

Corpus Driven Analysis of News Reports about Covid-19 in a Malaysian Online Newspaper

Nor Fariza Mohd Nor^a

fariza@ukm.edu.my

Center for Research in Language and Linguistics
Faculty of Social Sciences and Humanities
Universiti Kebangsaan Malaysia

Adlyn Syahirah Zulcafli

asyahirahzras@gmail.com

Center for Research in Language and Linguistics
Faculty of Social Sciences and Humanities
Universiti Kebangsaan Malaysia

ABSTRACT

Covid-19 was declared a global pandemic by the World Health Organization on 11th March, 2020. Scientists and researchers immediately took interests to find a vaccine. While researchers in various discipline, particularly in medicine, began to study about the virus to help us understand the situation from various perspectives, there is limited research on linguistics analysis to date about Covid-19. Therefore, the present study aims to contribute to research in linguistics about Covid-19, by employing corpus-driven approach for data analysis. Online newspaper reports about Covid-19 were downloaded from the Star online from March 1st to March 31st, 2020. A total of 1018 news reports were processed and analysed, with 140 themes being identified. 100 top collocates with the MI score set at a minimum of 5, using AntConc software revealed that verb and noun collocates co-occurred frequently with the node word Covid-19. Generally, the collocates reflect fear, anxiety and uncertainty that majority of Malaysians feel. It also shows how the virus is taking a toll on Malaysia and Malaysians with regards to economy and social life. On the other hand, some of the collocates portrayed the government to be in total control of the situation, despite the threat to health and economic situations in the country. Suggestions for future research about Covid-19 include comparison of several newspaper reports from English newspapers in Malaysia, analysis of news report during extended movement control order (ECMO) in Malaysia or discourse analysis of crisis communication by political leaders in ASEAN.

Keywords: Covid-19; corpus-driven; collocates; news report; Movement Control Order

INTRODUCTION

The world is regularly experiencing episodes of infectious diseases outbreaks in different parts of the world, such as Ebola, SARS, HINI influenza and other infectious diseases. The current 2019-ncCOV infodemic (hereinafter, Covid-19 infordemic) is taking a toll on every aspects of life and is spreading across the world at an unprecedented scale. The outbreak was reported in Wuhan, a major city of Hubei province in China. In a very short time, the disease spread very quickly and reached almost every part of China. Due to mass migration at that time for the Chinese New Year festival, the virus began to spread outside China.

Whenever disasters struck, researchers in various disciplines contribute to help us understand the situation from various perspectives. For instance, some researchers have studied information behavior in natural disasters (Lin et al. 2016; Pang & Anwar, in press) and crisis

^a Main & correspondent author

informatics researchers on natural disasters (Liu, Palen, & Giaccardi, 2012; Palen, Hiltz, & Liu, 2007; Palen & Hughes, 2018, in Soden & Pallen, 2018; Toppenberg-Pejcic et al., 2019). The outbreak of infectious diseases is no exception.

In fact, past studies have been carried out on infectious diseases, namely the three main infectious disease; Ebola (e.g. Trčková, 2015; Wong et al., 2015, Balteiro, 2016; Feijen, 2016;; Tshiswaka, Whembolua & Conserve, 2017, Mondragon, 2017; Tran & Lee, 2016) SARS (severe acute respiratory syndrome) (e.g. Chung, 2011; Chiang & Duang, 2007; Wallis & Nerlich, 2005; Joye, 2010) and MERS (Middle East respiratory syndrome) (Do et al. 2016; Choi, 2017). Linguistics analysis about these infectious diseases revealed how the disease is conceptualized metaphorically, public's reactions and concerns and government's reactions and responses to these diseases.

When coronavirus was first detected, it was identified as “severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)” due to its genetic similarity to SARS (the outbreak which was in 2003). On February 11th. 2020, the World Health Organization (WHO) (2020a) named it “COVID-19.” In the first two months, the virus extended from China to over 143 countries and territories (WHO, 2020b, Coronavirus disease 2019). Since then, its spread has been exponential. WHO was initially reluctant to declare an international health emergency “because of [a] limited number of cases abroad, and also considering the efforts which are presently made by China and the Chinese authorities in order to try to contain the disease” (WHO, 2020c, International Health Regulations (IHR) Emergency Committee). However, considering the global threat of the 2019-nCoV, the World Health Organization (WHO) declared it as a ‘public health emergency of international concern’ on January 30th, 2020. WHO later declared Covid-19 as a global pandemic on 11th March, 2020. Currently, more than 21 million cases have been reported worldwide, with more than 700,000 deaths (John Hopkins Coronavirus Resource Center, 16th August, 2020), in more than 188 countries.

Researchers all over the world immediately took interest in this virus. Readings about research done in Elsevier database (<https://www.elsevier.com/connect/coronavirus-information-center>), Clarivate Analytics database (<https://clarivate.com/coronavirus-resources/>) and Covid-19 bibliometrics database (<https://www.covid19bibliometrics.org/>) showed that majority of research on covid-19 focus on medical research. Therefore, the present study is deemed timely and pertinent to contribute to an understanding about Covid-19 from a linguistics perspective. This study also intends to contribute to past studies about infectious diseases or pandemic situations in linguistics. . This study aims at exploring how language is used to present news about covid-19 in a local online newspaper in Malaysia, using corpus driven analysis approach. The software employed to process the data is Antconc 3.5.8. The present study is guided by two research objectives; (1) to determine the themes which can be discerned from the headlines on news about Covid-19 and (2) to identify words which collocate with the node word Covid-19. The corpus driven analysis will show how Malaysians and the Malaysian government responded to this pandemic.

LITERATURE REVIEW

This section will review past studies based on linguistics analysis about Ebola, SARS and MERS to illustrate the types of linguistics analysis which have been done and how these infectious diseases are conceptualized.

PAST STUDIES ON EBOLA VIRUS

The Ebola virus was first discovered in 1976. On 8th August 2014, the outbreak was declared public health emergency of international concern by the WHO. Balteiro (2016) examined specific instances of metaphors which are used to describe Ebola virus in *Scientific American*; a magazine that focused on scientific discovery and technological innovation). Ten “hard-news” articles on Ebola (over 15,000 words altogether), published between November 2014 and February 2016 from the online version of *Scientific American* were analysed. Ten frames were identified, namely (1) the disease itself, (2) patients or victims, (3) the body and its parts, (4) symptoms and consequences, (5) recovery or healing, (6) health workers (namely, doctors, researchers and experts, (7) research, knowledge, information and findings, (8) treatment, vaccination and drugs, (9) other preventive measures and (10) authorities. In addition, metaphorical conceptualization of the virus was also pertinent. The Ebola virus was conceptualized as war, a hidden object which moves ferociously movement like a road or path inside the body and as light (for research on Ebola). Where lexical metaphors are concerned, some of the conceptualizations are *war*, *fire*, *burden*, *plague*, *storm* and *fire*.

Trčková (2015), on the other hand, not only used cognitive theory of metaphor, but employed critical discourse analysis to analyse the metaphorical representation of Ebola in hard copies of the New York Times and the New York Daily News, with the time span from 25th March to 31st December, 2014. The most prevalent metaphoric theme in both newspapers portrayed Ebola as *war* and as *aggressive enemy*. Two other metaphoric conceptualizations include Ebola as an *animate/human being* and a *natural catastrophe*. Tshiswaka, Whembolua and Conserve (2017) took a different stance by using thematic analysis to evaluate the construction of the Ebola outbreak narrative in a corpus that consisted of online and offline news articles from the mainstream Congolese print media, covering the time span from 26th August to 31st October 2014. A database of keywords associated with the outbreak, e.g., *Ebola*, *virus*, *Democratic Republic of the Congo* was created. Three themes emerged from a dominant metaphor of *warfare* as used in the selected articles, namely that the general population was depicted as a potential victim, while public institutions were portrayed as heroes or warriors fighting the Ebola Virus Disease (EVD).

