

基于 R 语言的汉语“被+N./V.+给”与葡萄牙语“ser+过去分词+por”句式对比分析

谢钰琪

湖南师范大学, 湖南长沙, 410006;

提要: 被动句是突出受事、调整信息结构的重要语法手段, 在不同语言中的构式和功能分布存在差异。本文以北京大学 CCL 现代汉语语料库和 Corpus do Português 为基础, 对比分析汉语“被+N./V.+给”与葡萄牙语“ser+过去分词+por”被动构式特征。研究通过人工标注施事、受事及介词短语类型, 并结合 R 语言对清洗后的语料进行量化分析, 统计并可视化了施事和受事的显隐性分布, 以及介词短语类型与二者显隐性的关系。研究问题聚焦于: (1) 两种语言被动句在施事和受事显隐性分布上的异同? (2) 介词短语结构类型如何影响施事和受事的显隐性? (3) 这些差异和联系背后反映了怎样的语言结构和语用机制?

关键词: 汉语被动句; 葡萄牙语被动句; 介词短语类型; 显隐性; R 语言

DOI: 10. 64216/3080-1516. 25. 05. 051

引言

在全球化背景下, 探讨不同语言的句法结构对语言学理论发展和学习具有重要意义。被动句作为汉葡两语中的重要句式, 承载着复杂的语法语义信息, 并反映了认知与表达差异。尽管汉语“被+N./V.+给”句式研究丰富, 且与英语被动句对比成果显著, 但国内针对其与葡萄牙语“ser+过去分词+por”被动句式的研究尚少。

1 文献回顾

1.1 汉语被动句与“被+N+给”句式研究

汉语被动句以“被”字句为典型, 传统结构为“被+施事+动词+受事”。近年来, “被+N./V.+给”句式受到关注, 如“他被我给打了”。其中“给”作为介词引入施事, 与动词结合, 可表达施予、转移等语义, 功能更丰富(赵元任, 1968; 李宇明, 2003)。语义上, 该句式常用于突出受事的被动性与施事的主动性, 具较强情感色彩和语用功能。句法上, 该结构允许施事、受事、动词和“给”灵活组合, 体现汉语被动句的构式多样性(Lü, 1980)

1.2 葡萄牙语被动句与“ser+过去分词+por”句式

葡萄牙语主要有两种被动语态: 助动词被动语态和代词被动语态。本文关注的助动词被动语态, 是助动词“ser”与主动词的过去分词形式构成, 如“Os traba-

hos foram feitos pelo Pedro.”(这些工作是由佩德罗完成的)。在此结构中, 助动词“ser”与动词过去分词搭配, 主语与过去分词在人称和数上保持一致, 施事由介词“por”引入, 强调动作承受者及施事显性表达。施事不重要或已知时, “por”短语可省略。

1.3 跨语言被动句式对比研究

汉葡被动句在构式、施事表达、受事突显及介词短语结构等方面存在显著差异(Givón, 1994; Keenan & Dryer, 2007)。汉语被动句更注重语用和情感色彩, 葡语则强调形式一致性和语法规则。此外, 葡语被动句形式多样, 语法形态丰富, 施事表达更显性; 汉语被动句使用频率较低, 语义多集中于不利或受害事件, 且对语境依赖度高, 施事推断需借助语境, 缺乏明确语法标记。

2 研究设计

2.1 研究方法

本研究以北京大学 CCL 语料库和 Corpus do Português 为数据源, 选取含目标被动句式的语料, 人工标注施事、受事及介词短语结构类型, 构建对比数据集。旨在通过定量统计与比较分析, 揭示汉葡被动句中施事与受事的表达方式、介词短语的类型分布及其句法功能的异同, 并阐释其构式特征。

方法上, 本文采用 Excel 整理数据, 并利用 R 语言进行统计分析与可视化, 充分利用编程平台的数据处理

与统计功能,实现量化对比。结合语料库语言学与句式语法理论,提升分析的科学性和客观性,为跨语言句法对比研究提供新路径。最终,基于实证数据,相对系统地总结汉葡被动句的句式特征异同,丰富被动句的跨语言认知研究。

