CORPUS-BASED ANALYSIS OF ONLINE HOAX DISCOURSE ON TRANSPORTATION SUBJECT PICTURING INDONESIAN ISSUE

Yune Andryani Pinem*

Transport Management, Sekolah Tinggi Teknologi Kedirgantaraan (STTKD)
Jl. Parangtritis KM 4,5 Sewon, Bantul, Daerah Istimewa Yogyakarta 55187, Indonesia
j.andryani@gmail.com

How to Cite: Pinem, Y. A. (2021). Corpus-based analysis of online hoax discourse on transportation subject picturing Indonesian issue. *Lingua Cultura, 15*(1), 39-50. https://doi.org/10.21512/lc.v15i1.7067

ABSTRACT

The research discussed Indonesian hoax on transportation, spread using chain messages over social media, a major hoax subject. Setting social structure as an element to consider in hoax discourse gave different perspectives on why hoax news took place. Contrasting identified online hoax discourse to corpus-derived prosody could be interpreted using a socio-pragmatic approach by Lin and Chung (2016). The research tried to look at how discourse on hoax correlated to real social issues among Indonesian people. Digital forensic was done by tracking discourse used as data from hoax news collection on turnbackhoax.com during 2017-2019. Discourse was later generated into clusters of subject idea of 'tilang', 'jalan tol', public land transport and infrastructure, 'kecelakaan pesawat', 'lowongan kerja', 'China', insecurity toward government and 'kecelakaan kapal'. The subject idea was later consulted to Indonesian Web Corpus (IWaC). Through the socio-pragmatic approach, prosody derived by contrasting process portrayed hoax on transportation as a reflection of unsolved current social problems among Indonesian community including distrust toward the government, unawareness on transport safety, law and regulation, economic status, and trend. The finding of the research is recommended to take action toward those problems to eliminate hoax spread in the transportation sector.

Keywords: web-based corpus, digital forensic, digital hoax, discourse prosody, socio-pragmatics

INTRODUCTION

Social media is a service product of advanced innovation in Information Communication of Technology (ICT), which can be considered beneficial as well as disruptive. The bipolar characteristic of social media is well said because regardless of its function to widely spread information close to real-time, its accountability is considered low since it has no editorial board and process. Thus, news spread by politics, ideology, economy, and popularity motives may lead to falsification or misleading in the form of hoax news. Newspapers as information media, for instance, by specific political and ideological purposes, can be employed and manipulated to describe a particular identity to build public perspective (Zottola, 2021).

Disruptive innovation affecting the emergence of false news always requires counter actions involving law and regulations, as in many digital and linguistics forensic media. Identifying hoax and their spread in cyberspace have become a concern of recent studies to perform this action. This identification is part of efforts to stop hoax spread through public education by looking at hoax patterns. By employing technology information systems and facilitated by algorithms, detection methods of hoax rely on identifying features of the user, content, and context. The reversed-tracking method of the articles posts on the cognitive system to successfully detects fake news possibility is up to 85%. Using digital forensic on the internet search engine, URL, and photo from the article to consult it with Indonesian regulation and law is also considered as an effective practical mechanism to confirm whether a news flow is classified as a hoax or not. Further action can also be done to identify the author of a hoax based on linguistic profile shown by bit-level based n-gram analysis on hoax discourse (Gjylbegaj, 2018; Huang & Chen, 2020; Ko et al., 2019; Kumar, 2017; Meinarni & Iswara, 2018; Peng, Choo, & Ashman, 2016; Singh et al., 2020; Sirajudeen, Azmi, & Abubakar, 2017; Tukina, Mozin, & Sanjaya, 2020; Zhang & Ghorbani,

2020).

Discourse is a substance used for every method of analysis to identify hoaxes. An analysis can also be performed using the discourse prosody approach by looking at what is written or portrayed in a text. This type of analysis facilitated by corpus study classifies interpretation of a neutral word into a binary distinction of negative versus positive or favorable versus unfavorable expression. It is known as John McHardy Sinclair's nature of the evidence, not only in collocation (Kotait, 2018; Praminatih & Nafiah, 2018) but also in internal affixes (Du, Wang, & Yang, 2018), metaphor (Indhiarti, 2019), and concordance (Erdentug & Vefalı, 2018).

The corpus-based study has been proven to provide insight in looking at semantics on certain words and transfer it into a deeper comprehension by giving prosody. The topic model of text corpora is used as a possible tool to evaluate media and classify the negative information on a socially significant topic. Thus, the media's objective, role, focus, and quality of a society can be measured (Mukhamediev et al., 2020). Finding in semantic prosody's study reveals that prosodic descriptors affect the impressions formed of others in five different circumstances for produced versus cause and utterly versus totally (Hauser & Schwarz, 2018). Corpus also defines ideology on Arabic political discourse related to Arab Israeli conflict (Sayaheen & Malkawi, 2019), which is possibly the result of one's recognition of word based on his/her inner psycholinguistics' element (Ellis, Frey, & Jalkanen, 2009) as a response toward generalized 'schema' constructed by contexts appearing in every discourse on corpora or product of delusionality, dogmatism, and religion fundamentalism, which results in wrong belief (Bronstein et al., 2019). Simply put, corpus-based prosody gives a larger picture involving a relatively broad neighboring field of study on a lexical item. In the previous research, there is false news in Romanian since the case of a wide arranged of attributes and typologies of fake news in the Romanian 2019 presidential election. It shows the contribution of social and mainstream media toward false narratives (Sutu, 2019).

Web-based corpus is an alternative online and practical way to process data using words. Data storage is classified based on country, such as ukWAC used in analyzing word CHALLENGE according to grammatical relation with its collocation environment (Lin & Chung, 2016). The research uses Indonesian Web Corpus (IWaC) to contrast words to those found on online hoax collections stored in turnbackhoax.com (TBH), particularly related to transportation issues as one of the major hoax subjects in Indonesia. By contrasting subject ideas found on TBH to IWaC and drawing prosody from each discourse, the research attempts to figure out why particular hoax news may occur in the first place by connecting it to real social problems in the Indonesian community for two years back before the presidential election 2019. The technique of comparing vocabulary using corpus has been proven to be faster than a manual technique to identify and describe spoken and written technical legal vocabulary used for specific purposes in a contract class and the general field (Bancroft-Billings, 2020).

