

基于语料库的中国形象历时对比研究 ——以领导人重要讲话为例

白植坪

摘要: 本文通过将习近平系列重要讲话数据库与三位美国总统涉华语料进行比较,运用LDA主题建模、词汇联想分析、用法波动分析和情感分析等语料库方法,考察了中美政治话语中的中国形象建构特征。研究发现,习近平主席讲话以人民为中心,形成了以发展、改革、合作与和平为基础的多维国家形象。相比之下,美国总统涉华话语中的中国形象经历了从“合作共赢”向“战略竞争”的转变。

关键词: 中国形象;语料库;主题建模

一、引言

国家形象历来是国际传播研究的中心话题,近年来中国的全球影响力显著提高,因此中国的国际形象已经成为学术界研究的热点。现有研究主要关注西方媒体他塑视角里的中国形象,关于中美领导人话语中的中国形象自塑和他塑的对比研究仍然相对稀少。鉴于中美关系在国际体系中的关键地位,系统地分析习近平主席讲话中的中国形象,并将其与奥巴马、特朗普和拜登的涉华讲话进行比较,有助于揭示不同视角下中国形象的话语建构机制。

二、文献综述

Fairclough(2003)指出,批评话语分析能够揭示话语中隐含的权力关系和意识形态。这为分析领导人讲话中的国家形象建构机制提供了一个宏观理论框架。而语料库语言学的发展推动了政治话语研究从以定性方法为主到定量方法并重的转变。Baker等人(2013)强调,使用语料库方法分析文本,可以揭示话语背后的意识形态。

在国家形象传播的研究中,Law(2009)从媒体他塑视角揭示了西方舆论有色眼镜下的中国形象,而许峰、高意(2023)则从自塑视角研究了习主席外交话语中的中国生态形象。总体而言,现有中国形象的相关研究成果丰富。然而,关于中美领导人话语中的中国形象历时对比研究仍然相对稀少。本文旨在采用语料库方法,探讨中美领导人话语中的中国形象建构模式的历时差异。

三、语料库设计

习近平系列重要讲话数据库收录了习主席的重要讲话和文章,涵盖了政治、经济、社会和文化等核心领域。本文收集了该语料库中2012年11月至2024年6月的新华网和新华社的新闻语料,共140个TXT文本(140个月),总字数85,510字,经过ICTCLAS2012的分词处理。本文采用LDA主题建模、词汇联想分析、搭配词分析、UFA(用法波动分析)及情感分析等语料库工具,探讨习主席讲话中的国家形象。

美国总统计划项目(The American Presidency Project)是由加州大学圣塔芭芭拉分校创建的一个在线实时更新的数据库。它收录了美国历史上所有总统的公开演讲、新闻发布会、访谈等文献。本文将从该语料库中提取三任美国总统语料里含有“China”的全部文本,使用Lancsbox软件中的GraphColl功能分析节点词“China”的搭配特征,揭示美国对华政治话语中的中国形象历时变化。通过自塑与他塑的双重视角,本文旨在呈现中国形象的差异化建构模式及其背后的意识形态动因。

四、话题分析

本文先使用KH Coder软件对习主席话语语料库进行了LDA主题建模分析,得到了以下11个主题:国家领导、中央决策、政策讨论、党的领导、国际合作、改革创新、中国人民、国家安全发展、国际事务、经济社会发展、历史文化。

这些主题清晰地反映了习主席政治话语的基本结构及其新时代国家治理的关注焦点。其中,“国家领导”“中央决策”“国际合作”等都属于讲话中对国家治理、外交布局的讨论,而“改革创新”“经济社会发展”等说明了中国新时代推进现代化建设、提高国家竞争力的具体方向,而“中国人民”“历史文化”等很好地体现了“以人为本”的发展理念。这些主题共同塑造出一个欣欣向荣、自信开放的中国形象。

五、词汇联想分析

本文选取了习主席话语语料库中出现频率最高的前200个词语,使用KH Coder软件的词汇联想分析功能,来揭示特定检索词的使用型式。第一步,以“中国”作为检索词进行词汇联想分析,以探索国内形象。通过分析“中国”共现的高频词汇,理解“中国”一词在习主席讲话中的使用频率和语境,以及与之相关联的词汇如何