2.2 研究步骤

2.2.1 语料搜集与预处理

本研究语料主要来源于北京大学中国语言资源库和 Corpus do Português 的子语料库 NOW (葡萄牙语, 2012-2019 年间语料)。通过关键词检索抽取数据: 葡语检索 “ser [vpp*] por” 获得 89 条 “ser+过去分词+por” 结构, 为避免动词时态影响, 仅选择 “ser” 动词形式; 汉语检索 “被(X, 1-4)给” 获得 70 条 “被+N./V.+给” 句式。

初步收集后, 人工预处理, 剔除非被动句、非 “被…给” 固定句型及重复句子。例如, 汉语中 “那一段时间, 大家一天当三天用, 硬是把工作的被动局面给扭转了过来” 因实际为主动语态而被剔除。清洗后, 最终保留 58 条汉语语料和 86 条葡萄牙语语料, 形成分析语料库。

2.2.2 语料标注与分析

语料预处理后, 采用人工标注, 基于 Excel 表格, 设置施事 (agent)、受事 (patient) 和介词短语类型 (pp_type) 三项核心字段。agent 和 patient 字段记录施事和受事的具体表达; 若未在句面出现, 则依据上下文推断, 标记为 “~” 表示隐性成分, 便于 R 语言自动分类统计, 保障分析客观性。

介词短语类型依语义角色分为施事介词短语 (P-AG)、工具介词短语 (P-INST) 和地点介词短语 (P-LOC) 三类, 便于分析其在被动句中的功能。此分类源于 Fillmore (1968) 格语法理论, 强调语义角色对句法结构的作用。P-AG 引入动作发出者, 是施事角色的显性表达, 如汉语 “给+施事”、葡语 “por+施事”; P-INST 表示动作所依赖工具, 如 “用笔”; P-LOC 描述动作空间背景, 如 “在门上”。

标注完成后, 利用 R 语言加载 “readxl”、“dplyr”、“ggplot2” 等包, 导入标注语料。为保证分析针对性, 首先清洗数据, 剔除 pp_type 列为 “~” 的句子, 排除非被动句或不符研究范围的实例, 确保后续统计聚焦目标句式。

3 研究发现

3.1 施事与受事显隐性比例

在分析施事和受事的显隐性比例时, 本文发现, 汉葡的施事隐性比例都要远高于其本身施事显性比例, 分别占比 67.2% 和 60.5%。两种语言的受事隐性比例也不约而同高于受事显性比例, 分别占比 91.4% 和 39.5%。虽然数据上显示, 葡萄牙语的受事显隐性比例相差别不会特别大, 相差 15.1%, 但我们仍然认为葡萄牙语的受事显性占优势。此外, 葡萄牙语对于施事省略或弱化的倾向会远大于汉语。

3.2 介词短语类型分布

本文也对介词短语类型 (pp_type) 进行了统计。结果显示, 汉葡被动句均偏好使用引入施事的介词短语 (P-AG), 使用上分别占比 82.8% 和 57.0%, 工具介词短语 (P-INST) 次之, 分别占 15.5% 和 40.7%。

3.3 施事和受事显隐性介词短语类型的关系

通过交叉统计和可视化分析, 本文发现汉语在工具介词短语 (P-INST) 中更倾向搭配显性施事 (55.6%), 在施事介词短语 (P-AG) 中更倾向隐性施事 (70.8%)。而葡萄牙语则相反, 在工具介词短语中施事倾向隐性表达 (91.4%), 在施事介词短语中则倾向显性施事 (63.3%)。

总体来看, 汉葡被动句中受事均倾向显性表达, 符合被动句中受事作为信息焦点的特点。但在不同介词短语类型下, 受事隐性比例存在差异。汉语施事介词短语 (P-AG) 下受事显性比例高达 91.7%, 而工具介词短语 (P-INST) 下为 88.9%, 可能与工具本身暗示受事有关。葡萄牙语施事介词短语 (P-AG) 下受事显性比例为 75.5%, 工具介词短语 (P-INST) 下为 74.3%, 相较汉语, 葡语在施事介词短语下受事隐性表达比例更高 (见图 1)。

