WILEY | Hindawi

Research Article

A Corpus-Based Study of Public Attitudes towards Coronavirus Vaccines

Ganlin Xia [ID],[1] Yiting Chen [ID],[2] and Lijing Lu [ID][3]

[1]*Shanghai University of Political Science and Law, Shanghai 200030, China*
[2]*Shanghai Meiyuan Primary School, Shanghai 200082, China*
[3]*Hebei University of Economics and Business, Shijiazhuang 050067, China*

Correspondence should be addressed to Ganlin Xia; xiaganlin@shupl.edu.cn

Since the beginning of 2020, COVID-19 has been sweeping the world on an unprecedented scale. As an important means of fighting the virus, vaccines have provoked heated discussions. Motivated by this practical concern, the present study aims at contributing to the understanding of public opinions about vaccines, which may provide implications for the government in their making and implementation of related policies. This research adopts a corpus-based approach in conjunction with a Critical Discourse Analysis (CDA). Data for this study drawn from the Coronavirus corpus show what people are actually saying in online newspapers and magazines in 20 different English-speaking countries. The collocation and frequency of the word "vaccine" are arranged in concordance contextually. Overall, this study reveals that the collocation of vaccine can be divided into several categories, and people's major concerns about COVID-19 vaccination include global progress, equality, and the latest development.

1. Introduction

Outbreaks of infectious diseases such as Ebola, SARS, and H1N1 flu are common around the world. Since the first half of 2020, Coronavirus has been sweeping the world on an unprecedented scale, affecting every aspect of people's lives. The outbreak is reported in Wuhan, a major city in China's Hubei province. In a very short time, the disease spread so fast that it reached almost every part of China. As a result of mass migration during the Chinese Lunar New Year, the virus began to spread abroad. It was proclaimed by the World Health Organization (WHO) as a pandemic on March 11, 2020. At present, the global Coronavirus is still ongoing, posing a serious threat to the life safety of people all over the world. To some extent, we have come into a new postcorona era.

With no particularly effective drugs to treat Coronavirus, the development of vaccines has been racing against time since the outbreak. As an important means of fighting the virus, the development of vaccines has attracted the attention of people all over the world. However, although

laboratories and medical institutions are vigilant in developing a vaccine, their attempts are not effective enough. Besides, because of the lack of knowledge and information about fatalities, diagnosis, medicines, and vaccines of COVID-19, unnecessary fear and nervousness are generated [1]. These factors led to a war of words and an exchange of blame. Governments, public health agencies, and awareness organizations must be prepared to tackle hesitancy and improve vaccine awareness so that, where possible, the population supports immunization. Antivaccination groups are now lobbying against the need for a vaccine in several nations, with others questioning the presence of COVID-19 entirely. Misinformation transmitted across different networks may have a major influence on acceptance.

However, there is only very limited research on public attitude towards COVID-19 vaccines. Therefore, motivated by this practical concern, the present study is aimed at contributing to the understanding of public opinions about vaccines, which may provide implications for governments in their making and implementation of related policies.

This research uses a corpus-based approach in conjunction with a Critical Discourse Analysis (CDA), following the Discourse Analytical Perspective in Dijk [2] to probe public attitudes towards the COVID-19 vaccine. Data for this study drawn from the Coronavirus corpus show what people are actually saying in online newspapers and magazines in 20 different English-speaking countries. In May 2020, the corpus was initially published. Today, there are about 901 million words in the corpus, and every day it keeps rising by 3-4 million words. It is the definitive record of the social, cultural, and economic impact of Coronavirus (COVID-19) in 2020 and beyond. Collocation and frequency of the word "vaccine" can be concordanced contextually. Through word frequency, word collocation, sentiment analysis towards adjectives, and other research methods, the public's attitude towards Coronavirus vaccines can be analyzed and perceived [3]. In addition, since the media is the main source for the public to acquire information about vaccine development, its role in describing the current phenomenon is also worth exploring. Specifically, this paper wants to focus on the following research questions:

(1) What are the most common adjective collocations for "vaccines"?

(2) How are "vaccines" portrayed contextually in online discourses?

(3) What is the general picture of the public attitude towards Coronavirus vaccines?

This study reveals that the collocation of the word "vaccine" can be divided into several categories, indicating that people's major concerns about COVID-19 vaccination include global progress, equality, and the latest development.

The rest of this paper is organized as follows. In Section 2, the theoretical framework and related studies are reviewed, introducing the history of vaccine, research methods of CDA and corpus linguistics, and their applications in studies related to vaccines. In Section 3, instruments used and data collection procedures for this study are described, while in Section 4, results of data analysis are reported. In Section 5, research questions put forward in Section 1 are answered, policy implications are proposed for governments and other organizations, and further research directions are suggested accordingly.

2. Literature Review

2.1. Theoretical Framework

2.1.1. Controversy over Vaccines and Vaccination. Vaccination ranked first among the ten greatest public health achievements of the 20th century [4], since vaccines have saved millions of lives by preventing contagious diseases from spreading [5]. Vaccines are regarded as an important guarantee to maintain social security and improve social productivity. Smallpox vaccination was mandatory in Europe and North America in the 19th century; in the 20th century, certain vaccinations were also included in the admission requirements of public schools [6].

However, vaccination is controversial. A lot of problems we have today can date back to the 1790s, when the first human vaccine was produced by Edward Jenner for curing smallpox ([6]: 612). Since then, the argument of safety, money, and proper immunization schedules has aroused a growing concern from the public, including parents, medical profession, policymakers and the media. Their attitudes vary from "awe of a seeming scientific miracle to skepticism and outright hostility" ([6]: 613). Margaret Chan, the Director-General of the World Health Organization, once described the wide distrust of vaccination as "worrisome" ([7]:1151).

