



Investigating Hate Speech Comments: A Forensic Linguistic Study

Rika Purnama Sari^{1*}, Miftahulhairah Anwar², Muhammad Kamal bin Abdul Hakim³

^{1,2,3} Applied Linguistics Department, Universitas Negeri Jakarta, Jakarta Timur 13220, Indonesia
Email: rika.purnama.sari@mhs.unj.ac.id, miftahulhairah@unj.ac.id, m.kamal.ah@unj.ac.id

ARTICLE INFO	ABSTRACT
Received: 02 December 2024	The meaning of hate speech in social media can shift because the way people use social media is constantly changing, and the social and political context in which social media is used can also change. This study aims to analyze hate speech comments on Instagram and the meaning and relevance of the implicature to forensic linguistics. This research employs a qualitative approach to collect and analyze the meaning of hate speech. The research data was obtained from Instagram through hashtag searches. The main findings of this study indicate that hate speech on social media tends to avoid direct vulgar linguistic forms and instead relies on implicit language strategies to convey negative sentiments towards individuals or groups. There is widespread use of diminutives such as “bocah” or “bocil,” social stigmatization such as “ndeso” and “bau kencur,” and institutional euphemisms such as “makkamah keluarga,” all of which are important linguistic indicators in the analysis of political hatred in the digital space. The intensity classification shows that hate speech with moderate and severe intensity does not only come from explicit insults but also from narratives that normalize harassment based on age, appearance, and political affiliation. The implications of this study indicate that a forensic linguistic approach is very effective in uncovering forms of hate speech disguised through language strategies such as sarcasm, stereotypes, and euphemisms. Theoretically, these findings enrich the study of pragmatic linguistics and speech acts by showing the importance of implicature analysis in detecting non-explicit hate intentions.
Accepted: 20 February 2025	
Published: 05 April 2025	
Keywords: Bad comments, Forensic linguistics, Hate speech, Implicature meaning, Language evidence, Speech acts, Social media	
Doi: https://doi.org/10.59011/austronesian.4.1.2025.47-64	

*Corresponding Author
E-mail: rika.purnama.sari@mhs.unj.ac.id
Copyright © 2025 The Author(s)
CC BY-NC-SA 4.0 License

1. Introduction

Hate speech in social media is any communication that attacks a person or group based on attributes such as race, religion, ethnic origin, national origin, sex, disability, sexual orientation, or gender identity (Teh et al., 2018). It can take many forms, including direct and indirect attacks, threats, and harassment.

Social media platforms have become breeding grounds for hate speech in recent years. It is due to several factors, including the anonymity that social media provides, the ease with which hate speech can be disseminated, and the algorithms that social media platforms use to show users content similar to what they have liked or interacted with in the past (Singer & Brooking, 2018).

The meaning of hate speech in social media can shift over time. It happened because the way that people use social media is constantly changing and because the social and political context in which social media is used can also change (Alsagheer et al., 2023; Aulia, 2022; Ibrohim & Budi, 2023; Mintowati & Dasion, 2020; Nave & Lane, 2023; Ria & Setiawan, 2023; Subramanian et al., 2023; Yuliyanti et al., 2020). For example, the term “triggered” was originally used to describe the psychological response of people who have experienced trauma. However, it has since been appropriated by internet trolls and used as a pejorative term to describe easily offended people. Another example is the term “cuck”. This term was originally used to describe a cuckold, which is a man whose wife has sex with another man. However, it has since been appropriated by white supremacists and used as a pejorative term to describe white men who they believe are betraying their race by supporting diversity and inclusion. The shifting meaning of hate speech on social media can make it difficult to identify and combat. It is important to be aware of how somebody’s hate speech can go on social media and to be vigilant in reporting it when we see it (see Harahap & Siregar, 2022). Hate speech is a complex notion, especially from a computational perspective (Sanguinetti et al., 2019; Finegan & Lee, 2020).

There has been a significant amount of research on hate speech on social media conducted previously using various data and methodologies (e.g., Al-Saqqa et al., 2024; Fonseca et al., 2024; Ghaly et al., 2024; Ghenai et al., 2025; Putra et al., 2024; Mubarok et al., 2024; Ramli et al., 2024; Wasi et al., 2024). Al-Saqqa et al. (2024) conducted a study specifically to detect hate speech on Arab social media by involving machine learning, deep learning, and transfer learning approaches. The results showed that real-time detection of harmful content is essential to protect vulnerable communities (see also Ghaly et al., 2024). Then, Fonseca et al. (2024) analyzed hate speech in Portuguese Twitter/X network conversations. The main findings of this study revealed that the number of users followed by an individual and the likelihood of reading were relevant predictors of a user’s tendency to post aggressive content.

Furthermore, Ghenai et al. (2025) analyzed the dynamics of hate speech on Twitter/x social media with a dataset of 6,002 users. The findings of their study revealed that users who uploaded hateful content showed significantly higher levels of anger, anxiety, and negative emotions, accompanied by an increase in the use of third-person pronouns.

Further research from Putra et al. (2024) on analyzing hate speech comments on the Indonesian House of Representatives' Instagram. The analysis results show that the comments contain negative and derogatory elements, using language that tarnishes a good name and makes baseless accusations. Some comments also violate the rules of politeness, especially the rules of sympathy and praise. Likewise, Mubarok et al. (2024) and Ramli et al. (2024), in the case of hate speech, which expresses a negative attitude. This can be seen in the act of insulting (expressive).

