

Collocation Networks and Covid-19 in Letters to the Editor: A Malaysian Case Study

Siti Aeisha Joharry
(Universiti Teknologi MARA)
Syamimi Turiman
(Universiti Teknologi MARA)

Joharry, S. A., & Turiman, S. (2020). Collocation networks and COVID-19 in letters to the editor: A Malaysian case study. *Asia Pacific Journal of Corpus Research*, 1(1), 1-30.

The present study examines language used to talk about the global coronavirus pandemic during a three-month period of movement control order in Malaysia. More specifically, a corpus of online letters to the editor of a local popular national newspaper was collected during the time in which the official quarantine instruction was initiated, resulting in a total of 303 online letters written by Malaysians that were analyzed through use of corpus linguistics techniques. For this purpose, the latest version of #LancsBox 5.0 (Brezina et al., 2020) is used to analyze patterns of language surrounding the portrayal of Covid-19 and further visualizing them by use of collocation networks. Findings present 25 statistically significant collocates that share an interesting relationship in revealing what the letters are about and thus, reflecting how Malaysians perceive and receive news about the pandemic during this time. Recurring topics and expressions include describing the virus in terms of metaphorical use of language (*Covid-19 does not discriminate*), preparing for an economic fallout (*Prihatin Economic Stimulus Package*), and preference to associate Covid-19 as a pandemic (*impacts of the Covid-19 pandemic*) rather than an outbreak (*first/second/third wave of the outbreak*). Implications of the study resonates with findings from

Azizan et al. (2020) where constructions of positive discourse among Malaysian writers may reflect the culture and society that make up the nation.

Keywords: Covid-19, Coronavirus, Corpus Linguistics, #LancsBox, Collocation Network

1. Introduction

Corpus linguistics has been around for centuries. Typically, these include research that analyzes language using sophisticated computer software such as AntConc, WordSmith, SketchEngine and many others. One particular function would be to do a collocational analysis. This means that words are further analyzed in terms of neighboring/surrounding words that reveal how words behave in context. Traditionally, these co-occurring words or also known as ‘collocates’ have been examined in terms of looking at patterns, two words at a time (the node or search word plus collocate). However, much development has seen the emerging trend of collocational *networks* (Brezina, 2016) and one tool in particular that builds collocation networks from user-defined corpora is GraphColl (Brezina et al., 2015). As Baker (2016) argues: “collocational networks give ‘added value’ to corpus analysis by indicating relationships between multiple words which can help to suggest equivalencies, synonyms, rewordings or related terms and concepts, which (in the case of a discourse-based analysis) may have ideological significance. They can also help to suggest relevant terms which may not have been considered for analysis in the first instance” (p. 14).

In recent studies, corpus techniques are being used to analyze language surrounding the coronavirus scare. Research has focused on issues related to Covid-19 in terms of media discourse (e.g. Katermina & Yachenko, 2020; Singer, 2020), news reporting/ headlines (e.g. Aslam, Awad, Syed, et al., 2020), and semantic prosody (e.g. Leo & David, 2020). What most, if not all of these corpus studies have in common would be the use of collocational analysis to examine how language is used in the coronavirus discourse. This paper adds to the existing pool of

research by examining an online public discourse on the pandemic through collocational networks provided using a more robust suite of tools, which is the #LancsBox 5.0.

2. Background: Collocation and Collocation Networks

2.1. Collocation

Halliday (1994) defines collocation as a simple tendency to co-occur. This means that when one word is used, there is a statistically high probability that a certain word or words will occur alongside it. Studying collocations allows us to have a deeper understanding of the meaning and the use of a word, than simply studying a word alone. But what constitutes a collocate? Brezina et al. (2015, p. 140) outlined three traditional criteria adopted for the automatic extraction of collocations as being based on: 1) distance, which is the maximum number of words around the node word which are to be searched for. This is often referred to as the window, or span, and is frequently set at four or five words to the left and right; 2) frequency, referring to how often a given type appears within the collocation window; and 3) exclusivity: the proportion between how often a given type occurs inside and outside the collocation window.

In addition, Brezina et al. (2015) also include three additional criteria arising from newer research (Gries, 2013), namely: 4) directionality, which tells about which word in the pair occurs before, or after; 5) dispersion, referring to how many cases of the collocates are occurring, and in how many texts in the corpus; and 6) type-token distribution, which measures the strength of collocation and the competition that a chosen collocate faces with other possible types of collocates for 'the slot' next, or to a specified distance, from the node word (Brezina et al., 2015, pp. 141-142).

Collocates can be identified using modern corpus linguistic software tools such as GraphColl in #LancsBox 5.0 and in this paper, the statistical measure to identify collocates will be based on the Mutual Information Score, or MI. As will be shown in the following sections, MI score has proven to be a reliable indicator of words that

are both frequently and infrequently associated, not likely due to chance.

2.2. Collocational Networks

Brezina et al. (2015) further expands the criteria for extracting collocation by proposing a seventh criteria -- connectivity. They argue that collocates do not occur in isolation, but in the broader context of the text. Additionally, collocates form chains, which in turn create semantic networks in the text. This idea that texts consist of “semantic networks of collocations” is originally conceived by Phillips (1983 as cited in Brezina et al., 2015, p. 139).

Collocational networks as defined by Williams (2001) is “a more objective methodology for the extraction of the lexis that typifies a field by making use of the lexical relationships into which words enter, that is through collocation” (p. 63). Thus, words are no longer seen simply as meanings in isolation, rather as forming larger units that may influence the cohesiveness of a particular discourse. The scholarship on collocational studies recognises two main traditions (Williams, 2001, p. 64): “a lexicographical school, which seeks to formalise collocation in functional and syntactic terms, and a contextualist school, which views the concept as a textual phenomenon”. In the present study, the latter school of thought is adopted as public opinion letters online are found to be an interesting register in its own right.

However, Williams (2001) cautions that “the more specialized the corpus, the more restricted the paradigm” (p. 67) and in turn, raises some limitations to the present study. Although the corpus is limited in terms of size, the reason behind collecting samples of writing during the three month containment period at a time of a health crisis in Malaysia provides researchers with a very situated type of discourse. As the internet became the only available channel to express oneself during the time, findings from this pool of data may generate highly-specialized associations in the discourse of online communities of practice and how people were dealing with the Covid-19 crisis.