Nevertheless, observing hoax discourse cannot be separated from language's external factors (Wijana, 2020). Interpreting meaning from what is given by context and connecting it to a social structure inside the particular community as language users can be further explained using a socio-pragmatic approach as in medical communication involving multi-party medical encounters which include patients with a range of ages. The result of the research explores family members' contribution in understanding the patient's condition to be beneficial to the treatment decision-making process (Fioramonte & Vásquez, 2019). Having this understanding in mind is likely to be beneficial for the public to look at hoax news from different points of view, from the perspective of current Indonesian society's problem. Thus, recommendations deduced from the finding of this result can be considered for future responsive actions.

Based on these explanations, the research aims to reveal what currently happens inside the Indonesian community based on types of hoax stories spread around transportation subjects. Using collections of hoaxes in transportation from turnbackhoax.com, subject ideas are generated, which later are compared to the Indonesian web-based corpus. Prosody driven from the comparison is observed and interpreted using a socio-pragmatic approach.

METHODS

Meinarni and Iswara (2018) have identified and described hoax in the research by tracking digital forensic related to the news which is found on internet search engines, URL, and photo from the article. Data is obtained online from two sources (TBH: Hoax news collection and IWaC: Indonesian web-based corpus). First, hoax news collected from turnbackhoax.com is everything related to transportation subject from 2017 to 2019, considering the Indonesian presidential election in 2019 and period of a campaign around two years earlier. Second, since hoax is usually spread using social media, discourse on hoax includes not only the title but also the body of the message and picture if available. Then, hoaxes in transportation which are divided into three different means of land, air, and sea, are categorized into clusters of the subject idea by generating 'scheme' through content and narration of each discourse. These subject ideas (in the form of word or phrase) are later consulted to IWaC to contrast its prosody. This is a qualitative study using a socio-pragmatic approach to explain findings on contrasting prosody from TBH and IWaC. Finally, finding on socio-pragmatics aspect is used to picture current social issues among Indonesian people.

RESULTS AND DISCUSSIONS

From a general outlook on hoaxes related to transportation subject, it is over 100 hoax-news found on turnbackhoax.com collection throughout 2017-2019. Transportation issues that emerged in the form of false news are categorized into three kinds of transportation means in Indonesia (air, land, and sea). These include airplanes, helicopters, cars, and motorcycles (public or private), trains, and all types of ships. On how to categorize the news is also including every aspect or facilities which may relate to or mention these means. Hoax news is collected from the source for two years, and the land transportation category dominates each year in numbers. This category exceeds air transportation by 7 points in 2019, 6 points in 2018, and 5 points in 2017. Extremely, it also surpasses sea transportation by nine times in 2019, 25 times in 2018, and 15 times in 2017. This fact makes land transportation the major spotlight on the public when it relates to hoax news compared to other transportation means. It can be seen in Figure 1.
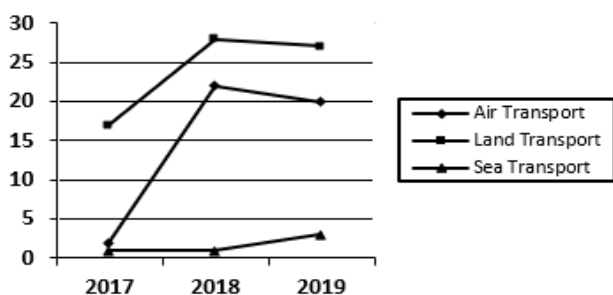

Figure 1 Number of Hoax-News on Transportation in Indonesia 2017-2019

To look at hoax news that is considered misleading, a comparison must be done to real news with similar topics. Corpus-based contrast analysis on discourse is done by generating and categorizing topics found on turnbackhoax.com (TBH), which is later analyzed in an algorithm using Indonesian Web Corpus (IndonesiaWaC/IWaC). This calculation covers all words recorded by the system and not necessarily on different headlines of news/source. By looking at text collected from the internet, certain components will be clustered and calculated based on the frequency of concordance, collocation with the first word to the left and the right (if possible), and discourse prosody. The cluster category in transportation is divided into three, which are clusters on land transport category, air transport category, and sea transport category.

The first cluster is the land transport category. An interesting finding over hoax news that occurs during 2017 to 2019 is that two later years are easier to cluster in topic due to similarities of substance than one earlier year. It seems that hoax news in 2018 and 2019 shows repetition in theme with only minor adjustments. The collection of news in 2017, on the other hand, shows more random subjects making it more difficult to classify. Table 1 shows the corpus data comparison on the word 'tilang'.

Table 1 Corpus Data Comparison on 'Tilang'

	IndonesianWaC (IWaC)	turnbackhoax.com (TBH)
Word Form	'tilang' (n=76)	'tilang' (n=10)
	'Tilang' (n=12)	'Tilang' (n=0)
	'TILANG' (n=10)	'TILANG' (n=1)
Collocation:	'surat' (n=12)	'surat' (n=5)
First word to the left	'kena' (n=4)	'kena' (n=6)

Hoax news on this category is dominated by the emergence of component 'tilang', which relates to violating traffic, including administrative regulation (2019=eight times; 2018=thirteen times; 2017=once). However, for this issue classification, the matching node word of 'tilang' and its variation are only 11 occurrences. This happens because the term 'tilang' arose not only from literal word statements but also interpreted from context. Additional terms are shown up in the form of 'razia' or police road checks (n=7) with four variants ('razia', 'Razia', 'RAZIA', and 'rajia').

As expected, the context that arose from this node word is consistent with the law with negative discourse prosody of cause and effect. However, this negative implication is materialized only for wrongdoers since there is a penalty and or fine, which are managed on the regulation for it. Therefore, whenever someone violates any traffic regulation, he/she will be given a traffic ticket. Some expanded contexts include *tilang oknum, membayar tilang, tilang pelanggaran*, and many others.