Mondragon et al. (2017) examined the collective image of the 2014 - 2015 Ebola outbreak to understand how people incorporate this epidemic in their everyday thinking from 700 newspapers articles, in Spain. Ebola was represented as inherently African and depicted as a global threat that created fear. Emotions such as anger and blame were pertinent. The people blamed the political authorities and the mass media for the failure to manage this crisis.

Since social media platforms have the potential to match this rapid transmission time by disseminating information quickly to a large audience, tweets were also used as data in research about Ebola. Feijen (2016) employed corpus linguistics approach to examine tweets by the Dutch about Ebola, with the aim of examining how Dutch people think about Ebola virus, versus the message that is spread about Ebola in the Dutch media. The analysis from the top 100 collocates found that Ebola was frequently mentioned with other diseases, such as gonorrhoea, leprosy, herpes, bubonic plague and MERS. The tweets revealed that the Dutch form their own opinion, instead of trusting the newspapers where subjective topics are concerned. Another study about the Ebola virus using tweeter was carried out by Wong et al. (2015), which analysed tweets send out by local health departments (LHD) in the U.S. Characteristics of local health departments (LHDs) tweets about Ebola, the type of information and when LHDs communicated to the public Ebola-related messages were examined. 287 LHDs were involved in the study and data was collected from September 3rd to November 2nd, 2014. A descriptive and bivariate inferential statistic to explore LHD characteristics, services, and Ebola-related tweeting were employed. From the 1648 Ebola tweets, 78.6% of

all tweets were information giving, 22.5% of all tweets were on preparedness, 20.8% of all tweets were news updates, and 10.3% of all tweets were event promotion tweets. Since the tweets were from LHDs, many tweets were intended to reduce fear, either indirectly or directly, by describing the minimal risk in contracting Ebola. Wong et al. (2015) concluded that the tweets that indirectly or directly address fear about Ebola are important to counteract inaccurate information about Ebola, which was extensively spread in social media.

Tran and Lee's study (2016) to understand the public's reactions and concerns during the Ebola crisis involved collecting a comprehensive data-drive analysis of the geotagged tweets in 90 languages. The study aims to understand language distribution and geographical distribution, using topic modelling analysis to develop topic models and establish the distribution of the tweets. Only tweets in English were analysed. Different concerns were raised about Ebola, such as *fear* and *death*. These concerns depended on which part of the world people are from. They also found that based on geolocation analysis of the tweet, information about some cases is locally propagated.

PAST STUDIES ON SARS

China's Southern province of Guangdong was affected by a 'strange contagious disease' (Associated Press, 2003) that infected and killed several people within the region, in November 2002. The contagious disease soon spread to Hong Kong, Taiwan, Singapore, Vietnam and Canada. The World Health Organization (WHO) responded to the global spread of SARS and declared the disease a 'worldwide health threat' (Associated Press, 2003) on 15 March 2003 and started to issue travel advisories. Scholars in linguistics and language took interest in how the disease was communicated to the public, the public's opinion and thoughts about it. Chung (2011) examined how SARS was reported in an English newspaper in Malaysia (New Straits Times) and in the United Kingdom (The Times) over several consecutive months, with the aim of comparing how the two newspapers from the two nations present news about SARS to its respective readers. A corpus-based study was carried out to compare the attitude of the two newspapers. The findings indicated that the two nations presented different messages to the public. The Malaysian press played the role of an encourager and a patriot in an effort to unite the nation by using language such as, 'must fight SARS', 'let's fight SARS' and reminded the public not to be discouraged by SARS. The Times, however, acted as a reporter and occasionally employed the "Self" metaphors where "Others" were blamed for the outbreak, such as the phrase 'the mishandling of SARS in China'. In another study Ching and Duann (2007), the naming strategies and conceptual metaphors of SARS were investigated in The Liberty Times (the US), The United Daily News (Taiwan) and People's Daily News (China) newspapers. The disease was constructed as *war*, *disaster* and *natural disaster* metaphors. Chang and Duann (2017) concluded that political agendas and underlying ideologies of the newspapers permeate their use of metaphors. Wallis and Nerlich (2005) study were also on the representation of SARS as a *disease* metaphor. Akin to Ching and Duann (2007), Wallis and Nerlich examined how the UK press framed the SARS epidemic in five major national newspapers, which are The Sun, The Daily Mail, The Daily Telegraph, The Mirror and The Guardian, from March 2003 until the end of July 2003. The findings revealed that SARS was conceptualized as a KILLER metaphor, such as 'killer virus', 'killer plague', or 'deadly bug', instead of the WAR metaphor. Struggle metaphor was also prevalent and was used to describe the human and economic impact of SARS. There were also occasional mentioned of other metaphors and analogical frames.

Using a case-based method, Joye (2010) investigated how two Belgian television stations covered reports on outbreak of SARS. Fairclough's (1992) CDA model was employed to examine the relation between discourse and social, and cultural phenomena. Four discursive

moments were identified, which were as follows - period of media silence, SARS as emergency news, three days of ecstatic news and SARS as adventure news. The findings demonstrated that the news coverage of SARS not only construct, but maintains the socio-cultural difference between 'us' and 'them'. The discourses of global power structures were prevalent in the news report. Further, a divide of the world zones of poverty and prosperity, danger and safety were imminent in the news reports.

PAST STUDIES ON MERS-COV

The Middle East Respiratory Syndrome (MERS) is a viral respiratory illness, which was first reported in Saudi Arabia in 2012. Those infected developed severe respiratory illness, including fever, cough, and shortness of breath. Choi et al. (2017), and Do et al. (2016) gathered data from the social media to examine the public's perceptions about the virus in South Korea. Choi et al. (2017) examined how social media could influenced people's risk perceptions during the MERS outbreak in South Korea and the role that individuals' information processing mode and self-efficacy have on the association between social media and risk perception. An online survey was conducted from a national online panel sample of South Korean adults. Negative words, such as anxiety, fear, uncertainty, risk and suspicion were common because people tend to focus more on negative information. The study sheds light on how social media can shaped people's risk perceptions during the MERS outbreak period. Do et al. (2016) were more specific with their analyses of tweets during the MERS outbreak in Korea. They employed emotion classification scheme, that is particularly attuned to Korean and Twitter domain, in order to gain insight into the emotion dynamics during the crisis situation, including trend of the emotional intensities and responses to different topics. The results showed that anger was predominant among people's emotions.

The reviewed past studies which employed corpus linguistics and critical discourse analysis highlighted how the mass media and social media presented Ebola, SARS and MERS-COV as something very unpleasant, hence, these infectious diseases were inherently associated with fear, catastrophe and death. The impact of the infectious diseases on social and economic situations in the affected countries are also prevalent. The three viruses were metaphorically conceptualized as *war*, *disaster*, *natural disaster*, *storm*, *fire*, *enemy* and killer metaphor which heightened feelings of fear among the public. In addition, the Ebola virus is also associated with other infectious diseases such as gonorrhoea, leprosy, herpes, bubonic plague and MERS. The negative presentation of the infectious diseases can subsequently lead to panic and heightened fear among the public, which may affect the government's efforts to respond appropriately during this pandemic crisis. Nevertheless, presentations by the mass media and postings in social media by the public did not hinder the government from giving assurances to the public. Studies by Wong et al. (2015) reported that tweets by the LHD in the U.S were intended to reduce fear among the public by describing the minimal risk in contracting Ebola. The aim of these LHDs is to counteract inaccurate information about Ebola, which was extensively spread in social media. Chung's (2011) contrastive study on reports about SARS in the New Straits Times (Malaysia) and The Times (United Kingdom) indicated that the Malaysian press indirectly urged the people to unite in the fight against SARS. Despite limitations in the reviewed studies such as scope of the study and employment of only one method or approach, the reviewed studies associate infectious disease with fear, death and crisis.

