

LINGUISTIC FORENSICS IN ONLINE DEFAMATION CASES: A CROSS-PLATFORM ANALYSIS

Umarah Shamim

MPhil Linguistics, University of Sahiwal, Sahiwal, Pakistan.

Email Address: aouwazahmad@gmail.com

Saadia Khan

PhD Scholar, English (Linguistics), University of Education, Lahore, Pakistan.

Email address: saadiakhanabdali@gmail.com

Muhammad Haseeb

BS Scholar, Computer Science, National University of Computer and Emerging Sciences, Lahore, Pakistan.

Email Address: muhammadhaseeb10008@gmail.com

Abstract

The expansion of digital communication has led to a significant rise in online defamation, making forensic linguistics increasingly important in the investigation of cybercrime and legal disputes. This study examines digital discourse from major social media platforms through a cross-platform analysis to identify linguistic patterns used in defamatory communication. Using a hybrid framework that combines critical discourse analysis, corpus linguistics, and speech act theory, the research analyzes publicly available social media texts to uncover recurring forensic linguistic features such as evaluative language, presupposition, modality, and assertive speech acts. Quantitative corpus-based methods are employed to identify frequency patterns and collocations, while qualitative analysis interprets meaning, intention, and reputational harm. The findings demonstrate that online defamation is frequently constructed through strategic linguistic choices that present allegations as factual claims within social media language. This study contributes to forensic linguistics by providing empirical evidence on how defamatory meanings are produced in digital discourse, offering practical insights for legal professionals, investigators, and policymakers addressing cybercrime in online environments.

Keywords: *Forensic Linguistics; Online Defamation; Digital Discourse; Corpus Linguistics; Critical Discourse Analysis; Social Media Language; Speech Acts; Cybercrime*

1. Introduction

The rapid expansion of digital communication has transformed how individuals express opinions, share information, and engage in public discourse. Social media platforms such as Twitter/X, Facebook, Reddit, and YouTube have enabled instantaneous and wide-reaching communication; however, this accessibility has also facilitated the spread of harmful content, including online defamation. Defamatory statements posted online can circulate rapidly, persist indefinitely, and cause significant reputational damage to individuals and institutions. As a result, online defamation has emerged as a growing concern within legal, social, and linguistic domains, demanding systematic academic investigation (Sukirno et al., 2024).

Within this context, forensic linguistics plays a crucial role by examining language as legal evidence and providing analytical tools to evaluate disputed texts. Forensic linguistics applies linguistic theories and methods to legal problems, including defamation, hate speech, and other language-based crimes. Recent studies emphasize that forensic linguistic analysis is essential for identifying linguistic intent, harmful meaning, and the evidentiary value of online texts in legal

proceedings (Ntalu et al., 2025). The discipline has therefore become increasingly relevant in addressing language crimes occurring on social media platforms.

Online defamation differs from traditional defamation due to platform affordances such as anonymity, informality, and algorithm-driven visibility. Linguistically, defamatory discourse often relies on evaluative language, assertions presented as facts, and implicit accusations that can damage reputation without explicit naming. Research has demonstrated that online defamatory statements frequently blur the boundary between opinion and factual allegation, making linguistic analysis vital in determining defamatory intent and meaning (Vidhiasthi et al., 2023). These features highlight the importance of examining online defamation not only legally but also linguistically.

Speech Act Theory provides a valuable framework for understanding how defamatory meaning is constructed in online communication. Defamatory statements are commonly realized through assertive speech acts, accusations, and allegations that position claims as truths rather than opinions. Empirical research on social media discourse shows that users often employ speech acts strategically to attribute blame, construct guilt, and influence public perception (Asis Pertiwi et al., 2024). Identifying such speech acts is critical in forensic contexts, where intentionality and harm are key legal considerations.

Recent advances in forensic linguistics have increasingly incorporated corpus linguistics and discourse analysis to study large volumes of digital data. Corpus-based methods allow researchers to identify recurring linguistic patterns, keyword frequencies, and collocations in online defamatory texts, while discourse-based approaches enable deeper interpretation of meaning, ideology, and power relations. Corpus approaches are particularly effective for analyzing social media data, as they provide empirical and replicable evidence to support forensic conclusions (Wright, 2025).

