Tersedia secara online
http://ejournal.iainmadura.ac.id/ghancaran

Ghancaran: Jurnal Pendidikan
Bahasa dan Sastra Indonesia
Volume 7, Nomor 2, 2026
Halaman: 643-663

GHÂNCARAN: JURNAL PENDIDIKAN BAHASA DAN SASTRA INDONESIA
http://ejournal.iainmadura.ac.id/ghancaran
E-ISSN: 2715-9132; P-ISSN: 2714-8955
DOI *10.19105/ghancaran.v7i2.21949*

The Discrusive Power of Verbal Violence in Indonesian Social Media: A Linguistic Corpus Study

Imanuella Anastasia Brigitha Carneliaputri Dewanty*, Djoko Saryono*, Azizatus Zahro*, Zainur Rijal Abdul Razak, Habibah Ismail****
*Pendidikan Bahasa Indonesia, Universitas Negeri Malang, Indonesia
**Fakulti Pengajian Bahasa Utama, Universiti Sains Islam Malaysia, Malaysia
Alamat surel: imanuella.anastasia.2402118@students.um.ac.id

Keywords:
Digital verbal violence; Power reproduction; Critical discourse analysis.

Abstract

The rise of gender-based, political, and identity-based verbal violence on social media highlights the urgency of understanding how language plays a role in the reproduction of power in digital culture. This study aims to analyze the linguistic and ideological structures that shape practices of verbal violence in digital interactions. Data were collected from July 2024 to January 2025 from various social media platforms, then analyzed using corpus linguistics to identify patterns of language use, as well as Fairclough's Critical Discourse Analysis to interpret the power relations that emerge in the texts. The results show that digital verbal violence is not merely an expression of individual emotion, but is connected to dominant ideologies such as patriarchy, toxic nationalism, and religious fanaticism. Violent speech contains both vertical and horizontal power relations, which allow users to assert symbolic authority, silence others, or negotiate certain identities. These findings confirm that verbal violence in digital spaces is not merely a matter of communication ethics, but part of the mechanism of power reproduction through linguistic practices. In conclusion, this study contributes to understanding how language in social media reproduces power and ideology in digital interactions. As a follow-up, future research may expand the data scope or examine counter-discursive strategies to challenge verbal violence online. The findings have implications for critical digital literacy by helping users recognize and resist ideological domination embedded in everyday social media discourse.

Kata Kunci:
Kekerasan verbal digital; Reproduksi kekuasaan; Analisis wacana kritis.

Abstrak

Meningkatnya kekerasan verbal berbasis gender, politik, dan identitas di media sosial menunjukkan urgensi memahami bagaimana bahasa berperan dalam reproduksi kekuasaan pada budaya digital. Penelitian ini bertujuan menganalisis struktur linguistik dan ideologis yang membentuk praktik kekerasan verbal dalam interaksi digital. Data dikumpulkan pada Juli 2024–Januari 2025 dari berbagai platform media sosial, kemudian dianalisis menggunakan linguistik korpus untuk mengidentifikasi pola penggunaan bahasa, serta Analisis Diskursus Kritis Fairclough untuk menafsirkan relasi kuasa yang muncul dalam teks. Hasil penelitian menunjukkan bahwa kekerasan verbal digital tidak hanya merupakan ekspresi emosional individual, tetapi terhubung dengan ideologi dominan seperti patriarki, nasionalisme toksik, dan fanatisme agama. Ucapan kekerasan mengandung relasi kuasa vertikal maupun horizontal, yang memungkinkan pengguna menegaskan otoritas simbolis, membungkam pihak lain, atau menegosiasikan identitas tertentu. Temuan ini menegaskan bahwa kekerasan verbal di ruang digital bukan semata persoalan etika komunikasi, melainkan bagian dari mekanisme reproduksi kekuasaan melalui praktik linguistik. Kesimpulannya, studi ini

berkontribusi pada pemahaman mengenai bagaimana bahasa dalam media sosial membentuk dan mempertahankan struktur ideologis yang mempengaruhi relasi sosial di era digital.

Dikirim: 15 September 2025; Revisi: 10 Januari 2026; Diterbitkan: 12 Januari 2026

©Ghâncaran: Jurnal Pendidikan Bahasa dan Sastra Indonesia
Tadris Bahasa Indonesia
Institut Agama Islam Negeri Madura, Indonesia

INTRODUCTION

The transformation of information technology has reshaped human communication practices while simultaneously expanding the space for verbal violence within Indonesia's social media ecosystem. Digital platforms that function as new public spaces, such as Twitter, Instagram, TikTok, and others, not only facilitate the exchange of information but also serve as arenas for the production and circulation of hate speech, insults, and identity-based attacks (Ariibah & Zhakiyyah, 2023). Indonesia's digital culture, which tends to be permissive toward verbal aggression—characterized by the harshness of netizen comments and the normalization of rudeness framed as humor or spontaneous expression (Naco, 2019) underscores the urgency of linguistic research into the patterns and mechanisms of verbal violence in online interactions.

The characteristics of social media such as anonymity, rapid dissemination, and the fluidity of identity shift the boundaries of communication ethics and encourage aggressive forms of expression (Wang, 2013). Digital verbal violence manifests in various forms, including insults, harassment, threats, ideological attacks, and the misuse of personal information (Akhvlediani & Moralishvili, 2021). At the same time, national regulatory frameworks, such as the ITE Law and the Criminal Code, have yet to provide adequate protection for victims due to limitations in interpretation and the potential misuse of legal enforcement mechanisms. This condition underscores the need for an academic approach capable of uncovering the linguistic structures and power relations that shape and sustain verbal violence in digital spaces.

Theoretically, this study is grounded in Critical Discourse Analysis (CDA) and the theory of language and power, which conceptualize language as a social practice that both shapes and is shaped by ideology and power relations (Seda et al., 2025; Saraswati et al., 2025). Within the context of Indonesian social media, verbal violence often embodies ideological orientations such as patriarchy, toxic nationalism, and religious fundamentalism, which are articulated through linguistic patterns that can be systematically traced. The CDA framework enables an examination of how digital discourse functions as a site for negotiating identity, enacting resistance, or reinforcing subordination.