STUDIES ABOUT COVID-19

When coronavirus began to spread rapidly in China and exponentially spread outside China, this infectious disease motivated the scientific community to work together in order to gather, organize, process and distribute data on the novel biomedical hazard. A review of the literature on research about Covid-19 revealed that majority of research on Covid-19 focus on medical research. To date, research using linguistics analysis approach about Covid-19 are beginning to emerge. This researches used approaches such as Natural Language Processing (NLP), machine learning, sentiment analysis and text mining, which employed software for analysis. This paper will present a few research on linguistics analysis about Covid-19 to highlight part of the results and insight about Covid-19. Katermina and Yachenko (2020) examines the linguistic phenomenon of Covid-19 in English mass media texts to determine if mass media creates, reproduces and transmits axiological values about Covid-19. Data comprised of The Economist issues (March-May 2020) obtained from the Economist website. The study is based on words that collocate with “covid-19”, ‘Coronavirus’, “virus” and ‘disease’. The results showed that Covid-19 is widely represented in the mass media in various ways, such as conceptualizing the virus as war through the use of the mentioned lexemes and warlike realities in the same context: “a frontline”, “a campaign”, “a weapon”, “battlefield conditions”, “wartime-like emergency measures” and negative connotations like “the blow from the virus on the public services”, “the costs of covid” and “the scars of covid-19”. The researchers concluded that that Covid-19 pandemic has affected culture and language, to the extent that it can be considered as considerable social calamity. A contrastive study of American Washington Post newspaper and the Chinese People’s Daily newspaper on COVID 19 by Mohamad Awad (2020) demonstrated that media bias is practiced through gatekeeping, coverage and statement made by the authors. CDA and CL approaches were employed in the analysis, which showed that Covid-19 is labelled as the Chinese virus, Wuhan Virus, and Kong Flu. This labelling was foregrounded in the People Daily’s article to create a negative hollow among Asian Americans, but it was backgrounded in the Washington Post’s article. It was instead replaced by the label “CCP”, with the intention of creating a positive impression among Asian and Chinese Americans. This contrastive study also reveals that the topics that are foregrounded by the American newspaper are backgrounded or filtered by the Chinese newspaper, and the topics that are backgrounded by the American newspaper are foregrounded by the Chinese newspaper. Mazlin Azizan, Hanita Hanim Ismail and Shatha Naiyf Qaiwer (2020) conducted a study using positive discourse analysis and critical discourse analysis approaches on 15 Facebook postings about Covid-19. Data was coded and categorised using a thematic analysis, which revealed construction of positive discourse. The postings highlighted the extensive use of collective pronouns such as “we” and “us” which suggest solidarity and empowerment among Malaysians in dealing with Covid-19. Four positive themes were found with regards to the postings in the Facebook, which are religion, patriotism, call for heroism and public awareness. The researchers concluded that the individual efforts have subsequently acted as a strong buffer against negative discourse and the emergence of subtle power as an effort to make a difference. Reflected in the 15 Facebook postings analysed, the discourse styles and linguistic strategies adopted have demonstrated the emergence of subtle power as an effort to make a difference.

The corpus linguistics approach employed in Katermina and Yachenko’s study (2020) provided insights into how the virus has affected culture and language. Results on word that collocate with Covid-19 and coronavirus resonates the results of past studies about Ebola, SARs and MERS-COV, whereby Covid-19 is also conceptualized as *war* and has negative connotations. The contrastive analysis of newspapers (Mohamad Awad, 2020) revealed how the writers in both newspapers foregrounded and backgrounded ideologies of the newspapers

and the country through careful selection of topics and the tone of the reports. Mohamad Awad's study is akin to Chang and Duann (2017) who reported that newspapers have political agendas and underlying ideologies in presenting news. Mazlin Azizan, Hanita Hanim Ismail and Shatha Naiyf Qaiwer (2020)'s study which employed positive discourse and critical discourse analysis approaches showed how Malaysians demonstrated solidarity and unity as a collective society in the battle against Covid-19. The results concurs with Chung's (2011), who reported that the Malaysian press acted as a patriot and encourager in the effort to unite the nation during the SARS outbreak.

METHODOLOGY

This study employed corpus driven approach using Antconc 3.5.8 (Windows) 2019 tool to analyse the corpus of the online news report on Covid-19. Discourse analysis of excerpts from the concordance list and news report was conducted to obtain insight and understanding about the discourse surrounding Covid-19 news. qualitative approaches. In doing the quantitative analysis, Antconc 3.5.8 (Windows) 2019 tool was used to analyse the corpus of the online news reports on Covid-19.

COVID-19 NEWS CORPUS

Data for this study was collected from The Star Online news portal (<https://www.thestar.com.my>). By using the keyword *Covid-19*, a total of 1030 news articles on Covid-19 issues in Malaysia, dated March 1st until March 31st, 2020 were extracted. An average of 40 short news report about Covid-19, were published everyday. It is a fact that one cannot possibly collect all texts, hence, for this reason, the corpora chosen is based on sampling. In addition, the researcher needs to ensure that "the corpora is maximally representative of that variety, in order to provide a picture as accurate as possible and avoid being skewed" (Agnostou and Weir, 2006: 1). The researchers, therefore, decided to collect data based on news report in March only, because it was during this time that cases of Covid-19 started to increase in Malaysia, which prompted the government to impose the first phase of 'lockdown' in the country, which is referred to as Movement Control Order (MCO). In addition, this particular time was the first wave of Covid-19 outbreak in Malaysia. This time is also considered significant because Covid-19 was spreading exponentially in countries around the world. In addition, Malaysia was experiencing political turmoil in March. A new government was formed when Tan Sri Muhyiddin Yassin was sworn in as the new Prime Minister of Malaysia on 8th March, 2020. Feelings of fear and uncertainty about Covid-19 and the economic situation was expressed by Malaysians, especially in social media. The new government, thus, is tasked with a huge responsibility to handle this pandemic as best as it can. News reports about Covid-19 in March is expected to be interwoven with news about Covid-19 in relation to politics and economy.

RESULTS

IDENTIFICATION OF THEMES

The headline was analysed manually, using Clarke and Braun (2013) six phases of thematic analysis. Clarke and Braun (2013) assert that thematic analysis is suitable for a wide range of research interest and theoretical perspectives because it works with (1) a wide range of research questions, (b) can be used to analyse different types of data, from secondary sources to transcripts of focus groups or interviews, (c) works with large or small data-sets and (d) can be

applied to produce data-driven or theory driven analysis. The news reports were manually ‘copied and pasted’ onto word document (.doc). An average of 40 short news report about Covid-19, were published every day in the newspaper, which made up a total of 1280 news report. The data filtration process was conducted by eliminating any overlapping and irrelevant article, leaving only 1018 to be used as the data for this study. The news reports were manually ‘copied and pasted’ onto word document (.doc). An example of how the themes were identified based on the news headline using Clarke and Braun (2013) thematic analysis is shown in table 1.

The headlines of the three news report are shown below (refer Appendix for the three news report).