Other forms of networks include Social Network Analysis (Brett, 2017) that uses Gephi (<https://gephi.org/>); a visualization software

that is available online for free, but is not considered a “one-stop” tool compared to GraphColl (p. 131). In exploring the patterns of relationships from networks, we find that collocation can go beyond this role in intra-textual cohesion to reveal patterns that are significant for texts emanating from a discourse community. These patterns may then be used to demonstrate the essential lexis of that community (Williams, 2001, p. 67). So in terms of the letters to the editor, collocation networks would tell us what the discourse is about during the time of the three-month movement control order (henceforth, MCO) period in Malaysia.

3. Methodology

The methodology used in this paper to analyze the type of discourse surrounding Covid-19 among the public online community during a specific containment period is that of collocation networks. Collocation networks as Brezina puts it are described as words that systematically co-occur in texts and discourse “to create a range of cross-associations that can be visualized as networks of nodes and collocates” (Brezina, 2016, p. 90). In other words, Williams (2001) states that the collocational networks are mainly statistical in that the associations are exploited in terms of what he calls ‘cohesive collocation’ - “the habitual and statistically significant relationship between word forms within a predefined window and for a defined discourse community, expressed through an electronic corpus of texts” (p. 67). Essentially, these networks not only provide a map that illustrates what different aspects may belong to texts or discourses (Brezina et al., 2015), but also how the texts/discourses are cohesive given Williams’ description of collocation.¹

For the present study, this idea has significant implications for understanding how Malaysians react to or view the current

¹ Williams reflects on the notion that “most frequent lexical words form obvious starting points, although in reality they rapidly join into one network that will take in central themes covered in the corpus” (2001: p. 71).

pandemic by examining their use of language in online letters to the editor of a widely read news portal: The Star Online. In so doing, letters that can be accessed on their website (<https://www.thestar.com.my/opinion/letters/>) are collected from the beginning of the official date for the MCO (18th March 2020) until the last day before the Malaysian Government lifted the ban (9th June 2020) and transitioned into a recovery movement control order (RMC0). This resulted in 303 letters that were written by Malaysians, which amounted to 169,306 words.

Using GraphColl under the freely available #LancsBox 5.0, collocates for the term ‘Covid-19’ were extracted from a span of 5 words to the left and right of the (node) word and the MI statistical measure was chosen because this “tends to privilege rare and specialized usage which makes it ideal for the exploration of special language corpora” (Williams, 2001, p. 68). More importantly, collocational networks reveal types of possible cross-associations (described as second-, third-order collocates) that may be unnoticed through previous one-way collocational analysis tools.

If we select the statistical value of more than 3.0 for MI stats as often used in the literature (Hunston, 2002), GraphColl extracts what Brezina calls an “overpopulated graph” as can be seen in Figure 1 below.

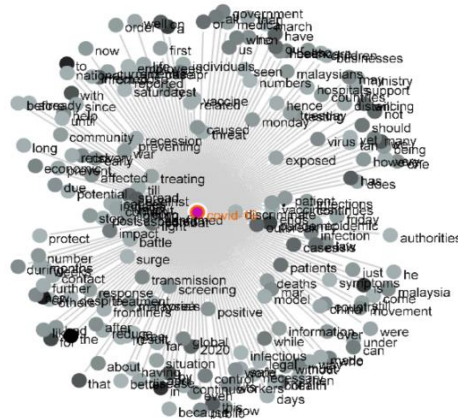


Figure 1. An Overpopulated Graph of ‘Covid-19’ Collocates

This allows us to increase the threshold of the MI-score to 6.0 where Baker (2016) has asserted to be more reasonable. Collocation patterns were also set to have collocates that occurred at least 5 times per text as per default. The new graph is presented in Figure 2 and in contrast to the previous figure, this one is more visible to the eye.

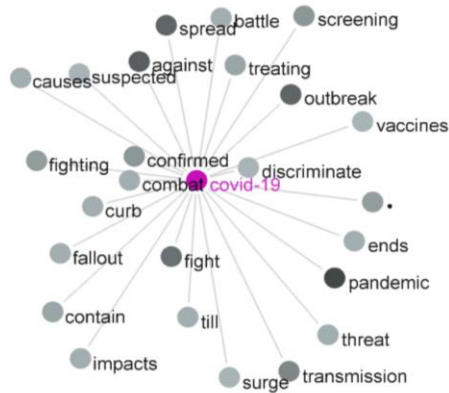


Figure 2. 25 Most Statistically Significant Collocates of ‘Covid-19’

Figure 2 shows 25 co-occurring words that were identified to be statistically significant and strongly associated with ‘Covid-19’, namely action verbs related to the virus (e.g. *spread*, *fighting*, *impacts*), nouns that are described by the virus (e.g. *battle*, *screening*, *outbreak*), and prepositional words like ‘against’ and ‘till’. In other words, these are what Brezina et al. (2015) call “first-order collocates” as defined through the MI statistic and respective cut-off points (threshold). The deeper colored collocates signal words with higher frequency, while the closer a collocate is to the node word (*Covid-19*), the stronger the association is (based on MI-value). It is also noteworthy to point that collocates appearing left of the node such as *curb*, *fallout* and *contain* indicate that these words occur before ‘Covid-19’ in the texts, whereas words like *vaccines*, *ends* and *pandemic* occur to the right or after ‘Covid-19’. This is similar to the presentation of these collocates in Table 1:

Table 1. 25 Top Collocates for ‘Covid-19’

No.	Collocate	MI
1	discriminate	7.20
2	confirmed	7.10
3	combat	7.10
4	fight	6.99
5	curb	6.81
6	treating	6.69
7	outbreak	6.66
8	against	6.61
9	till	6.57
10	fallout	6.46
11	fighting	6.42
12	suspected	6.42
13	spread	6.41
14	battle	6.36
15	ends	6.36
16	pandemic	6.36
17	vaccines	6.31
18	*	6.31
19	contain	6.16
20	screening	6.14
21	causes	6.10
22	threat	6.10
23	surge	6.10
24	transmission	6.02
25	impacts	6.02

