LLT Journal, e-ISSN 2579-9533, p-ISSN 1410-7201, Vol. 27, No. 1, April 2024, pp. 230-251

LLT Journal: A Journal on Language and Language Teaching
http://e-journal.usd.ac.id/index.php/LLT
Sanata Dharma University, Yogyakarta, Indonesia

HATE SPEECH ON SOCIAL MEDIA: INDONESIAN NETIZENS' HATE COMMENTS OF PRESIDENTIAL TALK SHOWS ON YOUTUBE

Ismail Tahir[1*] and Muhammad Gana Fajar Ramadhan[2]
[1]Beihang University, Beijing, China
[1]University of Bina Mandiri Gorontalo, Indonesia
[2]China University of Political Science and Law, Beijing, China
ismailtahir@ubmg.ac.id[1] and hanhafan@gmail.com[2]
*correspondence: ismailtahir@ubmg.ac.id
https://doi.org/10.24071/llt.v27i1.8180
received 13 February 2024; accepted 5 April 2024

Abstract
Hate comments on digital platforms, particularly social media, pose a significant threat to online investigations and present a complex issue for linguistic culture. Previous studies have mainly focused on the analysis of moderation strategies regarding this type of language use on the internet. In contrast, little attention has been given to identifying which hate comments on the internet are considered problematic. The current study investigates the phenomena of Indonesian netizens' hate comments regarding the presidential election in 2024 on YouTube talk shows uploaded in 2023. By analyzing 315 datasets using Nvivo software, it shows that early warning (43%) is mostly employed by Indonesian netizens, followed by dehumanization and demonization (21%), violence and incitement (19%), and offensive language (17%) respectively. Moreover, it is found that anonymity and personality traits are the factors contributing to hate comments related to the 2024 presidential election. The factors influence Indonesian netizens in early-warning hate comments. The implications of the present study highlight the importance of examining the online behavior and language usage in online communication among Indonesian netizens. This understanding can help in mitigating hate comments on online platforms and in society as a whole.

Keywords: hate comments, Indonesian netizens, online platforms, talk shows, YouTube

Introduction
Hate speeches or comments on social media are seen as a complex topic, and social media users frequently post hateful remarks in the comment section of social media platforms (Poletto et al., 2021). As the notion refers to any expression in texts, images, or videos used to offend, provoke, or attack a person or group on social media (Guo & Johnson, 2020), hate speech on online platforms has been the subject of decades' worth of research, leading to extensive research efforts aimed at addressing and understanding this issue. The prevalence of hate speech on social media platforms has highlighted the importance for effective detection methods. Al-Hassan and Al-Dossari (2019) conducted a study in the form of a multilingual

This work is licensed under CC BY-SA.
Creative Commons Attribution-ShareAlike 4.0 International License

corpus. In their study, using a machine learning model tailored to a specific language and region proved to be a significant approach for detecting hate speech. It presented different forms of anti-social behaviors, including cyberbullying, abusive and offensive language, radicalization, and hate speech. The study also explored some challenges that can be a guide for the implementation of the hate speech model. Therefore, it was possible to determine the statements that might cause hate speech and impolite words in a social media text (Esau, 2021; Siahaan et al., 2019).

Impoliteness and hate speech are related phenomena because of their role in attracting the attention of scholars, primarily due to their impact on social interactions. Impoliteness is a language behavior that forces the hearer to judge what the speaker does and says in terms of what the hearer considers acceptable in a conversational context (Culpeper et al., 2017). The examples of impoliteness are violations of the rules of courtesy or the use of mocking and sarcasm (Bitonti et al., 2023). Conversely, hate speech is characterized by discriminatory expressions, hate, and incitement to violence. In this regard, impoliteness and hate speech differ phenomenologically and conceptually (Niebuhr & Neitsch, 2022). The strategies of impoliteness have been investigated by many researchers in various areas, such as politics, health, and entertainment (Ardila, 2019; Han, 2021; Sinkeviciute, 2018; Vladimirou & House, 2018). In this respect, studies on impoliteness have highlighted the complexity of human interactions across diverse contexts and the significance of understanding the intentions and effects of impoliteness on various social behaviors. Thus, using hate comments as a commenting strategy on social media users is part of folk categorization systems in the netizens' comment behaviors (Shinta et al., 2018). It is because the use of impoliteness strategies on social media platforms may carry the complex dynamics of online interactions that highlight how relative anonymity and lack of physical presence can encourage more disrespectful behaviors than might be exhibited in face-to-face interactions. Furthermore, impoliteness is claimed to be an essential category to be investigated due to its complex and crucial consequences on the interpersonal communication and society that are linked to the hate speech in online communication.

The issue of hate speech on digital platforms, especially on social media, has been studied by scholars. Wilhelm and Joeckel (2019) investigated hate speech in online users' comments, focusing on flagging hate comments among women and sexual minorities. They reported that gender stereotypes influenced online interactions in which women were more concerned about fairness and avoiding harm to others than men were. Paasch-Colberg et al. (2021) offered a modularized approach to distinguish between various types of hate speech and offensive language. It experimentally mapped the range of abusive language in the area, going beyond the typical "hate/no-hate" dichotomy observed in prior studies, spanning from insults to call for hate crimes, and focusing on the ambiguity between hate speech and offensive language. Guo and Johnson (2020) also conducted an experiment study to investigate several types of hate speech among university students. They reported that students tended to think about the influence of hate speech on others was greater than themselves. Their perception of such messages' effect on themselves was a significant indicator of supportive attitudes toward hate speech. More recently, Paasch-Colberg and Strippel (2022) identified the categories of hate comments and how these categories were moderated practically in different

moderation. Their findings showed strong agreement regarding extreme cases of hate comments, whereby there was overlap with the concept of hate speech theoretically, but also forms of incivility.

Therefore, it may be claimed that previous researchers have shown an interest in studying hate speech on social media, as it is a complex issue that requires interdisciplinary research efforts to develop an understanding of hate speech dynamics and its impact on society. Hate speech in the form of hate comments on social media, nevertheless, has received little attention in the Indonesian context, primarily focusing on such political discourse aspects. It is crucial to investigate hate speech in the Indonesian context due to the country's complex socio-cultural landscape, its rapidly evolving digital environment, and the dynamics of its political scene. On the other hand, failure to address hate speech, particularly in the Indonesian context, could lead to increased misunderstandings and even violent clashes between various community groups on social media, undermining national unity and social cohesion. Moreover, unaddressed hate speech in the Indonesian context, where democracy is still evolving, could undermine trust in the context of political processes that can contribute to political instability.

In addition to political context, presidential talk shows play an important role in shaping public opinion and policy decisions, as they are closely related to greater social issues. They often address and provide a platform for discussing and addressing key societal challenges. Wrenn and Reed (2019) argued that presidential talks influence public discourse and opinion on important social and political matters that can impact the relationship between the government and the public and the government's decision-making process (Kilian, 2021). Thus, studies of hate speech in a political context should be elaborated more comprehensively, as the current study aims to provide insights into the types of hate comments and the factors employed by Indonesian netizens on presidential talk shows on YouTube. By addressing the types and factors of hate comments, this study aims to provide Indonesia's political context from a linguistic perspective that offers not only into the language itself but also into broader social, cultural, and Indonesian political dynamics, particularly on social media. It also aims to provide insights into the language practices that drive political movements, adding to the understanding of the role of language on social media in shaping political landscapes. Therefore, the research objectives of the current study are:

1) What are the types of hate comments employed by Indonesian netizens regarding the presidential talk shows on YouTube?
2) What are the factors affecting Indonesian netizens' hate comments regarding the presidential talk shows on YouTube?

