Literary Studies Conference PROCEEDINGS
E-ISSN: 2963-3532

Who are Affected by the COVID-19 Pandemic?
A Corpus Discourse Analysis in Indonesian News 2020

Arina Isti'anah
Universitas Sanata Dharma, Indonesia
arina@usd.ac.id

Abstract

The global COVID-19 pandemic has infected more than 6 million Indonesian people, and the number of deaths from the virus has surpassed 150,000. Since the first announcement on March 2, 2020, the number of COVID-19 cases in Indonesia has been increasing, followed by news on the number of cases, victims, regulations, and impact of the pandemic that dominated the discourse. As a form of ideology (re)production, news media can shape people's perspectives and behavior toward the pandemic. Through the lens of corpus discourse analysis, this paper explicitly addresses the lexeme terdampak 'affected' in Indonesian news corpora 2020. By the assistance of AntConc, a corpus analysis software tool, this paper involved 300,000 sentences or 4,557,173 tokens downloaded on corpora.uni-leipzig as the data to represent how Indonesian news media frame the impacts of COVID-19. The analysis shows that the lexeme terdampak occurs with noun, adjective, and verb collocates as follows: 'pekerja 'workers', miskin 'poor', PHK 'layoffs', informal 'informal', rentan 'vulnerable', UMKM '(Usaha Mikro, Kecil dan Menengah 'micro, small and medium enterprises (MSMEs), bantuan 'help (n)', meringankan 'to ease', dirumahkan 'laid off', and paket 'package'. Some of the noun and adjective collocates of the lexeme terdampak highlight the working class as the participants who faced the most consequences of the crisis. In contrast, the other noun and verb collocates show the government's compensation for the PSBB (Pembatasan Sosial Berskala Besar 'large scale social distancing) in the forms of the BST (Bantuan Sosial Tunai "cash social assistance") and staple food package. This article recommends further corpus-driven research to examine and compare how the affected sector of COVID-19 is portrayed in other countries, particularly how developing countries cope with the pandemic situation.

***Keywords**: corpus; COVID-19; affected; Indonesian; news*

Introduction

On December 31, 2019, SARS-CoV-2 was detected in Wuhan, China, and the virus was then named COVID-19 by WHO on February 11, 2020 (Nor & Zulcafli, 2020). In the first two months, the virus was spread to over 143 countries and territories, and on March 11, 2020, the WHO declared COVID-19 a pandemic (Labbé et al., 2022; Nor & Zulcafli, 2020). Research on the virus's impact has been conducted in various fields, including health, technology, education, the economy, and other social and scientific issues. It proves that the pandemic has forced many aspects of life to adjust their behavior and living to survive this global catastrophe. Besides the high number of deaths, poverty increased since many people had to give up their lives because of the layoffs. Many industries had to close their businesses because of the social restrictions to prevent the virus from spreading. Reports mention that the unemployment rate reached 7,07%, or 9,77 million people, in August 2020

(Ramadani, Sartika, & Herawaty, 2022).

The information about the number of COVID-19 cases, victims, regulations, and impacts of the pandemic has dominated the media, particularly the news media. Its role in disseminating information cannot be neglected as more than 2,9 billion people use social media, including news reports. Social media promptly updates information about the pandemic and popularizes the officials' knowledge, including government and medical professionals, to guide the public to understand correct information, particularly in developing countries (Tang, Zhang, & Li, 2021). Social media are also responsible for creating public sentiment, peace, and security during the pandemic (Kapoor & Dubey, 2020). However, news media are also blamed for spreading fake news that can create harmful effects, as found in South African media (Kabha, Kamel, Elbahi, Hafiz, & Dafri, 2020)

Positive and negative sentiments during the pandemic have been investigated by scholars from a linguistic point of view. For example, the UK and US tweets on the pandemic showed different sentiments. While the US tweets focused more on risk management and private funding, the UK tweets were more on health literacy, practitioners, and start-ups. The semantic relationship among the linguistic data in tweets revealed some dimensions of health technologies: technological priorities, professional skills, privacy issues, and privacy issues before and after the COVID pandemic (Reveilhac & Blanchard, 2022). Another research took a diachronic perspective that compared news frames in Iran from January to April 2020 (Jafarinejad, Rahimi, & Mashayekhi, 2021). Findings showed that the topics changed from social issues to health-related issues to the side effects of the virus. The Iran community was primarily affected by the pandemic, health care, entertainment, medical, and economic issues. There is also the general polarization of the discourse: mostly negative and shifts to neutral and positive since the government and officials' decision influences social morale.

