



ARTICLE



<https://doi.org/10.1057/s41599-023-02241-5>

OPEN

What is newsworthy about Covid-19? A corpus linguistic analysis of news values in reports by *China Daily* and *The New York Times*

Siyang Liu¹ & Hailing Yu¹✉

Integrating corpus linguistic analysis with news values, this article investigates how the Covid-19 pandemic in different countries is represented as newsworthy by *China Daily* (CD) and *The New York Times* (NYT), two news outlets with different cultural and political backgrounds. A distinction is made between reports on the pandemic in the newspaper's home country and reports on the pandemic in other countries. Three corpus linguistic techniques, i.e., keywords list, collocation, and concordance, are used to analyze the headlines and leads of 2572 news reports published in 2020 and 2021. Analysis results show that when presenting the pandemic in domestic news, CD tends to highlight Proximity, Positivity, and Personalization, whereas NYT gives more prominence to Eliteness and Personalization. When the pandemic in other countries is presented, CD foregrounds Negativity, Impact, Superlativeness, and Eliteness, whereas NYT focuses on Negativity, Impact, and Proximity. Compared with NYT, CD shows a stronger tendency to adopt positive self-representation and negative other-representation in its coverage of the Covid-19 pandemic. Apart from the analysis results, the significance of the study also lies in its demonstration of the applicability of a corpus linguistic approach to news values analysis.

¹School of Foreign Languages, Hunan University, 2 South Lushan Road, 410082 Changsha, China. ✉email: hailing.yu@hnu.edu.cn

Introduction

Since the outbreak of Covid-19 in early 2020, news media have been playing a vital role in communicating public health information on the coronavirus disease (Mach et al., 2021) and correcting Covid-19-related misinformation (Lwin et al., 2023). Meanwhile, news media also promote the discursive construction and reinforcement of ‘othering’ in their representations of the pandemic, with out-group foreigners or minority groups such as migrants, refugees, and asylum seekers often held responsible for the disease and its spread (De Rosa and Mannarini, 2020; Richter et al., 2021). This is often characterized by a polarized representation between positive Self and negative Others. For instance, the study by Martikainen and Sakki (2021) shows that Finnish press reporting of Covid-19 in Sweden constructs an image of Sweden as “a warning example of misguided COVID-19 policy” (p. 401) and portrays Swedes as arrogant and immoral in the face of the health crisis, against which the success of Finnish policy and the superiority of virtuous and responsible Finns are highlighted. Al-Salman and Haider (2021) compared the representation of Covid-19 and China in the headlines of *Reuters* and *Xinhua*. Their findings demonstrate that *Reuters* underlines the immediate and long-term repercussions, serious implications, and consequences of the health crisis, whereas *Xinhua* tends to downplay the effects of Covid-19 outbreak on the Chinese economy and emphasize China’s readiness and ability to contain the virus.

While some previous studies have noted the differentiation between Us and Them in Covid-19-related news discourse (Martikainen and Sakki, 2021; Richter et al., 2021), media representations of the national Self and international Others in the context of the Covid-19 pandemic are still empirically under-explored, especially how media of different cultural and political backgrounds represent the pandemic in domestic and international news. To fill the gap, this study will explore how *China Daily* (CD) and *The New York Times* (NYT), two mainstream newspapers from China and the US, represent the Covid-19 pandemic in their home countries and in other countries from the perspective of news values.

As with other kinds of discourse analysis, news values analysis can be undertaken manually through in-depth reading of relevant texts, or it can be assisted by corpus techniques. Recently, more and more researchers have combined corpus linguistic methods with news values analysis to investigate media coverage so as to compensate for the limitations of qualitative approaches in terms of representativeness and generalizability (Potts et al., 2015; Bednarek et al., 2021; Zhang and Cheung, 2022). Following this line of research, this study proposes a three-pronged corpus linguistic approach to news values analysis that combines keywords list, collocation, and concordance. Specifically, this study will examine (1) what news values and how they are constructed in the headlines and leads of the reports by CD and NYT on the pandemic in the newspaper’s home country and in other countries; and (2) what aspects of the Covid-19 pandemic are highlighted or downplayed in the headlines and leads of the reports. The significance of the study resides in its focus on the Self- versus Other-representation in the time of the Covid-19 pandemic and the integration of corpus linguistic analysis with news values.

In what follows, we start with a literature review of the representation of the Covid-19 pandemic by Chinese and Western media. This is followed by an introduction to the corpus linguistic approach to news values analysis. The data and methodology are then explained before we present the analysis results. A discussion of the results and a conclusion are given at the end.

Representation of Covid-19 by Chinese and Western Media

Research on news media has shown that news coverage is usually not unbiased but rather linked to various sociocultural and

political factors (Van Dijk, 2013). Influenced by culture, ideology, political positions and media systems, news outlets in different countries may choose distinctive frames to represent similar or identical issues (Guo et al., 2012). Existing studies have indicated that Chinese and Western media display different emphases in their representations of the Covid-19 pandemic (Liu et al., 2020; Hubner, 2021; Wirz et al., 2021; Sing Bik Ngai et al., 2022).

When reporting on the Covid-19 pandemic, Chinese media tend to highlight positive topics and themes, although the pandemic’s impact is frequently mentioned. According to Liu et al. (2020), prevention and control procedures, medical treatment and research are the two major themes identified in Chinese media’s early coverage of the Covid-19 outbreak, whereas the global/local social/economic influences theme is relatively less prominent. Gong and Firdaus’s (2022) investigation of the news coverage of Covid-19 by *China Daily* reveals that although Covid-19 is initially depicted as a ‘bane’ that results in numerous deaths, economic decline and political tensions, it is later re-framed as a relative ‘boon’ reflected by positive themes like prosociality, cooperation, pandemic recovery and informativeness.

In addition, scholars have found that Chinese media representation of the Covid-19 outbreak is characterized by a combination of globalism and nationalism (Yang and Chen, 2020). Chinese official media outlets promote globalism by emphasizing China’s active role in helping the international community combat Covid-19, but also hold a strong nationalist stance by praising China’s success in fighting against the virus while criticizing Western countries’ ineffective reactions to the pandemic. This construction of nationalism creates a “polarizing ‘self-versus-other’ relationship between China and other countries” (ibid, p. 105).

Meanwhile, consistent with the representations of previous infectious disease outbreaks such as Ebola, SARS, and AIDS, Western media’s coverage of Covid-19 is characterized by labeling the virus as a foreign disease and threat. For instance, the study by Georgalakis (2021) shows that similar to Western media coverage of SARS, BBC’s early representation of Covid-19 emphasizes China’s alleged culpability for the outbreak and Western countries’ priority to protect themselves from a foreign virus. In many European news reports on Covid-19, the coronavirus is presented as caused by the poor food and hygienic practices of the Chinese, who have the strange habit of eating exotic animals (Pietrzak-Franger et al., 2022).

Another obvious feature of Covid-19 coverage by Western media, US media in particular, is a high degree of politicization, with political topics, actions and actors frequently being the focus of pandemic-related reporting (Hart et al., 2020; Tejedor et al., 2021; Schmidt, 2022). According to Tzopopoulos (2020), the Covid-19 pandemic is regarded by most Western media as “an opportunity to criticize the Chinese government, condemn its practices, and advocate for media freedom rather than focus on potential repercussions for global public health” (p. 147). Besides, there is a tendency for Western media to highlight the economic impacts of the pandemic (Basch et al., 2020; Hubner, 2021), and to give special attention to ordinary people affected by the pandemic (Matua and Oloo Ong’ong’a, 2020; Hubner, 2021).