Surprisingly, collocation 'CCTV' or 'E-CCTV' for 'tilang' occurs repeatedly on TBH heading, and it cannot be found elsewhere on IWaC. The topic of implementing testing of E-CCTV to catch traffic trotters occurs almost half of the total amount of news about '*tilang*'. By year, this news spread on social media with similar content and discourse mentioning location of CCTV parameters, instruction to obey, and integrated authority, which is involved. What differentiates this hoax news is merely about location (West Jakarta, Bogor, Manado, and Depok). There are two modifications of discourse found on hoax news mentioning Malang (one of the district capitals in East Java) and a fine list for '*tilang*'. The absence of word collocation for '*tilang*' and 'E-CCTV' or 'CCTV' on IWaC can be described by clarification made related to this hoax. According to the authority, CCTV installment in several spots on the road is not done solely to monitor traffic but for other purposes. Even though E-CCTV can be used as evidence to bring

Table 2 IWaC Concordance on '*Tilang*'

Left	Kwic	Right
pemohon harus membuktikan bahwa tidak ada kewajiban-kewajiban khusus yang dibebankan pada mereka, termasuk kartu	tilang	dan pajak yang belum dibayar. </><s> Menurut laporan koresponden Washington Post Jackson Diehl: " Bagi orang-orang
, yang dalam beberapa kasus bahkan dijadikan sarana untuk kegiatan pelanggaran hukum lainny berupa pembayaran	tilang	terhadap oknum. </s><s> Ketika polisi secara kelembagaan sedang disorot karena pengakuan salah seorang oknum petingginya
: Mgr .</s><s> G. Manek, SVD Selama masa Mgr. </s><s> A. Thijssen di Mau mere didirikan satu paroki baru yaitu • Tahun 1953 Paroki	Tilang	dengan P. de Zwart, sebagai Pastor Paroki. </s><s> Pada tanggal 25 Januari 1961 Roma mendirikan Struktur Hiraikis Gereja di
. </s><s> Suku yang jumlahnya terbanyak adalah Sikka Krowe yang berdomisili pada paroki-paroki bagian tengah yaitu: Paroki	Tilang	, Paroki Bloro, Paroki Lela, Paroki Sikka, Paroki Nita, Paroki Wairpelit, Paroki Koting, Paroki Nele, Paroki Habi,
Pemantau Khusus Pilkada (PKP) Jakarta, mengusulkan agar pelaku kecurangan Pilkada bisa di adili segera seperti	tilang	ditempat bagi pelanggar lalulintas. </s><s> Jadi, terhadap pelaku kecurangan, Polri tidak perlu menunggu sampai proses
lapor dan diproses polisi. </s><s> ` ' Atas laporan itu polisi wajib segera meresponnya. </s><s> Mungkin bentuknya bisa seperti	tilang	ditempat. </s><s> Begitu melakukan pelanggaran, langsung bisa segera disidangkan. </s><s> Sebab, peradilan pasca pemilu,
Pemantau Khusus Pilkada (PKP) Jakarta, mengusulkan agar pelaku kecurangan Pilkada bisa diadili segera seperti	tilang	ditempat bagi pelanggar lalulintas. </s><s> Jadi, terhadap pelaku kecurangan, Polri tidak perlu menunggu sampai proses
lapor dan diproses polisi. </s><s> ` ' Atas laporan itu polisi wajib segera meresponnya. </s><s> Mungkin bentuknya bisa seperti	tilang	ditempat. </s><s> Begitu melakukan pelanggaran, langsung bisa segera disidangkan. </s><s> Sebab, peradilan pasca pemilu ,

traffic offenders to court sounds promising, it is still far from execution.

Different concepts of '*tilang*' have surfaced four times in TBH involving driving registration (STNK for vehicle and SIM for a driver). It is kind of '*tilang*' if someone drives or rides their vehicle without having one or both registration with him/her. STNK or vehicle registration is a legal document for vehicles along with an annual tax payment receipt, while SIM or driver's license is a document to legalize someone to drive or ride a vehicle. Being neglectful to have them while driving may cause the driver or rider to get a ticket. In IWaC, the joint between both words is mostly pictured by *mengurus, pengurusan, memperpanjang, perpanjangan,* and *motor*. All these words suggest the process of ensuring registration is set legally for a drive, especially for motorcycles (*motor*).

This process on TBH has become a basis to ignite false news. As a result of being careless in the process, a driver must face the law consequence of '*tilang*', which is done in a '*razia*' or regular police road checks. To pass from '*razia*' false news is spread including information on location and date schedule of '*razia*' to avoid being caught, how to do '*pemutihan sim*' to renew an expired driver license without passing test drive procedures as well as mob driver license making, '*pemutihan pajak kendaraan*' as pardon for late vehicle tax payment, additional agenda relates to drug raids, irrationally expensive annual vehicle tax fare, and how tax is not part of police department jurisdiction; therefore, they are not allowed to give

ticket on unpaid vehicle tax.

As seen from all collected false news on TBH on this subject, there is a tendency to give resources for the reader to avoid the law. The user of this type of information is those who have problem with their vehicle or driver registration. Instead of dealing with it procedurally, employing and sharing information from hoax news is a sign that either he/she chooses to solve the problem wrongly or he/she lacks information about the legal procedure on this particular matter. Even though there is no further explanation for this, however, one reason for those behaviors is manifested in the false news itself, which is the taxes fare issue.

Table 3 Corpus Data Comparison on 'Tol'

	IndonesianWaC (IWaC)	turnbackhoax.com (TBH)
Word Form	'*tol*' (n=1,666)	'*tol*' (n=61)
	'*Tol*' (n=285)	'*Tol*' (n=21)
	'*TOL*' (n=30)	'*TOL*' (n=6)
		'*toll*' <informal> (n=2)
Collocation:	'*jalan*' (n=922)	'*jalan*' (n=41)
First word to the left	'*Jalan*' (n=131)	'*Jalan*' (n=1)
		'*JALAN*' <head-line> (n=3)

The second lead in this hoax category is the component of the '*jalan tol*' (toll road) issue from 2017 to 2019 (three times, ten times, and six times). In concordance on node word '*tol*' only, nearly two thousand KWIC are found on IWaC, and almost a hundred on TBH. In terms of word form, there is a modification of '*toll*' that appears twice in TBH, which is interpreted as informal variation since it is not Indonesian but an English word. In word collocation, another variation shown in the word '*JALAN*' appears only on the news headline/title three times. There is no indication of this variation existing elsewhere on the news body. By comparing data collected from the Internet to all discourse found on hoax news relates to this word, intersection on collocation indicates the word 'Cipularang' (IWaC=24; TBH=1) and 'Cikampek' (IWaC=18; TBH=4). Both are known as the names of toll roads in Jakarta. Context found on IWaC about 'Cipularang', which is more frequent than 'Cikampek' is inversely compared to what is found in TBH making both words have different levels of attention on both media. Figure 2 shows the IWaC contrast concordance on Cipularang and Cikampek toll.

