

A CORPUS-ASSISTED SOCIO-COGNITIVE ANALYSIS OF POLITICAL BIAS ON YOUTUBE AND FACEBOOK

Sana Tafseer

Department of English, Division of Arts and Social Sciences, University of Education Lahore

Dr. Jabir Hussain

Lecturer (English) Department of English, Division of Arts and Social Sciences, University of Education Lahore.

Hunza Arshad

Visiting Lecturer at University of Education Lahore. Language Trainer at University of Engineering and Technology

Dr. Jahanzeb Jahan

Lecturer (English) Department of English, Division of Arts and Social Sciences, University of Education Lahore

*Corresponding author: ranaatafseer1234@gmail.com

Article Info



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license <https://creativecommons.org/licenses/by/4.0>

Abstract

This study aims to explore the linguistic markers of political bias in YouTube comment sections and Facebook posts through a corpus-assisted socio-cognitive analysis. In an era where social media platforms like YouTube and Facebook significantly shape public discourse, the study seeks to uncover the subtle ways political biases are encoded in online conversations. The primary objectives are to identify and categorize the linguistic markers of political bias and to analyze them through the framework of social cognitive theory. A specialized corpus, consisting of comments from 10 politically diverse YouTube videos, and posts from 4 Facebook groups was compiled and processed using AntConc software. By applying van Dijk's Socio Cognitive Theory, the research investigates how language in comment sections and posts reflects and reinforces political ideologies, cognitive biases, and group identities. The analysis reveals distinct patterns of ideological framing, polarizing language, and strategic discourse that shape political perceptions and align users with particular ideological stances. These findings underscore the power of online discourse in influencing political thought and identity construction. The study facilitates a deeper view of the intersection between language, social media, and political polarization, suggesting avenues for research in future on the impact of digital platforms on democratic processes and public opinion.

Keywords:

Political bias, corpus-assisted analysis, socio-cognitive theory, ideological framing, AntConc, political polarization, online discourse.

Introduction

Political identity encompasses the affiliation with a specific political party or the adoption of an ideological label that defines one's worldview and self-concept. In today's digital era, platforms such as YouTube and Facebook serve as influential arenas where individuals publicly assert, strengthen, or even challenge their political identities. These platforms, far beyond being spaces for casual interaction, function as powerful mechanisms where language profoundly shapes and molds public opinion and ideological beliefs (Abrams, 1994). According to Sapir Whorf hypothesis, our thought process is affected by the language we use (Hussein, 2012). Social psychology is created through the perception of language on social media that sometimes may create bias or prejudice among some groups. The way people use language determines the thinking of that society and the way they deal with certain issues (Jahan & Alvi, 2023). The utterance of any individual carries specific intentions and are specifically used in order to elicit desired results as mentioned in speech act theory (Austin, 1975).

Researchers suggest that language affects the behavior of people in everyday life ranging from normal to sudden violent responses (Chen, 2013). Sometimes certain groups on social media use violent language that leads to aggressive behavior and eventually perpetuates the spread of nonstop hatred among its users (Bernstein et al., 2011). The constant use of hatred in large communities leads to the habit formation of mocking and othering other communities (Huesmann & Taylor, 2006). The sarcastic language has been considered as the fashion among people especially among youth who are in touch of social media all time (Kalaba, 2014). Social media interactions also result in the formation of new linguistic terms like the word "Tabdeeli" was itself correct, meaning "change," however, its usage and context on social media is sometimes aggressive resulting in sudden outrage among its users (Romaine, 1983). Comment section of most platforms serve as firestorm for the political discourse among the supporters of various political parties (Pfeffer et al., 2014).

Freedom of speech has been interpreted in wrong way since centuries and when it is manipulated under the garb of personal benefits for achieving political goals or advantages, then it destroys the image of the country, globally (Khan, 2013). As long as the ratings and demand of people increases, the violent discourse among social media users keeps on gaining momentum in an unchecked manner (Haque, 2014). The rise of social media usage in an unchecked way can thus affect community like a virus in negative manner forming large number of violent and politically charged groups (Subrahmanyam & Greenfield, 2008). Various platforms on social media including facebook, twitter, youtube etc sometimes become a place of spreading hate speech more appropriately as "cyber hate" (Mehboob & Alvi, 2021). Thus, this study aims to analyze the relationship between political discourse and identity construction in youth through socio-cognitive lens.

Statement of the Problem

In an increasingly digital world, social media has become a vital medium for political communication among youth in Pakistan, where traditional media often fails to represent diverse voices. This study seeks to explore the usage of politically biased language in the YouTube comment section. By examining how these linguistic strategies construct, reflect, and transform political identities, this research aims to uncover their impact on political engagement. Understanding these dynamics is crucial for educators, policymakers, and activists aiming to leverage social media for fostering informed and active citizenship. This study addresses a significant gap in sociolinguistic literature, providing insights into the intersection of language, identity, and politics in Pakistan's evolving socio-political landscape.

Objective of the Study.

- To identify and categorize linguistic markers of political bias in YouTube comment sections.

- To analyze linguistic markers of political bias in YouTube comments through a sociocognitive lens.

Research Questions.

- What are the key linguistic markers that indicate political bias in YouTube comment sections?
- How does social cognitive theory help in analyzing linguistic markers of political bias in YouTube comments?

Importance of the Study.

The study will explore the lexical choices that contribute in spreading hatred among Pakistani youth. The use of corpus to analyze the interactions of youth on YouTube and other social media platforms serves as a nuanced tool to explore the relation between social media language and identity constructions.

Delimitation of the Study

Large amount of data gathering from social media for corpus creation, is a factor that may make this study limited.

Geographic Delimitation

This research will limit its analysis to Pakistani youth, excluding individuals outside of Pakistan or those who identify with other nationalities.

Social Media Platform Delimitation

Only data from YouTube will be analyzed. Other platforms like Twitter, Instagram, or TikTok will not be included, to keep the focus on platforms where political discourse is frequently observed among youth.

Linguistic Delimitation

The study will concentrate on specific types of lexical choices, such as metaphors, emotive vocabulary, and catchy phrases. Broader linguistic features (e.g., grammar structures or phonetics) will not be analyzed.

