

Emotional Triggers and Cognitive Manipulation in Romanian Social Media: A NLP Analysis

Ștefania-Elena Stoica

Carol I National Defense University, Bucharest, Romania

stoica.stefania@myunap.net

Abstract: Emotionally charged speech has become a key component of disinformation efforts in the era of social media. To elicit strong emotional responses, speakers strategically employ cultural and historical allusions in their highly engaging media appearances on various internet platforms and so is the case for Romania as well. This work uses Natural Language Processing (NLP) tools to identify and analyze these speeches' emotional signals and language patterns, especially those intended to evoke negative emotional reactions like hatred, fury, and fear. We aim to highlight how historical allusions are used to influence public opinion by analyzing the interaction between quick, automatic emotional reactions (System 1 thinking) and more thoughtful, rational replies (System 2 thinking) when culture and history are embedded. The results of this investigation show how speakers adeptly lead listeners from an initial emotional connection to controversial beliefs that are intellectually defined and frequently aligned with polarized societal or political goals. Our findings contribute to the understanding of how emotionally charged language fosters ideological division and perpetuates disinformation in the digital space. Moreover, the research offers a critical lens on the cognitive mechanisms that make individuals susceptible to such manipulative tactics, providing insights that could inform strategies to counter disinformation and promote more informed public discourse. This paper is part of a broader PhD research program focused on analyzing narratives and narrative strategies in disinformation, with all findings contributing to this purpose.

Keywords: Cognitive Manipulation, Disinformation, Romanian Social Media, NLP, Emotional Trigger.

1. Introduction

In the online age, disinformation campaigns have leveraged the vast reach of social media to manipulate public opinion on an unprecedented scale. Social media platforms, while fostering connectivity and democratizing access to information, have simultaneously become fertile ground for spreading disinformation among the Romanian population. These campaigns rely heavily on emotionally charged rhetoric, a tactic that taps into audiences' psychological and cognitive vulnerabilities. Such rhetoric is particularly effective in shaping perceptions and driving engagement, as it triggers instinctive emotional responses like fear, anger, and hate, bypassing critical reasoning processes.

This phenomenon takes on a culturally specific dimension in online Romanian-speaking communities, with disinformation campaigns strategically exploiting collective historical memories and cultural narratives. Events with profound societal significance, such as the 1989 Romanian Revolution and the communist regime, are invoked to evoke strong emotional reactions. These historical references are weaponized to manipulate public perception, using emotions as a tool to deepen ideological divides and align audiences with polarized social or political agendas. By leveraging such emotionally resonant themes, disinformation campaigns skillfully guide audiences from initial emotional engagement to more deliberate, cognitively framed ideologies that perpetuate division.

Despite the increasing body of research on disinformation, significant gaps still exist in understanding how emotionally charged language functions within specific cultural contexts. Vosoughi, Roy, and Aral (2018) demonstrated that false information spreads more quickly when it evokes emotions such as fear or anger, but their study does not emphasize cultural nuances. Similarly, Wojcieszak et al. (2021) emphasize the role of emotional manipulation in driving engagement on platforms like Facebook, yet they overlook how cultural and historical references enhance these effects. While Kahneman's (2011) dual-process theory explains how emotional (System 1) and analytical (System 2) thinking influence decision-making, its application to disinformation remains underexplored, particularly in culturally specific contexts.

This study addresses these gaps by analyzing the linguistic patterns and emotional cues embedded in disinformation campaigns on Romanian online platforms. Utilizing Natural Language Processing (NLP) techniques, including Named Entity Recognition (NER) to extract and categorize cultural and historical references, the study identifies how speakers strategically use language to evoke negative emotions. By calculating offsets for temporal occurrences and examining their contextual significance within the text, the research traces how emotionally charged references—such as those to Romania's 1989 Revolution or the communist regime—are positioned to maximize psychological impact. Furthermore, the study employs models for detecting emotionally charged language, isolating patterns of fear, anger, and hate embedded in the discourse.

The findings contribute to a deeper understanding of the role of emotionally charged language in fostering ideological division and spreading disinformation in the digital space. Additionally, this research offers critical insights into the cognitive mechanisms that underlie susceptibility to manipulative narratives. By shedding light on these processes, the study provides a foundation for developing strategies to counter disinformation and promote informed, critical public discourse. Furthermore, the research emphasizes the importance of educational initiatives and awareness programs to enhance digital literacy and build resilience against emotionally charged manipulation on social media. This work aims to foster a more discerning and informed population capable of resisting disinformation by equipping individuals with the tools to evaluate emotionally manipulative content critically.