Increasing anxiety and panic leads to some quitting. The antivaccination movement began to gain support in the United States at the end of the 19th century when smallpox outbroke. Parents' anxiety and protest were triggered by the smallpox vaccine itself, which violates their personal rights. Tensions had escalated when a mandatory vaccination program was issued by the government. More than 100 years ago, Jacobson, a citizen of the city of Cambridge in Massachusetts, declined to be vaccinated for smallpox as he argued that the law violated his right to care for his own body. The Court dismissed the challenge raised by Jacobson. In order to protect the public's welfare, the U. S. Supreme Court ruled in 1905 that, in the case of infectious disease, the state of Massachusetts had the authority to protect the public with mandatory laws. That is the first U.S. Case on the power of states in public health law [8]. There is often a dispute between preserving individual liberties and protecting public health, and a major challenge is how to balance individual rights with community needs. Certain religions and value systems also fostered alternate vaccine views. In general, moral opposition to vaccination is based on the ethical dilemmas involved in the use of human tissue cells to manufacture vaccines and the conviction that such drugs or blood or tissues should not be received from animals and that they should be cured by God or by natural means [9].

In the mid-1970s, an international dispute arose in Europe, Asia, Australia, and North America over the immunological protection of DTP [10]. The UK is once again the center of antivaccination protests nearly 25 years after the DTP controversy, this time against the MMR vaccine. As Johnston [11] observed, most opponents of vaccines are not entirely opposed to immunization. What they are concerned about is the safety and efficacy of vaccines. Therefore, they resist only some specific vaccines with potential safety hazards.

2.1.2. CDA and Corpus Linguistics. Following the Discourse Analytical Approach in van Dijk [12], the CDA framework analyzes media texts not just in the text but also from the context. It was described as a sort of study of dialogue that focuses on the examination of social motivations or the philosophy of the speaker behind his or her language choice in a debate. The reason for using CDA to analyze these texts is that CDA is an approach to discourse research aimed at analyzing the use of language, whether spoken or written, in daily conversation. From this perspective, language is seen as a social activity [13] that reflects not only other social

practices but also elements such as domination, opposition, control, and ideology [14].

Much of the research that is combined CDA with Corpus Linguistics (henceforth CL) has placed their emphasis on "the implications of lexical choices in the text" ([15]:16) after analyzing ideological ideas in corpora. According to Hunston [16], such studies usually choose high-frequency words as their research emphasis. The use of CL tends to establish integrity and trustworthiness for the research results since the researcher tries to gather a specialized corpus that is representative of the text type [16]. The textual study would analyze the language used in the texts in expressing the view of writers and their semantic and pragmatic meaning [17]. However, the qualitative method still needs to be adopted to analyze the newspaper's fiscal, political, and cultural values. The purpose of CL is not to discover the meaning in isolated words that are out of context. Instead, in a specific sense, the purpose is to identify patterns of word items used in discourse (consistent lines) and words that occur together (collocations) [17].

2.1.3. CDA and Media Discourse. This research focuses on the analysis of the reports in online newspapers, a typical discourse type. As the main function of news reports is to give definitions and labels to people's lives, their expression is the key to portraying real life. From the angle of linguistics, the most integrated study was done in Bednarek [18], who believes that the essence of media discourse is "manufactured." Bednarek and Caple [19] also show the embedded bias of newspapers. The media discourse of this study refers to vaccination in online magazines and newspapers that are deposited in the corpus. Their creators are journalists, and readers are the general public. This research will concentrate on the discourse, adopt the method of Van Dijk, and highlight the theme.

2.2. Related Studies. As Coronavirus started to spread rapidly, the scientific community was motivated to work together in order to gather, organize, and analyze data. At the beginning of the outbreak, most of the scientific research was focused on biology or medicine, and linguistic research was rare.

Katermina and Yachenko [20] and Kim [21] examined the linguistic phenomenon related to COVID-19 in English mass media texts to determine if mass media creates, reproduces, and transmits axiological values about COVID-19. The research studied the word collocation of "COVID-19," "Coronavirus," "virus," and "disease." Their results revealed that COVID-19 is widely discussed in the mass media in many ways. For example, some conceptualize the virus as war. The researchers concluded that the COVID-19 pandemic had affected culture and language to the extent that it can be considered a social calamity.

A contrasting analysis conducted in Awad AlAfnan [22] studied the coverage of COVID-19 in two newspapers, America's Washington Post and China's People's Daily. It is found that mass media expressed biased opinions in their coverage and declaration. In the study, both CDA and CL approaches were used, which revealed that the Chinese virus, Wuhan Virus, and KongFlu were used to refer to the COVID-19. In Washington Post, these labels were focused on creating a derogatory hollow among Asian Americans.

Azizan et al. [23] performed a study based on 15 Facebook posts about COVID-19. They used Constructive Discourse Analysis and Critical Discourse Analysis approaches. A thematic analysis was made to codify and categorize data, which revealed the creation of positive discourse. The posts highlighted the extensive use of collective pronouns such as "we" and "us" that symbolize unity and empowerment in dealing with COVID-19 among Malaysians. Regarding Facebook posts, four positive trends have been identified, which are faith, patriotism, call for heroism, and public awareness. The researchers concluded that individual actions subsequently served as a powerful buffer against negative discourse and as an attempt to make a difference to the emergence of subtle influence.