Yu *et al.* compared news corpus in China and UK from January to July 2020 to figure out the media's attitudes before and after the official announcement of the pandemic by WHO. From more than 2 million tokens from *China Daily* and more than 26 million words from *The Guardian*, Yu et al. found some similarities and differences between the keywords used in both corpora. UK media show a more negative attitude than Chinese media since *China Daily* has consistently shared news for experience and cooperation. On the other hand, *The Guardian* focused more on medical prevention and physical restriction even though the news after the pandemic announcement became more positive (Yu, Lu, & Hu, 2021). In the Malaysian context, negative sentiment was also framed by *Star Online* from March 1 to March 31, 2020 (Nor & Zulcafli, 2020). By involving 1,018 news reports and 100 top collocates, the research found that the media reflected the fear, anxiety, and uncertainty the Malaysian feel because of the economic and social life affected by the pandemic. It also showed the government's effort to control the situation despite the threat to health and the economic situation (Nor & Zulcafli, 2020).

In the Pakistani context, Zahra & Abbas (2022) examined socio-cultural attitudes in news editorials from February to April 2020. Involving 138 editorials or 215,842 tokens, the authors used AntConc to identify the top twenty-five wordlist and compared the frequency based on different months. The themes that emerged from February to April are different. While February's news themes were dominated by racism and stigmatization, criticism of health facilities and the implication of globalization, March's news themes were loaded with increased fear and anxiety, debate on the Chinese virus, and world leadership. On the other hand, April's news themes were about the efforts to stabilize the vulnerable world, conspiracy theories, lack of facilities and people's

awareness, orientation from global to local, and the contribution of WHO. The findings also showed some metaphorical use to represent COVID-19, some racist phrases, and criticism of the government for health facilities and the implication of globalization. The research summed up that the Pakistani editorial corpus is loaded with a negative discourse that shapes the attitudes and behavior of the people (Zahra & Abbas, 2022). From a multimodal perspective, cartoon editorials in Canada also showed a negative attitude toward the pandemic by stigmatizing particular groups of the origins of COVID-19 transmission. The cartoons were polarized into two groups: virtuous people that are labeled as selfless and smart and those who respect the public health preventive measure, and 2) immoral, self-centered, silly, and stupid people who do not respect the prevention measures of the COVID-19 transmission (Labbé et al., 2022).

Triggered by the above studies, the present paper aims to fill the gap by focusing on the impacts of COVID-19 in the Indonesian news discourse 2020. With the same method, corpus linguistics, this paper is expected to enrich the recent discussion on COVID-19 in media, particularly in the Indonesian context. Specifically, this paper investigates a lexeme that directly refers to the affected elements of Indonesian society as the victims of the COVID-19 pandemic. Furthermore, this paper also involved more data than the studies above since it includes more than 4 million tokens as the target corpus. This paper also emphasizes the collocation and concordance analysis to prove the occurrence of the lexeme *terdampak* in the Indonesian news corpora. Therefore, this paper aims to answer the following questions:

1) What are the keywords of the Indonesian news corpus 2020?
2) What are the collocates of the lexeme *terdampak*?
3) How does the collocation tell us about the victims of the pandemic?

Methodology

This study presumes the Indonesian news corpus as the discourse that (re)produces the ideology of the Indonesian media. As a social element that aims to report information and facts on COVID-19, news discourse represents how Indonesian media shows the attitudes and behaviors of its people, including the government. The experience, memory, and belief form the cognition that bridges ideology and discourse. The individual news reporter is a part of a society with similar norms and values, so its position in a society cannot be neglected. The choice of semiotic modes, exceptionally written forms of language, represents not only individual but also the news companies, symbolic power supporting the news circulation, and the country in a much broader context. Therefore, this paper positions Indonesian news discourse from a socio-cognitive approach (van Dijk, 1998). The mental model articulated in the discourse reveals the ideology that bases the media on framing the COVID-19 pandemic.