Figures 1 and 2 present visual representations that would otherwise have been represented by “word clouds” in other software. They indicate the first-level order collocates that are extracted using MI-score and do not tell us much of the discourse without further examining the concordance lines. To create a collocation network using GraphColl, two or more collocates can be clicked on to see their respective collocates and how this connects to the bigger picture of the original node word (Covid-19). To illustrate this, the first two most frequent collocates of ‘Covid-19’ are

chosen: *discriminate* and *confirmed*

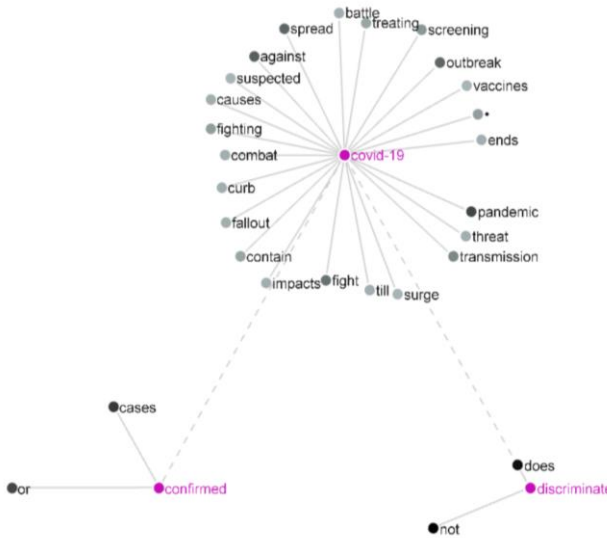


Figure 3. A Collocation Network of ‘Covid-19’ with ‘Discriminate’ and ‘Confirmed’

Discriminate, which was ranked as the highest significant collocate for the node word ‘Covid-19’ in Table 1 was found to be strongly co-occurring with words ‘does’ and ‘not’. In fact, further examination of the concordance lines reveal that the phrase “Covid-19 does not discriminate” were found, as seen in the following concordance lines:

- 1) in this fight against an invisible enemy Covid-19 *does not discriminate* by age, race, colour
- 2) their health and susceptibility to disease. Since Covid-19 *does not discriminate* based on ethnicity, neither
- 3) the spotlight on underlying socio-economic inequalities. While Covid-19 *does not discriminate* who it infects, the
- 4) the greater good of all Malaysians. Remember, Covid-19 *does not discriminate*. The virus does not
- 5) important for us to remind ourselves as Covid-19 *does not discriminate*, our response should not

Discriminate, which was ranked as the highest significant collocate for the node word 'Covid-19' in Table 1 was found to be strongly co-occurring with words 'does' and 'not'. In fact, further examination of the concordance lines reveal that the phrase "Covid-19 does not discriminate" were found, as seen in the following concordance lines:

The repeated lines show a metaphorical use of language in which Covid-19 is being personalized with animate actions such as to discriminate. This brings about a persuasive technique in writing where we argue in the later part of the paper how Malaysian writers strategize in their attempt to engage with their readers. Interestingly, this first look at the language highlights how Malaysian letters do not present discriminatory messages that one may tend to find in other texts related to fear and racism surrounding the coronavirus news (e.g. Leo & David, 2020; Noel, 2020).

What is also interesting are the second-order collocates around the node *confirmed* - 'cases' and 'or'. Upon clicking on these two words, the graph becomes livelier with extended words, co-occurring and even overlapping with some other words. Here, you can find instances of 'cases' that are written about the virus, specifically in the use of 'surge' whereas 'confirmed' is also described as related to either cases or someone that is infected with 'Covid-19' or as shown in the third-order collocate - 'suspected' of the virus (as shown in Figure 4). This in turn, demonstrates how use of a collocation network can show more connectivity between collocates at various levels of the collocational relationship and ultimately, revealing how a discourse is constructed -especially within public online letters to the editor in this context.

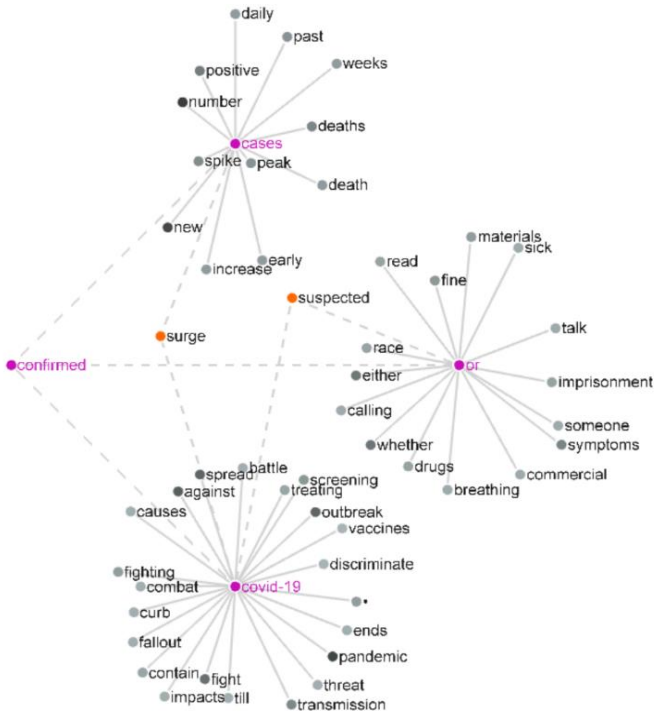


Figure 4. A Collocation Network of ‘Covid-19’ Showing Second-order Collocates of ‘Confirmed’

With this in mind, the following section presents findings that are a result of multiple collocation networks found to be interesting for further analysis. These include the use of metaphorical language related to war or conflict on Covid-19, the impacts or effects of the virus on the nation, as well as the framing of Covid-19 as a pandemic rather than an outbreak.

4. Research Findings

Covid-19 was the main interest of this study not only because it is among the most recent topic to be discussed, but because #LancsBox 5.0 revealed that it occurred frequently in the corpus, occurring almost 79% across the 303 letters and thus reasonable for further analysis. For reasons of space, the collocational network

analysis is carried out for some of the 25 significant collocates identified earlier (see Figure 2). We have categorized our findings into three sub-categories; each portraying a particular type of Covid-19 discourse in these letters.

