词是视听翻译和语料库语言学; 2016 年的研究热点是翻译态度和解释(explication)、大众科学; 2017 年的关键词包括医学翻译、多模态、感官知觉和第二语言; 2018 年的关键词与社会现象结合, 包括气候变化、音频描述和新闻翻译; 2019 年的关键词包括术语能力、批评话语分析、平行语料库、定量分析和机器翻译; 2020 年的关键词是词块和制度变迁; 2021 年的关键词是共现数据和亚马逊语言; 2022 年的关键词是政治话语、神经机器翻译和受众接受度。从关键词的发展脉络可以发现, 近十年来, 语料库翻译研究逐渐向多考察点、多语种、多学科的发展模式转变, 并且关注点也逐渐从语言表面向深层发展。

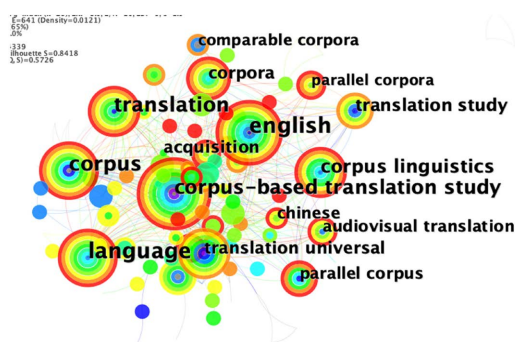


Figure 6. High-frequency keywords (1993~2022)
图 6. 高频关键词共现图谱(1993~2022)

Table 2. High-frequency and high-centrality keywords (1993~2022)
表 2. 高频关键词和高中心度关键词(1993~2022)

高频关键词		高中心度关键词	
关键词名称	频次	关键词名称	中心性
language	42	English	0.26
English	39	translation	0.20
translation	36	language	0.18
corpus-based translation study	30	corpus-based translation study	0.15
translation universal	17	corpus linguistic	0.14
audiovisual translation	15	translation study	0.09
literary translation	12	acquisition	0.06
translation study	12	discourse	0.05
machine translation	11	comparable corpora	0.05
parallel corpus	11	corpus analysis	0.04
Chinese	11	translation universal	0.04

3.3. 研究发展态势

总体而言,从关键词的发展轨迹来看,非文学翻译领域的研究已经扩展到音频描述和科普读物,这表明非文学翻译已成为一个新的研究领域。同时,同声传译的兴起也预示着口译语料库将成为未来的研究热点。此外,气候变化、音频描述和新闻翻译的出现表明跨学科研究将成为语料库翻译学未来的发展趋势,与心理语言学和认知语言学的结合也将成为主流。根据科学发现的理论,许多重要的科学贡献都源自于跨学科研究[9]。语料库翻译学的研究方法不断被应用于法律翻译、译者技能、教育、新闻和医疗等多个领域。这些新的研究视角的形成和发展既是该学科自身发展的必然要求,也是各学科之间相互协作和相互影响的结果。

4. 结论

本文对 Web of Science 数据库中收录的 610 篇国际语料库翻译学文献进行了分析和梳理,旨在揭示语料库翻译学的研究力量分布、知识来源以及热点和前沿研究,以此为基础预测该领域的发展趋势。新的前沿研究领域包括机器翻译、语料库数据、法律翻译、语言接触、医学翻译和政治话语等。

1) 在研究力量方面,语料库翻译学的分布在全球呈不均衡态势。欧洲国家是该领域的研究阵地,中国近年来逐渐成为了该领域的中坚力量之一。

2) 从研究方法来看,语料库的使用正逐渐成为翻译研究的趋势。通过对该领域的考察,本文发现了许多新的研究思路和视角,从多个方面拓展了实证研究的范围。同时,随着新领域的发展,语料库的研究技术和方法也在不断更新。

3) 从跨学科的角度来看,翻译文本特征研究和语言对比研究仍然是语料库翻译研究的重点。语料库翻译学作为一门具有发展前景的学科,其跨学科特点决定了研究视角的多样性,需要专家学者进行深入挖掘和探索。语料库翻译需要从语言学、认知语言学、传播学等多个学科的角度出发,对语料库的核心概念进行精炼,并从理论、方法和观点等方面进行细致探讨。

参考文献

- [1] Baker, M., Francis, G. and Tognini-Bonelli, E. (1993) Text and Technology: In Honour of John Sinclair. John Benjamins Publishing Company, Amsterdam, 233-250. <https://doi.org/10.1075/z.64>
- [2] 张继光. 国内语料库翻译学研究状况的科学知识图谱分析(1993-2014) [J]. 上海翻译, 2016(3): 34-40+61+93.
- [3] 何春艳. 国内语料库翻译学研究团队可视化分析(1993-2020)——基于科学知识图谱[J]. 现代交际, 2021(1): 76-78+75.
- [4] De Sutter, G. and Lefer, M.-A. (2020) On the Need for a New Research Agenda for Corpus-Based Translation Studies: A Multi-Methodological, Multifactorial and Interdisciplinary Approach. *Perspectives*, **28**, 1-23. <https://doi.org/10.1080/0907676X.2019.1611891>
- [5] Prieels, L. and De Sutter, G. (2017) Between Language Policy and Language Reality: A Corpus-Based Multivariate Study of the Interlingual and Intralingual Subtitling Practice in Flanders. *Perspectives*, **26**, 322-342. <https://doi.org/10.1080/0907676X.2017.1380053>
- [6] Collard, C. and Defrancq, B. (2018) Predictors of Ear-Voice Span, a Corpus-Based Study with Special Reference to Sex. *Perspectives*, **27**, 431-454. <https://doi.org/10.1080/0907676X.2018.1553199>
- [7] Plevoets, K. and Defrancq, B. (2016) The Effect of Informational Load on Disfluencies in Interpreting. Translation and Interpreting Studies. *The Journal of the American Translation and Interpreting Studies Association*, **11**, 202-224. <https://doi.org/10.1075/tis.11.2.04ple>
- [8] Pan, F., Russo, M., Bendazzoli, C. and Defrancq, B. (2019) Making Way in Corpus-Based Interpreting Studies. *Across Languages and Cultures*, **20**, 147-151. <https://doi.org/10.1556/084.2019.20.1.9>
- [9] Chen, C. (2017) Science Mapping: A Systematic Review of the Literature. *Journal of Data and Information Science*, **2**, 1-40. <https://doi.org/10.1515/jdis-2017-0006>