Specifically, this paper applied corpus discourse analysis that selected data from the Indonesian news corpus 2020. This research took 300,000 sentences or 4,557,173 million tokens as the target corpus, accessed on https://corpora.uni-leipzig.de/en?corpusId=ind_news_2020. The corpus, in a .txt file, was then uploaded to AntConc 4.0 developed by Laurence Anthony from Waseda University (Anthony, 2021). This paper used keywords, collocations, and concordances tools provided by the corpus software. This paper used a reference corpus to find the keywords to compare the target corpus. The reference corpus was taken from a more general Indonesian mixed 2013 comprising 1 million sentences, compared to the Indonesian news corpus 2020, which only contains 300,000 sentences. Besides keywords, this paper also investigated the collocation of the lexeme *terdampak* as one of the keywords of Indonesian news 2020. The collocation set in this paper was

from 5L and 5R from the node, separated from their parts of speech. This step is urged to figure out how the media frame the affected participants and efforts to deal with the situation. Concordance lines would be helpful to display the context in which keywords appear that presents information as the keyword in context (KWIC).

Results and Discussion

There are three main aims of this paper. First, it aims to figure out the keywords of the Indonesian news corpus 2020. Using AntConc 4.0 to find out the keywords of Indonesian news 2020, this paper involved a more general corpus. I used Indonesian news mixed 2013 containing 1 million sentences or more than 15 million tokens as the reference corpus to be compared with the target corpus, Indonesian news 2020. The corpus tool provided keywords by a statistical value called Likelihood value, even though other statistical results using MI score and p value are also familiar in the corpus study. When the Likelihood value is >6,68, we can be 99% sure that the word characterizes the target corpus, so we can call it a keyword. Therefore, keywords can reveal the characteristics or themes of a corpus. Table 1 displays the first 50 keywords found in the target corpus and found both lexical and function words. However, to focus on the lexical analysis, this paper ignored the function words like *com*, *saat*, *di*, *enggak*, *tersebut*, and *hingga* as the keywords. Table 1 below displays the 50 keywords in the Indonesian news corpus 2020.

Table 1. List of 50 Keywords in Indonesian News Corpus 2020

Rank	Type	Freq	Keyness (Likelihood)	Rank	Type	Freq	Keyness (Likelihood)
1	covid	16048	48.298.845	26	isolasi	1190	2.819.488
2	corona	7202	21.229.001	27	terdampak	962	2.542.055
3	virus	7955	16.527.725	28	jumat	3562	2.540.408
4	pandemi	5589	16.502.899	29	instagram	1031	2.523.551
5	pasien	5359	7.728.028	30	rapid	992	2.517.152
6	masker	2651	6.712.745	31	jokowi	1729	2.496.210
7	protokol	2800	6.659.236	32	trump	848	2.413.883
8	positif	5175	5.927.793	33	enggak	1479	2.382.759
9	kasus	8482	5.484.068	34	gugus	1026	2.284.836
10	kesehatan	7377	5.163.591	35	kamis	3431	2.275.816
11	com	3754	5.034.988	36	senin	3533	2.261.565
12	psbb	1441	4.333.307	37	video	2821	2.247.492
13	penyebaran	2144	4.078.771	38	mengatakan	9156	2.229.665
14	wabah	1646	3.958.025	39	liga	2575	2.198.989
15	pelaku	4228	3.449.091	40	terkonfirmasi	753	2.197.279
16	saat	19366	3.342.506	41	rabu	3417	2.189.871
17	di	96846	3.167.444	42	tersebut	23124	2.129.856
18	swab	1065	3.162.816	43	penularan	1067	2.111.988
19	korban	5580	3.125.954	44	diketahui	3366	2.097.424

Rank	Type	Freq	Keyness (Likelihood)	Rank	Type	Freq	Keyness (Likelihood)
20	penanganan	2378	3.095.597	45	selasa	3324	2.095.682
21	vaksin	1477	3.060.220	46	sabtu	2755	2.095.587
22	meninggal	2922	2.994.303	47	ujar	4622	2.006.798
23	warga	7498	2.884.017	48	anies	775	1.993.842
24	motogp	1327	2.877.133	49	karantina	851	1.946.032
25	sembuh	1691	2.823.027	50	hingga	8928	1.929.830

The table above displays keywords from different parts of speech and topics. Topics are essential for discussion in Critical Discourse Analysis (CDA) as they are parts of the macro discourse structure. Macro discourse structure explains common ideological practices from a communicative event that affects readers' perspectives of an issue. The topics show the global meaning of discourse representation in constructing their mental model of an event (van Dijk, 1998). The topics presented in Indonesian news 2020 invite readers to process information from the texts and construct their subjective and social mental models. As a part of society, the news represents how Indonesian media shape the discourse of the COVID-19 pandemic. The linguistic choice, then, (re)produces ideology that underlies the social cognition of Indonesian people.