Previous studies can be classified into three main strands. First, international research has examined social bias and comment toxicity (Cho & Moon, 2021), the gap between legal definitions and the linguistic realization of hate speech (Lepoutre et al., 2024), as well as annotation constraints in NLP-based corpora (Poletto et al., 2020). However, these studies tend to overlook an in-depth analysis of linguistic patterns underlying verbal violence. Second, studies conducted in the Indonesian context have identified patriarchal patterns in verbal violence on TikTok (Dewanty & Saryono, 2024) and highlighted the limitations of manual techniques in detecting hate speech (Ibrohim & Budi, 2023), yet they have not comprehensively employed corpus-based approaches. Third, research situated in the domains of education and gender (Eliasson et al., 2005; Lau et al., 2021; Yusri et al., 2024) has successfully documented various forms of verbal violence, but has not explicitly connected these findings to empirically measurable linguistic structures (Fatim et al., 2024).

From this review, it is evident that no previous study has specifically mapped the linguistic patterns of verbal violence on Indonesian social media through an integrated approach combining corpus linguistics and Critical Discourse Analysis. Therefore, this study addresses a significant gap in the existing literature.

This study aims to identify and analyze the linguistic patterns of verbal violence in Indonesian social media interactions by integrating a corpus-based approach with the Critical Discourse Analysis (CDA) framework. The analysis focuses on the forms, frequencies, collocational patterns, and ideological contexts that produce digital verbal violence, as reflected in recurrent expressions such as gendered insults, dehumanizing labels, religiously framed accusations, and nationalist derogatory terms commonly used in online debates.

The contribution of this research is twofold. Theoretically, it extends the study of language–power relations in digital contexts by mapping linguistic patterns through an objective and replicable corpus-based approach. Practically, the findings provide an empirical foundation for the development of critical digital literacy initiatives and hate speech detection systems that are sensitive to local sociocultural contexts.

Thus, this research positions digital verbal violence not merely as a communicative phenomenon, but as a linguistic practice that actively reproduces power structures within Indonesian digital culture. This perspective affirms the role of linguistics not only as an analytical discipline, but also as an interventional one capable of contributing to social critique and transformation.

METHOD

This research employed a descriptive qualitative design with a corpus-based orientation. The descriptive qualitative approach was chosen to enable an in-depth examination of linguistic phenomena as they naturally occur in digital interactions, particularly forms of verbal violence in social media discourse. The corpus-based orientation allowed for the systematic identification and mapping of recurring language patterns, ensuring that the analysis was grounded in empirical linguistic evidence rather than subjective or anecdotal interpretations.

This study integrated corpus linguistics techniques with Fairclough's Critical Discourse Analysis (CDA). Corpus linguistics was employed to identify patterns of frequency, collocation, and lexical association related to verbal violence in an objective and replicable manner. CDA functioned as the interpretive framework to examine how these linguistic patterns operate within broader discursive and social contexts. The analysis focused on three interrelated levels—textual features, discursive practices, and social practices—to reveal the ideological structures and power relations reproduced through digital discourse.

The data sources consisted of publicly accessible written texts in the form of captions, comments, and hashtags collected from Twitter (X), Instagram, and TikTok. These platforms were selected due to their high levels of user interaction and the frequent emergence of public debates. The study did not involve human participants directly and did not collect personal or private data. Accordingly, private messages, closed-group content, and multimodal elements were excluded to maintain ethical standards and analytical consistency.

Data collection employed purposive sampling based on predefined keywords related to verbal violence, including expressions associated with gender, political identity, religion, and nationalism. Text extraction was conducted using Python-based natural language processing (NLP) tools. The initial corpus comprised approximately 120,000 words prior to data cleaning. To ensure consistency and reliability, all visual and audio elements, as well as textual content embedded in images, were excluded from the corpus.

The research procedures involved several stages: keyword-based identification of relevant posts, text extraction, removal of duplicate entries and non-linguistic symbols, and preliminary categorization of verbal violence. The cleaned corpus was then analyzed using AntConc software through keyword analysis, N-gram profiling, collocation analysis, and concordance examination to identify recurrent linguistic patterns. The corpus findings were subsequently interpreted using Fairclough's CDA framework to explain how verbal

violence functions as a discursive practice that reproduces power relations in digital spaces.

Research validity was ensured through cross-platform source triangulation to compare linguistic patterns across different social media platforms, as well as analyst triangulation involving two researchers during the coding and interpretive processes. An audit trail was maintained to document all stages of data extraction, annotation, and analysis. Ethical considerations were addressed by restricting the data to publicly available content and anonymizing any identifiable user information in the analysis.

RESULTS AND DISCUSSION

The results and discussion of this study are presented as follows.

Overview of Keyword Analysis

From a total of 60,578 extracted keywords, the top 100 positive keywords were selected for further analysis. Positive keywords are defined as lexical items that occur with significantly higher frequency in the target corpus—the *Verbal Violence in Social Media Corpus*—than in the reference corpus, *Indonesian Web 2024 (idTenTen24)*. This statistical salience indicates that these lexical items play a central role in the discourse of verbal violence on social media (Culpeper & Demmen, 2015).

The selected keywords were subsequently organized into thematic categories adapted from the Five Contextual Types of Harassment framework. These categories were used to classify patterns of verbal violence based on their communicative and ideological functions. Table 1 presents the distribution of keywords across the thematic categories, with keywords in each category ranked according to their position in the keyword list and their raw frequency of occurrence.

No.	Types of Verbal Violence	Keywords (Rank/Frequency)
1.	Profanity (21 keywords)	anjeng (2/6,238), bangsat (3/9,225), ngentot (6/5,435), kontol (12/7,794), ajg (36/385), anjg (39/339), anjir (42/530), bacot (43/417), anj (50/272), kampret (63/316), taik (64/179), asu (66/250), tai (67/713), cok (70/398), bajingan (72/379), anying (75/139), mampus (89/220), bgst (97/110), bjir (98/105), njir (99/115), jir (100/114)
2.	Sexual (11 keywords)	lonte (8/1,910), tobrut (13/1,216), pulen (16/1,959), perek (24/624), boti (29/560), pepek (34/389), iclik (35/367), murahan (37/1.387), topita (38/343), ewe (41/321), memek (51/774)
3.	Intelligence (8 keywords)	cogil (10/1.525), goblok (11/3.120), cegil (14/1.172), tolol (20/2.378), bloon (27/575), dongo (65/172), bego (78/286), dungu (96/203)
4.	Political (8 keywords)	fufufafa (5/3.243), kadrun (7/3.071), cebong (17/1.757), pansos (18/1.325), buzzer (23/2.396), pencitraan (25/3.759), mulyono (28/2.439), penjilat (60/236)

5.	Appearance (2 keywords)	nuruls (26/533), banci (44/585)
6.	Racial (2 keywords)	kampungan (15/1.802), kafir (82/2.228)

Table 1. Keyword Analysis of the Verbal Violence Corpus on Social Media

Verbal Violence Keywords

The corpus documentation used in this study is fully accessible at https://bit.um.ac.id/fG6Q5Qb5NN, with additional access to the Sketch Engine–hosted *Verbal Violence Corpus* available upon request. From this dataset, one or two highly frequent and representative keywords from each category of verbal violence were selected for in-depth analysis, based on their frequency, contextual salience, and ideological relevance within social media discourse.