Literature Review

Online hate speech role in shaping public perception

Online hate speech has become a significant concern on social media. Its dissemination in communicative spaces represents one of the major societal challenges of digitalization. According to a model of moderation variables, it is postulated that factors at various analytical levels influence comment moderators' perceptions of hate speech, which in turn influences the decisions they make regarding moderation to fight abusive content (Kunst et al., 2021). For instance, it is anticipated that legal definitions at the level of social institutions would affect the

kind of content comment moderators deem objectionable and how they respond to such comments. On the other hand, recent discussions about the advantages of diversity in newsrooms highlight how a journalist's personal history may also have an impact on their decisions regarding moderation. At each level, several ideas and expressions are pertinent to the management of objectionable material in user comments. This is because they are based on various (theoretical) presumptions and come from a variety of contexts.

Most definitions of hate speech state that it is directed at particular groups based on particular traits. The traits that characterize a group, however, might differ greatly. Some definitions emphasize traits like gender, and others emphasize that hate speech can target any group. It is because online users frequently direct hate speech at oppressed social groups like women, immigrants, and sexual minorities (Wilhelm & Joeckel, 2019); along with the general seriousness of such violations, social media users' individual traits affect how hateful remarks are judged. Online hate speech is a growing concern due to the increasing social media content (MacAvaney et al., 2019). The willingness of people to express their opinions online has contributed to the propagation of hate speech (Fortuna & Nunes, 2018). Witnessing and perpetrating online hate speech among adolescents is associated with factors such as moral disengagement and empathy. The spread of hate speech is a challenge that requires effective detection and solutions (Chiril et al., 2022).

In their user guidelines, news organizations and social networking sites like Facebook, Google, and Twitter define communication norms and parameters for the content they accept, whereas social media has been considered an indispensable and dominant means of communication (Nwozor et al., 2022). As a result, they reserve the right to remove comments that do not adhere to these rules, such as hateful ones, given the enormous growth of user-generated material in online forums and social media. Incorporating users to recognize expressions that go against societal standards is a workable strategy for platforms. Therefore, practically every social media network has some sort of reporting or "flagging" feature. Using these features, users can alert platform administrators to content they believe violates social norms. Options range from buttons that merely allow users to "report a post" to mechanisms that allow users to specify why the content was reported (like Facebook). In addition, efforts to curb online hate speech include content moderation by governments and social media companies (Hangartner et al., 2021). Cross-national studies have examined the commonalities and differences in online hate speech content, exposure, and emotional reactions (Reichelmann et al., 2021). Analyzing the targets of hate in online social media can provide a broader understanding of the phenomenon and guide prevention and detection approaches (Silva et al., 2016). Hate speech prevalence among adolescents has also been studied, with a significant percentage of students witnessing hate speech both in school (Taradhita & Putra, 2021) and online (Castellanos et al., 2023), particularly an online discourse (Aporbo, 2023).

Moreover, by reporting unfavorable remarks on refugees, Wilhelm et al. (2020) examined how comment traits and neutralization strategies in potentially hateful comments affected users' attribution of deviance. They concentrated on subtlety, type of victim, norm violation type, and justifications as remark characteristics. In addition, they looked at how descriptive and injunctive norms affected how people reported suspected hate speech. In addition to serving the

technological purpose of alerting platform providers to potentially problematic content, reporting (or flagging) user comments also serves as a barometer for what is deemed appropriate and what is not. If a user flags a comment, it can be assumed that they do not agree with it and, more importantly, that the comment is so contrary to their beliefs that action needs to be taken to have it removed since handling and perception of hate comments could be linked to explanatory factors at the individual and professional routine levels (Paasch-Colberg & Strippel, 2022). Reporting as such might be viewed as an audible proclamation of disagreement with a claim or remark. In addition, hate speech is an example of deviant communication when it comes to the context of digital media. Deviance implies that offensive behaviors change depending on how harmful people consider them to be, and people evaluate them differently depending on who they are intended for. Therefore, in social group contexts or across society, hate speech statements are seen as aberrant because they go against established cultural standards, laws, or conventions of social interaction. These remarks include those that spread animosity and intolerance, encourage prejudice and hostility, and, in extreme circumstances, may even incite violent behavior.

(Im)politeness of hate speech on digital platforms

The (im)politeness of hate speech on digital platforms represents significant challenges to online communication norms, underscoring the need for a multifaceted way to foster constructive discourse online. As the current study is founded on the theory of politeness, politeness theory was first introduced by Brown and Levinson (1987). To demonstrate "politeness" in a broad sense, they propose the concept of "face" in the theoretical portion of their work. The philosophy of Grice (1975), who established politeness through the four cooperative principal maxims in "logic and politeness," serves as the foundation for the politeness theory. These maxims are veracity and supporting proof; minimum of quantity—information should be provided concisely or sufficiently long to ensure no detail is missed; relevance maxim: one should constantly be pertinent and always aim to be concise, well-organized, and free of confusion and ambiguity. When it comes to acting politely toward others, the cooperative principle makes the assumption that most human interaction is cooperative. The participants in a discourse can see one other's courteous cues. Because each speech act has meaning, the speech acts collectively have significance. According to Searle (1969), "the chief motivation—though not the only motivation—for using these indirect forms of politeness" is the indirectness of speech. Four politeness strategies—positive, negative, going off the record, and not performing a face-threatening act (FTA)—are suggested by Brown and Levinson and can be adopted by the speaker.

Regarding (im)politeness in the digital era, online comments on social media are increasingly characterized by impoliteness, incivility, and hate speech since hate speech is considered an expression stimulating people from certain social groups (Subyantoro et al., 2019). Moreover, name-calling, insults, exaggeration, language that indicate non-cooperation, derogatory terms, vulgarity, and sarcasm are examples of impolite messaging because impoliteness tends to show people’s contempt and dissatisfaction (Zhao, 2022). While incivility also refers to a "set of behaviors that threaten democracy, deny people their personal freedoms, and stereotype social groups," impoliteness refers to generally impolite methods of

acting and communicating. Impoliteness is language-based conduct that the hearer judges based on how the speaker's actions and words are interpreted by the hearer in relation to the context of the conversation (Culpeper, 2011). In addition, the common definition of impoliteness is "behavior that is face-aggravating in a particular context" (Bousfield & Locher, 2008).

It is commonly understood that the face is dyadic, formed, and maintained through social interaction. In order to preserve each other's faces (Brown & Levinson, 1987), the interactants must cooperate. Otherwise, an attack on someone's face would inevitably be met with retaliation, and the target of the attack would feel justified in launching a counterattack. This could be interpreted as a "reciprocity" of rudeness between parties (Culpeper, 2011). Such communication takes place over the course of lengthy discourse rather than occurring in discrete statements. In addition, the same fundamental set of impoliteness output strategies has been used in numerous investigations, suggesting that Culpeper's (2011) recommendations have stood the test of time. Impoliteness strategies are a well-known hypothesis in the field of linguistics that researchers or linguists frequently investigate in a variety of societal contexts. According to Culpeper (2011), impoliteness is theoretically defined as language or actions negatively regarded in a given situation. Therefore, the (im)politeness concept is dynamic and necessitates a "process-oriented view of conversation."