4.1. Covid-19 and the Language of War or Conflict

In most linguistic descriptions of the pandemic, we have seen a rise in the reporting of the coronavirus in terms of war and combat (Pfrimer & Barbosa Jr, 2020). News reports have been portraying the current pandemic scene as a battle against a cruel enemy that must be defeated and while Pfrimer and Barbosa Jr (2020) contend that war-framings of COVID-19 in Brazil administrative discourse is dangerous in that it may lead to a militaristic response, Jaworska (2020)² discovered that in other parts of the world (in Germany for example), such language description or war rhetoric are not prevalent. Instead, she reveals that in the German press discourse, words collocating around the coronavirus are more methodological (e.g. 'informieren'/to *inform*, 'untersuchen'/to *investigate*, 'entwickeln'/to *develop*) and communicative in nature (e.g. 'vergleichen'/to *compare*, 'testen'/to *test*).

Based on the 25 first-order collocates shown in Figure 2, further examination was carried out on words related to conflict such as 'combat', 'fighting', 'fight', and 'battle'. This resulted in a collocation network for words related to war and conflict presented in Figure 5.

² <https://viraldiscourse.com/2020/04/13/is-the-war-rhetoric-around-covid-19-an-anglo-american-thing/>

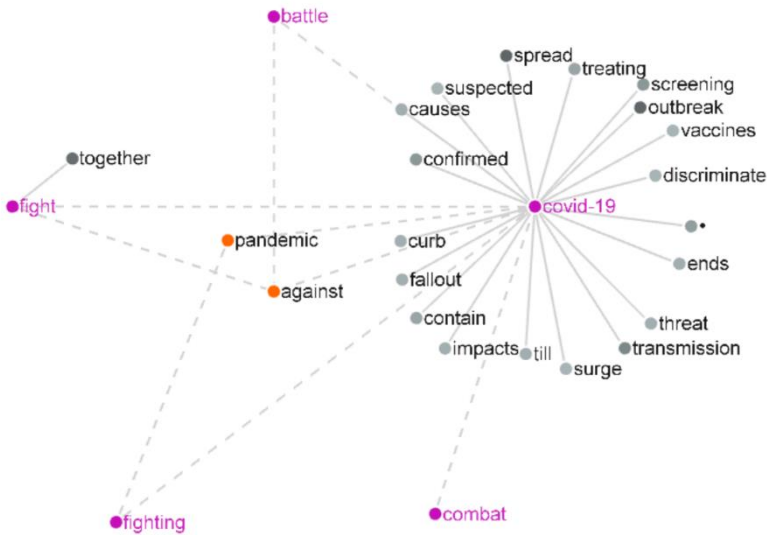


Figure 5. Collocation Networks for War/Conflict-Like Language

It can be seen that two collocates are shared: ‘against’ is found to be closely associated with three words (the node word ‘Covid-19’, ‘fight’, and ‘battle’), while ‘pandemic’ can be found to occur with ‘Covid-19’ and ‘fighting’. These result in what appears to be another metaphorical use of the language, especially where the agent or doer/subject is described as “fighting against Covid-19” (lines 1-4). More specifically, letters show a big response on how Malaysians view the virus as being the “toughest battle” (in line 8) the nation has had to face, at the same time hopeful for a “successful battle” (line 6) and that the “fight against” (lines 8-10) the virus was mostly referred to when calling readers’ attention to each personal message.

- 6) government's efforts so far in *fighting* the Covid-19 *pandemic* in the country. However, as a
- 7) concerned. As we focus on *fighting* the Covid-19 *pandemic*, we should also not neglect the
- 8) confident our healthcare system has been *fighting* Covid-19 as effectively they can. But this is

- 9) everyone can do their part in *fighting Covid-19*. These policies must be based on evidence
- 10) mark and pull our weight in *fighting Covid-19*. We must be disciplined and united in
- 11) political stability and a successful *battle* against Covid-19.
GAN CHEE KUAN Ipoh
- 12) Network (InciSiON), contributes to the *battle* against Covid-19 by organising its members from around the
- 13) are all aware that the *fight* against Covid-19 is the toughest *battle* we and most
- 14) at the workplace. In the *fight* against Covid-19, OSH committees should look into the following:-
- 15) *fight* against racism together! Let's *fight* against Covid-19 *together*! TAN SING PEI School of Humanities
- 16) the front lines of the *war* against Covid-19, which has spread to over 100 countries
- 17) all mental health issues. The *war* against Covid-19 is not over. We have not won
- 18) *fight* till we win this *war* against Covid-19 together, but the demotivation we feel now
- 19) in this *fight* against an *invisible enemy*. Covid-19 does not discriminate by age, race, colour

Upon further expansion of the network, it was revealed that the collocate ‘against’ also co-occurred with ‘together’ that were found mostly in instances alluding to a call to arms (e.g. *Let’s fight against Covid-19 together!* in line 10). Others include third-level collocates like ‘war’ and ‘enemy’ where the coronavirus is perceived to be an “invisible (common) enemy” (in line 14) and that Malaysians are describing it as a “war against Covid-19” (lines 11-13).

War rhetoric that most often makes use of metaphors to frame the coronavirus in terms of war and combat has for the most part function to persuade and appeal to readers in these letters. More specifically, by using this style of writing, we may know what writers’ evaluation and stance are on the subject matter. This concurs with Biber and Finegan (1989) who argue that personal

stance can be marked by expressions of attitudes, feelings, judgments, or commitment concerning the propositional content of a message. Additionally, McEntee-Atalianis (2013) mentions that metaphor marks attitude and orientations to context, propositions and social and political structure or relationships. Findings from the concordance lines indicate that while certain war-like discourse can be found, most instances implied positive discourse and solidarity strategies leading towards empowering the nation to “soldier on”, which suggest an uplifting role of these online communities of practice and their significant role in propagating messages of goodness (Azizan et al., 2020).

4.2. Impact or Effects of Covid-19

Another first-order collocates to look at would be the impacts or effects of Covid-19 that would be expressed in these letters. Letters expressed negative ‘impacts’ that were a result of the pandemic (e.g. *negative, disastrous, unprecedented*) and more specifically, ones related to the economy. Further inspection of the second-order collocates ‘economic’ revealed that it was also a shared collocate for ‘fallout’, adding to the negative connotation of the Covid-19 impact. This led to the connection/connectivity between what was described as “economic impacts” in some letters and “economic fallout” in others.

