



ADVANCE SOCIAL SCIENCE ARCHIVE JOURNAL

Available Online: <https://assajournal.com>
 Vol. 05 No. 01. Jan-March 2026. Page# 3669-3685
 Print ISSN: [3006-2497](https://doi.org/10.5281/zenodo.20989410) Online ISSN: [3006-2500](https://doi.org/10.5281/zenodo.20989410)
 Platform & Workflow by: [Open Journal Systems](https://doi.org/10.5281/zenodo.20989410)
<https://doi.org/10.5281/zenodo.20989410>



Linguistic Markers of Fake News in Pakistani Political Discourse: A Corpus-Based Forensic Linguistic Study

Sadaf Rashid

M.Phil Scholar, Department of English, University of Okara, Okara, Pakistan
[sadaf.b217786@gmail.com.pk](mailto:sadaf.b217786@gmail.com)

Tariq Usman

Assistant Professor of English, University of Okara, Pakistan
tariq.usman@uo.edu.pk

Hina Riaz

M.Phil Scholar, Department of English, University of Okara, Okara, Pakistan
hina.riaz25000@gmail.com

Afsha Rana

M.Phil Scholar, Department of English, University of Okara, Okara, Pakistan
afshirana80@gmail.com

Sana Riaz

M.Phil Scholar, Department of English, University of Okara, Okara, Pakistan
rsana4776@gmail.com

ABSTRACT

As digital media platforms have grown in scale, so has the amount of fake news and political misinformation spreading throughout the world. Emotionally manipulative, sensational and ideologically polarizing content has become the hallmark of the political discourse on social media like Facebook, X (formerly known as Twitter), YouTube and online news portals in Pakistan. There is little to no research that explored the linguistic features of fake political discourse in the Pakistani context using forensic linguistic and corpus-based approaches despite the remarkable impact of fake political news on peoples' opinion and democratic processes. In this regard, the present quantitative study aims to explore the linguistic indicators of fake news in the context of political discourse in Pakistan based on forensic linguistic approach. The study was of quantitative research and created a special corpus of 200 political news texts, of which 100 were fake political news and 100 were authentic political news texts taken from Pakistani digital media platforms in the year 2023 to 2025. The collected data were analyzed using an antConc corpus analysis software, which includes lexical frequency, emotional vocabulary, certainty markers, usage of pronouns, sensational lexical items and modality patterns. The results showed substantial differences between fake and authentic political discourse in the language used. Fake political news showed higher levels of emotionally charged lexicon, hyperbole expressions of certainty, ideological pronouns, and sensational lexical options used for persuasion and political manipulation. The lexical patterns as well as the structures of modality and factual reporting in political news were, however, rather neutral in the authentic ones. The study finds that linguistic markers have a considerable impact in the distinction between fake political discourse and real political communication in Pakistan. The results are relevant to forensic linguistics, corpus linguistics, digital misinformation research and have implications for media literacy, fake news identification, and political communication research.

Keywords: Fake News, Forensic Linguistics, Corpus Linguistics, Political Discourse, Misinformation, Pakistani Media Discourse

INTRODUCTION

Digital communication technologies and social media have quickly revolutionized political communication around the globe. Social media has emerged as a powerful tool for sharing political information, debate, and ideology, including social media platforms like Facebook, X (Twitter), YouTube, TikTok, and WhatsApp. These digital platforms have given more people access to information and to democratic participation, but also to fake news, misinformation and politically manipulated information. In the last few years, fake political news has become one of the most serious threats to the democratic processes, media credibility, and public trust in information systems (Sousa-Silva, 2022).

The fact that emotionally provocative and sensational topics get more attention from people than factual reports in digital networks makes political misinformation spread quickly in digital networks (Vosoughi et al., 2018). Consequently, ideological manipulation and deceptive communication tactics are increasingly becoming a threat to digital political discourse. Fake news is usually used to describe false, manipulated, or misleading information that is made to appear as if it is news media, but is used for the purposes of manipulating people's opinions, or inducing emotional responses (Allcott & Gentzkow, 2017).

Fake news is not the same as misinformation; it also has elements of linguistic and discursive strategies that are deliberate and attempt to look believable and convincing. Political fake news is especially used to influence political opinions, help polarize, to attack opponents of ideas, and to manipulate the perception of the electorate. Political misinformation has been found to have affected election campaigns, political protests, public policy debates, and social stability in many countries. This trend towards fake news has, therefore, emerged as a prominent topic in the fields of media studies, communication studies, political science, linguistics and forensic discourse analysis.

In the last decade Pakistan has witnessed significant growth in digital media consumption. In Pakistan, political influence and communication have become more and more techno literate from traditional TV and print to the internet and social networking sites. Political parties, ideological groups, journalists, social media influencers and anonymous digital users are all involved in online political discussions. But as political communication has increasingly become digital, it has also been accompanied by the proliferation of fake political narratives, propaganda, conspiracy theories and manipulated visuals and false claims. Political misinformation is common in Pakistan, especially when elections are held, when political crises emerge, when protests occur, when institutional conflicts emerge, or during ideological debates. This is often done by using rhetoric, nationalism, religious sentiment, and anti-opposition narratives to sway public opinion and generate further polarization (Ahmed & Nabi, 2021).

Fake news is not only technological or political, it is linguistic/discursive. Fake political narratives are created through effective use of emotions, intensifiers, ideological framing, sensational vocabulary, and manipulative discourse strategies, and persuasive lexical choices. Language is a main player in the perception of audience, the formation of political identity and the formulation of ideological statements. Hence, the study of fake news in the field of forensic linguistics has gained significance in modern times. Forensic linguistics is the application of linguistic theories and techniques to contexts of communication that are investigative, evidential, juridical and deceptive (Coulthard et al., 2020).