The difference between Chinese and Western media in reporting Covid-19 becomes more prominent in comparative studies. For instance, after examining the main topics in the Covid-19 coverage by traditional news media in the US and China, Wirz et al. (2021) concluded that the Chinese coverage gives prominence to positive domestic responses to the pandemic, whereas the American coverage pays more attention to political conflicts. Sing Bik Ngai et al. (2022) also investigated the differences between US and Chinese mainstream news media’s

Table 1 News values and their definitions in this study.

News value	The pandemic or related news actors/events are constructed as...
Consonance	(stereo) typical (limited here to news actors, social groups, organizations, or countries/nations)
Eliteness	of high status or fame (including but not limited to the people, countries, or institutions involved)
Impact	having significant effects or consequences (not limited to impact on the target audience)
Negativity	negative
Personalization	having a personal or 'human' face (involving non-elite actors, such as patients)
Positivity	positive
Proximity	geographically or culturally near (in relation to the publication location/target audience)
Superlativeness	being of high intensity or large scope/scale
Timeliness	timely in relation to the publication date: as new, recent, ongoing, about to happen, current, or seasonal, etc.
Unexpectedness	unexpected, unusual, strange, or rare

coverage of Covid-19, with their particular focus on coping strategies and emotions. They found that *The New York Times* reports fewer themes on coping strategies in a more negative tone and highlighted negative emotions like fears and anxiety, while *People's Daily* presents a wider range of themes in a more positive way and foregrounds positive emotions such as cheer, gratitude and good wishes.

While news media function as important sources of information during a global health crisis, news narratives about the pandemic may influence and be influenced by other social discourses. For instance, users' comments on the Chinese community question-answering (CQA) site Zhihu showcase their positive assessments of the Chinese government and criticism of the British (and other Western) governments (Peng et al., 2020). Although many news outlets in the US adopt a critical stance against former president Donald Trump, they also share his politicization of the Covid-19 pandemic when dealing with China (Prieto-Ramos et al., 2020; Yaqub, 2020).

According to the ideological square framework proposed by Van Dijk (2006; 2013), the 'Self' is usually described in neutral or positive terms, whereas the 'Others' tend to be described in neutral or negative terms. Recent studies have found that the discourse strategy of positive Self-representation and negative Other-representation is adopted in news reporting on the Covid-19 pandemic by media from different countries (e.g., Dezhkameh et al., 2021; Abbas, 2022). However, most of these studies are limited to the representation of Covid-19 by newspapers from the same country (Yang and Chen, 2020; Tareen and Dilawer, 2021; Martikainen and Sakki, 2021). Comparative research on Self- and Other-representation by media of different cultural and political backgrounds during the time of a global health crisis is still rare. In view of this, the present study compares how CD and NYT represent the outbreak of Covid-19 in their own countries and in other countries in terms of new values. For this purpose, we will draw on a corpus linguistic approach to new values analysis.

Corpus linguistic approach to news values analysis

In the field of journalism and media studies, news values have been conceptualized from the material, social, and cognitive perspectives (Galtung and Ruge, 1965; Bell, 1991; Cotter, 2010; Harcup and O'Neill, 2017). To complement these studies, Bednarek and Caple (2012; 2014; 2017) proposed a discursive perspective for newsworthiness, whereby news values are defined as "the 'newsworthy' aspects of actors, happenings and issues as existing in and constructed through discourse" (2014, p. 137). Within the framework proposed by Bednarek and Caple (2017), there are ten news values that are discursively constructed through linguistic resources in news reporting. They include Consonance, Eliteness, Impact, Negativity, Personalization, Positivity, Proximity, Superlativeness, Timeliness and

Unexpectedness. Definitions of the news values in the context of Covid-19 news reports are provided in Table 1.

While analysis of newsworthiness can be undertaken via content analysis and manual coding (Dahl and Fløttum, 2017; Makki, 2019), many previous studies have demonstrated the usefulness of corpus linguistic techniques in news values analysis, especially when dealing with a large corpus. For example, Bednarek and Caple (2014) applied word/bigram frequency lists, keywords, and concordance analysis to a 70,000-word corpus of British news discourse; Potts et al. (2015) used a 36 million-word corpus of news reporting on Hurricane Katrina to test and evaluate the four corpus techniques of tagged lemma frequencies, collocation, key POS tags and key semantic tags. Different combinations of corpus techniques have been employed by researchers to examine news values in reports on various themes, such as cycling and cyclists (Bednarek and Caple, 2017), Australian National Day (Bednarek et al., 2021), Libyan civil war and violence (Attia, 2022) and the Kunming terrorist attack (Guo et al., 2022).

Of all the corpus techniques, automatic keyword analysis proves to be especially useful in comparative studies as keywords help reveal the differences between a target corpus and a reference corpus. Drawing on computer-assisted keyword analysis, prior studies have compared the discursive construction of newsworthiness between Chinese and Western media (Zhang and Caple, 2021; Zhang and Cheung, 2022). Although keywords can work as potential 'pointers' (Bednarek and Caple, 2014, p. 145) to specific news values, additional qualitative analysis is considered crucial as the construction of news values is highly sensitive to context (Potts et al., 2015; Bednarek, 2016; Bednarek et al., 2021). This can be undertaken through collocation and concordance analysis.

With insights from these previous studies, this paper aims to further explore the potential of computer-assisted approaches (keyword, collocation, and concordance analyses) in revealing the construction of news values in the reports on the Covid-19 pandemic by CD and NYT.

Data and methodology

The data used for this study were collected by visiting the websites of CD (<https://www.chinadaily.com.cn/>) and NYT (<https://www.nytimes.com/>). The keywords 'virus', 'coronavirus', and 'Covid-19' were used to search for the reports. The method of constructed week sampling was adopted. According to the specification of Song and Chang (2012), we used the calendar as the sampling frame. Within the time frame from 1 January 2020 to 31 December 2021, we randomly selected one day from each week to 'construct' a week to account for the cyclic variation of the news content. Since there is no clear distinction between weekday and weekend editions of online newspapers in the two selected media, both weekdays and weekends are counted. Altogether, 15 constructed weeks were obtained. News reports collected from both CD and NYT were

Table 2 Data composition of the study.

Newspapers	Sub-corpora	No. of reports	No. of words
<i>China Daily</i>	Reports on the pandemic in China	719	28,539
	Reports on the pandemic in other countries	641	24,921
	Sum	1360	53,460
<i>The New York Times</i>	Reports on the pandemic in the US	775	27,322
	Reports on the pandemic in other countries	437	15,343
	Sum	1212	42,665
	Total	2572	96,125

classified into two groups respectively: reports on the Covid-19 pandemic in the newspaper's home country (i.e., CD's reports on the pandemic in China and NYT's reports on the pandemic in the US) and reports on the pandemic in other countries. For both newspapers, reports that contain the keywords but do not focus on the pandemic and those not related to any specific country were excluded. Also excluded from the data were reports that compare the pandemic in the home country to that in other countries (which are very limited in number). The resulting corpus consists of 1360 news reports from CD and 1212 from NYT. The composition of the data is shown in Table 2.

In this study, we restricted our data collection to the headlines and leads for the following two reasons. First, the presentation of online news headlines and leads can have a considerable effect on news selection by the audience. Different from print newspapers, online news portals feature an entering page where only headlines and leads hyperlinked to full news reports are presented. In order to attract audiences to read the whole story, journalists may present the most 'newsworthy' information in the headlines and leads. Second, English-language hard news reporting features the "inverted pyramid" structure, an arrangement by which "what is 'most important information' comes first and what is less important comes after" (Thomson et al., 2008, p.8). The headline/lead nucleus therefore constitutes the "most important news element of the story" (Cotter, 2010, p. 162) and sets up the newsworthy focus or 'angle' of a reported event (White, 1997). The headline (including the sub-heading if there is any) of a news report can be identified without any difficulty, and the lead is usually located immediately below the headline and before the main text.