Finally, research from Wasi et al. (2024) used bibliometric analysis to examine forensic linguistics. The results show that forensic linguistics is a significant field of research, education, and professional practice worldwide because this research has implications for researchers, linguists, lawyers, higher education policymakers, and university administrations. Hargreaves (2022) studied the appropriation of the term "woke" by the right. He found that the term "woke" has been increasingly used as a pejorative term to attack people who support social justice causes. The study found that the term is often used to dismiss the concerns of marginalized groups as frivolous or illegitimate. Yeagley (2020) studied the snowflake generation, A pejorative term used to silence and discredit dissent. He found that the term "snowflake" has been increasingly used as a pejorative term to attack people perceived as sensitive or easily offended. The study found that the term is often used to silence and discredit people speaking out against injustice. This research highlights the importance of awareness of the shifting meaning of hate speech in social media. By understanding how hate speech is used, we can be better equipped to identify and combat it According to Manne (2017), hate speech is divided into two types: expressive and performative. Expressive hate speech conveys hatred, hostility, or contempt for a particular group of people. It can be direct and overt, such as calling someone a racial slur, or it can be more indirect and implicit, such as using stereotypes or generalizations. Performative hate speech is speech that is intended to harm or silence a particular group of people. It can include threats of violence, calls for discrimination, or other acts that seek to exclude or marginalize a group. Therefore, this study aims to analyze hate speech comments on Instagram, the meaning of the implicature, and its relevance to forensic linguistics.

The forensic linguistic approach becomes a relevant and strategic framework for deeply understanding and classifying hate speech. Forensic linguistics allows the analysis of language in legal and social contexts, including the identification of speech that is discriminatory, inciting, or contains violence explicitly or implicitly (Coulthard, 2016; Coulthard & Johnson, 2011; Gibbons, 2003; McMenamin, 2008). This approach is based on several main theories. First, Olsson's (2012) forensic text analysis theory emphasizes the importance of analyzing linguistic structures (word choice, syntactic patterns, and rhetorical style) in determining the intention and effect of a speech. Second, van Dijk's (2008) Discourse and Power approach explains that hate speech cannot be separated from the power relations that shape and reproduce discriminatory ideologies through language. Third, Austin's (1962) and Searle's (2011) speech act theory provides a pragmatic framework for studying speech as action, including actions that hurt, discriminate against,

or marginalize certain groups. Fourth, Grice's (1975) implicature theory allows for exploring implied meanings in hate speech, especially implicit speech disguised through irony, metaphor, or euphemism.

Based on this theory, this study utilizes forensic linguistics as an approach to analyze hate speech not only textually but also contextually and ideologically. Through forensic analysis of hate speech on social media, researchers can analyze the types of speech (expressive and performative), language structure, and the implicature meaning contained therein (see Aulia, 2022; Borg, 2022; Gyuró, 2013; Kryk-Kastovsky, 2006). Thus, this study not only enriches the study of hate speech but also contributes to understanding forensic linguistics in a broader digital and social context.

2. Literature Review

2.1 Forensic Linguistics

Forensic linguistics is a branch of applied linguistics that studies and applies language analysis in legal, criminal, and judicial contexts. This discipline is growing rapidly due to the need to understand, identify, and analyze the use of language in various spoken and written legal cases. Forensic linguistics serves as a bridge between linguistics and law enforcement, offering important insights into criminal cases and resolving issues while comprehending the perspectives of linguists and attorneys (Varney, 1997). Within applied linguistics, forensic linguistics is a specialized study of language usage in various modalities. Its primary goal is to offer a strict legal framework for closely examining language patterns (Wasi et al., 2024). In addition to supporting academic research and incorporating technology into linguistic investigations, forensic linguistics improves language and communication, which benefits official, legal, and judicial domains (Sakakini, 2020). A significant rise in corpus-based research in forensic linguistics has made it possible to analyze big datasets in previously unthinkable ways and spot patterns in text (Wright, 2021). With important words like human, linguistics, legal translation, and linguistic rights, forensic linguistics has broadened to cover speech, text, plagiarism, verbal violence, discrimination, and social media detection (Alduais et al., 2023). Forensic linguistics can contribute to cybercrime investigations through applications like authorship analysis, sociolinguistic profiling, and determination of meaning (Perkins, 2018). The analysis of speech and its settings is aided by forensic linguistics, which provides linguistic evidence relevant to the case without establishing guilt or innocence (Kusumawardhani, 2024). According to Ali (2020), forensic linguistics is a promising discipline that plays a vital role in court procedures and aids in identifying language crimes. Examining different text forms and concentrating on auditory and acoustic phonetics, semantics, discourse, and pragmatics, forensic linguistics defines, analyzes, and investigates language in court as evidence for law enforcement, judges, and attorneys (Umiyati, 2020).

In conclusion, forensic linguistics is an interdisciplinary field that connects linguistics with law, playing an important role in analyzing language evidence in the

realm of justice. This discipline continues to develop with technological advances and the need for justice based on strong linguistic evidence.

2.2 Meaning

In recent years, semantic change has become a focal point in empirical Computational Linguistics/Natural Language Processing (NLP), as highlighted by Tredici et al. (2019). This increasing attention is paralleled by the demands of our modern era, where cross-cultural communication has become more prevalent (Sumiati et al., 2020).

The swift and ongoing development of information and communication technology has propelled society into what is now termed the fourth industrial revolution or the era of disruption, as articulated by Dewi et al. (2021). This rapid evolution signifies a transformative phase driven by technological advancements.

The pervasive presence of digital media in various aspects of daily life, particularly in research fields, has witnessed a recent surge. Febriyanti and Lustyantje (2020) highlight the diverse array of digital media, including the internet, websites, blogs, vlogs, videos, and platforms, contributing to this upsurge.

As indicated by Hasibuan et al. (2023), examining meaning change offers insights into the expansion, narrowing, connotation, synesthesia, and association of a word's meaning. In meaning shift, the initial reference does not necessarily change but experiences an expansion within the same phonetic symbol. The analysis of meaning shift, elucidated by Liswahyuningsih et al. (2020), reveals various forms, such as gain of information, loss of information, and skewing of information. These types of meaning shifts are discussed further:

- a. Gain of information: in this shift, additional information not present in the source language is introduced in the target language.
- b. Loss of Information: translators may omit certain information if a word or expression in the source language is deemed less crucial. This omission aims to enhance the communicative aspect of the translation.
- c. Skewing of Information: translators often adjust information when direct word-for-word translation is impractical for conveying equivalent meanings between the source and target languages. The intent of the source text is retained, ensuring a more communicative translation.