Forensic linguistic analysis can be used to determine linguistic features that are related to manipulation, deception, persuasion and ideological construction, in the context of fake news studies. In recent years, the field of forensic linguistics has become more focused on digital communication and deception online. Studies also suggest that there are linguistic patterns that can be identified in deceptive discourse and are distinct from genuine discourse (Sousa-Silva, 2022).

In the construction of fake political news, false language can be emotionally charged, overconfident, ideologically charged with pronouns, sensational in choosing words, and repetitive in using propaganda constructions. The use of these linguistic techniques is deliberate in an attempt to sway the audience's emotions and to construct effective political narratives. Fake news rhetoric has a surprising tendency to focus on negativity, emotionality, and dramatization, in order to gain the attention and engagement of the audience (Baissa 2025), and create a sense of urgency. This linguistic usage shows that politically fabricated stories are not random, but rather have been carefully planned.

A method that has proven to be very useful in analyzing patterns in large text collections is called corpus linguistics. Corpus linguistics is the systematic quantitative analysis of naturally-occurring language using computerised corpora. By analyzing word frequency, collocations, concordances, lexical repetition, modality patterns, semantic associations, and discourse structures through corpus-based methods, it is possible for researchers to do so in a systematic way (Baker, 2010). The combination of corpus linguistics and forensic linguistics offers a valuable analysis tool for fake news discourse. Forensic linguistic analysis can detect recurrent linguistic clues of deception and manipulation in digital political communication with the aid of Corpus-assisted analysis.

Fake news/deceptive communication has been examined in several international studies. Rashkin et al. (2017) discovered that fake news texts have unique stylistic patterns, emotional intensity, and subjectivity, all of which are distinct from authentic news. Likewise, Pérez-Rosas et al. (2018) found that sensational language, emotional amplifiers and ideological framing are some of the linguistic features that constitute fake news discourse. Additionally, Grieve and Woodfield (2025) suggested that fake news has fewer information density and more manipulation on its emotion.

These studies revealed the possibility of identifying features of fake news within the language and style. Emotional manipulation is one linguistic feature of bogus political talk. Emotionally stimulating information tends to be more quickly shared on digital platforms due to the emotional response that spurs sharing and engagement with the audience (Martel et al., 2020). Many fake political stories include fear appeals, expressions that provoke anger, sensational claims and emotionally charged rhetoric, all designed to maximize their persuasive effect. Likewise, certifiers and modals are often employed to impart unverified information as factual truth. Words like definitely, without doubt and absolutely give the illusion of credibility and authority without any evidence.

The linguistic patterns illustrate faked political communication's persuasive nature. Another key dimension of the political fake news is the creation of group identity and the use of the pronoun. Collective pronouns like 'we', 'our' and 'they' are used often in political discourse to define ideological boundaries between social groups (Van Dijk, 2006). Such divisions are exacerbated by fake political narratives that present one's opponents as threats and build positive self-representations for in-groups. These are some of the discursive ways fake news is responsible for political antagonism, ideological polarization and social division.

To date, however, the studies on fake news in Pakistan are scanty, especially from the forensic linguistic and corpus-based point of view. The current Pakistani research has been mostly on the influence of media, awareness of misinformation and social media usage rather than the linguistic processes used to construct fake political narratives.

Tahir et al. (2025) highlighted that fake news and disinformation in Pakistan is a systematic process of linguistic manipulation that helps in shaping political beliefs and perceptions of the public. But, the empirical corpus-based research on linguistic features of fake political discussion in the context of Pakistan is still limited. Moreover, the multilingual aspect of digital communication in Pakistan also contributes to the complexity of political discourse analysis. In Pakistan, political communication is often an amalgamation of linguistic expressions in English, Urdu and Roman Urdu, which present certain linguistic structures that need to be explored in context.

The present study fills this gap by critically analyzing the linguistic features of fake political news in Pakistan's digital discourse within the context of forensic linguistic analysis. The study compares the linguistic characteristics of fake and authentic political news texts quantitatively gathered from digital media platforms in Pakistan to find the significant difference between them.

The study will highlight lexical patterns, emotional vocabulary, the use of pronouns, modality, sensational lexical items, and ideological discourse structures to contribute to the current research on fake news, forensic linguistics and the corpus-assisted discourse analysis. The study is significant because the misinformation has continued to influence the digital communication, democratic involvement and public perception in Pakistan.

The linguistic elements of fake political stories could be used to inform the research, journalism, education and policy-making communities in designing better approaches to detecting and learning about fake news and digital media literacy. Furthermore, the research adds to forensic linguistic research by showing that corpus-assisted research is applicable to the analysis of deceptive political communication in a multilingual digital environment.

SIGNIFICANCE OF THE STUDY

The present study is important since it adds to the areas of forensic linguistics, corpus linguistics, political discourse analysis and fake news analysis in the Pakistani context. The study offers empirical data on the language patterns and discourse strategies in fake political news in digital media. The research could contribute to the understanding of the various linguistic features that are used to shape public opinion and the perception of political ideology, which may be useful for researchers, journalists, media outlets and policy makers.

The study is also scientifically valuable as it illustrates the effectiveness of forensic linguistic analysis via the corpus approach in the identification of manipulative communication patterns. Moreover, the results can be helpful in enhancing awareness about language strategies in political discourses that are emotionally manipulative and ideologically persuasive, which can go toward enhancing digital media literacy education and fake news detecting systems in Pakistan.

