

# 语料库驱动下三江源国际传播的话语分析

杨 洋, 戴延红

青海大学基础部, 青海 西宁

收稿日期: 2021年11月20日; 录用日期: 2021年12月23日; 发布日期: 2021年12月30日

## 摘 要

随着中国第一批国家公园正式设立, 三江源再次成为世界关注的焦点。如何推进三江源国际传播, 提高其国际知名度和美誉度应受到高度重视。本文基于语言学视角, 采用语料库驱动与批评话语分析相结合的方法, 通过对AntConc软件检索超大型语料库NOW中75篇三江源国际报道, 分析了三江源国际传播的重要时间节点、区域分布、英媒关注热点、话语构建方式以及海外受众接受度, 并根据研究发现提出了建立积极的传播话语体系、搭建国际传播平台以及建设平行语料库的三江源国际传播的话语策略。希望本研究为三江源国际传播课题提供不同的理论视角和数据支撑。

## 关键词

三江源, 国际传播, 批评话语分析, 语料库驱动

# Study on the International Dissemination of Sanjiangyuan Based on Corpus-Driven Discourse Analysis

Yang Yang, Yanhong Dai

Department of General Education and Research, Qinghai University, Xining Qinghai

Received: Nov. 20<sup>th</sup>, 2021; accepted: Dec. 23<sup>rd</sup>, 2021; published: Dec. 30<sup>th</sup>, 2021

## Abstract

With the official establishment of the first-batch national parks in China, Sanjiangyuan has attracted worldwide attention once again. Thus, how to popularize Sanjiangyuan in the world and improve its international reputation is worth exploration. This paper starts from a linguistic perspective and performs combined corpus-driven and critical discourse analysis. AntConc software is used to retrieve 75 Sanjiangyuan international reports in the super-large corpus NOW. On this basis, important time nodes, regional distribution, presenting the focus, discourse construction of English media and audience acceptance in terms of Sanjiangyuan international dissemination are

examined. Based on the analysis results, the discourse strategy of Sanjiangyuan international dissemination is put forward, *i.e.*, establishing a positive communication discourse system, an international communication platform and a parallel corpus. It is hoped to provide different theoretical perspectives and qualitative data support for Sanjiangyuan international dissemination.

## Keywords

Sanjiangyuan, International Dissemination, Critical Discourse Analysis, Corpus-Driven

Copyright © 2021 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

“三江源”是指长江、黄河、澜沧江的发源地，被誉为“中华水塔”。作为我国重要的生态安全屏障和高原物种资源库，其生态保护价值对全国乃至全球都具有重大意义。习总书记走访青海时强调，三江源保护应是青海生态文明建设的重中之重，青海要承担好维护生态安全、保护三江源、保护“中华水塔”的重大使命。

从建立国家公园体制试点方案颁布，到青海省委“一优两高”重大战略部署，再到三江源国家公园正式设立，我国在推进生态文明建设中的每一次举措都引起了不同领域的关注，其中也不乏一些国内的学者。以中国知网为例，检索“三江源”主题词，得到自 1986 年至 2021 年 2857 篇中文文献，其中，2017 年至 2020 年年发文量都超过 200 篇，并成逐年上升趋势(见图 1)。