The headlines and leads were analyzed by using the corpus linguistic software Wmatrix (Rayson, 2003), which allows for keyword, concordance, and collocation analyses. Automatic keyword analysis provides the researcher with a list of keywords or clusters that are frequent (or infrequent) in a target corpus as compared to a reference corpus (Scott and Tribble, 2006). While most studies choose a very general corpus as a reference, such as the British National Corpus (BNC), it is more desirable for scholars to design a particular reference corpus in order to answer specific research questions. Since this study focuses on how news media cover Covid-19 in their home countries and in other countries, we reciprocally compared the two sub-corpora of CD as well as the two sub-corpora of NYT. As a result, four keyword lists were obtained. The keywords were then ranked according to their keyness values, which were calculated by using the log-likelihood (LL) statistical test. The higher the keyness value, the more frequently a keyword appears in the target corpus than in the reference corpus, and the more statistically significant it is (Partington, 2010). Following the method adopted in previous research (Zhang and Cheung, 2022), we retained the top 100 keywords in each sub-corpus for further qualitative analysis.

It is worth mentioning that we don't rely on the keyword itself to determine the news value because sometimes a word alone doesn't reveal any newsworthiness. In fact, we draw on the 'concordance tool' of Wmatrix to check how a keyword is used in

its immediate context, while at the same time consulting the inventory of linguistic devices constructing newsworthiness (Bednarek and Caple, 2017). Only in this way can a keyword be linked with a specific news value. For example, 'rate' is a keyword obtained through automatic analysis. Its concordance (Fig. 1) shows that in most cases (12 out of 17), 'rate' is used to describe negative situations, such as 'a general rise in infection rate', 'low vaccination rate', and 'Covid-19 test positivity rate tops 9 pct'. Therefore, the keyword 'rate' serves as a potential pointer to Negativity in this context. That is to say, a keyword will not be grouped into a news value category until it is verified through its concordance lines. In cases where a keyword points to two news values simultaneously, it will go to two different categories. Those keywords that are not clearly related to a specific news value or cannot reveal any potential differences between two sub-corpora were excluded from our analysis. Altogether 118 keywords from the CD corpus and 111 from the NYT corpus were retained. To complement keyword analysis, we also applied collocation analysis whereby the analyst can identify which other words a particular word typically co-occurs with within a given co-textual span. For example, in CD's reports on the pandemic in China, the keyword 'cases' usually co-occurs with words such as 'confirms', 'adds', 'asymptomatic', 'imported', and 'additional', etc. In accordance with the method taken by Potts et al. (2015), the log-likelihood (LL) statistic was employed here for identifying statistically significant collocates.

For identifying news values constructed in media coverage, it is necessary to consider the target readership of the news outlet. As is argued by Bednarek and Caple (2017), news values are dependent on target audiences, and "what is newsworthy to one target audience is not automatically newsworthy to another" (p.68). In the analysis, we tentatively regard American readers as the target audience of NYT and Chinese readers who are proficient in English as the target audience of CD. This is confirmed by the statistics from Alexa Traffic Rank, which shows the core audience of NYT comes from the US and the core audience of CD comes from China.

Findings

This section presents the analysis results of the news values constructed through keywords in the news reports by CD and NYT. The keywords are grouped into six categories according to the news values they are associated with. Negativity and Impact are included in the same category here since impact tends to be constructed as negative in public health emergency news reporting. No keywords pointing to Unexpectedness, Consonance, and Timeliness are found, and these three news values will not be discussed in the following.

News values in the reports by CD. Table 3 presents the keywords that contribute to establishing specific news values in CD's reports on the Covid-19 pandemic in China and in other countries. The keyness value of each keyword is given in the bracket.

17 occurrences.

surpassed 4,000 despite a significantly lower rate of transmission . Latvia toughens social dista
 showed that on a national level the infection rate has slowed to a trickle . Virus-hit Singapore
 successfully reduced its COVID-19 daily case rate from a peak of over 700 in early August to jus
 VID-19 cases amid a general rise in infection rate , three months after all municipal elderly car
 impact of the coronavirus . ECB refrains from rate cut , steps up liquidity measures amid COVID-1
 vaccine . India 's daily COVID-19 positivity rate doubles in 12 days . India 's daily COVID-19 p
 12 days . India 's daily COVID-19 positivity rate has doubled to 16.69 percent in 12 days amid a
 vaccinated . As Germany 's COVID-19 infection rate has risen to its highest recorded level yet ,
 ocks opened strongly Tuesday on news that the rate of new infection has declined recently in New
 ny Fauci said Sunday . NYC COVID-19 infection rate hits new high of 3.9 pct : mayor . New York Ci
 : mayor . New York City 's COVID-19 infection rate on 7-day average has reached a new height of 3
 ospital system . NYS COVID-19 test positivity rate , hospitalizations trend up : governor . Of th
 e on Wednesday . NYC COVID-19 test positivity rate up to 6.29 pct : mayor . New York City 's COVI
 r . New York City 's COVID-19 test positivity rate on a seven-day average went up to 6.29 percent
 sitive results . NYC COVID-19 test positivity rate tops 9 pct : mayor . New York City 's COVID-19
 r . New York City 's COVID-19 test positivity rate on a seven-day average topped 9 percent , reac
 ses traced to Delta variant , low vaccination rate : media . With the increase of confirmed COVID

Fig. 1 A screenshot shows the keyword in its immediate context. Concordance lines of 'rate'.

Table 3 Keywords and news values in the reports by CD.

CD-China sub-corpus	CD-Other countries sub-corpus	News values
epidemic (61.14), pneumonia (45.22), carriers (18.63), virus (17.93), cases (12.00), case (11.18), outbreak (10.97)	lockdown (44.59), deaths (42.77), positive (38.54), variant (31.32), restrictions (29.85), pandemic (28.33), toll (26.84), death (20.25), state of emergency (20.05), wave (18.57), fears (18.50), delta (16.75), rate (15.91), hospitalizations (12.34), concerned (12.34), concerns (12.01)	Negativity and Impact
province (205.71), mainland (163.92), Chinese (110.87), Beijing (106.42), China (106.11), Hubei (74.32), Hong Kong (67.06), Wuhan (65.37), Shanghai (62.22), city (44.69), Guangdong (40.98), Heilongjiang (34.77), East China (32.32), autonomous region (31.05), Tianjin (25.38), Yunnan (24.84), Shandong (24.84), Hebei (23.60), Northeast China (22.35), Jilin (19.87), Shenzhen (18.63), Inner Mongolia (18.63), Jiangsu (16.64), Dalian (16.14), Taiwan (16.14), Zhejiang (16.06), Shanxi (14.90), North China (14.90), Liaoning (14.90), HK (14.90), central China (13.81), Yantai (13.66), Hechi (13.66), HKSAR (13.66), South China (12.42), Shijiazuhang (11.18), Qingdao (11.18), Nanjing (11.18), Macao (11.18), Harbin (11.18)	Japan (44.72), South Korea (20.05), India (17.78), Philippines (15.42), Singapore (13.88)	Proximity
-	surge (38.67), spike (16.96), record (16.75), nationwide (14.08), surpass (13.88), topped (12.34), surpassed (11.32), reached (11.32)	Superlativeness
nucleic (40.98), fight (39.11), acid (37.26), administered (37.05), control (28.77), prevention (27.20), testing (18.35), medical (17.25), consumption (13.66), TCM (13.66), medicine (12.42), platform (11.18)	vaccine (37.89), vaccination (28.36)	Positivity
commission (136.19), provincial (42.22), Xi (34.77), health (31.00), national (30.67), local authorities (21.11), foreign (17.88), municipal (12.69), health authority (12.42)	prime minister (50.89), Trump (38.55), Donald Trump (33.92), Johns Hopkins (27.76), President (24.92), White House (20.05), PM (18.50), federal (16.96), FDA (16.96), ministry of health (15.42), Biden (15.42), administration (14.18), minister (13.94), CDC (12.01)	Eliteness
medical (17.25), patient (13.95), residents (13.50), patients (11.01)	Americans (23.13), adults (12.34)	Personalization