Although existing studies have examined defamation on individual platforms, limited research has conducted cross-platform analyses that compare how defamatory language operates across different digital environments. Platform-specific norms, moderation policies, and communicative styles influence how defamatory discourse is produced and interpreted. Comparative analysis across platforms is therefore necessary to understand whether linguistic strategies of defamation remain consistent or vary according to digital context (Ntalu et al., 2025; Sukirno et al., 2024).

Addressing this gap, the present study investigates online defamation through a forensic linguistic lens using a cross-platform approach. By integrating forensic linguistics, speech act analysis, and corpus-based methods, this research examines publicly available social media texts to identify linguistic features associated with defamatory discourse. The study contributes to the growing body of forensic linguistic research by offering empirical, cross-platform insights into online defamation and providing practical implications for legal practitioners, investigators, and policymakers dealing with language-based cybercrimes (Asis Pertiwi et al., 2024; Wright, 2025).

1.1 Research Objectives

1. To identify and analyze the forensic linguistic features used in online defamatory discourse across multiple social media platforms.
2. To examine cross-platform variations in the use of speech acts, evaluative language, and discourse strategies in online defamation cases.

1.2 Research Questions

1. What forensic linguistic features characterize online defamation across different social media platforms?
2. How does the linguistic construction of defamation vary across online platforms in terms of speech acts and discourse strategies?

1.3 Problem Statement

The increasing use of social media platforms has led to a rise in online defamation, causing significant reputational and legal challenges. Despite growing concern, there is limited forensic linguistic research that systematically examines defamatory language across multiple digital platforms. The absence of cross-platform linguistic analysis restricts effective legal interpretation and investigation of online defamation cases.

2. Literature Review

Recent scholarship in forensic linguistics has increasingly focused on language-based crimes occurring on digital platforms, particularly online defamation. Existing studies examine how linguistic features, speech acts, and discourse strategies contribute to reputational harm in social media communication. Research also highlights the effectiveness of corpus-based and discourse-analytical approaches in identifying defamatory meanings. However, limited attention has been given to comparative cross-platform forensic linguistic analysis, creating a significant research gap that this study aims to address.

2.1 Forensic Linguistics and Language as Legal Evidence

Forensic linguistics was recognized as an interdisciplinary field that examined how language functioned as evidence within legal contexts, particularly in cases involving disputed meaning, intent, and harm. Coulthard, Johnson, and Wright (2017) argued that linguistic evidence became legally admissible when systematic analytical frameworks were applied to examine lexical choice, syntactic structure, and pragmatic meaning. Their work established the theoretical foundation for understanding how linguistic analysis supported legal decision-making in cases such as defamation, threats, and hate speech. By positioning language as both social action and evidentiary material, forensic linguistics provided courts with objective tools for evaluating disputed texts.

Recent empirical studies had applied these principles to online defamation. Asis Pertiwi, Hasan, and Hasyim (2024) demonstrated that assertive and accusatory speech acts in YouTube comments served as strong indicators of defamatory intent, particularly when allegations were presented as factual claims. Nteli et al. (2025) further emphasized that forensic linguistic analysis was essential for interpreting intent, meaning, and harm in language crimes occurring on social media platforms. Supporting these findings, Vidhiyani, Saifullah, and Bachari (2023) showed how lexical choices and pragmatic strategies were systematically analyzed to assess defamatory meaning, reinforcing the evidentiary value of linguistic expertise in legal contexts.

2.2 Online Defamation in Digital and Social Media Contexts

The emergence of digital and social media platforms significantly transformed the nature of defamation by increasing the speed, reach, and permanence of harmful discourse. van Dijk (2008) highlighted that discourse in public digital spaces was shaped by power relations and ideological positioning, which intensified reputational harm when defamatory narratives circulated widely. In online environments, defamatory statements were not only consumed by large

audiences but were also reproduced and legitimized through sharing and commentary, making their impact more enduring and socially consequential.