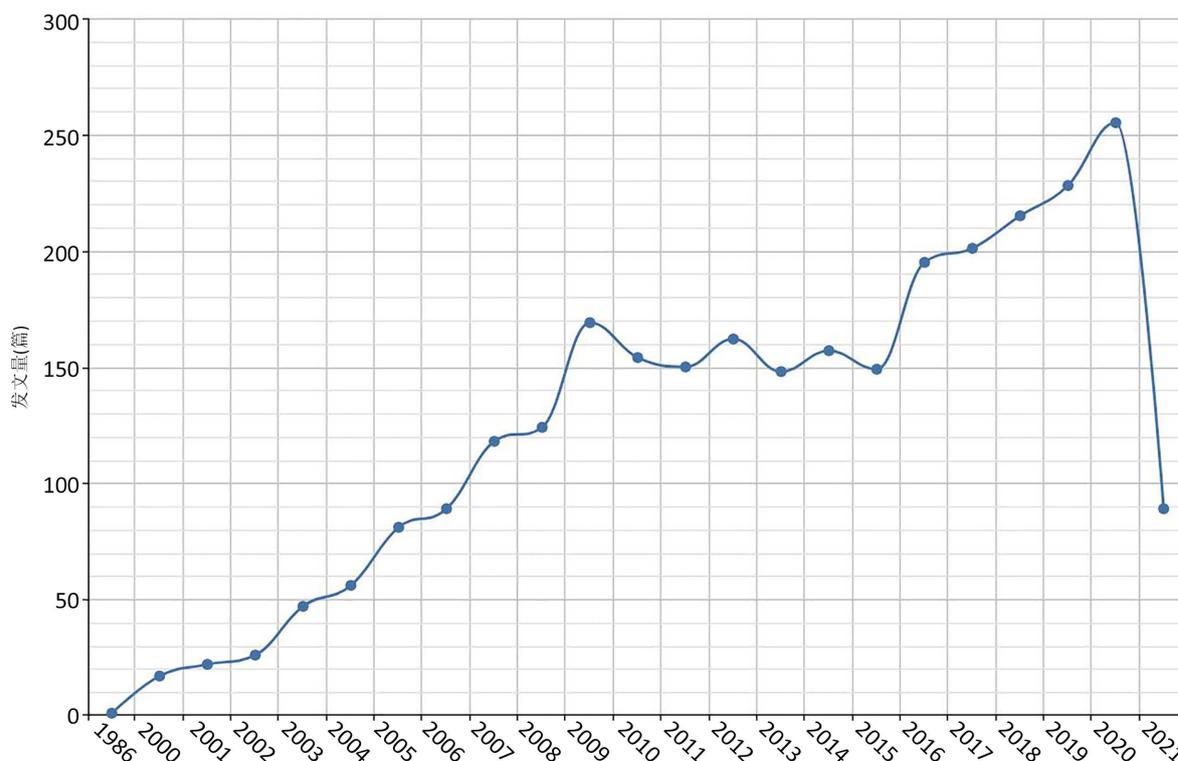


Figure 1. Quantity of Sanjiangyuan themed papers published from 1986 to 2021

图 1. 1986~2021 年“三江源”主题论文发文量

在已发表的文献中, 以环境、草学、生态等自然科学领域研究为主。在社会科学领域, 以 CSSCI 为例, 键入主题词“三江源”, 可检索到自 2000 年至 2021 年 287 篇文献, 主题词共现频次主要集中在生态保护(40), 国家公园(25), 生态环境(25), 生态移民(20)等领域, 其中, 2012 年发文量 30 篇, 关注度最高。但如果进一步检索“三江源 + 海外传播”“三江源 + 国际传播”“语料库 + 三江源”“三江源 + 话语分析”都未找到相关文献, 说明目前基于语言学和传播学视角的三江源研究鲜有人关注(见图 2)。