2.3 Hate Speech

People communicate with each other through language to share ideas or emotions (Sumiati et al., 2020). Hate speech, as defined by Mafeza (2016), refers to verbal conduct or expressions that are abusive, insulting, intimidating, or harassing, and/or incite violence, hatred, or discrimination against groups identified by specific characteristics. Jay (2009) further describes hate speech as bias-motivated speech directed at individuals based on gender, sexual orientation, race, ethnicity, religion, national origin, or disability.

The contemporary social behavior of individuals, who increasingly communicate through social media, contributes to the prevalence of digital crimes. Social media platforms such as Instagram often host content containing elements of pornography, hoax news, hate speech, and more (Riadi et al., 2022).

To provide a comprehensive understanding, Sanguinetti et al. (2019) outline additional categories related to online communication:

- a. Aggressiveness
- b. Offensiveness
- c. Irony
- d. Stereotype

3. Method

This research employs a qualitative approach to collect and analyze the meaning of hate speech. The research data was obtained from Instagram through hashtag searches: *#gibran* and *#gibranrakabumingraka*. The search for *#gibran* brought up 507 thousand posts, while *#gibranrakabumingraka* brought up 373 thousand posts. The highlighted data were 20 comments containing elements of hate speech determined randomly (see Appendix 1). Then, the Comments that had been collected were arranged in a table for analysis. There are three main theories used in this study: Olsson's (2012) theory emphasizes analyzing linguistic structures, Searle's (2011) theory for studying speech as action, and Grice's (1975) theory for the exploration of implied meanings.

The steps in analyzing hate speech are as follows. First, identify the utterances that are used in hate speech. It can be done by collecting a corpus of hate speech texts and identifying the most common utterances used. Second, identify types of hate speech. Third, the speech acts and implicature of the meaning of the utterances are analyzed (see Miles et al., 2014). It can be done using various methods, such as looking at the dictionary definitions and examining how the words are used in context. Finally, identify the implicature meaning in the words and phrases.

4. Results and Discussion

The following explanation focuses on disclosing the results of the analysis of hate speech comments, which is divided into several parts. First, the data description and the theory of hate speech analysis will be presented. Second, the analysis of the meaning of implicature and the relevance of forensic linguistics will be presented. Hate speech is a complex phenomenon that includes language or expressions that demean, mock, or damage the image of a person or group based on certain characteristics.

4.1 Results

1. Data Description and Theory of Hate Speech Analysis

This study analyzed 20 comments containing hate speech on social media, especially Instagram. All data are closely related to the presidential election, especially in terms of politics, age, and national leadership. The analysis was conducted using a forensic

linguistic approach, using theories from Olsson (2012), Searle (2011), and Grice (1975). The following are the results of the comments classification based on the category of hate speech, intensity, linguistic characteristics, types of speech acts, and implicature meaning.

Table 1. Results of 20 Hate Speech Comments Analysis

No	Comments/ utterances	Target & Categories (Olsson)	Intensity	Linguistic Features	Speech Act (Searle)	Implicature (Grice)
1	<i>ksian di hina mulu si monyong</i>	Racial/physical appearance	Light	Implicit insults, irony	Expressive	False sympathy to insult; implying ridicule
2	<i>Anjir males bngt denger si gibran dan si kaesang</i>	Political figures/family	Light	Expression of disgust, mild vulgarity	Expressive	They are not fit to appear in public
3	<i>Dasar kutu loncat</i>	Ideological (traitor)	Medium	Pejorative label	Assertive	Accused of being disloyal, opportunistic
4	<i>Efek adanya MAHKAMAH KELUARGA...</i>	Legal institutions/criticism of nepotism	Medium	Sarcasm, negative euphemism	Assertive	MK was controlled for family interests
5	<i>INI LAH HASIL NYA ANAK MUDA...</i>	Youth/dynasty	Heavy	Generalization, hyperbole	Assertive + Expressive	Power-hungry youth; systemic critique
6	<i>Pak Jokowi BIG MOUSE...</i>	President/ruler	Medium	Satire metaphor, capitalization	Assertive + Expressive	Jokowi is a dominant and authoritarian figure
7	<i>emangnya dia pikir dia ni siapa sih anying</i>	Personal/offensive	Heavy	Explicit swearing, harsh rhetoric	Directive	Rejecting authority/figures head-on
8	<i>jangan smpe planga plongo jilid 2</i>	Leader figure	Light	Satire, physical stereotypes	Directive	An appeal not to repeat the previous leader
9	<i>ketipu tampang ndeso...</i>	Appearance/social	Medium	Stereotypes, sarcasm	Assertive	Appearances can deceive people
10	<i>bisa apa tu bocah kl jdi cawapres...</i>	Young age & survey	Medium	Diminutive, delegitimization	Assertive + Expressive	Not fit to be vice president; survey institute invalid
11	<i>satu arahnya dinasti... pemerintahan bobrok</i>	Political system	Heavy	Stigmatization, intense repetition	Assertive	Immoral dynasty government
12	<i>Jin setan iblis pun menangis...</i>	Politics (religious hyperbole)	Heavy	Hyperbole, negative metaphor	Expressive	The political situation is very bad/sad