Theoretical Delimitation

The research will be guided by a socio-cognitive approach, focusing on how linguistic choices construct and negotiate identity. Other theoretical frameworks, like critical discourse analysis or sociolinguistics, will not be included in the analysis.

Content Focus

The analysis will specifically examine language within political discourse. Non-political content, even if posted by Pakistani youth, will not be included, to ensure a focused exploration of political identity formation.

Temporal Delimitation

The data collection will cover a specific period (e.g., the last two years), allowing the research to capture recent trends in political discourse among Pakistani youth on social media. These delimitations ensure that the study is feasible, relevant, and clearly focused on the role of social media language in shaping political identities within a specific population and context.

Review of Relevant Literature.

Social media as a pulse of society and culture

Social media has surged as an essential means of human expression in our day and age, permeating politics, culture, lifestyle, and perception. These online platforms such as Facebook, Instagram, and Twitter serve as channels for connections and information sharing, creating a ubiquitous network for social engagement and the sharing of changing perspectives. Social media is much more than just a tool for communication; it acts as a ferment for cultural reorientation, quietly shifting beliefs, attitudes, and conventions in previously unheard-of ways ([Razaq et al., 2023](#)).

Language: The Instrument of Social Perception Construction

The highly calibrated frequencies of society are embodied in language, a social trigger. It is a tool that can be used for more than only communicating ideas; it can also be used for impression and influence. Language is a tool that people use to change perceptions and perspectives in addition to being a means of self-expression. Such deliberate linguistic manipulation produces "discourse"—a prism through which subconscious messages and nuances of belief are propagated and assimilated. Discourse analysis approaches differ; some researchers look for structural hints in linguistic textures, while others reveal hidden ideologies integrated into everyday conversations ([Razaq et al., 2023](#)).

According to Amedie (2015), social media is a "potent conduit" that blurs traditional lines and gives everyone a forum to openly express a wide range of opinions. This openness encourages expression, but it also gives people a way to strengthen or change their beliefs. As a result, the stories that are shared on social media have a significant impact on the opinions of the audience, making the platform a place for both ideological influence and connection.

Power and Subtext: Applying Critical Discourse Analysis

According to Paltridge (2006), Critical Discourse Analysis (CDA) acts as a lens through which latent prejudice and ideological undercurrents in language can be found. CDA uncovers the hidden codes and presumptions present in speech by analyzing complex phrases. In order to understand the influence of linguistic choices in Pakistani news headlines on public attitude, Nawaz et al (2013) investigated "Media Discourse and Their Implied Dogmas." Through the models developed by Van Dijk and Kress, the researchers investigated how seemingly neutral statements can actually influence judgements. By quietly influencing the reader's understanding and casting doubt on the media's objectivity, the study found that word choice frequently reflects prejudices ([Nawaz et al., 2013](#); [Van Dijk, 1993](#)).

The study "Power, Ideology, and Identity in Digital Realities," conducted by Baig et al (2019), highlighted the significance of digital media in identity development and ideological dissemination. Relationships, family, and personal development topics frequently reflect societal ideas, according to their qualitative analysis of a few Facebook pages. Thus, language, authority, and culture work together to create prevailing societal narratives in digital space ([Baig et al., 2019](#)).

The Digital Matrix's Growth in Literacy

The idea of "literacy" has evolved beyond conventional definitions in the digital age. According to ÖZPOLAT & KÜLEKÇİ (2022), digital literacy today includes contextual knowledge that is adaptive and enables people to understand complicated information in a digitally dominated environment ([ÖZPOLAT & KÜLEKÇİ, 2022](#)). Literacy has historically required rhetoric and oratory skills. Following Gutenberg's invention of the press, literacy evolved into textual understanding, then visual and informational literacies that honed abilities in critical thinking and image analysis ([Baig et al., 2019](#)). In

line with Lanham's (1995) thesis that literacy encompasses the capacity to interact with information in a variety of media, digital literacy today denotes the ability to interpret multi-modal content fluently. Literacy now encompasses both the ability to critically digest content and to influence the conversation on digital platforms as people move from being passive observers to active participants. A broader definition of literacy and society's adjustment to digital realities are reflected in this contemporary form of literacy.

Social Belief Formation and Linguistic Powerplay

The ability of language to discreetly incorporate social beliefs into common utterances is what makes it such a powerful ideological tool. Social media users utilize well-chosen language to influence others and incorporate their beliefs into popular conversation. By examining how language contains ideological frameworks and determines societal norms, this study seeks to unravel the complexities of language. By shedding light on these relationships, the study promotes a sophisticated understanding of the complex relationships between language and power.

Van Dijk's Cognitive Prism

Discourse and Ideological Perception A cornerstone of discourse analysis, Van Dijk's socio-cognitive theory explores the three-way relationship between discourse, cognition, and social structure. Similar to Fairclough's concept, Van Dijk highlights how mental schemas can connect individual discourse with societal constructs (Youssefi & Baghban Kanani, 2013). This model clearly illustrates how ideological matrices shape language by using cognition as a crucial conduit to bring personal speech into line with the larger societal architecture. Discourse analysis goes beyond superficial interpretation with Van Dijk's paradigm, exploring the small and large factors that shape both individual and societal ideology. This is demonstrated by social media sites such as Instagram, where comments, captions, and shared photos represent and support society opinions (Razaq et al., 2023). Specifically, research employing Van Dijk's technique to analyze editorials in the British press demonstrates how ideological bias is softly expressed, subtly influencing public opinion and conversation (Van Dijk, 1998).

In accordance with Huddy (2001), widely held societal ideologies have a significant impact on social identity theory, especially when it comes to issues like politics, religion, and ethnicity. By using this paradigm to analyse David Cameron's speeches, Ghachem (2015) demonstrated how political discourse shapes public opinion by conforming to common cultural beliefs. Therefore, Van Dijk's thesis can be used as a lens to study how people's engagement with political and social discourse is shaped by their shared ideals. The description discloses the revolutionary influence of language and social media on the ideational framework of society. Discourse analysis reveals the hidden beliefs buried in language, while social media's democratized discourse provides a platform for the free interchange of ideas. By dissecting linguistic nuances, critical discourse analysis sheds light on the prejudices influencing public opinion. Furthermore, Van Dijk's socio-cognitive model illuminates the cognitive processes that underlie discourse and provides a sophisticated viewpoint on how language, ideology, and identity match. When taken as a whole, these frameworks clarify how language, particularly in social media ecosystems, both influences and reflects society values, making it a crucial field of study for comprehending contemporary culture and identity ([Razaq et al., 2023](#)).

