

A Corpus-Based Discourse-Historical Analysis of China-related Coverage of COVID-19 in Chinese and American Mainstream Media

Lou Lingling*

School of English Language, Zhejiang Yuexiu University, Shaoxing, Zhejiang, P.R. China

ABSTRACT

Through employing Ruth Wodak's discourse-historical approach, one of the most influential schools of Critical Discourse Analysis, as the guiding theory, the paper conducted a corpus-based study on China-related coverage of COVID-19 in Chinese and American mainstream media. With the help of the various functions of the corpus such as concordance, word clusters, keyword list, and so on, both qualitative and quantitative research methods are adopted to explore the changes in the focus of Chinese and American media in different periods and their respective discursive strategies. It is hoped that by combining synchronic research with diachronic research, this paper can effectively explore the deep-rooted reasons behind such differences and changes, and provide insights for the high-quality construction of China's national image.

Key words: Discourse-historical approach, Corpus, COVID-19, News coverage

HOW TO CITE THIS ARTICLE: Lou Lingling, A Corpus-Based Discourse-Historical Analysis of China-related Coverage of COVID-19 in Chinese and American Mainstream Media, J Res Med Dent Sci, 2022, 10(2): 645-656

Corresponding author: Lou Lingling
e-mail ✉: rolin.0-0@qq.com
Received: 14/02/2021
Accepted: 28/02/2022

INTRODUCTION

Research background

At the beginning of 2020, a sudden outbreak of COVID-19 swept across China, attracting the attention of people around the world. As a public health emergency, the epidemic has a wide impact, therefore, the public is in urgent need of a large amount of information to guide their personal lives and stabilize their mentality. The public's interest in information related to the epidemic exceeded that of any other information in the same period. According to the survey data, 47% of the respondents follow the epidemic information every day, 43% pay close attention to the epidemic information all the time, and the remaining 10% pay attention to the epidemic-related news at least 2-3 times a week on average. All kinds of epidemic information, in different ways and carriers, affect people's emotions and lives. Understanding the dynamics of the epidemic has become an important part of public life in this special period. People pay attention to the changes of epidemic situations released through different channels and follow updates on various topics related to anti-epidemic and epidemic prevention through different carriers.

Scientific communication is of great value in epidemic control. As the most permeable and disseminating media, news has the dual attributes of objectivity and sociality

and is an important source for the public to obtain information. The media dissemination of news reports is an important means to shape the national image [1], besides; it is also a powerful carrier of ideological output [2]. In the process of news reporting, journalists tend to express ideology implicitly, which exerts a subtle influence on readers' cognition and judgment. Therefore, the discourse mode of a news report and the ideology hidden behind the discourse are worth discussing. This paper collected the news reports of mainstream media in China and the United States in the first three months after the outbreak of the epidemic and analyzed the news reports of the two countries from multiple perspectives by combining the methods of corpus linguistics (CL) and discourse-historical approach (DHA) to explore the ideological differences behind the discourse.

Research significance

Taking DHA as the theoretical framework and CL as the analytical tool, the paper tried to make a comparative analysis of the reports on China's COVID-19 epidemic at different stages in the mainstream media in China and the United States, find out the deep-seated reasons behind the differences, so as to provide hints for China to build a good national image.

In view of this, the significance of this study is as follows.

First, in terms of theory, this paper used DHA to analyze news discourse, which enriches the application scenarios of DHA to a certain extent and provides a new research perspective of coronavirus news discourse at the same time. In addition, this study is not limited to a single

research method, but innovatively combines quantitative research method with qualitative research method to analyze news discourse, making the analysis results more objective and convincing; what's more, it is a new attempt to integrate diachronic study and synchronic study into the discussion of the research results, which is quite meaningful for the future research.

Second, on the practical side, analyzing China-related reports on COVID-19 by mainstream media in China and the United States can reveal the ideology behind the language, allowing people to go beyond text reading and gain a more comprehensive understanding of this public health emergency of international concern.

In addition, people can clearly see what kind of national image the mainstream media in China and the United States try to construct respectively. More importantly, this study reveals how these national images are constructed through discussing the discursive strategies chosen by the mainstream media in the two countries, thereby providing suggestions for the optimization of China's national image construction.

THEORETICAL FRAMEWORK

Discourse-historical approach

Critical Discourse Analysis (CDA) is a linguistic trend that developed from the late 1970s to the early 1980s and originated from critical linguistics. Later, CDA has attracted much attention because of its "interdisciplinary vision, strong political concern and its purpose of social criticism through linguistic analysis" [3]. The representatives mainly include several anti-mainstream linguists and sociolinguists in Britain, France, and Germany, such as M. Foucault, N. Fairclough, G. Kress, R. Fowler, and so on. It is a social-oriented discourse analysis method based on Halliday's Systemic Functional Linguistics [4].

As a branch of linguistics or a method of discourse analysis, the linguistic analysis of CDA is mainly based on various theories of modern linguistics, which emphasizes the connection between language and society. What differs from the traditional discourse analysis is that CDA regards language as a social practice and emphasizes the comprehensive analysis of the social and cultural aspects of discourse. On the one hand, it reveals the relationship between language, power, and ideology from the perspectives of linguistics, sociology, psychology, and communication through the forms of language expression; On the other hand, it also studies the reaction of discourse to social ideology [5]. Therefore, It can be seen that CDA views language as a power that intervenes in society, and believes that language not only reflects social reality but also reacts to society.