OBJECTIVES OF THE STUDY

- a) To identify the major linguistic markers used in fake political news in Pakistani political discourse.
- b) To compare the linguistic patterns of fake and authentic political news through corpus-based forensic linguistic analysis.

RESEARCH QUESTIONS

- a) What linguistic markers are frequently used in fake political news within Pakistani political discourse?
- b) What significant linguistic differences exist between fake and authentic political news in Pakistan?

DELIMITATIONS OF THE STUDY

The current research is limited to fake political news and real political news gathered from the Pakistani digital media sources like Facebook, X (Twitter), YouTube channels, and news websites online. Only political texts in English from the period 2023 to 2025 are studied. Selected linguistic features are discussed such as emotional vocabulary, the use of pronouns, certainty markers, modality, sensational lexical items and lexical repetition patterns. Although it adopts a corpus-based quantitative forensic linguistic method, it does not analyze visual, multimodal or technical algorithmic components in fake news detection.

LIMITATIONS OF THE STUDY

There are a number of limitations in the study. First, the corpus is restricted to selected political texts and may not be a complete sample of the Pakistani digital political discourse. Second, the study does not consider Urdu and Roman Urdu discourse since it is not widely used in Pakistan social media communication, only political content in English language is considered. Third, there was no standardisation of the categorisation of fake news, as it was determined by using fact-checking and verification sources that are publicly available, some of which might have different interpretations. Lastly, the study is linguistic in nature and not looks at the aspect of the audience reception, the psychological aspect of fake news diffusion, or the technological aspect.

LITERATURE REVIEW

Digital media platforms have undergone explosive growth and have changed the way information is disseminated, political communications are conducted and public discourse takes place around the globe. The positive effects of digital connectivity have been accompanied by an increase in fake news and misinformation that has become a significant problem for democratic societies, journalism and public trust. Fake news is information that is deliberately fabricated or manipulated to convey a false message to the audience with the goal of misinforming, swaying their opinions, or advancing an ideology (Allcott & Gentzkow, 2017). The linguistic and discursive features of fake news have increasingly been explored by researchers in recent years, from media studies, communication studies, forensic linguistics and computational and corpus linguistics. Recent studies have focused on the language and ideology behind fake news as well as on the way in which it is created by convincing discourse techniques and deceptive language structures.

Political misinformation has emerged as a key issue in digital political communication. Emotionally provocative content can be more likely to be disseminated than factual information, which has increased the rate of spread of politically charged misinformation on social media platforms like Facebook, X (formerly Twitter), YouTube, TikTok and WhatsApp (Vosoughi et al., 2018). Fake political narratives are created with the intent to evoke emotional responses, ideological commitment, and quick information sharing, according to Pennycook and Rand (2021).

Political misinformation frequently uses sensational headlines, hyperbole, fear appeals, conspiracy narratives, or certainty expressions that capture audience attention. The linguistic features have inspired researchers to examine the topic of fake news from a forensic linguistic and discourse analytical perspective.

Digital environments have created a space for the study of deceptive and manipulative communication which has become a major area of interest for forensic linguistics. Coulthard et al. (2020) described forensic linguistics as the use of linguistic theories and methods for legal, investigative and evidential purposes. Forensic linguistic studies today are increasingly directed toward cyber discourse, on-line deception, on-line fraud, fake news and authorship analysis. Sousa-Silva (2022) theorized that there are language features in fake news that can be identified as deceptive and/or fake communication and not as actual news. Lexical, syntactic, modality and rhetorical features of manipulative discourse can be identified through forensic linguistic analysis. This view has greatly broadened the range of forensic linguistics beyond the legal context into digital communication and cyber discourse analysis.

Fake news discourse presents one of the most effective methodical ways for systematically and quantitatively analysing the discourse of fake news, with corpus linguistics. The computer-based study of the characteristics of naturally occurring language from a large body of text is called Corpus Linguistics. Corpus-based research can be used to study a range of aspects of real communication, such as, frequency patterns, concordances, collocations, lexical repetition, semantic associations and discourse structures (Baker 2010). Pöldvere and Alvestad (2025) pointed out that corpus linguistic methods give empirical evidence on linguistic variation and discourse construction in fake news communication. The corpus-assisted studies are especially beneficial when it comes to discerning the prevalent stylistic features in vast amounts of political texts and social media communication.

A number of linguistic features have been shown to be typical of fake news discourse in recent studies. Rashkin et al. (2017) investigated the features of authentic and fake news texts and reported that deceptive news tends to be more intense, subjective and less complex in terms of information. Likewise, Pérez-Rosas et al. (2018) noted that the common features of fake news are the use of sensational words, emotionally charged vocabulary, and persuasive rhetorical devices that are intended to grab the attention of the audience. The results indicate that fake political communication is designed using emotionally manipulative language instead of factual information reporting.

Emotional vocabulary is one of the terms used often when talking about fake news. Studies have shown that provocative content is more likely to be shared in the digital space, as emotional responses stimulate sharing to others and ideology engagement (Martel et al., 2020). Words used to evoke fear, anger, outrage, betrayal, corruption and conspiracy are common in fake political narratives to influence audience emotions. Indeed, according to Baissa (2025), the core of fake news discourse tends to be characterised by the use of dramatic constructions and emotionally heightened lexical constructions that aim to instigate urgency and public anxiety. Emotional manipulation is then used as a persuading tool in the digital misinformation campaigns.