In most of the studies, Natural Language Processing (NLP), machine learning, sentiment analysis, and text mining were used for analysis by computer software. Many different aspects of COVID-19 and its impacts have been analyzed. There are some studies on the public opinions about the introduction of the COVID-19 vaccines. Some of these studies are based on the text of social media, seeking to understand the general public's acceptance or hesitance of vaccines and the influencing factors (e.g., [24–26]). There are also studies designed to understand the attitude of medical professionals towards vaccines, aiming to provide advice and guidance to nonprofessionals (e.g., [27, 28]). Compared with most of the above studies, the most remarkable characteristic of the current study is as follows. The text style in the corpus selected for this study is news reports, and the corpus includes the news media, journals, and magazines of 20 English-speaking countries in the world, so it can better represent the global public attitude.

3. Methodology

3.1. The Coronavirus Corpus. The discourse analyzed in this study is from English Corpora (https://www.english-corpora.org/). It is the most widely used collection of corpora in the world. Currently, the corpora are visited by more than 130,000 users every month from over 140 countries worldwide. Besides, hundreds of universities worldwide have applied for academic licenses, which allow their students and scholars to use the expanded and additional functions of the corpora.

The Coronavirus Corpus selected for this study is a subset of the NOW Corpus (News on the Web), a 14.3-billion-word corpus based on online newspapers and magazines from 2010 to the present time, with 180–200 million words added each month. The Coronavirus Corpus, with 1,350 million words by January 12, 2022, is also characterized by daily updates. It is selected for this study mainly because it is constructed centering around the news content related to COVID-19 on the Internet. Articles meeting the following two requirements are selected from the news websites: first, each article should have at least two occurrences of one of the three words,

i.e., Coronavirus, COVID, or COVID-19; Second, each article should have one of the following words/strings in the title: at-risk, cases, confirmed, contagious, containm*(words start with "containm," the same below), Coronavirus, covid*, curbside, curve, deaths, disinfect*, distanc*, epicenter, epidemic, epidemiol*, flatten*, flu, high-risk, hoard*, hospital*, hydroxychloroquine, infect*, influenza, isolat*, lockdown, lockdown, mask*, nursing, outbreak, pandemic, panic, patient*, pneumon*, preventative, preventive, quarantin*, reopen*, reopen*, respiratory, sanitiz*, self-isolat*, shelter*, shutdown, spread, spreading, stay-at-home, stay at home, stockpil*, testing, vaccine*, ventilator*, virus.

As claimed in the introduction of the corpus on the website, the Coronavirus Corpus can be regarded as "the definitive record of the social, cultural, and economic impact of the coronavirus (COVID-19) in 2020 and beyond" (https://www.english-corpora.org/corona/).

3.2. Data Collection and Analysis. Data between January 2020 and the end of April 2021 was collected from the above-mentioned Coronavirus Corpus. Two functions of this corpus are used to analyze this study, frequency list, and collocation analysis.

The frequency list displays the frequency of words or phrases searched in 10-day increments since Jan 2020, which can not only provide the total frequency of the retrieved words but also indicate the changing characteristics of the frequencies in a certain period of time.

Collocation analysis was conducted to investigate people's attitudes towards COVID-19 vaccines, and the results list words that usually appear together with the searched word in context. Baker emphasizes that collocation has an ideological nature and that the recurrent collocation of lexical items demonstrates the association of two concepts in people's minds (2006: 114). These words can be found on the node word's left (L) side and right (R) side. The span of the word was set to search for the 2L and 2R around the node word "vaccine*" (words starting with vaccine including the word vaccines and other compounds) in this study. The minimum frequency of collocation was controlled to be 40, so the words shown in the list must have appeared around the node word over 40 times in the corpus. These settings are meant to filter out words that have no real meaning or do not occur frequently enough. Besides, in this research, mutual information3 (MI3) is adopted to measure the effect size of associations between a node word and its collocates (see [29] for detailed discussion).

4. Results

4.1. Frequencies. Figure 1 shows that the frequencies of the word "vaccine" in NOW Corpus have increased significantly since 2020. In 2020, the PER MIL (per million) frequency of it was 203.06, 8–11 times that of the previous years, while in 2021, the frequency per million words was 613.87, three times that of 2021. There is no doubt that while the epidemic is raging all over the world, people's attention to vaccines is bound to increase.

From Figure 2, it is clear that the frequency of the word "vaccine" has been increasing since January 2020. Vaccination has been put on the agenda in Asian countries such as China, European and American countries, and people around the world are paying increasing attention to and widely talking about vaccines. In December 2020, many countries, including China and the United States, successively approved the conditional marketing of vaccines [32, 33]. Therefore, since December 2020, the frequency of the word in the corpus has increased sharply and maintained this high frequency so far.

4.2. Collocation Analysis. The result of collocation concordance is shown in Figure 3. Adjectives modifying the word "vaccine" are grouped into three emotional categories, i.e., positive, negative, and neutral.

4.2.1. Positive Vaccine Portrayals. Based on the concordance results, the top100 adjectives are taken to illustrate the point. Among the top 100 most frequently used adjectives, positive and optimistic one are: "*available*," "*effective*," "*safe*," "*successful*," "*promising*," "*leading*," "*good*," "*viable*," "*positive*," and "*eligible*." Their specific rankings and frequency of use are shown in Table 1.