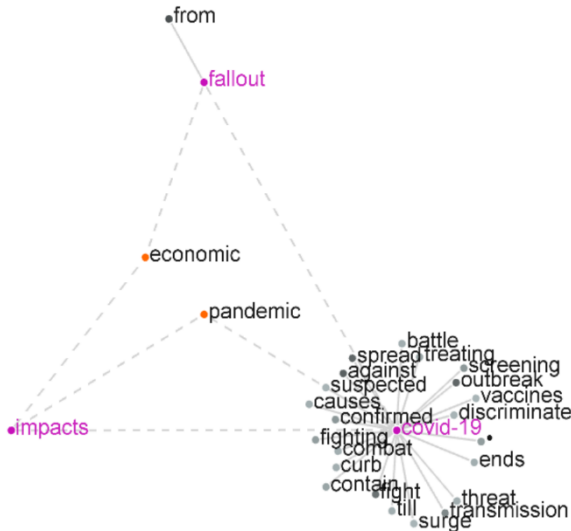


Figure 6. Collocation Networks for Collocates ‘Impacts’ and ‘Fallout’

A further look into ‘economic’ revealed 12 third-order collocates ranging from words describing an entity/organization (*institute*, *sectors*), monetary scheme (*package(s)*, *stimulus*, *recovery*) in lines 15, 16 and 19, condition (*recession*, *growth*, *environmental*, *activities*, *impact*) in lines 17, 20, and 23; and communication (*announced*) that co-occurred with ‘economic’. It can be argued that during the MCO period, people were taking stock of the social, political and to a greater extent -economical- fallout that followed the realization of the Covid-19 news. And while fear and panic can be felt in these letters, numerous suggestions and expressions were found from this growing reality of an economic fallout, mainly in relation to the government’s RM250bil economic stimulus package, which received much praise from the nation.

Given this, collocational networks between *Covid-19*, *impact*, *economic* and *fallout* have particularly shown how lexical connections in discourse point to letters that are more focused in terms of discussing the economic effects of the pandemic, usually providing a critical review on governmental actions or responses on a continuous basis. This, then, contributes to the shaping of a Covid-19 discourse during the three-month MCO period in Malaysia that

goes in tandem with the Malaysian government's actions of managing the coronavirus crisis at the time.

- 15) most exposed to the *fallout* from the Covid-19 outbreak. The scope of the stimulus package
- 16) alleviating the economic *fallout* caused by the Covid-19 crisis. These initiatives include a RM600mil allocation
- 17) are convinced that the economic *fallout* from Covid-19 would be far more disastrous than the
- 18) really prepared for the economic *fallout* from Covid-19? Monday, 30 Mar 2020 The extension of
- 19) to cushion the economic *fallout* from the Covid-19 outbreak. This staggering boost is more than
- 20) stressed out by the economic *fallout* from Covid-19. Whilst we deal with Covid-19's economic fallout,
- 21) inevitably be winners and losers from the economic *fallout* of the crisis.
- 22) In fact, some are convinced that the economic *fallout* from Covid-19 would be far more
- 23) healthcare system, while minimising the social and economic *fallout*. If this pandemic is war, mass

4.3. Covid-19 Pandemic Versus Covid-19 Outbreak: Pseudo-synonyms?

Finally, we take a look at the portrayal of the coronavirus as either more 'pandemic' or 'outbreak' within these letters. Figure 7 clearly shows that 'pandemic' has more shared collocates with 'Covid-19' (*impacts, till, ends, *, fighting*) compared to 'outbreak' and thus, suggests that the former collocate co-occurs more widely and frequently with 'Covid-19'.

With the nodes *Covid-19* and *pandemic*, it is firstly interesting to point out that six repeated lines were found with shared collocates 'till' and 'ends', appearing in one single text. The phrase "from April till Covid-19 pandemic ends" was identified to be made up of strongly associated words (that collocated together) mainly because they are used in such a way. Further inspection of the letter reveals

that the phrase refers to a stimulus package for health and medical officers suggested by the writer (Dr. Rais Hussin) and in turn, tells us a lot about how repetition can be used to create a particular style in writing. It was also found that the bullet point (•) was identified as a second-order collocate and this refers to the structural organization of his writing that made it easier for readers to read his points. Dr. Rais Hussin, who recently became Chairman of the Malaysia Digital Economy Corporation (MDEC) sounded hopeful to have shared bold recommendations (i.e. to ensure monetary compensation till pandemic ends) since we have yet to find out whether Covid-19 can be resolved completely.

- 24) dichotomised. Priorities are being reset. When this pandemic ends, and yes it will, everything we
- 25) RM750 to RM1,000 from April until the pandemic ends. • RM500 monthly allowance for medical
- 26) hospitals and clinics from April till Covid-19 pandemic ends. (Sarawak Model) • RM500 monthly allowance
- 27) points throughout Malaysia from April till Covid-19 pandemic ends. (Sarawak model) • RM500 monthly allowance
- 28) security of Malaysians from April till Covid-19 pandemic ends. (Sarawak Model) • RM500 monthly allowance
- 29) law and order from April till Covid-19 pandemic ends. (Sarawak Model) • Free food: Providing
- 30) risk of Covid-19 from April till Covid-19 pandemic ends. The expected economic stimulus to be

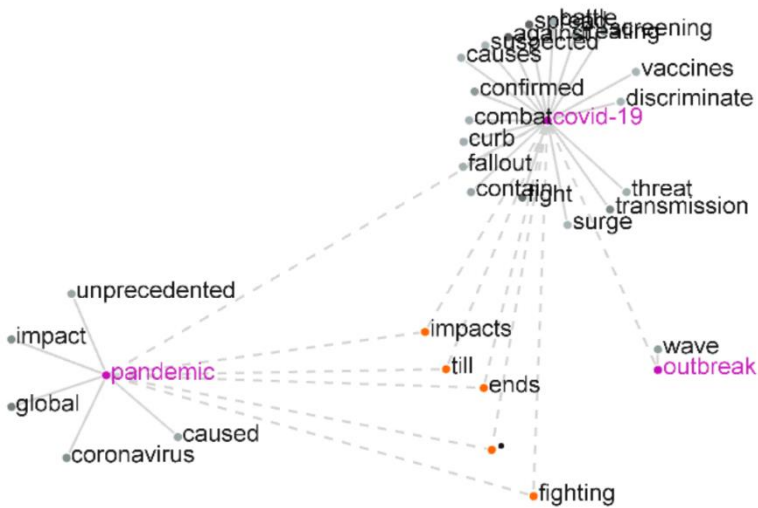


Figure 7. Collocation Networks for Collocates ‘Pandemic’ and ‘Outbreak’