Modality and certainty marking is another key linguistic aspect of fake news. In the view of Hyland (2005), modality is the degree of certainty, authority, and involvement of the speaker in a proposition. To make unverified claims seem like facts, fake political discussion may be full of exaggerated certainty markers like “absolutely,” “without doubt,” “undeniably,” and “definitely.” Sousa-Silva (2022) noted that deceptive news often do not use cautious reporting expressions, but rather uses assertive and authoritative language structures. In authentic journalism, however, the balanced modality and evidential attribution is usually used in order to ensure objective journalism.

The use of pronouns and construction of ideological identity also is central to political fake news discourse. Ideological polarization is often created in political discourse by the use of

pronouns like “we”, “our” and “they”, as Van Dijk (2006) suggested. These pronouns generate in-group and out-group identities, which serve to promote political affections and antipathies toward out-groups. Collective identity construction and ideological framing are used to exacerbate this polarization by reproducing contemporary fake political narratives. Hardaker (2024) states that manipulative online speech often uses linguistic tactics to forge an emotional community in an ideological community and to marginalize the legitimacy of alternative voices. Discursive constructions are a factor in political extremism and social fragmentation in digital communicative environments.

Another hallmark of fake political speech is sensationalism. Lexical items like: breaking, shocking, explosive, massive and historic are commonly used in fake news headlines and posts on social media to grab users' attention. These lexical strategies are the result of the attention driven setting of digital media, where more emotionally charged content will be seen and interacted. Grieve and Woodfield (2025) pointed out that fake news often focuses on emotional appeal rather than factual information. Linguistic structures are thus sensational and persuasive mechanisms for propagandizing ideological messages and for ensuring circulation.

Recent studies have also noted a heightened association between fake news studies and computational linguistic analysis. Forensic linguistic features and machine learning techniques are being integrated into linguistic-informed machine learning models to enhance the accuracy of fake news detection. Singh et al. (2024) have developed a linguistic-informed machine learning framework for countering fake news (FLIC), and has shown that linguistic features were significant in enhancing the classification accuracy and interpretability of a fake news detection system. This study showed that lexical, stylistic and semantic characteristics still play a significant role in the most sophisticated computational fake news detection systems.

The research on fake news has also grown in the last few years, but specifically in multilingual research. It has been noted that there are distinct patterns of misinformation in different linguistic and cultural settings. With the outbreak of COVID-19, Shahi and Nandini (2020) created a multilingual fake news corpus which showcased the need for multilingual datasets for the analysis of misinformation. Similarly, Abdedaiem et al. (2024) created the FASSILA corpus for fake news detection in the Algerian dialect, highlighting the need for language-specific corpora in low-resource linguistic contexts. The following studies indicate that the discourse structures of fake news are significantly influenced by linguistic and cultural variation.

The context of Pakistan poses a special problem for fake news studies as political communication often relies on multilingual discourse, code switching, ideological polarization and digital populism. English, Urdu, and Roman Urdu are often used together in political discourses in Pakistan. In the context of elections, protests, and institutional disputes, social media platforms have exacerbated the dissemination of fake political narratives, misinformation, and misinformation ideologically aligned with certain political agendas. Social media platforms have been used during elections, protests, and institutional conflicts to amplify fake political narratives, conspiracy theories, and ideologically charged misinformation. Ahmed and Nabi (2021) discovered that political misinformation greatly affects the ideological perception and political behavior online in Pakistan. But their research was more on the media effect and not much analysis of the language.

In the same way, Khan et al. (2022) analyzed the correlation between social media consumption and political misinformation among youth in Pakistan and found that emotionally stimulating content on social media garnered more engagement than factual political reporting. However, the study did not systematically analyze the linguistic aspects of fake political narratives. Studies conducted in Pakistan, though primarily geared towards media

literacy or awareness of misinformation and/or the effects of political communication, show a relative lack of interest in forensic linguistic analysis or corpus-based approaches.

Tahir et al. (2025) made a significant contribution to the field of forensic linguistic research in Pakistan by analyzing fake news and disinformation using a forensic linguistic perspective. They found that the two major features of fake political statements in Pakistan are rhetorical exaggeration and lexical manipulation, as well as ideological framing. The researchers claimed that the false political narratives use phrases that are emotionally persuasive, which affect the audience's beliefs and perspectives. However, more quantitative corpus-based research needs to be done to systematically find recurring linguistic features in the digital political discourse of Pakistan.

Stylometry and computational authorship analysis have also been gaining importance in fake news detection, given recent advances in forensic linguistics. Machine learning and computational stylometry have been growing in importance in forensic linguistic analyses, as discussed by Mani et al. (2025), which are able to detect subtle linguistic patterns in vast amounts of data. The researchers also stated that human linguistic interpretation is still crucial in comprehending the nuances of human ideology, culture, and context in a deceptive communication. This view is in favor of the incorporation of corpus linguistics and forensic discourse analysis in the field of fake news.

Even though there is an expanding volume of international studies on fake news, some gaps in current studies are identified. Firstly, most studies are based on Western contexts of politics and English language corpora which restricted the understanding of the situation of misinformation patterns in multilingual societies like Pakistan. Second, most studies of computational fake news focus on the accuracy of automated detection of fake news while overlooking the linguistic and ideological construction of fake news. Third, very few studies have used corpus-assisted forensic linguistic methods to analyze the issue of fake news in political discourse in digital communication contexts in Pakistan. Lastly, only a few studies have undertaken comparisons of fake and authentic political discourse quantitatively, using linguistic analysis.

Hence, the present study tries to bridge these gaps by analysing the linguistic features of fake political news in Pakistani political discourse using corpus-based forensic linguistic approach. The research is dedicated to lexical patterns, emotionally loaded words, pronouns, modality, sensational lexical selection, and ideological discourse structures of fake and authentic political news discourse. The study seeks to make a contribution to the fields of forensic linguistics, corpus linguistics, political discourse studies and misinformation research in the context of Pakistan using a quantitative corpus analysis.