The discourse-historical approach (DHA), which was proposed by Ruth Wodak, is one of the most important branches of CDA. DHA, also known as the Vienna School of CDA, involves knowledge of linguistics, philosophy, politics, history, sociology, etc. It belongs to the forefront of interdisciplinary research and has an important

influence in the field of CDA. According to Reisigl & Wodak (2001), DHA mainly analyzes discourse from three aspects: (1) determine the content or theme of a specific discourse, for example, racism; (2) analyze the discursive strategies used in the discourse; (3) analyze the language means used in the specific context. This paper will try to analyze the news discourse during the epidemic from the first two aspects.

Discursive strategies are a series of practical behaviors adopted by people to achieve specific social, political, psychological, or linguistic goals [6,7]. These strategies include the nomination strategy (gaining support for an argument by constructing in-group and out-group identities), the predication strategy (assigning positive or negative attributes to specific social actors), the argumentation strategy (confirming positive and negative descriptions), the perspectivization strategy (through which speakers can express their views on the content of the discourses), the intensification and mitigation strategy (through which the discourse can reproduce the reality sharply or obscurely) [8]. The corpus studied in this paper mainly involves the nomination strategy, the predication strategy, and the perspectivization strategy.

Corpus-based discourse-historical approach

In the development process of CDA as well as DHA, the objectivity and credibility of its analysis have been questioned due to the small number of research texts, the randomness of corpus selection, and the lack of sufficient and systematic linguistic evidence for the interpretation of discourse. To make up for the defects of the research method, corpus-based CDA has gradually emerged. The approach combines the advantages of quantitative-based CL and qualitative-based CDA, so as to avoid the subjectivity and one-sidedness of researchers, and improve the objectivity, credibility, and persuasiveness of the research. At the same time, through the data processing of the corpus, researchers can pay attention to language phenomena that were previously ignored and focus on analyzing high-frequency information, which provides a large amount of corpus basis for people to explore the realistic construction of language and reveal the ideology behind it.

The aid of corpus greatly enriches the research of CDA and DHA. In the past 20 years, relevant research on CDA based on corpus has been in the ascendant. Scholars mainly use the word frequency, concordance, collocations, word clusters, and other functions of corpus to identify representative research questions, and then conduct detailed analysis and in-depth explanation based on the perspective of CDA, deconstructing the meaning and rules of discourse according to linguistic features. For example, Baker et al. (2008) and Gabrielatos & Baker (2008) were the first to carry out diachronic analysis of refugee and immigration issues in the news under the framework of DHA and CDA and integrated discursive strategies with the description of social, cultural and historical background into the interpretation, which further improved the reliability of the illustration of the

results. Since then, Baker's team has carried out a series of diachronic studies on British media discourse [9-15]. The above research broke through the limitation of emphasizing synchrony and neglecting diachrony (history) in CDA research and carried out a more comprehensive analysis on the construction of specific topics and images in news, which not only gave consideration to the completeness of corpus quantification but also fully presented the depth and systematization of CDA framework analysis.

In contrast, domestic scholars' research in this area is still relatively limited. Jiang Jinhui & Lin Zhangmei (2018) used five mainstream American newspapers to establish a corpus of anti-corruption topics and analyzed the social and historical reasons for the formation of reports and the ideology behind the news discourse of western mainstream media by using the data of high-frequency words, collocation and index lines in the corpus. Wang Peijuan et al.(2020) explored the regional and temporal distribution of the English media's attention to the Shanghai Cooperation Organization through the self-built corpus and explored the reasons behind it. There are also scholars who have conducted in-depth research on corpus-based discourse analysis from the DHA or diachronic perspective [16-19], but such research is still in its infancy, and the research topics still need to be further expanded.

The introduction of the corpus provides a broad space for CDA. Although the research on news discourse is deepening in depth and expanding in scope, at present, it mainly focuses on the related topics of economy, military, and politics, and not much research has been done on Chinese health care coverage. In addition, few scholars have conducted corpus-based discourse-historical studies on the news coverage of the ongoing COVID-19 epidemic. This study will establish two special corpora, and use corpus linguistics as an analytical tool to carry out discourse-historical analysis on the discourse of the epidemic reports in Global Times and CNN, in an attempt to interpret the ideology hidden behind such reports and explore the relationship between language, power, and ideology.

RESEARCH METHODOLOGY

Research question

By analyzing the mass media reports on the epidemic in China by Global times and CNN with the help of a corpus-based discourse-historical approach, the author aims to answer the following three questions:

How have the reporting themes of China-related epidemic reports changed at different stages?

What are the differences between Chinese and American mainstream media's coverage of China-related epidemics?

Why is there such a big difference in the news reports of the mainstream media in China and the United States? What are the deep-seated reasons behind it?

Corpus data

In this study, reports related to the COVID-19 outbreak in China from Global Times and CNN between January 1 and March 31, 2020 were selected and two small corpora were built. The reason for choosing the reports of these two media is that Global Times and CNN are respectively one of the most authoritative media in China and the United States. Known as the world news leader, CNN was called "the 16th member of the UN Security Council" by former UN Secretary-General Boutros Boutros-Ghali. Renowned for its timeliness, exclusiveness, and coverage of major and breaking news events around the world, it has become one of the most influential news media in the United States and around the world. Global Times, an English-language Chinese newspaper under the People's Daily, serves as an important window for China to understand the world and the world to understand China. Therefore, they not only have a broad readership and influence but also can well represent the political positions and ideologies of China and the United States respectively.