Looking at discourse prosody found related to both node words on IWaC, 'Cipularang' is described positively based on its benefit toward the economy, tourism, politics, education, and traffic between Jakarta-Bandung in vice versa. Additional context on this word is shown on advice to keep road layout green along with the toll. Meanwhile, the context around 'Cikampek' shows a different vibe since mentioned problems on rest area toilets and traffic jams. Additional context on this word is found mentioning detail and integrated access on this toll. Differently, discourse on both words in TBH relates many times to accidents, including pile up and land collapsing. Observing

context around toll roads in TBH, hoax news tends to exaggerate high-risked traveling by toll road access. Factors influencing these mentioned accidents are two-way traffic system changing, road maintenance, land collapsing, and extreme weather conditions. Looking up '*kecelakaan*' (traffic accident) as a supplementary word to '*tol*' in concordance on IWaC, 13 contexts are found. However, most of all accidents mentioned in the context are related to research findings on how to prevent accidents from occurring on toll roads. Extra contexts on this are on how supernatural force may be involved, on the news about minimizing accident victims in Arizona, and how someone experiences witnessing an accident. There is none about the factors mentioned on TBH.

Interestingly, regardless of the small number of occurrences (n=4) in IWaC, 'Semarang' as one of the toll road names appears four times in TBH. The number of occurrences in TBH is considered significant in percentage considering the overall number of hoax news. In concordance given in IWaC, '*tol*' and 'Semarang' collocation only give research findings as context, which is mentioned three times. The other context is about toll road project developers. In TBH, '*tol Semarang*' is linked to sensitive issues such as religion (*perobohan masjid* n=3) and culture (*sultan Jogja* n=1) as negative impacts of toll road development. According to clarification made by the authority on TBH, this is caused by misinterpretation and misleading information given toward particular matters. This kind of information is later used by irresponsible people or groups for their unknown interests.

Referring to the negative side of the toll road, other contexts recorded in TBH also mention an environmental concern, human error on toll fare, and

Figure 2 IWaC Contrast Concordance on Cipularang and Cikampek Toll

traffic jam (refers to common transportation issue). The different context in TBH also misleads information about toll fares on new or nearly operated and integrated toll roads. Misleading and misinterpreted toll fare is around free toll, newly integrated toll access, and newly operated toll fare (Ngawi-Sragen, Jakarta-Surabaya, and Soreang-Soroja).

The next category of hoax is related to public land transport and its supporting facilities. One of them relates to '*kereta api Indonesia*' (Indonesian railway) facilities. What highlighted on seven related news to this matter are officer abuse of power toward punk kids inside the train and other road users at the railroad crossing, electronic ticketing, and operating system error. Newly launched train route passing Soekarno-Hatta airport and Indonesia-China partnership impact negatively on developing train systems and purchasing. The 3716 KWIC found on IWaC gives abundant context to '*kereta api*' (word form: *kereta api, Kereta Api, Kereta api, kereta Api*). In IWaC, *kereta api* is in collocation mostly with *jalur* and *Jalur* – railway line (n=300), *stasiun* and *Stasiun* – railway station (n=248), and *rel* - railroad (n=227). In comparison to discourse on TBH, *stasiun* is the most relevant topic since it can be interpreted largely. *Stasiun* comprises a train operation system as well as regulations.

Instead of prosody on a mystical railway station and some stories on several stations history in Indonesia from Dutch government era and railroad history from all over the world which appear many times, looking at *stasiun* on IWaC, there is no intersection found on system error or on how the involvement of China affects Indonesia negatively in whatsoever. To connect Indonesian railway enhancement three times with China include the Chinese language on train operating system, Router project partnership with Chinese authorities who are engaged in corruption, and Chinese project bribing using ecstasy pill portray insecurity toward Indonesian authority for being individualist and vulnerable to corruption based on clarification made on TBH.

Following news related to infrastructure is transport access, including drainage sabotage, blocked or limited road access, and collapsed or broken bridges. Most location setting takes place around Indonesian capital city, Jakarta, and its surrounding area. All kinds of news links to this matter are mentioned, such as landfall and flood or even accident. Material of the news portrays fault in road access, which leads to traffic jams and casualties. According to clarification made on TBH, misleading facts on hoax news are exaggerated by mentioning a greater number of victims or pairing captions with wrong pictures. Limited road access is also mentioned as an alternative to unravel traffic jams as a regular problem in Jakarta. Hoax news relates to this topic is odd-even access systems for motorcycles at several roadway and Electronic Road Pricing (ERP). Even though the odd-even access system functions in several roadways, it only applies to cars and free of charge. This is the type of misleading news by substituting a regular applied

system with the nearest context.

The last classification made for hoax news on public land transport is related to the busway and online taxi or taxi-bike, and it appears only on TBH 2017 collection. It seems like the traffic problem and its alternative solution in Jakarta and other metropolis cities still become a hot issue and potential for hoax news. Trans Jakarta or Jakarta busway which is functioned to facilitate mass in public transport to reduce the number of private vehicles using is affected by the occurrence of fabricated news on job vacancy, authority abuse of power, free usage for public 212 Islamic prayers in Monas (Jakarta National Monument), and foreign car body contractor for the bus. Regardless of all positive impressions shown on 24 KWIC found in IWaC, context taken from TBH on three news inputs is considered political, which links to both sides of public figures that once fought for Jakarta governor position. Each news input tends to discriminate against one party and elevate the other one. On the other hand, online taxi or taxi-bike contexts cannot be found elsewhere in 503 KWIC of '*ojek*' in IWaC. However, TBH files negative gist of this online version of '*ojek*' when hoax news emphasizes dispute toward traditional '*ojek*' or '*angkot*' by its existence. Considering the minimal number of news records, this issue's contribution can be considered low on TBH.

The next is clusters in the air transport category. A general look at the news record on TBH for this category shows a significant gap in frequency between the later two years (2018 and 2019) and earlier year (2017). For example, with more than 20 records of news from 2018 to 2019, two archives on hoax news in 2017 are only one percent. Elements clustered in this category are airports, airlines, planes or helicopters, and other related companies.