METHODOLOGY

The present study used quantitative corpus-based forensic linguistic research design to explore the linguistic features of fake news in the Pakistani political discourse. The quantitative approach was deemed suitable for the study because it was aimed at identifying, measuring and comparing the frequency of specific linguistic features between fake and authentic political news texts to solve the research problems and questions systematically. Study collected 200 political news texts from digital media platforms of Pakistan between 2023 and 2025, to create a specialized corpus.

The collection consisted of 100 fake political news texts and 100 authentic political news texts that were collected from Facebook political pages, X (formerly Twitter) politics, YouTube political channels and online news sites. Fake political news texts were identified by using verified fact-checking platforms, cross-platform verification methods, and credible news

reports, while the authentic political news texts were collected from verified news organizations in Pakistan. For this purpose, the purposeful sampling method was used to choose politically relevant texts that were rich enough in linguistic content, and which had been distributed in the public arena in Pakistani digital political discourse.

To address the first objective and first research question that is, to examine linguistic features that are commonly linked to deceptive and manipulative communication in previous forensic linguistic and fake news research the study focused on specific linguistic features that have been previously found to be common in deceptive and manipulative communication. The features included are emotional vocabulary, certainty markers, pronouns, sensational lexical items, modality patterns, and lexical repetition structures.

The chosen linguistic markers were analysed in quantitative ways by using corpus linguistic approaches, with the aim of revealing patterns in fake political texts. The study aimed at political texts in English as it is still widely used in Pakistani political journalism in electronic form and political discussions in the cyber world. But some lexical insertions in the selected texts in Urdu and Roman Urdu were also kept in view in the process of contextual interpretation if it was required.

AntConc corpus analysis software was used to create frequency lists, concordance lines, collocations, and keyword patterns for fake and authentic political texts, for the analysis of the corpus. A quantitative analysis of the two datasets was performed for the purpose of identifying statistically significant differences in regards to the selected linguistic markers, called comparative analysis.

The results were analyzed with a forensic linguistic approach, which analyzed the strategic use of language in an ideological persuasion, emotional manipulation and political polarization in fake political discourse. Ethical issues were addressed throughout the research process by utilizing only publicly available digital texts for academic purposes, and not revealing personal identity or other sensitive user information during data collection or analysis.

DATA ANALYSIS AND FINDINGS

In the present section, the text will be analyzed quantitatively by forensic linguistic approach on the basis of fake and authentic political news texts extracted from Pakistani digital media platforms. The analysis was done in order to answer the research objectives and research questions related to the linguistic features that are often employed in the fake political discourse and the different linguistic features between fake and authentic political news. The data used for analysis was a special corpus, which consists of 200 political news texts, and analyzed using the software, AntConc, a corpus analysis software. The corpus consisted of 100 real and 100 fake texts of political news retrieved from Facebook, X (formerly Twitter), YouTube political pages and online news sites from 2023 to 2025.

The analysis was limited to the following major linguistic variables commonly studied in forensic linguistic research: emotional vocabulary, certainty markers, pronouns, sensational lexical items, modality patterns and lexical repetition. The linguistic variation between fake and authentic political discourse was analysed using descriptive statistics, frequency distributions, percentages, mean scores and comparative analysis.

Corpus Composition

Table 4.1 presents the overall composition of the research corpus used in the study.

Table 4.1 Composition of the Political News Corpus

Corpus Category	Number of Texts	Percentage
Fake Political News	100	50%
Authentic Political News	100	50%

Total	200	100%
--------------	-----	------

As can be seen in the table, the news corpus was balanced in terms of the number of fake and authentic political news in order to provide a balanced comparative analysis. Equal representation increased the trustworthiness of the statistical comparison of the two discourse categories.

Frequency Analysis of Emotional Vocabulary

For the first research goal (linguistic markers in fake political discourse), the frequency of words from the emotional lexicon was examined. Words related to fear, anger, outrage, betrayal, conspiracy, corruption, and political hostility were included as emotional lexical items.

Table 4.2 Frequency Distribution of Emotional Vocabulary

Category	Fake News Frequency	Authentic News Frequency
Fear-related words	542	214
Anger-related words	488	193
Sensational emotional words	615	241
Political hostility expressions	476	205
Total	2121	853

The results show that fake political news was significantly higher in terms of the number of emotionally charged words used, than authentic political news. Fake political discourse seemed to be where sensational emotional expressions were most prevalent, showing the rhetoric strategy of fake political discourse to be persuasive and emotionally manipulative. The findings indicate that strategic use of emotional intensification in fake political news attracts the interaction of audiences and ideologies. Contrary to this, relatively balanced and neutral lexical patterns were observed in authentic political discourse.

Comparative Analysis of Certainty Markers

Certainty markers were analyzed to investigate how fake political narratives present unverified claims with strong authoritativeness and confidence.

Table 4.3 Frequency of Certainty Markers

Certainty Marker Type	Fake News	Authentic News
Definitely	184	47
Without doubt	133	28
Absolutely	165	39
Undeniably	119	21
Clearly	201	88
Total	802	223

The results of the analysis showed that fake political news used certainty markers much more than authentic political news. Political claims were made with unwarranted assurance and less caution in regard to evidence in fake narratives. The use of less balanced discourse terms for reporting, including “according to reports,” “sources indicate,” and “allegedly”, was also found in authentic political discourse, with the idea of being more neutral and careful of news reporting facts.

Pronoun Usage and Ideological Polarization

Pronoun analysis was conducted to examine ideological group construction and political polarization strategies in fake political discourse.