In terms of news discourse collection, this study took "China + COVID-19/ Coronavirus/virus" as the search keywords to collect reports from the official websites of Global Times and CNN respectively, and a total of 583 reports were obtained from Global Times and 391 from CNN. Then, manual screening was conducted according to the relevance of the reports, and finally, the author got 266 related reports from Global Times and 132 related reports from CNN. Due to the relatively short length of its news reports, more texts were selected from Global Times to ensure that the two corpora were of the same size. The total number of word tokens in the Global Times corpus was 144,717, while that of CNN was 134,661. Although there was a large gap in the number of news discourse in the two corpora of China-related epidemic reports, the two corpora had the unified collection standard and similar size, thus ensuring the comparability of the analysis results.

Corpus tool

AntConc is a multi-platform tool for carrying out corpus linguistics research, introducing corpus methods, and doing data-driven language learning. AntConc contains eight tools that can be accessed, among which the most frequently used are concordance tool, cluster tool, collocate tool, word list, and keyword list. The concordance tool allows you to see how words and phrases are commonly used in a corpus of texts. The cluster tool shows adjacent word groups based on the search condition, it allows you to see common phrases that appear in the target texts. In some cases, this tool can be seen as summarizing the results generated in the KWIC tool. The collocate tool shows words that appear frequently within a certain distance of the search term (i.e., collocates), which allows users to find which words co-occur with other words in a corpus. The word list tool counts all the words in the corpus and presents them in an ordered list. This allows users to find which words are the most frequent in a corpus. The keyword list tool

shows words that appear unusually frequently in the target corpus in comparison with the words in the reference corpus based on a statistical measure (i.e., keywords). These words can be considered to be characteristic of the target corpus. The settings can also be changed to show words that appear unusually infrequently in the target corpus compared with the reference corpus (i.e., negative keywords).

By using the analysis tools such as concordance, keyword list, and words cluster in AntConc3.5.8, this paper made a comparative analysis of China-related epidemic reports in Global Time and CNN and used DHA to explore the features of Chinese and American news discourse and the ideology behind the differences in discourse features.

Research procedures

The study was divided into three steps.

The first step was the construction of the two target corpora and the determination of the reference corpus. As mentioned in 2.2, this study collected China-related epidemic reports from Global Times and CNN from January 1 to March 31, 2020, and completed the construction of two special corpora through manual screening. The reason why news reports in this period were selected is that these three months basically cover the whole process from the outbreak to the dynamic clearing of the epidemic in China. Studying the changes in the themes and the discursive strategies of China-related epidemic reports during this period is of great significance to interpret the deep-seated ideology behind the news. In addition, the British National Corpus (BNC) was selected as the reference corpus. BNC is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of British English from the later part of the 20th century. The latest edition is the BNC XML Edition, released in 2007. In this study, BNC was used to show words that appear unusually frequently in the target corpus in comparison with the words in the reference corpus based on a statistical measure.

Table 1: General information of the two corpora.

	Corpus GT	Corpus CNN
Number of Texts	266	132
Word Types	8963	8638
Word Tokens	145058	135356

Through the analysis of the keywords, topics, and discursive strategies of the two self-built corpora, it was found that the mainstream media in China and the United States differ in their position when reporting the epidemic in China, which provided data support for further discussion.

Theme analysis

The second step was the synchronic and diachronic study of the corpus. Firstly, the number and topics of epidemic reports in Global Times Corpus (Corpus GT) and CNN Corpus (Corpus CNN) at different stages were compared and analyzed, and the reasons for the changes in the number and topics of epidemic reports were discussed. Secondly, by taking Corpus GT and Corpus CNN as the target corpora and comparing them with the reference corpus BNC respectively, the keyword lists were obtained by using AntConc 3.5.8, and then the keywords were ranked according to their keyness after deleting function words.

By comparing the keywords of the two corpora, the different reporting tendencies of the two countries' media on the COVID-19 in China were analyzed.

The third step was the examination of the discursive strategies used by Global Time and CNN. On the aspect of nomination strategy, this study focused on analyzing naming patterns of the virus, since it can directly reveal the underlying ideology in news coverage.

With the help of AntConc 3.5.8, different naming patterns for the virus were identified and lists of frequency of the naming patterns were obtained. As for the prediction strategy, the very word "China" was chosen as the social actor, aiming to reveal what kind of national image of China the Global Times and CNN have constructed.

In terms of perspectivization strategy, this paper focused on the analysis of direct and indirect language, in this way, the writers' point of view could be clearly seen.

RESULTS AND DISCUSSION

Taking Global Times and CNN as examples, this study selected China-related epidemic reports of the two media in the critical period of China's epidemic development (January 1, 2020 to March 31, 2020) to build two special corpora, and compare them with the reference corpus BNC respectively (Table 1).

Overall data of China-related epidemic reports

Since the outbreak of COVID-19, the overall volume of China-related reports in the United States has reached a record high.

As seen in Figure 1, the two media outlets showed different temporal trends in their coverage of the outbreak in China. In early January, as the number of confirmed cases in China gradually increased, CNN began

to report sporadically, the first of which appeared on January 6, 2020. In mid-January, as more and more COVID-19 cases were reported in China, CNN began to pay close attention to the development of the epidemic in China, and the number of its coverage reached its peak. At this stage, reports mainly focused on the popularization of COVID-19, China's COVID-19 prevention and control measures, the harm caused by COVID-19 to the human body, and the Chinese government's information disclosure transparency. In February, CNN continued to maintain a high degree of attention to the development of the epidemic in China. During this time, CNN took advantage of the death of Dr. Li Wenliang to criticize the Chinese government's "strict internet censorship" and China's deliberate "cover-up" of the emergence of the virus at the beginning of the outbreak. In March, COVID-19 broke out on a large scale in the United States, while the epidemic in China was basically under control. As a result, CNN began to gradually reduce its coverage of the epidemic in China and focus on the development of the epidemic in Europe and the United States. The focus of the China-related epidemic coverage at this stage included the heavy blow of the epidemic to the Chinese economy, the impact of the epidemic in China on the world, and the traceability of the epidemic.