Impoliteness strategies are discussed in terms of the change in the way of communication in the linguistic aspect. Regarding the language used on social media, internet users are considered to participate actively in online communication. Their online interactions are unpredictable and can sometimes be polite or impolite (Rabab'ah & Alali, 2020). Unpolished remarks of language users on social media are insulting, superfluous (since they do not add anything to the conversation), and directed towards discussion forums, participants, or issues. Moreover, comments that advocate "incitement to harm (particularly, discrimination, hostility, or violence) based upon the target's being identified with a certain social or demographic group" are often considered hate speech and go beyond being rude and incivility. Social deviance is a broad concept that includes all activities, from minor norm violations to breaking the law. Deviant behavior is defined as being inconsistent with accepted societal standards and values. Additionally, this distinction is seen as menacing and aggressive. Currently, digital communication, including traditional media (online newspapers, televised interactions), contemporary social media, and post-performance interviews, have complemented (im)politeness analysis (Reiter, 2021).

Several studies have investigated the impoliteness strategies of digital media, particularly on social media. Anwar (2019) predicted that speaking obscene, hurtful, sarcastic, and frequently offensive Indonesian on Facebook is a reality since it explored the impoliteness phenomena of the Indonesian language on social media. Andersson (2021) examined impoliteness and homophily value in the context of YouTube videos. The analysis reveals a startling degree of homophily as the postings exhibit consistent patterns of the face and social rights attacks that share the same viewpoint. Consequently, it also provides insight into how the phenomena help to consolidate and standardize views through social comparison, even though impoliteness has been acknowledged as pervasive in social media for reasons such as anonymity and social detachment. Impoliteness in ideological debates on

YouTube could operate as the adhesive for impromptu social interactions amongst like-minded people, ultimately leading to social identification in relevant groups and the emergence of homophilous online communities. Vladimirou and House (2018) also explored the impoliteness strategies in globalized social media, particularly Twitter. The results indicated that participants used a variety of analogies and juxtapositions while primarily utilizing locally relevant resources, such as mixed Greek-English script and intertextual cultural references. The communal, artistic parody of Tsipras' public character served as a means of amusement, neighborhood cohesion, and political criticism.

Hate comments on social media

Hate comments on social media related to political discourse are a pervasive issue that reflects the broader polarization and tension in political climates worldwide. It is important to address the issue of hate comments on social media by considering the types and factors contributing to such behavior's prevalence in political discourse (Wilhelm et al., 2020). Individuals are more likely to engage in negative behavior when they share similar beliefs or attitudes in an online group or community, particularly on social media. Individuals are allowed to express and employ hateful sentiments that can be generated positively and negatively on online communication platforms. Elfrida and Pasaribu (2023) proposed three main types of hate comments on social media, including early warning, dehumanization, and violence and incitement. These strategies are important to address in the present study since they provide the current situation of online hate comments among Indonesian netizens.

In relation to online hate comment issues, individuals tend to feel more emboldened to engage in hateful behavior since the viral nature of social media can amplify hateful messages that contribute to the rapid escalation of online hate speech among individuals, professional groups, and the organizations (Paasch-Colberg & Strippel, 2022). This rapid escalation leads to highlighting the transmission of traumatic imagery that dehumanizes racially marginalized individuals on social media platforms. In this situation, individuals are more likely to dehumanize others, which can lead them to contempt and unfavorable attitudes (Abdalla et al., 2021).

Furthermore, hateful comments attract a great number of politically affiliated people (Erjavec, 2014) on social media and can undermine the credibility and trust in early messages, where individuals tend to decrease public engagement and compliance with safety instructions. This situation can perpetuate discrimination and prejudice, especially towards marginalized groups who may already face barriers to accessing timely information on social media platforms. This early message or warning can easily be spread in online environments since hate speech on social media can spread misinformation and create panic during situations that negatively impact cognitive notions of information, beliefs, and knowledge (Đorđević, 2020).

Hate comments on social media can also have detrimental effects on individuals and society as a whole, particularly in the exposure to violent content. Online hate speech, which is in the form of hate comments, can contribute to a culture of intolerance and normalizing violence in which exposure to online violent content can lead to increased aggression and desensitization to violence among

individuals. This process also allows individuals to use offensive language that can significantly negatively impact themselves and others (Paasch-Colberg et al., 2021). Exposure to offensive language on social media can lead to negative psychological outcomes that can create a hostile online environment in various aspects, such as religion, culture, and language issues (Chekol et al., 2023). Thus, hate comments on social media are serious issues to be addressed that can have far-reaching consequences, as the current study highlights and presents the use of hate comments and their factors on social media.

Method

Data collection

To answer the research questions, the researchers collected data from Indonesian netizens' comments regarding the recorded talk shows uploaded on YouTube. The instances of hate comments in the current study were obtained from this accessible domain on YouTube. This platform under consideration shares the feature of allowing users to comment as part of their participatory engagement. The data were gathered and de-identified in compliance with the Association of Internet Researchers' ethical guidelines. This study aims to investigate the hateful comments employed by Indonesian netizens on presidential talk shows on YouTube. "Mata Najwa" talk shows were used to collect data on Indonesian netizens' hate comments on the presidential election 2024.

The data comprised a three-month period (August-October 2023) following the search for "Mata Najwa" talk shows. It has been confirmed that there are three candidates for the Indonesian presidential election in 2024. The exclusive talk shows among these candidates were conducted in 2023 before the election, which was conducted in 2024. Three videos were collected, and each video consisted of almost one and a half hours as a result. The average number of views is from two to three million, and each video has more than ten thousand comments. As of October 2023, there was a total of 7.7 million views and 46.098 comments in the collected material (average ratio: 15.366 comments per video). To obtain the data, the researchers manually collected Indonesian netizens' comments on each talk show by watching the full content of the talk shows to understand the given topic discussed in each talk show session. Then, the researchers read comments written by Indonesians. Most of the comments are written in Indonesian language. The researchers obtained 315 datasets that met the current study criteria. These criteria include linguistic characteristics of hate speech, including the use of ambiguous language, incitement, and taboo words. To define these criteria, the use of ambiguous language in the current study refers to abusive or offensive language that leads to ambiguity employed by Indonesian netizens regarding the presidential talk shows on YouTube. Incitement is defined as a language that encourages or provokes violence, discrimination, or hatred against presidential candidates, including their race, ethnicity, religion, gender, or other identifiable traits. Taboo words in the current study refer to the use of language that is considered inappropriate or offensive and can contribute to the radicalization of online comments and associate with negativity. Therefore, these 315 datasets were manually typed into a document format and ready to be imported to NVivo software for analysis.

Data analysis

The starting point of the analysis was assigning the hate comments among Indonesian netizens on YouTube regarding the presidential election in 2024. The processes for producing codes are outlined below using the NVivo software. The following description and classification of hate comments on the presidential election in 2024 published on YouTube in 2023 must first be understood and known by the researchers who acted as a coder. The researchers then coded all information related to the research topic. In this scenario, each piece of collected data was independently coded by entering it into the NVivo program.