- i. Trainee nurse among four new Covid-19 cases in Malaysia, bringing total to 29 (1st March)
- ii. First healthcare worker infected (2nd March)
- iii. Politicians test negative for Covid-19, says Health DG (2nd March)

TABLE 1. Clarke and Braun (2013) six phases of thematic analysis

No	Phases	Description	Example of the thematic analysis process
1	Familiarisation with data	The researcher must be familiar with the data which involved reading and re-reading the data (and listening to audio-recorded data at least once, if relevant) and noting any initial analytic observations.	Each news report has detailed information about the individual who are affected by Covid-19.
2	Coding	This process captures both a semantic and conceptual reading of the data. The researcher codes every data item and ends this phase by collating all their codes and relevant data extracts.	Identified keyword: new covid cases, healthcare worker infected, politicians test negative
3	Searching for themes	Coding the identified codes to identify similarity in the data.	Covid cases Covid report positive negative
4	Reviewing themes	Involves checking that the themes ‘work’ in relation to both the coded extracts and the full data-set. It may be necessary to collapse two themes together or to split a theme into two or more themes, or to discard the candidate themes altogether and begin again the process of theme development.	Covid-19 cases and report are the same, hence, the researcher decided to label news reports such as this as <i>Covid report</i> . News report about Covid-19 under this theme will be about positive and <i>negative cases</i> , and <i>death</i> .
5	Defining and naming themes	The researcher must conduct and write a detailed analysis of each theme. Questions which are asked at this point include the ‘what story does this theme tell?’ and ‘how does this theme fit into the overall story about the data?’	The final theme is identified as <i>Covid report</i> .
6	Writing up	Writing-up involves weaving together the analytic narrative and (vivid) data extracts to tell the reader a coherent and persuasive story about the data, and contextualising it in relation to existing literature	Covid report (positive, negative, death)

CORPUS LINGUISTICS ANALYSIS

As mentioned earlier, the corpus analysis utilized the Antconc 3.5.8 (Windows) 2019 tool, which is a computer software that allows researcher to run concordancing and text analysis. In order to carry out the corpus analysis, the articles were converted into plain text format (.txt)

before they were uploaded into the software. The corpus of this study has a total of 13122 word types, comprising 317,473 word tokens.

COLLOCATION ANALYSIS

The definition of collocation adopted in the present study is the above-chance frequent co-occurrence of two words within a pre-determined span, usually five words on either side of the word under investigation (the *node*) (see Sinclair, 1991). The statistical calculation of collocation is based on three measures: the frequency of the node, the frequency of the collocates, and the frequency of the collocation.

The processing of collocation analysis involved gathering the text files and uploading them into the Antconc 3.5.8 software. Then, the node word *Covid-19* was keyed into the system for collocation analysis. This statistical analysis was done using the Mutual Information (MI Score) to measure the strength of the collocation. The MI Score gives researchers information on the collocate likelihood with the node word in this study (Nor Fariza Mohd Nor et al., 2019). However, it could not determine the collocates statistical significance; hence, the researchers employed the Log Likelihood (LL) measure. The two statistical measures, LL together with the MI Score were used to obtain the statistical significance. Therefore, by using MI Score and LL in the analysis, the top 100 of Covid-19 collocates were identified. According to Xiao and McEnery (2006), the MI Score of 3 and above indicates the collocates of the node words. For the present study, the researchers set the minimum MI score to 5 with the minimum frequency greater than 4, within the 5-5 window span to be considered as the collocates of the node word *Covid-19*. The top 100 collocates were chosen based on this parameter.

PART OF SPEECH (POS) TAGGING

POS tagging was done to ensure that the collocates were tagged according to the part of speech (verb, noun and adjective). This was done using English POS tagger (<https://parts-of-speech.info>). Each POS tag was reviewed in order to check the accuracy of the tagging (verb and noun). Any inaccurate tags were corrected manually with reference of English words dictionary (Cambridge Online Dictionary (<https://dictionary.cambridge.org/>

RESULTS

This section begins with the findings of the themes discerned from the headline of the news report. Each headline was scrutinised using Clarke and Braun (2013) six phases of thematic analysis, which resulted in 140 themes from a total of 1080 news reports. The ten top frequent themes are depicted in table 1. The most frequent theme, at a frequency of 66, refers to *statement about Covid-19* situations, such as suspected covid cases, health information, preventive measures and Covid-19 situations at designated hospitals. These statements were made by the government, the Health Ministry, ministers and other individual with authority. The other frequent themes are statements about Covid-19 which has political nuances, hence, identified as *politics*, that occurred 49 times, *Covid-19 and economy* (48), *reports on Covid-19, specifically about positive cases and death* (40), *Movement Control Order* (39), *Seri Petaling gathering* (29), *face masks* (23), *cancelled events* (21), *covid disinfection* (17) and *statement about Covid-19 and education* (15). The pattern that emerged can somehow be predicted based on the headline and news report, for instance, news report about the Seri Petaling gathering will have recurring words and phrases such as Seri Petaling, religion, religious gathering, *tabligh*, 16,000, spread, infected.

TABLE 2. Themes and frequency of headlines

No	Themes	Frequency
1	Covid Statement	66
2	Politics	49
3	Economy	48
4	Covid report (positive cases, death)	40
5	MCO	35
6	Seri Petaling gathering	29
7	Face masks	23
8	Cancelled events	21
9	Covid disinfection	17
10	Covid and education	15

COLLOCATION ANALYSIS

The collocation analysis focus on verb and noun. To reiterate, the node word is ‘Covid-19’. To reiterate, the top 100 collocates were generated using LL and MI score. The collocates generated from the collocate list were reclassified into two lexical groups which are verb and noun. Figure 3 shows the frequency of each lexical group in the news reports, from 1st May until 31st May, 2020. The collocation rank is done based on MI and Log-likelihood score. Information in Figure 3 illustrated that verb collocate recorded the highest occurrence (50), followed by noun (36).

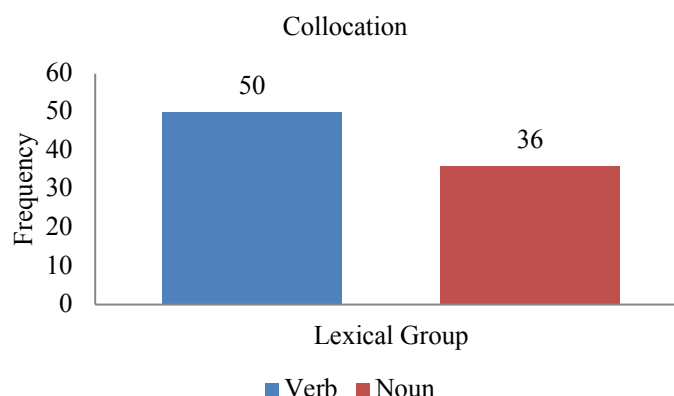


FIGURE 3. Graph chart for frequency of verb and noun collocates

COLLOCATION ANALYSIS OF VERB COLLOCATE

To reiterate, the minimum MI score of 5 with the minimum frequency greater than 4, within the 5-5 window span was considered as the collocates of the node word *Covid-19*, in which the top 100 collocates were chosen based on this parameter. Verb collocates occurred frequently with the node Covid-19, which is at 50 (see figure 3). The top 10 verb collocate is shown in table 3. The top three most frequent verb collocates will be analysed as examples.

TABLE 3. Frequency count for the top ten verb collocates

Freq	Freq L	Freq R	MI	Collocate
226	212	14	6.01919	spread
85	85	0	6.38126	curb
47	42	5	6.32532	fight
43	43	0	6.25754	suspected
38	36	2	5.27775	prevent
31	8	23	5.59283	detected
23	13	10	6.0918	died

19	14	5	5.0792	recovered
17	12	5	6.3756	battling
16	16	0	5.3201	combat

The observation made is that the majority of the verb collocates refer to actions taken by the government with regards to health preventive measures, government's health procedures and the state of Covid-19 in Malaysia. The analysis of the top three verb collocates, which are *spread* (226) *curb* (85) and *fight* (47) illustrated this observation.