One of the reasons for *pandemic* to be highly, frequently occurring with ‘Covid-19’ might be after the World Health Organization (WHO) declared the virus as a pandemic in March 2020. The node ‘pandemic’ has five second-order collocates that can be discussed into two parts: 1) adjectives describing the pandemic (*coronavirus, global, unprecedented*) and 2) cause and effects of the pandemic (*caused, impact*). Basically, these words are connected with each other to further describe Covid-19 as a pandemic - it is one that has not happened in the country before and is affecting worldwide. ‘Coronavirus’ is used interchangeably with ‘Covid-19’ and therefore may function as synonyms of each other. The connection between ‘pandemic’ and second-order collocates ‘caused’ and ‘impact’ is similar to the description of the impacts (especially those related to the economy) of the virus discussed in the previous section and therefore, is not repeated here.

- 31) warning: "We have never before seen a pandemic sparked by a *coronavirus*. This is the
- 32) all in this fight against the *coronavirus* pandemic. Together,

- we will emerge from this crisis
- 33) in response to climate change. The *coronavirus pandemic* is a grave wake-up call to humanity.
 - 34) in the current climate of the *coronavirus pandemic*, charlatan MLM cure-alls may seem like an
 - 35) However, with the impact of the *coronavirus pandemic* hitting Britain hard, we had lost 21
 - 36) streets in the middle of the *coronavirus pandemic* Before we are goaded into the fray,
 - 37) Street Journal (April 3) titled The *coronavirus pandemic* will forever alter the world order, America's
 - 38) public on matters pertaining to the *coronavirus pandemic*. The role of the Yang di-Pertuan Agong
 - 39) with regards to dealing with the *coronavirus pandemic* and its effects. Unfortunately, due to fiscal
 - 40) by a coronavirus. This is the first *pandemic* caused by a *coronavirus*. And we have
 - 41) international and national obligations. The *global Covid-19 pandemic* has entered a severe period, and not
 - 42) workers Saturday, 02 May 2020 The Covid-19 *pandemic* has hit the *global* economy hard as
 - 43) due to *global* recession and the Covid-19 *pandemic*. The truth is many elderly citizens are
 - 44) World Health Organisation declared *Covid-19* a *global pandemic*, my wife received a memo from the
 - 45) risk as well? Covid-19 is a *global pandemic* so how can one ascertain whether the
 - 46) Saturday, 04 Apr 2020 The Covid-19 *global pandemic* has caused worldwide panic as it continues
 - 47) collapse during and after the Covid-19 *global pandemic*. In reality, Malaysia's economy is actually very
 - 48) but have consequences nonetheless. The Covid-19 *global pandemic* is the epitome of a black swan
 - 49) jobless individuals! Prior to the Covid-19 *global pandemic*, human resources management was focused on the
 - 50) for non-citizens in this time of *global pandemic*. Life is life,

- irrespective of whether we
- 51) scientific rationale behind these rules. The *global pandemic* of Covid-19, which is caused by a
- 52) hopes to successfully recover from this *global pandemic*.
SALIM BASHIR President, Malaysian Bar
- 53) the dust has settled from this *global pandemic* and its after-effects. This includes the unfair
- 54) you big time! However, in this *global pandemic*, we seem to have somehow missed the
- 55) Malaysia is facing an unprecedented *global health pandemic* and domestic economic crisis. As we continue
- 56) especially during the initial phase of the *pandemic*. Due to the *global* shortage, many countries
- 57) is facing an *unprecedented* crisis. The Covid-19 *pandemic* occurs amid the backdrop of ongoing uncertainties
- 58) Mar 2020 The impacts of the Covid-19 *pandemic* in some respects are *unprecedented* in recent
- 59) a whole-of-society response to this *unprecedented* Covid-19 *pandemic*. StayHome/DudukRumah! DATUK DR MUSA MOHD NORDIN (Paediatrician)
- 60) to consider that even before this *unprecedented pandemic* emerged, every time we stepped out of
- 61) for us to reflect on this *unprecedented pandemic* and lockdowns of various forms implemented globally.

On the other hand, ‘outbreak’ may not function similarly to ‘pandemic’ and therefore, would be worth investigating. Firstly, it does not have any shared collocates with ‘Covid-19’ except for occurring 46 times as a first-order collocate. In the many instances of ‘Covid-19 outbreak’, these occurrences (lines 62-81)³ mainly describe the virus in relation to the definition of an outbreak (i.e. a sudden rise in the number of cases of a disease) and are used to show the cause/effects of the virus as well as express emotive

³ The remaining 28 occurrences are shown in Appendix

language. Outbreak carries the same definition of epidemic,⁴ but is often used for a more limited geographic area, whereas pandemic refers to an epidemic that has spread over several countries or continents, usually affecting a large number of people.

- 62) still within the first *wave* of the outbreak. The Philippines (17 deaths) and Indonesia (19
- 63) us facing the third *wave* of the outbreak with the second highest number of positive
- 64) ride into the second *wave* of the outbreak with positive cases doubling or even tripling
- 65) for the worst during the current Covid-19 outbreak in Malaysia: a third *wave* of infection
- 66) During the first *wave* of the Covid-19 outbreak in Malaysia, from Jan 25 to Feb
- 67) current Covid-19 *outbreak* in Malaysia: a third wave of infection that could see cases surging
- 68) Philippines and Indonesia ride into the second wave of the *outbreak* with positive cases doubling
- 69) to Italy which us facing the third wave of the *outbreak* with the second highest
- 70) 07 May 2020 The unexpected *outbreak* of Covid-19 and implementation of the movement control order
- 71) impacts of the *outbreak*, with people seeing Covid-19 as an economic crisis first and a
- 72) under strain even before the *outbreak* of Covid-19 as Malaysians are highly dependent on public
- 73) used to address this *outbreak*. Containment for Covid-19 is aimed at preventing it from spreading
- 74) teachers and the school. Before the *outbreak* of Covid-19 Malaysians were vehemently debating the different approaches
- 75) are consumed by the vagaries of the Covid-19 outbreak and its effects on societal medians

⁴ Epidemic refers to an increase, often sudden, in the number of cases of a disease above what is normally expected in that population in that area (Source: <https://www.merriam-webster.com/dictionary/epidemic>).