Profanity

The analysis of profanity focuses on the lexemes *anjeng* and *bangsat*, both of which occur with high frequency in the corpus (6,238 and 9,225 tokens, respectively). A salient linguistic feature emerging from their usage is syntactic transformation, particularly adjectivalization. Collocational patterns such as *bangsat banget*, *banget bangsat*, *banget anjeng*, and *bgt anjeng* indicate that lexemes originally functioning as nouns undergo a functional shift into evaluative predicates, operating in a manner similar to adjectives that intensify the speaker's stance. This shift reflects an ongoing grammaticalization process driven by the expressive demands of digital communication.

In addition to adjectivalization, the data also reveal a process of interjectionalization. Constructions such as *ya bangsat*, *ya anjeng*, and forms combined with pragmatic particles (e.g., *lah*, *sih*) demonstrate that these lexemes function as affective exclamations largely detached from their original referential meanings. These usages position profanity as a pragmatic resource for emotional expression rather than solely as a referential insult.

Both processes underscore that verbal violence in online contexts is not merely offensive behavior but also a form of creative lexical refunctionalization shaped by digital spontaneity and emotional immediacy. These transformations align with broader theoretical perspectives on grammaticalization (Heine & Kuteva, 2007) and pragmaticalization (Diewald, 2011), whereby lexical items acquire new structural and pragmatic functions over time.

The findings further support the argument that social media environments foster linguistic hybridity, enabling vulgar lexicon to operate as expressive, evaluative, and stance-marking devices that become normalized through repetition, memetic circulation, and platform affordances. Consequently, profanity in Indonesian digital discourse indexes affective stance, social alignment, and ideological evaluation, revealing a complex interplay between aggression, creativity, and identity performance.

Sexual

This study aims to foreground a crucial aspect within the domain of sexual verbal aggression. Certain lexemes inherently carry negative connotations—for instance, *iclik* and *ewe* (referring to sexual acts), *memek* and *pepek* (referring to female genitalia), and *perek* and *lonte* (referring to sex workers). However, lexical negativity alone does not automatically render a word an act of verbal violence. What transforms these terms into instruments of aggression is their *deployment within sexually charged discourse* with the intent to demean, shame, or objectify individuals (Felmlee et al., 2020; Dehingia et al., 2023). In such contexts, these words are *strategically mobilized not for neutral description*, but to assert dominance, mock, or exert control. Hence, the violence lies not within the word per se, but in the *discursive function* it fulfills within *misogynistic and patriarchal communicative practices* (Marwick & Caplan, 2018).

From a discourse perspective, the repeated use of *lonte* serves not only as an individual insult but also as a performative act that reproduces patriarchal and moralistic worldviews. The term taps into cultural scripts that regulate women's behavior and sexual expression, marking them as moral deviance. In critical discourse analysis (Fairclough, 1992), this lexeme operates on three levels: textually, it appears with high emotional intensity and vulgarity; discursively, it is embedded in practices of moral judgment and shaming; socially, it contributes to the reproduction of misogynistic ideology in digital public spheres. As such, *lonte* is not merely a vulgar label but a discursive tool systematically used to assert dominance, express hatred, and regulate gender norms in online interactions. Its frequent appearance within syntactic and emotional clusters of verbal violence affirms its status as a keyword in the lexicon of digital misogyny.

At the level of socio-cultural practice, the use of sexual and gendered lexemes reveals underlying patriarchal ideologies and normative gender relations deeply embedded in Indonesia's digital culture. Many lexemes classified under sexual verbal violence stem from social contexts already marked by unequal power relations and conservative moral norms. In both popular culture and everyday speech, these words

arise from collective practices that reproduce gender stereotypes and justify the surveillance of women's bodies and expressions, as well as those of non-dominant gender groups.

Within this context, social media can be understood as a new institution with its own discursive conventions. As with institutions like courts, schools, or hospitals—which possess their own rules, roles, and discursive forms (Fairclough, 1995; Roberts, 2011; Freed, 2015)—social media constructs and regulates user behavior and language use through implicit norms, technological affordances, and the dynamics of online interaction. Unlike face-to-face interaction, where ethical norms such as politeness, empathy, and personal accountability are upheld due to direct contact and higher social risk, social media—due to anonymity, physical distance, and screen mediation—blurs ethical boundaries (Puspitasari, 2022; Crystal, 2011; Wang et al., 2012). This fosters the release of internal censorship and heightens the expression of violent, sarcastic, and discriminatory language (Budiawan, 2024).

Languages that might not be uttered in offline spaces due to ethical norms or shame appear with high intensity on social media. This phenomenon characterizes social media as a discursive institution that enables more vulgar, extreme, and open utterances, forming a discourse landscape distinct from other social institutions (Crystal, 2011; Mayr, 2015; Graham, 2005; Jones & Hafner, 2021). Social media, through algorithms, comment sections, and virality features, not only facilitates the circulation of hate speech but also legitimizes certain speech styles—such as sarcasm, irony, or sexual slurs—as part of a 'discourse habit' considered normal, humorous, or even authentic (Mayr, 2015).

Words such as *perek*, *lonte*, or *boti*, which might be avoided in face-to-face situations due to taboo or social risk, circulate widely in meme formats, comment sections, and viral content. When collocated with personal referents like *kau* or *lu*, these words function not only as insults but as what Fairclough (1995) terms *ideological-discursive formations*—discourse structures that reflect and reinforce unequal power relations. In this sense, social media is not neutral—it actively participates in reinforcing discursive structures that normalize verbal violence and symbolic exclusion.