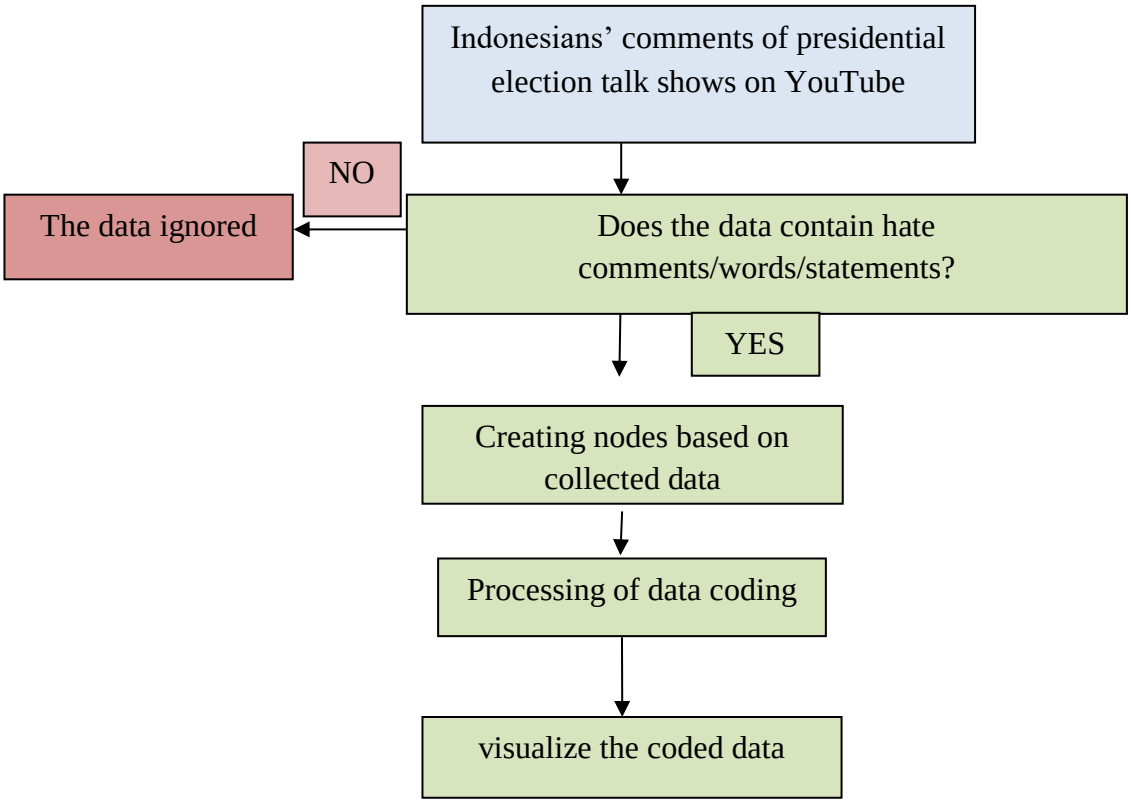


Figure 1. Steps of data analysis of hate comments in the current study

Furthermore, the data coding is illustrated as follows. First, the researchers created the data segments in the form of nodes. Second, the researchers named each node and described it in relation to the types of hate comments and their factors. Once all the data were coded, the coded data were customized for visualization by selecting the settings for the type of data visualization. Finally, the analyzed data were ready to be reported by interpreting their visualization to draw insights into the research data.

In addition, adopting Elfrida and Pasaribu's (2023) work on hate speech on social media has been useful for supporting the "coding" process of the data into the "types and factors of hate comments employed by Indonesian netizens." Thus, the process of coding was divided into four types of hate comments and two types of their factors. These types and factors are briefly defined in the working definitions below.

Table 1. The working definition of types and factors of hate comments

Types of Hate Comments	**Working Definition**
Dehumanization and demonization	It is defined as a form of hate speech that has the purpose of bringing down, criticizing, humiliating, or harassing someone. It is usually done physically: skin color, ethnicity, race, religion, equating a person with an animal, and so on.
Early warning	Early warning in the present study refers to comments containing offensive or derogatory language aimed at warning or threatening individuals or groups based on race, religion, or other characteristics.
Violence and incitement	Violence and incitement comments in the current study involve explicit calls for violence or incitement of harm towards individuals or groups. These comments go beyond mere insults and actively encourage or promote violent actions. Incitement in the current study is an act such as speech that is considered to cause riots regarding the talk shows.
Offensive language	In the current study, offensive language includes offensive slurs, derogatory terms, and discriminatory language targeting the presidential candidates of Indonesia for 2024.
Factors	**Working Definition**
Anonymity	It refers to the state of being unidentified while engaging in hate comments on YouTube regarding the presidential candidate's talk shows.
Personality Traits	It refers to the utterance of Indonesian netizens' thoughts, feelings, and behavior towards presidential talk shows on YouTube.

In summary, this part employs the methods of collecting data, including the data criteria of hate comments/speech on recorded talk shows of the presidential election in 2024 that were uploaded in 2023, data analysis, and constructing and presenting the research results. These results are elaborated on in the findings section.

Findings and Discussion

Findings

In this section, the answers to two research questions are provided. The findings concerning RQ1 provide the categories of hate comments employed by Indonesian netizens on presidential talk shows. Moreover, the findings of RQ2 present the factors affecting the hate comments employed by Indonesian netizens on YouTube.

Hate comments categories employed by Indonesian netizens on presidential talk shows

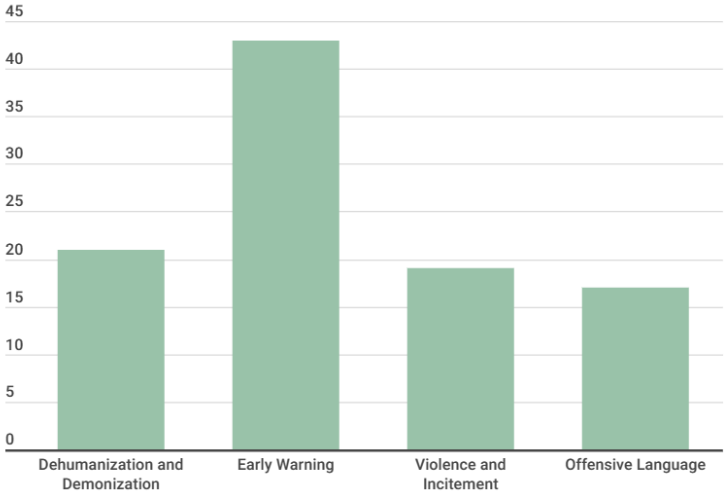


Figure 2. Hate comments types among Indonesians regarding the presidential election

As illustrated, the study's findings showed that Indonesian netizens employed four types of hate comments regarding the presidential election in 2024. Early warning (43%) was the most frequently employed by Indonesian netizens, followed by dehumanization and demonization (21%), violence and incitement (19%), and offensive language (17%). These findings are illustrated in sub-sections below.

Dehumanization and demonization

The context of dehumanization and demonization in the current study was related to the topics of three presidential candidates' speeches, which are related to corruption issues in Indonesia, as illustrated in the examples below.

Example 1
Dimana ada kebaikan, disitu pasti ada setan yang kepanasan. Betul gak?
(Where there is goodness, there must be demons in heat. Isn't that, right?)

Example 2
Ngomongin masalah cerdas, mungkin yang ngomong cerdas Anies cerdas matanya buta kayaknya… Jakarta aja dibikin hamburadul.. pohon mahuni dah cantik nan hijau ditebangin semua… maksudnya cerdas bikin patung bambu jaringan hitam kasus E-KTP patung sepatu atap rumah warna warni maksudnya… beras aja korupsi.. herannya kok bisa aja dapat tiket nyapres… kita lihat aja nanti siapa yang terpilih
(Speaking of intelligent matters, those who say Anies is smart, you have blind eyes. Even Jakarta is made wasted by him. Beautiful and green trees were cut down as well. Do you mean smart by creating a black network bamboo statue, E-KTP issue, shoe statue, or even a colorful house? What do you mean? The rice was even being corrupted. I am just wondering how he could get a chance to be a presidential candidate. We will see who gets elected.)

Hate comments employed by Indonesian netizens tended to show how they mentioned a presidential candidate was seen as evil (*there must be demons in heat*).

These included threats of insults and comments that discriminated against a certain presidential candidate who was not supposed to be elected in the presidential election. Some Indonesian netizens mentioned attacking presidential candidates in terms of their intelligence (*who is saying Anies is smart?*), indicating that this presidential candidate is not smart enough to be elected to lead the country. Meanwhile, Indonesian netizens tended to attack others physically (*you have blind eyes*), which seemed to engage discussion among netizens by attacking one another in online discussion forums by letting their arguments in comment sections. By analysing these examples, the hate comments involving dehumanization and demonization can lead to the justification of harsh treatment in online political discourse. The spread of hate comments by Indonesian netizens as political supporters are seen to be employed during political campaigns before the presidential election in 2024.