The organization of the socio-cognitive approach of discourse

Table 1

Level of structure	Cognition	Society
Macro	Socially shared knowledge, Attitudes, ideologies, norms, values	Communities, groups, organizations
Micro	Personal mental models of (experiences of) social members	Interaction/discourse of social members

In accordance with Van Dijk's socio-cognitive theory, mental frameworks of individuals influence how speech is formed and understood in society. How people interact with and assess political or other types of discourse is greatly influenced by their cognitive processes, which include their knowledge base and ideological convictions. The ideology that a person has internalised plays a major role in how well they understand any political issue, establishing cognition as the conduit between discourse dynamics and societal impacts ([Wodak, 2014](#)).

In concise editorials across the UK press, the news about government actions was presented positively, highlighting how mental models significantly influence the perception of information. ([Van Dijk, 1998](#)). Social identity theory has been suggested as highly applicable in both religious and political spheres, as well as in shaping ethnic identity, due to the collective nature of these identities (Huddy, 2001). An analysis of David Cameron's pre-election speeches from a sociocognitive perspective further demonstrated that individuals' ideologies and shared beliefs form the foundation of their comprehension of certain political issues ([Ghachem, 2015](#)).

Methodology and Procedure

This research utilizes a mixed-methods design, combining both qualitative and quantitative approaches to examine how language on social media platforms shapes political identities in Pakistani society. The primary research methods employed are Corpus-Assisted Discourse Analysis (CADA) and Van Dijk's Socio-Cognitive Theory, which together provide an insightful framework for analyzing how language influences political ideologies and group affiliations.

Corpus Creation

For this study, a custom corpus was constructed using content from two major social media platforms: YouTube and Facebook. Ten YouTube videos were chosen to represent a diverse array of political viewpoints, ensuring that the selected material captures a wide spectrum of opinions. In addition, posts from four different Facebook groups, each known for their political engagement, were incorporated into the corpus. This selection process aimed to reflect the range of political discourse prevalent in Pakistani social media. The content from these videos and posts was extracted and formatted into plain text files for further analysis. These files were then processed using AntConc software, enabling a detailed exploration of the linguistic patterns that signify political bias, ideological positioning, and identity construction within the online discourse. This curated corpus serves as the core dataset for examining how language shapes and reflects political ideologies in the digital public sphere.

Table 2
Metadata of YouTube for corpus:

	Source		Date	Title
1	YouTube	Al Jazeera	6 August 2024	Thousands of pakistan `s imprisoned ex-PM Imran khan `s supporters rally demanding his release
2	YouTube	Firstpost	10 February 2024	Why has pakistan `s democracy failed?
3	YouTube	WION	4 March 2023	Gravitus plus: pakistan `s economic crisis explained
4	YouTube	WION	3 April 2022	Gravitus plus: The many shades of Imran khan
5	YouTube	Al Jazeera	15 June 2023	Who is imran khan ? start here
6	YouTube	Al Jazeera	7 December 2023	Why are so many afghans being kicked out of pakistan ?
7	YouTube	Dhruv Rathee	17 May 2023	Imran khan vs pakistan army. Who will win ?
8	YouTube	Dawn news english	22 oct 2024	Dissecting Pakistan's 26th Constitutional Amendment Explained Dawn News English
9	YouTube	DW news	Oct 27 2024	Pakistan swears in new chief justice - What does this mean for democracy? DW News
10	YouTube	Dawn news english	22 oct 2024	Dissecting Pakistan's 26th Constitutional Amendment Explained Dawn News English

Metadata of Facebook for corpus:

Table

S. No	Statements	Likes
1	I Hate Imran Khan (The Master of U-Turn	(+237000)
2	کنجروں کی پہچان ، ہلے کا نشان (The symbol of the Bat represents prostitutes)	(+500000)
3	I Hate Nawaz Sharif- The Master of Noora Kushti	(+323000)
4	غیرتوں کی پہچان، شیر کا نشان (The symbol of the Lion represents Shameless people)	(+500000)

Tools for the data collection and procedure of administration of tools.

Manual data gathering from YouTube videos and Facebook posts for the creation of corpus. The analysis of data is done using Antconc.

3. Data Analysis

For data analysis corpus tools will be employed supplemented by socio-cognitive theory of Van Dijk for detailed analysis.

● **Figure: 1**

The following results suggests that commenters see current political dynamics through a prism of resistance against perceived oppression and control, emphasizing a sense of battle between the public and the system. This approach focuses on how social media speech conveys and reinforces ideological positions, in-group solidarity, and sociocognitive biases.

File	Left Context	Hit	Right Context
1 Music.docx	imran khan is the only honest PM of Pakistan our	army	is not for Pakistan the original name of Pakistan
2 Music.docx	are wrong you does not say all truths about pakistan	army	is not involved in politics are you shure confirm
3 Music.docx	for second quaid azam or imran khan. This na pak	army	is not leaving us soon Reply @ibrahimrana-h9n 4
4 Music.docx:months ago 11:37 11:37	👍 Reply @kingkhan-cj1pb 2 months ago Our pak	Army	is not pak this is na pak Army in
5 Music.docx:Reply @rameenqueshi7643 1 month ago	The quoted word u used "Pakistan	Army "	is not reasonable actually U could say traitors instead
6 Music.docx:Imran Khan Zindaabad	👍👍👍👍👍 Reply @HassanHabib-m2n 2 months ago	Army	is not responsible but some generals are responsible army
7 Music.docx	of a martyred captain, I must admit that our entire	army	is not the main reason. It is the generals
8 Music.docx	must say you have done a very well analysis ,Our	Army	is responsible for all this instability in Pakistan 347 Reply 7
9 Music.docx	are absolutely right Whatever is happening in Pakistan, only the	army	is responsible for it👍👍👍👍 Reply @SAF_1YT 2 months ago
10 Music.docx	ago Dhruz rthee saying correct about our army that the	army	is responsible for all the big crises happened since 1947
11 Music.docx	ago Dhruz rthee saying correct about our army that the	army	is responsible for all the big crises happened since 1947
12 Music.docx	my army. This army is a curse on Pakistan. This	army	is responsible for the murdered of Liaquat Ali Khan.
13 Music.docx	month ago Whatever is happening in Pakistan only the dollar	Army	is responsible and their Asim muneer. 1 Reply @honeybeebadger 1 month
14 Music.docx	I have never been more ashamed of my army. This	army	is a curse on Pakistan. This army is responsible

Figure: 1

The comments reveal a polarized view of the army's role in Pakistani politics, with language reflecting strong distrust and blame. Words like "traitors" and "curse" highlight a perceived conflict between the public and the establishment, aligning with van Dijk’s socio-cognitive model, which shows how mental models of resistance against authority shape collective identity and reinforce ideological divides.