As is shown in Table 3, there are more keywords pointing to Proximity, Positivity, and Personalization in CD's domestic reports. By contrast, CD's reports on the pandemic in other countries contain more keywords that construct Negativity, Impact, Superlativeness, and Eliteness. How these news values are constructed through keywords will be discussed in the following.

Negativity, Impact, and Superlativeness. Negativity, Impact, and especially Superlativeness are highlighted more in CD's reports on the pandemic in other countries than in its reports on the pandemic in China. In reporting the pandemic's negative impact on China, CD prefers to use negative lexis within the semantic domain of the novel coronavirus, including the term 'virus', the

11 occurrences.

the number of COVID-19 cases in Australia has surpassed 4,000 despite a significantly lower rate of US state by both area and population, has surpassed New York state in total confirmed coronavirus a positive coronavirus test in Britain has surpassed 100,000 after another 1,631 have been confirmed 11,800 as more testing was done, and Italy surpassed China for the most coronavirus-related death advocates masks as virus surges. As the US surpassed 3.3 million cases of the novel coronavirus of opkins University. US COVID-19 deaths have surpassed the 150,000 mark to reach 150,034 as of 3:35 p.m. the number of COVID-19 cases in the United States surpassed the 6 million mark on Monday, according to the CDC. COVID-19 deaths top 350,000. US COVID-19 deaths surpassed 350,000 on Saturday evening, according to the CDC. the number of COVID-19 cases in the United States surpassed 29 million on Monday, according to the CDC. the number of COVID-19 cases in the United States surpassed 30 million on Wednesday, according to the CDC. continue. US COVID-19-related deaths on Monday surpassed 675,000, the estimated US fatalities from the

Fig. 2 A screenshot shows the keyword in its immediate context. Concordance lines of 'surpassed'.

11 occurrences.

ed positive for the new coronavirus in Italy reached 12,839 on Thursday, Italian authorities said the number of COVID-19 cases in Indonesia had reached 34 with one death, the Indonesian government ABA - The number of confirmed COVID-19 cases reached 10,075 in Africa and the death toll reached 4 reached 10,075 in Africa and the death toll reached 487, the Africa Center for Disease Control and Prevention reported. The total number of confirmed cases in the country reached 4,421. Libya reports 1st COVID-19 recovery center who tested positive for the novel coronavirus reached 12,107 as of Monday afternoon, a daily increase of COVID-19 pandemic on the African continent reached 1,242 as the number of confirmed cases hit 25 million. The vaccination drive against COVID-19 on Wednesday reached the population of healthy adults after nearly 100 million people had been vaccinated. On Tuesday they hoped an agreement can be reached with the United States on the reciprocal lift of the number of COVID-19 cases in the United States reached 19,624 as of 11 pm on Friday (0300 GMT on Saturday). The COVID-19 infection rate on 7-day average has reached a new height of 3.9 percent, compared with 3

Fig. 3 A screenshot shows the keyword in its immediate context. Concordance lines of 'reached'.

relevant consequences of virus infection ('epidemic', 'pneumonia', 'outbreak'), as well as references to people who are infected ('cases', 'carriers'). In contrast, CD's reports on other countries put less emphasis on the pandemic per se, but more on the negative outcome brought by the pandemic. This can be seen through frequent references to the pandemic's negative effects on people's health and life ('deaths', 'positive', 'hospitalizations'), different countries' negative social and political reactions to the pandemic ('lockdown', 'restrictions', 'state of emergency') as well as people's negative emotions ('fears', 'concerns').

Keywords pointing to Superlativeness are absent from CD's domestic reports, whereas in CD's international news keywords such as 'surge', 'spike', and 'surpassed' are frequently used to describe the severity of the pandemic. A close examination of the concordance lines shows that nearly all instances of 'surge' and 'spike' construe the news value of Superlativeness through descriptions of the sharp increase in Covid-19 infections or deaths (see Examples 1 and 2).

- (1) US hospitals overwhelmed with Covid-19 patients as cases surge: media
- (2) Italy sees biggest daily spike in Covid-19 deaths

Collocated with numerals, the keywords 'surpass', 'surpassed', 'topped', and 'reached' are also used to construct Superlativeness. Concordance lines of the keywords 'surpassed' and 'reached' in Figs. 2 and 3 show that of the twenty-two occurrences, sixteen are followed by a large number. This functions to represent the Covid-19 pandemic in other countries as affecting a large number of people. Besides, many occurrences are used to describe the serious status quo of the pandemic in the US. The majority of instances (9 out of 13) of the keyword 'nationwide' are used to

indicate the large scope of the reported event, evidenced by such expressions as 'nationwide lockdown', 'capacity is overwhelmed nationwide', and 'nationwide virus resurgence', etc.

Positivity. Positivity is more prominent in CD's reports on the pandemic in China than in its reports on the pandemic in other countries. In CD's reports on the pandemic in China, Positivity is mainly constructed through references to anti-virus actions and measures taken by the Chinese government, such as rounds of mass nucleic acid testing being conducted or millions of Covid-19 vaccine doses being administered, as can be seen through the concordance lines of the keywords 'nucleic', 'acid', 'administered' and 'testing'. Another two keywords foregrounded in the construction of Positivity are 'prevention' and 'control'. Concordancing shows that the two keywords are often used together to refer to China's scientific and effective prevention and control efforts to contain the spread of the deadly disease. Although the keyword 'fight' might be considered negative in many cases, concordances show that it is metaphorically used in the data to refer to the efforts made for pandemic control, as shown in the following two examples.

- (3) Quanzhou businesspeople help fight coronavirus outbreak
- (4) Volunteers sew masks for fight against novel coronavirus outbreak

Also included in the CD-China sub-corpus are keywords 'consumption', 'TCM', 'medicine', and 'platform', which are mentioned in reports on newly launched online platforms offering information and treatment services, on measures unveiled to promote consumption growth or on traditional

Table 4 Key collocates of ‘medical’.

Key collocates	workers	supplies	institutions	teams	team	staff
Log-likelihood	77.84	62.29	52.76	41.46	36.34	32.08

Chinese medicine’s effectiveness in fighting Covid-19. This is illustrated in Examples 5 to 7.

- (5) Internet medical service *platform* helps coronavirus control
- (6) China details measures to boost *consumption*, and mitigate virus impact.
- (7) *Traditional Chinese Medicine (TCM)* is being adopted as a means of treatment for novel coronavirus pneumonia patients in Wuhan...