ANALYSIS OF VERB COLLOCATE *SPREAD*

A few examples of the verb collocate *spread* from the concordance list and the news reports (NR, henceforth) have been extracted for analysis purpose.

Examples of verb collocate *spread* from concordance list:

- 1 ...personal hygiene as individuals and as a country. This is to slow down the **spread** of **Covid-19** within our communities if or when it were to get into our communities
- 2 banned from rom entering Sarawak after the state government decided to extend its entry ban to foreigners to prevent the **spread** of **Covid-19** in the state.
- 3 the Armed Forces will be deployed to help the police to enforce the movement control order to curb the **spread** of the **Covid-19** pandemic
- 4 “The reason we are implementing this is to prevent the **spread** of **Covid-19** from the affected areas..

EXCERPTS OF NEWS REPORT

NR1: He commented on the personal hygiene as individuals and as a country. This is to slow down the *spread* of **Covid-19** within our communities if or when it were to get into our communities.

NR2: Foreign travellers from four countries (China, South Korea, Italy and Iran) are now banned from entering Sarawak after the state government decided to extend its entry ban to foreigners to prevent the *spread* of **Covid-19** in the state.

NR3: Starting Sunday, March 22 at 12.01am, the Armed Forces will be deployed to help the police to enforce the movement control order to curb the *spread* of the **Covid-19** pandemic.

NR4: “The reason we are implementing this is to prevent the *spread* of **Covid-19** from the affected areas. This will enable the tracking operation of cases from house to house for 14 days, said Ismail Sabri, adding that all together, 83 cases of Covid-19 were detected in Kluang”.

The first example from the concordance list for the collocate *spread* refer to hygiene measures which everyone must practice. The next three examples (condordance list) and news reports is about mitigation measures taken by the government, which includes travel ban to foreign travellers, deploying the Armed forces to assist in the MCO enforcement and enhanced movement control order (EMCO) in Kluang (identified as red zone due to high positive cases of Covid-19). The occurrence of the collocate *spread* is more frequent to the left of the node word (see table 3), which means that the spread of the Covid-19 had affected individual and the country in various ways. An observation of the concordance list for the verb collocate *spread* indicated that references are made to the government's actions to assure the people that the government is doing whatever that is necessary to control and prevent Covid-19 from spreading. In NR2, the army has been deployed to assist in the MCO enforcement, thus, construing the fact that the government is now in total control of the situation.

ANALYSIS OF VERB COLLOCATE *CURB*

This section presents analysis for the verb collocate *curb*.

Examples of verb collocate *curb* from concordance list:

- 1 “We are doing contact tracing to **curb Covid-19 outbreak**”, says the Datuk Dr Noor Hisham.
- 2 Unreasonable Malaysians are making life difficult for medical frontliners who are already working around the clock to **curb the Covid-19 outbreak**
- 3 The government is doing its best to **curb the Covid-19 disease outbreak**, says Datuk Seri Dr Wee Ka Siong
- 4 Let’s deliver a proper message so that the community will take the necessary steps to help **curb Covid-19 transmissions**,” he said

EXCERPTS OF NEWS REPORT

NR1: More than 3,000 individuals who attended the tabligh programme in Masjid Jamek Sri Petaling have come forward to be tested after 4,942 of them were identified, says the Health Ministry. “We are doing contact tracing to **curb Covid-19 outbreak**”, says the Datuk Dr Noor Hisham.

NR2: Unreasonable Malaysians are making life difficult for medical frontliners who are already working around the clock to **curb the Covid-19 outbreak**. Refusing to be quarantined, demanding unnecessary screenings, lying about their travel history, and concealing the fact that they have had contact with a positive patient – these are just some of the issues medical personnel are facing, said the health director.

NR3: The government is doing its best to **curb the Covid-19 disease outbreak**, says Datuk Seri Dr Wee Ka Siong. In giving this assurance, the Transport Minister also reminded the people to refrain from unnecessary movements during the enforcement of the movement control order, including interstate travel.

NR4: We have worked out safety measures according to the timeline. “Let’s deliver a proper message so that the community will take the necessary steps to help **curb Covid-19 transmissions**,” he said

The verb *curb* from the concordance list and the news reports refer to statements made by the director general from the Ministry of Health (NR1, NR2 and NR4) and the Minister of Transport (NR3) to remind the people about travel restrictions during MCO. The verb *curb* can have positive and negative semantic connotations. According to Halliday (1977) the denotative, connotative or expressive meaning of a text is related to the meaning of the text and the construction of the text, where one aspect that must be taken into account is the meaning of the text and the construction of the text. Based on the news reports, the verb *curb* has positive connotation, for instance, in example three and four ‘doing its best’ and ‘to help’ imply attempts taken to curb the Covid-19 virus. With reference to negative connotations, in the four examples, Covid-19 is associated with nouns ‘outbreak’ and ‘transmission’. The association of these nouns with Covid-19 indicated that Covid-19 virus is dangerous and is spreading rapidly and widely, thus, efforts should be taken immediately and in various ways to curb the outbreak of Covid-19. The four examples illustrated the government’s call to the people to assist the government in controlling the spread of Covid-19, by cooperating with the Health department (NR1 and NR2), adhering to MCO travel rules (NR3) and practising safety measures, most

probably where hygiene practices is concerned (NR4). The message from the government is that their effort to curb the virus can only be successful with help from Malaysians.

ANALYSIS OF VERB COLLOCATE *FIGHT*

The analysis for the verb collocate *fight* is presented in this section.

Examples of verb collocate *fight* from concordance list

- 1 will be sending medical equipment to help Malaysia **fight** against the Covid-19 outbreak
- 2 The foundations' goal is to aid the **fight** against the Covid-19 virus in the country, with more support
- 3 government will allocate RM5mil in special allocation to **fight** against Covid-19, says Kelantan Menteri Besar Datuk Ahmad Yaakob
- 4 strengthen the Health Ministry efforts in the **fight** against the Covid-19 outbreak, say Health director-general Dr Noor Hisham Abdullah.

Examples of news excerpt:

NR1: A statement issued today confirmed China will be sending medical equipment to help Malaysia *fight against* the **Covid-19** outbreak. This pledge is made by the China's Ambassador to Malaysia today.

NR2: A statement issued by Jack Ma Foundation said the foundations had reached out to the governments of these four South-East Asian countries and proposed the donation of medical supplies. ... The foundations' goal is to aid the *fight against* the **Covid-19** virus in these nations, with more support to follow for other neighbouring countries that are in need.

NR3: Kelantan state government will allocate RM5mil in special allocation to *fight* **Covid-19**, says Kelantan Menteri Besar Datuk Ahmad Yaakob.

NR4: The government's decision to place the country under nationwide movement control will strengthen the Health Ministry efforts in the *fight* the **Covid-19** outbreak, says Health director-general Dr Noor Hisham Abdullah.

The four examples for the collocate *fight* demonstrated actions taken to stop Covid-19 outbreak. China's promise to send medical equipment to Malaysia (NR1) and aid from the Jack Ma's foundation to fight Covid-19 in Malaysia (NR2) obviously echoed the observation made about the health mitigation measures to curb the virus. NR3 and NR4 are about effort by the Kelantan's state government and the Ministry of Health's actions to control the outbreak. The verb *fight* is used with preposition *against* (example 1 and 2 – the concordance list), to emphasize the action of 'going against something'. The object of *fight* is obviously Covid-19. The three news reports highlighted the governments' actions and commitment to curb the Covid-19 outbreak. All the news report highlighted the governments' actions and commitment to curb the Covid-19 outbreak, that is also supported by Malaysia's main trade partner; China and the Jack Ma foundation (Jack Ma is a Chinese citizen).