2. Literature Review

2.1 Disinformation and Emotional Manipulation

The significance of emotionally charged language in the dissemination of disinformation has been extensively studied within academic literature, underscoring its potential to shape behaviors and exacerbate ideological divides. Research conducted by Vosoughi, Roy, and Aral (2018) illustrates that false information proliferates more rapidly and extensively on social media platforms than factual content, with emotionally evocative material serving a pivotal function. The study reveals that emotions like fear, disgust, and surprise enhance the chances of content being shared, suggesting that emotional manipulation plays a vital role in the spread of disinformation. This insight directly guides the current study's emphasis on using emotionally charged language as a key strategy in disinformation efforts. In a similar vein, Wojcieszak et al. (2021) investigated the link between emotional manipulation and political disinformation on platforms such as Facebook and YouTube. Their research highlights how emotions like fear and anger drive engagement with politically motivated falsehoods, acting as mechanisms to intensify ideological divides. This complements the present study's focus on how emotional triggers, especially in historical contexts, are deliberately employed in Romanian disinformation initiatives.

In addition to promoting disinformation, emotional appeals are well-known for their powerful impact on political communication, influencing public sentiment and inspiring collective action. Nabi (2021) thoroughly examines how emotions such as fear and anger are deliberately utilized in political campaigns to sway public opinion. Her analysis highlights how political rhetoric, often referencing historical or cultural contexts, engages emotions at the cost of rational discourse. This observation is particularly pertinent to this study's investigation into how historical trauma—like the collective memory of the 1989 Romanian Revolution and the communist regime—is exploited to shape public views (Wojcieszak et al., 2021; Neumann, 2019). Similarly, Brader (2006) showed in *Campaigning for Hearts and Minds* that emotional stimuli, such as fear and hope, influence public attitudes by resonating with people's fears and aspirations. His results resonate with this study's focus on how references to Romania's cultural and historical background elicit emotional responses, aligning audiences with divided ideologies.

The strategic use of emotional appeals is also central to populist rhetoric, as Hameleers et al. (2020) observed in their examination of the dichotomy between "the people" and "the elite." They found that such rhetoric fosters in-group/out-group dynamics by employing anger and resentment to rally support, a tactic particularly effective in polarized political environments. These findings raise critical questions about the ethical boundaries of exploiting societal divisions for political gain, a theme that resonates in the Romanian context, where similar tactics are often used to reinforce collective identity or discredit opponents.

Nai and Maier (2021) explored emotional appeals in political campaigns, revealing that negative emotions like fear and anger sway voter behavior more significantly than positive emotions. They highlighted a trade-off in this approach, noting that while emotional engagement can drive votes, it often detracts from rational discussions about policies. This perspective aligns with Pagliaro et al.'s (2021) research on trust and emotional language, which indicated that politicians who effectively employ emotional appeals can appear more trustworthy, even if their messages lack strong policy content. Additionally, Soroka et al. (2022) showed that emotionally charged political rhetoric increases engagement on digital platforms, albeit frequently contributing to greater polarization. Collectively, these studies illustrate the complex nature of emotional appeals; they can effectively rally support but may also weaken democratic discourse and intensify societal divisions.

In Romania, cultural and historical references hold a significant position within the realm of emotional political rhetoric. Stoica and Florea (2022) examined how political leaders integrate historical narratives into their speeches to provoke emotional responses and substantiate their agendas. Their research reveals that these approaches resonate profoundly with audiences, drawing upon shared cultural memories to strengthen collective identity. This observation is consistent with Berinsky's (2017) analysis regarding the lasting influence of historical allusions on the

formation of collective political memory, underscoring the critical importance of contextual factors in comprehending emotional dynamics appeals.

2.2 Summary and Relevance to the Study

The reviewed literature highlights emotionally charged language's significant role in disinformation and political communication. Existing research emphasizes using emotional manipulation to increase engagement and shape public opinion, particularly by evoking fear, anger, or hope. However, while these studies provide valuable frameworks, they often lack a culturally specific perspective, especially in contexts where historical trauma and collective memory are employed to amplify emotional responses. This study builds on these foundational works by investigating the use of emotionally charged language in Romanian disinformation campaigns, focusing on how historical and cultural references provoke negative emotions to manipulate public perception and cultivate ideological influence alignment.