It can be discovered from Table 1 that the frequencies of the first and the second adjectives "*available*" and "*effective*" are very high, and the word "*safe*" ranked sixth and the word "*successful*" ranked 10, which are also the focus of public attention. It is quite reasonable that vaccine, as a special public security product, its safety and effectiveness are two of its characteristics that cannot be ignored. In order to gather a deeper understanding of how these words are used contextually, we referred to the concordance lines. Following Barnbrook [34], the power of concordance lines lies in the way "of placing each word back in its original context, so that the details of its use and behaviour can be properly examined" (p. 65). The following example shows a random sampling of concordance lines from a total of 2042 occurrences for the word "*successful*:"

(1) Given that at least 200,000 in the U.S. and more than a million globally have already died from COVID-19, we desperately need as many of the vaccines under development to succeed as soon as possible. # Success in dealing with COVID-19 requires far more than *successful* vaccine development, however. We need to eradicate the happy talk that, somehow, the pandemic will end when a vaccine is approved. #

(2) "A *successful* vaccine against Sars-Cov-2 could be used to prevent infection, disease, and death in the whole population, with high-risk populations such as hospital workers and older adults, prioritized to receive the vaccination." #

(3) The global race to find a *successful* coronavirus vaccine continues at full pace. Meanwhile, the virus has claimed over 8 lakh lives across the world. #
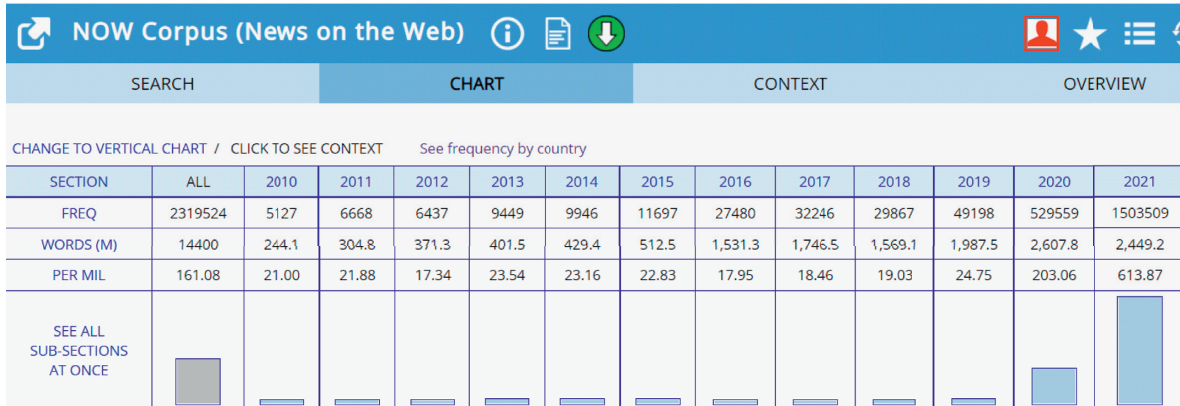
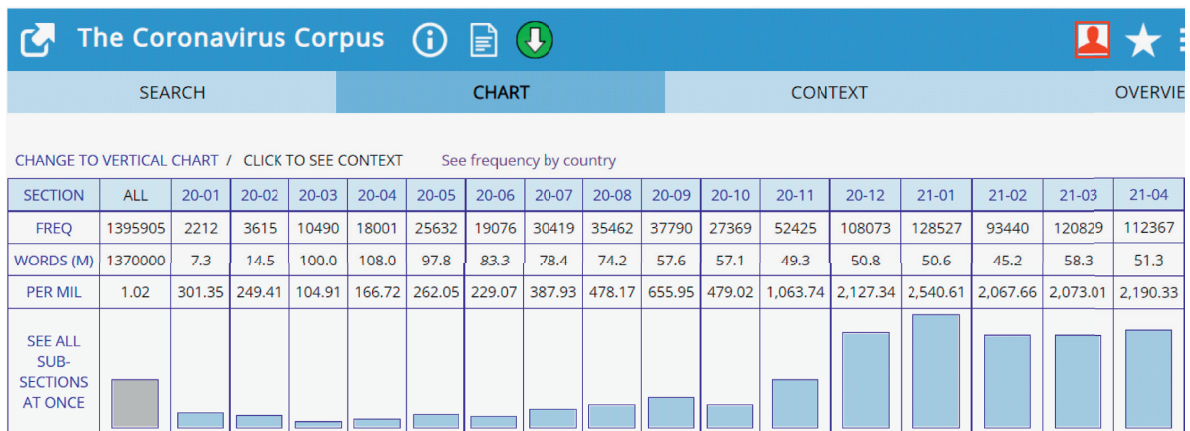FIGURE 1: The frequencies of "VACCINE" in NOW Corpus (news on the web) [30].