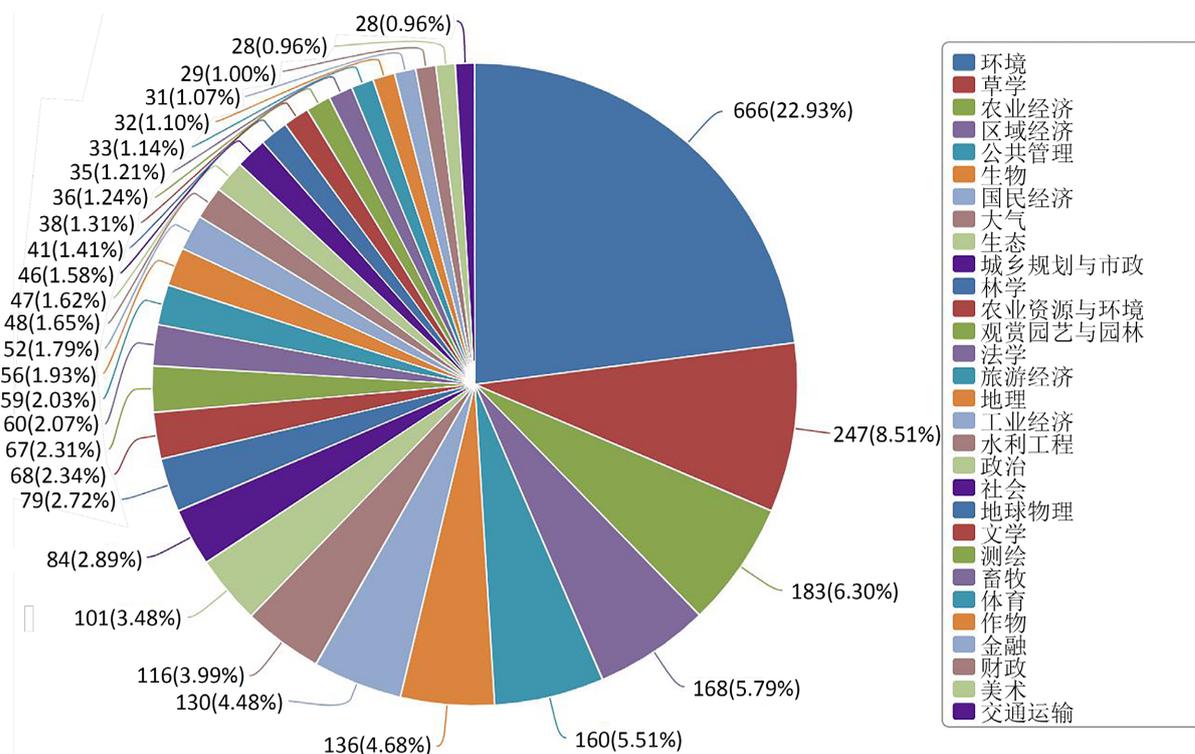


Figure 2. Discipline distribution of Sanjiangyuan themed papers from 2000 to 2021

图 2. 2000~2021 年“三江源”主题论文学科学科分布

地域形象是一个地区的地理商标, 地域形象的知名度和美誉度是地域价值的体现, 同时也是一个地区对内塑造自己、对外展示自己的载体[1]。随着三江源国家公园的正式设立, 除了对其环境、生态等自然科学领域的研究, 对如何推进三江源国际传播, 扩大三江源国际影响力的社会科学领域思考也应受到高度重视。

## 2. 研究思路和语料收集

本文基于语料库驱动, 结合批评话语分析法进行研究。批评话语分析(CDA)是由 Fowler 等人在 *Language and Control* 一书中提出, 旨在揭示话语、权力及意识形态之间的内在联系[2]。它不仅关注话语和篇章, 更聚焦于社会和文化。CDA 主要以定性分析为主, 相对片面和主观的研究方法, 使得其研究结论失去了部分可信度, 引起了学术界的一些争议和批评。语料库语言学是以真实语言数据为研究对象, 从宏观角度对大量的语言事实进行分析, 从中寻找语言的使用规律[3]。研究者可通过收集大量语料, 以词频统计、词汇的搭配强度、索引呈现、类连接等分析工具为批评话语分析在意识形态的抽象阐释方面提供事实依据, 以寻求真实有效的研究结果。Hardt-Mautner 是第一位倡导将批评话语分析和语料库语言学相结合的学者, 他认为语料库强大的技术能够支持对大量语篇进行定性研究[4]。因此, 语料库批评话

语分析法不但可以保证研究结果的客观性和可验证性, 而且还可以关注到语篇的更高层次, 如其社会属性和文化属性。

为确保语料的丰富性和研究的科学性, 本文依托超大型在线监控语料库——NOW 语料库(News on the Web)进行检索。NOW 语料库包括全球 20 个英语国家的 2000 多个网站和报报纸, 库容近 130 亿词。具体操作如下: 一提取主题为“sanjiangyuan”的生文本语料, 二利用 Emedior 和 CorpusWordParser 文本编辑器将生文本进行清洁和标注处理, 按年代分类自建三江源国际传播话语语料库 6 个子库, 三使用 Ant Conc 3.5.9 语料库检索工具对自建语料库进行统计和分析。

### 3. 三江源国际报道文本分析

#### (一) 三江源国际报道数量分析

通过检索“sanjiangyuan”在 NOW 语料库中的频次, 发现自 2013 年至 2020 年三江源在英媒报道中共出现 75 次(见表 1), 自 2018 年以来报道数量呈现逐年上升的趋势, 2019 年报道数量最多。从整体趋势看, 自 2016 以来英媒对三江源的关注度逐年递增, 2019 年关注度最高。其原因主要是 2016 年我国政府正式启动了三江源国家公园体制试点, 明确了五年建成国家公园的工作目标, 此项举措受到了世界的关注。但相比在 CCL 中文语料库中检索“三江源”词频的结果(500 次)来说, 其关注度呈现出国内关注度高、国外关注度低的总态势。