13	<i>anak bau kencur</i>	Youth/leadership	Medium	Diminutive, age stereotype	Assertive	Young people are not yet fit to lead
14	<i>jalan aja susah prabowo patriot sejati</i>	Military/political figure	Light	Irony, conflict of meaning	Assertive	Prabowo's patriotism is in doubt
15	<i>MK = MAHKAMAH KELUARGA...</i>	Legal institution	Medium	Sarcastic redefinition	Assertive	MK is considered nepotistic
16	<i>sdh sprti kerajaan muka tembok...</i>	Political dynasty	Heavy	Monarchical metaphor, implicit insult	Expressive	Government like an authoritarian kingdom
17	<i>Wapresnya Jangan bocil caper...</i>	Youth	Light	Infantile labels, subtle criticism	Directive	Reject young candidates who are attention seeker
18	<i>keras ngomongnya ketemu planga-plongo</i>	Debate figures	Light	Contrast stereotypes, banter	Assertive	A satire of two incompetent candidates
19	<i>raja yg sdh loyo... tulul plongo naik tahta</i>	Old vs young	Heavy	Biological sarcasm, degradation	Assertive + Directive	Old leaders are just transitional tools
20	<i>tertipu lagi dgn wajah deso nya...</i>	Social appearance	Medium	Visual stereotypes, invitations	Directive	Don't be fooled by the simple face

Table 1 above reveals that hate speech is mostly directed at aspects of age, political power, physical appearance, and state institutions. The intensity of the speech ranges from mild to severe, with dominant linguistic characteristics in the form of satire, stereotypes, and aggressive metaphors. It proves that hate speech can be veiled and requires linguistic analysis to reveal its meaning accurately.

Table 2. Classification of Hate Speech Comments Based on Intensity Level

Intensity level	Comments/utterances	Target & Categories (Olsson)	Linguistic Features (Olsson)
Light	Comments 2	Political figures (family)	Mild vulgarity, expression of disgust
	Comments 8	Leader figure (speaking style)	Visual stereotypes, satire, light teasing
	Comments 14	Military/political figure	Irony, contrast of meaning
	Comments 17	Young figures	Label infantile, delicate imperative
	Comments 18	Debate figures (two presidential candidates)	Stereotypical contrast, visual mockery
Medium	Comments 1	Physical appearance (racism)	Irony, implied insults, derogatory labels
	Comments 3	Ideological (political opportunism)	Pejorative labels, delegitimizing expressions
	Comments 4	Legal institutions (nepotism)	Sarcasm, conspiratorial euphemism
	Comments 6	President (authoritarianism)	Animal metaphors, capitalization, satire

	Comments 9	Social appearance	Stereotypes, social euphemisms, sarcasm
	Comments 10	Young age & survey institutions	Diminutive, rhetorical question
	Comments 13	Young age (child of leader)	Local idioms, age stereotypes
	Comments 15	Legal institutions and young figures	Sarcastic re-labeling, negative personification
	Comments 20	Social appearance and political history	Facial stereotypes, political appeals
Heavy	Comments 5	Age & dynasty of power	Generalization, hyperbole, systemic criticism
	Comments 7	Personal (direct attack)	Explicit swearing, aggressive rhetoric
	Comments 11	Dynasty system of government	Stigmatization, repetition of insults
	Comments 12	Government (religious metaphor)	Emotional hyperbole, negative metaphor
	Comments 16	Political dynasty	Monarchy metaphor, implied swearing
	Comments 19	Old vs young leaders	Biological sarcasm, age and position degradation

Table 2 above shows the targets & categories referring to the target groups of hate speech and the types of discrimination: age, status, institution, appearance, ideology. Targets & categories are classified based on Olsson's theory (2012) within the framework of forensic linguistics. Intensity is determined based on emotional burden, the purpose of the attack, and the level of explicitness of the insults. Then, linguistic features show language strategies: from diminutives, visual stereotypes, explicit insults, to institutional sarcasm. In this case, linguistic features show the language techniques used to disguise or emphasize hate speech: irony, stereotypes, metaphors, to direct vulgarity.

2. Analysis of Implicature Meaning and Forensic Linguistic Relevance

The following analysis focuses on classifying types of speech acts based on Searle's theory (2011) and the disclosure of implicatures based on Grice's principles (1975) as shown in Table 3.

Table 3. Analysis of Implicature Meaning, Speech Acts and Forensic Linguistic Relevance

Comments	Implicature	Explanation of Implicature	Forensic Linguistics Relevance	Speech Acts
Comments 3	Conventional	<i>Kutu loncat</i> 'Jumping flea' implies betrayal	Explicit meaning as an accusation of character	Assertive
Comments 4	Conversation – Sarcastic	Comparing MA (Constitutional Court) to family → nepotism satire	Criticism of the legal system through metaphor; need for analysis of implied meaning	Assertive
Comments 5	Conventional + Conversational	Mixing facts and judgments → excessive criticism	Strong accusations and systemic judgment; potential violation of hate speech laws	Assertive