3. Methodology

This study uses a mixed-methods approach that combines Natural Language Processing (NLP) techniques with rhetorical analysis to examine how Romanian speakers on high-engagement media platforms utilize emotionally charged language and historical or cultural references to influence public perception.

3.1 Media Platforms and Content Selection

In this study, the selection of video content from Romanian video-sharing platforms—including TikTok, YouTube Shorts, Facebook Reels, and video posts on public pages—was guided by several key criteria to ensure the analysis's relevance, reliability, and depth. The rationale behind these criteria is outlined below:

Engagement Metrics as Indicators of Public Impact - Videos were selected based on their ranking in the top 10 for likes, shares, comments, and views. Engagement on social media serves as a crucial indicator of a video's impact, reflecting how the content resonates with audiences. High engagement levels suggest that a video has captured public interest, sparked discussions, or influenced opinions. By prioritizing videos with strong engagement, the study ensures that the analysed content represents messages and narratives that actively shape online discourse rather than being isolated or overlooked occurrences.

Content Duration for Reliable Pattern Detection - Only videos lasting between 10 and 30 minutes were included in the analysis. This duration was chosen for two reasons:

- Ensuring sufficient textual data for Natural Language Processing (NLP) techniques: Applying NLP techniques requires a substantial volume of text to detect linguistic and narrative patterns accurately. A minimum of 1000 words ensures the dataset is large enough for meaningful pattern recognition, reducing the risk of false positives or misleading results based on limited content.
- Avoiding extreme brevity or excessive length: Very short videos (e.g., under 10 minutes) may lack the depth necessary for comprehensive analysis, while excessively long videos (over 30 minutes) can introduce unnecessary complexity without corresponding analytical benefits. The chosen range strikes a balance between depth and manageability.

Historical and Cultural References as Emotional Triggers - The study also prioritized content featuring references to Romanian historical and cultural events, particularly those capable of eliciting strong emotional responses such as anger, fear, nostalgia, or national pride. These emotions play a significant role in shaping public perception and engagement with media content.

3.2 NLP Techniques and Rhetorical Analysis

The study utilized Named Entity Recognition (NER), temporal offset measurements, and emotion detection models to analyze the language and references embedded within the selected videos.

- Named Entity Recognition (NER): The multilingual “spaCy” model “ro_core_news_lg” was used to detect and categorize direct references to dates (e.g., 1989, 1991) and entities such as names, institutions, or events

- Indirect References: Phrases like “dictatorship,” “oppression,” or “totalitarian regime” were extracted because they evoke the communist era, particularly Nicolae Ceausescu’s rule, without explicitly referencing specific dates.

3.3 Interplay of Historical References and Accusations

This component explores how speakers strategically connect historical references to accusations against individuals, institutions, or policies:

- Offset Measurement: Temporal offsets were calculated to assess where and how often direct (e.g., specific years) and indirect references appear in relation to accusations.
- Contextual Mapping: Extracted references were analyzed to identify rhetorical targets such as people (e.g., political figures), institutions (e.g., government, political parties), and policies (e.g., restrictive measures, reforms).

This methodology integrates video content analysis, NLP techniques (NER and emotion detection), and rhetorical analysis to uncover how historical and cultural references are weaponized to elicit strong emotional responses. By measuring temporal offsets and correlating historical references with emotionally charged terms, the study reveals strategic patterns of emotional manipulation used to shape public discourse and deepen ideological divisions.

4. Emotional Manipulation Through Historical References

Using historical references to evoke strong emotional responses is a recurring strategy in Romanian political and media discourse. This section highlights two prominent examples of how such references are employed to manipulate public perception, drawing connections to existing research on post-communist narratives.

4.1 The 1989 Romanian Revolution

Recent speeches and media posts frequently invoke a pivotal moment in the nation's history to create or reinforce societal divisions. Speakers strategically weaponize the memory of the Revolution by casting certain political groups as "enemies of freedom" due to their perceived historical or ideological associations with communism. This rhetoric exploits collective memories of oppression and sacrifice, triggering emotions such as anger, distrust, and nostalgia for democratic ideals.