共同塑造了国内形象。结果显示,以“中国”为中心的词汇网络里,“发展”“人民”“国家”“社会”“经济”和“政府”等词语高度共现,形成了以全面发展和民生治理为中心的语义框架。“人民”的高频率共现体现了“以人为本”的发展理念。同时,“国际”“合作”和“伙伴”等词的共现说明中国在国际事务中一贯奉行开放共赢的外交方针。

第二步,以“中方”为检索词,分析与“中方”共现的词汇,并构建一个词汇共现网络,着重探索外交形象。观察“中方”一词在国际交流和外交场合中的使用情况,以及它如何与相关词汇一起构建中国的外交形象。结果显示,“合作”“发展”“交流”“战略”“和平”和“伙伴”等共现词语突出了中国在国际事务中的建设性角色。“安全”“稳定”“协调”“投资”和“贸易”等词揭示了中国积极参与国际事务合作。“支持”“深化”和“感谢”等词语则强化了平等友好的外交立场。

第三步,同时将“中国”和“中方”作为检索词,以一个更全面的视角来探索国内形象和外交形象之间的联系。结果显示,以“中国”和“中方”为中心的词汇网络里,“中国”关联“发展”“改革”“经济”和“人民”等词,强调国内建设和民生稳定,而“中方”则与“战略”“合作”和“国际”等词相关联,反映了中国积极参与全球治理的政策导向。

六、搭配和索引行分析

以上词汇网络表明,“人民”在习主席讲话中始终占据中心位置。笔者使用Lancsbox的Collocate功能,在习主席话语语料库中检索了以“人民”为节点词的高频搭配词。结果显示,在政治层面上,“人民”与“共和国”“党”和“政府”等词的密切联系强调了“人民是根本”的政治逻辑。党和政府通过“服务”“造福”和“团结”不断提升民生。在社会层面上,“发展”“健康”和“美好”等搭配词表明,人民福祉是制定政策的目标。在国际关系层面,“世界”“共同”“和平”等词反映出中国政府共同构建人类命运共同体的国际姿态。同时,“伟大”“奋斗”“实现”等搭配词突出了中国人民在中华民族伟大复兴中发挥着核心作用。

本文仅对2013至2023年“人民”节点词的搭配词进行UFA分析。首先,使用PinyinTagger软件将“人民”的搭配词转换为拼音。然后,使用Lancaster Stats Tools online在线网站的UFA工具对搭配词进行分析。UFA工具自动对节点词的搭配词分为四类:持续搭配词(长期保持的搭配)、终止搭配词(一段时间内保持,后来消失的搭配)、起始搭配词(一段时间内生成且后来保持的搭配)和短暂搭配词(短暂出现的搭配)。

分析结果显示,在2013年至2023年期间,“人民”的搭配词只有持续和短暂搭配词,共748个。“服务”“保障”“贡献”“幸福”和“政府”等持续搭配词反映了党为人民服务的根本宗旨。与此同时,短暂搭配词主要出现在特定的历史节点。例如,“安全感”集中在2017年至2020年,反映了国家在社会治理方面的政策导向;“抗击”“疫情”“保护”等词与新冠肺炎防控有关,体现了“人民至上,生命至上”的治理原则。总体而言,UFA结果揭示了“人民”是习主席话语体系中社会情感和民族叙事的凝聚核心。以下是习主席话语语料库中含有“人民”节点词的典型索引行。

2013年1月习主席在全国政协新年茶话会上的讲话:“中国特色社会主义事业是造福人民的美好事业,也是需要我们为之付出智慧和汗水的征程。”这句话将中国特色社会主义事业定义为“造福人民”的事业,通过强调“智慧和汗水”,习主席将国家发展与人民奋斗联系起来,表明人民既是发展的动力源,也是实现中华民族伟大复兴的主力军。

七、情感分析

下一步使用LIWC工具,对2012年11月至2024年6月期间的习主席话语语料,按月份为单位,分别提取积极情感词和消极情感词的百分比,结果发现,2024年1月的语料中积极情感词和消极情感词比例的差值是1.2,是差值最大的月份,最具有代表性。将该月份含有“人民”节点词的话语语料和以下指令输入GPT:“分析文本的情感。在分析过程中,充分考虑语境,并根据-10(非常负面)至10(非常正面)的标准进行评分。结果应以三列展示,第一列为文本,第二列为情感,第三列为分数。第二列的情感须明确指出具体的情感类型”。表1展示的是3个典型的示例。