At the level of socio-cultural practice, the use of sexual and gendered lexemes reveals underlying patriarchal ideologies and normative gender relations deeply embedded in Indonesia's digital culture. Many lexemes classified under sexual verbal violence stem from social contexts already marked by unequal power relations and conservative moral norms. In both popular culture and everyday speech, these words arise from collective practices that reproduce gender stereotypes and justify the

surveillance of women's bodies and expressions, as well as those of non-dominant gender groups.

Within this context, social media can be understood as a new institution with its own discursive conventions. As with institutions like courts, schools, or hospitals—which possess their own rules, roles, and discursive forms (Fairclough, 1995; Roberts, 2011; Freed, 2015)—social media constructs and regulates user behavior and language use through implicit norms, technological affordances, and the dynamics of online interaction. Unlike face-to-face interaction, where ethical norms such as politeness, empathy, and personal accountability are upheld due to direct contact and higher social risk, social media—due to anonymity, physical distance, and screen mediation—blurs ethical boundaries (Crystal, 2011; Wang et al., 2012). This fosters the release of internal censorship and heightens the expression of violent, sarcastic, and discriminatory language.

Intelligence

Verbal violence targeting intellectual capacity is also evident in the analyzed data. Referring to Rezvan et al. (2020), this category encompasses insults aimed at an individual's cognitive abilities, intelligence level, or educational status. In this study, eight key lexical items were identified as markers of intellectual verbal violence.

Within this framework, the lexeme *cogil* emerges as a prominent feature in the corpus of verbal violence on Indonesian social media. Although its frequency is lower compared to several other aggressive terms, its consistent usage pattern and rich semantic implications render it a significant object of analysis in studies of informal linguistic aggression. This term originates from the acronym *cowok gila* ("crazy guy") and functions as a pejorative label carrying both gendered connotations and stigma related to cognitive conditions.

Within the framework of institutional discourse analysis (Mayr, 2015; Roberts, 2011; Freed, 2015), social media can be understood as a distinct discursive institution—a social space with its own structure, norms, values, and even characteristic lexicon produced, circulated, and collectively negotiated. Social media develops a variety of distinctive language styles and interaction forms, including differentiated insults and in-group coded expressions such as *cogil* and *cegil* (Staples et al., 2015; Holmes, 2015). These terms exemplify how the social media institution creates an autonomous linguistic economy—where meaning, emotional resonance, and speech legitimacy are internally produced by the user community.

The emergence of *cogil* and *cegil* reflects the participatory and memetic nature of this digital institution (Mayr, 2015). Rather than borrowing terms from oral or conventional written traditions, these terms arise as affective and evaluative products within very specific social contexts: opinion battles, debates, and comments that are often fast-paced, spontaneous, and competitive. Social media users are not merely discourse consumers but active producers who develop new jargon that can strengthen their epistemic position and group identity. Accordingly, terms like *cogil* function not only as tools for mockery or belittlement but also as markers of membership, social sorting devices, and even ideological symbols reinforcing power structures within online discourse.

Moreover, the use of these terms demonstrates how power relations are enacted through language. Speakers position themselves epistemically superior, framing themselves as rational, objective, or credible while portraying their interlocutors as deviant, irrational, or foolish (Fricker, 2007; Kidd et al., 2017). In this sense, insults linked to intellectual capacity are not merely offensive utterances but strategies to assert dominance in communicative events.

At the social practice level, this pattern reflects broader societal attitudes toward intelligence, mental health, and rationality. In Indonesian digital discourse, insults targeting intellectual capacity are not only common but normalized as legitimate tools in social contestation. This normalization reflects deeper ideological structures, including stigmatization of mental disorders, anti-intellectualism (Chen et al., 2023), and exclusion based on class and education (Holman, 1980).

This phenomenon reinforces Fairclough's (1995) concept of discourse as ideology—such verbal acts are not neutral or isolated but reproduce systems of marginalization, particularly against those with differing views, lower educational backgrounds, or mental health issues. Consequently, these actions uphold hierarchies of knowledge, rationality, and legitimacy in the digital public sphere (Fricker, 2007; Kidd et al., 2017).

Verbal violence targeting intelligence—especially through lexemes like *cogil* and *cegil*—is deeply embedded in the institutional discourse practices of social media. These terms form part of a social vocabulary collectively shaped and maintained by users, serving simultaneously as expressive means and tools of social exclusion in online communication. Their widespread and creative use across collocations, concordances, and N-grams confirms the role of language as a tool of insult with cultural and ideological

resonance, normalized through humor, memes, and participatory practices characteristic of social media as a discursive institution.

Political

In the Corpus of Verbal Violence on Social Media, the word *fufufafa* ranks fifth in the keyword analysis, with a frequency of 3,243 occurrences, renßdering it one of the most prominent lexical items in the dataset. This high ranking and frequency signal not only its referential significance but also its conspicuous discursive presence within online verbal aggression, particularly in political conversations on Indonesian social media platforms. The term *fufufafa* operates as a euphemistic or parodic placeholder, implicitly referencing specific political figures. Its widespread usage suggests a deliberate linguistic strategy—employed to convey sharp critique and ridicule while circumventing direct defamation or potential legal repercussions. See the illustration in the following figure.

At the level of discourse practice, the semantic transformation within politically violent utterances is reinforced by the status of social media as an autonomous digital discursive institution (Mayr, 2015; Fred, 2015). Social media does not merely function as a communication medium but as an institutional space that produces and normalizes certain linguistic styles, including pejorative utterances. Within this institution, distinct lexemes such as *fufufafa*, *mulyono*, and *owo* emerge, which are not only digitally exclusive but also discursively functional. These lexemes are rarely found in face-to-face interactions due to social norms, ethical constraints, and the risks of identity exposure in offline public spaces (Boyd, 2014). In contrast, the digital sphere permits anonymity, performativity, and virality, making it fertile ground for transgressive political expression.

Fufufafa, originally coined as a satirical jab at a political figure, rapidly underwent semantic pejoration through associations with expressions deemed unsubstantial, childish, or manipulative. Collocations such as *fufufafa ganyang* and *fufufafa pantat* intensify this degradation through associations with bodily imagery and symbolic violence. A similar pejoration process is evident in the lexeme *mulyono*. In digital practice, *mulyono* has evolved beyond a mere name into a transformed identity laden with insinuations against individuals perceived as politically opportunistic. Its use in collocations like *mulyono pengkhianat* or *mulyono bangsat* turns the name into a symbol of betrayal or opportunism. This process constructs new meanings through symbolic resignification, transforming *mulyono* from a personal identifier into a political archetype (Fairclough, 1995).