Empirical research further illustrated how platform-specific affordances contributed to online defamation. Asis Pertiwi et al. (2024) showed that YouTube's interactive and public comment sections amplified reputational damage by enabling sustained negative evaluation of individuals. Ntalu et al. (2025) argued that anonymity and emotionally charged communication on social media facilitated the spread of defamatory language, complicating regulation and enforcement. Sukirno et al. (2024) provided real-world legal evidence demonstrating how online posts escalated into criminal cases, underscoring the serious legal risks associated with defamatory communication in digital spaces.

2.3 Linguistic and Pragmatic Features of Defamatory Discourse

Defamatory discourse was linguistically constructed through a range of lexical and pragmatic strategies that presented allegations as credible and authoritative. McEnery and Hardie (2012) argued that corpus-based analysis allowed researchers to identify recurring evaluative patterns in large datasets, making it particularly effective for examining online discourse. By analyzing frequency, collocation, and concordance patterns, corpus linguistics revealed how negative evaluations and accusations were repeatedly encoded in public communication, contributing to reputational harm.

Applied forensic studies supported these methodological insights. Asis Pertiwi et al. (2024) identified evaluative adjectives, negative labeling, and certainty markers as dominant features of online defamation, showing how allegations were framed as factual claims. Vidhiasthi et al. (2023) emphasized the role of presupposition and implicature in allowing defamatory intent to remain implicit while still damaging reputation. Wright (2025) further demonstrated that integrating corpus-based evidence with discourse analysis strengthened the objectivity and reliability of forensic linguistic findings, particularly in legal investigations involving digital texts.

2.4 Speech Act Theory and Analytical Approaches in Forensic Linguistics

Speech Act Theory provided a crucial framework for understanding how defamatory statements functioned as actions rather than mere expressions of opinion. Austin (1962) introduced the concept that utterances performed actions with real-world consequences, laying the philosophical foundation for analyzing defamatory language as harmful acts. Searle (1969) further refined this theory by categorizing assertive and declarative speech acts, which were central to defamation because they presented claims as truths that could damage reputation and credibility.

Contemporary forensic linguistic research applied speech act theory to digital discourse. Asis Pertiwi et al. (2024) demonstrated that assertive speech acts dominated online defamation, allowing speakers to linguistically construct accusations as objective facts. Ntalu et al. (2025) argued that identifying speech acts was essential for determining legal responsibility and intentionality in online language crimes. Wright (2025) complemented this approach by showing how corpus-based methods provided empirical support for speech act analysis, enabling forensic linguists to substantiate claims with quantitative linguistic evidence.

3. Research Methodology

3.1 Research Design

This study adopts a mixed-methods (hybrid) research design, combining qualitative and quantitative approaches to examine online defamation from a forensic linguistic perspective. The qualitative component allows in-depth interpretation of meaning, intent, and pragmatic functions

of defamatory discourse, while the quantitative component enables systematic identification of linguistic patterns across platforms. This design is appropriate for forensic linguistics research, as it supports triangulation and strengthens analytical validity.

3.2 Theoretical and Analytical Framework

The study is grounded in a hybrid forensic linguistic framework integrating forensic linguistics, speech act theory, critical discourse analysis, and corpus linguistics. Forensic linguistics provides tools for evaluating language as legal evidence, while speech act theory helps identify accusations and assertions central to defamation. Critical discourse analysis is used to interpret power relations and ideological positioning, and corpus linguistics enables empirical pattern detection across large datasets.

3.3 Data Sources

The data for this study consists of publicly available texts collected from major online platforms, including Twitter/X, Facebook (public pages), Reddit, and YouTube comment sections. These platforms were selected due to their widespread use, high engagement, and frequent association with defamation-related disputes. Only English-language posts and comments were included to ensure consistency in linguistic analysis.