FIGURE 2: The frequencies of "VACCINE" in The Coronavirus Corpus [31].
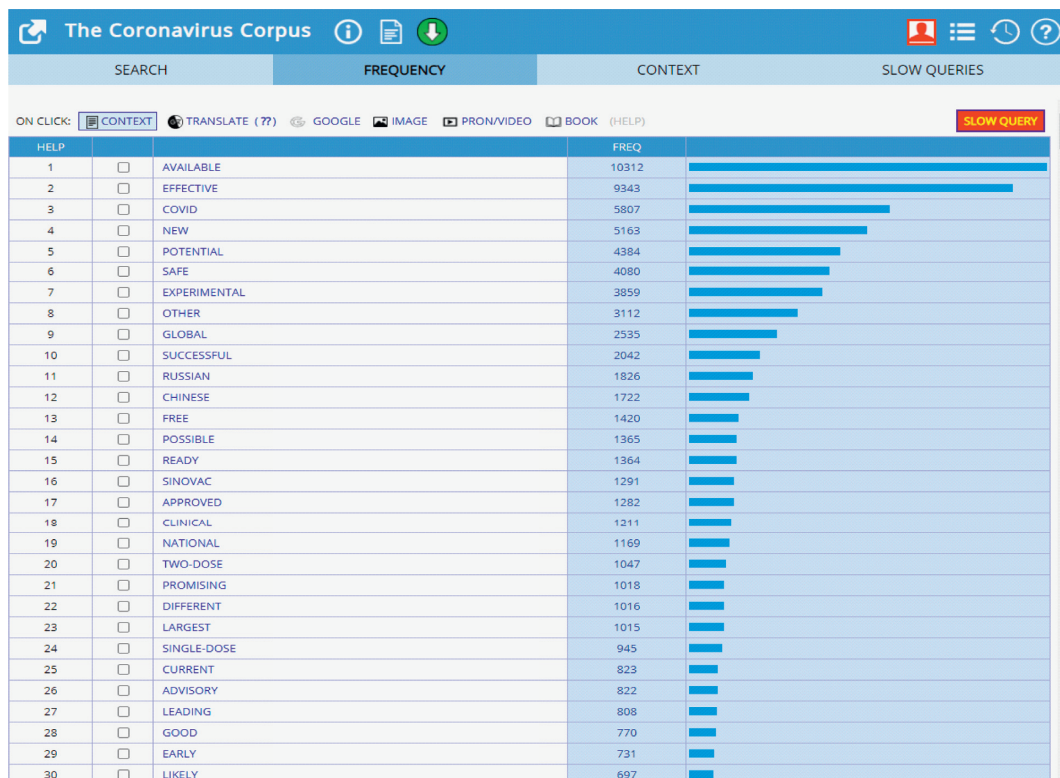

FIGURE 3: The collocation list in The Coronavirus Corpus [31].

TABLE 1: Positive collocation of the word *vaccine*.

Word	Rank	Freq
Available	1	10312
Effective	2	9343
Safe	6	4080
Successful	10	2042
Promising	21	1018
Leading	27	808
Good	28	770
Viable	34	653
Positive	46	538
Eligible	73	356

TABLE 2: Negative collocation of the word "vaccine."

Word	Rank	Freq
Potential	5	4384
Mandatory	53	456
So-called	61	410
Adverse	84	327
Fake	90	315

All the three examples above illustrate the public's desire for an effective vaccine, but they are slightly different. In the first example, we can see that the speaker wants to warn people not to rely entirely on the development of vaccines but to be prepared for a long battle.

When analyzing the specific context, we can dig out the deeper inner meaning behind the surface phenomena, which is also an important reason why this study combines the method of CDA with that of corpus linguistics so that both the concordance function of corpus tools and more detailed and critical analysis of CDA can be made full use of.

4.2.2. Negative Vaccine Portrayals. Among the top 100, there are only five adjectives with negative ideas and attitudes (Table 2). However, it is noteworthy that the use of "*potential*," which ranks fifth, indicates that people still pay critical attention to the potential risks of vaccines. In addition, the use of words such as "*so-called*" and "*adverse*" reflects a high degree of public distrust of vaccines, especially in terms of the effect and quality of vaccines. This point is also well confirmed in the following specific context.

(1) # However, one controversial part of this potential "vaccine bubble" is that people will feel like their freedom will still be restricted unless they are part of the "bubble." #

(2) # In recent days, China has vowed to deploy a potential vaccine as a "global public good" that would be accessible and affordable. Toward that end, the World Health Organization also laid out principles to encourage collaboration and information sharing on a COVID-19 treatment. #

In the analysis of adjectives indicating negative emotions, the word *potential* is taken as an example. In one context, the word "*potential*" *is used together with words such as* risk or threat. In another context, the word "*potential*" suggests that vaccines will be developed in the future, and China will use vaccines for global health. The difference between these two examples lies in the use of the potential for two different meanings of the word. That is to say, the contextual results of the corpus will not distinguish the meanings, and at this time, we need to make semantic division manually to obtain more precise and accurate meanings.

4.2.3. Neutral Vaccine Portrayals. Most adjectives do not have a clear tendency of emotion; thus, they are classified as neutral (Table 3). These neutral adjectives are divided into four categories (Table 4).

The first category can be labeled as "region, country, and organization." Countries on the top of the list are "*Chinese*" and "*Russian*." In fact, China and Russia have been among the fastest in the world in terms of vaccine research and development, and their progress has attracted the attention of many countries. It is worth noting that SINOVAC, the major vaccine research company in China, is mentioned frequently. Also, Global Alliance for Vaccines and Immunization (GAVI) is the only organization in the top 100. GAVI is a crucial public-private global health alliance established in 1999 that works with governments and nongovernmental organizations to promote global health and immunization. During the period selected by this study, it actively updated the latest vaccine research situation on their website and encouraged global vaccination. GAVI is coleading COVAX raised by WHO, the vaccines pillar of the Access to COVID-19 Tools (ACT) accelerator. It has established a global risk-sharing mechanism for pooled procurement and equitable distribution of COVID-19 vaccines. COVAX has transported 53 million COVID-19 vaccines to 121 participants so far and has played an important role in the transportation and distribution of vaccines around the world.

The second category is labeled as "medical terms," including information about testing and production of vaccines and some of the words that are specific to the medical field of vaccines. This shows that the media consciously used medical terms and principles related to vaccines to inform the public of the basic information about vaccines so that the public can better understand the working mechanism of vaccines.

The third category is labeled as "development process," mainly including adjectives related to the research and development process and time points of the vaccine. This indicates that the media has been paying close attention to the research and development process of vaccines all the time. This also fully reflects the urgent need for effective vaccines and the expectation that vaccine development can achieve a breakthrough.

5. Conclusions and Discussions

5.1. Conclusions. Research questions put forward here are as follows:

TABLE 3: Neutral collocation of the word "vaccine."