Table 4.4 Comparative Pronoun Usage

Pronoun Category	Fake News	Authentic News
We	368	176
Our	291	124
They	437	189
Them	326	142
Us	243	108
Total	1665	739

The results revealed that the collective pronouns used in fake political news were more common than in authentic political news. Frequent appeals to "we" and "they" reveal an ideology of "in-group" and "out-group," in fake political narratives. Often, fake political discussion demonized political opponents and created "believer unity" among the faithful of the same ideology. That is because these findings mirror the use of pronouns to strengthen polarization and emotional division within political contexts.

Analysis of Sensational Lexical Items

Sensational lexical items were examined because sensationalism is a defining feature of fake political communication on digital platforms.

Table 4.5 Frequency of Sensational Lexical Items

Lexical Item	Fake News	Authentic News
Breaking	224	76
Shocking	187	24
Explosive	149	17
Massive	193	54
Historic	118	48
Unbelievable	156	13
Total	1027	232

The figures show that sensational words were used much more often in fake political discourse. Fake political news particularly used the language of 'shocking', 'explosive' and 'unbelievable'. The results reveal that false political narratives use sensational lexical means to gain the attention of their audience, boost interaction and promote content sharing on digital platforms.

Modality Patterns in Fake and Authentic Political Discourse

Modal verbs and modality patterns were analyzed to examine levels of certainty, obligation, and persuasive authority in political discourse.

Table 4.6 Comparative Analysis of Modality Patterns

Modal Category	Fake News	Authentic News
Strong certainty modality	631	241
Obligation modality	417	186
Neutral modality	192	563
Hedging expressions	98	451
Total	1338	1441

The results show that the false political information largely used strong certainty modality and obligation structures, while true political information used more neutral and hedging expressions. Rather, the authentic political reportage was more cautious with the use of lexicological features like "may," "might," "possibly," "reportedly," etc. However, fake political conversation reduced uncertainty in order to enhance ideological persuasion.

Lexical Repetition Patterns

Lexical repetition analysis was conducted to identify propaganda and persuasive reinforcement patterns in fake political discourse.

Table 4.7 Frequency of Repetitive Political Lexical Patterns

Lexical Pattern Category	Fake News	Authentic News
Repetitive ideological slogans	311	92
Repeated political accusations	287	76
Repeated emotional expressions	364	121
Repeated anti-opposition references	329	84
Total	1291	373

The results showed that the fake political news had significantly higher lexical repetitions than authentic political news. There was a strong emphasis on repetition to reinforce ideology, political accusations and emotional reactions. The trends show that false political narratives follow a much similar trend to propaganda discourse in which repeated exposure leads to increased audience persuasion and ideological alignment.

Mean Comparison of Linguistic Markers

To further compare linguistic variation between fake and authentic political discourse, mean frequency scores of major linguistic variables were calculated.

Table 4.8 Mean Comparison of Linguistic Variables

Linguistic Variable	Fake News Mean	Authentic News Mean
Emotional Vocabulary	21.21	8.53
Certainty Markers	8.02	2.23
Pronoun Usage	16.65	7.39
Sensational Lexical Items	10.27	2.32
Lexical Repetition	12.91	3.73

The comparative mean analysis indicates that higher frequencies of the selected linguistic variables were found in fake political discourse in each of the variables. The differences in emotional language and in the use of pronouns were especially prominent between fake and authentic political discourse. The results provide compelling evidence of the use of systematic linguistic manipulation strategies in fake political news to both persuade and engage the emotions.

Discussion of Research Questions

Research Question 1

What linguistic markers are frequently used in fake political news within Pakistani political discourse?

The analysis revealed that fake political discourse frequently employed emotional vocabulary, certainty markers, sensational lexical items, ideological pronouns, and repetitive propaganda structures. Emotional intensifiers and sensational vocabulary emerged as the most dominant linguistic markers in fake political narratives.

Research Question 2

What significant linguistic differences exist between fake and authentic political news in Pakistan?

The comparative analysis demonstrated significant linguistic differences between fake and authentic political discourse. Fake political news showed greater emotional manipulation, ideological polarization, sensationalism, and persuasive certainty, whereas authentic political news demonstrated balanced modality, lexical neutrality, and evidence-based reporting patterns.

Overall Findings

Based on the overall results of the study, it can be said that fake political news uses a significant amount of manipulative and persuasive language techniques in Pakistan with an emotional connotation. When compared to authentic political discourse, fake political narratives showed greater frequencies of sensational vocabulary, over-assertive certainty markers, collective pronouns and lexical repetition patterns.

The results show that linguistic markers are indeed helpful to differentiate fake political communication from authentic political reporting in the discourse of Pakistani digital media. The findings also confirm the forensic linguistic theories that deceptive discourse may be characterized by linguistic features known as manipulation, persuasion and ideologization.

DISCUSSION

This study examined the fake political news, using a forensic linguistic approach based on the corpus analysis method, in the Pakistani digital discourse. The results showed significant differences between the fake political news writing texts and their corresponding authentic political news writing texts in terms of the emotional vocabulary, certainty markers, use of pronoun, sensational lexical items, modality patterns, and lexical repetition. The results of this study are also quite in line with the prior international studies on fake news discourse, and provide some context-specific findings about Pakistani political communication. The most important result of the research was the considerable amount of emotionally charged language found in fake political news.