Frunză (2019) analyzes Romanian post-communist narratives to demonstrate how the Revolution's legacy has been co-opted in political discourse to frame opponents as threats to freedom or democratic progress. By aligning contemporary political actors with the negative imagery of the former regime, speakers manipulate public sentiment and deepen ideological divides. This emotional framing capitalizes on the trauma and symbolism of 1989, creating a strong emotional anchor that influences audience perceptions and political choices.

4.2 Fear of Communism's Return

Another common rhetorical strategy involves exploiting post-communist anxieties to generate fear and distrust. Speakers frame contemporary political parties or movements as potential threats to democracy, suggesting a symbolic "return to communism." This fear-driven narrative resonates deeply with populations that experienced the hardships of the communist regime, evoking emotions of insecurity, anger, and vigilance.

Stan and Turcescu (2017) explore how anti-communist rhetoric remains a powerful tool in Romania's political landscape. Their study highlights how politicians and media figures leverage collective anxieties about authoritarianism and economic instability to mobilize support and discredit opponents. By emphasizing phrases like "protecting democracy" or "fighting against dictatorship," these narratives exploit historical fears to frame contemporary events or groups as existential threats. This form of emotional manipulation positions speakers as defenders of freedom while portraying opponents as dangerous regressors.

4.3 COVID-19 Pandemic and Restrictive Measures

The COVID-19 pandemic created fertile ground for emotionally charged narratives, particularly those invoking fears of government overreach and the loss of individual freedoms. Public figures and conspiracy theorists weaponized comparisons between pandemic-related restrictions (e.g., lockdowns, mandatory vaccination policies) and historical authoritarian measures under communism, provoking deep societal distrust toward governments and institutions.

Anti-vaccination campaigns, for instance, often framed the promotion of vaccines as a "global conspiracy" or a "threat to individual autonomy," tapping into public anxieties about control and misinformation. This rhetoric frequently connected the perceived authoritarianism of pandemic restrictions to historical trauma from communist rule, provoking anger, fear, and resistance. Such narratives spread rapidly on Romanian social media platforms like Facebook and TikTok, increasing distrust in health authorities and government institutions.

Research by Baines and Elliott (2020) highlights how misinformation during the pandemic exploited emotional triggers—particularly fear and distrust—to enhance engagement and propagate conspiracy theories. Similarly, Vosoughi et al. (2018) demonstrated that emotionally charged false information spreads more rapidly than factual content, which further accounts for the swift dissemination of pandemic-related disinformation narratives.

4.4 Patterns of Emotional Manipulation in Romanian Online Communities

All three examples illustrate a recurring pattern of emotional manipulation:

- **Historical Anchoring:** Utilizing shared historical traumas (e.g., the communist regime, the Revolution) as emotional touchpoints
- **Framing Opponents:** Casting political adversaries as threats to freedom, democracy, or national identity.
- **Emotional Triggers:** Deploying emotionally charged language (e.g., "enemies of freedom," "return to dictatorship") to evoke fear, anger, and distrust.

These rhetorical strategies align with studies such as Wojcieszak et al. (2021) and Vosoughi et al. (2018), emphasizing that emotionally charged narratives are more likely to engage audiences and spread rapidly on digital platforms. By linking historical references to current political events, speakers bypass rational deliberation and provoke instinctive emotional responses, resulting in polarized public opinion.

5. Cognitive Mechanisms: System 1 vs. System 2

The strategic use of emotionally charged language in media and political discourse exploits the cognitive mechanisms outlined in Kahneman's (2011) *Thinking, Fast and Slow*. His dual-system model serves as the foundation for understanding how individuals process information and respond to stimuli, especially in the context of disinformation and emotional manipulation rhetoric.

5.1 System 1: Fast, Emotional, and Instinctive Thinking

System 1 operates quickly, automatically, and with minimal effort. It is driven by emotions, instincts, and heuristics, enabling rapid responses to external stimuli. When speakers use emotionally charged language—such as references to fear, anger, or nostalgia—they trigger System 1 thinking, bypassing rational deliberation. For instance, invoking historical traumas like the 1989 Romanian Revolution or the fear of communism's return provokes emotional reactions, such as anger or distrust, that resonate deeply with audiences.