- 76) keep afloat in the face of the Covid-19 outbreak and the imposition of the movement
- 77) do not want a repeat of the Covid-19 outbreak associated with airconditioning in a restaurant
- 78) measures Monday, 27 Apr 2020 As the Covid-19 outbreak continues to grip the world, and
- 79) slide into an economic recession if the Covid-19 outbreak continues to worsen and the MCO
- 80) go up and society will suffer. The Covid-19 outbreak has brought everyone to their knees!
- 81) other adults talk to them about the Covid-19 outbreak in a way that is informative,

The node ‘outbreak’ however, can be seen to have one second-order collocate (*wave*) and this is shown to be used in a metaphorical description of temporal and spatial deictic that expresses different times of the outbreak (lines 62-69). Since the wide community containment plan was in effect, Malaysians were informed of the recurring “waves” of the infection that may be a result of the transmission of the disease. As a result, *wave* is used in close proximity to ‘outbreak’ compared to ‘pandemic’ in order to describe the sudden rise or increase of cases in the country. It can also be argued that by examining these collocation networks, we can safely say that ‘pandemic’ and ‘outbreak’ are pseudo-synonyms for ‘Covid-19’ in that they are not equivalently the same.

5. Discussion and Conclusions

Generally, this paper sets out to demonstrate new ways to analyze frequently co-occurring words or collocates, following collocational networks put forth by Brezina and his team behind #LancsBox 5.0. By using GraphColl, which is a tool in the suite, a set of online letters to the editor of The Star Online during the movement control order period in Malaysia was investigated to see how online communities of practice viewed and responded to the recent Covid-19 pandemic in the country.

This study presented a corpus-based analysis of 'Covid-19', particularly through use of a collocational network. Using GraphColl with MI as the statistical measure and a span of five words to the left and right, it was found that 'Covid-19' strongly occurs with words that namely describe action related to the virus (e.g. *spread, fighting, impacts*), nouns that are described by the virus (e.g. *battle, screening, outbreak*), and prepositional words like 'against' and 'till'. Building networks of these words thus, reveals the thematic content of the corpus, which has shown to be useful for the present study on Covid-19. As Williams (2001) notes, "[i]n exploring the lexis we gradually reveal terms, lexical phrases and lexicographic collocations that typify the lexis of the domain" (p. 71), which in our case: COVID-19. It is also feasible now to use GraphColl that allows for a "one-stop" methodology to build collocation networks that can then be studied in relation to individual letters of an online commentary group of writers, which in turn shows us how the Malaysian public respond to the pandemic during times of crisis.

In terms of our dataset, public online letters to the editor represent online commentary journalism as a particular register and to a certain extent, is what Rosenbaum and Shachaf (2010) describe as "online communities of practice". Examination of collocation networks have exemplified a form of relationship between words associated with 'Covid-19' that revealed some type of cohesiveness in the shaping of Covid-19 themes identified, which illustrates the discourse of this particular group of writers as representing the Malaysian public voice. As McNair (2008) points out, 'citizen journalism' describes the circulation of public opinion, and to a certain extent influences the way the mainstream media covers important events, such as COVID-19. According to Nguyen (2009), citizen journalism is able to generate significant changes in public life via a healthy public sphere, created by three elements that constantly interact with each other, namely journalism, social movements, and the discourses between citizens, experts and policymakers. The mainstream media carries the role of initiating an issue which will be 'amplified via online citizen discussions, and then echoed back to the mainstream agenda' (p. 160). These we have

seen to be highlighted in the collocational networks of 'Covid-19' that were also positively constructed to a certain degree such as in soft directives (*everyone can do their part in fighting Covid-19*) and giving advice (*Remember, Covid-19 does not discriminate*).

Studying language used in public letters to the editor is thus, important for many reasons. On a general scale, the corpus of letters written during the three-month MCO period has shown much interest in the talk about Covid-19. Zooming in via collocational networks, we have seen that Malaysians mainly write about the nature of the virus (in that it is non-discriminatory), the call for solidarity against the virus (*Let's fight against Covid-19 together!*) and the impacts it has had on the economy (particularly the "economic fallout" post-Covid-19). More importantly, it can be safely argued that although the world may be discussing about the coronavirus in a less positive light, Malaysian online writers (at least for these 303 letters), have shown to "highlight the goodness of the community, sharing similar notions on solidarity and empowerment" (Azizan et al., 2020). It is essential to reflect on Azizan et al's (2020) argument that "[u]nlike the West, Malaysians remain largely as a collective society that pays attention to conforming and standing with each other" and so while there are some similarities to the description or expression of Covid-19 as a problem with inevitable consequences, the call for action among Malaysians in the letters can be felt to be more of an encouragement rather than a complaint. "Language not only reflects the world around us; it can also reinforce reasoning and influence the ways people act and behave" (Jaworska, 2020), and hopefully, Malaysians turn to these letters in times of distress to find the right ways to respond to crises like the Covid-19.

References

- Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., & Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities and Social Sciences Communications*, 7(23), 1-9.