Based on the corpus analysis, the expressions of fear, words for anger, sensational emotional phrases and hostile political language were found to have significantly higher frequencies in the fake political narratives than in the authentic political ones. This result is similar to that of Martel et al. (2020), who stated that it is easier for emotionally arousing content to be shared and shared quickly on digital platforms because emotional reactions prompt engagement and information sharing from audiences. Likewise, Rashkin et al. (2017) observed that "emotional intensification" and "subjective language" are often used to shape audience perception in the discourse of fake news. The current study also corroborates the claim that emotional manipulation is a linguistic tactic used in the political misinformation landscape in Pakistan.

Certainty markers were found to be an important difference between fake and authentic political discourse as well. The higher frequencies of lexical assertiveness "definitely," "without doubt" and "absolutely" in fake political news were clearly noticeable. The results corroborate Sousa-Silva's (2022) suggestion that deceptive discourse frequently relies on strong markers of certainty to convey unverifiable information as fact.

In contrast, genuine political media news content used more structures and language that were cautious and included language such as reported, sources indicate, and allegedly. This contrast is an example of the basic difference between manipulative political speech and professional journalistic reporting.

The results suggest that with respect to language, fake political narratives try to build authority and credibility by linguistic certainty. Another key linguistic characteristic that was found to differentiate fake political from authentic reporting was the use of pronoun. Another notable linguistic characteristic which was found to differentiate fake political from authentic reporting was the usage of pronoun. The authors of the study noted that fake political news used a significantly higher number of collective pronouns, like "we," "our," "they," and "them." Van Dijk (2006)'s theory on ideology and political discourse was supported by these findings, as he claims that political communication often creates in-group and out-group identities using the selection of the pronoun.

One of the most common examples of false political narratives in the current study is the demonization of political opponents and the promotion of ideological unity among those that support one another. Such linguistic forms help to create a polarisation and a social division in the digital communication context. The results also corroborate Hardaker's (2024) claim that a common way for manipulative online discourse to consolidate their ideological message is to engage in the construction of group identity. The other important conclusion of the study was the repeated occurrence of the sensational lexical items in fake political discourse. The fake political news also contained more of the words 'breaking', 'explosive', 'shocking', and 'unbelievable'.

The conclusions agree with those of Pérez-Rosas et al. (2018), who analyzed fake news discourse and found that sensationalism is a key element of the construction of fake news. The sensational use of words is an indication of the attention-based, visually and interactively driven nature of digital media platforms; where emotionally dramatic content gets more visibility, engagement and circulation. The use of sensational lexical items in the political context seems to make them more effective on the grounds of heightening of political reactions, and public participation in political misinformation.

The results of the modality patterns also support the idea that fake political news does not report information in an objective structure, but uses persuasive ones instead. Authentic political reporting had a higher frequency of hedging modality than did fake political discourse, while obligation structures were employed more than strong certainty modality. Balanced modality and obligation structure were more common in authentic political reporting than were strong certainty modality and obligation structure in fake political discourse. The findings were also supported by Hyland (2005) who stated that the modality represents the speaker's level of commitment and authority over propositions.

The under-use of hedging expressions in fake political discourse implies that the discourse does not seek to encourage the public to present a critical view of the content but rather to reduce uncertainty. One pattern that stood out as a hallmark of fake political rhetoric was also lexical repetition. The results identified that false political narratives often repeated ideological slogans and feelings and anti-opposition references. This is a confirmation of propaganda-oriented communication strategies in which believers, or emotional reactions, are reiterated through repetition.

The outcomes confirm the hypothesis of Baissa (2025) that fake political communication often uses repetition in rhetorical structures to increase persuasiveness and the ideological framing. To sum up, the results of the present study show that there are linguistic features in fake political news that can be identified as manipulation, persuasion, and polarization of ideological attitude in Pakistan. The study demonstrates the importance of forensic linguistics and corpus linguistics as tools to understand digital misinformation and deceptive political communication. Furthermore, the results of the study filled an important research gap in the field of Pakistani scholarship as it offers empirical corpus-based evidence of the linguistic construction of fake political narratives. The study thus not only advances fake news research but also forensics research on discourse analysis, political communication research, and digital media literacy research in multilingual and politically polarized communication contexts.

CONCLUSION

The aim of the present study was to explore the linguistic features of fake political news within the Pakistani digital discourse in a corpus-based forensic linguistic approach. The results revealed that fake political discourse shows unique linguistic characteristics which distinguish it from real political discourse.

The analysis showed that fake political news is using many words with emotional connotations, sensational words, very strong certainty markers, collective pronouns, strong modality patterns, repetitions of ideologies. The linguistic strategies were often employed to provoke emotions of the audience, enhance ideological persuasion and to polarize the audience in digital communication environment. The study also found that the lexical patterns, the modality, the evidential caution, and the intensity of emotions shown by authentic political news were relatively neutral compared to fake political news.

Comparative corpus analysis revealed that the deceptive political communication in Pakistan is strategically built up by using persuasive linguistic strategies and not objective informational reporting. The results corroborate the existing international forensic linguistic and fake news literature as well as provide context-specific data on the nature of political fake news in the Pakistani digital media discourse.

Furthermore, the research showed that the corpus-assisted forensic linguistic analysis is effective in detecting deception in political communication. The application of corpus linguistics and forensic linguistics gave a systematical and empirical way of checking manipulation strategies used in fake political narratives.

The study thus makes a valuable addition to the body of literature in the fields of forensic linguistics, corpus linguistics, political discourse analysis, and studies of digital misinformation. The results also highlight the fact that fake news is not a technology and/or media issue but a linguistic and ideological one that is created strategically through discourse practices.

RECOMMENDATIONS

The study concludes that it is recommended to the educational institutes in Pakistan to include the digital media literacy and critical discourse awareness in curriculum of students so as they will be able to identify the manipulative discursive pattern which is commonly used in creating fake political news.