COLLOCATION ANALYSIS OF NOUN COLLOCATE

The next collocate analysis is for noun collocate, which is the second lexical group that frequently collocates with the node Covid-19. The concordance analysis revealed that noun collocate occurred 36 times (see figure 3). The top 10 noun collocate is shown in table 4. The

top three most frequent noun collocates – cases, outbreak and patients will be analysed as examples.

TABLE 4. Frequency count for the top ten noun collocates

Freq	Freq L	Freq R	MI	Collocate
401	46	355	5.61084	cases
250	14	236	6.3527	outbreak
123	17	106	5.36004	patients
31	10	21	5.45532	tests
25	11	14	5.30063	deaths
18	0	18	5.17629	infections
16	7	9	5.67516	transmission
16	4	12	5.53429	surveillance
13	13	0	6.26868	battle
11	0	11	5.61264	zones

The concordance analysis indicated that the noun collocates are used to refer to cases of Covid-19 infections, death and places that have Covid-19 cases, and actions taken to curb Covid-19. The news reports suggest negative consequence and results.

ANALYSIS OF NOUN COLLOCATE CASES

This section presents the analysis for the noun collocate cases, based on a few examples extracted from the concordance list and news reports.

- 1 Malaysia recorded two deaths from **Covid-19 cases** on Tuesday (March 17), the first fatalities due to the pandemic
- 2 out of Malaysia's 673 confirmed **Covid-19 cases**, nearly two-thirds are reportedly linked to the four-day meeting
- 3 There are 212 new **Covid-19 cases** in the country, with the death toll now at 14
- 4 Sabah has tightened its Immigration restrictions following an increase in **Covid-19 cases** in the state

Examples of news report:

NR1: Malaysia recorded two deaths from **Covid-19 cases** on Tuesday (March 17), the first fatalities due to the pandemic in the country.

NR2: Out of Malaysia's 673 confirmed **Covid-19 cases**, nearly two-thirds are reportedly linked to the four-day meeting, Dr Adham said. The reality is that it will be difficult to trace all the participants because of the mobility of people...

NR3: There are 212 new **Covid-19 cases** in the country, with the death toll now at 14. Health director-general Datuk Dr Noor Hisham Abdullah said that as of Monday (March 23) noon, 212 new Covid-19 cases were recorded.

NR4: Sabah has tightened its Immigration restrictions following an increase in **Covid-19 cases** in the state. State secretary Datuk Safar Untong said that state, federal civil servants and officials seeking to enter Sabah...

The four examples for the noun collocate *cases* made references to positive Covid-19 cases. Words associated with Covid-19 cases, which are 'deaths', 'confirmed', 'new', 'increase' attested the fact that the situation is alarming and Covid-19 cases in Malaysia are increasing. Therefore, Malaysians must be adhere to the MCO rules, to help curb the outbreak of Covid-19.

ANALYSIS OF NOUN COLLOCATE *OUTBREAK*

Analysis for the collocate noun *outbreak* is presented in this section.

Examples of noun collocate outbreak from the concordance list

- 1 Currently Malaysia is facing various threats from the **Covid-19 outbreak** to trade war and as such policies and reform
- 2 ...the Health Ministry efforts in the fight against the **Covid-19 outbreak**, say Health director-general Dr Noor Hisham Abdullah
- 3 ...appointed senior ministers to coordinate their responses on the **Covid-19 outbreak** that has resulted in
- 4 additional measures to ensure “the effective delivery of financial resources to combat the **Covid-19 outbreak**”

Examples of news report:

NR1: Currently, Malaysia is facing various threats, from the Covid-19 *outbreak* to trade war, and as such policies and reforms are needed to stimulate the economy and strengthen the country's foundations.

NR2: The government’s decision to place the country under nationwide movement control will strengthen the Health Ministry efforts in the fight against the Covid-19 *outbreak*, say Health director-general Dr Noor Hisham Abdullah.

NR3: Singapore and Malaysia have appointed senior ministers to coordinate their responses on the **Covid-19 outbreak** that has resulted in Malaysia imposing a movement control order.

NR4: Tax deductions will be given to those who donate to the Covid-19 fund and Health Ministry, said the Finance Ministry. Finance Minister Tengku Datuk Seri Zafrul Tengku Abdul Aziz said this was part of additional measures to ensure “the effective delivery of financial resources to combat the **Covid-19 outbreak**”.

The noun *outbreak* refers to something dangerous and unpleasant, hence, the four examples from the concordance list and news reports refer to actions taken to control and reduce the impact of the Covid-19 outbreak. Furthermore, ‘threat’, ‘fight’ and ‘combat’ (in the concordance list and news reports) which is associated with Covid-19 outbreak is meant to depict the severity of the outbreak. The government’s decision to impose MCO is supported by the Health Ministry and considered timely due to the increase in Covid-19 cases (NR2). The MCO restrictions has also affected individual who work in Malaysia and Singapore, thus, appropriate travel arrangements between the two countries have to be made (NR3). What can be observed from the four examples are calamity that can happen if mitigation actions are not taken because the outbreak is a threat to Malaysia’s economy (NR1 and NR4). Therefore, safety measures should be the priority (NR2 and NR3).

ANALYSIS OF NOUN COLLOCATE *PATIENTS*

This section presents the analysis for the noun collocate *patients*.

Examples of noun collocate *patients* from the concordance list.

- 1 On Sunday, a meeting on **Covid-19 patients** and suspected case management among private hospitals and clinics
- 2 The Ministry will do the necessary and contact all close contact of **Covid-19 patients**, he explained
- 3 All cases currently tested positive were all linked to travel history and close contact of **Covid-19 patients**, said Health director general
- 4 Azizah Aminah Maimunah Iskandariah conveyed their condolences Wednesday (March 18) to the family members of two **Covid-19 patients** who died Tuesday (March 17).

Examples of news report:

NR1: On Sunday, a meeting on Covid-19 *patients* and suspected case management among private hospitals and clinics were held at the national level to increase understanding of the disease among private health practitioners as well as getting them prepared to face the Covid-19 situation, said Dr Noor Hisham.

NR2: The ministry hopes everyone can remain calm and continue to cooperate with us. The Ministry will do the necessary and contact all close contact of **Covid-19 patients**, said Dr Noor Hisham.

NR3: All cases currently tested positive were all linked to travel history and close contacts of **Covid-19 patients**, said Health director-general Datuk Dr Noor Hisham Abdullah.

NR4: The Yang di-Pertuan Agong Al-Sultan Abdullah Ri'ayatuddin Al-Mustafa Billah Shah and the Raja Permaisuri Agong Tunku Hajah Azizah Aminah Maimunah Iskandariah conveyed their condolences Wednesday (March 18) to the family members of two **Covid-19 patients** who died Tuesday (March 17).

The noun collocate *patient* is associated with hospital and hospital is a place where injured or ill people go to seek treatment. Thus, the noun collocate *patient* connotes an unpleasant situation. The first three examples from the concordance list and the news reports made references to confirmed cases of Covid-19 patients, while the final example refers to death of Covid-19 patients. It is apparent that noun collocate *patient* refer to incidents about Covid-19 cases and death. The four examples taken from the concordance list made references to Covid-19 patients' situations in hospitals (example one and three) and the health ministry's action to trace and test individual who have close contact with Covid-19 patients (example two and four). It is obvious that the four examples are actions taken to curb and control the outbreak of Covid-19.

DISCUSSION

The collocate analysis of Covid-19 revealed that the top two most frequent collocates are verb and noun. Although the corpus linguistics analysis in the present study is based on newspaper reports in only one Malaysian newspaper, some of the collocates identified in this study are in the list of top twenty collocates of *coronavirus* identified by the Oxford Dictionary, as part of its effort to continually monitor the linguistics developments of some of the words and phrases which have become increasingly familiar in the context of the current global crisis. A list of top twenty collocates of coronavirus in January, February and March has been compiled by the Oxford Dictionary (<https://public.oed.com/blog/corpus-analysis-of-the-language-of-covid-19/>). The collocates from this study which are also in the Oxford Dictionary are *outbreak, spread, test, positive, fight, case, impact, infection, patient, death, died, suspected*. This confirms that there are common and universal collocates of Covid-19 with reference to news about Covid-19.