The ten files are related to airplane accidents (involving causality) or incidents. Words used for this context are '*pesawat jatuh*' (plane crash). There is no positive implication for this situation. However, a contrast made between input on IWaC and TBH gives a different focus on public perspective. In context derived from 37 KWIC found in IWaC, '*pesawat jatuh*' gives information about plane crashes in the past, mostly abroad and only one in Indonesia. Referring to TBH, there is no intersection to IWaC on Lion Air or any other plane crash mentioned on TBH. Lion Air JT610, which crashed on the sea in 2018, becomes the subject of hoax news six times by claiming a baby photo as a survivor and pre/after crash video or photo. Such accidents attract national attention that results in regular updates on the media. Examining the timeline of hoax news appearance to the official authority statement on media, photo or video claiming is considered too close to be completely contradictory. Adding religion and political sympathy by praying for the victim and writing a poem by a closely linked public figure to shore up even more community support toward certain figure/party or even linking plane crash to mystical power to avoid a disaster which mostly is used to collect likes and comment for personal income

give the color of compassion and purpose as news modification on this matter. Claiming tragic videos or photos of plane crashes is one way to spread hoaxes by matching them with different captions and heading. Reference photos or video on TBH is fabricated from a video game, regular practice simulation, and other similar accidents from the past.

Compared to 'lowongan kerja' or job vacancy on other transport means, the air transport category stores many hoaxes making it a second major issue in this field. The only appearance on land transport means is found once in 2017 as a hoax in job vacancy for Trans Jakarta busway. The seven times hoax on this topic spread through 2017 to 2019 using the word 'lowongan kerja' with some modifications on subject ideas including 'rekrutmen' (n=1), 'recruitment' (n=1), 'Recruitment' (n=2) or 'Penerimaan Karyawan Baru' (n=1). It can be seen in Table 4.

Table 4 Corpus Data Comparison on 'lowongan kerja'

	IndonesianWaC (IWaC)	turnbackhoax.com (TBH)
Word Form	'lowongankerja' (n=268)	'LOWONGAN KERJA' (n=1)
	'LowonganKerja' (n=129)	'LOWONGAN PEKERJAAN' (n=1)
	'LOWONGAN KERJA' (n=16)	'lowongankerjanya' (n=1)
	'Lowongankerja' (n=12)	
First word to the left	'iklan' (n=29)	'DI BUKA' (n=1)
	'informasi', 'INFORMASI', 'INFO', 'Informasi', INFORMATION (n=22)	'Info' (n=1) PenerimaanKaryawanBaru

The implication derived from over 400 times 'lowongan kerja' appeared on IWaC mostly expresses general problems and recommendation on making application letters, analysis on some potential job occupations including qualification needed, few revealed discrimination facts on recruitment on the past and many job vacancy advertisements from agency or few directly from the institution. However, when matching subject idea to air transport means, there is not even one intersection to Garuda Indonesia Airways, AngkasaPura (Airport Operator), or PTDI (Aerospace Company) as mentioned on TBH. Three of these institutions are state-owned enterprises and their employees are considered civil servants with well-known supporting facilities, allowance, and fixed future assurance. Job vacancy information on these companies draws great interest as on other CPNS or civil servant job vacancy advertisement in general.

Some indicators of hoax news can be observed in the discourse, such as chained-letter as media to distribute the message, incorrect email or site address whether using slightly different or too general word and code, free email domain, and salary mentioning. Based on clarification made on TBH, when the advertisement directs the applicant to send a letter to a fabricated email address, he/she will be given an interview location that is far from his/her original place. Later, he/she will be asked to travel by plane using a selected fake travel agent. This is how a deceiver collects money illegally from those who fall on the trick. As suggested by those companies whose names are exploited by the hoax, rechecking to the official site is the best way to validate a job vacancy.

The air transport hoax category on node word 'China' is equally in second place with 'lowongan kerja' since it has a similar number of appearances, and it can be considered as a political issue based on its substance. This issue links to the handover process or planning on state-owned enterprises, including Kertajati Airport, PTDI, Nurtanio, and Garuda Indonesia Airways to China. As a result, another hoax emerges highlighting Chinese reign as seen from the crowd picture of what is claimed to be Chinese immigrant flock at Sultan Syarif Kasim, Pekanbaru airport, and China airlines direct flight operating Soekarno-Hatta Jakarta and Sultan Hasanuddin Makasar airport. Slightly pairing node word 'China' to every state-owned enterprise in Indonesian air transport means and facilities gives negative connotation on TBH. Combining pictures and some Chinese words to a provocative narration mentioning Jokowi's winning is how this hoax works. Further, with a new obtained subject idea of Jokowi, more hoaxes are found about selling Sepinggan Balikpapan airport as part of moving Indonesian capital of Jakarta and removing previous president trace by changing the inauguration inscription in Lombok airport as the result of changing Lombok International Airport name (BIL) to Zainuddin Abdul Madjid.

There are a few that can be obtained from the IWaC intersection between 'China' and jual, yet it can give insight by spotting some words of impor, murah, produk, perdagangan, and perjanjian. From these words, it is seen that China is known as a country with cheap product export trading all over the world. However, there is not even once found an exclusive partnership between Indonesia and China related to air transport means on the news. Instead, it is incorporated on bigger forums such as the ASEAN-China Working Group (Julaika, 2020) and ASEAN STOM (Dinas Perhubungan Aceh, 2020). Thus, what relates to the Indonesia-China partnership on transportation sectors is only supporting the main purpose of strengthening the economic sector on export trading, which can be found in purchasing superfast railways and maritime shafts (Amalia, 2018; Uyun, 2017; Yamin & Windymadaksa, 2017). In brief, no evidence relates to the hand-over of big companies in the transportation sector mentioned on TBH.

Being insecure toward government policy related to the partnership with China is similarly shown in the previous implicature finding on the land transport category. Corruption is linked to another hoax in the air transport category involving former president and CEO of Garuda Indonesia, caught in smuggling cases and case legalizing uncertified Indonesian pilots to fly planes. In addition, government discrimination is highlighted on immigration policy forbidding certain Islamic names to take a flight and a plane to land in Syamsudin Noor, Banjarmasin, for political reasons. On the other hand, hoax news associated with power abuse is pictured in a single passenger flight of Citilink to Surabaya, changing Soekarno-Hatta airport name to PegiPegi and Traveloka, which are online applications, canceling all flight routes from/to Palu due to the earthquake and skyrocketed the price of flight during *Lebaran* holiday.