File	Left Context	Hit	Right Context
14 Music.docx	I have never been more ashamed of my army. This	army	is a curse on Pakistan. This army is responsible
15 Music.docx	t inevitable". Refusing to be a puppet in hands of	army	is a positive. I have respect for army but
16 Music.docx	ghar sa hi utha letay hain👍👍 Reply @shahfibre 2 months ago	Army	is a shit here. More power to Imran Khan
17 Music.docx	King Reply @Ahmedjarral-y7v 3 months ago No Doubt Our	Army	is Corrupt But i Believe Nawaz Shreef is still
18 Music.docx	Here it's important to add that Not all Pakistan	army	is corrupt but there are few "generals" that always
19 Music.docx	SafdarTiwana 1 month ago Being a Pakistani i wanna say ... our	Army	is corrupt and dont like a great and a
20 Music.docx	Being a Pakistani what Dhruv Rathee is saying true. Our	Army	is main cause of this situation 1 Reply @MUHAMMADQasim-1crl
21 Music.docx	what Dhruv Rathee is saying is 100% true and accurate. Our	army	is main cause of this situation we r facing 8.9
22 Music.docx	A Pakistani, I 100% Agree with what Dhruv Rathee said. Our	army	is main cause of this situation we r facing 50
23 Music.docx	ago No more nonsense about being selected. Even when the	army	is against him he received the most votes. Reply 1
24 Music.docx	just on Khan's side at that time and now	army	is against him but if army didn't interfere
25 Music.docx:es sense 92 Reply 11 replies @syedabuhairaira9022 10 months ago	Pakistan	Army	is All Behind This👍 32 Reply 4 replies @PeleRana-pp6zc 10
26 Music.docx:es sense 92 Reply 11 replies @syedabuhairaira9022 10 months ago	Pakistan	Army	is All Behind This👍 32 Reply 4 replies @PeleRana-pp6zc 10
27 Music.docx	a helpless Pakistani. There is no civil freedom in Pakistan.	army	is occupying everywhere. Reply @utv6250 9 months ago It's

Figure: 2

Using van Dijk's socio-cognitive approach, the comments on this image reveal deep-rooted beliefs about the "establishment" and its role in Pakistani politics. Many commenters use "establishment" to suggest powerful institutions, like the military, are influencing politics, particularly against figures like Imran Khan. This language reflects a view that organized forces oppose certain political groups, with commenters aligning themselves as supporters of Khan or as "the public," building an in-group identity in opposition to an out-group they see as corrupt or repressive. Expressions like "scared," "destroyed," and "losing respect" reinforce a strong us-versus-them sentiment, suggesting that the establishment is actively suppressing political dissent. The comments are ideologically divided, portraying Khan as a symbol of integrity and the establishment as upholding a negative status quo.

Words like "shameful" and "supporting PTI" convey strong emotional biases, favoring Khan while critiquing the establishment. Repetitions of terms like "establishment" and "poor people have been destroyed" underscore public dissatisfaction, portraying the elite as disconnected or even adversarial to the public's interests. References to issues like "free Balochistan" and "judges wearing khaki" reflect a shared awareness of political controversies, using these topics to highlight mistrust. This analysis shows a discourse rooted in distrust of powerful institutions, polarized ideologically, and shaped by emotionally charged language that fosters group division and solidarity.

The screenshot shows a search interface with the following details:

- Search Query:** establishment
- Context Size:** 10 token(s)
- Total Hits:** 67
- Page Size:** 100 hits
- Page:** 1 to 67 of 67 hits

File	Left Context	Hit	Right Context
1 Music.docx	Army Staff. The head of the army. Today, this Pakistani	establishment	is after Imran Khan with its full might. They
2 Music.docx	Army would never allow it. Reply @MasabAhmads 8 months ago military	establishment	is going to punish Pakistani Poor Nation Reply @BrindaBaskar 7
3 Music.docx	39 Reply 13 replies @tousifqammar3153 1 year ago As a Pakistani Military	Establishment	is responsible for this whole unrest this thing happened
4 Music.docx	Thanks for the video bro. I am Pakistani and the	establishment	is ruining this country badly. IK is the true
5 Music.docx	Whole country stands by Imran Khan. Only national and international	establishment	is scared of his upright position. He is a
6 Music.docx	no one accept Allah A man from whom whole corrupt	establishment	is scared badly 75 Reply 2 replies @AbdulRehman-eq3x 1 year
7 Music.docx	for every pakistani to watch and understand the background behind	establishment. 2	Reply @aarthyo9 8 months ago Democracy is like a good
8 Music.docx	The only difference is, now he being supported by the	establishment.	Reply @fahmad7194 10 months ago Judges are wearing khaki underwear
9 Music.docx	based on reality. Pakistani politicians often rely on the Pakistani	Establishment.	Reply @malzahoor7265 8 months ago love from pk 5 Reply @seshu4434 8 mor
10 Music.docx	Thanks WSK, very detailed video and shameful behavior of our	establishment	Reply @shahzarsah1829 2 years ago The power pf a Man.
11 Music.docx	poor people have been destroyed by the current government and	establishment	Reply @SHOAIBSAICH 1 month ago اس وقتوں میں شعیب اختر
12 Music.docx	ago free balochistan Reply @Waqas12-z6o 1 month ago Pakistani	establishment	and their imported government are losing out their respect
13 Music.docx	the public 92% Pakistani are supporting PTI . Feel shame for the	establishment	and their poppets government. 7 Reply 2 replies @JainalAbedin-f4n 2
14 Music.docx	Reply @aymaakram2380 8 months ago Now, Pakistani people are tired of	establishment.	And are trying to change things. One day Pakistan