Comments 6	Conversation – Metaphorical	“Big Mouse” means controlling & saying a lot of things	Implies accusation of excessive power; important in analysis of motive and image	Assertive
Comments 9	Conversation – Visualization	“Tampang ndeso” implies visual deception	Accusing manipulation of public image with stereotypes; essential for framing analysis	Assertive
Comments 10	Conventional + Conversational	Calling out lies without evidence → intent to doubt legitimacy	Explicit & implicit delegitimization; relevant for identifying age bias	Assertive
Comments 11	Conventional	Express extreme opinions about the dynasty	Rough and systemic accusation; can be used in the analysis of the rhetoric of power	Assertive
Comments 13	Conversation – Age/low	Implicit criticism of age → attributing inappropriateness	Demonstrating age-based impropriety; important in discrimination cases	Assertive
Comments 14	Conversation – Irony	The irony of “true patriots” → questioning reputation	Indirectly criticizing a figure’s reputation	Assertive
Comments 15	Conversation – Sarcastic	<i>Mahkamah Keluarga</i> ‘Family Court’ implies nepotism	Re-labeling institutions and individuals; it is important to point out implicit insults	Assertive
Comments 16	Conversation – Metaphorical	“Kerajaan” = absolute power	Systematic critique in a royal style; important analysis of political perception	Assertive
Comments 19	Conversation – Dark Sarcasm	Death and ascendancy predictions → sarcastic	Inserting the intention to overthrow with biological metaphors; relevant for symbolic threat analysis	Assertive
Comments 18	Conversation – Satire Humor	Contrast between “keras” and “planga-plongo”	A mockery full of laughter to undermine the opponent’s authority	Expressive
Comments 12	Conversation – Hyperbole	Mythological reference → shows absurdity	Religious metaphors to strengthen denial; an important analysis of persuasive power	Expressive
Comments 1	Conversation – Sarcastic	Irony shows social satire	Implying insult through irony; essential for evidence of implicit meaning	Expressive
Comments 2	Conversation – Evaluative	Not explaining the reason → implied meaning disgusted	Showing the dislike; indicator of speaker’s emotional motivation	Expressive
Comments 17	Conversation – Soft Imperative	“Caper” → implies ulterior motives	Calls to stop performing; relevant for tracking social pressure	Directive

Comments 20	Conversation – Implicit Invitation	<i>Wajah desa</i> ‘simplicity face’ = symbol of deception	Calls not to vote disguised in stereotypes; relevant in digital political discourse	Directive
Comments 8	Conversation – Satire	Invitation to opt out → implies previous bad experiences	An implied invitation for the public to reject certain figures	Directive
Comments 7	Conventional	Rude rhetoric → doesn’t ask for an answer	Explicit swearing; direct evidence for strengthening of linguistic profile	Expressive/ Directive

Table 3 above highlights that assertive and expressive speech acts dominate hate Comments, indicating that social media users often express opinions, assessments, and negative emotional expressions towards individuals and institutions. Assertives, which were divided into 12 Comments, show “Statements,” “accusations,” claims,” or negative opinions.” Meanwhile, expressives divided into 5 Comments show personal emotions (disgust, anger, ridicule, frustration). In addition, directive speech acts in three Comments appear as a form of invitation or warning containing a veiled political message. The implicature side found that many Comments violate Grice’s principles, especially in the maxims of quality and relevance, to convey hidden meanings that indirectly strengthen hate sentiments. Conventional implicature: meaning is conveyed directly and can be interpreted without special context (e.g., explicit cursing). Conversational Implicature: requires context to understand the real intention, e.g., irony, sarcasm, metaphor, and implied invitation. From the perspective of forensic linguistic relevance analysis, it shows that linguistic features help reveal hidden meanings in hate speech, are useful in analyzing motives, profiling speakers, and potential law violations, and provide valid linguistic evidence for online conflict trials or mediation.

4.2 Discussion

The results of the analysis show that hate speech on social media is not only present in the form of direct insults, but also through veiled forms such as sarcasm, stereotypes, hyperbole, and aggressive metaphors. This finding is in line with Olsson (2012) who explained that hate speech has a distinctive linguistic structure such as diminutives *bocah*, *bocil* [kids], pejorative euphemisms *mahkamah keluarga* [family court] and institutional sarcasm as a form of symbolic degradation of certain parties. In terms of intensity, 6 out of 20 Comments are categorized as severe, which is characterized by explicit insults or character degradation (for example Comments 7 and 12). A total of 9 Comments are moderate, showing subtle mockery or stereotypes, and the rest are light, such as ordinary cynical expressions. This variation shows that hate speech on social media is very adaptive to context and audience, as revealed in a study by Sanguinetti et al. (2019) that hate speech experiences an expansion of meaning in digital space.

In the analysis of speech act types, most Comments fall into the assertive and expressive categories, supporting Searle’s (1979) findings that utterances in an ideological context often function to convey attitudes and assessments. Comments such as *anak kemarin sore bisa nyapres* [a child who was born yesterday can run for president]

or *kerajaan muka tembok* [a kingdom facing a wall] are forms of value assessment of the quality of leadership and the government system.

The directive type appears in the form of a subtle warning or invitation, such as in the Comments *planga-plongo jilid 2* [planga-plongo volume 2] which implicitly invites the public not to repeat their previous political choices. In the context of hate speech, this has the potential to form polarization and mass opinion in a veiled manner.

From a pragmatic perspective, Grice's (1975) theory helps reveal the implicatures hidden behind Comments. Violations of the maxim of quality (hyperbole) and the maxim of relevance (exaggerated assumptions) produce implied meanings such as calls for boycotts, accusations of corruption, or age stigmatization. For example, the Comments *tulul plongo naik tahta* imply a narrative that the leader is only a puppet of the old power.

This discussion is reinforced by research by Yeagley (2020) and Hargreaves (2022) which shows that political language on social media often hijacks neutral terms into veiled insults. Words such as "snowflake" or "woke" in their study are in line with the findings on terms such as "ndeso," "bocil," or "bau kencur."

The implications of this study indicate that the forensic linguistic approach is very effective in uncovering forms of hate speech that are veiled through language strategies such as sarcasm, stereotypes, and euphemisms. Theoretically, these findings enrich the study of pragmatic linguistics and speech acts by showing the importance of implicature analysis in detecting non-explicit hate intentions. Practically, the results of this study can be utilized by law enforcement, digital platform providers, and policy makers to develop a hate speech detection system that is more sensitive to context, not just keywords. The social implication is the importance of language literacy in society to understand that hate speech can be present in subtle forms, but still has the potential to be systemically damaging

5. Conclusion

This study This study examines hate speech on social media using a forensic linguistic approach that combines Olsson's (2012) classification of hate speech, Searle's (2011) speech act theory, and Grice's (1975) implicature principle. The analysis of 20 Comments shows that hate speech is not always explicit and harsh. However, subtle satire, visual stereotypes, aggressive metaphors, and sarcastic euphemisms often express it. The main finding of this study is that hate speech on social media tends to avoid direct vulgar linguistic forms and instead relies on implicit language strategies to convey negative sentiments towards individuals and groups. There is the widespread use of diminutives such as "bocah" or "bocil," social stigmatization such as "ndeso" and "bau kencur," and institutional euphemisms such as *mahkamah keluarga* [family court], all of which are important linguistic indicators in the analysis of political hatred in the digital space. In addition, the intensity classification shows that moderate and severe hate speech comes not only from explicit insults but also from narratives that normalize harassment based on age, appearance, and political affiliation. Pragmatically, violations of the

maxims of quality and relevance in Grice's theory are prominent features because they convey hidden messages that actually strengthen the communicative effects of hatred.