**Table 1.** Capacity and frequency of Sanjiangyuan related reports in the NOW corpus

**表 1.** NOW 语料库中三江源报道库容及频次

年份	库容(百万)	频次
2013	401.5	5
2016	1531.5	11
2017	1746.5	8
2018	1569.1	11
2019	1987.5	26
2020	2607.8	14
		总计 75

#### (二) 三江源国际报道区域分布分析

三江源的国际报道热点地区为印度、美国、香港、新加坡、菲律宾、爱尔兰、加纳、加拿大等。总体来看, 北美洲、欧洲、亚洲英语媒体对三江源报道较多, 非洲相对较少, 大洋洲和南美洲未见报道(见表 2)。结果说明, 三江源国际传播的辐射范围以中国及其周边地区为主, 呈离中国近报道多, 离中国远报道少的总趋势; 但在北美地区报道居多, 以美国报道为主。结合文本分析, 得到两个主要原因: 一是美国媒体对比报道了三江源国家公园与美国黄石公园在国家公园制度政策、生态研究和经济价值等方面的异同, 二是美国媒体介绍了三江源国家公园建设过程中中美双方进行的经验借鉴和学术交流活动。

**Table 2.** Regional distribution and frequency of international reports on Sanjiangyuan

**表 2.** 三江源国际报道区域分布及频次

地区	库容(百万)	频次
印度	1314.3	24
美国	4182.5	11

Continued

香港	64.3	7
新加坡	502.2	7
菲律宾	387.8	4
爱尔兰	938.7	4
加纳	153.1	3
加拿大	1498.6	1

总计 61

#### 4. 三江源国际传播的批评话语分析

##### (一) 三江源国际传播的话语聚焦点分析

主题词统计是研究语篇和语域特征的有效方法,它能够在语篇层面上反映词汇的显著分布和文本主题之间的关系,即语料库中凸显的话语事件和话语主题。主题词表可以使研究者有依据地判断语篇的意义和说话者的观点。在生成的主题词词表中,越靠前出现的词关键值越高,关键值高说明该信息在文本中出现的次数多、意义大。值得注意的是,和高频词不同,主题词分析的关键因素不是绝对频率,而是通过对比观察语料库和参考语料库,得到语篇层面上词语的显著分布和文本的凸显主题。研究表明主题词的数量与参照语料库的大小成正比,参照语料库要至少大于观察语料库五倍以上时,对比两个语料库产出的主题词才最大限度地具有有效性和稳定性[5]。为此,笔者以 BNC 语料库的主题词为参照,生成自建语料库的主题词词表,进而观察三江源国际传播中的聚焦点。

笔者发现,除去在任何文本中都高频出现的介词、连词、助动词、代词等,在自建语料中 china, tibetan, national, qinghai, park 等词的关键性表现出凸显的行为(见表 3)。其中,名词或名词短语的使用多于动词或动词短语的使用。语篇中的名词或名词词组常常承载着新的或有效的信息,它们本身就能表达特定的话题,从而带动整个语篇的话题走向。

**Table 3.** Analysis of keywords of Sanjiangyuan international dissemination

**表 3.** 三江源国际传播的话语主题词分析

序号	频次	关键性	关键词	序号	频次	关键性	关键词
1	230	1249.59	china	16	57	302.47	new
2	133	717.79	tibetan	17	57	302.47	world
3	107	575.51	national	18	56	297.02	snow
4	106	570.04	qinghai	19	54	286.13	first
5	104	559.1	park	20	50	274.06	yushu
6	96	515.37	local	21	51	269.79	species
7	84	449.79	conservation	22	48	253.46	three
8	82	449.56	sanjiangyuan	23	46	242.58	land
9	74	395.19	province	24	46	242.58	wildlife
10	71	378.82	people	25	44	231.7	nature
11	69	367.91	area	26	44	231.7	protection
12	68	362.45	government	27	44	231.7	river

Continued

13	67	356.99	plateau	28	43	226.27	animals
14	63	335.18	chinese	29	43	226.27	ecological
15	62	329.73	region	30	43	226.27	more