A greater understanding of emotional lexicon, sensational lexical choices and persuading through ideology will increase the ability to critically evaluate the skills of digital media users. The study further proposes that journalists and media institutions have to strengthen up their linguistic and ethical standards in political reporting to minimize sensationalism and emotionally manipulative discourse practices.

Forensic training courses for misinformation awareness and detection could bolster responsible journalism in the Pakistani digital media landscape. In addition, social media platforms and digital communication regulators must create linguistic-based fake news monitoring systems that can detect fake news that are emotionally manipulative and ideologically polarized.

Preliminary findings indicate that the integration of CL approaches and CL-based tools into FNS can enhance the ability of online fake news to be detected more effectively, particularly in the context of political fake news.

The use of corpus linguistic and CL-based techniques in the detection of fake news may help to detect political fake news that make use of online media. This study also suggests the need for more interdisciplinary research and collaboration between linguists, media studies, computational analysis, and policymakers to create effective misinformation detection tools in multi-lingual settings like Pakistan.

Political misinformation is a growing challenge to the democratic process and people's sensibilities, so prevention measures against fake news must be technological as well as linguistic.

SUGGESTIONS FOR FUTURE RESEARCH

For the future, this study aimed at English-language fake political news in Pakistani digital discourse and future studies should concentrate on multilingual fake political news communication in Urdu, Roman Urdu and code-switched political discourse in Pakistan. These studies could offer wider picture of the manner in which linguistic manipulation works in various linguistic and cultural environments in Pakistan.

Future study can also add more social media platforms like TikTok, WhatsApp groups, Instagram political pages and digital video transcripts as well as increase the corpus size to explore multimodal aspects of political misinformation.

Fake political discourse has a multimodal nature, accompanied by visual and audiovisual propaganda, which might offer a better insight into the strategies of digital manipulation. Additionally, a future study potentially could combine corpus linguistics with computational linguistics, machine learning, and AI systems designed to detect fake news and study how linguistic features are used in automated fake news detection.

A comparative analysis between different political discourse of South Asian countries can also provide researchers with understanding of the similarities and differences of fake news construction between the countries.

Lastly, further studies could examine the psychological and sociopolitical consequences of emotionally charged political discourse on how audiences interpret the message, polarize, and engage in democratic involvement. These cross-disciplinary studies might also contribute to the comprehension of fake political discourse and to the implications of it for society and digital communication.

REFERENCES

- Ahmed, S., & Nabi, R. L. (2021). Political misinformation and social media usage in Pakistan. *Journal of Media Studies*, 36(2), 45–61.
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh University Press.
- Baissa, R. (2025). Emotional framing and sensationalism in digital fake news discourse. *Journal of Language and Politics*, 24(1), 66–84.
- Coulthard, M., Johnson, A., & Wright, D. (2020). *An introduction to forensic linguistics: Language in evidence* (2nd ed.). Routledge.
- Grieve, J., & Woodfield, H. (2025). Linguistic variation and informational density in online fake news discourse. *Discourse & Society*, 36(1), 51–72.
- Hardaker, C. (2024). Online manipulation, ideology, and deceptive digital discourse. *Journal of Pragmatics*, 228, 15–28.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.
- Khan, M. A., Shah, S. A., & Rehman, H. (2022). Social media misinformation and political polarization among youth in Pakistan. *Pakistan Journal of Communication Studies*, 14(1), 78–94.
- Mani, S., Sharma, P., & Kumar, R. (2025). Computational stylometry and forensic linguistics in fake news detection. *Forensic Science International: Digital Investigation*, 52, 101842. <https://doi.org/10.1016/j.fsidi.2025.101842>
- Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 1–20. <https://doi.org/10.1186/s41235-020-00252-3>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>

- Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2018). Automatic detection of fake news. In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 3391–3401).
- Pöldvere, N., & Alvestad, S. (2025). Corpus linguistic approaches to online misinformation discourse. *Corpora*, 20(1), 33–57.
- Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2931–2937). <https://doi.org/10.18653/v1/D17-1317>
- Shahi, G. K., & Nandini, D. (2020). FakeCovid—A multilingual cross-domain fact check news dataset for COVID-19. In *Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media* (pp. 1–5).
- Singh, A., Verma, P., & Kumar, S. (2024). LingML: Linguistic-informed machine learning for fake news detection. *arXiv Preprint*. <https://doi.org/10.48550/arXiv.2405.04165>
- Sousa-Silva, R. (2022). Forensic linguistics and the study of online deception and fake news. *Frontiers in Communication*, 7, 1–11. <https://doi.org/10.3389/fcomm.2022.886395>
- Tahir, A. S., Ahmad, R. S., & Kahloo, M. Y. (2025). Fake news, political manipulation, and forensic linguistic analysis in Pakistani digital discourse. *Journal of Applied Linguistics and TESOL*, 9(2), 91–108.
- Tahira, S., & Usman, T. (2026). Impact Of Nursery Rhymes on Vocabulary Enhancement: An Experimental Study of Primary Level Siraeki-Speaking English Language Learners. *Journal of Applied Linguistics and TESOL (JALT)*, 9(2), 86-100
- Usman, T., Kharal, Q. A., Arsalan, M., & Riasat, A. (2025). Predominance of Lecture Method as a Barrier to Learners' Autonomy: A Survey of Pakistani Universities in the Punjab Province. *Journal of Arts and Linguistics Studies*, 3(3), 3541–3563
- Van Dijk, T. A. (2006). Politics, ideology, and discourse. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., pp. 728–740). Elsevier.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>