The keywords in CD’s domestic news also include ‘medical’. This keyword, as is shown in Table 4, frequently co-occurs with ‘workers’, ‘supplies’, ‘institutions’, ‘team(s)’, and ‘staff’. An investigation of the co-text of ‘medical supplies’ shows such expressions as ‘overseas Chinese communities send medical supplies to fight coronavirus’, ‘HK firms donate money, medical supplies to mainland’, indicating people’s support and help for each other amid the pandemic. When ‘medical workers/teams/staff’ appear in the reports, they are often associated with positive actions, thus constructing Positivity and Personalization at the same time (Examples 8 and 9).

- (8) *Medical workers make inroads* in battle against coronavirus
- (9) As of Feb 17, over 2000 *medical workers* in 17 medical teams *have been fighting hard* for a week since they took over 17 intensive care units...

In comparison, in CD’s reports on the pandemic in other countries, there are only two keywords contributing to the construal of Positivity: ‘vaccine’ and ‘vaccination’. Concordancing shows that Positivity is established through descriptions of new Covid-19 vaccine development or approval and the launch of the mass vaccination campaign.

Eliteness and personalization. Eliteness is less emphasized in CD’s domestic news than in its international news. References to the names and titles of political figures such as ‘Xi’, ‘Trump’, ‘Biden’, ‘prime minister’, ‘president’, and ‘ministry of Health’ are frequently used in both sub-corpora to construe Eliteness. Eliteness is sometimes combined with Positivity or Negativity, as elites can be judged negatively or positively. For example, the Chinese President Xi Jinping is usually mentioned with positive or neutral descriptions whereas the former American President Trump is often represented in a negative way (Examples 10 and 11). Eliteness is also construed through references to central and local governmental institutions such as ‘commissions’, ‘local authorities’, ‘The White House’, ‘The FDA’ and ‘CDC’, which often serve as sources of information.

- (10) *Xi* extends condolences to people who died from coronavirus
- (11) *Trump* still doesn’t understand the severity of COVID-19 pandemic: CNN

Personalization is slightly more emphasized in CD’s domestic reports than in its international reports. The most frequent keywords categorized as pointers to Personalization are general references such as ‘patients’ and ‘residents’. The concordance analysis shows that of 60 instances of ‘patients’, 19 are correlated with positive expressions like ‘recovered’, ‘discharged’, ‘released’, and ‘cured’, highlighting the patients’ recovery from the disease

(Example 12). As mentioned before, the keyword ‘medical’ can be combined with other words to refer to ‘medical workers/team/staff’, thus constructing the news value of Personalization. Different from the positive tendency observed in CD’s domestic reports, Personalization tends to co-occur with Negativity in CD’s reports on the pandemic in other countries. This is shown in Example 13, where the keyword ‘Americans’ is mentioned with descriptions of their negative situations.

- (12) A total of 328 *patients* infected with the novel coronavirus *had been discharged from hospital* after recovery...
- (13) More than 5 million *Americans, ..., have missed their scheduled second dose*, US media reported.

Proximity. Proximity is constructed when places familiar to the target audience are mentioned. As can be seen in Table 4, Proximity is a prominent news value in CD’s domestic reports. In contrast, only a few keywords pointing to Proximity are found in its coverage of the pandemic in other countries. Keywords that construct Proximity in CD’s reports on the pandemic in China are mainly references to different places in China, such as Beijing, Hong Kong, Wuhan, etc. Country names that construct Proximity in CD’s reports on the pandemic in other countries include Japan, South Korea, India, the Philippines, and Singapore.

News values in the reports by the NYT. Table 5 presents the keywords that contribute to establishing specific news values in NYT’s reports on the Covid-19 pandemic in the US and in other countries. The keyness value of each keyword is given in the bracket. It can be seen that Eliteness and Personalization are highlighted more in NYT’s domestic news, and Negativity, Impact, and Proximity are more salient in its international news. Superlativeness and Positivity are constructed with similar (low) frequencies. The construction of these news values through keywords will be discussed in the following.

Negativity, Impact, and Superlativeness. Negativity and Impact are less emphasized in NYT’s reports on the pandemic in the US than in its reports on the pandemic in other countries, as fewer keywords pointing to the two news values are identified in the NYT-US sub-corpus. The keywords ‘school’ and ‘schools’ serve as pointers to Negativity and Impact, because the two words are frequently used to refer to the negative effects of the Covid-19 outbreak on the opening of schools. The keyword ‘schools’ is frequently collocated with ‘closed’ (Log-Likelihood=40.96) in the data, establishing the news values of Negativity and Impact (Example 14). The same is true of the keyword ‘travel’, as concordancing reveals that it is often used to refer to the negative economic impact of the pandemic on the traveling industry (Example 15).

- (14) Gov. Andrew M. Cuomo of New York announced that *schools would stay closed* through the end of the school year...
- (15) South Africa detects a new variant, prompting new international *travel restrictions*.

Figure 4 shows that out of nine occurrences of the keyword ‘led’, eight construct cause-effect relations and are followed by references to negative consequences, as in ‘led to unnecessary deaths’ and ‘led to a chaotic virus crackdown’, etc. Although the keyword ‘rates’ is not a candidate for Negativity at first glance, its concordance lines display such expressions as ‘alarming rates of coronavirus infection’, ‘disparate rates of sickness and death’, ‘rising virus rates’, and ‘low vaccination rates’, thus representing the pandemic in the US as negative.

Table 5 Keywords and news values in the reports by NYT.

NYT-US sub-corpus	NYT-Other countries sub-corpus	News values
school (23.02), schools (17.21), rates (12.87), led (8.06), hospitalized (7.16)	lockdown (42.14), restrictions (26.28), protest (18.30), travel (15.07), flouting (8.15), ill (8.15), blow (8.15), limit (8.15)	Negativity and Impact
New York (65.51), state (57.68), California (37.61), New York City (32.70), States (31.90), NYC (24.18), US (19.92), Texas (16.12), Florida (16.12), Los Angeles (13.43), New Jersey (10.75), county (8.95), Chicago (7.16), NY (6.27), south (5.37)	Britain (77.45), Europe (45.67), Italy (36.69), Australia (36.69), France (34.65), UK (32.61), Canada (30.97), European (24.46), Germany (23.99), Spain (22.42), European Union (20.38), England (20.38), EU (20.38), New Zealand (14.27), Melbourne (12.23), Brazil (12.13), Latin America (10.19), London (10.19), Ireland (8.15), Greece (8.15)	Proximity
continued (8.06)	far (10.56)	Superlativeness
testing (12.16), guidance (11.64), plan (6.10)	inoculation (10.19), approves (10.19), celebrate (8.15)	Positivity
Trump (74.04), Biden (55.22), Cuomo (41.19), Gov. (39.40), White House (35.82), President (31.35), the mayor (27.76), administration (26.77), governor (25.97), federal (21.43), CDC (19.70), Andrew (17.91), republican (15.22), office (12.54), House (11.64), Bill de Blasio (11.64), Democrats (10.75), FDA (10.75), governors (9.85), officers (8.95), Congress (8.95), Senate (8.06), centers (7.23), Jr. (7.16), Moderna (6.47), Newson (6.27), Fauci (6.27), Joseph. R. Biden (6.27), senator (5.37), Republicans (5.37)	prime minister (79.49), Boris Johnson (50.95), WHO (26.50), government (23.14), astrazeneca (20.38), leader (16.51), official (14.42), parliament (12.23), UN (12.23), Angela Merkel (10.19), World Health Organization (9.70), Xi (8.15), minister (8.15), Communist Party (8.15), Narendra Modi (8.15), Merkel (8.15), Macron (8.15), Emmanuel Macron (8.15)	Eliteness
Americans (22.87), students (22.44), (New) Yorkers (22.39), children (20.74), black (10.75), employees (10.32)	citizens (24.46), Italians (10.19)	Personalization

9 occurrences.

elieve she passed away , " said a relative who led the family 's increasingly frantic search . " B
 handling of the coronavirus outbreak in China led to unnecessary deaths . The organization has de
 tion Fell Apart in Weeks . The coronavirus has led to the collapse of New York 's usual admissions
 crisis caused by the coronavirus outbreak has led young people from elsewhere to leave the city ,
 ing bars and bringing back football teams have led to new outbreaks . Communities that evolved aro
 istry . How a Feud Between Cuomo and de Blasio Led to a Chaotic Virus Crackdown . The mayor pushed
 fferent approaches for calculations . That has led to confusion and has a big impact on which rest
 f outbreaks connected to youth ice hockey have led many states in the Northeast to restrict it . P
 eve Americans overseas in the initial outbreak led to safety risks , a new report says . The feder

Fig. 4 A screenshot shows the keyword in its immediate context. Concordance lines of 'led'.