The statistical result displayed in table 1 reveals some important topics of Indonesian news 2020: COVID-19-related terms, health aid and care, case report, victims, and government. Those topics are constructed from different parts of speech. For example, nouns may fall for COVID-19-related terms and health aid and care, such as in the words *covid, corona, virus, pandemi, wabah, masker, vaksin, isolasi,* and *karantina.* Though nouns seem to dominate the keywords, some verbs also refer to government, particularly the actions of saying, as found in the keywords *ujar* and *mengatakan.* Some adjective keywords also appear in the data, referring to victims in the words *positif* and *meninggal.*

Below are examples of concordance lines of the lexeme *terdampak* in Indonesian news 2020. The keyword in context here is set in 80 characters on the left and right sides of the keyword.

Selatan (BS) membuktikan janjinya untuk mulai menyalurkan beras kepada warga	*terdampak*	korona atau Covid-19 mulai kemarin, Senin (4/5). 140902 KOTA MANNA, bengkuluekspress.
makanan fakir miskin. 26681 Beras itu diserahkan 10 kg beras setiap KK	*terdampak*	Covid-19, berikut uang tunai, zakat harta yang dihimpun warga,
dan Lembaga Pendidikan keagamaan menyampaikan bermacam usulan dan keluhan akibat	*terdampak*	COVID-19. 185792 Padahal, sudah sepekan terakhir Boyolali tak ada tambahan
kemanusiaan melalui pembagian bantuan kepada tenaga kesehatan dan masyarakat yang	*terdampak*	COVID-19. 153230 Melainkan, juga dari konsumen di negara-negara nonmuslim

upaya pemerintah dalam membantu mengurangi tingkat pengangguran di masyarakat yang	*terdampak*	pandemi Covid-19 secara ekonomi. 207969 PK telah menahan Juliari untuk 20
yang terkumpul juga disalurkan untuk berbagai kelompok pekerja informal yang	*terdampak*	pandemi Covid-19. 248484 Selain upaya tersebut, Pemerintah terus mendorong agar

In general, Indonesian news 2020 has a more positive tone than keywords found in UK media and shares some similarities to media in China (Yu et al., 2021). The keywords of Indonesian news also did not contain any keywords referring to fear, threat, anxiety, racism, stigmatization, and self-centered tones as found in Malaysia, Pakistan, and Canada (Labbé et al., 2022; Nor & Zulcafli, 2020; Zahra & Abbas, 2022). The Indonesian media focused more on daily case reports, including the number of cases, victims, and government efforts to prevent the virus from spreading. Starting March 2020, the Indonesian spokesperson of COVID-19, Achmad Jurianto, reported the daily case, so the lexemes like *senin, selasa,* and *rabu* also appear in the keyword list. The spokesperson reported the number of newly confirmed patients, the number of dead victims, and those who recovered. Moreover, he also warned the citizens to follow the regulations and stay at home without asserting words relating to fear. The language choice by Achmad Jurianto was also proven to improve the public's optimism and build solidarity amongst citizens as a moral force to face the pandemic (Sultan & Rapi, 2020). In a nutshell, Indonesian news media focused on building solidarity and trust in the government.

This paper examined how the keyword *terdampak* '*affected'* co-occurred with the surrounding lexemes. Morphologically, the lexeme *terdampak* is not yet found in the Indonesian dictionary. Instead, the dictionary lists the word *berdampak* meaning 'to cause particular effects'. The prefix *ter-* in the word *terdampak* is assumed to carry the meaning of *terkena dampak* '*affected'.* Therefore, this paper assumes that the collocation of the lexeme *terdampak* would reveal the social aspects related to the impacts of the pandemic. Statistically, there are 72 types of collocations realized in 3,225 tokens of the lexeme *terdampak*. This paper also ignored function words that collocate with the lexeme *terdampak*. In general, verb, noun, and adjective collocates of the lexeme *terdampak* emerged in the data. Below is the discussion of each collocate category.