Conflict of interest

None

Authors' contribution

The authors made substantial contributions to the conception and design of the study. The authors took responsibility for data analysis, interpretation and discussion of results. The authors read and approved the final manuscript.

References

- Al-Saqqa, S., Awajan, A., & Hammo, B. (2024). A Survey of Hate Speech Detection for Arabic Social Media: Methods and Datasets. *Procedia Computer Science*, 251, 224–231. <https://doi.org/10.1016/j.procs.2024.11.104>
- Alduais, A., Al-Khulaidi, M. A., Allegretta, S., & Abdulkhalek, M. M. (2023). Forensic linguistics: A scientometric review. *Cogent Arts and Humanities*, 10(1). <https://doi.org/10.1080/23311983.2023.2214387>
- Ali, J. H. M. (2020). Forensic Linguistics: A Study in Criminal Speech Acts. *Beni-Suef University International Journal of Humanities and Social Sciences*, 2(1), 39–65. <https://doi.org/10.21608/buijhs.2020.91333>
- Alsagheer, D., Mansourifar, H., & Shi, W. (2023). Statistical Analysis of Counter-Hate Speech on Voice-based Social Media. *Procedia Computer Science*, 220(2019), 1009–1014. <https://doi.org/10.1016/j.procs.2023.03.140>
- Aulia, N. & F. K. L. (2022). Hate Speech Comments on Joe Biden'S Instagram Posts: Pragmatic Analysis. ... *Journal of English Language Teaching and ...*, 11(1), 12–20.
- Austin, J. L. (1962). *How to Do Things With Words*. Oxford: University Press.
- Borg, E. (2022). Implicatures. *Linguistics Meets Philosophy*, 441–469. <https://doi.org/10.1017/9781108766401.024>
- Coulthard, Malcolm. (2016). *An Introduction to Forensic Linguistics. An Introduction to Forensic Linguistics*. <https://doi.org/10.4324/9781315630311>
- Coulthard, Malcolm, & Johnson, A. (2011). *An Introduction to Forensic Linguistics: Language in Evidence*. London: Routledge Taylor & Francis Group. <https://doi.org/10.1201/b11275-2>
- Dewi, E. R., Lustyantje, N., & Isnain, F. N. (2021). The Needs of Technology Transformation in French Literature Learning. *Ijlecr - International Journal of Language Education and Culture Review*, 7(1), 59–68. <https://doi.org/10.21009/ijlecr.071.05>
- Finegan, E., & Lee, B. T. (2020). Legal writing: attitude and emphasis. *The Routledge Handbook of Forensic Linguistics*, 48–63. <https://doi.org/10.4324/9780429030581-6>
- Fonseca, A., Pontes, C., Moro, S., Batista, F., Ribeiro, R., Guerra, R., ... Silva, C. (2024). Analyzing hate speech dynamics on Twitter/X: Insights from conversational data and the impact of user interaction patterns. *Heliyon*, 10(11), 1–10. <https://doi.org/10.1016/j.heliyon.2024.e32246>
- Ghaly, R., Elkorany, A., & Ezzat, C. A. (2024). Hate Speech Detection in Arabic Text: Survey.