Early warning

Early warning context in the current study highlighted the vision and mission of the three presidential candidates' speeches. These candidates tried to deliver their point of view about the future of Indonesia. However, some hate comments were found to be related to this aspect, as illustrated in the following examples.

Example 3
Setelah saya tonton sampai selesai, saya sepenuhnya belum yakin gagasan yang akan dilakukan cawapres Ganjar Pranowo dapat diimplementasikan. Semua gagasannya bagus tapi tetap saja masih dikendalikan oleh partai pengusungnya
(After I watched it till the end, I was not completely sure that the idea that Ganjar Pranowo would carry out as the presidential candidate could be implemented. All his ideas are good but still controlled by his party.)

Example 4
Pak Prabowo terlalu tua, gak lincah lagi kalau harus terjun langsunng ke Masyarakat. Anies?
(Mr. Prabowo is too old. He is not too young anymore if he has to go directly into society. What about Anies?)

The idea of early warning included the practice of early evaluation by seeing the results of the presidential candidates before applying for the elections. Indonesian netizens clearly stated that the ideas provided by the candidate should not be accurate and valid if this candidate is under the control of a particular political party in Indonesia (*all his ideas are good, but he is still under the control of his political party*). This early warning indicated that Indonesians should have a future mindset about the vision and mission delivered by the candidates through the talk shows. In addition, physical condition (*he is not too young anymore*) can be considered one of the early warnings for all Indonesians to select who would be the president of Indonesia in 2024. The physical condition may impact the results of the leadership. Moreover, these examples were prevalent in YouTube comment sections employed by Indonesian netizens, particularly during significant talk shows of online political discourse in Indonesia.

Violence and incitement

Violence and incitement context in the current study highlighted Indonesian netizens' disagreement regarding the notion of incitement to hatred. This was partly due to disagreements about what has been delivered by the three presidential candidates related to the government policies and their implementation, as illustrated in the following examples.

Example 5
Jadi Menteri penyerangan dong kalo harus agresif, presidennya aja diem diem jual pulau, bawahan bisa apa? Mungkin hal yang sama juga terjadi nanti
(Be the Minister of attack if you have to be aggressive. The president just secretly sells the island. What can subordinates do? The same thing might happen in the future.)

Example 6
Prabowo nggak usah cerita yang akan datang.. nggak masuk akal blas.. sekarang yang dikerjakan aja sudah bertele-tele.. anggaran belanja alutsista aja nggak jelas..
(Prabowo does not need to tell the story to come. It doesn't make sense. Now, what is done is long-winded. The budget for defense equipment spending is not clear.)

As described, the hate comments in terms of violence and incitement confirmed the context of sensitivity of the phenomenon related to aggressively attacking the candidate by assuming a fact without providing proof (*the president just secretly sold the island*). In this case, Indonesian netizens clearly attacked a specific person to do the same thing that probably happened when he was elected president. Furthermore, Indonesian netizens did not expect illogical ideas, so they illustrated them using stories told by the candidates during the talk shows (*it does not make sense*). It was clearly stated that the ideas should be provided in an orderly and logical manner. These examples potentially lead to the rejection of the government policies and their implementation. In addition, Indonesian netizens tended to call for increased regulations and monitoring to prevent the escalation of tensions.

Offensive language

The context of the offensive language employed by Indonesian netizens was in the form of insults because of the protests that the three presidential candidates explained during the talk shows. This had the effect of anger and annoyance among Indonesian netizens, where anger and annoyance were a form of hate speech that aimed to release their negative feelings, as illustrated in the following examples.

Example 7
Apa yang dibanggakan dari orang seperti Anies? 0%
(What should we be proud of someone like Anies? 0%)

Example 8
Ganjar gak bisa kasih solusi dan gak punya gagasan
(Ganjar cannot give solutions and does not have any ideas)

During the talk shows, Indonesian netizens tended to assess and evaluate each candidate's capability. As a country with various kinds of cultures, the way of criticizing somebody should be delivered in a good manner by illustrating it using polite words. However, the example above showed that Indonesian netizens employed bad words to attack one of the presidential candidates for his lack of capability and intelligence (*Ganjar cannot give solutions*). Moreover, Indonesian netizens tended to use hate comments to rely on bringing down the candidate. They noticed that they should not be proud of having a candidate who did not have many achievements. The political discourse of these examples highlights the prevalent political spheres that potentially contribute to the disintegration in which the offensive language employed by Indonesian netizens extends beyond just the targeted individuals, particularly those three presidential candidates for the election in 2024.

Factors affecting hate comments on YouTube

The issue of hate comments in the current study highlighted the factors affecting the hate comments employed by Indonesian netizens toward the presidential candidates' talk shows on YouTube. The results showed that Indonesian netizens employing hate comments were mainly affected by anonymity and personality traits, where these factors were highly used in the form of early warning hate comments, as illustrated in the following figure.

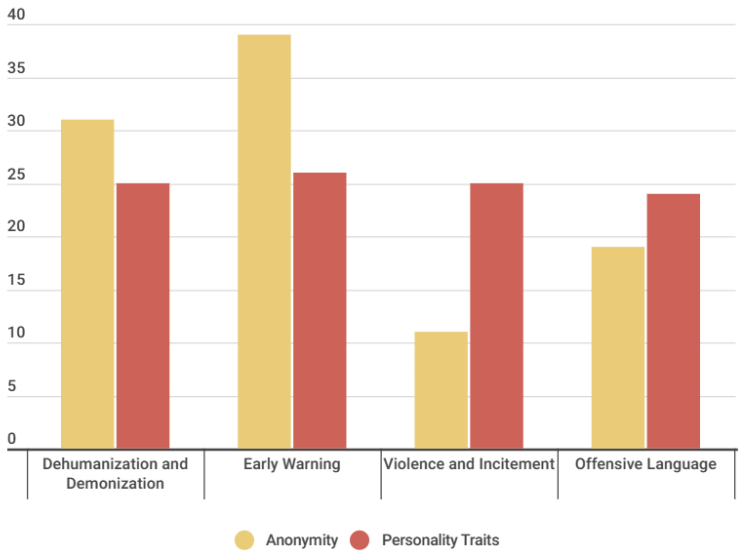


Figure 3. Factors contributing to hate comments among Indonesians regarding the presidential election

One factor contributing to the prevalence of hate comments is the anonymity found on YouTube. It suggests that Indonesian netizens may feel less inhibited online due to the lack of face-to-face interaction and the ability to remain anonymous. This anonymity can lead to a decrease in self-regulation and an increase in aggressive behavior. The current study's findings show that anonymity contributes to the prevalence and impact of hate comments employed by Indonesian netizens on presidential talk shows on YouTube. Their hate comments are performed to enable victims' presidential candidates to be involved in hateful

comments threatening the candidate's freedom *(I just wonder how come he can get a chance to be a presidential candidate)*. Anonymity found in the current study is also considered an attractive hate comment leading to hate speech, which can be categorized as masking online criminal activities. Indonesian netizens highly employ this factor to show their hate speech to a particular candidate.

Certain personality traits are associated with aggressive and violent behavior, which can manifest in hate comments. As described in the current study's findings, Indonesian netizens high in traits such as extraversion, neuroticism, and low agreeableness are more likely to engage in aggressive behavior, including online hate speech. The findings of the current study also show that this factor may predispose the presidential candidates to engage in hate speech and cyberbullying regarding the talk shows delivered by each of the candidates. It is due to Indonesian netizens tend to show their anger *(now what is done is long-winded)* as a personality trait that may be more likely to engage in hate comments and cyberbullying.