In their study on political communication within Cognitive Psychology, Todorov and Uleman (2022) demonstrate how political figures exploit System 1 by employing emotionally charged and repetitive language. This repetition reinforces emotional triggers and ensures rapid, automatic processing, prompting audiences to react instinctively rather than analytically. Similarly, Haidt (2012), in *The Righteous Mind*, argues that moral and emotional appeals dominate rational arguments, especially in politically charged or divisive contexts.

5.2 Transition from System 1 to System 2

While System 1 thinking dominates initial emotional reactions, skilled communicators often guide audiences toward System 2 thinking—slower, more deliberate, and analytical reasoning—to solidify ideological beliefs or disinformation narratives. This transition is facilitated through rhetorical strategies, such as:

- **Framing:** Contextualizing emotional triggers within broader ideological or political arguments. For instance, after invoking fears of communism or authoritarianism, speakers position themselves as defenders of democracy freedom.
- **Repetition and Reinforcement:** Repeated references to historical or cultural events, along with emotionally charged phrases, anchor these messages in the audience's memory and direct focus toward a deliberative 'problem-solution' approach framing (Windahl et al., 2009)."

- Appeals to Identity and Values: Linking emotional rhetoric to shared values or group identity encourages audiences to rationalize their instinctive emotional responses as aligned with their beliefs or moral frameworks (Tajfel & Turner, 1986).

For example, during the COVID-19 pandemic, fears of government overreach and comparisons to communist-era restrictions initially triggered System 1 responses of distrust and anger. Subsequently, these emotional reactions were reinforced through logical arguments (System 2), such as claims about individual autonomy, freedom of choice, and resistance to global conspiracies. This strategic blend of emotional and rational appeals strengthened the audience's ideological alignment with anti-government or anti-vaccination narratives (Kahneman, 2011).

5.3 Exploiting Cognitive Mechanisms to Promote Ideologies

The interaction between System 1 and System 2 thinking is often utilized to advance specific ideologies or disinformation narratives. By initially provoking emotional responses such as fear, anger, or hatred, speakers effectively 'prime' their audiences for ideological alignment. System 1's quick and emotional processing makes individuals more susceptible to messages that resonate with their existing beliefs or anxieties, while System 2 provides a post-hoc rationalization to reinforce those emotional reactions (Kahneman, 2011).

Kahneman (2011) explains that System 2, although capable of critical reasoning, often validates System 1. This phenomenon is especially pronounced in politically polarized environments where emotional manipulation aligns with group identity and values. Haidt (2012) supports this perspective, emphasizing that moral and ideological reasoning is frequently driven by emotions, with rational arguments acting as tools to justify instinctive emotional responses judgments.

In the context of Romanian disinformation campaigns, emotionally charged references to historical events or societal crises (e.g., the Revolution, communism, and COVID-19) are skillfully exploited to create ideological divides. Audiences are first drawn in through emotional triggers and then guided toward polarized stances, solidifying their positions through instinctive and deliberative reasoning.

6. Findings and Analysis

This chapter presents key patterns observed in emotionally charged language, the framing of divisive ideologies, and the cognitive shifts identified through Natural Language Processing (NLP) analysis. The analysis was conducted using LLaMA 2 (70B), a state-of-the-art large language model designed to process long-form text and detect rhetorical patterns within Romanian digital discourse.

Unlike traditional NLP models that rely on predefined classification labels and explicit performance metrics (e.g., F1-score, accuracy, precision, and recall), LLaMA 2 operates as a context-aware generative model, allowing for more nuanced detection of implicit accusations, emotional saturation, and historical references. Its efficiency was validated through:

- Consistency across multiple speeches - Similar rhetorical patterns were observed across the 10 reviewed videos, confirming the stability of the detected trends.
- Manual validation - To ensure accuracy, key emotional triggers, accusations, and historical references were cross-checked against manually annotated transcripts.
- Comparative analysis - Results were aligned with traditional keyword-based text analysis to verify LLaMA 2's ability to detect explicit and implicit rhetorical strategies.

Although this paper illustrates and graphs only one representative video speech, all ten reviewed speeches yielded similar findings. The results highlight how historical references, emotional triggers, and accusations are systematically employed to manipulate public perception, particularly in Romanian political speeches and media appearances.