Search Options: Words Case Regex **Results Set** All hits **Context Size** 10 token(s) Adv Search

Sort Options: Sort to right | Sort 1 1R | Sort 2 2R | Sort 3 3R | Order by freq

Figure: 3

According to Teun A. van Dijk's socio-cognitive framework, bias in remarks is analyzed through linguistic, social, and cognitive elements. Here's a summary based on his approach: **Mental Models:** Comments reflect individual ideologies, with words like "biased" indicating a belief that media follows set agendas aligned with or against these ideologies. **Social Identity:** In-group and out-group biases appear, especially through nationalistic terms like "Pakistan" and "India," signaling biases based on cultural and political affiliation. **Group Relations:** Statements often express tension between groups, particularly when one group feels misrepresented by the media. Terms like "biased" and "controlled" highlight ideological divides, with individuals dismissing opposing perspectives as propaganda. **Lexical Choices:** Negative terms like "biased" and "propaganda" are used to undermine opposing views, reinforcing ideological divides. **Structure and Coherence:** Comments often lack nuanced arguments, favoring short, emotionally charged statements. **Contextualization:** References to people or events (e.g., "IK" or "Dharna politics") imply shared political understandings among the audience, reinforcing specific viewpoints.

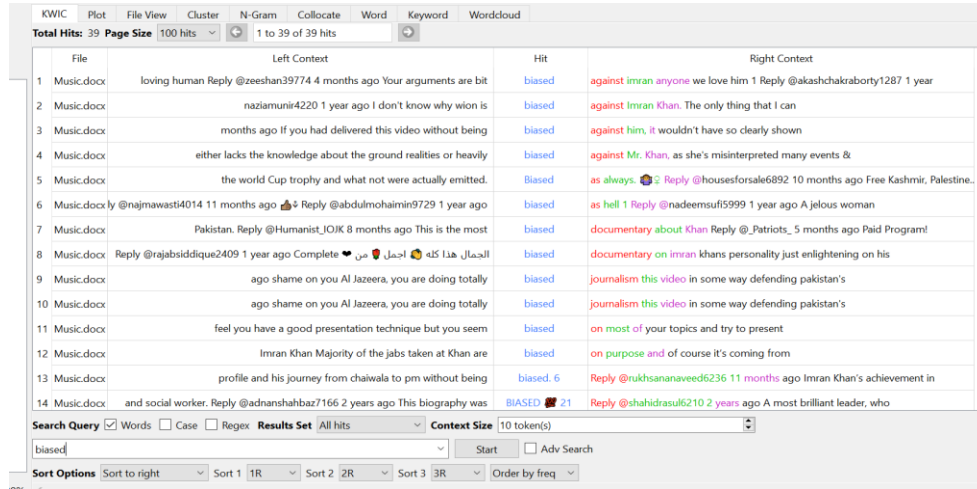
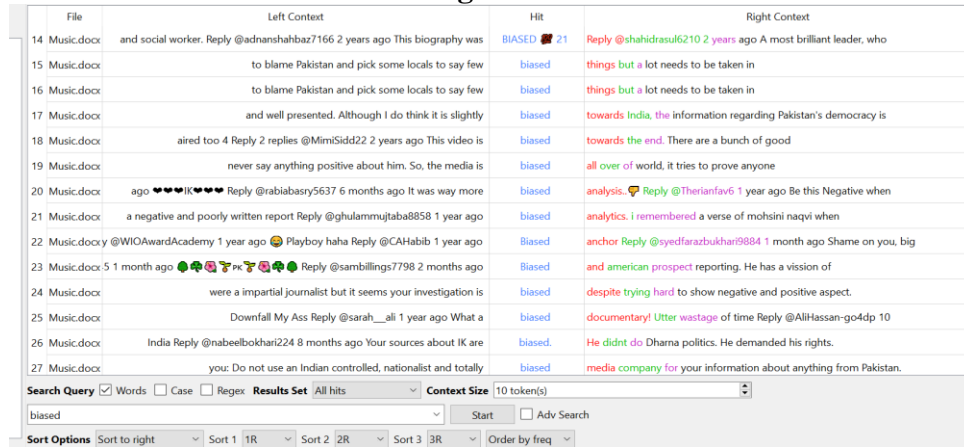


Figure: 4



This analysis highlights a discourse marked by nationalistic and ideological bias, simplified language, and emotionally charged terms, all fostering in-group loyalty and rejection of out-group views. Social media amplifies these socio-cognitive biases, creating echo chambers that reinforce shared beliefs.