Discussions

The present study provided valuable insights into the language use employed by Indonesian netizens. The findings showed that most Indonesian netizens employed early warning of hate comments on presidential talk shows, followed by dehumanization and demonization, violence and incitement, and offensive language. Understanding these hate comments to counter hate comments on online platforms was crucial because online platforms have changed how people interact and respond to online discourse (Aporbo, 2023).

The present study showed that Indonesian netizens mainly employed the early warning category on the presidential talk show on YouTube. Indonesian netizens tended to use strong, offensive, or even aggressive language to provoke others, as provocation was considered a particular reference to the issues of impoliteness (Hall, 2022). Provocation on social media may involve rude or uncivil online behavior and can take many different forms, such as humorous provocations, including teasing, ribbing, making fun of someone, and even pulling someone's leg (Hasell & Weeks, 2016; Kotthoff, 2009; Wati, 2023).

Furthermore, the incitement on social media employed by Indonesian netizens tended to engage in harmful behavior in which incitement on social media could violate not only the rules of the particular platform but it could also violate legal and ethical standards (Etaywe & Zappavigna, 2023; Jaconelli, 2018). The current findings indicated that Indonesian netizens often manifested the direct and indirect encouragement of online violent behaviors against others on social media. This process might involve promoting hate comments by employing impolite utterances of derogatory language (Taradhita & Putra, 2021). As the present study found that Indonesian netizens employ the violence and incitement category, it showed that Indonesian netizens tended to use offensive or derogatory language to attack or demean others in online communications. The hate comments they employ can take various forms and provide different purposes, as insults can be in the form of in/direct insults, explicit and implicit insults, non-/pure, and non-vocatives (Baczkowska, 2021). On the other hand, they generally emotionally and socially aimed to harm others as they tended to express impolite words through sarcasm, irony, or even mockery by making fun of others in online hate comments. Impoliteness in insulting others, particularly in relation to hate comments on social

media, can have serious consequences, such as cyberbullying, emotional distress, or even a toxic online environment and hate crime (Culpeper et al., 2022). Employing impoliteness of insults that constitute intentionally threatening acts (Bernal, 2022) can contribute to a decline in the quality of online communications.

Furthermore, the current study showed that Indonesian netizens tended to use offensive language that is intended to mock, insult, or even ridicule others. This can be in cultural and religious contexts, as the Indonesian language on social media is full of offensive language and tends to be provocative (Anwar et al., 2021). Both cultural and religious contexts have quite distinct perspectives on offensive language and can have very varied effects on online social interactions that can lead to blasphemy. Due to the wide reach and diversity of social media platforms, blasphemy in cultural and religious contexts (Belyutina, 2023) on social media can be a particularly delicate aspect by considering its impact and consequences in the online environment as comments or anything that is disrespectful to religious or traditions can elicit outrage and calls for action. In other words, when discussing religious or cultural contexts, cross-cultural relationships and idea exchanges can cause misunderstandings and potentially offensive conflicts (Coats, 2021). Moreover, dehumanization and demonization in the form of hate comments on social media can be a source of hate speech as social media has provided a platform for individuals to raise about the vital topics or issues in which it intends to describe painful or absurd situations (Agustina, 2023).

The current study findings showed that dehumanization and demonization were seen as instances of impoliteness and contentious behavior on social media platforms. Indonesian netizens have shown online behavior that resorts to harassment or cyberbullying tactics. In addition, defamation language, such as forensic linguistics (Shuy, 2022), lies in disseminating false information about something or someone that can include allegations of wrongdoing or damaging claims involving social emotions, harming an individual's dignity and reputation in the social community (Nieto, 2020). The current findings indicated that Indonesian netizens intended to tarnish an individual's image or credibility, potentially leading to social, professional, or even personal aspects.

In relation to online political discourse, the present study provided insights into the presidential election in 2024, focusing on the hate comments employed by Indonesian netizens. The use of hate comments on YouTube allowed Indonesian netizens to create a space to engage themselves in online discussions related to political issues, particularly the presidential election. This situation contributed to particular perspectives that ultimately hinder the various points of view, which are crucial for political discourse (Wilhelm et al., 2020). The findings also suggested that Indonesian netizens tended to expose their opinions, which allowed them to express their own point of view. Surprisingly, they tended to employ hate comments in accordance with the increased feelings of anger (Vehovar & Jontes, 2021).

In addition, the current study highlighted the factors affecting hate comments on presidential talk shows on YouTube in which anonymity and personal trait factors were found in this study. These factors contributed to the use of hate comments among Indonesian netizens toward presidential candidate talk shows on YouTube. Maintaining this anonymity factor allowed Indonesian netizens to attack others more easily than through face-to-face interaction (Limba & Šidlauskas, 2018). Moreover, this anonymity factor was seen to protect Indonesian netizens'

privacy and their freedom of expression. Mossie and Wang (2018) emphasized that the anonymity of social networks made it attractive for hate speech, as Indonesian netizens tended to express their opinions without fear of retaliation. However, Hsueh et al. (2015) argued that comfort and lack of accountability provided by anonymity did not uniformly cause online commenters to respond in antisocial ways, but the impact of anonymity on prejudice depended on social norms. It can be argued that Indonesian netizens in the current study tended to express hateful sentiments without facing the social, professional, or legal consequences they might have if their identities were known.

Furthermore, personal traits contributed to hate comments among Indonesian netizens who commented on presidential talk shows uploaded on YouTube. The research findings have shown that certain personality traits, including dark personality traits, aggression, and moral personality, were associated with the occurrence of hate comments and cyberbullying. By employing these certain personal traits, Indonesian netizens tended to create an environment where they felt more emboldened to express hate comments. Therefore, personality traits concomitantly present a pivotal explanation of speech in online communication, where Indonesian netizens with certain authoritarian tendencies were likely to engage in hate speech (Bilewicz et al., 2017).

Conclusions

This paper aims to answer the research question of what hate comment types and what factors Indonesian netizens employ toward the presidential talk shows on YouTube. The types of hate comments and the factors were addressed in this research. The findings of the present paper show that the type of hate comment is mostly early warning employed by Indonesian netizens, followed by other hate comment types, including dehumanization and demonization, violence and incitement, and offensive language. The following factors contribute to the use of hate comments in the present paper, including anonymity and personality traits. The classifications of these factors were explained and discussed, and they provided a shield for what Indonesian netizens feel and the language used to influence its characteristics exhibited in hate comments employed by Indonesian netizens on YouTube. Therefore, the present study highlighting hate comments on presidential talk shows has revealed significant insights into this pervasive and concerning phenomenon in online communication, particularly the prominent issue in the current digital age of political discourse.

By addressing this phenomenon, this study suggests that social media platforms, decision-makers, and users themselves are encouraged to pay attention to this issue. Promoting a culture of respect and empathy in the online environment is also important for a digital space to be safer and more inclusive, particularly on online talk shows that provide audiences with comments.

Since the current investigation into the hate comments employed by Indonesian netizens on presidential talk shows has yielded valuable insights, it is important to acknowledge certain limitations that might affect the applicability and generalizability of the current study's findings. The current study only investigated hate comments employed by Indonesian netizens on particular talk shows uploaded to YouTube. Future research may consider other investigations in different social

media platforms such as Twitter, Instagram, Facebook, or even WhatsApp applications, as these platforms seem popularly used by Indonesians.