6.1 Key Patterns in Emotional Triggers and Accusations

The NLP analysis revealed the following insights:

- Word Count: Over 10,500 words were extracted from a single video, demonstrating the density and richness of the rhetoric analyzed.
- Frequency of Emotional Triggers: Emotional language—especially terms invoking fear, betrayal, and anger—occurred at consistent intervals, maintaining audience engagement and emotional saturation.

- Temporal Anchors: Historical references occurred an average of once every two minutes, sustaining high cognitive and emotional tension throughout the speech.

Accusations: A total of 80 accusations were identified, directed at a wide range of entities, including:

- Political figures and parties include the PSD, PNL, USR, and UDMR, as well as government leaders such as Ciolacu and Ciucă
- Judicial and Security Institutions: CCR (Romanian Constitutional Court), SRI, Mossad, SIE, and judicial courts.
- Global Entities and Nations include the USA, Russia, France, Israel, Benjamin Netanyahu, NATO, and EU allies.
- Abstract and Systemic Targets: “Globalists”, “Marile puteri” (great powers), “Sistemul” (the system), and “Teatru” (theater, used metaphorically).

These accusations consistently portrayed the speaker’s adversaries—be they institutions, political parties, or international actors—as enemies of Romania, democracy, or national interests’ values.

Time References: A total of 55 references to time periods were detected, strategically interwoven with accusations to amplify emotional impact:

- Communism Era (1947–1989): The most frequent indirect references invoked trauma, oppression, and dictatorship associated with Nicolae Ceausescu’s regime.
- 1989 Romanian Revolution: Used as a symbolic anchor for freedom and sacrifice and as a point of societal betrayal or division.

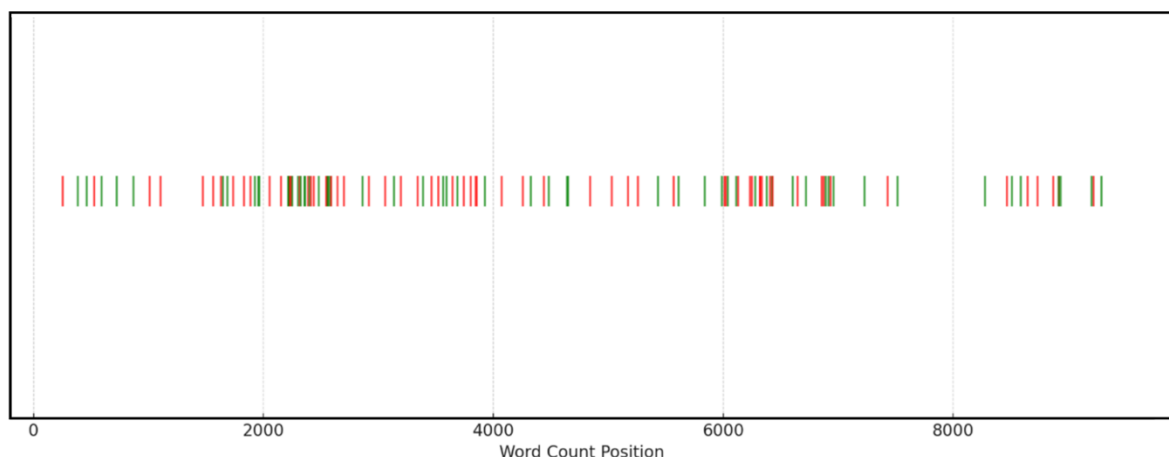


Figure 1: Occurrences of TIME (green) and ACCUSATIONS (red) in text

Table 1: Sample of analysed data

Word Offset	Type	Word	Details
253	Accusation	PSD and PNL	Political parties
386	Indirect time	dictatorship	communism 1947 - 1989
466	Indirect time	dictatorship	communism 1947 - 1989
527	Accusation	Constitutional Court of Romania	institution
592	Time	in 2019	Elections 2019
723	Time	5 years ago,	COVID 2020
868	Time	2019	COVID 2020
1010	Accusation	Constitutional Court of Romania	institution
1106	Accusation	Constitutional Court of Romania	institution
1475	Accusation	Constitutional Court of Romania	institution
1563	Accusation	Constitutional Court of Romania	institution
1631	Accusation	Constitutional Court of Romania	institution