- Azizan, M., Ismail, H. H., & Qaiwer, S. N. (2020). Power and solidarity in positive Facebook postings amidst COVID-19 in Malaysia. *Journal of Nusantara Studies*, 5(2), 329-364.
- Baker, P. (2016). The shapes of collocation. *International Journal of Corpus Linguistics*, 21(2), 139-164.
- Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, 9(1), 93- 124.
- Brett, D. (2017). Collocate networks in the language of crime journalism. *Studii de Lingvistică*, 7, 125-144.
- Brezina, V., Weill-Tessier, P., & McEnery, A. (2020). #LancsBox v. 5.x. [software]. Available at: <http://corpora.lancs.ac.uk/lancsbox>.
- Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, 20(2), 139-173.
- Brezina, V. (2016). Collocation networks. In Baker, P., & Egbert, J. (Eds.), *Triangulating Methodological Approaches in Corpus Linguistic Research* (pp. 102-119). New York/London: Routledge.
- Gries, S. (2013). 50-something years of work on collocations: What is or should be next. *International Journal of Corpus Linguistics*, 18(1), 137-165.
- Halliday, M.A.K. (1994). *Introduction to Functional Grammar*. London: Edward Arnold.
- Hunston, S. (2002). Methods in corpus linguistics: Beyond the concordance line. In Hunston, S. (Ed.), *Corpora in Applied Linguistics* (pp. 67-95). Cambridge: Cambridge University Press.
- 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.
- Leo, A. R., & David, M. K. (2020). A critical metaphor analysis on Malaysia's gazetted metaphors amid the movement control order: A COVID-19 episode. *Horizon J. Hum. & Soc. Sci.*, 2(S), 193-204.
- Manderson, L., & Levine, S. (2020). COVID-19, risk, fear, and fall-out,

- Medical Anthropology*, 39(5), 367-370.
- McEntee-Atalianis, L. J. (2013). Stance and metaphor: Mapping changing representations of (organizational) identity. *Discourse and Communication*, 7(3), 319-340.
- McNair, B. (2008). I, columnist. In Franklin, B. (Ed.), *Pulling Newspapers Apart: Analysing Print Journalism* (pp. 106-114). London/New York: Routledge.
- Nguyen, A. (2009) Globalization, citizen journalism, and the nation state: A Vietnamese perspective. In Allan, S., & Thorsen, E. (Eds.), *Citizen Journalism: Global Perspectives* (pp. 153-162). New York: Peter Lang.
- Noel, T. K. (2020). Conflating culture with COVID-19: Xenophobic repercussions of a global pandemic. *Social Sciences & Humanities Open*, 2, 1-7.
- Pfrimer, M. H., & Barbosa Jr, R. (2020). Barzil's war on COVID-19: Crisis, not conflict – Doctors, not generals. *Dialogues in Human Geography*, 10(2), 137-140.
- Rosenbaum, H., & Shachaf, P. (2010). A structuration approach to online communities of practice: The case of Q&A communities. *Journal of the American Society of Information Science and Technology*, 61(9), 1933-1944.
- Williams, G. (2001). Mediating between lexis and texts: Collocational networks in specialised corpora. *Asp, la revue du GERAS*, 31, 63-76.

Appendix

Remaining 28 occurrences of 'Covid-19' and 'outbreak'

- 1) only know the true extent of the Covid-19 outbreak in our country many months later.
- 2) for patients while trying to curb the Covid-19 outbreak in our country. As the pandemic
- 3) stimulus packages Monday, 06 Apr 2020 The Covid-19 outbreak is likely to become the defining

- 4) legislation was already in place when the Covid-19 outbreak reached Europe. Made under a principal
- 5) will not be sufficient to bring the Covid-19 outbreak to a safe level. Knowing this,
- 6) second stimulus package to cope with the Covid-19 outbreak, announced last week by Prime Minister
- 7) the magnitude of effects caused by the Covid-19 outbreak, governments should introduce laws to save
- 8) the frontliners in the fight against the Covid-19 outbreak, he dedicated his service to the
- 9) to the workers' long-term concerns. Since the Covid-19 outbreak, the government has announced stimulus packages
- 10) of the movement control order following the Covid-19 outbreak. Even if they receive government assistance,
- 11) being stretched to the seams by the Covid-19 outbreak. Healthcare practitioners are on the verge
- 12) time of having to cope with the Covid-19 outbreak. However, we do not know how
- 13) services should not be compromised during the Covid-19 outbreak. One of these medical emergencies is
- 14) from the worst economic impacts of the Covid-19 outbreak. SHAMINI MANIKAM Managing director, TinkTank
- 15) is dictated by the lifespan of the Covid-19 outbreak. So please allow two persons instead
- 16) critical period, a period far beyond the Covid-19 outbreak. TAN SRI LEE LAM THYE Kuala
- 17) measure to "flatten the curve" of the Covid-19 outbreak. The aim is to decelerate infection
- 18) are involved in the fight against the Covid-19 outbreak. The basis of this suggestion is
- 19) Canadians cope with the impact of the Covid-19 outbreak. The House, having been suspended on
- 20) online at bit.ly/star) while dealing with the Covid-19 outbreak. The necessary movement control order to
- 21) most exposed to the fallout from the Covid-19 outbreak. The scope of the stimulus package

- 22) for dealing with deadly conflicts amidst the Covid-19 outbreak. The world seems to be overwhelmed
- 23) to cushion the economic fallout from the Covid-19 outbreak. This staggering boost is more than
- 24) that has been imposed to curb the Covid-19 outbreak. We are unsure of how many
- 25) in major cities there due to the Covid-19 outbreak. Worldwide, leaders are experimenting with novel
- 26) bank loans for individuals affected by the Covid-19 outbreak.
 - A six-month waiver on interest
- 27) bank loans for SMEs affected by the Covid-19 outbreak. • A six-month waiver on interest
- 28) recognises that since the *outbreak* of the Covid-19 pandemic, the legal profession cannot go back

THE AUTHORS

Siti Aeisha Joharry is a senior lecturer in the Department of English Language and Linguistics at Akademi Pengajian Bahasa, Universiti Teknologi MARA, Shah Alam. Her research interests include Corpus Linguistics, Corpus-assisted discourse analysis, and English for Professional Communication.

Syamimi Turiman is a senior lecturer in the Department of English Language and Linguistics at Akademi Pengajian Bahasa, Universiti Teknologi MARA, Shah Alam. Her research interests include Corpus Linguistics, Discourse Analysis, and English for Professional Communication.

THE AUTHORS' ADDRESSES

First and Corresponding Author

Siti Aeisha Joharry

Senior Lecturer

Department of English Language and Linguistics

Akademi Pengajian Bahasa

Universiti Teknologi MARA

40450 Shah Alam, Selangor, MALAYSIA

E-mail: aeisha@uitm.edu.my

Co-author**Syamimi Turiman**

Senior Lecturer

Department of English Language and Linguistics

Akademi Pengajian Bahasa

Universiti Teknologi MARA

40450 Shah Alam, Selangor, MALAYSIA

E-mail: syamimituriman@uitm.edu.my

Received: 15 July 2020

Received in Revised Form: 9 August 2020

Accepted: 14 August 2020