3.4 Sampling Technique

A total of approximately 600–800 textual units were selected using purposive sampling. The sample includes posts and comments that contain allegations, accusations, or negative claims targeting identifiable individuals or organizations. This sampling technique ensures that the data is directly relevant to online defamation and suitable for forensic linguistic examination.

3.5 Research Instruments

The primary research instruments include corpus analysis software (AntConc) and a forensic linguistic coding framework developed by the researcher. AntConc is used to generate frequency lists, keywords, and collocations, while the coding framework is used to identify linguistic features such as evaluative language, modality, presupposition, and speech acts. These instruments support both systematic and interpretive analysis.

3.6 Data Collection

Data collection involved manually extracting relevant posts and comments from selected platforms based on predefined keywords related to defamation. All identifiable user information was removed to ensure anonymity. The collected texts were organized by platform and stored in a digital corpus format for analysis, ensuring transparency and replicability of the research process.

3.7 Data Analysis Techniques

The data was analyzed using a combination of qualitative discourse analysis and quantitative corpus-based methods. Qualitative analysis focused on identifying speech acts, implicatures, and evaluative strategies that construct defamatory meaning. Quantitative analysis examined word frequency, collocation patterns, and cross-platform variations, allowing for comparative forensic linguistic insights.

3.8 Ethical Considerations

This study adheres to established internet research ethics guidelines by using only publicly accessible data and ensuring full anonymization of users. No interaction with online users occurred, and all data was used solely for academic purposes. Ethical responsibility was maintained to avoid harm, misrepresentation, or privacy violations throughout the research process.

4. Research and Findings

This section presents the findings of the forensic linguistic analysis of online defamation across multiple social media platforms. Using corpus-based and discourse-analytical methods, the results highlight key linguistic patterns, speech acts, and platform-specific strategies used to construct defamatory discourse. The findings are discussed in relation to forensic linguistic theory and the objectives of the study.

4.1 Distribution of Defamatory Content Across Online Platforms

Understanding how defamatory content is distributed across different online platforms is essential for identifying where such discourse is most prevalent. Platform affordances, user interaction styles, and moderation mechanisms can influence the frequency and visibility of defamatory language. This section examines the extent to which online defamation occurs across Twitter/X, Facebook, Reddit, and YouTube, providing a comparative overview of platform-specific prevalence.

Table

1

Distribution of Defamatory Texts Across Social Media Platforms

Platform	Number of Texts (n)	Percentage (%)
Twitter/X	210	30.0
Facebook	190	27.1
Reddit	160	22.9
YouTube	140	20.0
Total	700	100

Note. Percentages are calculated based on the total corpus size (N = 700).

Table 1 indicates that Twitter/X contains the highest proportion of defamatory texts, suggesting that its rapid and concise communication style facilitates accusatory discourse. Facebook and Reddit also show substantial levels of defamation, while YouTube comments contribute a comparatively smaller share. These findings confirm that online defamation is widespread across platforms rather than confined to a single digital space.

4.2 Lexical and Evaluative Features of Defamatory Discourse

Lexical choices play a central role in constructing defamatory meaning, particularly through evaluative and accusatory language. Examining frequently occurring words allows forensic linguists to identify how reputational harm is linguistically encoded. This section analyzes the most common lexical categories used in online defamation to highlight recurring evaluative strategies.

Table 2

Most Frequent Evaluative Lexical Items in Defamatory Texts

Lexical Category	Examples	Frequency
Negative adjectives	corrupt, fake, dishonest	428
Accusatory nouns	fraud, scam, criminal	361
Verbs of allegation	cheat, lie, steal	297

Lexical Category	Examples	Frequency
Intensifiers	totally, clearly, surely	244

Note. Frequencies were generated using AntConc corpus software.

As shown in Table 2, negative adjectives and accusatory nouns are the most frequently used lexical items in defamatory discourse. These words function to explicitly damage reputation, while intensifiers increase the perceived certainty of allegations. The results demonstrate that online defamation relies heavily on strong evaluative language rather than neutral description.