Different from CNN, Global Times has always maintained high-density coverage of the epidemic in China since the outbreak in January, but the focus of coverage has changed over time. In the early days of the outbreak, global Times' coverage focused on the popularization of COVID-19, the anti-epidemic measures taken by Chinese governments at all levels, the supply of anti-epidemic materials, and the concerted efforts of the people across the country to fight the epidemic. In February, the number of newly confirmed COVID-19 cases began to decline day by day after the peak on February 12. The rapid rise of the epidemic in Hubei province and Wuhan city was curbed, and the epidemic situation in the whole country except Hubei Province was generally stable. Therefore, the focus of coverage also began to change. Apart from continuing to focus on the government's anti-epidemic measures and the selfless dedication of medical staff, Global Times began to refute western media's negative coverage of the epidemic in China with reasonable grounds. Besides, the focus has gradually shifted to promoting the orderly resumption of work and production. In March, the epidemic was effectively

brought under control in China, but at the same time, the pandemic began to emerge in Europe and the United States. In this context, in addition to continuing to focus on the resumption of work and production of enterprises, the promotion of national economic recovery, and the prevention of imported cases and domestic resurgence, our report also began to report the development of the epidemic abroad and China's foreign medical assistance.

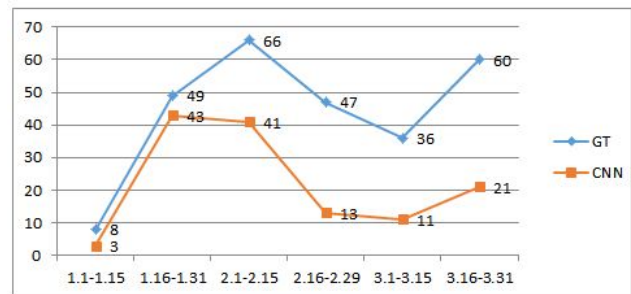


Figure 1: Distribution of China-related epidemic reports from Global Times and CNN.

Comparison of keywords

After a brief introduction to the Global Times and CNN's news coverage, the keyword list tool of AntConc 3.5.8 was used to discover the focus of the two media's coverage.

Word frequency statistics is one of the basic applications of corpus analysis technology. By using high-frequency words and repeatedly mentioning a certain word or concept, news reports can leave a deep impression on readers and euphemistically express the author's tendency, attitude, and ideology. Therefore, starting with the analysis of high-frequency words could help us grasp the salient features of lexical use in news discourse. With the assistance of AntConc 3.5.8, the two self-built corpora were sorted for high-frequency words, and in the end, the top 40 high-frequency words were obtained after removing the words with no practical meaning.

As can be seen from Table 2, there are many common keywords between the two corpora. Therefore, to further explore the similarities and differences between the top 40 keywords in the two corpora, the author conducted a synchronic comparative analysis of them. For the sake of comparison, Table 3 visually shows the shared keywords between the two corpora and the keywords unique to each corpus, which are listed for further discussion.

Table 2: Top 40 keywords in Corpus GT and Corpus CNN.

Corpus Rank	Corpus GT	Corpus CNN	Corpus Rank	Corpus GT	Corpus CNN
1	China	Wuhan	21	confirmed	confirmed
2	coronavirus	China	22	efforts	city
3	Chinese	coronavirus	23	pneumonia	epidemic
4	epidemic	virus	24	pandemic	Trump
5	Wuhan	Chinese	25	supplies	lockdown
6	COVID	outbreak	26	patients	government

7	virus	said	27	control	pandemic
8	said	cases	28	experts	deadly
9	outbreak	Sars	29	fight	CDC
10	global	health	30	Sars	Hong Kong
11	medical	people	31	online	media
12	Hubei	Hubei	32	officials	disease
13	cases	quarantine	33	authorities	patients
14	countries	Xi	34	cooperation	passengers
15	health	officials	35	vaccine	medical
16	prevention	Beijing	36	government	mainland
17	measures	authorities	37	public	infectious
18	Beijing	spread	38	support	public
19	masks	Li	39	international	hospital
20	people	infected	40	imported	measures

As can be seen from the list of high-frequency words, some words appear significantly more frequently than others in the two corpora, such as China, coronavirus, Wuhan, virus, outbreak, health, epidemic, indicating that both reports focused on the outbreak of COVID-19 in Wuhan and its impact, which was in line with the requirement of objectivity of news reports.

However, it is worth noting that in Table 3, words such as quarantine, Xi, spread, infected, deadly, etc. appear frequently in Corpus CNN, but they do not appear in the top 40 keywords of China Daily, indicating that CNN has shown a high degree of concern for every move of the Chinese government represented by President Xi Jinping, and has made continuous reports. In the introduction of

the novel coronavirus, the high transmissibility and mortality rate of the virus was again and again emphasized, and the negative emotions and critical comments could be clearly felt between the lines.

In sharp contrast with the CNN's report, the reports of Global Times put more emphasis on words such as global, preventions, masks, efforts, supplies, control, fight, cooperation, vaccine, supported, international, from which it is clear that the Chinese media paid more attention to the series of measures taken by the Chinese government to effectively control the epidemic and the international cooperation in fighting against the epidemic.