Word	Rank	Freq
Covid	3	5807
New	4	5163
Experimental	7	3859
Other	8	3112
Global	9	2535
Russian	11	1826
Chinese	12	1722
Free	13	1420
Possible	14	1365
Ready	15	1364
SINOVAC	16	1291
Approved	17	1282
Clinical	18	1211
National	19	1169
Two-dose	20	1047
Different	22	1016
Largest	23	1015
Single-dose	24	945
Current	25	823
Advisory	26	822
Early	29	731
Likely	30	697
Additional	31	681
Inactivated	32	680
Single	33	653
Indigenous	35	653
Federal	36	598
Oral	37	591
Slow	39	577
Only	40	575
International	41	570
Future	42	568
Novel	43	551
Local	44	550
One-shot	45	546
One-dose	51	461
Similar	52	459
Major	54	446
Eventual	55	440
Viral	56	435
Universal	57	435
Biggest	58	434
Investigational	59	412
Latest	60	412
Particular	62	407
Developed	63	407
Widespread	64	407
Domestic	65	403
Mass	66	396
Two-shot	67	393
Pandemic	68	392
High	69	390
Multiple	70	387
Live	71	386
Ongoing	72	360
Indian	74	355
Rapid	75	353
Public	76	352
Coronvavac	77	352
Existing	78	348

TABLE 3: Continued.

Word	Rank	Freq
Therapeutic	79	341
Actual	80	340
Specific	81	340
Recent	82	334
Equitable	83	331
Seasonal	85	327
Top	86	320
Proven	87	319
Developing	88	319
Sure	89	318
Recombinant	91	312
Pneumococcal	92	312
Various	93	309
Working	94	298
GAVI	95	298
Real	96	289
Important	97	288
African	98	280
Late	99	278

(1) What are the most common adjective collocations for "vaccines"?

(2) How are "vaccines" portrayed contextually in online discourses?

(3) What is the general picture of the public attitude towards Coronavirus vaccines?

Next, these questions will be answered based on the analysis results obtained in this study.

5.1.1. Adjective Collocations for "Vaccines". To answer the first two research questions, a total of 447 adjectives were included, and the lowest frequency of use was more than 40 times. The classified percentage of these adjectives is shown in Table 5.

Among the 447 adjectives, the number of positive adjectives was about twice as many as the number of negative adjectives. Therefore, it can be concluded that people are more inclined to express positive and hopeful emotions when expressing their opinions on COVID-19 vaccines. Many positive expressions have appeared in reports describing the latest progress of the Coronavirus vaccines, which is bound to play an optimistic role in the fight against the Coronavirus. At the same time, the study found that many of the positive adjectives found their antonyms in the negative adjectives, indicating that there are different views on the same issue among the public.

5.1.2. Public Attitude towards Vaccine. The third research question, "What is the general picture of the public attitude towards Coronavirus vaccines?" can be answered based on the results of collocation analysis. Public concerns about vaccines fall into three categories.

The first is the globalization of vaccines. As the first category of adjectives mentioned in Section 4, among the 447 adjectives, frequent occurrences of *global, national*

TABLE 4: The classification of neutral adjectives.

1	Region, country, organization	Global, Russian, Chinese, SINOVAC, national, indigenous, federal, international, local, universal, domestic, Indian, GAVI, African
2	Medical terms	Covid, two-dose, single-dose, inactivated, oral, one-dose, one-shot, two-shot, viral, live, coronaVac, therapeutic, recombinant, pneumococcal
3	Development process	New, experimental, ready, current, early, slow, future, eventual, latest, ongoing, working, rapid, recent, seasonal, developing, late, clinical
4	Descriptions	Other, free, possible, approved, different, largest, advisory, likely, additional, limited, only, similar, major, biggest, investigational widespread, mass, pandemic, mass, high, multiple, public, existing, actual, specific, equitable, proven, top, sure, various, real, important

TABLE 5: Adjective collocations for "vaccines."

Positive	41	9.1%
Negative	22	4.9%
Neutral	384	85.9%
Total	447	100%

worldwide, nationwide rank high in the list, indicating that people are concerned not only about the development of vaccines in their own country but also about the development of vaccines in the world. In the face of a huge crisis like COVID-19, all human beings should fight together. Here is an example in this sense.

Equally clear is that the drive for *a global COVID-19 vaccine regimen and the global surveillance grid* are moving ahead in conPcert to transform the world as we know it–if we allow it to happen. # Professor Michel Chossudovsky said in his recently republished timeline' COVID-19 Coronavirus "Fake" Pandemic.

The second category is about the equal distribution of vaccines. Such adjectives as *equitable, equal, and fair* frequently appeared in the context mostly reflect people's worry about whether everyone can be immunized. As a special public health measure, the distribution of vaccination can be a test to the governments in different countries and to WHO, and the improper sequence of vaccination may even lead to social disorder. The following example indicates concern of this type.

Climate activist Greta Thunberg says governments, vaccine developers, and the international community must "step up their game" to fight global vaccine inequity. The Swedish teen who inspired the Fridays for Future environmental movement cites estimates that one in four people in high-income countries have received coronavirus vaccines, compared with one in 500 in the middle- and lower-income countries. # Thunberg says it is "*completely unethical* that high-income countries are now vaccinating young and healthy people if that happens at the expense of people in risk groups and on the front lines in low- and middle-income countries." #

Finally, most people pay attention to the effect of vaccines. A group of frequently used adjectives is words like *latest, novel, update,* which express concerns about the latest situation and progress. After studying the context of these adjectives carefully, it is not hard to find that discourses with these words are mainly about whether the vaccine developing now can resist the attack from the variant of the virus. In fact, this is a big problem we are facing. The Coronavirus is mutating at a rapid rate that even vaccinated people may be infected again. The timeliness of vaccines has become one of the concerns of the public. The following statement is an example of this type.