Word Offset	Type	Word	Details
1650	Indirect time	dictatorship	communism 1947 - 1989
1685	Indirect time	dictatorship	communism 1947 - 1989
1740	Accusation	Constitutional Court of Romania	institution
1831	Accusation	Israel	country
1889	Accusation	Constitutional Court of Romania	institution
1928	Indirect time	dictatorship	communism 1947 - 1989
1957	Indirect time	dictatorship	communism 1947 - 1989
1964	Indirect time	dictatorship	communism 1947 - 1989

6.2 Emotional Manipulation Through Historical References

The analysis revealed a deliberate interplay between historical references and emotionally charged language:

6.2.1 *Traumatic Past as a Weapon*

The speaker frequently referenced traumatic events spanning over 80 years, including the communist period, the 1989 Revolution, and the post-revolution violence of 1991. These references were often paired with emotionally charged terms such as “betrayal,” “oppression,” “dictatorship,” and “failure,” provoking feelings of anger, fear, and distrust audience.

6.2.2 *Constant Temporal Shifts*

The speech demonstrated over 30 shifts between historical and contemporary events within an hour. The audience was repeatedly “thrown” from the present (e.g., 2024) to key moments such as 1989, then back to 1950 or 2020—averaging about once every 2 minutes. These abrupt temporal jumps amplified cognitive dissonance, creating a sense of instability and emotional exhaustion while reinforcing a narrative.

6.2.3 *Accusation Alignment*

Historical references were often used to frame current opponents (politicians, institutions, and global entities) as continuations of past oppressors. For example, COVID-19 restrictions were compared to communist authoritarianism, while judicial institutions were accused of perpetuating systemic betrayal.

6.2.4 *Target Audience*

The rhetoric primarily targeted individuals aged 40 and older, who personally experienced the late communist regime, the 1989 Revolution, and the turbulent post-communist transition. The emotional references resonated with their lived experiences, creating a sense of familiarity and shared context trauma.

7. Conclusions

This study investigates how emotionally charged language, historical context, and accusatory statements influence public perception in Romanian digital discourse. An extensive analysis of various political speeches and media appearances reveals that emotional manipulation is a frequent strategy and a key tool for crafting polarizing narratives and deepening ideological divides. The findings illustrate how speakers intentionally incorporate references to historical traumas and national challenges, like the 1989 Revolution and communist dictatorship, to portray present political adversaries as existential threats to national identity, sovereignty, and democratic values. The results of this study align with broader global trends identified by Bradshaw and Howard (2019), confirming that emotionally charged disinformation is a persistent feature of contemporary political communication. In Romania, such rhetorical strategies work by mobilizing collective trauma, exploiting societal fears, and reinforcing a divisive “us-versus-them” mentality. By intertwining present-day political conflicts with deep-seated historical anxieties, public speakers manipulate audience emotions to discredit adversaries, inspire distrust, and consolidate power. This deliberate evocation of past oppression and betrayal serves to amplify fear, resentment, and hostility, ensuring that audiences remain emotionally invested and susceptible to partisan narratives.

These findings highlight the need for enhanced public resilience against manipulative rhetoric through education and critical media literacy. According to Pennycook and Rand (2020), fostering deliberative thinking and encouraging audiences to engage with political content critically can reduce the impact of emotionally charged disinformation. An effective strategy is to incorporate digital and media literacy programs directly into popular social media platforms in Romania, such as TikTok, YouTube, and Facebook. We can cultivate a more informed and resilient populace by teaching audiences to recognize manipulative language, identify emotionally charged distortions, and critically evaluate political narratives.

Additionally, this research underscores the value of Natural Language Processing (NLP) techniques in the early identification of emotionally charged disinformation. Publicly accessible NLP tools can systematically scrutinize public speeches and media appearances for signs of emotional manipulation, temporal distortions, and accusatory framing. By incorporating these tools into content moderation strategies, it would be possible to identify emerging disinformation trends in real-time, allowing fact-checkers, media regulators, and policymakers to step in before these narratives spread widely. Incorporating NLP-based monitoring systems could act as an early warning system, highlighting content to foster distrust, fuelling political polarization, or eroding public trust in democratic institutions.