Figure: 5

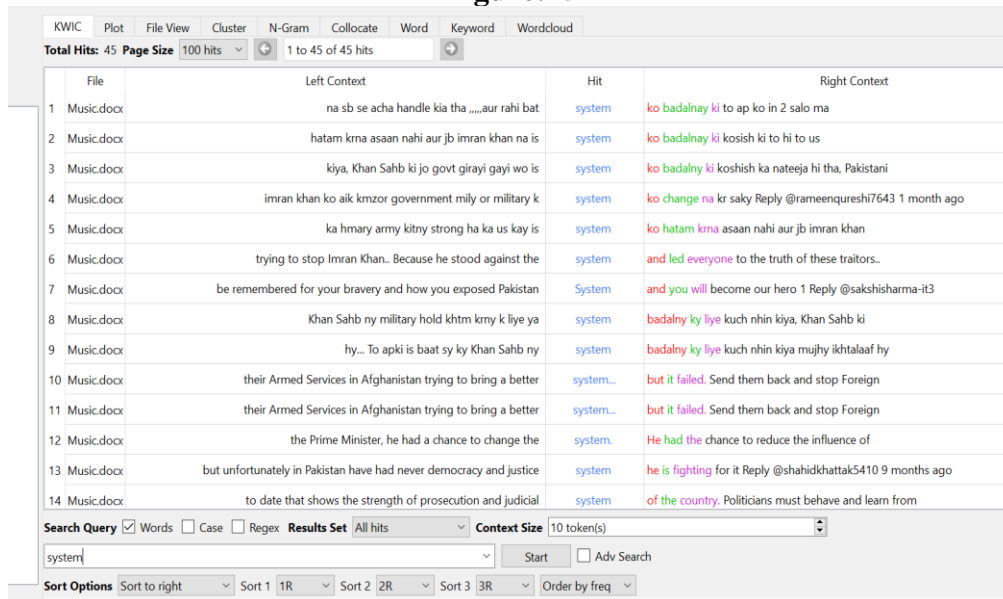


Figure: 6

File	Left Context	Hit	Right Context
1 Music.docx	country to be just our country, neither political nor Army	puppet	government 21 Reply @sameersamoon8024 3 months ago Well explained...@Dhruv R
2 Music.docx	to "even change isn't inevitable". Refusing to be a	puppet	in hands of army is a positive. I have
3 Music.docx	randomyt113 2 years ago Indians who say Khan was Army's	puppet	now saying he is fighting with army. I mean
4 Music.docx	when even your enemy said "He refuse to be a	Puppet" 394	Reply 24 replies @faseehasif2746 1 year ago His downfall?????? He's

The use of the word “army” in a negative context suggests that the public is having a negative image of the army. The word “system” is also of major concern among youth that highlights their negative views about the system as the constant demand of changing it is made. Thus, the political ideologies are conveyed negatively rather through argumentations or critical lens.

Table 4

Date	Post	Analysis
9 sep 2023	<p>نک دا کو کا پٹواریاں دا کھوتا</p> <p>Someone worthless, like the donkey of the land record officers (patwaris).</p>	<p>Blending humor with political satire, likely conveys a criticism of political leadership, specifically regarding corruption or ineffectiveness, represented through the humorous metaphor of a donkey, which can symbolize simplicity or stubbornness in South Asian cultures.</p> <p>Social Context: The humorous caption written in Punjabi, combined with a politician named as s donkey, connects with local cultural humor, appealing to a sense of shared discontent or skepticism towards leadership. Such posts often arise in societies where political figures are perceived as unproductive or corrupt, reflecting public sentiment through satire.</p> <p>Cognition: On a cognitive level, this meme may trigger associations with shared frustrations, stereotypes, and local idioms. People interpret the donkey symbolically, perhaps as a "burdened" or "stubborn" leader, which activates a shared mental model among viewers. This shared understanding reinforces the meme's impact and ensures its humor and critique are widely accessible.</p> <p>In summary, this meme uses familiar symbols, local language, and humor to create a shared understanding of political critique. By leveraging van Dijk's socio-cognitive model, we see how this post both reflects and reinforces social attitudes towards leadership in Pakistan.</p>
10 jan 2023	<p>یہ یوتھیا اپنے گھر والوں کی کفالت شہباز شریف کی بنائی ہوئی میٹرو بس پروجیکٹ میں ملازمت کر کے کر رہا ہے اور گن ایسے شخص کے گار رہا ہے جس نے اس قوم کو انڈے مرغی اور کٹے کی رام کتھا سنائی تھی۔ اور جب پرویز سندھیہ صاحب نے ہوش ٹھکانے لگاتے ہوئے اس بھائی صاحب کی درستگی کرنے کی کوشش کی اور آگے سے یہ جاہل دھمکیوں پہ اثر آیا۔ نہ صبر ، نہ تحمل، نہ برداشت اور نہ ہی سچ سننے کی ہمت اور حوصلہ۔ یہ شعور کا معیار ہے تاریخ انتشار کے یوتھیوں میں</p>	<p>In the post, “the child of tawaif” shows the level of hatred that people have for the ones who have different political views than them. The word “youthya” is also used to address this person. This illustrates the cognitive schemas formed in minds of individuals leading to “us” vs “them” mentality. Moreover, the use of emotional appeals (e.g., "no patience," "no tolerance") aims to elicit a moral judgment from the audience, guiding their cognitive processing toward condemning the actions of the individual being criticized. These framing techniques reflect how power dynamics influence people's understanding and responses to social issues.</p> <p>The speaker's use of phrases like "ignorant" and "threats" points to an othering process, where the opposing individual is framed as unworthy of rational discourse, reinforcing social divisions.</p> <p>These framing techniques reflect how power dynamics influence people's understanding and responses to social issues.</p> <p>Pronouns: The use of pronouns like "we" and "they" by different groups helps</p>

	<p>Translation The youthya is supporting his family by working in the metro bus project created by Shehbaz Sharif, and is singing the praises of a person who narrated the fable of eggs, chickens, and goats to the nation. When Mr. Pervez Sindhela tried to bring this man to his senses, this ignorant individual resorted to threats. There is neither patience, nor tolerance, nor the ability to hear the truth, nor the courage to face it. This is the level of awareness among the confused, chaotic youth.</p> <p>طوائف کا بچہ</p> <p>The child of a courtesan</p>	<p>distinguish between themselves and others, fostering an "us vs. them" mentality. This can lead to the formation of prejudice and, if unchecked, escalate into hatred.</p>
1 Nov 2023	<p>یوتھیے صرف بھونک سکتے ہیں</p> <p>PTI supporters can only bark</p>	<p>The phrase "یوتھیے صرف بھونک سکتے ہیں" uses strong, dismissive language to demean a particular group, presumably referring to youth who are critical of certain political figures or situations. From a socio-cognitive perspective, this statement constructs a negative identity for the youth, casting them as powerless and ineffectual ("only able to bark"). The use of "bhonk sakte hain" creates an image of these individuals as loud but insignificant, reinforcing stereotypes of immaturity or inability to contribute meaningfully to discussions. This type of framing serves to discredit the group, drawing a clear line between the speaker and the critic, which is a key element of van Dijk's "us vs. them" cognitive model. The language here simplifies complex social dynamics into easily digestible, polarized categories, influencing how audiences interpret the youth's actions and role in society.</p>
25 Aug 2023	<p>چوروں کا سردار قیدی نمبر اٹھ سو چار</p> <p>The leader of thieves, prisoner number 804.f</p>	<p>Polarization: Polarized ideologies arise through the use of language that divides people into ingroups and outgroups. This division impacts all levels of discourse, reinforcing social boundaries.</p> <p>Identification: Political groups form specific beliefs and ideologies by aligning with one group while disregarding the other. Group identification itself is not inherently harmful, as long as it is not rooted in prejudice or hatred.</p>
28 July 2023	<p>وہ سائفر نہیں تھا بلکہ یوتھیوں کی بہن کی ریٹ لسٹ تھی (👯) جو لہرائی تھی</p> <p>"It wasn't a cipher; rather, it was the price list of youth supporters' sister (👯) that was waved."</p>	<p>The phrase "وہ سائفر نہیں تھا بلکہ یوتھیوں کی بہن کی ریٹ لسٹ تھی" uses sarcastic and provocative language to discredit a previous event or claim (presumably the "cipher" incident) by equating it to something trivial and dishonorable. From a socio-cognitive perspective, this statement employs framing to create a negative association between the "youthya" (referred to as "youths" or "یوتھیے") and a morally questionable situation. By suggesting that the event was not a serious political matter ("سایفر نہیں تھا") but instead a "rate list" of their sisters, an attempt to made to condemn the other group.</p>
28 July 2023	<p>ایک جھوٹے کاغذ پر یوتھیوں کو پورا سال نچانے والا جوکر</p> <p>A joker who made PTI</p>	<p>The portraying of the political leader of another party as a joker leads to the discrediting of their emotions. These posts are made to devalue the views of the supporters and make them admit that they are on the wrong track. Furthermore, the supporters of Imran Khan are displayed as puppets who dance in the direction of their leader without having proper knowledge of anything.</p>