表1情感分析

序号	文本	情感	分数
1	“媒体在推动各国人民心灵相通方面发挥着越来越重要的作用。”	积极(友善、包容)	8
2	“中国人民立足国情、放眼世界,勇于自我革命、自我革新。”	积极(自信、奋进)	10
3	“禁毒工作事关国家安危、民族兴衰、人民福祉,毒品一日不除,禁毒斗争就一日不能松懈。”	负面(忧虑、警醒)	-6

对话料的情感分析表明,习主席讲话整体呈现出积极、理性与责任并重的情感特征,积极情感话语占主导地位,消极情感话语主要用于强调国家安全、社会风险和责任担当。对2024年1月语料的进一步分析表明,习主席在表达对人民的关切时,兼具温度和力度。

八、美国总统涉华话语分析

笔者使用“China”作为检索词,在The American Presidency Project语料库中下载了包含该检索词的三位美国总统的所有话语语料,并创建了三个独立的总统话语语料库:奥巴马中国话语库(2012,11—2017,1)特朗普中国话语库(2017,1—2021,1)和拜登中国话语库(2021,1—2024,6)。然后,将这三个语料库分别导入LancsBox中,使用“China”作为节点词,运用软件的GraphColl功能,设定节点词的跨距范围为左右5个词,提取了“China”的搭配词。

结果显示,奥巴马话语中的“China”一词经常与“co-operation”“agreement”“relations”“constructive”“commitment”“together”“world”“influence”等词语共现。这反映出在奥巴马的话语中,中国被定位为“全球合作伙伴”。特朗普话语中的“China”搭配词明显转向对抗框架,包括“virus”“tariffs”“ban”“outbreak”“stopped”“never”等词语。特朗普的话语侧重于贸易摩擦,中国被描绘成“挑战者”。此外,“China”常常与“I”“we”“our”等词语共现,揭示出其强烈的个人主义话语特征。与奥巴马时期的合作叙事相比,特朗普对中国的描绘明显更加单极化和情绪化。在拜登话语中,“China”的高频搭配词包括“competition”“strategic”“challenges”“market”“military”“outcompete”等,可以看出拜登描绘中国形象时的叙述框架是以“竞争性共存”为基调的。总体来看,三位总统涉华话语反映了美国政府对华叙述的历时演变:从奥巴马时期的全球合作伙伴,到特朗普时期的贸易挑战者,再到拜登时期的战略竞争对手。

九、结语

本文通过将习主席话语语料库与三位美国总统中国话语语料库进行比较,运用LDA主题建模、用法波动分析等语料库方法,揭示了中国形象在不同政治主体中的差异化建构模式及其背后动因。习主席的讲话围绕着发展、改革、合作与和平等主题,展示了以人为本的发展理念,构建了一个开放、负责任的大国形象。相比之下,三位美国总统的涉华话语揭示了从“合作伙伴”向“战略竞争对手”的转变。总体而言,中美之间关于“中国形象”的叙事差异反映了不同的国家利益和意识形态。

参考文献:

- [1] Baker, P., Gabrielatos, C., & McNery, T. Discourse analysis and media attitudes: The representation of Islam in the British press[M]. United Kingdom: Cambridge University Press, 2013.
- [2] Fairclough, N. Analysing discourse: Textual analysis for social research[M]. United Kingdom: Routledge, 2003.
- [3] Law, K.Y. The Portrayal of Chinese top Party-State Leaders in the American Mainstream News Media, 1978-2008: Thirty Years of Reform Have Not Changed China Yet?[J]. China Report, 2009, 45(1), 35-51.
- [4] 许峰,高意. 话语分析视角下中国国家生态形象自塑研究——以习近平主席外交话语为例[J]. 中国地质大学学报(社会科学版),2023,23(05).

基金项目:2024年度杭州市哲学社会科学规划常规性课题“基于语料库的中国形象历时对比研究——以中美两国领导人重要讲话为例”(编号:Z24JC011)。

作者系浙江理工大学科技与艺术学院副教授