The noun collocate of the lexeme *terdampak* dominates the Indonesian news 2020. Some topics of the collocates can be summarized as COVID-19-related terms, working class, and government compensation for the economic crisis. Table 2 below summarizes the noun collocate of the lexeme *terdampak.*

Table 2. Noun Collocate of the lexeme *terdampak*

Freq (Scaled)	Freq L	Freq R	Collocate	Freq (Scaled)	Freq L	Freq R	Collocate
160480	9	345	covid	32120	27	12	sosial
74980	170	9	warga	16460	0	26	wabah
55890	0	134	pandemi	34780	26	8	wilayah
96360	141	6	masyarakat	3040	10	1	bansos

Freq (Scaled)	Freq L	Freq R	Collocate	Freq (Scaled)	Freq L	Freq R	Collocate
30640	59	32	bantuan	1890	8	1	kepedulian
72020	3	84	corona	9630	15	1	umkm
31010	33	27	ekonomi	21120	21	1	usaha
79550	3	82	virus	3230	0	10	phk
17070	33	8	sektor	1050	7	0	debitur
12130	33	2	pekerja	9010	10	4	paket
9370	2	30	bencana	1180	6	0	restrukturisasi
24420	10	34	akibat	7910	9	3	pariwisata
4360	23	0	beban	31820	5	17	program
5940	19	4	sembako	1300	0	5	imbas

The noun collocates of the lexeme *terdampak* contain some COVID-19-related terms: *covid, corona,* and *virus*. Table 2 above displayed that the Indonesia news corpus did not contain any racist terms of COVID-19, different from the racist phrases like *china virus* as found in Pakistan and Canada (Labbé et al., 2022; Zahra & Abbas, 2022). The neutral terms of COVID-19 in the collocate of the lexeme *terdampak* shows a neutral tone of Indonesian media. The term "covid" that dominated the collocate co-occur with the term "COVID-19" and "virus". The interchangeable use of those terms did not confuse readers on the issue as the collocations appear in the same context or concordance lines. The virus is portrayed as the entity that causes the crisis. The word *akibat* '*effect'* found in the context explained the meaning of the prefix *ter-* in the word *terdampak.* Therefore, it can be said that the phrase *akibat terdampak covid* places *the virus* as the Agent and *society* as the Affected participants.

The second topic of noun collocates of the lexeme *terdampak* refers to the working class affected of the pandemic, as shown by the lexemes *pekerja* '*workers'*, *masyarakat* '*society'*, *phk 'layoff'*, and *umkm 'micro, small and medium enterprises (MSMEs)'*. Besides that, a more extensive economic sector, *pariwisata* '*tourism*', was also highlighted explicitly as the affected due to the pandemic. Below are examples of the concordance lines of the keyword *terdampak* collocated with the above lexemes.

Fauziyah bersama Mensos Juliari Batubara menyerahkan paket bantuan sosial kepada	*pekerja/*	buruh terdampak PHK maupun dirumahkan, di kantor Kemensos Jakarta,
yang paling terdampak dari pandemi, di mana sebagian besar merupakan	*pekerja*	informal atau terdampak PHK. 245735 Selain itu, kata dia, Perindo
tipikal yang speak-up? 129515 "Kenapa soal pesangon pekerja yang terdampak	*PHK*	pasti akan dibayar? 129516 Kenapa tagihan listrik makin melonjak? 129517 Kenapa,

terkadang datang dari proses pelaksanaannya. 265252 Slamet menjelaskan buruh yang terdampak	*PHK*	dan dirumahkan akibat pandemi wabah virus corona juga diusulkan
akan digunakan untuk membantu 12 juta pelaku Usaha Mikro Kecil Menengah (	*UMKM)*	yang terdampak pandemi Covid-19. 43719 Dana tersebut akan diperuntukkan penanganan
sambung Yusri. 154580 Memang pemerintah telah menganggarkan Rp123,46 triliun untuk bantuan	*UMKM*	terdampak Covid-19, tetapi realisasinya masih sangat kecil. 154581 Memang pemilih
dengan karang taruna. 37329 COM -- Kementerian Sosial mempercepat penyaluran bantuan sosial	*sembako*	bagi masyarakat terdampak COVID-19 di Jabodetabek. 37330 COM - Kemerdekaan Republik
ton meski tak ada penugasan khusus dari pemerintah untuk menyalurkan	*sembako*	untuk penerima manfaat yang terdampak Covid-19. 286007 Tri Purnama mengatakan,