The last cluster is the sea transport category. With a total of only five appearances throughout 2017-2019, sea transport turns to be the least sector affected by hoax news in transportation. As in two other transport means, this category also contains issues in four hoax news about accidents involving casualty. When it relates to a tragedy, the public commonly gives condolence in the form of prayer for the victims. Similarly, prosody deducted from news discourse by pairing pictures or video with fabricated caption or narration is to raise hope for the victims' family and or supporter on shipwreck finding, to wish for victims' safety on self-rescue efforts, to give false information about loss of life and material in an accident, and again to collect online like and comment for business purpose using prayer.

It seems like TBH reflects IWaC in terms of the percentage of news appearance. When other means of transport show a great number of interests in IWaC, the number of appearances in TBH is also considered great. In the case of sea transport, small numbers of appearance in TBH (n=5) are mirroring IWaC (n=24). It means whether hoax or non-hoax, only a little information can be found online regarding this matter. Table 5 shows the IWaC concordance on '*kecelakaankapal*'.

Table 5 IWaC Concordance on 'kecelakaankapal'

Left	KWIC	Right
<s> Kirimkanlah impian itu kepada Halsione, istrinya, dan	kecelakaankapal	Seiks .</s>
<s> Tudingan itu agaknya dapat dipahami, karena berdasarkan anization (IMO), 80 %	kecelakaankapal	didunia disebabkan oleh kelalaian manusia. </s>
<s> Banyaknya	kecelakaankapal	laut tersebut menginspirasi Widya Hapsari dan Rama Widya Is
<s>	Kecelakaankapal	yang terjadi di Bali, Kalimantan Barat, Aceh, kecelakaan pesawat
<s> Pada pelatihan water rescue ini para peserta diberikan simulasi penyelamatan korban di dalam air, mulai dari korban yang hanyut,	kecelakaankapal	, sampai pertolongan umtuk korban yang sudah diselamatkan kedar
<s> Ya, karena Si Penulis, Kinloch, ternyata mengalami asi saat itu, bahkan sampai dua kali	kecelakaankapal	. </s>
<s> Otto Kretschmer meninggal dunia pada tanggal 5 Agustus	kecelakaankapal	saat merayakan ulangtahun pernikahannya yang ke-50. </s>
<s>	Kecelakaankapal	laut sering terjadi dan umumnya dijelaskan dengan alasan human er
<s> Biasanya satu atau lebih jenasah akan hanyut kepantai setelah	kecelakaankapal	, tetapi ini tidak pernah terjadi di Segitiga Bermuda. </s>
<s> Berdasarkan penyelidikan para ahli, bahwa banyak terjadi kecelakaan pesawat udara, tabrakan, kebakaran ,	kecelakaankapal	laut yang menewaskan banyak orang. </s>
<s> Sebanyak 116 orang selamat dan 8 lainnya tewas dalam	kecelakaanKapal	Motor (KM) Sarimulya di perairan Sungai Barito, Tapin, Kalim
<s> Pencarian terhadap korban	kecelakaanKapal	Motor Sarimulya di perairan Desa Batalas, Kecamatan Candilaras
<s> Biasanya satu atau lebih jenasah akan hanyut kepantai setelah	kecelakaankapal	, tetapi ini tidak pernah terjadi di Segitiga Bermuda. </s>
<s> Tokoh sejarah lain yang disebut dalam buku termasuk Adolf 1951; Heinrich Himmler, yang konon sudah meninggal dalam s	kecelakaankapal	terbang misterius, mungkin karena sabotase, pada 1962; Joseph G

Table 5 IWaC Concordance on 'kecelakaankapal' (Continued)

Left	KWIC	Right
\<s> Gereja sebagai` papan penyelamatan kedua sesudah	kecelakaan \</s>\<s> kapal	\</s>
\<s> 10	KecelakaanKapal	Laut Terburuk Sepanjang Masa December 11 , 2009 — evanswina
\<s> Feng meninggal tahun 1948 dalam	kecelakaankapal	laut di Laut Hitam, saat itu ia sedang dalam perjalanan menuju Sov
\<s> Kalimat aslinya sebenarnya jauh lebih panjang sehingga kejanggalan itu tidak disadari pembuatnya: jumlah korban yang ditemuk	kecelakaankapal	penumpang itu berjumlah 356 orang. \</s>
\<s> - Korban yang ditemukan meninggal dalam	kecelakaankapal	penumpang itu berjumlah 356 orang. \</s>
\<s> - Jumlah korban yang ditemukan meninggal dalam	kecelakaankapal	penumpang itu mencapai 356 orang. \</s>
\<s> Selanjutnya 80 %	kecelakaankapal	disebabkan oleh kesalahan manusia dan kesalahan ini sebagai hasil
\<s> Kedua asuransi ini menjamin adanya pembayaran ganti-rugi hingga mencapai jumlah US $ 30 juta untuk setiap	kecelakaankapal	tangki raksasa. \</s>
\<s> Ketiga negara merasakan bahwa jumlah ini mungkin tidak akan mencukupi untuk mengganti kerugian bagi	kecelakaankapal	tangki. \</s>
\<s>	KecelakaanKapal	Motor (KM) Dumai Express 10 di perairan Tukong Hiu, Kabupa

Human error is mentioned in IWaC as a major factor causing a shipwreck along with a bad management system. A different context found is about some public figures who died in a shipwreck in the past. In contrast with TBH, even though IWaC shows several accidents related to motorboats, there is no data crossing found on location, date, and name of the ship. The IWaC is in Bali, West Kalimantan, Aceh, South Kalimatan, and Riau Islands; while TBH is in South Sulawesi, Madura, and North Sumatera. Thus, news on TBH cannot be clarified on IWaC.

Regardless of the difference in its means (land, air, or sea) and component cluster based on what appearance is dominant, hoax news on transportation depicts similarity in the topic, which can be interpreted into real problems in this sector over Indonesian society. Thus, it can be said that hoax news is one approach to elevate social conditions based on public perspective. The first social issue is predominantly related to public distrust toward Indonesia's recent government effort to develop transport means and access. Aside from all progress made to materialize national planning to overcome traffic jams and problems, political dispute drags down and questions government reliability on doing it. Bringing two parties' rivalry on presidential and regional elections during 2017 to 2019 as an object, numbers of hoax news emerge in negative prosody centralized on the capital city of Indonesia and some from other big cities in Java Island. Started from critics on the toll road, system utilization to unravel the jam, and unstandardized/malfunctioning access or transport system (bridge, road, or ship loading), discourse on hoax tries to direct public perspective on the incompetence of the government. Certain news to discriminate one side is a beneficial way to raise support of the opponent, which is considered a technique to elevate massive vote such in the issue of *inmigración* (immigration) and *inmigrante* (immigrant) presented in the main Spanish written media (Alcaraz-Mármol & Almela, 2016). Issues used in discourse are pictured in the form of corruption, including bribing, prejudice toward a particular ethnic or group, and misplaced loss-making partnerships.