词频统计是语料库检索中最基本的统计方法, 这种方法的优势在于研究者可以直观地观察语言的显著特征。在新闻报道中, 通过高频词的运用, 即对某个词或某个概念的重复使用, 能给读者留下深刻印象, 同时作者也以委婉的方式表达其倾向态度与意识形态。笔者在自建语料库 6 个子库中分别进行高频词检索, 每个子库选取前 100 个高频词, 按年份归类生成 6 个词云图, 获取三江源国际传播的特征词汇, 从而了解英媒关于三江源报道的概貌。笔者分析发现, 2013 年英媒关注了三江源地理位置和地域特点的描述, 2016 年至 2017 年聚焦到三江源生物物种、当地居民、政府行为等, 2018 年开始初步涉及了三江源文化艺术、投资发展、国际化进程等社会人文领域, 2019 年之后, 主要围绕建立三江源国家公园和生态保护机制展开报道。总体来说, 英媒对三江源的聚焦点是动态、变化的, 报道由浅入深、逐步丰富, 并持续关注着三江源国家公园及我国生态文明的建设和发展进程(见图 3)。

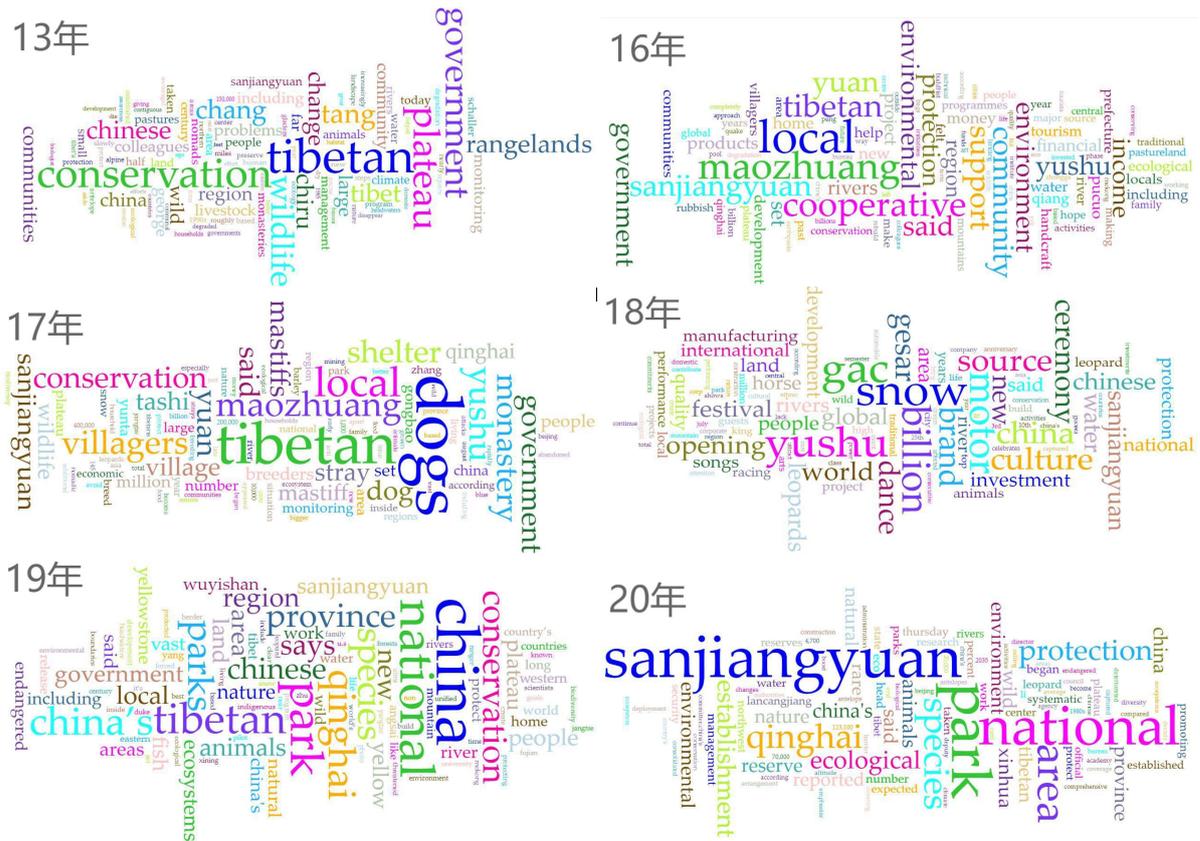


Figure 3. Cloud map of high-frequency words related to Sanjiangyuan international dissemination from 2016 to 2020  
图 3. 2016 年~2020 年三江源国际传播话语高频词云图

由表 3 和图 3 可以看出, 除去在任何语料库都高频出现的词, 本语料库中检索到的高频词和主题词高度吻合, 说明了英媒对三江源的主要关注点、三江源相关的重要议题以及英媒的认知倾向三者高度一致。