Table 3: Shared and unique keywords between the two corpora.

Corpus GT	Corpus CNN
China, coronavirus, Chinese, epidemic, Wuhan, virus, said, outbreak, medical, Hubei, cases, health, measures, Beijing, people, confirmed, pandemic, patients, Sars, officials, authorities, government, public	quarantine, Xi, spread, Li, infected, city, Trump, lockdown, deadly, CDC, Hong Kong, media, disease, passengers, mainland, infectious, hospital
COVID, global, countries, prevention, masks, efforts, pneumonia, supplies, control, experts, fight, online, cooperation, vaccine, support, international, imported	

Discursive strategies analysis

To dig deeper into the ideology hidden behind the news reports of the two media, this study further analyzed the target news discourses by using the corpus linguistic analysis tool under the DHA framework. DHA mainly contains five strategies, and this study involves the first three, namely nomination strategy, prediction strategy, and perspectivization strategy.

The nomination strategy

According to Chapter 2, the nomination strategy is designed to study how people, phenomena, objects, actions, processes, and events are named linguistically.

The analysis of the nomination of the virus in the news coverage was divided into two steps. The first step was to

compare and analyze the similarities and differences between Global Times and CNN in naming this new virus and disease by using the concordance function of AntConc 3.5.8. The second step was to use the cluster tool to make a comparative analysis of the word "virus" to deepen the understanding of the attitude of the reporter.

In terms of the nomination of the new virus, Table 4 shows the retrieval results of each corpus.

As can be seen from Table 4, on the premise that the two corpora have little difference in size, CNN is significantly more likely to mention the virus in the report than the Global Times, and its naming of the virus is more diversified. In terms of virus naming, in addition to the common neutral words adopted by the two media outlets, such as "virus", "coronavirus", "new coronavirus",

“new virus”, “novel virus”, CNN also used some discriminatory words in its reports, such as “Wuhan Coronavirus”, “Wuhan virus”, “Chinese virus”, “China virus”, “Chinese coronavirus”, and the proportion of such words is as high as 10.84%. Besides, CNN also used some virus names that are easy to cause public panic, such as “deadly virus”, “highly contagious virus”. According to the “Best Practices for the Naming of New Human Infectious Diseases” issued by the World Health Organization in 2015, the naming of infectious diseases should avoid using terms related to geographical location, type of animal or food, person’s name, culture, population, industry or occupation, and terms that evoke undue fear.

Therefore, the naming of the virus in the CNN report clearly violated WHO rules.

It is clear that CNN’s report violated the principle of fairness and objectivity in news reporting and attempted to arouse readers’ fear of the COVID-19 virus and even China through its reports to achieve its ulterior political purpose. In contrast, Global Times was more standardized and unified in naming the virus, which reflects its professional quality.

Next, by using the cluster tool of AntConc 3.5.8, the high-frequency word “virus” was compared and analyzed to further verify the previous analysis results.

Table 4: Naming patterns of the virus in Corpus GT and Corpus CNN.

Media	Corpus CNN	Corpus GT
Naming Patterns	virus (814)	virus (528)
	coronavirus (745)	Covid-19 (384)
	Wuhan coronavirus (135)	coronavirus (183)
	Covid-19 (53)	new coronavirus (28)
	new coronavirus (44)	new virus (5)
	new virus (39)	novel virus (1)
	Wuhan virus (34)	
	Chinese virus (20)	
	China virus (16)	
	deadly virus (15)	
	Chinese coronavirus (4)	
	mysterious pneumonia (3)	
	mysterious virus (2)	
	mysterious new virus (1)	
	novel virus (1)	
	mysterious respiratory virus (1)	
highly contagious virus (1)		
Total	1928	1129

As shown in Table 5, the reports presented by the media of the two countries are different in that: in the reports of CNN, the word “virus” is mainly used with words such as “death”, “severe”, “crisis”, etc., to reflect the severity of the epidemic and its serious impact; while in Global Times, “virus” is often matched with words such as “fight”, “prevention”, “containment”, “control”, “detection” and “vaccine”, whose clusters focus on epidemic prevention and control. This difference reflects the different focus of

Chinese and American media: Chinese media focused on the treatment of patients and the great efforts made by all sectors of society, which was in line with the reality and expectations of the current society; On the contrary, the American media paid more attention to the serious problems and damage caused by the epidemic, and repeatedly stressed the inescapable responsibility of the Chinese government.

Table 5: Cluster list of word “virus”.

Corpus CNN			Corpus GT		
Total No. of Cluster Types: 411			Total No. of Cluster Types: 265		
Total No. of Cluster Tokens: 944			Total No. of Cluster Tokens: 555		
Rank	Freq	Cluster	Rank	Freq	Cluster

1	32	virus spread	1	29	virus outbreak
2	12	virus originated	2	9	virus fight
3	8	virus continues	3	8	virus prevention
4	8	virus outbreak	4	6	virus containment
5	7	virus death	5	5	virus spread
6	7	virus carriers	6	4	virus control
7	5	virus control	7	4	virus infection
8	5	virus emerged	8	3	virus epidemic
9	2	virus severe	9	3	virus detection
10	2	virus crisis	10	3	virus vaccine

The prediction strategy

The second strategy to be examined was the prediction strategy, which was used to clarify the characteristics, qualities, and images of social actors, phenomena, objects, events, and processes. This paper selected “China” as the social actor for analysis, and used the concordance tool of AntConc 3.5.8 to retrieve the common high-frequency keyword “China” in the two

corpora, in an attempt to further reveal the national image of China constructed by the two media.