The noun collocations referring to workers, society, and UMKM in the corpus reveals that Indonesian news 2020 paid more attention to the country's economic situation. Other than focusing on the patients or the number of victims, the media articulated the challenge of the economic crisis that the country might face. Similar to Malaysian news discourse, theme about economy was also emphasized on the news (Nor & Zulcafli, 2020). In Indonesian context, the economic theme was triggered by the number of unemployment and layoffs, particularly in the microeconomics, because of the *PSBB (Pembatasan Sosial Berskala Besar 'large scale social distancing)'* from April to June 2020 (Andriani, 2020). The government paid more attention to the workers, UMKM, and society who got layoffs and tried to compensate them by distributing cash social assistance *bansos (bantuan sosial tunai).*

The government compensation to the working class and UMKM is proven by the verb collocates of the lexeme *terdampak*. The verbs below were categorized into both passive and active verbs. The passive verbs refer to the affected sector, while the active verbs refer to the government actions, such as *meringankan* '*to ease*', *menyalurkan* '*to distribute*', and *menyasar* '*to target*'.

Table 3. Verb Collocate of the lexeme *terdampak*

Freq(Scaled)	Freq L	Freq R	Collocate
23550	52	52	membantu
2000	23	23	meringankan
980	5	5	dirumahkan
15990	7	7	diberikan
2330	4	4	disalurkan
2530	6	6	menyalurkan
1390	4	4	menyasar

The government is portrayed as the social actor responsible for the unemployment rate because of social restrictions. The Indonesian government also launched the Employment Card Programs for the workers who lost their jobs and for the MSMEs that closed their businesses. The government even

prepared 20 trillion rupiahs for around 5,6 million citizens (Andriani, 2020). Therefore, the collocates *membantu* '*to help', meringankan* '*to ease', menyalurkan* '*to distribute',* and *menyasar* '*to target'* emphasize the government compensation to those affected by the virus. The concordance lines below display how the lexeme *meringankan* is used in context.

mewabahnya virus Corona atau Covid-19 di Tanah Air, Pemerintah berusaha	*meringankan*	beban masyarakat yang terdampak melalui bantuan sosial. 38227 COM - Selama
sektor informal jauh berkurang, sehingga Dapur Umum diharapkan bisa membantu	*meringankan*	beban masyarakat terdampak Covid-19. 258207 "Seperti diketahui bersama, selain sulit
Saling bantu berikan pitulungan kepada sesama penting dijalankan agar bisa	*meringankan*	mereka yang terdampak secara ekonomi. 223542 Saling berbagi makanan dengan
Kepala Diskoperindag Kabupaten Gresik, mengatakan, diharapkan kegiatan gotong royong untuk	*meringankan*	beban masyarakat terdampak pandemi Covid-19 sangat penting. 254101 Sementara, Kepala

There were not only government agencies positioned as the Actor who helped and eased the affected participants (society, workers, and UMKM), but also society or citizens in general as shown by the word *sesama.* Therefore, it can be said that the Indonesian media also articulated solidarity. The government and citizens should go hand in hand to prevent the economic crisis. The balue of *gotong royong* also emerged in the context to highlight a positive attitude to face the pandemic. The verb *meringankan* is followed by noun phrase *beban masyarakat 'society's burden'* and *mereka yang terdampak secara ekonomi 'those affected economically'.*

Table 4. Adjective Collocate of the lexeme *terdampak*

Freq(Scaled)	Freq L	Freq R	Collocate
710	11	1	informal
3790	5	8	parah
2410	9	1	miskin
2930	6	2	rentan
2900	7	7	mikro

The other collocate of the lexeme *terdampak* is in the form of adjective. There were only five adjective found: *informal 'informal', parah 'severe', miskin 'poor', rentan 'vulnerable',* and *mikro 'micro'.* Different from the discourse of cartoon in Canada that stigmatized travelers in the beginning of the pandemic (Labbé et al., 2022), Indonesian news discourse paid more attention to the working class, poor and vulnerable people, as well as the people working in informal sectors or UMKM. The dominant discourse of economic impact of the pandemic also emerged in Twitter which included terms referring to people's financial situation (Ngo, 2022). It means that the language choice in Indonesian news is a reproduction of the other social media. This fact discloses facts that the COVID-19 pandemic has affected many aspects of life that influence social media in general as platforms to share their mental cognitions of the issue.