## (二) 三江源国际传播的话语建构分析

词语搭配是对于某话语进行建构的典型方式, Firth 认为搭配词之间是一种“结伴关系”, 即词的意义与其结伴出现的词之间密切相关[6]。词与词之间的互相依存和互相期待行为表明二者的共现概率与其语义粘合程度成正比。共现的频率越高, 两者搭配提供的语境意义也就越明确[7]。为了提炼语料库中三江源的相关搭配, 本文采用互信息分析(MI)进行搭配检索。一般来说 MI 高于 3 就被认为是强搭配的象征。笔者以“sanjiangyuan”作为节点词, 观察跨距为正负 2 的全部搭配, 得到 84 组搭配的 MI 值大于 3。其中, 有 10 组搭配的 MI 值大于 8, 也就是说以下搭配是三江源国际传播中话语构建的典型方式(见图 4)。

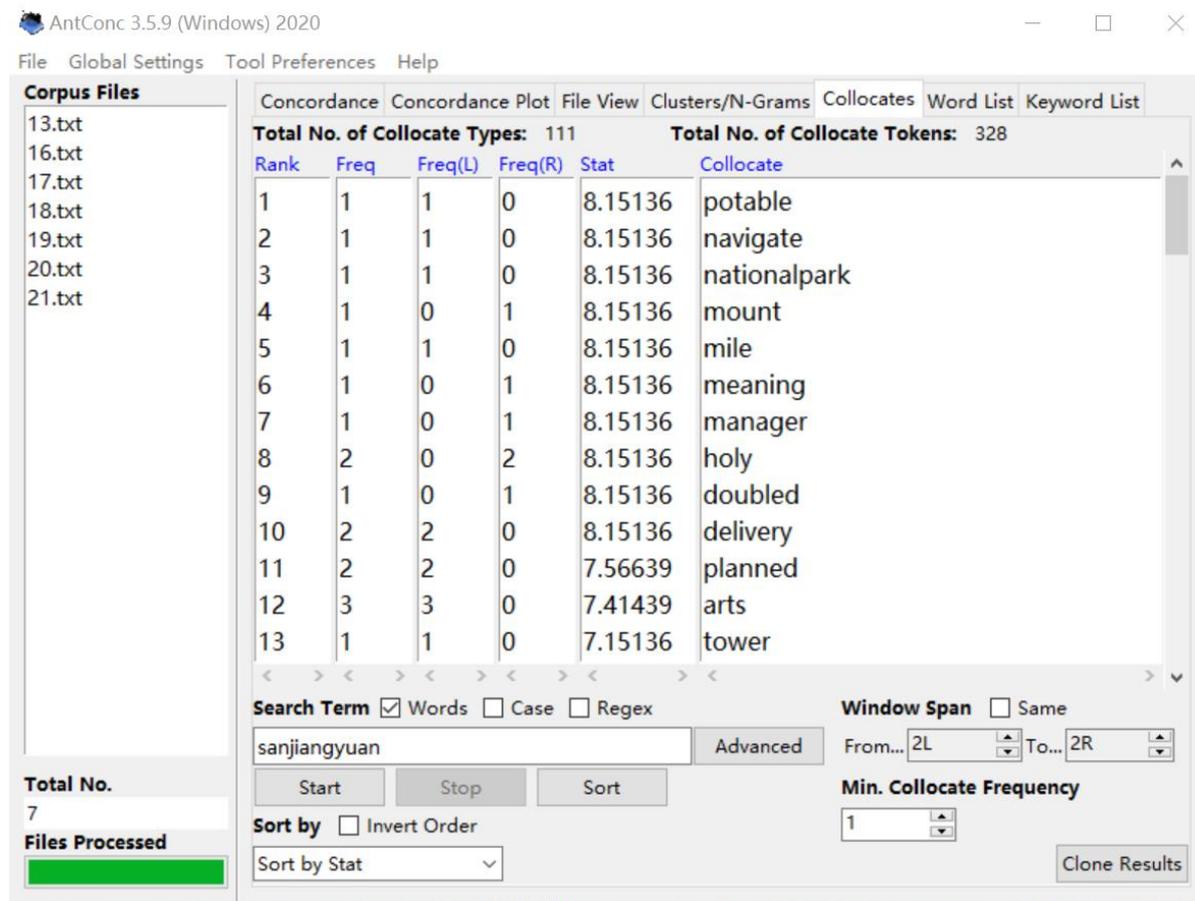


Figure 4. Collocability and frequency of Sangjiangyuan  
图 4. “三江源”主题词搭配强度与频次

结合数据和报道原文分析, 英媒对三江源的话语描述有以下特点: 对三江源地理地貌、环境保护、国家政策描述较多, 对当地民俗文化、经济发展、合作创新描述较少。但这种话语建构方式未能将三江源的当地特色和生态优势进行深入宣传, 使受众对其知之甚少。