	supporters dance for a whole year over a fake document.	
28 July 2023	Every lie is being exposed, and the coward is being unmasked in every street ہر جھوٹ کا پردہ فاش ہو رہا ہے گیدڑ گلی گلی بے نقاب ہو رہا ہے	Emphasis of Positive Self-Description and Negative Other-Description: Group identities often emerge through emphasizing the positive traits of one's own group while highlighting the negative traits of others. This distinction between ingroups and outgroups is largely based on identity, with ideologies reinforcing the positive self-image of the ingroup. Ideological Square: The ideological square is a result of exaggerated focus on positive self-representation and the negative portrayal of others. This leads to viewing the actions of other political groups in a consistently negative light.
18 July 2023	"Seema went to her lover with four kids and is considered despicable, but the one who left five kids for her lover is a spiritual guide. Salute to you, Pakistanis." سیما چار بچے لیکر عاشق کہ پاس چلی گئی تو قابل نفرت ہے لیکن جو پانچ بچے چھوڑ کر عاشق کیساتھ آگئی وہ مرشد ہے۔ سلام ہے آپ پاکستانیوں کو	Cognition: On a cognitive level, this meme may trigger associations with shared frustrations, stereotypes, and local idioms. People interpret "the use of Murshad" for the wife of Imran Khan as negative as it was not her first marriage, which activates a shared mental model among viewers. This shared understanding reinforces the meme's impact and ensures its humor and critique are widely accessible. This post tries to build polarization in the audience through defaming the character of the wife of a leader ultimately in an attempt to defame their political party.
11 July 2023	The translation of the sentence "انتخابی نشان بلا ران کنجری خصم دلا" in English is: "Election symbol: Bat, slogan: Progeny of a courtesan, husband of Dalaa."	Discourse: language choice and imagery play a crucial role in delivering the message. By using relatable visual humor, this meme simplifies complex political criticism, making it accessible to a broader audience. The informal tone, paired with exaggerated visuals, indicates a playful yet critical stance, reflecting the public's disillusionment with leadership in a non-confrontational, humorous way. The use of the word "kanjari" and "dla" are showing the level of hatred that people assume for those who are against their views.
11 July 2023	فیصلہ آپ کا پاکستان کو ترقی کی "طرف لے کے جانا ہے" "The decision is yours, G. We have to take Pakistan towards progress." Or we have to make a dance at D-Chowk." "یا ڈی چوک پر مجرہ کروانا ہے"	Norms and Values: Ideologies are often communicated through the values, beliefs, and norms of a group. In the context of political groups, the norms and values of opposing groups are frequently portrayed negatively. Interests: The primary motivation behind ideological conflicts is the competition for both basic resources (such as food, shelter, and health) and symbolic resources (such as status, knowledge, and access to public discourse).
10 July 2023	جھوٹ اور نفرت پھیلانے والا شیطان خان The devil Khan who spreads lies and hatred.	The expression "جھوٹ اور نفرت پھیلانے والا شیطان خان" (The devil Khan who spreads lies and hatred) utilizes potent linguistic tools to craft a derogatory representation of Khan, manipulating both cognitive and social processes. Van Dijk's Socio-Cognitive Theory illuminates how such discourse serves to create mental representations that align with specific ideological stances. The term "شیطان" (devil) is a pejorative metaphor, imbued with moral condemnation, casting Khan as a malevolent force synonymous with deceit

	<p>and destruction. This lexical choice is designed to evoke strong affective responses, further entrenching negative biases in the audience's cognitive framework. By associating Khan with "lies" and "hatred," the speaker positions him as a corrupt, divisive figure, creating an unequivocal moral dichotomy between the in-group and the out-group.</p> <p>From a cognitive standpoint, this discourse seeks to influence the listener's schemas and mental representations, instilling a prejudiced perspective of Khan as a threat to social cohesion. It functions within the realm of ideological cognition, where language is leveraged to sustain and amplify polarized views.</p> <p>In the social domain, this rhetoric enforces group polarization and fosters social exclusion. By demonizing Khan, the speaker solidifies their own group's moral high ground, emphasizing virtue and truth in contrast to the perceived malevolence of the other. This stratification nurtures a sense of collective identity while perpetuating social and political fragmentation, reinforcing Van Dijk's notion of discourse as a mechanism for ideological and identity construction.</p>
--	--

Conclusion

In conclusion, this study sheds light on the crucial role of language in manifesting political biases within YouTube comments and Facebook posts. By employing a corpus-assisted sociocognitive methodology, it unveils the subtle ways in which online discourse influences political ideologies, cognitive biases, and group affiliations. Through the lens of van Dijk's Socio Cognitive Theory, the research identifies key patterns of ideological framing and polarizing language that shape political viewpoints and align users with specific ideological positions. The results underline the influential power of social media in shaping political thought and identity. This study offers valuable insights into the intersection of language, social media, and political polarization, suggesting potential directions for future research on the effects of digital platforms on public opinion and democratic engagement.