Meanwhile, the lexeme *owo* represents a form of infantilization and distortion of the name Prabowo, a senior Indonesian political figure. This infantilization entails a dual-layered meaning: first, an act of denigration by associating the subject with immaturity or unseriousness; second, a deliberate diminishment of elevated social status through linguistic miniaturization. This highlights dimensions of feminization and symbolic degradation, blending humor, irony, and aggression into a single utterance. In this context, *owo* is not merely a distorted name but a method for deconstructing authority through politically and emotionally charged linguistic othering (Foucault & Carrette, 1999).

As an institution, social media enables the emergence of new semantic conventions that intensify polarization and facilitate epistemic violence, functioning as a generative arena for informal political language where identities can be simultaneously constructed and dismantled (Mayr, 2015; Herring, 2004; Wodak, 2021). By reducing ideological differences to jokes and identity-based ridicule, this digital discourse not only marginalizes certain actors but also instrumentalizes language as a hegemonic tool. More broadly, this dynamic reflects an ethical crisis in electoral democracy, wherein language no longer serves deliberative purposes but becomes a weapon in the battle of narratives and emotions in the digital public sphere (Hukmi & Taufiqurrahman, 2024).

Socioculturally, this semantic shift illustrates how digital institutions reproduce new norms within Indonesia's affective politics. Metaphorical oppositions like *kadrun* and *cebong* reduce ideology to identity-based antagonism, reinforcing epistemic violence—namely, the suppression of another's cognitive agency through delegitimization of meaning and representation (Alfian, 2021; Syarif, 2024). As a result, political language ceases to function as a deliberative tool and instead becomes an instrument of power within a hyperrealistic and speculative media arena (Hukmi & Taufiqurrahman, 2024; Papacharissi, 2010). Thus, social media institutions do not merely facilitate expression but actively engineer the ethical, emotional, and ideological frameworks of digital political discourse in Indonesia.

Appearance

Verbal violence targeting physical appearance—including practices of body shaming and derogatory comments on bodily features—was also found to be significant (Rezvan et al., 2020; Boukemidja, 2018). Within this category, two keywords were repeatedly identified as indicators of appearance-based verbal abuse (Himawan et al, 2020). The terms in focus—*nuruls* and *banci*—originate from different lexical-cultural and

sociolinguistic domains, yet both undergo resignification in the context of digital discourse.

From both a semantic and pragmatic perspective, *nuruls* originated as a neutral onomastic term, but has undergone a process of pejoration, acquiring persistent negative semantic components through frequent collocational use. This transformation reveals a simultaneous process of semantic bleaching and semantic intensification: the literal meaning of the name fades, while its ideological load intensifies.

This process shifts *nuruls*' semantic field from something neutral to a metonymic stand-in for a variety of traits perceived as misaligned with dominant values in particular digital communities. As such, the word is not only associated with the religious expressions of Muslim women, but also with symbols of conservatism, ignorance, and even fanaticism. At this juncture, collocations with terms like *banci* reveal how the digital semantic field fuses divergent forms of femininity perceived as deviant—whether ideologically (*nuruls*) or in terms of gender/sexual identity (*banci*)—into a unified, discursively legitimized pejorative field.

Following Bourdieu (1991), this process constitutes a form of symbolic violence: a power exerted through symbolic structures that appear natural and go unquestioned. Through the repeated use of *nuruls* as joke, insult, or meme material, symbolic violence becomes institutionalized in digital culture. This is a form of *banal othering* that requires no direct intervention from the state or official authority—only algorithmic amplification and the active participation of netizens in generating normative microcultures.

In this light, social media users are not merely passive consumers of discourse, but co-producers of meaning within an informal yet highly influential institutional structure. Hence, digital institutionalization does not necessarily operate through legal-formal mechanisms but is maintained via discursive conventions, affective economies (Ahmed, 2004), and participatory regulation—where norms are enforced not through laws, but through social sanctions, symbolic exclusions, and viral repetition.

By viewing social media as a discursive institution, it becomes evident that lexemes like *nuruls* and *banci* not only reflect social realities but actively shape them. These words function as nodes within meaning networks that regulate who may speak, who is silenced, and how identities are constructed. This process is deeply embedded in broader power relations—patriarchy, digital capitalism, and political conservatism—which operate through symbolic and semiotic mechanisms that are highly structured yet appear spontaneous. In this regard, social media is not merely a platform for message exchange, but rather a digital ideological apparatus (adapting Althusser, 1971), one that amplifies,

produces, and circulates new forms of verbal violence based on gender, ideology, and religious expression.

Racial

Lastly, racially charged verbal violence was also found in the form of ethnic slurs or xenophobic remarks. Rezvan et al. (2020) define racial verbal violence as expressions that target a person's racial, ethnic, or national identity. Two key lexical items have been identified as markers of verbal violence within this category.

The lexeme *kampungan* ranks 15th among the most frequently occurring words in the corpus of verbal violence on Indonesian social media, with a total of 1,802 occurrences. This high frequency indicates the significant role of the term in articulating social disgust and perceptions of impoliteness or incivility in online interactions.

Within the framework of institutional discourse analysis, social media is understood not merely as a platform but as a digital institution that shapes and governs how people speak, feel, and think (Mayr, 2015; Roberts, 2011; Freed, 2015). Social media fosters its own discursive logic that differentiates it from offline spaces, particularly in the reproduction and circulation of social values (Herring, 2004; Marwick & boyd, 2011). As a digital institution, social media produces new linguistic norms—such as the use of *kampungan* to demean or *kafir* to exclude—which are legitimized and amplified through likes, retweets, virality, and algorithmic processes (Gillespie, 2018).

In this context, *kampungan* no longer functions as a spatial or cultural descriptor but has become a symbol of inferiority, one that is inscribed upon specific social bodies—those deemed digitally unworthy. This discourse reinforces the urban–rural hierarchy through digital indexicality, wherein words and their associations construct social identity and value within online communication (Blommaert, 2005). Unlike its offline connotations (e.g., *kampung halaman* or "hometown"), in digital spaces *kampungan* operates as a pejorative instrument driven by urban middle-class digital logic that valorizes modernity and aesthetic homogeneity.