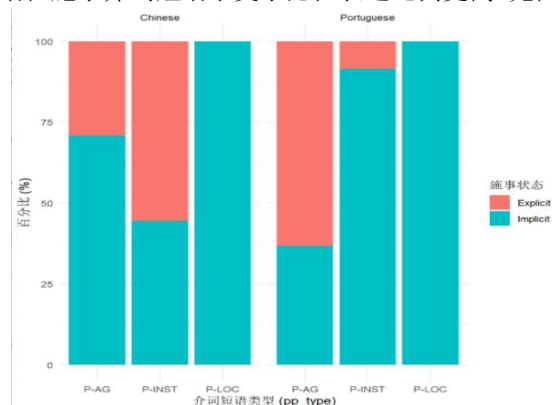


图 1. 介词短语结构类型与施事显隐性关系

4 结语与展望

本文以汉语“被+N./V.+给”和葡萄牙语“ser+过去分词+por”结构为研究对象,结合人工标注与R语言量化分析,系统探讨了两种语言被动句中施事与受事的显隐性及其与介词短语类型的关系,旨在揭示句式共性与差异,为跨语言研究提供实证支持。

分析显示,汉葡被动句中受事成分均呈强烈显性,尤其在地点介词短语(P-LOC)中必须显性表达,反映其语法语义稳定性。施事显隐性则差异显著:汉语“被+N./V.+给”中施事隐性比例高,倾向省略或弱化施事,突出受事被动性与情感色彩;葡语被动句中施事显性多,尤其在施事介词短语中表达更明确,强调施事参与。

介词短语类型分布上,两语言均偏好施事介词短语,但葡语工具介词短语使用明显多于汉语,体现其在信息补充与语义细化上的不同策略。施事与受事的显隐性与介词短语类型存在互动,使得两语在被动句搭配上呈现不同规律,有助于语言学习者理解被动句内部结构的功能分化。

本文为理解汉葡被动句信息结构及语用功能提供数据基础,也为跨语言句式对比提供实证依据。结合现代数据分析技术,可更深入揭示被动句成分间互动及语言表达的共性与差异。本研究仍存在语料库规模偏小的局限,尤其地点介词短语(P-LOC)数据不足,限制了对该类别与施事受事显隐性的深入分析。葡萄牙语完整语料收集难度大,样本量有限,影响结论稳健性。未来研究应扩大语料库,结合更细致的句法分析、语用因素

与信息结构,全面揭示不同语言被动句的复杂语言现象。

参考文献

- [1]Chao,Yuen Ren(赵元任).现代汉语语法[Modern Chinese Grammar].商务印书馆[The Commercial Press],1968.
- [2]Givón,Talmy,editor.Voice and Inversion.John Benjamins,1994.
- [3]Keenan,Edward L.,and Matthew S. Dryer. “Passive in the World’s Languages.” *Language Typology and Syntactic Description*. Edited by Timothy Shopen, vol. 1, Cambridge University Press, 2007, pp. 325 – 361.
- [4]Li,Yuming (李宇明). “现代汉语‘被’字句的多样性 [The Diversity of ‘Bei’ Sentences in Modern Chinese].” *中国语文[Studies of the Chinese Language]*, no. 6, 2003, pp. 481 – 491.
- [5]Lü, Shuxiang(吕叔湘). *Xiandai Hanyu Babai Ci [Eight Hundred Words in Modern Chinese]*. The Commercial Press, 1980.
- [6]Wang,Hongjun(王洪君). “‘被+给’句式的语法化及其语用功能[The Grammaticalization and Pragmatic Functions of the ‘Bei+Gei’ Construction].” *语言研究[Linguistic Research]*, no. 2, 2012, pp. 23 – 29.

作者简介:谢钰琪(2000.4),女,汉族,广东潮州,硕士在读,研究方向:多模态语义学、语言对比