Apart from references to the governments' negative reactions to the pandemic, such as 'lockdown' and 'restrictions', keywords that construct Negativity and Impact in NYT's reports on other countries also include references to people's negative actions triggered by the preventive measures ('protest', 'flouting'), as is shown in Examples 16 and 17.

- (16) Thousands *Protest* New Zealand's Covid-19 Mandates
- (17) Londoners rallied in the face of the deadly German Blitz, many are *flouting* the rules aimed at preventing the spread of the novel coronavirus.

Superlativeness is constructed, but not highlighted in NYT's domestic and international news. While the word 'continue' is classified by Bednarek and Caple (2017) as an implicit time reference pointing to Timeliness, concordance lines with 'continued' as the node provide evidence that the word in this context actually constructs the news value of Superlativeness. As is shown in Fig. 5, seven out of nine instances of the keyword 'continued' are used to describe a negative action or situation such as 'The outbreak's toll continued to rise' and 'the continued spread of the virus'. In this way, the news values of Superlativeness and Negativity are constructed simultaneously. Concordance lines containing 'far' show that it can be used with comparative forms to intensify the severity of the reported news

event (Examples 18 and 19). Negativity is co-constructed here in that negative situations are highlighted.

- (18) Symptomless transmission makes the coronavirus *far harder* to fight.
- (19) As bleak a milestone in the pandemic as the new figures are, unconfirmed cases are believed to be *far higher*.

Positivity. Positivity is also constructed, but not highlighted in NYT's domestic and international news. In NYT's reports on the pandemic in the US, keywords that help construct Positivity include 'testing', 'guidance' and 'plan'. Concordancing shows that they are used in the data to refer to active measures taken to deal with the pandemic, including regular virus testing, governments' relief and rescue plans, and CDC's health guidance, etc. This is illustrated in the following three examples.

- (20) NYC Unveils New *Testing* Program To Keep Children in School
- (21) Bill de Blasio unveiled a \$635 million *plan* to restore the system post-lockdown....
- (22) Biden Celebrates New CDC Mask *Guidance*

The keywords identified in NYT's reports on the pandemic in other countries include 'inoculation', 'approves' and 'celebrate',

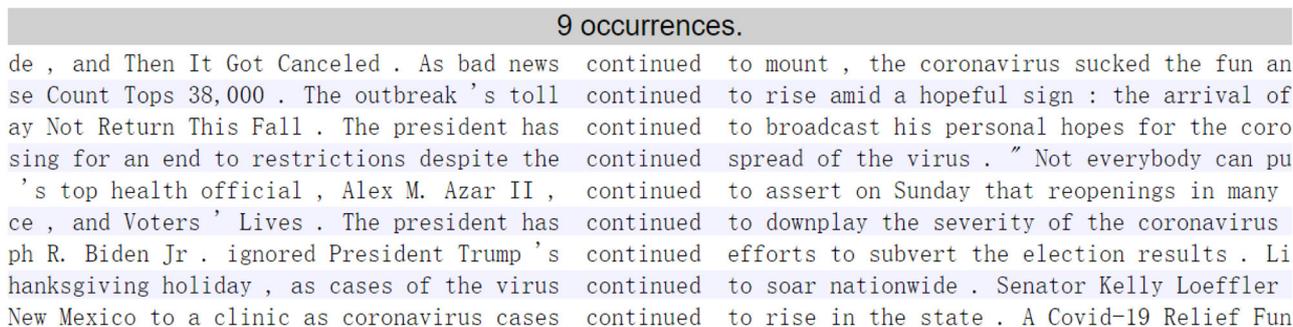


Fig. 5 A screenshot shows the keyword in its immediate context. Concordance lines of 'continued'.

which contribute to constructing Positivity in such instances as 'The country is beginning a large inoculation effort', 'India Approves Oxford-AstraZeneca Covid-19 Vaccine', 'Israelis celebrate a festival of freedom', etc.

Eliteness and personalization. Eliteness is slightly more highlighted in NYT's domestic news than in its coverage of the pandemic in other countries. References to national leaders (e.g., 'Trump', 'Boris Johnson'), recognized organizations (e.g., 'WHO', 'UN'), as well as coronavirus vaccine companies ('Moderna', 'AstraZeneca') contribute to establishing Eliteness in all reports by NYT. Concordances for the keyword 'Trump' show that Trump is often criticized by NYT, as is manifested in such headlines as 'Trump Attacks WHO and Ousts Watchdog for Pandemic Fund' and 'Trump Falsely Tries to Tie Obama to CDC's 'Inadequate' Testing System', etc., where his negative behaviors are highlighted. The names and titles of city mayors and state governors are also frequently mentioned (e.g., 'Cuomo', 'Gov.') in NYT's domestic reports. Besides, references to non-government actors of high status and fame such as the American health expert 'Fauci' also construct Eliteness. It is worth noticing that references to party politicians such as 'Democrats' and 'Republicans' are unusually frequent in the NYT's domestic coverage and hence become keywords.

As is shown in Table 4, Personalization is more prominent in NYT's domestic reports as more keywords are found in the NYT-US sub-corpus. General references to ordinary people such as 'Americans', '(New) Yorkers', 'children', 'students' and 'citizens' construct Personalization. This is a type of weak Personalization (Bednarek, 2016) since no personal information is provided. An examination of all concordance lines for the keyword 'Americans' shows that the negative situations or emotions of American people are presented in many headlines and leads (Examples 23 and 24). This gives rise to the co-occurrence of Personalization and Negativity.

- (23) More than a third of *Americans* live in areas where hospitals are running critically short of intensive care beds.
- (24) Native *Americans* Feel Devastated by the Virus Yet Overlooked in the Data

Proximity. In NYT's reports, Proximity is constructed through references to locations that are geographically or culturally close to American readers. It is more foregrounded in NYT's coverage of the pandemic in other countries than in its domestic news. In NYT's reports on the pandemic in the US, Proximity is construed when the names of domestic states or cities ('New York', 'Texas', etc.) are mentioned. In NYT's coverage of the pandemic in other countries, countries that are culturally close to the US are mentioned more frequently than those that are geographically

proximate. The former group includes Britain, Europe, Italy, Australia, France, Germany, Spain and New Zealand, among which Britain receives the most frequent coverage. The latter group includes Canada, Brazil, and Latin America. This kind of 'proximity-created newsworthiness' (Joye, 2010, p. 594) reflects the Euro-American-centric nature of NYT coverage where stories in the non-Western world are underrepresented.