The collocation of *kampungan* with terms like *norak*, *kuno*, and *miskin* (see Figure 4.52) demonstrates how digital institutions engineer semantic fields, shifting meanings from positive to negative. This semantic shift occurs through the insertion of evaluative components that attach negative connotations. Grammatically and semantically, this indicates both pejoration and discriminatory collocativization. While in face-to-face or traditional cultural contexts *kampung* may evoke tranquility, greenery, and nostalgia

(Sutrisno, 2019), in digital spaces it becomes a marker of backwardness, ignorance of trends, and even unworthiness of visibility.

Furthermore, from a global perspective, comparison with Western cultural contexts clarifies this sociocultural dimension. In American culture, there is no direct semantic equivalent to *kampungan*. Terms such as *slum* or *ghetto* refer more to physical or economic conditions, not cultural inferiority (Wacquant, 2008). This suggests that the pejoration of *kampung* is a localized product of power relations and modernism in Indonesia, distorted and internalized within social media discourse.

Meanwhile, the lexeme *kafir* functions within religious discursive realms as a tool of identity marking, strategically deployed in digital public spaces. Although etymologically neutral in classical Islamic tradition (Madelung, 1997), in contemporary social practice—particularly in digital arenas—the term has undergone resemanticization and is used to aggressively identify and marginalize the "other." In the Indonesian context, *kafir* no longer merely denotes differing belief systems but has been commodified as a labeling device to stigmatize and silence dissent or divergent identities.

Discursively, the use of *kafir* on social media exhibits a process of epistemic closure, the foreclosure of interpretive religious plurality in favor of a singular hegemonic meaning. This aligns with the concept of epistemic imperialism, wherein one ideological worldview dominates and erases others (Sonevytsky, 2022; Mazur, 2021). The collocation of *kafir* with terms such as *murtad* (apostate), *membunuh* (kill), or *neraka* (hell) underscores that it is no longer a passive identity marker, but an active linguistic weapon used to assert moral superiority and eliminate difference.

Kampungan and *kafir* do not operate independently in social media discourse but rather as discursive pairs that reinforce social and symbolic exclusion. The former marginalizes culturally, the latter theologically. In discursive practice, this reveals how digital institutions foster a symbiosis of verbal violence, sanctioned by algorithms and online community practices (Mayr, 2015). Within this context, language ceases to be a mere tool of communication and instead becomes a machine of stigma production (Fairclough, 2003; Pennycook, 2001).

Together, they constitute a mechanism of digital inequality production and reproduction, where dominant groups consolidate their positions through language that appears spontaneous but is ideologically structured. As Foucault (1972) noted, discourse is never neutral: it is always implicated in relations of power. Social media, as a digital institution, provides a stage where such power is displayed, negotiated, and reproduced in the form of symbolic violence—often hidden beneath ordinary comments or humor.

Typology of Verbal Violence

The typological mapping of Indonesian verbal violence using Waseem et al.'s (2017) framework reveals complex intersections of power, identity, and social order in digital discourse. By categorizing utterances along two axes—target specificity and linguistic explicitness—this analysis demonstrates how language functions not merely as a tool for direct domination but also as a vehicle for covert delegitimization within socio-political spaces.

Textually, expressions like *anjing* (dog), *bangsat* (bastard), and *kontol* (penis) (Targeted + Explicit) operate through direct aggression toward interlocutors, often in bidirectional exchanges (*lu anjing*). As Fairclough (1992) explains, such utterances reproduce hierarchical relational structures by positioning targets in morally inferior roles, transforming emotional outbursts into acts of discursive subjugation.

Conversely, truncated and euphemistic forms like *anj, anjg,* and *bgst* (Targeted + Implicit) reflect what Warner and Hirschberg (2012) and Nobata et al. (2016) term encoded incivility. These retain offensive potency while evading content moderation systems through deliberate ambiguity—a discursive adaptation to digital platform logics where users navigate algorithmic constraints while preserving aggressive intent.

In discursive practice, terms like *cogil* (slow-witted), *tolol* (idiot), and *goblok* (moron) (Targeted + Explicit, Intellectual) enact epistemic marginalization. Such utterances not only question cognitive capacity but discredit opposing views by linking intelligence to moral and civic worth. Observe in microaggression studies, this weaponizes intellect as a tool of social exclusion particularly against those perceived as politically naïve or lower-class (Sue, 2007). Meanwhile, expressions like *murahan* (cheap) and *perek* (slut) (Targeted + Explicit, Sexual) regulate female bodies and choices through gendered verbal violence, reinforcing heteropatriarchal norms (Butler, 1997; Hasan, 2023).

Political epithets like *kadrun* (hardline Muslim), *cebong* (tadpole/pro-government), and *penjilat* (sycophant) (Targeted + Explicit) reflect increasingly polarized digital landscapes where ideology becomes a primary site of verbal violence. These labels function dually as identity markers and slurs—naming while reducing individuals to stereotypes. Terms like *mulyono* (used sarcastically/repeatedly) (Targeted + Implicit) exemplify critique through personification, mediating social grievances via mockery of public figures. As Waseem and Hovy (2016) note, such language embeds ideological implicatures indirect yet trenchant institutional criticism conveyed through satire.

Socioculturally, expressions like *kampungan* (provincial) and kafir (infidel) (Targeted + Explicit, Racial) reinforce systemic hierarchies by normalizing assumptions of cultural inferiority and religious deviance. These constitute symbolic violence (Bourdieu, 1991) that reasserts ethno-religious boundaries and justifies exclusion under the guise of everyday speech. Meanwhile, phrases like *pencitraan* (image-polishing) and buzzer (political influencer) (Targeted + Implicit, Political) reflect discursive fatigue toward perceived state propaganda, where protest manifests not through rational argument but sarcastic cynicism an increasingly characteristic pattern in digital political critique.

This deliberative crisis aligns with Fairclough's (2003) concept of aestheticized violence: language transformed into spectacle, where verbal mockery supplants argumentative depth. Humorous or hyperbolic slurs like *topita* (coconut-head), *nuruls* (hijabi stereotype), and *banci* (f*g) (Appearance/Sexuality categories) blend entertainment with identity denigration, transmuting conflict into viral content.

Ultimately, this typology reveals Indonesian digital language as a non-neutral medium of symbolic governance. Verbal violence—both overt and covert—reproduces power asymmetries across social class, gender, religion, and politics. Through euphemism, ambiguity, and humor, users create performative distance that enables violence without accountability, forging discursive ecosystems where symbolic domination converges with plausible deniability. This constitutes not mere incivility, but a rhythmic, ritualistic choreography of linguistic mastery—performances of symbolic control unfolding within comment sections, timelines, and digital threads.