- Procedia Computer Science*, 244, 166–177. <https://doi.org/10.1016/j.procs.2024.10.222>
- Ghenai, A., Noorian, Z., Moradisani, H., Abadeh, P., Erentzen, C., & Zarrinkalam, F. (2025). Exploring hate speech dynamics: The emotional, linguistic, and thematic impact on social media users. *Information Processing and Management*, 62(3), 1–28. <https://doi.org/10.1016/j.ipm.2025.104079>
- Gibbons, J. (2003). *Forensic Linguistics: An Introduction to Language in the Justice System*. Hoboken: Wiley-Blackwell.
- Grice, H. P. (1975). Logic and Conversation. In Cole and JL Morgan (Ed.), *Syntax and Semantic Vol 3: Speech Act*. New York: Academy Press.
- Gyuró, M. (2013). Legal discourse, power and pragmatics. *Discourse and Interaction*, 6(2), 5–13. <https://doi.org/10.5817/DI2013-2-5>
- Harahap, Y., & Siregar, N. (2022). A Forensic linguistic Approach to interpreting a hate speech case on social media. *Vision*, 18(1), 58–65. Diambil dari <http://dx.doi.org/10.30829/vis.v18i1.1414>
- Hargreaves, B. (2022). *The Appropriation of the Term “woke” by the Right*. Berkeley: Berkeley Center for Right-Wing Studies, University of California.
- Hasibuan, A., Ningsih, A. M., & Wahyuni, S. (2023). Metaphors in Minangkabau Language: Meaning Shift and Change. *International Journal of Research and Review*, 10(1), 520–527. <https://doi.org/10.52403/ijrr.20230160>
- Ibrohim, M. O., & Budi, I. (2023). Hate speech and abusive language detection in Indonesian social media: Progress and challenges. *Heliyon*, 9(8), e18647. <https://doi.org/10.1016/j.heliyon.2023.e18647>
- Jay, T. (2009). Do Offensive Words Harm People? *Psychology, Public Policy, and Law*, 15(2), 81–101
- Kryk-Kastovsky, B. (2006). Legal Pragmatics. In K. B. T.-E. of L. & L. (Second E. Brown (Ed.) (hal. 13–20). Oxford: Elsevier. <https://doi.org/https://doi.org/10.1016/B0-08-044854-2/04720-9>
- Kusumawardhani, P. (2024). Language and Forensic Linguistics. *International Journal of English and Applied Linguistics (IJEAL)*, 4(1), 48–53. Diambil dari <https://doi.org/10.47709/ijeal.v4i1.3706>
- Liswahyuningsih, N. L. G., Artawa, K., & Udayana, I. N. (2020). Form and Meaning Shifts in Indonesian-English Translation of Noun Phrases. *AWEJ for Translation & Literary Studies*, 4(2), 247–262.
- Mafeza, F. (2016). Preventing Genocide by Fighting Against Hate Speech. *International Journal of Advanced Research*, 4(3), 117–132.
- Manne, K. (2017). *Down Girl: The Logic of Misogyny*. Oxford: Oxford University Press
- McMenamin, G. R. (2008). *Forensic Linguistics: Advances in Forensic Stylistics. The Handbook of Linguistics*. New York: CRC Press. <https://doi.org/10.1002/9780470756409.ch30>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (3 ed.). United States of America: SAGE Publications, Inc.
- Mintowati, M., & Dasion, H. Y. T. (2020). Hate Speech: Forensic Linguistics Study. In *3rd Social Sciences, Humanities, and Education Conference (SoSHEC 2019)* (Vol. 380, hal. 266–268). Surabaya: Atlantis Press. <https://doi.org/10.2991/soshec-19.2019.59>
- Mubarok, Y., Sudana, D., Yanti, D., Sugiyo, Aisyah, A. D., & Af'idah, A. N. (2024). Abusive Comments (Hate Speech) on Indonesian Social Media: A Forensic Linguistics Approach. *Theory and Practice in Language Studies*, 14(5), 1440–1449.

- <https://doi.org/10.17507/tpls.1405.16>
- Nave, E., & Lane, L. (2023). Countering online hate speech: How does human rights due diligence impact terms of service? *Computer Law and Security Review*, 51(September), 105884. <https://doi.org/10.1016/j.clsr.2023.105884>
- Olsson, J. (2012). *Wordcrime: Solving Crime through Forensic Linguistics*. London: Continuum International Publishing Group.
- Perkins, R. C. (2018). The Application of Forensic Linguistics in Cybercrime Investigations. *Policing: A Journal of Policy and Practice*, 15(1), 68–78. <https://doi.org/10.1093/police/pay097>
- Putra, V. G. R., Febriyatko, A., & Busri, H. (2024). Hate Speech in Instagram Comments of DPR RI: A Forensic Linguistic Analysis. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya*, 10(1), 173–196. <https://doi.org/10.22219/kembara.v10i1.28128>
- Ramli, R. B., Puspita, L. E., & Sauhenda, A. F. (2024). Language Crime Practices: Cyberbullying in the TikTok Comment Section (Review) – A Forensic Linguistics Analysis. *Musamus Journal of Language and Literature*, 7(1), 288–295. <https://doi.org/10.35724/mujolali.v7i1.6241>
- Ria, R. N., & Setiawan, T. (2023). Forensic Linguistic Analysis of Netizens' Hate Speech Acts in Tik-Tok Comment Section. *Britain International of Linguistics Arts and Education (BIO LAE) Journal*, 5(2), 141–152. <https://doi.org/10.33258/biolae.v5i2.894>
- Riadi, I., Herman, H., & Rafiq, I. A. (2022). Mobile Forensic Investigation of Fake News Cases on Instagram Applications with Digital Forensics Research Workshop Framework. *International Journal of Artificial Intelligence Research*, 6(2). <https://doi.org/10.29099/ijair.v6i2.311>
- Sanguinetti, M., Poletto, F., Bosco, C., Patti, V., & Stranisci, M. (2019). An Italian twitter corpus of hate speech against immigrants. *LREC 2018 - 11th International Conference on Language Resources and Evaluation*, 2798–2805
- Sumiati, A., Lustyantje, N., & Iskandar, I. (2020). Integrating 21st Century Skills Into Translation Classroom: a Brief Perspective on Its Syllabus. *Ijlecr - International Journal of Language Education and Culture Review*, 6(2), 87–96. <https://doi.org/10.21009/ijlecr.062.11>
- Sakakini, adel. (2020). Forensic Linguistics: an Applied Theory. *BAU Journal - Society, Culture and Human Behavior*, 1(2), 1–10. <https://doi.org/10.54729/2789-8296.1022>
- Searle, J. R. (2011). *Speech Acts: An Essay in The Philosophy of Language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781107415324.004>
- Singer, P. W., & Brooking, E. T. (2018). *LikeWar: The Weaponization of Social Media*. Boston: Houghton Mifflin Harcourt.
- Subramanian, M., Easwaramoorthy Sathiskumar, V., Deepalakshmi, G., Cho, J., & Manikandan, G. (2023). A survey on hate speech detection and sentiment analysis using machine learning and deep learning models. *Alexandria Engineering Journal*, 80(May), 110–121. <https://doi.org/10.1016/j.aej.2023.08.038>
- Teh, P. L., Cheng, C. Bin, & Chee, W. M. (2018). Identifying and categorising profane words in hate speech. *ACM International Conference Proceeding Series*, 65–69. <https://doi.org/10.1145/3193077.3193078>
- Tredici, M. Del, Fernández, R., & Boleda, G. (2019). Short-term meaning shift: A distributional exploration. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 1, 2069–2075. <https://doi.org/10.18653/v1/n19-1210>