Discussion and conclusion

Analyses of news values through keywords show similarities and differences between CD and NYT in representing the Covid-19 pandemic in their home countries and in other countries. Both media foreground the news values of Proximity, Eliteness, Personalization, Negativity, and Impact in presenting the pandemic in their home countries. That is to say, the pandemic in the domestic news tends to be represented as proximate, negative, impactful, involving many elites and affecting the lives of ordinary people. Moreover, as fewer keywords pointing to Negativity and Impact are identified in the domestic news than in the international news, both CD and NYT represent the pandemic in their own countries as less negative and impactful than that in other countries. This echoes the finding of a previous study that the negative impact of domestic crises tends to be played down by both Chinese and American media (Yu and Chen, 2023).

Both CD and NYT foreground the news value of Eliteness, but construct it in different ways. When covering the pandemic in China, CD gives more prominence to government institutions than to politicians. For example, the word 'commission', which is used to refer to '(local/national) health commission', appears as a keyword with a rather high keyness value of 136.19 as compared to that of the keyword 'Xi' (34.77, cf. Table 3). Conversely, in NYT's reports on the pandemic in the US, political figures such as Trump and Biden receive more coverage than governmental institutions, pointing to the politicization in Covid-19 news coverage by the American media (Hart et al., 2020).

In addition, the two newspapers show differences in terms of Positivity, which is foregrounded in CD's reports on the pandemic in China but not in NYT's reports on the pandemic in the US. As is discussed in the previous section, Positivity is frequently constructed in CD's domestic coverage through war metaphors (Yang, 2021) such as 'fight against the virus', which highlights the actions taken to control the health crisis and also boost the morale of the Chinese people. That the Covid-19 pandemic in China is presented by CD as less negative and more positive than the pandemic in other countries confirms prior findings that there is a tendency for the Chinese media to create a contrast between positive Self and negative Others in its representation of the Covid-19 outbreak (Yang and Chen, 2020). However, this positive self-representation is less prominent in NYT's reports,

which may be influenced by the consistent negativity in US major media coverage of Covid-19 (Sacerdote et al., 2020). As mentioned before, this negative trend is also found in the construction of Eliteness and Personalization in NYT's domestic reports, where the keywords 'Trump' and 'Americans' are often associated with negative descriptions.

As for the representation of the pandemic in other countries, both newspapers highlight Eliteness, Negativity, and Impact, portraying the pandemic in other countries as having a negative influence on society and involving many elites. Besides, the pandemic in other countries is represented as more negative and impactful than in the newspapers' home countries. When constructing Eliteness in international news, both CD and NYT pay more attention to political figures than to health experts. Although previous studies have found that Covid-19 coverage in US newspapers is highly politicized (Schmidt, 2022), this finding demonstrates that CD follows suit in giving much attention to political actors when reporting on the pandemic in other countries. In addition, both CD's and NYT's international news de-emphasize Positivity and Personalization. While Positivity has been recognized as a salient feature in most health-related stories published by Chinese newspapers (Peng and Tang, 2010), it is shown in our study that such salience might be restricted to domestic news.

In spite of these similarities, CD and NYT show differences in Superlativeness, which is highlighted in CD's international news through the linguistic resources of intensification and qualification, such as 'surge' and 'surpass(ed)', yet backgrounded in all the reports by NYT. Since Superlativeness is often correlated with Negativity and Impact, its prominence in the international news by CD further intensifies the negative Other-representation. It can be said that compared with NYT, CD shows a stronger tendency to adopt positive Self-representation and negative Other-representation in its coverage of the Covid-19 pandemic.

To conclude, this study contributes new insights into mainstream media story-telling of Covid-19 in different countries by expanding beyond existing work to look at the Self- versus Other-representation. It demonstrates how US and Chinese media converge and diverge in covering Covid-19 in their home countries and other countries. Despite differences in media systems, culture, and political background, both CD and NYT represent the pandemic in their own countries as less negative than in other countries, which is consistent with previous cross-cultural comparative studies on crisis coverage (Yu and Chen, 2023). However, CD and NYT differ in terms of Positivity and Superlativeness, and CD portrays a more positive Self and a more negative Other than NYT does.

Methodologically, this study combines automated quantitative analysis (identification of keywords and collocations) with qualitative concordance analysis, showcasing the effectiveness of corpus linguistic techniques for analyzing news values. This allows the analysis of large datasets with more time efficiency than manual content analysis, which can hopefully be applied to future corpus-based news value studies.

However, there are a number of limitations in this study. For instance, we only examined news values constructed in the immediate co-texts of certain keywords without analyzing the whole text, which resulted in omissions of news values that are established in other parts of the text. Second, we limited our analysis to the headlines and leads of reports and there is no guarantee that the results obtained in this analysis are fully representative of the results from analysis of entire news reports. Also, the data of the study is limited to two newspapers, and only three corpus techniques are adopted. It is recommended that further studies apply more techniques to a wider range of data to validate the corpus linguistic approach to news values

analysis. Notwithstanding these limitations, it is worth noting that the overall result of this study is consistent with that of another study (Yu and Liu, 2023) that adopts full-text manual analysis, proving the reliability of the methodology in the current study.

Data availability

All data analyzed during this study can be accessed at <https://www.researchgate.net/publication/371701680>.

Received: 14 October 2022; Accepted: 11 October 2023;

Published online: 26 October 2023

References

- Abbas AH (2022) Politicizing the pandemic: a schemata analysis of COVID-19 news in two selected newspapers. *Int J Semiot Law-Revue Internationale de Sémiotique Juridique* 35(3):883–902
- Al-Salman S, Haider AS (2021) The representation of Covid-19 and China in Reuters' and Xinhua's headlines. *Search (Malaysia)* 13(1):93–110
- Attia S (2022) Libya, the media and the language of violence: a corpus-assisted discourse analysis. *Res Corpus Linguist* 10(1):84–116
- Basch CH, Kecojevic A, Wagner VH (2020) Coverage of the Covid-19 pandemic in the online versions of highly circulated US daily newspapers. *J Community Health* 45(6):1089–1097
- Bednarek M (2016) Investigating evaluation and news values in news items that are shared through social media. *Corpora* 11(2):227–257
- Bednarek M, Caple H (2012) 'Value added': language, image and news values. *Discourse Context Media* 1(2-3):103–113
- Bednarek M, Caple H (2014) Why do news values matter? Towards a new methodological framework for analysing news discourse in critical discourse analysis and beyond. *Discourse Soc* 25(2):135–158
- Bednarek M, Caple H (2017) The discourse of news values: how news organizations create newsworthiness. Oxford University Press, New York
- Bednarek M, Caple H, Huan C (2021) Computer-based analysis of news values: a case study on national day reporting. *J Stud* 22(6):702–722
- Bell A (1991) The language of news media. Blackwell, Oxford
- Cotter C (2010) News talk: investigating language journalism. Cambridge University Press, Cambridge
- Dahl T, Fløttum K (2017) Verbal-visual harmony or dissonance? A news values analysis of multimodal news texts on climate change. *Discourse Context Media* 20:124–131
- De Rosa AS, Mannarini T (2020) The "invisible other": social representations of COVID-19 pandemic in media and institutional discourse. *Pap Soc Represent* 29(2):5.1–5.35
- Dezhkameh A, Layegh N, Hadidi Y (2021) A critical discourse analysis of Covid-19 in Iranian and American newspapers. *GEMA Online J Lang Stud* 21:3
- Galtung J, Ruge MH (1965) The structure of foreign news: the presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *J Peace Res* 2(1):64–90
- Georgalakis J (2021) Countdown to a pandemic: BBC news representation of the response to Covid-19 as a global crisis approached. Available at SSRN: <https://doi.org/10.2139/ssrn.3931855>
- Gong J, Firdaus A (2022) Is the pandemic a boon or a bane? News media coverage of Covid-19 in China Daily. *Journalism Pract.* <https://doi.org/10.1080/17512786.2022.2043766>
- Guo J, Mast J, Vosters R et al. (2022) Convergence or divergence? A computer-assisted analysis of how Chinese state-sponsored and market-oriented newspapers discursively construct the newsworthiness of the Kunming terrorist attack. *Discourse Context Media* 45:100578
- Guo L, Holton A, Jeong SH (2012) Transnational comparative framing: a model for an emerging framing approach. *Int J Commun* 6(1):1918–1941
- Harcup T, O'Neill D (2017) What is news? News values revisited (again). *J Stud* 18(12):1470–1488
- Hart PS, Chinn S, Soroka S (2020) Politicization and polarization in Covid-19 news coverage. *Sci Commun* 42(5):679–697
- Hubner A (2021) How did we get here? A framing and source analysis of early Covid-19 media coverage. *Commun Res Rep* 38(2):112–120
- Joye S (2010) News discourses on distant suffering: a critical discourse analysis of the 2003 SARS outbreak. *Discourse Soc* 21(5):586–601
- Liu Q, Zheng Z, Zheng J et al. (2020) Health communication through news media during the early stage of the Covid-19 outbreak in China: digital topic modeling approach. *J Med Internet Res* 22(4):e19118