Verb and noun collocates which occurred with the node word *Covid-19*, such as *spread, curb, fight, combat, battling, outbreak, transmission, death* and *battle* is associated with fear because the virus is spreading quickly and is affecting Malaysians in all aspects of life; health, economy, social, psychology and so forth. The noun *outbreak*, for instance, depicts/connotes that Covid-19 is a negative event. In fact, *outbreak* refers to something unpleasant and difficult to control, such as a war or disease. Observation from the concordance list revealed that noun collocates *outbreak* refer to death, positive cases of Covid-19 and mitigation measures. This means that the virus is spreading rapidly and widely, hence, actions should be taken immediately and in various ways to curb the outbreak of Covid-19.

Some of the verb collocates and noun collocates found in this study resonates the results in Trčková's (2015) study about Ebola virus, who reported that the general population was depicted as potential victim. This is demonstrated by verb collocates such as *spread, fight, prevent, died, combat and battling* that implied the threat of Covid-19 to the health of Malaysians, while noun collocates *patients* and *cases* refer to Malaysians who have been infected. Trčková (2015) also mentioned that the public institutions were portrayed as heroes or warriors. Analysis of the verb collocates *curb* and *fight* revealed that the Malaysian government is taking actions through various mitigation measures to stop the spread of Covid-19 and to control its outbreak. This finding is akin to Chung's (2011) study about the SARS virus. Chung's (2011) analysis of news report in the New Straits Times (a Malaysian newspaper) and The Times (a newspaper in the United Kingdom) revealed that the Malaysian press acted as a patriot in an effort to unite the nation by encouraging the nation to fight against SARS.

The word *outbreak* means something unpleasant and difficult to control, such as a war or disease. Analysis of the noun collocates *outbreak* revealed that *outbreak* is associated with words such as 'threat', 'fight' and 'combat' to depict the severity of the virus. The results echoes that of Balteiro's (2016) study on metaphoric portrayal of the Ebola virus. Balteiro (2016) found that the Ebola virus was associated with *fear* and *catastrophe*. The findings in this study also confirm the reason why there are about 40 news reports every day in the Star about Covid-19. The daily updates about Covid-19 from the Malaysian government and the Malaysian Ministry of Health in the newspapers were meant to provide accurate information to the public and to address fear about Covid-19. The result resonates a study by Wong et al. (2015) who reported that the African government used tweets frequently to address fear about the Ebola virus and to counteract inaccurate Ebola information in social media.

CONCLUSION

The present study describes how a corpus driven approach has revealed insights into the collocates of Covid-19. The top frequent verb and noun collocates revealed what the Malaysian government is doing, how Malaysians feel about the virus and updates about Covid-19 situation in Malaysia. Feelings of fear and threat of the virus are evident based on the collocates analysis. Although the data was only from a month of newspaper reports, the corpus analysis has revealed collocates which significantly co-occurred with the mention of Covid-19, such as *spread, fight, prevent, died, combat and battling* (verb collocates) and *outbreak, cases, patients, and tests* (noun collocates).

Even though this study does not exhaust the analysis of newspapers reports about Covid-19 in other English newspapers in Malaysia, we conclude that the corpus driven approach has provided some insights into the discourse of the government and Malaysians actions, views and responses to this pandemic and the global crisis attributed to this pandemic. Freedman and Wilder-Smith (2020), in fact, argue that one of the important protocols during this crisis must include engaging the public on the seriousness of, and their role in, the response

[to COVID-19]. In addition, this study contributes to the limited number of studies about infectious diseases using a corpus driven approach, since previous research about infectious diseases mostly focused on metaphorical conceptualization of infectious diseases which has enriched our understanding about metaphors. The present study is timely in view of the contribution to linguistics about Covid-19. We recommend several future researchers about this pandemic such as comparison of newspaper reports in English in Malaysia, analysis of newspaper reports during MCO or the extended movement control order period (EMCO) and discourse analysis or critical discourse analysis of crisis communication in pandemic situations.

REFERENCES

- Anagnostou, N.K. & Weir, G.R.S. (2006). From Corpus-Based Collocation Frequencies to Readability Measure. <https://www.semanticscholar.org/paper/From-corpus-based-collocation-frequencies-to-Anagnostou-Weir/be513abe8f3aef31c372af18505c227db71984e0>
- Balteiro, I. (2017). Metaphor in Ebola's popularized scientific discourse. *Iberica*. 34, 209-230. Cambridge Online Dictionary (<https://dictionary.cambridge.org/>)
- Chiang, Wen-yu & Ren-Feng Duann. (2007). Conceptual metaphors for SARS: 'War' between whom? *Discourse and Society*. 18, 579-602.
- Choi, D-H., Woohyun Yoo, W., Ghee-Young Noh, G-Y. & Park, K. (2017). The impact of social media on risk perceptions during the MERS outbreak in South Korea. *Computers in Human Behavior*. 72, 422-431.
- Clarivate Analytics Coronavirus Resource. <https://clarivate.com/coronavirus-resources/>
- Clarke, V. & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *Psychologist*. 26(2), 120-123.
- Corpus analysis of the language of Covid-19. <https://public.oed.com/blog/july-2020-update-scientific-terminology-of-covid-19/>
- Covid-19 bibliometrics Research Trends. <https://www.covid19bibliometrics.org/research-trends>
- Chouliaraki, L. & Fairclough, N. (1999). *Discourse In Late Modernity: Rethinking Critical Discourse Analysis*. Edinburgh: Edinburgh University Press.
- Chung, Siaw-Fung. (2011). A Corpus-Based Study of SARS in English News Reporting in Malaysia and in the United Kingdom. *International Review of Pragmatics*. 3, 270-293.
- Do, H.J., Chae-Gyun Lim, C-G., You Jin Kim, Y.J. & Choi, H-J. (2016). Analyzing Emotions in Twitter During a Crisis: A Case Study of the 2015 Middle East Respiratory Syndrome Outbreak in Korea. *IEEE 2016*. 415-418
- Elsevier Coronavirus Information Centre. <https://www.elsevier.com/connect/coronavirus-information-center>.
- English POS tagger. <https://parts-of-speech.info>
- Fairclough, N. (1992). *Discourse and Social Change*. Cambridge: Polity Press.
- Fairclough, N. (1995c). *Critical Discourse Analysis: The Critical Study of Language*. New York: Longman.
- Feijen, H.L.M. (2016). The opinion Dutch people have on Ebola on Twitter versus what message about Ebola is spread in Dutch newspapers: a comparison. Unpublished MA Thesis.
- Freedman, D.O. & Wilder-Smith, A. (2020). Isolation, Quarantine, Social Distancing and Community Containment: Pivotal Role for Old-Style Public Health Measures in the Novel Coronavirus (2019-Ncov) Outbreak. *Journal of Travel Medicine*, 27 (2), 1-4. <https://doi.org/10.1093/jtm/taaa020>
- Halliday, M.A.K (1977). 'Text as semantic choice in social context'. In Teun A. van Dijk and János S. Petöfi (ed.) *Grammars and Descriptions*. Walter de Gruyter, 176-226.