Conclusion

This paper limits its discussion on the keywords and collocation of the lexeme *terdampak* to find out the victims affected by the COVID-19 pandemic in Indonesian context. The first 50 keywords found in Indonesian news 2020 tell us about COVID-19-related terms, health aid and care, case report, victims, and government. The Indonesian news is loaded with the case reports including the victims and government action to prevent the economic crisis. As one of the keywords, the lexeme *terdampak* collocates with nouns, verbs, and adjectives to show the working class as the victims and government's compensation of the unemployment rate. This paper also proves that corpus linguistics is a beneficial tool to investigate the macro discourse structure of big data. From the themes emerged in the data, Indonesian news has a positive tone in portraying the pandemic. It shows the victims and government compensation of the financial situation and asks for people's solidarity. Since this paper only limits its data from 2020 news, further research can examine Indonesian news 2021 and 2022 to compare the similarities and differences of the themes related to COVID-19 pandemic in social context.

References

Andriani, H. (2020). Effectiveness of Large-Scale Social Restrictions (PSBB) toward the New Normal Era during COVID-19 Outbreak: A Mini Policy Review. *Journal of Indonesian Health Policy and Administration*, *5*(2). doi: 10.7454/ihpa.v5i2.4001

Anthony, L. (2021). *AntConc (Version 4.0.0)*. Tokyo: Waseda University. Retrieved from http://www.antlab.sci.waseda.ac.jp/

Jafarinejad, F., Rahimi, M., & Mashayekhi, H. (2021). Tracking and analysis of discourse dynamics and polarity during the early Corona pandemic in Iran. *Journal of Biomedical Informatics*, *121*, 103862. doi: 10.1016/j.jbi.2021.103862

Kabha, R., Kamel, A. M., Elbahi, M., Hafiz, A. M. D., & Dafri, W. (2020). Impact of Fake News and Myths Related to COVID-19. *Journal of Content, Community & Communication*, *12*(6), 270–279.

Kapoor, S., & Dubey, M. (2020). The Impact and Role of Media for Peace and Human Security during COVID-19 Crises. *Journal of Content, Community & Communication*, *12*(6), 210–218.

Labbé, F., Pelletier, C., Bettinger, J. A., Curran, J., Graham, J. E., Greyson, D., … Dubé, È. (2022). Stigma and blame related to COVID-19 pandemic: A case-study of editorial cartoons in Canada. *Social Science & Medicine*, *296*, 114803. doi: 10.1016/j.socscimed.2022.114803

Nor, N. F. M., & Zulcafli, A. S. (2020). Corpus Driven Analysis of News Reports about Covid-19 in a Malaysian Online Newspaper. *GEMA Online® Journal of Language Studies*, *20*(3). doi: 10.17576/gema-2020-2003-12

Ngo, V. M. (2022). Dataset of discourses about COVID-19 and financial markets from Twitter. *Data in Brief*, *43*, 108428. doi: 10.1016/j.dib.2022.108428

Ramadani, A. N., Sartika, D., & Herawaty, H. (2022). Increase in Unemployment Rates During the Covid-19 Pandemic. *Jurnal Ilmiah Ilmu Administrasi Dan Manajemen*, *15*(3), 111–120.

Reveilhac, M., & Blanchard, A. (2022). The framing of health technologies on social media by major actors: Prominent health issues and COVID-related public concerns. *International Journal of Information Management Data Insights*, *2*(1), 100068. doi: 10.1016/j.jjimei.2022.100068

Sultan, S., & Rapi, M. (2020). Positive Discourse Analysis of the Indonesian Government

Spokesperson's Discursive Strategies during the Covid-19 Pandemic. *GEMA Online® Journal of Language Studies*, *20*(4), 251–272. doi: 10.17576/gema-2020-2004-14

Tang, Q., Zhang, K., & Li, Y. (2021). The Important Role of Social Media During the COVID-19 Epidemic. *Disaster Medicine and Public Health Preparedness*, *15*(4), e3–e4. doi: 10.1017/dmp.2020.330

van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. London ; Thousand Oaks, Calif: Sage Publications.

Yu, H., Lu, H., & Hu, J. (2021). A Corpus-Based Critical Discourse Analysis of News Reports on the COVID-19 Pandemic in China and the UK. *International Journal of English Linguistics*, *11*(2), p36. doi: 10.5539/ijel.v11n2p36

Zahra, T., & Abbas, A. (2022). Corpus-Driven Analysis of Pakistani Newspaper Editorials on COVID-19 Discourse. *GEMA Online® Journal of Language Studies*, *22*(1). doi: 10.17576/gema-2022-2201-02