References

Abdalla, M., Ally, M., & Jabri-Markwell, R. (2021). Dehumanisation of 'outgroups' on Facebook and Twitter: Towards a framework for assessing online hate organisations and actors. *SN Social Sciences, 1*(9), 1-28. https://doi.org/10.1007/s43545-021-00240-4

Agustina, T. (2023). The figurative language sarcasm and satire on the song "We can't stop" and "WTF do i know" by Miley Cyrus and "Reckless" by Madison Beer. *Journal of Language, Literature, and Teaching*, *4*(3), 84-90. https://doi.org/10.35529/jllte.v4i3.84-90

Al-Hassan, A., & Al-Dossari, H. (2019). Detection of hate speech in social networks: A survey on multilingual corpus. *Academy and Industry Research Collaboration Center (AIRCC)*, 83-100. https://doi.org/10.5121/csit.2019.90208

Andersson, M. (2021). The climate of climate change: Impoliteness as a hallmark of homophily in YouTube comment threads on Greta Thunberg's environmental activism. *Journal of Pragmatics*, *178*, 93-107. https://doi.org/10.1016/j.pragma.2021.03.003

Anwar, M. (2019). Impoliteness in Indonesian language on Facebook as a representation of cultural blindness. *Journal of Multicultural Education*, *5*(1), 88-91.

Anwar, M., Murtadho, F., Boeriswati, E., Yarmi, G., & Rosa, H. T. (2021). Analysis model of impolite Indonesian language use. *Linguistics and Culture Review*, *5*(S3), 1426-1441. https://doi.org/10.21744/lingcure.v5ns3.1840

Aporbo, R. J. (2023). Sociopragmatic analysis of Filipino celebrities "Posts and fans" comments. *World Journal of English Language*, *13*(5), 544-558. https://doi.org/10.5430/wjel.v13n5p544

Ardila, J. A. G. (2019). Impoliteness as a rhetorical strategy in Spain's politics. *Journal of Pragmatics*, *140*, 160-170. https://doi.org/10.1016/j.pragma.2018.11.017

Baczkowska, A. (2021). "you're too thick to change the station" - Impoliteness, insults and responses to insults on Twitter. *Topics in Linguistics*, *22*(2), 62-84. https://doi.org/10.2478/topling-2021-0011

Belyutina, Y. (2023). Taboo language as a means of the category of impoliteness in the English language. *Izvestia of Smolensk State University*, *4*(60), 106-115. https://doi.org/10.35785/2072-9464-2022-60-4-106-115

Bilewicz, M., Soral, W., Marchlewska, M., & Winiewski, M. (2017). When authoritarians confront prejudice: Differential effects of SDO and RWA on support for hate-speech prohibition. *Political Psychology*, *38*(1), 87-99. https://doi.org/10.1111/pops.12313

Bitonti, A., Marchetti, R., & Mariotti, C. (2023). Did COVID-19 change the political communication of polarizing leaders? The case of Salvini's campaigning before and after the pandemic. *European Journal of Communication*, *38*(4), 380-397. https://doi.org/10.1177/02673231221140697

Bousfield, D., & Locher, M. A. (2008). *Impoliteness in language: Studies on its interplay with power in theory and practice*. Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110208344

Brown, P., & Levison, S. (1987). Politeness: Some universals in language usage. Cambridge: Cambrigde University Press.

Castellanos, M., Wettstein, A., Wachs, S., Kansok-Dusche, J., Ballaschk, C., Krause, N., & Bilz, L. (2023). Hate speech in adolescents: A binational study on prevalence and demographic differences. *Frontiers in Education*, *8*, 1-14. https://doi.org/10.3389/feduc.2023.1076249

Chekol, M. A., Moges, M. A., & Nigatu, B. A. (2023). Social media hate speech in the walk of Ethiopian political reform: Analysis of hate speech prevalence, severity, and natures. *Information Communication and Society*, *26*(1), 218-237. https://doi.org/10.1080/1369118X.2021.1942955

Chiril, P., Pamungkas, E. W., Benamara, F., Moriceau, V., & Patti, V. (2022). Emotionally informed hate speech detection: A multi-target perspective. *Cognitive Computation*, *14*(1), 322-352. https://doi.org/10.1007/s12559-021-09862-5

Coats, S. (2021). 'Bad language' in the Nordics: Profanity and gender in a social media corpus. *Acta Linguistica Hafniensia*, *53*(1), 22-57. https://doi.org/10.1080/03740463.2021.1871218

Culpeper, J. (2011). *Impoliteness: Using language to cause offence*. Cambridge: Cambridge University Press.

Culpeper, J., Iganski, P., & Sweiry, A. (2022). Linguistic impoliteness and religiously aggravated hate crime in England and Wales. *Journal of Language Aggression and Conflict*, *5*(1), 1-29. https://doi.org/10.1075/jlac.5.1.01cul

Culpeper, J., Haugh, M., & Kadar, D. Z. (2017). *The pulgrave handbook of linguistic (im)politeness*. London: Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-37508-7

Đorđević, J. P. (2020). The sociocognitive dimension of hate speech in readers' comments on Serbian news websites. *Discourse, Context and Media*, *33*, 1-9. https://doi.org/10.1016/j.dcm.2019.100366

Elfrida, R., & Pasaribu, A. N. (2023). Hate speech on social media: A case study of blasphemy in Indonesian context. *English Review: Journal of English Education*, *11*(2), 433-440. https://doi.org/10.25134/erjee.v11i2.7909

Erjavec, K. (2014). Readers of online news comments: Why do they read hate speech comments? *ANNALES*, *24*(3), 451-462.

Esau, K. (2021). Impoliteness (hate speech/incivility). *DOCA - Database of Variables for Content Analysis*, *1*(5), 1-5. https://doi.org/10.34778/5b

Etaywe, A., & Zappavigna, M. (2023). The role of social affiliation in incitement: A social semiotic approach to far-right terrorists' incitement to violence. *Language in Society*, 1-26. https://doi.org/10.1017/s0047404523000404

Fortuna, P., & Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys*, *51*(4), 1-30. https://doi.org/10.1145/3232676

Grice, H. P. (1975). "Logic and conversation". In P. Cole & J. Morgan (eds.), *Studies in syntax and semantics: Speech acts* (Vol 3, pp. 41-58). New York: Academic Press.

Guo, L., & Johnson, B. G. (2020). Third-person effect and hate speech censorship on Facebook. *Social Media and Society*, *6*(2), 1-12. https://doi.org/10.1177/2056305120923003

Hall, J. (2022). Banter, teasing and politeness in Varro's De Re Rustica. In L.U. Gómez & Ł. Berger (eds.), *Politeness in ancient Greek and Latin* (pp. 273-292). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009127271.012

Han, Y. (2021). Situated impoliteness revisited: Blunt anti-epidemic slogans and conflicting comments during the coronavirus outbreak in China. *Journal of Pragmatics, 178,* 31-42. https://doi.org/10.1016/j.pragma.2021.03.004

Hangartner, D., Gennaro, G., Alasiri, S., Bahrich, N., Bornhoft, A., Boucher, J., Demirci, B. B., Derksen, L., Hall, A., Jochum, M., Munoz, M. M., Richter, M., Vogel, F., Wittwer, S., Wüthrich, F., Gilardi, F., & Donnay, K. (2021). Empathy-based counterspeech can reduce racist hate speech in a social media field experiment. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(50), 1-3. https://doi.org/10.1073/pnas.2116310118

Hasell, A., & Weeks, B. E. (2016). Partisan provocation: The role of partisan news use and emotional responses in political information sharing in social media. *Human Communication Research*, *42*(4), 641-661. https://doi.org/10.1111/hcre.12092