- Joye, S. (2010). News discourses on distant suffering: a Critical Discourse Analysis of the 2003 SARS outbreak. *Discourse and Society*. 21(5), 586-601. doi: 10.1177/0957926510373988
- John Hopkins Coronavirus Resource Center. (2020). <https://coronavirus.jhu.edu/map.html>
- Katermina, V. & Yachenko, E. (2020). Axiology of COVID-19 as a Linguistic Phenomenon in English Mass Media Discourse. *Advances in Journalism and Communication*. 8, 59-67. <https://www.scirp.org/journal/ajc>
- Lin, L. McCloud, R.F., Bigman, C.A. & Viswanath, K. (2016). Tuning in and catching on? Examination of the relationship between pandemic communication and awareness and knowledge of MERS in the USA. *Journal of Public Health*. 39(2), 282-289.
- Mazlin Azizan, Hanita Hanim Ismail & Shatha Naiyf Qaiwer. (2020). Power and Solidarity in Positive Facebook Postings Amidst Covid-19 in Malaysia. *Journal of Nusantara Studies*. 5(2) 329-364. <http://dx.doi.org/10.24200/jonus.vol5iss2pp329-364>
- Mohammad Awad AlAfnan. (2020). COVID 19-The Foreign Virus: Media Bias, Ideology and Dominance in Chinese and American Newspaper Articles. *International Journal of Applied Linguistics & English Literature*. 56-60. <http://dx.doi.org/10.7575/aiac.ijalel.v.9n.1p.56>
- Mondragon, N.I., Gil de Montes, L. & Valencia, J. (2017). Ebola in the Public Sphere: A Comparison Between Mass Media and Social Networks. *Science Communication*. 3(1), 101-124.
- Nor Fariza Mohd Nor, Anis Nadiah Che Abdul Rahman, Azhar Jaluddin, Imran Ho Abdullah, Sabrina Tiun. (2019). A Corpus Driven Analysis of Representations Around the word 'ekonomi' in Malaysian Hansard Corpus. *GEMA Online® Journal of Language Studies*. 19(4), 66-95.
- Pang, N., Karanasios, S. & Anwar, M. (2019). Exploring the information worlds of older persons during disasters. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.24294>.
- Tannen, D., Hamilton, Heidi, E. & Schiffrin, D. (2015). *A Handbook of Discourse Analysis*. (2nd edition). Oxford: Wiley-Blackwell.
- The Star Online. <https://www.thestar.com.my>
- Toppenberg-Pejcic, D., Noyes, J. Allen, T., Alexander, N., Vanderford, M. & Gamhewage, G. (2019). Emergency Risk Communication: Lessons learned from a rapid view of recent gray literature on Ebo, Zika and yellow fever. *Health Communication*. 34(4) 437-455. Doi: 10.1080/10410236.2017.1405488
- Tran, T. & Lee, K.(2016)Understanding citizen reactions and Ebola-related information propagation on social media. *In Advances in Social Networks Analysis and Mining (ASONAM), IEEE/ACM International Conference*. IEEE: 106-111
- Trčková, D. (2015). Representations of Ebola and its victims in liberal American newspapers. *Topics in Linguistics*. 16. doi: 10.2478/topling-2015-0009
- Tshiswaka, Daudet, L., Whembolua, Guy-Lucien, S. & Conserve, Donaldson, F. (2017). A Qualitative Analysis of Newspaper Response to the Ebola Outbreak in Central Africa. *Africology: The Journal of Pan African Studies*. 10(1), 224-239
- Wallis, P. & Nerlich, B. (2005). Disease metaphors in new epidemics: the UK media framing of the 2003 SARS epidemic. *Social Science & Medicine*. 60, 2629 - 2639.
- Wong, R., MPH, MSW; Harris, Jenine. K., Staub, M., MPH, MSW, Bernhardt, Jay. M. (2015). Local Health Departments Tweeting About Ebola. *Journal of Public Health Management Practice*. 1-9.
- Siaw-Fong. (2011). A Corpus-Based Study of SARS in English News Reporting in Malaysia and in the United Kingdom. *International Review of Pragmatics*. 3, 270-293.

- Van Dijk, T. A. (2001). Multidisciplinary CDA: A plea for diversity. In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis* (pp. 95–120). London: Sage.
- Waller, L. (2006). Introducing Fairclough's Critical Discourse Analysis Methodology for Analyzing Caribbean Social Problems: Going beyond systems, resources, social action, social practices and forces of structure or lack thereof as units of analysis. www.academia.edu/28963268/
- Weiss, G. & Wodak, W. (Ed.). (2002). *Critical Discourse Analysis: Theory and Interdisciplinarity*. New York: Palgrave Macmillan.
- Soden, R. & Palen, L. (2018). Informing crisis: Expanding critical perspectives on crisis informatics. Paper presented at the *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, New York City.
- Sinclair, J. (1991). *Corpus Concordance Collocation*. Oxford University Press.
- World Health Organization (WHO). (2020a). WHO Emergencies Press Conference on Novel Coronavirus. Geneva, Switzerland: WHO.
- World Health Organization (WHO). (2020b). Coronavirus Disease 2019 (COVID-19) Situation Report 55. Geneva, Switzerland: WHO.
- World Health Organization (WHO). (2020c). International Health Regulations (IHR) Emergency Committee regarding the Outbreak of Novel Coronavirus (2019-ncov). Geneva, Switzerland: WHO.
- Xiao, R. & McEnery, T. (2006). Collocation, Semantic Prosody, and Near Synonymy: A Cross-Linguistic Perspective. *Applied Linguistics*. 27(1), 107.

APPENDIX

NEWS REPORT 1

Trainee nurse among four new Covid-19 cases in Malaysia, bringing total to 29

Nation

Sunday, 01 Mar 2020 5:36 PM MYT

By LOH FOON FONG

KUALA LUMPUR: A trainee nurse has been infected with Covid-19, the first healthcare worker to be infected with the disease in Malaysia. The 20-year-old was one of the four new Covid-19 cases reported at noon Sunday (March 1), said the Health Ministry.

All the four cases were Malaysians and this brings the total number of those infected with Covid-19 to 29. The trainee, who was labelled as case 27, was undergoing short-term training at a private hospital and had treated case 24. She was identified as a close contact during tracing by the health district office.

NEWS REPORT 2

First healthcare worker infected

Nation

Monday, 02 Mar 2020

By LOH FOON FONG

KUALA LUMPUR: A 20-year-old Malaysian trainee nurse has been infected with Covid-19, the first healthcare worker to have the illness in the country.

Labelled as case number 27, she was undergoing short-term training at a private hospital and had treated case number 24. She had handled the patient without wearing any personal protective equipment (PPE) on Feb 21.

NEWS REPORT 3

Politicians test negative for Covid-19, says Health DG

Nation

Monday, 02 Mar 2020 8:30 PM MYT

By JO TIMBUONG

PETALING JAYA: Two high-profile individuals suspected of having Covid-19 have tested negative for the illness.

Health director-general Datuk Dr Noor Hisham Abdullah (pic) said both individuals had undergone nasopharyngeal swabs and the virus was not detected in any of the tests.

The Health Ministry confirmed when contacted on Monday (March 2) that both individuals were politicians. It was recently reported that former entrepreneur development minister Datuk Seri Mohd Redzuan Yusof and his deputy Datuk Dr Hatta Ramli had been in close contact with a Covid-19 patient.

They are said to have attended an appreciation event for them at a hotel here last Thursday that was also attended by a member of the UDA Holdings Berhad board of directors who has since tested positive for Covid-19.

ABOUT THE AUTHORS

Nor Fariza Mohd Nor (Ph.D) is an Associate Professor at the Center for Research in Language and Linguistics, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia. Her area of interests are critical discourse analysis, corpus linguistics and digital humanities. She is currently the Chair of the Centre for Research in Language and Linguistics.

Adlyn Shahirah Zulcafli is pursuing her MA in English Language Studies at the Center for Research in Language and Linguistics, Universiti Kebangsaan Malaysia.