The next social issue is about public less awareness on transport safety, law, and regulations. Highly rated hoax news on traffic tickets and regular police road checks presents community ignorance on the importance of driving documents. The fact that possessing these documents portrays not only obedience to regulation but also an assurance to the ownership and personal legal to drive or ride a vehicle physically and mentally. Assurance is important to help minimizing illegal vehicle trading, which usually has no legal registration and potentially comes from a stolen or problematic vehicle and traffic accident caused by the incompetent driver due to imperfect condition or age. Illegal ways to obtain these registrations to avoid fines or procedures, including tax payment and tests, seem to become the preference. Thus, offering this way helps to spread hoax news to the public.

The economic level comes next as a public social problem in transportation. Indeed, discourse on

hoax news reveals a gap in economic status among the Indonesian community. In the case of toll roads, a problem arises because this facility is usually for middle-high community users considering the rate. Low-level communities are pictured to become victims of toll construction, and development for its facility is unreached by the poor. Further, the process of toll construction, especially on the condemnation, is pictured as unjust because of the price rate or land acquisition dispute on certain public religious facilities. Another case is related to the need for job availability. Undeniably, being occupied with a job is a way to raise one's prosperity. Sharing a job vacancy by including a high salary rate is intriguing and likely to be used as a hoax news topic. Job vacancy on state-owned enterprises is still highly favored by the Indonesian community. Government bond school is a way to ensure graduates have a job there. For those who are not from these special schools, being succeeded to get this job is greatly anticipated. In desperate need of a job, whenever a rarely published job vacancy appears for a position at the airport or airlines, many will lower consciousness to do verification and shove the application directly. Instead of improving one's well-being, this unawareness may lead to criminal prey where an applicant is asked to transfer an amount of money for any reason.

Last but not least, social problems in any community are linked to the trend. Whenever a trending topic appears, seasonal hoax news can emerge in the company. Shocking news with national or international impact requiring proper closing will stay the center of attention for some time. An accident or incident in transport means, access, and system (land, air, and sea) assemble public concern, which requires quick response and handling. Great sympathy toward the victims, their families, and every involved party is easily driven to an exaggerated narration. Usually, the nature of evidence leads to a positive sense to console the loss of grieving victims and family in the form of prayer or even negative one in chunks of individual experience preceding the unfortunate event. Achievement to catch in hand smuggling cases of an airlines' official or everything surfaced from one candidate of the party competing in regional or presidential election attracts every information connected to the subject. Therefore, sometimes hoax news rises circumstantially. Once again, the semantic association added by socio-pragmatic aspects is proven to be practical in identifying and describing phenomena pictured in a discourse. Since discourse prosody is recommended to be applied even in learning language (Omidian & Siyanova-Chanturia, 2020), it is not excessively to resolve that it is also suggested in learning more about hoax news and its link to the actual social issue in this country.

CONCLUSIONS

Digital information is beneficial since it is stored in the mass collection through time in the cloud, making it easier to verify with almost real-time updating. To clarify a chain message spread using social media, it is easier for a receiver to consult the internet as an online source. Once clarification is made by a receiver, hoax news spreading can be stopped right away.

When compared to general online sources on the internet, Hoax discourse in transportation can be generated and interpreted. Dominated by hoax on land transport means, which majorly is focused in Indonesian capital city, news on this category is clustered into fabricated news on tilang, jalan tol, and public land transport and infrastructure. Taking the second lead in number, the hoax on air transport category is colored by cluster kecelakaan pesawat, lowongan kerja, China, and insecurity toward government. Finally, with only five numbers of news, hoax on sea transportation highlights some shipwreck news in kecelakaan kapal. Those results on hoax clustering reflect real social issues among the Indonesian community mostly in negative prosody explained by corpus evident on distrust toward Indonesian government, unawareness on transport safety, law and regulation, economic status, and trend.

Indeed, hoaxes in transportation are a reflection of unsolved social problems surrounding Indonesian people, as shown by linguistic features in its discourse. Paying attention to the real problem is presumed to be one way to eliminate hoax news spread related to it. Social issues found in the research can be used as recommendations for further action of those who are involved. The research data are collected from two different sources (real news and hoax news). However, both media are limited by online collection only. Many other collections may be existing in printed (offline) versions. By considering these remaining untouched media, the findings of the research may result differently. Therefore, further observation involving these sources is possible be done in the future.

REFERENCES

Alcaraz-Mármol, G., & Almela, J. S. (2016). The semantic prosody of the words inmigración and inmigrante in the Spanish written media: A corpus-based study of two national newspapers. *Revista Signos, 49*(91), 145-167. https://doi.org/10.4067/S0718-09342016000200001.

Amalia, R. (2018). Kerja sama maritim Indonesia-Tiongkok menuju visi Indonesia sebagai poros maritim dunia. *E-Jurnal lmu Hubungan Internasional, 6*(3), 1271-1280.

Bancroft-Billings, S. (2020). Identifying spoken technical legal vocabulary in a law school classroom. *English for Specific Purposes, 60*, 9-25. https://doi.org/10.1016/j.esp.2020.04.003.

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G.,

& Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition, 8*(1), 108-117. https://doi.org/10.1016/j.jarmac.2018.09.005.

Dinas Perhubungan Aceh. (2020). *Usulan Kemenhub jadi pemantik perkuat kerja sama transportasi negara ASEAN.* Retrieved from https://dishub.acehprov.go.id/informasi/usulan-kemenhub-jadi-pemantik-perkuat-kerja-sama-transportasi-negara-asean/.

Du, Y., Wang, B., & Yang, L. (2018). A corpus-based study on the semantic prosody of quasi-affix 'Zu'. *Lecture Notes in Computer Science, 11173*, 445-453. https://doi.org/10.1007/978-3-030-04015-4_37.

Ellis, N. C., Frey, E., & Jalkanen, I. (2009). The psycholinguistic reality of collocation and semantic prosody. In U. Römer & R. Schulze (Eds.), *Exploring the Lexis-Grammar Interface* (pp. 89–115). Amsterdam: John Benjamins Publishing.