The results show that there are 1206 concordance lines of “China” in Corpus CNN, while 2138 in Corpus GT. In CNN, some typical concordance lines of “China” are shown in Table 6.

Table 6: Some typical concordance lines of “China” in Corpus CNN.

No.	Context on the Left	Search Term	Context on the Right
1	When asked about	China,	Trump said he was not happy with China
2	some discussion about	China	and I’ve been very critical
3	sparked outrage across	China	where a backlash is growing against state censorship
4	a storm of outrage across	China	film industry are taking a major hit
5	autocratic regimes	China	are using disinformation
6	Amid boiling anger	China	public criticism of authorities
7	Anti- large public gatherings are now forbidden around	China	coronavirus death toll overtakes SARS

The results show that CNN’s coverage of the epidemic in China is relatively negative. Among these expressions related to China, some compared the epidemic to “trauma”, emphasizing the heavy blow to China’s economy; Some blindly played up the death of Li Wenliang, describing the Chinese government as a dictatorial ruler, magnifying the contradiction between the government and the people; Some disregarded the facts, raised strong doubts about the information transparency of the Chinese government, and stigmatized

China by labeling it as the “culprit”. In other words, the use of these negative evaluation words had obvious ideological significance.

By contrast, in the report of Global Times, the expressions related to “China” are diametrically opposed to those in CNN. In addition to indicating the name of the country, there are many positive expressions that can be paired with “China”, and some of the concordance lines are shown in Table 7.

Table 7: Some typical concordance lines of “China” in Corpus GT.

Context on the Left	Search Term	Context on the Right
Economic activity in	China,	has picked up significantly
Disinfection operations carried out across	China	to curb epidemic
WHO delegation highly appreciated the actions	China	has implemented
	China	and Italy are sharing the same fate

		and UK to go through epidemic
	China	has assigned medical groups to epidemic-stricken countries
		will strengthen efforts on supporting African countries
	China	has greatly helped Europe fight the virus.
		stays alert on spreading pneumonia
Japan and	China	cooperate in disease prevention.

Looking at the concordance lines related to “China”, we can conclude that the attitudes of the two media on China’s performance in the epidemic have a distinct ideological tendency. The former completely blamed the Chinese government for the outbreak of the epidemic and its impact on the world economy, and even directly named the virus as “China virus” or “Wuhan virus”. In fact, the outbreak was an accidental event that could have happened in any country. Artificially adding attributes to the epidemic is not conducive to western readers’ correct understanding of the epidemic. The latter’s reports reflect that the Chinese government, in the face of the sudden outbreak, not only rose to the challenge and took the initiative to respond but also actively cooperated with other countries in the fight against the epidemic, which fully embodies the vision of a community with a shared future for mankind advocated by General Secretary Xi Jinping. Van Dijk pointed out that the choice of words in news discourse actually reveals the reporters’ values and judgment. The choice of words will indirectly affect foreign readers’ perception of China, and the words used by CNN are obviously not conducive to the construction of China’s image.

The perspectivization strategy

After analyzing the prediction strategy used by the two media, the perspectivization strategy was studied. An analysis of this strategy aims to discover from what angle the reporters integrated their points of view into the news stories. The paper focuses on analyzing direct and indirect speech to explore the news sources.

As is shown in Table 2, “said” is the most frequently used word in direct and indirect speech. Therefore, it is chosen as the search term to retrieve direct and indirect speech in each corpus. With the help of the concordance tool of AntConc 3.5.8, the differences between the quotations of the two media could be seen clearly.

Different media cite different sources for news reports. By analyzing 1158 Concordances hits from Corpus CNN, four types of information sources were identified, namely “officials and experts” (41.1%), “organizations and departments” (38.8%), “ordinary people” (15.1%), and “media” (5.0%). It can be seen that the category “officials and experts” accounts for the largest proportion of total instances of news sources, followed by the category “organizations and departments”. In the category of “officials and experts,” CNN preferred to cite experts from foreign universities and US Centers for Disease Control and Prevention (C.D.C.).

The Corpus GT contained 1105 Concordances hits, which were also divided into four information sources: “organizations and departments” (41.3%), “officials and experts” (35.2%), “media” (9.5%) and “ordinary people” (14.0%). Unlike CNN, “organizations and departments” is the most popular news source for Global Times, while “experts and officials” is the second most popular one. Some retrieval results of “said” are shown in Figures 2 and Figure 3.

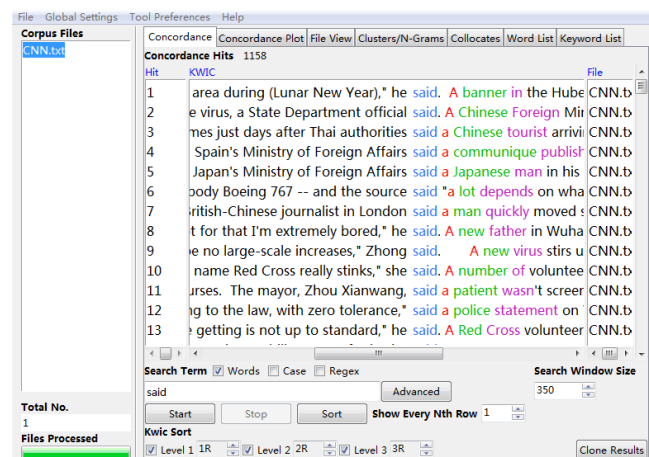


Figure 2: Some concordance lines of “said” in Corpus CNN.