" The problem with a novel virus is it has got so much room to grow and shift and change to optimize itself, " said Harvard School of Public Health's Dr. Michael Mina. " The question is how quickly is it going to keep updating itself " " Whatever mutation happens, the vaccine manufacturers can keep up with it, but given the multiplicity of strains circulating at one time, that may be a challenge, " said Dr. Schaffner. # Dr. Mina, from Harvard, said he is less confident that the existing vaccines can be updated for any and all future variants.

5.2. Discussions. This study mainly studied the public's attitude towards COVID-19 vaccines and their focus of attention. It confirms the positive portrayal of those who choose to be vaccinated. In addition, this study analyzed the neutral adjectives and divided them into four categories. Such classification can help policymakers and healthcare practitioners to understand people's concerns about vaccines and help the media to understand what people most want to know. Based on these findings, it is recommended that governments and public health agencies improve their strategies to better communicate the benefits of vaccination to the public, i.e., to increase vaccine confidence. Such strategies could include methods like increasing transparency of data and introducing data from scientific studies to make vaccines more credible. For example, Chinese media has constantly reported the development of the COVID-19 vaccines in China and around the world at a very early stage, and there are a lot of discussions about this issue on social media platforms like Sina Weibo (a blogging platform where people leave comments and express attitudes) [25]. Moreover, public attitudes are influenced by many factors. Thus, it is suggested that strategies should be adopted to eliminate negative emotions of the public on vaccination and improve positive feelings at the same time [35]. A strategy called "prosocial motivation" proved effective in making people take preventive measures by Jordan et al. [36] and Heffner et al. [37] belong to this category, which emphasizes how vaccination will protect the community and will bring life back to a state where everyone could be intimately connected. This position is supported by the study of Lyu et al. [38], which indicates that improving the public's pandemic experience and increasing their sentiment scores can promote their acceptance of the COVID-19 vaccines.

This study is limited to examining data from January 2020 to April 2021. With the continuous development of the epidemic, human beings have entered the postvaccine era, with various virus variants emerging in an endless stream. It is believed that people's attitudes towards vaccines will also continue to evolve. At the same time, this corpus only contains data from a few countries, so a corpus composed of international newspapers and media will also be a valuable supplement to the study.

Now, the practice of vaccination has entered a new stage, and the governments of all countries are vigorously advocating and encouraging people to be vaccinated against COVID-19. Although China adheres to the principle of voluntary vaccination, people's enthusiasm about vaccination is very high. However, at the same time, there are still people who refuse to be exposed to the COVID-19 vaccine because of concerns about the safety of the vaccine. Lin et al. [40] investigate and summarize the steps taken by the Chinese government to control the spread of COVID-19 and reopen lockdown cities, emphasizing the effect of awareness diffusion in this process. Therefore, further research can be conducted to probe into factors underlining different attitudes and ways to promote people's awareness so that both vaccines and vaccination policies can be improved accordingly.

Data Availability

The data analyzed in this study were taken from Coronavirus Corpus from English Corpora (https://www.english-corpora.org/).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Zhao, Z. Ahmad, S. K. Khosa, M. Yusuf, O. A. Alamri, and M. S. Mohamed, "The Role of Technology in COVID-19 Pandemic Management and its Financial Impact," *Complexity*, vol. 2021, Article ID 4860704, 12 pages, 2021.

[2] T. A. V. Dijk, *News as Discourse*, University of Groningen, Groningen, Netherlands, 1988.

[3] V. Raghupathi, J. Ren, and W. Raghupathi, "Studying public perception about vaccination: a sentiment analysis of tweets," *International Journal of Environmental Research and Public Health*, vol. 17, no. 10, p. 3464, 2020.

[4] Centers for Disease Control and Prevention (Cdc), "Ten great public health achievements--United States, 1900-1999. MMWR," *Morbidity & Mortality Weekly Report*, vol. 48, no. 12, pp. 241–243, 1999.

[5] H. Markel, "Taking Shots: The Modern Miracle of Vaccines. Medscapem," 2004, https://www.medscape.com/viewarticle/481059.

[6] A. M. Stern and H. Markel, "The history of vaccines and immunization: familiar patterns, new challenges," *Health Affairs*, vol. 24, no. 3, pp. 611–621, 2005.

[7] A. Kennedy, K. LaVail, G. Nowak, M. Basket, and S. Landry, "Confidence about vaccines in the United States:

[8] T. Baroiant, "Combining Critical Discourse Analysis and Corpus Linguistics to Examine the Term "vaccine" in Online News media," Unpublished master's thesis), San Francisco State University, San Francisco, USA, 2015.

[9] R. A. Goodman, R. E. Hoffman, W. Lopez, G. W. Matthews, M. A. Rothstein, and K. L. Foster, Eds., *Law in Public Health Practice*, Oxford University Press, Oxford, UK, 2007.

[10] "The History of Vaccines. (n.d.)," 2020, http://www.historyofvaccines.org/content/articles/history-anti-vaccination-movements.

[11] R. D. Johnston, *The Politics of Healing: Histories of Alternative Medicine in Twentieth-century North America*, Psychology Press, Hove, East Sussex, UK, 2004.

[12] T. A. Van Dijk, *News as Discourse*, University of Groningen, Groningen, Netherlands, 1988.