CONCLUSION

The findings of this study demonstrate that verbal violence in Indonesian social media operates through identifiable linguistic patterns, as revealed by corpus analysis. The results show that verbal violence is not limited to explicit insults or vulgar expressions, but also emerges through the recontextualization of lexicon that is denotatively neutral or even positive. These patterns confirm that lexical shifts and collocational structures play a central role in producing aggressive meanings in digital discourse.

At the discursive level, the analysis indicates that verbal violence functions through strategic language use, such as satire, labeling, and symbolic naming, which enables users to construct identities, negotiate social positions, and reinforce polarization. The use of personal names as political or ideological markers exemplifies how discursive practices transform everyday language into tools of symbolic domination.

From a critical discourse perspective, these linguistic strategies reproduce layered and flexible power relations. Verbal violence targets not only individuals, but also social groups and symbolic identities, often through implicit and indirect forms. Such veiled

aggression reinforces stereotypes and dominant ideologies, making it particularly effective in sustaining social and political hierarchies in digital spaces.

Overall, this study confirms that digital verbal violence is not merely an ethical issue of communication, but a linguistic mechanism that reproduces power and ideology in contemporary digital culture. Future research may expand the range of platforms or incorporate multimodal analysis to examine interactions between text, images, and digital symbols. Further studies may also develop NLP-based detection models to support digital literacy initiatives and context-sensitive moderation systems.

ACKNOWLEDGEMENTS

We thank Universitas Negeri Malang for funding this research, Universiti Sains Islam Malaysia for their collaborative support, and Mr. Aat Sadewa for his essential technical assistance in data crawling and collection.

REFERENCES

Ahmed, S. (2004). Affective Economies. *Social text*, *22*(2), 117-139.

Akhvlediani, Maia, & Moralishvili, Shopio. (2021). Verbal Abuse of Children. *Language and Culture, (26)*, 154–158.

Althusser, L. (1971). Ideology and ideological state apparatuses. In L. Althusser (Ed.), *Lenin and philosophy and other essays*. Monthly Review Press.

Ariibah, A. A., & Zakiyyah, H. (2023). Verbal Abuse by Netizens on Social Media Against A Youtuber's Childfree Decision Statement Named Gita Savitri. In *RSF Conference Series: Business, Management and Social Sciences, 3*(2), 29-36.

Austin, J. L. (1962). *How to do Things with Words*. Oxford University Press.

Bawden, D., & Robinson, L. (2012). *Introduction to Information Science*. Facet Publishing.

Blommaert, J. (2005). *Discourse: A critical introduction* (Key topics in sociolinguistics). Cambridge University Press.

Blommaert, J. (2010). *The Sociolinguistics of Globalization*. Cambridge University Press.

Boukemidja, N.B. (2018). Cyber Crimes Against Women: Qualification and Means. *European Journal of Social Sciences*, *1*(3), 33-44.

Bourdieu, P. (1991). *Language and Symbolic Power*. Harvard University Press.

Bousfield, D. (2008). *Impoliteness in Interaction*. John Benjamins Publishing Company.

Boyd, D. (2014). *It's Complicated: The Social Lives of Networked Teens*. Yale University Press.

Budiawan, R. Y. S., Fatin, A. N., & Mualafina, R. F. (2024). Kolokasi pada Kata Cinta Berbasis Linguistik Korpus. *Sasindo: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, *12*(1), 186-195.

Chen, Y., Zhou, Y., Zhu, S., & Xu, H. (2012). Detecting Offensive Language in Social Media to Protect Adolescent Online Safety. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (71-80). IEEE.

Cho, W. I., & Moon, J. (2021, December). How Does the Hate Speech Corpus Concern Sociolinguistic Discussions? A Case Study on Korean Online News Comments. In *Proceedings of the Workshop on Natural Language Processing for Digital Humanities* (13-22).

Crystal, D. (2011). *Internet linguistics: A Student Guide*. Routledge.

Culpeper, J. (2011). *Impoliteness: Using Language to Cause Offence*. Cambridge University Press.

Culpeper, J., & Demmen, J. (2015). Keywords. In D. Biber & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (90–105). Chapter. Cambridge University Press.

Dehingia, N., McAuley, J., McDougal, L., Reed, E., Silverman, J. G., Urada, L., & Raj, A. (2023). Violence against women on Twitter in India: Testing a taxonomy for online misogyny and measuring its prevalence during COVID-19. *PLoS one*, *18*(10),1-13.

Dewanty, I. A. B. C., & Saryono, D. (2024). Verbal violence against women on TikTok: A feminist linguistic study. *LITERA, 23*(2), 247-260.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. Longman.

Fairclough, N. (Ed.). (1992). *Critical language awareness*. Longman.

Fatim, A. L. N., Arifa, Z., Basid, A., & Lestari, S. W. (2024). Potret Perempuan pada Media Berita Daring Republika: Analisis Linguistik Korpus. *Suar Betang*, *19*(2), 155-165.

Felmlee, D., Inara Rodis, P., & Zhang, A. (2020). Sexist Slurs: Reinforcing feminine stereotypes online. *Sex roles*, *83*(1), 16-28.

Foucault, M., & Carrette, J. R. (1999). Sexuality and power. In *Religion and culture* (1st ed., pp.16). Routledge.

Freed, A. F. (2015). Institutional Discourse. *The International Encyclopedia of Language and Social Interaction*, 1-18.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. Yale University Press.

González, I. C. (2019). *Sentiment Analysis: An Approach to Verbal Violence & Abuse Against Women on Twitter*. Universidad Nacional de Educación a Distancia.

Graham, L. J. (2005). Discourse Analysis and The Critical Use of Foucault. In *The Australian association of research in education annual conference*.

Hassan, H. (2023). Kajian linguistik bahasa Melayu Bersumberkan Korpus. Kekuatan dan kelemahannya. *PENDETA*, *14*(1), 23-34.

Herring, S. C. (2004). Computer-Mediated Discourse Analysis: an Approach to Researching Online Behavior. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for Virtual Communities in the Service of Learning* (pp. 338–376). Cambridge University Press.

Himawan, R., Fathonah, E. N., Heriyati, S., & Maslakhah, E. N. I. (2020). Analisis Kesalahan Berbahasa Bidang Semantik pada Karangan Teks Deskripsi Siswa Kelas VII A SMPIT Ar-Raihan Kabupaten Bantul. *Pena: Jurnal Pendidikan Bahasa dan Sastra*, *10*(1), 1-9.