## (三) 三江源国际传播的话语情感分析

在 AntConc3.5.9 中进行主题词检索, 可以提取共显语境和检索行内容, 采用贝叶斯算法或人工甄别法可对话语进行情感归类, 从而得出英媒对此主题的情感态度和情感倾向结果。笔者提取了以“sanjiangyuan”为主题词报道的共显语境和检索行, 发现英媒对其话语有正向、负向和中性的情感态度, 其中正面评价多于负面评价, 以下是典型的正面评价和负面评价检索行(见表 4):

**Table 4.** Analysis of audiences' emotion attitude in Concordance line**表 4.** 检索行受众情感分析

序号	检索行	情感倾向
1	what will be the world's largest land-based national park, Sanjiangyuan. While there, the team met and spoke with ethnic Tibetan community rangers who expressed their pride in contributing to the creation of the park	正向
2	are all enthusiastic and dedicated in the long-term. "Elsewhere in the Sanjiangyuan region, according to Peng Kui, seven to eight communities have set up their	正向
3	the ecological environment of the Sanjiangyuan area has greatly improved in recent years, with ecological and water conservation functions consolidated, and wetland areas expanded	正向
4	one of four major prefectures in the Sanjiangyuan area, said in June. "The water quality of the Yangtze River in Yushu is generally maintained at Level I	正向
5	a healthy population of snow leopards in the region. Moreover, it means the Sanjiangyuan region has a high conservation value in biological diversity besides headwaters	正向
6	Qinghai Province issued guidelines for promoting the establishment of Sanjiangyuan National Park on May 9, which include the comprehensive deployment and systematic arrangement for the preliminary work of promoting the official establishment of the park	正向
7	Berkeley, who has toured the Qinghai pilot park, called Sanjiangyuan. China has previously undertaken vast resettlement programs to clear land for large infrastructure	正向
8	These have a better chance of survival, but they have damaged the ecosystem of the Sanjiangyuan River Basin and pose a great threat to local fish stocks and water habitats.	负向

由表 4 可见, 英媒对三江源国际传播的话语构建大多为正向、积极的, 肯定了我国三江源国家公园建立的重大成果, 有力地证明了我国各级政府监管得力, 保障了政策落实到位, 使三江源生态环境得到了改善和优化, 为全世界生态事业做出了重要的贡献。但其中也存在英媒对三江源话语建构中的负面情感。以表 4 中第 8 条为例, 报道原文大意是当地一些居民向江河中放生非本地原产鱼群, 使本原地生物受到外来物种威胁, 破坏了三江源的生态系统和生态链条。这种负向情感的出现, 除了客观和理性的思考, 另一个主要原因是英媒缺失了对三江源民风民俗、宗教文化的深入认知。其实放生是在当地藏民族中出现的一种特殊文化现象, 藏民族本着尊重自然、尊重生命的生态意识和人与自然和谐共生的朴素生态理念, 世代代的维护着三江源的生态安全。因此, 话语建构中负向情感的出现也会源于认知中的不全面和不对等。

## 5. 三江源国际传播的话语策略

批评话语分析人士认为, 话语权是一种与话语相关的社会权利。它是一个群体对另一个群体的行为和思想的控制。如果滥用话语权, 就会形成一种消极的支配权力。值得思考的是, 海外公众对三江源的认知主要源自英文媒体的传播话语, 如果这些媒体在传播过程中对三江源形象进行了曲解、误传或者刻意扭曲, 那么公众对三江源的认知就是不完整的、有偏差的, 甚至是混淆黑白的, 以至于影响到三江源的国际知名度和美誉度。因此, 我们应积极谋划、向世界传播积极、正面的三江源国际形象, 从内容生产、评价体系、技术创新和人力支撑等方面为着力点, 思考三江源国际传播的话语策略, 讲好中国故事和青海故事。