- Umiyati, M. (2020). A Literature Review of Forensic Linguistics. *IJFL (International Journal of Forensic Linguistics)*, 1(1), 23–29.
- van Dijk, T. A. (2008). *Discourse and Power*. Michigan: Macmillan Education UK.
- Varney, M. H. (1997). Forensic Linguistics: A New, Varied and very Practical Application of Linguistic Science. *English Today*, 13(4), 42–47.
<https://doi.org/10.1017/S0266078400010014>
- Wasi, M., Zaki, S., & Huma Anwar, R. (2024). A Bibliometric Analysis of Research in Forensic Linguistics. *SAGE Open*, 14(4), 1–18. <https://doi.org/10.1177/21582440241299280>
- Wright, D. (2021). Corpus approaches to forensic linguistics: Applying corpus data and techniques in forensic contexts. In M. Coulthard, A. May, & R. Sousa-Silva (Ed.), *The Routledge Handbook of Forensic Linguistics* (2 ed., hal. 611–627). London: Routledge.
- Yeagley, D. S. (2020). The Snowflake Generation: A Pejorative Term to Silence and Discredit Dissent. *Journal of Language and Social Psychology*, 39(1), 82–96
- Yuliyanti, T., Subyantoro, S., & Pristiwati, R. (2020). Form of Hate Speech Comments on Najwa Shihab Youtube Channels in The General Election Campaign of President and Vice President of The Republic of Indonesia 2019. *Seloka: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 9(3), 254–263. <https://doi.org/10.15294/seloka.v9i3.42513>

Appendix

1. Classification of Comments Obtained from Instagram through Hashtag Search (Primer Data)

#hashtag search via Instagram URL	Comments/utterances
#gibran	1. <i>ksian di hina mulu si monyong</i> [poor thing, monyong always being insulted]
	2. <i>Anjir males bngt denger si gibran dan si kaesang</i> [Damn, I really hate listening to Gibran and Kaesang]
	3. <i>Dasar Kutu loncat</i> [Bastard, jumping flea]
	4. <i>Efek adanya MAHKAMAH KELUARGA, anak kemaren sore udah bisa ikut nyapres, walaupun tanpa jejak pengalaman yg mumpuni</i> [The effect of the FAMILY COURT, yesterday's child can now run for president, even without any experience.]
	5. <i>Ini lah hasil nya anak muda, yg tua kelakuan dinasti sangat rakus kekuasaan, keluarga mk perlu di telusuri</i> [This is the result of young people, the old dynasty's behavior is very greedy for power, the mk family needs to be investigated]
#gibranrakabumingraka	6. <i>Pak Jokowi BIG MOUSE NYA INDONESIA. INDONESIA AKAN DIKUASAI BAPAK JOKOWI DITAHUN 2030 nanti</i> [Mr. Jokowi is INDONESIA'S BIG MOUSE. INDONESIA WILL BE CONTROLLED BY MR. JOKOWI IN 2030]
	7. <i>emangnya dia pikir dia ni siapa sih anying</i> [Who does he think he is?]
	8. <i>Plis lah jangan smpe planga plongo jilid 2</i> [Please don't allow planga plongo volume 2 again]
	9. <i>Pda byk ketipu tampang ndeso, sederhana n Kya merakyat ya?</i> [Many people are fooled by his rustic, simple and popular appearance, right?]
	10. <i>Jiaahh bisa apa tu bocah kl jdi cawapres...semua surve cari nafkah dengan cara berbohong tidak patut sebagai acuan sesuatu lembaga yg kerjanya ngibul</i> [Jiaahh what can that kid do to become vice presidential candidate? All surveys about making a living by lying are not worthy of being used as a reference by an institution whose job is to lie.]
	11. <i>Intinya satu arahnya dinasti pemerintahan!!! Dan gobloknya dinasti itu pemerintahan yg bobrok dan tidak tau malu!!!</i> [The point is that the government is in one direction, the dynasty!!! And the stupid thing about the dynasty is that it is a government that is rotten and shameless!!!]
	12. <i>Jin setan iblis pun menangis...disuruh pensiun dini</i> [Even the devil and the demon cried....told to retire early]
	13. <i>Kalo prabowo minang gibran jadi wakilnya, fix partai yg ikut berkoalisi pada gak punya malu. Masa sekelas ketua partai kalah sama anak bau kencur</i> [If Prabowo taken Gibran becomes his vice (president), the parties that are in the coalition will definitely have no shame. How can a party leader lose to a kid who smells of ginger?]
	14. <i>jalan aja susah prabowo dikatain patriot sejati</i> [Prabowo is called a true patriot even though he has difficulty walking]

	<p>15. <i>MK = MAHKAMAH KELUARGA..dan sicungkring gak tau punya malu soal ini</i> [MK = FAMILY COURT..and sicungkring has no shame about this]</p>
	<p>16. <i>Gilaaaaaa SDH sprti kerajaan muka TEMBOK gak tau malu</i> [Crazy it's like a wall-faced country with no shame]</p>
	<p>17. <i>Wapresnya Jangan bocil caper lah pliss</i> [Vice President, please don't be a jerk.]</p>
	<p>18. <i>Yang keras ngomongnya ketemu yang planga-plongo hahaha LOL!!!!</i> [Who talk loudly meet those who are planga-plongo (blank face) hahaha LOL!!!!]</p>
	<p>19. <i>Makanya di cari raja yg sdh loyo tua renta biar cpt k0lt..dan si tulul plongo naik tahta</i> [That's why they're looking for a king who is old and frail so he can quickly dead... and the stupid fool will ascend the throne]</p>
	<p>20. <i>Kita jangan tertipu lagi dengan wajah deso nya, sdh cukup 2 priode kita sengsara....</i>[We should not be fooled again by his country bumpkin face, we have suffered enough for two periods]</p>