Erdentug, F., & Vefalı, G. M. (2018). What is 'old' and 'past' in new age discourse? A qualitative analysis of corpus evidence. *Discourse, Context, and Media, 24*, 85-91. https://doi.org/10.1016/j.dcm.2018.01.009.

Fioramonte, A., & Vásquez, C. (2019). Multi-party talk in the medical encounter: Socio-pragmatic functions of family members' contributions in the treatment advice phase. *Journal of Pragmatics, 139*, 132-145. https://doi.org/10.1016/j.pragma.2018.11.001.

Gjylbegaj, V. (2018). Fake news in the age of social media. *IJASOS - International E-Journal of Advances in Social Sciences, 4*(11), 383-391. https://doi.org/10.18769/ijasos.455663.

Hauser, D. J., & Schwarz, N. (2018). How seemingly innocuous words can bias judgment: Semantic prosody and impression formation. *Journal of Experimental Social Psychology, 75*, 11-18. https://doi.org/10.1016/j.jesp.2017.10.012.

Huang, Y. F., & Chen, P. H. (2020). Fake news detection using an ensemble learning model based on self-adaptive harmony search algorithms. *Expert Systems with Applications, 159,* 113584. https://doi.org/10.1016/j.eswa.2020.113584.

Indhiarti, T. R. (2019). A corpus-driven discourse analysis of metaphor in climate change: A preliminary study. *ICEL 2019: First International Conference on Advances in Education*. Malang, Indonesia. https://doi.org/10.4108/eai.23-3-2019.2284962.

Julaika, H. (2020). *Kemenhub tawarkan proyek transportasi IKN ke Tiongkok dan ASEAN*. Retrieved from https://mediaindonesia.com/ekonomi/329050/kemenhub-tawarkan-proyek-transportasi-ikn-ke-tiongkok-dan-asean.

Ko, H., Hong, J. Y., Kim, S., Mesicek, L., & Na, I. S. (2019). Human-machine interaction: A case study on fake news detection using a backtracking based on a cognitive system. *Cognitive Systems Research, 55,* 77-81. https://doi.org/10.1016/j.cogsys.2018.12.018.

Kotait, R. (2018). On translating semantic prosody: A corpus-based cognitive-semantic approach. *Proceedings of the 1st International Conference of the Department of English in Literature, Linguistics and Translation – Travelling Theories: Origins and Manifestations.* Cairo, Egypt. pp. 1-59.

Kumar, S. (2017). Deciding metrics for detecting false news and influence on Twitter. *International Journal of Research Publications in Engineering and Technology (IJRPET), 3*(7), 96-100. https://doi.org/10.5281/ZENODO.1420905.

Lin, Y. Y., & Chung, S. F. (2016). A corpus-based on the semantic prosody of Challenge. *Taiwan Journal of TESOL, 13*(2), 99-146.

Meinarni, N. P. S., & Iswara, I. B. A. I. (2018). Hoax and its mechanism in Indonesia. *Proceedings of the International Conference of Communication Science Research (ICCSR 2018)*. Surabaya, Indonesia. pp. 183-186. https://doi.org/10.2991/iccsr-18.2018.39.

Mukhamediev, R. I., Yakunin, K., Mussabayev, R., Buldybayev, T., Kuchin, Y., Murzakhmetov, S., & Yelis, M. (2020). Classification of negative information on socially significant topics in mass media. *Symmetry, 12*(12), 1-23. https://doi.org/10.3390/sym12121945.

Omidian, T., & Siyanova-Chanturia, A. (2020). Semantic prosody revisited: Implications for language learning. *TESOL Quarterly, 54*(2), 512-524. https://doi.org/10.1002/tesq.557.

Peng, J., Choo, K. K. R., & Ashman, H. (2016). Bit-level n-gram based forensic authorship analysis on social media: Identifying individuals from linguistic profiles. *Journal of Network and Computer Applications, 70*, 171-182. https://doi.org/10.1016/j.jnca.2016.04.001.

Praminatih, G. A., & Nafiah, H. (2018). [Woman]'s world portrayed in literary works of Jane Austen. *Lingua Cultura, 12*(1), 77-82. https://doi.org/10.21512/lc.v12i1.4040.

Sayaheen, B., & Malkawi, M. N. (2019). Understanding the nature of Arab-Israeli conflict: A corpus-based study of semantic prosody in Arabic political discourse. *Lebende Sprachen, 64*(2), 309-322. https://doi.org/10.1515/les-2019-0017.

Singh, M., Bhatt, M. W., Bedi, H. S., & Mishra, U. (2020). Performance of Bernoulli's naive bayes classifier in the detection of fake news. *Materials Today: Proceedings.* https://doi.org/10.1016/j.matpr.2020.10.896.

Sirajudeen, S. M., Azmi, N. F. A., & Abubakar, A. I. (2017). Online fake news detection algorithm. *Journal of Theoretical and Applied Information Technology, 95*(17), 4114-4122.

Sutu, R. M. (2019). Fake news, from social media to television case study of the Romanian presidential elections 2019. *Styles of Communication, 11*(2), 81-92.

Tukina, Mozin, A. R., & Sanjaya, M. (2020). Disruptive innovation: A case of solving hoax information in Indonesia. *Humaniora, 11*(1), 7-11. https://doi.org/10.21512/humaniora.v11i1.6088.

Uyun, F. (2017). Kondisi hubungan triangular Indonesia-Tiongkok-Jepang setelah terbentuknya konsorsium PT KCIC. *Analisis Hubungan Internasional, 6*(2),

57-64.

Wijana, I. D. P. (2020). Pragmatik sebagai penangkal hoaks dan peran ilmuwan bahasa di era digital. *Jala Bahasa: Jurnal Ilmiah Kebahasaan, 16*(2), 171-178. https://doi.org/10.36567/jalabahasa.v16i2.653.

Yamin, M., & Windymadaksa, S. (2017). Pembangunan kereta cepat Jakarta-Bandung sebagai mercusuar hubungan Indonesia-Tiongkok. *Jurnal Politik Profetik, 5*(2), 200-218. https://doi.org/10.24252/PROFETIK.V5I2A4.

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing and Management, 57*(2), 102025. https://doi.org/10.1016/j.ipm.2019.03.004.

Zottola, A. (2021). *Transgender identities in the press: A corpus-based discourse Analysis.* London: Bloomsbury Publishing.