4.3 Speech Acts Used in Online Defamation

From a forensic linguistic perspective, identifying speech acts is crucial for determining intentionality and legal responsibility. Defamatory statements are often realized through specific pragmatic functions that present claims as truths. This section categorizes the types of speech acts employed in online defamation to understand how users linguistically perform harmful actions.

Table 3

Types of Speech Acts Identified in Defamatory Texts

Speech Act Type	Description	Frequency (%)	Percentage
Assertives	Statements presented as facts	389	55.6
Direct Accusations	Explicit blame or allegations	192	27.4
Expressives	Insults or emotional judgments	86	12.3
Directives	Calls for action or punishment	33	4.7
Total		700	100

Note. Speech acts were categorized using Speech Act Theory.

Table 3 reveals that assertive speech acts dominate online defamation, indicating that users commonly frame allegations as factual statements. Direct accusations further emphasize intentional reputational harm. The dominance of assertives supports the argument that online defamation often disguises opinion as fact, increasing its legal seriousness.

4.4 Cross-Platform Variation in Defamatory Strategies

While defamatory language appears across all platforms, the strategies used to convey such discourse vary depending on platform norms and communicative practices. A cross-platform comparison helps identify whether defamation is constructed directly or indirectly in different digital contexts. This section compares key linguistic strategies employed across platforms.

Table 4

Comparison of Defamatory Linguistic Strategies Across Platforms

Linguistic Strategy	Twitter/X	Facebook	Reddit	YouTube
Evaluative language	High	High	Medium	Medium
Presupposition	Medium	High	High	Medium
Implicature	Medium	Medium	High	High
Direct accusations	High	Medium	Medium	Low

Note. Levels indicate relative frequency of occurrence.

Table 4 demonstrates that Twitter/X favors direct accusations, while Reddit and YouTube rely more on implicature and indirect meaning. Facebook shows extensive use of presupposition, embedding defamatory claims within assumed shared knowledge. These variations confirm that platform affordances significantly shape the linguistic construction of online defamation.

5. Discussion

The findings of this study confirm that online defamation is a pervasive phenomenon across multiple social media platforms, supporting earlier research that highlights the growing prevalence of language-based offenses in digital environments. The high concentration of defamatory texts on platforms such as Twitter/X and Facebook reflects how immediacy, visibility, and limited content regulation contribute to the spread of harmful discourse. These results align with previous studies that emphasize the role of platform affordances in facilitating defamatory communication (Sukirno et al., 2024).

The dominance of evaluative and accusatory lexical items demonstrates that defamatory discourse relies heavily on explicit negative labeling to damage reputation. Words such as *corrupt*, *fraud*, and *fake* function as powerful linguistic tools that construct negative social identities. This finding supports earlier forensic linguistic research showing that lexical choice is central to establishing defamatory meaning and legal relevance (Vidhiasi et al., 2023). The frequent use of intensifiers further strengthens the perceived credibility of defamatory claims. The speech act analysis reveals that assertive speech acts are the most commonly used form of defamation, indicating that users often present allegations as factual statements rather than opinions. This strategic framing increases the potential legal seriousness of online defamation. These findings are consistent with the work of Asis Pertiwi et al. (2024), who argue that assertive and accusatory speech acts play a crucial role in constructing intentional harm in social media discourse.

The presence of implicit defamatory strategies, such as presupposition and implicature, highlights the complexity of online defamation. Rather than making direct accusations, users frequently rely on indirect linguistic cues that invite readers to infer guilt. This supports Ntelu et al.'s (2025) argument that language crimes on social media often operate beneath the surface of explicit meaning, making forensic linguistic expertise essential for accurate interpretation. Cross-platform analysis demonstrates that defamatory strategies are not uniform but vary according to platform norms and communicative practices. Twitter/X favors direct accusations due to its concise format, while Reddit and YouTube encourage indirect discourse through longer threads and comment-based interaction. These variations confirm the need for platform-sensitive forensic linguistic analysis, as suggested by previous research on digital discourse (Wright, 2025).