Holman, R. H. (1980). Clothing as Communication: An Ampirical Investigation. *Advances in Consumer Research*, *7*(1).

Holmes, J. (2015). Discourse in the Workplace. *The Handbook of Discourse Analysis*, 880-901.

Hukmi, R., & Taufiqurrahman, T. (2024). Framing Radicalism: Anticipatory Epistemic Injustice in Indonesian Religious Discourse. *Religiovedenie*, *2024*(2), 179-187.

Ibrohim, M. O., & Budi, I. (2023). Hate Speech and Abusive Language Detection in Indonesian Social Media: Progress and Challenges. *Heliyon*, *9*(8). 1-16.

Isam, H., & Abd Mutalib, M. (2019). Pemanfaatan Analisis Korpus sebagai Teknik Alternatif Pengajaran dan Pembelajaran Tatabahasa. *International Journal of Language Education and Applied Linguistics, 9*, 13–31.

Jay, T. & Janschewitz, K. (2008). The Pragmatics of Swearing. *Journal of Politeness Research*, *4*(2), 267-288.

Jones, R. H., & Hafner, C. A. (2021). *Understanding Digital Literacies: A Practical Introduction*. Routledge.

Kidd, I. J., Medina, J., & Pohlhaus, G., Jr. (Eds.). (2019). *The Routledge Handbook Of Epistemic Injustice* (1st ed.). Routledge.

Kilgarriff, A., & Grefenstette, G. (2003). Introduction to the Special Issue on The Web as Corpus. *Computational Linguistics, 29*(3), 333–347.

Kumar, R., Zaman, T., & Mehdad, Y. (2019). A Multi-level Annotation Scheme for Harassment in Hindi–English Code-Mixed Social Media Text. *Language Resources and Evaluation, 53*(3), 625–645.

Lepoutre, M., Vilar-Lluch, S., Borg, E., & Hansen, N. (2024). What is Hate Speech? The Case for a Corpus Approach. *Criminal Law and Philosophy*, *18*(2), 397-430.

Madelung, W. (1997). *The Succession to Muhammad: A study of the Early Caliphate*. Cambridge University Press.

Marwick, A. E., & Caplan, R. (2018). Drinking Male Tears: Language, the Manosphere, and Networked Harassment. *Feminist Media Studies*, *18*(4), 543-559.

Mayr, A. (2015). Institutional Discourse. *The Handbook of Discourse Analysis*, 753-774.

Mazur, L. B. (2021). The Epistemic Imperialism of Science. Reinvigorating Early Critiques of Scientism. *Frontiers in Psychology*, *11*, 1-12.

Naco. (2019). Dalih Menyesal, Guru Honor Terdakwa Ujaran Kebencian Menangis di Depan Hakim. *Borneo News*. Diakses 10 Februari 2024.

Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). Abusive language detection in online user content. In *Proceedings of the 25th International Conference on World Wide Web* (pp. 145–153).

Pennycook, A. (2001). *Critical Applied Linguistics*. Routledge.

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*, *55*, 477-523.

Puspitasari, D. A., & Sukma, B. P. (2022). Memotret Hoaks Covid-19 di Awal Pandemi melalui Analisis Wacana Berbasis Linguistik Korpus. *Ranah: Jurnal Kajian Bahasa*, *11*(2), 243-261.

Rezvan, M., Shekarpour, S., Alshargi, F., Thirunarayan, K., Shalin, V. L., & Sheth, A. (2020). Analyzing and Learning the Language for Different Types of Harassment. *PLOS ONE, 15*(3), e0227330.

Roberts, C. (2011). Institutional discourse. In *The Routledge Handbook of Applied Linguistics* (pp. 81-95). Routledge.

Saraswati, E., Adam, R., & Sohyb, M. N. (2025). Wacana Literasi Keuangan pada Media Sosial: Pendekatan Korpus Linguistik. *Jurnal Manajemen, Akuntansi, Ekonomi*, *4*(2), 369-382.

Seda, K. A., Kurniawan, E., & Supriadi, R. (2025). Makna Politik Istilah “Antek”: Kajian Ilokusi dan Korpus Linguistik dalam Jurnal Akademik. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, *11*(3), 2723-2737.

Sinani, B. (2023). The Power of Words-Unveiling the Depths of Verbal Violence. *SEEU Review*, *18*(2), 136-147.

Sonevytsky, M. (2022). What is Ukraine? Notes on Epistemic Imperialism. Философско-культурологический журнал Topos, (2), 21-30.

Staples, S., Egbert, J., Biber, D., & Conrad, S. (2015). Register Variation a Corpus Approach. *The handbook of discourse analysis*, 505-525.

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial Microaggressions in Everyday Life: Implications For Clinical Practice. *American Psychologist, 62*(4), 271.

Suhardijanto, T., & Dinakaramani, A. (2018). Korpus Beranotasi: ke Arah Pengembangan Korpus Bahasa-Bahasa di Indonesia.

Tagg, C., Seargeant, P., & Brown, A. (2017). *Taking Offence on Social Media: Conviviality and Communication on Facebook*. Palgrave Macmillan.

Wacquant, L. (2008). *Urban Outcasts: A Comparative Sociology of Advanced Marginality*. Polity.

Wang, W. Y. (2013). *Understanding Offensive Language in Social Media*. *Proceedings of NAACL*.

Wang, W., Chen, L., Thirunarayan, K., & Sheth, A. P. (2012). Harnessing Twitter "big data" for automatic emotion identification. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (pp. 587–592). IEEE.

Warner, W., & Hirschberg, J. (2012). Detecting hate speech on the world wide web. In *Proceedings of the Second Workshop on Language in Social Media* (pp. 19–26).

Waseem, Z., & Hovy, D. (2016). Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter. In *Proceedings of NAACL-HLT 2016* (pp. 88–93).

Waseem, Z., & Hovy, D. (2017). Understanding Abuse: A Typology of Abuse on Social Media. In *Proceedings of the 26th International Conference on World Wide Web Companion* (pp. 773–782).

Waseem, Z., Davidson, T., Warmsley, D., & Weber, I. (2017). Understanding Abuse: A Typology of Abusive Language Detection Subtasks. *Proceedings of the First Workshop on Abusive Language Online*.