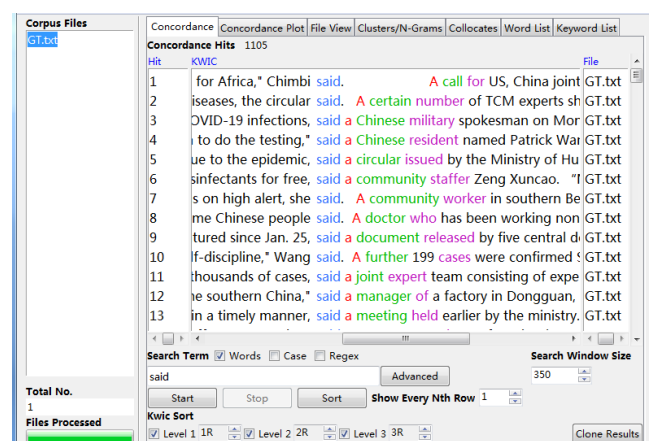


Figure 3: Some concordance lines of “said” in Corpus GT.

A careful study of the search results shows that Global Times and CNN have repeatedly quoted government officials, critics, experts, and scholars in their news reports. The Global Times quoted government leaders more often to increase the authority of the reports and

the credibility of the data sources, showing that leaders at all levels are highly concerned about the epidemic.

CNN's quotes barely supported the Chinese government's anti-epidemic measures, but mostly cited the doubts of observers, critics, and unidentified so-called experts and authorities, focusing on indicating that the epidemic was uncontrollable and wantonly undermining the image of the Chinese government on the grounds of violating the people's personal rights. It intended to use the stereotype of western readers on the Chinese government and media to further mislead them. Below are some examples.

Example 1

"What we've seen instead is that it's just become a new arena for this great power rivalry to play out," he said, adding that autocratic regimes, and China in particular, are using disinformation to try to advance other geopolitical goals, namely positioning itself as a responsible global leader. (CNN, Mar 24, 2020, "Blame game escalates between US and China over coronavirus disinformation").

Example 2

There are also human rights implications when Wuhan and more than a dozen other Chinese cities are placed under lockdown, Gostin said. "I don't think you can enforce a mass quarantine of 30 million people without violating human rights." (CNN, Jan 27, 2020, "China's unprecedented quarantines could have wider consequences, experts say").

Example 3

"Ironically, the Chinese leadership's keen efforts to push for accountability from bureaucrats and promise stiffer punishment for those who shirk responsibilities have contributed to their propensity to cover up disasters," Wang said. (CNN, Jan 27, 2020, "China's unprecedented reaction to the Wuhan virus probably couldn't be pulled off in any other country").

From the above examples, it is clear that CNN also relayed and amplified negative voices with ulterior motives, deliberately portraying China as a country with an autocratic dictatorship and a backward emergency response system, in an attempt to undermine the good image of the Chinese government and society. Through the interweaving of various voices, the media infiltrated personal views and opinions into the reports, and influenced the audience's value judgment, so as to achieve the purpose of controlling the readers' cognition and understanding.

DISCUSSION

Based on the corpus analysis of Chinese and American mass media coverage of China during the COVID-19 outbreak, the author finds that there are significant differences between Chinese and American media coverage of relevant topics.

To begin with, the two media chose different angles of news reporting. The coverage of the Global Times tended to focus on the Chinese government and people's anti-epidemic efforts, showing a moving scene that the Chinese government has always put the safety of the people in the first place, and the whole country has made concerted efforts to overcome the difficulties. However, CNN deliberately ignored the Chinese government's efforts to fight the epidemic, and blindly magnified a series of problems such as the government's neglect of vulnerable groups, restrictions on the personal freedom of citizens, and the fragility of the medical emergency system.

Besides, the two media adopted different discursive strategies in news reports. In terms of the naming of the virus, the naming of Global Times was more in line with WHO's naming rules for infectious diseases. Therefore, the naming of the virus was more standardized and systematic; Although CNN's naming methods for the virus were more diversified, some of them were not only informal but also full of discrimination, reflecting the attempt of some American media to politicize scientific issues by stigmatizing China, making China the scapegoat of causing the epidemic disaster. As Tedros Adhanom Ghebreyesus, director-general of the World Health Organization, said, "Stigma is worse than the virus itself." Since the outbreak, the Chinese government has been releasing information on the epidemic in an open, transparent, and responsible manner. Among the cases of unknown origin that have emerged in many countries, the first two patients confirmed in Iran had no travel history to China, and the first two cases in California had neither travel history to hard-hit areas nor contact with known COVID-19 cases... These also confirm the view that Academician Zhong Nanshan has always emphasized, "The virus first appeared in China, but it does not necessarily originate in China." Obviously, the claim that the virus originated in China is groundless, and the remarks of some American media are extremely irresponsible.

As far as the language expression of the relevant reports is concerned, the words used in the Global Times reports were positive, affirming the efforts made by all walks of life in China to fight the epidemic, while CNN's news reports tended to use negative words, highlighted the destructive power and serious impact of the epidemic, emphasized the pressure the epidemic brought to China and used a lot of derogatory words to create a negative image of the country for American readers.

In terms of direct and indirect speech, both of them repeatedly quoted the words of government officials, experts, scholars, and critics. CNN preferred to cite the comments of US government officials and scholars to emphasize the suffering caused by the poor leadership of the Chinese government. The Global Times, on the other hand, had a more comprehensive source of news. It not only quoted the words of government officials and experts to help eliminate panic and establish a scientific concept of epidemic prevention but also paid more attention to the feelings of patients, vulnerable groups,

and volunteers, which shows the determination and confidence of the Chinese government to care for and serve the people.