[13] M. A. K. Halliday, "Modes of meaning and modes of saying: types of grammatical structure and their determination by different semantic functions," in *Context in Linguistic Analysis. Essays Offered to William Haas*, D. J. Allerton, E. Carney, and D. F. Holdcroft, Eds., Cambridge University Press, Cambridge, England, 1979.

[14] A. Atkins, "Critical Discourse Analysis: A Letter to Expatriates from the Rt. Hon. Sir Norman Fowler MP," 2002, http://www.cels.bham.ac.uk/resources/essays/Andrew%20Atkins%20Assignment%205.pdf.

[15] M. A. Kandil, *The Israeli-Palestinian Conflict in American, Arab, and British media:Corpus-based critical discourse analysis*, Georgia State University, Atlanta, Georgia, USA, 2009.

[16] S. Hunston, *Corpora in Applied Linguistics*, Cambridge University Press, Cambridge, UK, 2002.

[17] M. Stubbs, *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*, Blackwell, Oxford, UK, 1996.

[18] M. Bednarek, *Evaluation in media Discourse: Analysis of a Newspaper Corpus*, A&C Black, London, UK, 2006.

[19] M. Bednarek and H. Caple, "Why do news values matter? Towards a new methodological framework for analyzing news discourse in Critical Discourse Analysis and beyond," *Discourse &Society*, vol. 25, 2014.

[20] V. Katermina and E. Yachenko, "Axiology of COVID-19 as a linguistic phenomenon in English mass media discourse," *Advances in Journalism and Communication*, vol. 8, no. 02, p. 59, 2020.

[21] Y. Kim, "On media coverage of the COVID-19 outbreak: a corpus-based collocation study," *SNU Working Papers in English Language and Linguistics*, vol. 17, 2020.

[22] M. Awad AlAfnan, "COVID 19-The foreign virus: media bias, ideology and dominance in Chinese and American newspaper articles," *International Journal of Applied Linguistics & English Literature*, vol. 9, no. 1, pp. 56–60, 2020.

[23] M. Azizan, H. H. Ismail, and S. N. Qaiwer, "Power and solidarity in positive Facebook postings amidst covid-19 in Malaysia," *Journal of Nusantara Studies (JONUS)*, vol. 5, no. 2, pp. 329–364, 2020.

[24] S. Das and A. Dutta, "Characterizing public emotions and sentiments in COVID-19 environment: a case study of India," *Journal of Human Behavior in the Social Environment*, pp. 1–14, 2020.

[25] H. Gao, D. Guo, J. Wu, Q. Zhao, and L. Li, "Changes of China's public attitudes to domestic COVID-19 vaccination after the vaccines approved: a semantic network and

sentiment analysis based on Sina Weibo texts," *Frontiers in Public Health*, p. 1677, 2021.

[26] A. Wisniak, H. Baysson, N. Pullen et al., "COVID-19 Vaccination Acceptance in the canton of Geneva: A Cross-Sectional Population-Based Study," *medRxiv*, vol. 151, 2021.

[27] K. H. Nguyen, D. Yankey, P. J. Lu et al., "Report of health care provider recommendation for COVID-19 vaccination among adults, by recipient COVID-19 vaccination status and attitudes—United States, april–september 2021," *Morbidity and Mortality Weekly Report*, vol. 70, no. 50, p. 1723, 2021.

[28] S. Bell, R. M. Clarke, S. A. Ismail et al., "COVID-19 vaccination beliefs, attitudes, and behaviours among health and social care workers in the UK: a mixed-methods study," *PLoS One*, vol. 17, no. 1, Article ID e0260949, 2022.

[29] D. Gablasova, V. Brezina, and T. McEnery, "Collocations in corpus-based language learning research: identifying, comparing, and interpreting the evidence," *Language and Learning*, vol. 67, no. 1, pp. 155–179, 2017.

[30] M. Davies, "Corpus of news on the Web (NOW)," 2016, https://www.english-corpora.org/now/.

[31] M. Davies, "The coronavirus corpus," 2019, https://www.english-corpora.org/corona/.

[32] Xinhuanet, "China approves first self-developed COVID-19 vaccine. Xinhuanet," 2020, https://www.xinhuanet.com/english/2020-12/31/c_139632053.htm/.

[33] FDA, "FDA approves first COVID-19 vaccine," 2021, https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine.

[34] G. Barnbrook, *Language and Computers*, Edinburgh University Press, Edinburgh, Scotland, 1996.

[35] W. Y. S. Chou and A. Budenz, "Considering emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence," *Health Communication*, vol. 35, no. 14, pp. 1718–1722, 2020.

[36] J. J. Jordan, E. Yoeli, and D. G. Rand, "Don't get it or don't spread it: comparing self-interested versus prosocial motivations for COVID-19 prevention behaviors," *Scientific Reports*, vol. 11, no. 1, pp. 1–17, 2021.

[37] J. Heffner, M. L. Vives, and O. FeldmanHall, "Emotional responses to prosocial messages increase willingness to self-isolate during the COVID-19 pandemic," *Personality and Individual Differences*, vol. 170, Article ID 110420, 2021.

[38] H. Lyu, J. Wang, W. Wu et al., "Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination," *Intelligent medicine*, 2021.

[39] S. F. Chung, "A corpus-based study of SARS in English news reporting in Malaysia and in the United Kingdom," *International Review of Pragmatics*, vol. 3, no. 2, pp. 270–293, 2011.

[40] L. Li, Y. Fan, A. Zeng, and Z. Di, "Understanding the Anticontagion Process and Reopening of China during COVID-19 via Coevolution Network of Epidemic and Awareness," *Complexity*, vol. 2021, Article ID 6623427, 11 pages, 2021.