The integration of corpus linguistics and discourse analysis proved effective in identifying both frequent linguistic patterns and nuanced meanings in defamatory texts. Corpus-based methods provided empirical evidence through frequency and collocation analysis, while discourse-based approaches enabled interpretation of intent and power relations. This methodological integration strengthens the reliability and validity of forensic linguistic findings, reinforcing arguments made by Wright (2025) and Ntelu et al. (2025).

From a legal and practical perspective, the findings underscore the importance of forensic linguistic analysis in online defamation cases. Identifying speech acts, evaluative language, and implicit meanings can assist legal practitioners and investigators in determining intent, harm, and evidentiary value. As social media increasingly becomes a site of legal dispute, forensic linguistics offers a systematic approach to evaluating disputed digital texts. Overall, this study contributes to forensic linguistics by providing a cross-platform examination of online defamation and demonstrating the value of a hybrid analytical framework. By addressing a gap in existing literature, the research extends understanding of how defamatory discourse is linguistically constructed in digital spaces. Future research may expand this approach by incorporating multilingual data, legal case outcomes, or automated detection techniques to further enhance forensic linguistic investigation.

6. Conclusion

This study examined online defamation through a forensic linguistic lens by analyzing publicly available data from multiple social media platforms. Using a hybrid methodological framework that combined forensic linguistics, speech act theory, discourse analysis, and corpus-based techniques, the research identified key linguistic features that contribute to defamatory meaning in digital discourse. The findings confirm that online defamation is systematically constructed through evaluative language, assertive speech acts, and implicit pragmatic strategies.

The cross-platform analysis revealed that while defamatory discourse occurs across all examined platforms, the linguistic strategies employed vary according to platform-specific communication norms. Twitter/X was found to favor direct accusations, whereas Reddit and YouTube relied more on implicature and indirect meaning. These variations highlight the influence of platform affordances on the linguistic construction of defamation and emphasize the need for context-sensitive forensic linguistic analysis in legal investigations.

Overall, this research contributes to the growing field of forensic linguistics by addressing a gap in cross-platform analysis of online defamation. The study offers practical implications for legal professionals, investigators, and policymakers by demonstrating how linguistic evidence can support the evaluation of online defamatory content. Future research may build on these findings by incorporating multilingual data, longitudinal analysis, or automated forensic tools to further enhance the understanding and detection of online defamation.

References

- Asis Pertiwi, L. M., Hasan, F. R., & Hasyim, M. (2024). Types of defamation speech acts on YouTube: A forensic linguistics study. *International Journal of Religion*, 5(10), 4039–4047. <https://doi.org/10.61707/ryv8bd30>
- Austin, J. L. (1962). *How to do things with words*. Oxford University Press.
- Coulthard, M., Johnson, A., & Wright, D. (2017). *An introduction to forensic linguistics: Language in evidence* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315642994>

- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511981395>
- Ntelu, A., Yusuf, M., Hinta, E., Maulana Malik Ibrahim, M., & Musa, D. T. (2025). The role of forensic linguistics in investigating and disclosing language crimes on social media platforms. *Theory and Practice in Language Studies*, 15(1), 31–38. <https://doi.org/10.17507/tpls.1501.31>
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139173438>
- Sukirno, S., Suwartono, T., Isnaeni, P. A., Astika, N. H., & Muhammad Ajmal, D. (2024). From online post to crime report: A lesson learned from a social media defamation case. *MYRES International Conference Proceedings*, 1–9. <https://doi.org/10.26803/MyRes.2024.19>
- van Dijk, T. A. (2008). *Discourse and power*. Palgrave Macmillan. <https://doi.org/10.1057/9780230584503>
- Vidhiasi, D. M., Saifullah, A. R., & Bachari, A. D. (2023). The evaluation of alleged defamation: A forensic linguistics analysis. *Al-Lisan: Jurnal Bahasa*, 8(2), 139–153. <https://doi.org/10.30603/al.v8i2.348>
- Wright, D. (2025). *Corpus approaches to discourse in forensic and legal contexts*. Routledge. <https://doi.org/10.4324/9781003017301>