- Lwin MO, Lee SY, Panchapakesan C et al. (2023) Mainstream news media's role in public health communication during crises: assessment of coverage and correction of COVID-19 misinformation. *Health Commun* 38(1):160–168
- Mach KJ, Salas Reyes R, Pentz B et al. (2021) News media coverage of COVID-19 public health and policy information. *Hum Soc Sci Commun* 8(1):1–11
- Makki M (2019) 'Discursive news values analysis' of Iranian crime news reports: perspectives from the culture. *Discourse Commun* 13(4):437–460
- Martikainen J, Sakki I (2021) Boosting nationalism through COVID-19 images: multimodal construction of the failure of the 'dear enemy' with COVID-19 in the national press. *Discourse Commun* 15(4):388–414
- Mutua SN, Oloo Ong'ong'a D (2020) Online news media framing of Covid-19 pandemic: probing the initial phases of the disease outbreak in international media. *Eur J Interact Multimed Educ* 1(2):e02006
- Partington A (2010) Modern diachronic corpus-assisted discourse studies (MD-CADS) on UK newspapers: an overview of the project. *Corpora* 5(2):83–108
- Peng AY, Zhang IS, Cummings J et al. (2020) Boris Johnson in hospital: a Chinese gaze at Western democracies in the COVID-19 pandemic. *Media Int Aust* 177(1):76–91
- Peng W, Tang L (2010) Health content in Chinese newspapers. *J Health Commun* 15(7):695–711
- Pietrzak-Franger M, Lange A, Söregi R (2022) Narrating the pandemic: covid-19, China and blame allocation strategies in Western European popular press. *Eur J Cult Stud* 25(5):1286–1306
- Potts A, Bednarek M, Caple H (2015) How can computer-based methods help researchers to investigate news values in large datasets? A corpus linguistic study of the construction of newsworthiness in the reporting on Hurricane Katrina. *Discourse Commun* 9(2):149–172
- Prieto-Ramos F, Pei J, Cheng L (2020) Institutional and news media denominations of COVID-19 and its causative virus: between naming policies and naming politics. *Discourse Commun* 14(6):635–652
- Rayson P (2003) Matrix: a statistical method and software tool for linguistic analysis through corpus comparison. Dissertation, Lancaster University
- Richter C, Al-Shami A, Khalifa S et al. (2021) The virus of the 'others'? Corona and discursive othering in Arab media. *J Arab Muslim Media Res* 14(1):3–24
- Sacerdote B, Sehgal R, Cook M (2020) Why is all COVID-19 news bad news? National Bureau of Economic Research. <https://doi.org/10.3386/w28110>
- Schmidt H (2022) Pandemics and politics: analyzing the politicization and polarization of pandemic-related reporting. *Newspaper Res J*. <https://doi.org/10.1177/07395329221095850>
- Scott M, Tribble C (2006) Textual Patterns: keyword and corpus analysis in language education. John Benjamins, Amsterdam
- Sing Bik Ngai C, Yao L, Gill Singh R (2022) A comparative analysis of the US and China's mainstream news media framing of coping strategies and emotions in the reporting of COVID-19 outbreak on social media. *Discourse Commun* 16(5):572–597
- Song Y, Chang TK (2012) Selecting daily newspapers for content analysis in China: A comparison of sampling methods and sample sizes. *Journal Stud* 13(3):356–369
- Tareen SR, Dilawer S (2021) Exploring ideological construction through COVID-19 related headlines in selected Pakistani newspapers. *Review of Applied Management and Social Sciences* 4(3):703–714
- Tejedor S, Cervi L, Tusa Jumbo F et al. (2021) The information of the Covid-19 pandemic on the front pages of newspapers. Comparative study of Italy, the United Kingdom, Spain, France, Portugal, the United States, Russia and Germany. *Revista mexicana de ciencias políticas y sociales* 66(242):251–291
- Thomson EA, White PR, Kitley P (2008) "Objectivity" and "hard news" reporting across cultures: Comparing the news report in English, French, Japanese and Indonesian journalism. *J Stud* 9(2):212–228
- Tzogopoulos GN (2020) Coronavirus and the western media. The Begin-Sadat Center for Strategic Studies. BESA Center Perspectives Paper 1492. <https://besacenter.org/perspectives-papers/coronaviruswestern-media/>. Accessed 15 Aug 2022
- Van Dijk T (2006) Ideology and discourse analysis *J Polit Ideol* 11(2):115–140
- Van Dijk T (2013) News as discourse. Routledge
- White P (1997) Death, disruption and the moral order: the narrative impulse in mass-media 'hard news' reporting. In: Christie F, Martin JR (eds) *Genres and institutions: social processes in the workplace and school*. Cassell, London, pp. 101–133
- Wirz CD, Shao A, Bao L et al. (2021) Media systems and attention cycles: volume and topics of news coverage on Covid-19 in the United States and China. *J Mass Commun Q* 99(4):1048–1047
- Yang Y, Chen X (2020) Globalism or nationalism? The paradox of Chinese official discourse in the context of the Covid-19 outbreak. *J Chin Polit Sci* 26:89–113
- Yang Z (2021) War metaphors in Chinese digital media coverage of Covid-19. *Network Knowl* 14(1):9–27
- Yaqub U (2020) Tweeting during the Covid-19 pandemic: sentiment analysis of Twitter messages by President Trump. *Digit Govern Res Pract* 2(1):1–7
- Yu H, Chen G (2023) Their floods and our floods: news values of flood photo galleries of Associated Press and Xinhua News Agency. *Journalism* 24(6):1362–1381
- Yu H, Liu S (2023) The pandemic in our country, the pandemic in their countries: news values and media representation of the COVID-19 pandemic. *Journalism Stud* 24(10):1257–1276
- Zhang L, Caple H (2021) The newsworthiness of Li Na—A critical comparative analysis of Chinese and international news media. *Lang Commun* 77:70–80
- Zhang W, Cheung YL (2022) The hierarchy of news values—a corpus-based diachronic and cross-cultural comparison of news reporting on epidemics. *J Stud* 23(3):281–301

Author contributions

The two authors jointly collected and analyzed the data. SL wrote the manuscript and HY supervised the study. They contributed equally to this work.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to Hailing Yu.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023