Hsueh, M., Yogeeswaran, K., & Malinen, S. (2015). "Leave your comment below": Can biased online comments influence our own prejudicial attitudes and behaviors? *Human Communication Research*, *41*(4), 557-576. https://doi.org/10.1111/hcre.12059

Jaconelli, J. (2018). Incitement: A study in language crime. *Criminal Law and Philosophy*, *12*(2), 245-265. https://doi.org/10.1007/s11572-017-9427-8

Kilian, S. (2021). From the "Paris agreement" to the "London pact": Political aspects of the democratisation of Poland's April constitution of 1935. *Historia i Polityka*, *36*(43), 99-113. https://doi.org/10.12775/hip.2021.016

Kunst, M., Porten-Cheé, P., Emmer, M., & Eilders, C. (2021). Do "Good citizens" fight hate speech online? Effects of solidarity citizenship norms on user responses to hate comments. *Journal of Information Technology and Politics*, *18*(3), 258-273. https://doi.org/10.1080/19331681.2020.1871149

Limba, T., & Šidlauskas, A. (2018). Peculiarities of anonymous comments' management: A case study of Lithuanian news portals. *Entrepreneurship and Sustainability Issues*, *5*(4), 875-889. https://doi.org/10.9770/jesi.2018.5.4(12)

MacAvaney, S., Yao, H. R., Yang, E., Russell, K., Goharian, N., & Frieder, O. (2019). Hate speech detection: Challenges and solutions. *PLoS ONE*, *14*(8), 1-16. https://doi.org/10.1371/journal.pone.0221152

Mossie, Z., & Wang, J.-H. (2018). Social network hate speech detection for Amharic language. *Academy and Indsutry Research Collaboration Center (AIRCC)*, 41-55. https://doi.org/10.5121/csit.2018.80604

Niebuhr, O., & Neitsch, J. (2022). The truth below the surface. *Journal of Speech Sciences*, *11*(00), 1-27. https://doi.org/10.20396/joss.v11i00.16153

Nieto, V. G. (2020). Defamation as a language crime-a sociopragmatic approach to defamation cases in the high courts of justice of Spain. *International Journal of Language & Law*, *9*, 1-22. https://doi.org/10.14762/jll.2020.001

Nwozor, A., Ajakaiye, O. O. P., Okidu, O., Olanrewaju, A., & Afolabi, O. (2022). Social media in politics: Interrogating electorate-driven hate speech in Nigeria's 2019 presidential campaigns. *EJournal of EDemocracy and Open Government*, *14*(1), 104-129. https://doi.org/10.29379/jedem.v14i1.683

Paasch-Colberg, S., & Strippel, C. (2022). "The boundaries are blurry…": How comment moderators in germany see and respond to hate comments. *Journalism Studies*, *23*(2), 224-244. https://doi.org/10.1080/1461670X.2021.2017793

Paasch-Colberg, S., Strippel, C., Trebbe, J., & Emmer, M. (2021). From insult to hate speech: Mapping offensive language in german user comments on immigration. *Media and Communication*, *9*(1), 171-180. https://doi.org/10.17645/mac.v9i1.3399

Poletto, F., Basile, V., Sanguinetti, M., Bosco, C., & Patti, V. (2021). Resources and benchmark corpora for hate speech detection: A systematic review. *Language Resources and Evaluation*, *5*(2), 477-523. https://doi.org/10.1007/s10579-020-09502-8

Rabab'ah, G., & Alali, N. (2020). Impoliteness in reader comments on the Al-Jazeera channel news website. *Journal of Politeness Research*, *16*(1), 1-43. https://doi.org/10.1515/pr-2017-0028

Reichelmann, A., Hawdon, J., Costello, M., Ryan, J., Blaya, C., Llorent, V., Oksanen, A., Räsänen, P., & Zych, I. (2021). Hate knows no boundaries: Online hate in six nations. *Deviant Behavior*, *42*(9), 1100-1111. https://doi.org/10.1080/01639625.2020.1722337

Reiter, R. M. (2021). How can ethnography contribute to understanding (im)politeness? *Journal of Politeness Research*, *17*(1), 35-59. https://doi.org/10.1515/pr-2020-0040

Shinta, V. M., Wahyuni, D., & Padang, U. N. (2018). Impoliteness strategies used by supporters and detractors of Ahok in their online comments by gender. *E-Journal of English Language & Literature*, *7*(1), 225-236.

Shuy, R. W. (2022). Forensic linguistics. *Encyclopedia of Forensic Sciences*, *2*, 620-629. https://doi.org/10.1016/B978-0-12-823677-2.00005-2

Siahaan, I. P. S., Rangkuti, R., & Ganie, R. (2019). hate speech used by haters of Lady Gaga on social media. *Nusa: Jurnal Ilmu Bahasa Dan Sastra*, *14*(4), 573-582. https://doi.org/10.14710/nusa.14.4.573-582

Silva, L., Mondal, M., Correa, D., Benevenuto, F., & Weber, I. (2016). Analyzing the targets of hate in online social media. *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, *10*(1), 687-690. https://doi.org/10.1609/icwsm.v10i1.14811

Sinkeviciute, V. (2018). "Ya bloody drongo!!!": Impoliteness as situated moral judgement on Facebook. *Internet Pragmatics*, *1*(2), 271-302.

Subyantoro, S., Apriyanto, S., Siroj, M., Nurhadi, R., Masykuri, E. S., Jumanto, Sulistyorini, H., Alabi, T. O., Ayeloja, A. K., Polytechnic, T. F., State, E., & State, O. (2019). Hate speech based on pragmatics studies in social. *Journal of Advances in Linguistics*, *3*(12), 324-333.

Taradhita, D. A. N., & Putra, I. K. G. D. (2021). Hate speech classification in Indonesian language tweets by using convolutional neural network. *Journal of ICT Research and Applications*, *14*(3), 225-239. https://doi.org/10.5614/itbj.ict.res.appl.2021.14.3.2

Vehovar, V., & Jontes, D. (2021). Hateful and other negative communication in online commenting environments: Content, structure and targets. *Acta Informatica Pragensia*, *10*(3), 257-274. https://doi.org/10.18267/j.aip.165

Vladimirou, D., & House, J. (2018). Ludic impoliteness and globalisation on Twitter: 'I speak England very best' #agglika_Tsipra, #Tsipras #Clinton. *Journal of Pragmatics*, *134*, 149-162. https://doi.org/10.1016/j.pragma.2018.05.005

Wati, D. R. (2023). The ethics of social media communication in the perspective of the al-qur'an. *Journal of Islamic Communication and Counseling*, *2*(2), 93-107. https://doi.org/10.18196/jicc.v2i2.28

Wilhelm, C., & Joeckel, S. (2019). Gendered morality and backlash effects in online discussions: An experimental study on how users respond to hate speech comments against women and sexual minorities. *Sex Roles*, *80*(7–8), 381-392. https://doi.org/10.1007/s11199-018-0941-5

Wilhelm, C., Joeckel, S., & Ziegler, I. (2020). Reporting hate comments: Investigating the effects of deviance characteristics, neutralization strategies, and users' moral orientation. *Communication Research*, *47*(6), 921-944. https://doi.org/10.1177/0093650219855330

Wrenn, M., & Reed, A. (2019). Developing academic discourse through literacy and the 2016 U.S. presidential election: A design-based approach. *Journal of Adolescent and Adult Literacy*, *63*(2), 189-200. https://doi.org/10.1002/jaal.983

Zhao, Y. (2022). A pragmatic analysis of politeness and impoliteness theory in TikTok. *Journal of Higher Education Research*, *3*(1), 83-86. https://doi.org/10.32629/jher.v3i1.661