Recommendations

The widespread, unregulated presence of hate speech on social media calls for urgent intervention. Implementing oversight measures is essential to safeguard the country's image and foster a positive, constructive environment for youth development. This approach would not only curb harmful narratives but also promote healthier discourse on social media.

The following are some unique and concise future recommendations for the study:

1. **Examine Broader Social Media Platforms:** Future research could expand the scope to include other social media platforms, such as Twitter and Instagram, to explore how political bias is manifested in different online environments.
2. **Longitudinal Studies:** Conduct longitudinal studies to track changes in linguistic markers of political bias over time, offering insights into the evolving nature of online discourse.
3. **Cross-Cultural Comparison:** Investigate political bias across different cultural or regional contexts to understand the global variation in online political conversations and ideological framing.
4. **Impact on Political Behavior:** Study how exposure to biased content on social media influences political participation, voting behavior, and civic engagement.

5. **Focus on Counter-Narratives:** Explore how counter-narratives or politically neutral language challenge existing biases and contribute to more balanced online discussions.
6. **Incorporate Multilingual Analysis:** Include multilingual perspectives, particularly in non-English online spaces, to understand how language diversity influences political bias in social media.
7. **Automated Detection Tools:** Develop and test advanced automated tools to detect political bias in social media content, enhancing the scalability of research and real-time analysis.
8. **Influence of Algorithmic Curation:** Investigate how social media algorithms contribute to the amplification or suppression of politically biased content, and their role in shaping user exposure to ideological viewpoints.
9. **User-Generated Content and Bias:** Analyze the role of user-generated content, such as memes or user comments, in reinforcing or challenging political biases within comment sections and posts.
10. **Psychological Impact on Users:** Explore the psychological effects of consuming politically biased content on social media, particularly regarding cognitive dissonance, emotional responses, and political polarization among users.

References

- Ghachem, I. (2015). A Sociocognitive Approach to Agency Framing in David Cameron's 2010 Pre-election Discourse. *Critical Approaches to Discourse Analysis across Disciplines*, 7(2).
- Mehboob, N., & Alvi, U. (2021). An analysis of hate speech tweets against women from a socio-cognitive discourse perspective. *University of Chitral Journal of Linguistics and Literature*, 5.
- Van Dijk, T. A. (1998). Ideology: A multidisciplinary approach. In: Sage.
- Wodak, R. (2014). Critical discourse analysis. In *The Routledge companion to English studies* (pp. 302-316). Routledge.
- Abrams, D. (1994). Political distinctiveness: An identity optimising approach. *European Journal of Social Psychology*, 24(3), 357-365.
- Baig, F. Z., Yousaf, W., Aazam, F., Shamshad, S., Fida, I., & Aslam, M. Z. (2019). Power, ideology and identity in digital literacy: a sociolinguistic study. *International Journal of English Linguistics*, 9(4), 252-264.
- Huddy, L. (2001). From social to political identity: A critical examination of social identity theory. *Political psychology*, 22(1), 127-156.
- Nasir, M., Nazar, F., Jamil, M., Ahmad, J., & Ameer, Z. (2023). Reconstruction of Political Dogmas: A Critical Discourse Study of Political Pages on Facebook. *Al-Qanṭara*, 9(4), 11-41.
- Nawaz, S., Bilal, H. A., Kalsoom, M., Fayyaz, Z., & Nayyar, H. (2013). Media discourse and their implicit ideologies. *Asian Journal of Social science and humanities*, 2(2), 328-337.
- ÖZPOLAT, E. A., & KÜLEKÇİ, G. (2022). Discovering Current Digital Literacy Knowledge of EFL Prep-School Learners: An Explanatory Study. 2nd INTERNATIONAL CONFERENCE ON EDUCATIONAL TECHNOLOGY AND ONLINE LEARNING- ICETOL 2022,
- Razaq, H. R., Aslam, S., Mukhtar, A., & Shahid, C. (2023). An Exploration of Various Ideologies in Pakistani Social Media: A Critical Discourse Analysis. *Review of Education, Administration & Law*, 6(2), 227-240.
- Sekaran, U. (2016). Research methods for business: A skill building approach. In: John Wiley & Sons
- Van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & society*, 4(2), 249-283.
- Jahan, J., & Alvi, U. F. (2023). A Corpus Analysis of Hate Speech in Pakistani Political Discourse on Twitter. *Bulletin of Business and Economics (BBE)*, 12(3), 260-267.
- Austin, J. L (1975). *How to do things with words*. Harvard university press.
- Bernstein, M., Monroy-Hernández, A., Harry, D., André, P., Panovich, K., & Vargas, G. (2011). 4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community. Proceedings of the international AAAI conference on web and social media,
- Chen, M. K. (2013). The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *American Economic Review*, 103(2), 690-731.
- Haque, J. (2014). Hate Speech-A study of Pakistan's cyberspace. *Islamabad, Pakistan: bytes4all*.

- Huesmann, L. R., & Taylor, L. D. (2006). The role of media violence in violent behavior. *Annual review of public health*, 27(1), 393-415.
- Hussein, B. A.S. (2012). The sapir-whorf hypothesis today. *Theory and Practice in Language Studies*, 2(3), 642-646.
- Jahan, J., & Alvi, U. F. (2023). A Corpus Analysis of Hate Speech in Pakistani Political Discourse on Twitter. *Bulletin of Business and Economics (BBE)*, 12(3), 260-267.
- Kalaba, J. (2014). The subversive role of verbal aggression in the sarcastic language of Njuz. net. *Sic: časopis za književnost, kulturu i književno prevođenje*, 4(2), 0-0.
- Khan, A. (2013). *Pakistan and the Narratives of Extremism*. JSTOR.
- Pfeffer, J., Zorbach, T., & Carley, K. M. (2014). Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of marketing communications*, 20(1-2), 117-128.
- Romaine, S. (1983). Historical linguistics and language change: progress or decay? *Language in society*, 12(2), 223-237.
- Subrahmanyam, K., & Greenfield, P. (2008). Online communication and adolescent relationships. *The future of children*, 119-146.