### (一) 建立积极的三江源传播话语体系

在三江源国际传播过程中, 要正确理解三江源的生态价值和民俗文化, 建构积极的三江源传播话语体系。具体而言, 传播过程中要倡导积极、公开、健康的价值观, 用更贴切的话语和方式阐释三江源民俗文化和人与自然和谐相处的绿色生态理念, 充分发挥情感传播认同力, 增强海外民众产生情感共鸣,

使三江源报道既通俗易懂, 又不失真实客观。既传播三江源的生态价值、地域文化和经济建设, 也传播我国生态文明进程中的新理念、新举措和新突破, 消除海外媒体和民众对三江源认知中的偏见和误解, 以提升三江源的海内外影响力和美誉度。

## (二) 搭建三江源国际传播平台

随着《三江源国家公园体制试点方案》的印发和青海省委省政府“一优两高”重大战略部署的实施, 三江源国际传播已然成为了生态文明建设和发展中不可或缺的部分。要积极利用国家政策带来的大好局面, 大力构建政府主导、社会协调、人民参与的多层次、全方位、立体式沟通体系, 打造全球治理、多媒体融合的国际新闻报道。各级政府拓展顶层设计维度, 创新三江源国际传播机制, 充分发挥语言学者和记者的“桥梁”作用, 充分收集海外受众的反馈, 改进三江源国际传播效果, 加强传播内容的客观性和准确度。深入研究海外社交平台运行规律, 拓宽传播渠道, 重视民间传播、第三方传播和新媒体传播。

## (三) 建设三江源英汉双语平行语料库

英汉双语平行语料库是一个英汉双语一一对应, 科学取样和加工的大规模电子文本库语料资源。语料库可提供大量有据可寻的真实语料, 其权威性、动态性、真实性是一般辅助工具和人力无法替代的。建库过程中将对三江源资源进行整合和归类, 对相关文本及术语翻译进行规范和统一, 为三江源英文翻译提供良好范式。建成的语料库提供在线检索功能, 不仅为三江源的国际传播, 也为政府政策宣传、生态文明教育以及其他学术领域提供规范、翔实的素材和切实可靠的数据支撑。

## 6. 结语

笔者在语料库驱动下对三江源国际传播进行话语分析, 发现自 2016 以来英媒对三江源的关注度逐年递增, 但相比国内关注度而言, 呈现出国内关注度高国外关注度低的总态势。传播的热点国家以北美洲、欧洲、亚洲等为主; 英媒对三江源国际传播的话语构建大多为正向、积极的, 主要肯定了三江源国家公园的建立和我国生态文明建设的重大成果, 但也存在一些负面情感, 批评了影响和破坏生态系统的做法。要提高三江源的国际知名度和美誉度, 树立其良好的国际影响力, 还需要在三江源国际传播策略上进行更多的思考和实践。

受时间、人力和学科等制约因素, 笔者只使用了 NOW 语料库作为数据来源, 对其他英语语料库和对非英语国家和地区的语料未做文本研究; 另外笔者对三江源国际传播策略问题未作深入探讨, 如若将此问题进行语言学、传播学、心理学、认知科学等跨学科研究, 相信会对本研究带来更多的启示和参考。

## 注释

本文中的英媒特指海内外英文媒体。

## 基金项目

国家社会科学基金项目“三江源国家公园宣传教育功能研究”(18BSH085)。

## 参考文献

- [1] 韩喜玉. 多媒体时代青海形象的构建与提升研究[J]. 新青海研究, 2020(9): 12-18.
- [2] Wodak, R. and Meyer, M. (2009) *Methods of Critical Discourse Analysis*. Sage Publication, London.
- [3] 杨惠中. 语料库语言学导论[M]. 上海: 上海外语教育出版社, 2002.
- [4] Hardt-Mautner, G. (1995) *Only Connect: Critical Discourse Analysis and Corpus Linguistics*. UCREL Technical Papers, Vol. 6.
- [5] Baber-Sardinha, A. (2000) Comparing Corpora with Word Smith Tools: How Large Must the Reference Corpus Be? *Proceedings of Workshop on Comparing Corpora*, 9, 7-13. <https://doi.org/10.3115/1117729.1117731>

- [6] Firth, J. (1957) A Synopsis of Linguistic Theory, 1930-1955. In: *Selected Papers of J.R. Firth 1952-1959*, Longmans, Green and Co., Ltd., London and Harlow.
- [7] 张淑静. 语料库在批评话语分析中的应用[J]. 郑州大学学报, 2014(5): 131-133.