Through the analysis, it can be seen that the two media have constructed completely different national images through different discourse expressions, showing completely different ideologies. In fact, the US media has a long history of negative bias in China-related reports. Since the reform and opening-up, China's rapid economic and social development has made it an important force in the existing international order. With the improvement of China's international status and influence, the US media's reports on China have increased year by year. Although there are some objective, fair and positive reports, due to the long-term accumulation of ideological prejudice and cultural stereotypes, some US media reports on China still show strong negative characteristics. Western media's coverage of China is more inclined to choose controversial topics, such as human rights, democracy, rule of law, corruption, market barriers, economic threats, ethnic affairs, etc., and question and criticize the relevant national governance behaviors of the Chinese government through one-sided reports. The reasons are generally as follows:

Firstly, the rise of China and the intensification of the game of interests between China and the US have led to the growing voices of the US media to suppress China's development. It should be noted that although the reports of the news media can cross national borders, the media themselves still have nationalities. This means that a country's news media must take its national interests as an important starting point to carry out news reporting activities. The US media's one-sided coverage of China's fight against the epidemic is a typical example of media using public opinion to suppress China's development.

Secondly, the ideological differences between China and the United States have also greatly affected American society and its people's perception of China. Although some reports cannot withstand close scrutiny, they cater to the ideological bias of the American society to some extent, which is the fundamental reason why these reports can be widely spread in American society.

Thirdly, racial and cultural biases also play a crucial role. In its coverage of COVID-19, the Wall Street Journal did not shy away from calling the Chinese people the "sick man of Asia" and using the epidemic to make racist attacks against China. This apparent stigmatization is not only related to the two reasons mentioned above, but also reflects the racial and cultural bias of some American media against China, as well as the lack of basic understanding of China, Chinese people, and Chinese culture today.

CONCLUSION

News reports not only convey journalists' position and attitude but also indirectly convey their values to readers. With the help of corpus-based discourse-historical analysis, it is found that CNN used a large number of

negative terms in its coverage of the epidemic in China, exaggerating the severity of the epidemic in China and criticizing the Chinese government's measures to combat the epidemic. The Global Times, on the other hand, highlighted the joint efforts of the Chinese people to fight the epidemic, showing the world the image of China as a responsible major country.

By comparing and analyzing the images of China in Chinese and American mainstream media, we can objectively understand how American media manipulate language and spread ideology in their news reports. In addition, exploring the ideology behind the discourse can help us gain a deeper understanding of the values of American society and its attitudes towards China. Chinese media should strive to tell China's stories well while actively strengthening its external communication. One can believe and expect that when the real China and Chinese people are fully presented in front of the world, the China that the world sees must be credible, lovely, and respectable.

REFERENCES

1. Lu G, Chen J, Zhang L. The image construction of the city Yiwu in overseas media: Corpus-driven research based on 2010-2015 and 2016-2021. *Linguistics Culture Rev* 2022; 6:162-76.
2. Cao Q, Chilton P, Tian H. *Discourse, politics and media in contemporary China*. John Benjamins Publishing Company 2014.
3. Tang L. Transitive representations of China's image in the US mainstream newspapers: A corpus-based critical discourse analysis. *Journalism* 2021; 22:804-20.
4. <https://benjamins.com/catalog/ct.7>
5. Xin Bing. Critical linguistics and critical analysis of English news discourse. *Foreign Language Education* 2000; 4: 44-47.
6. Reisigl M, Wodak R. *Discourse and discrimination*. London/New York: Routledge 2001.
7. Reisigl M, Wodak R. *The discourse-historical approach*. Beijing: Beijing University Press 2014.
8. Wodak R. *The discourse-historical approach*. London: Sage 2001.
9. Baker P, Gabrielatos C, Khosravini M, et al. A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and Society* 2008; 3:273-306.
10. Gabrielatos C, Baker P. *Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK Press 1996-2005*. *J English Linguistics* 2008; 1: 5-38.
11. Baker P. Representations of Islam in British broadsheet and tabloid newspapers 1999-2005. *Language and Politics* 2010; 2: 310-338.
12. Baker P. From gay language to normative discourse: A diachronic corpus analysis of

- Lavender Linguistics conference abstracts 1994-2012. *J Language Sexuality* 2012; 2: 179-205.
13. Baker P, Gabrielatos C, McEnery T. Sketching muslims: A corpus-driven analysis of representations around the word "Muslim" in the British press 1998-2009. *Appl Linguistics* 2013; 3:3255-3278.
 14. Baker P, Brookes G, Atanasova D, et al. Changing frames of obesity in the UK press 2008-2017. *Social Sci Med* 2020; 264.
 15. Brookes G, Baker P. Fear and responsibility: Discourses of obesity and risk in the UK press. *J Risk Res* 2021.
 16. Jinhui J, Zhangmei L. A corpus-based critical discourse analysis of China's anti-corruption campaign reported by western media. *Foreign Languages and Culture* 2018; 3: 124-134.
 17. Yu H, Lu H, Hu J. A corpus-based critical discourse analysis of news reports on the COVID-19 pandemic in China and the UK. *Int J English Linguistics* 2021; 11:36.
 18. Unvar S, Rahimi A. A critical discourse analysis of discursive structures in a political text. *Int J Sci Adv Technol* 2013; 3:12-20.
 19. Luqian Z. The international image of Chinese higher education: A diachronic discourse study of British media (2001-2017). *Higher Edu Exp* 2019; 4:30-36.