Despite its contributions, this study acknowledges several limitations in the NLP approaches used. One significant challenge is that LLaMA 2 (70B), as a generative model, does not operate within the same evaluation framework as traditional NLP classifiers. Unlike models that produce quantifiable accuracy, precision, recall, or F1-scores, LLaMA 2 functions primarily as a context-aware text generation system, so its effectiveness must be assessed through alternative validation methods. To address this limitation, the study relied on manual verification and comparative analysis, ensuring that the model's outputs aligned with human-labelled patterns of emotional manipulation and rhetorical distortion. Additionally, while LLaMA 2 demonstrated strong capabilities in recognizing explicit accusations and emotionally charged language, it faced challenges in detecting more subtle forms of rhetorical manipulation, such as sarcasm, indirect insinuations, or coded ideological references. Further refinements, such as fine-tuning the model on Romanian-specific datasets and expanding its training corpus with political speech transcripts, could significantly improve its accuracy in detecting linguistic nuances, subtextual framing, and evolving narrative structures.

Another limitation of the study is its focus on isolated speech events rather than long-term narrative evolution. While the analysis successfully identified recurring rhetorical strategies within individual political speeches, it did not explore how these patterns develop and shift over time. Future research could address this gap by conducting longitudinal studies that track the evolution of emotionally manipulative discourse, examining how specific political actors, media outlets, or social movements adapt their rhetorical techniques in response to public reactions, political crises, or shifting ideological landscapes. Expanding the dataset to include a wider range of speeches, social media posts, and digital interactions could provide a more comprehensive view of how disinformation narratives propagate across different platforms and audience demographics.

Moreover, a promising opportunity for future studies involves merging NLP with real-time misinformation detection systems. Developing automated workflows that incorporate deep learning models and discourse analysis techniques could significantly enhance the ability to identify and mitigate manipulative speech patterns on a large scale. This strategy would not only improve content moderation policies on social media but also support fact-checking efforts, journalistic inquiries, and regulatory oversight. Furthermore, a comparative approach across different nations could provide valuable insights into how emotional manipulation varies in diverse political and cultural contexts, shedding light on common global disinformation tactics and unique adaptations specific to each country.

Ultimately, this study underscores the pressing need for enhanced media literacy, critical thinking skills, and data-driven content moderation mechanisms to combat the influence of emotionally manipulative political rhetoric. In an era where digital platforms serve as primary channels for political engagement, the ability to critically evaluate emotionally charged content, recognize historical distortions, and resist divisive narratives has become an essential component of democratic resilience. By equipping the public with the tools to identify and challenge manipulative discourse, societies can foster a more informed, engaged, and critically aware citizenry capable of navigating complex digital landscapes without falling prey to polarization and disinformation.

References

- Baines, D., & Elliott, R. (2020). Defining misinformation, disinformation, and malinformation: An urgent need for clarity during the COVID-19 infodemic. *Media International Australia*, 177(1), 47-51.
<https://doi.org/10.1177/1329878X20961301>
- Berinsky, A. J. (2017). *The power of historical narratives in political communication*. New York: Princeton University Press.

- Brader, T. (2006). *Campaigning for hearts and minds: How emotional appeals in political ads work*. University of Chicago Press.
- Bradshaw, S., & Howard, P. N. (2019). The global disinformation order: 2019 global inventory of organized social media manipulation. Working Paper 2019.3. Oxford Internet Institute, University of Oxford.
<https://comprop.oii.ox.ac.uk/research/the-global-disinformation-order-2019/>
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon Books.
- Hameleers, M., Bos, L., & de Vreese, C. H. (2020). The appeal of populist communication: Explaining the effects of populist messages on political engagement and polarization. *European Journal of Communication*, 35(5), 463–481.
<https://doi.org/10.1177/0267323120909368>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Nai, A., & Maier, J. (2021). Do negative emotions mobilize voters? Fear and anger in political campaigns. *Political Psychology*, 42(3), 411–430. DOI: 10.1111/pops.12731
- Nai, A., & Maier, J. (2021). Emotional politics: How negative emotions drive voter behavior. *Electoral Studies*, 72, 102328.
<https://doi.org/10.1016/j.electstud.2021.102328>
- Neumann, I. B. (2019). *Cultural trauma and collective identity: Comparative perspectives on political communication in Europe*. Routledge.
- Neumann, I. B. (2019). *Historical trauma and memory politics: Global perspectives*. Cambridge: Cambridge University Press.
- Pagliaro, S., Sacchi, S., Brambilla, M., & Giacomantonio, M. (2021). The impact of emotional language on trust in political leaders. *Personality and Social Psychology Bulletin*, 47(12), 1714–1728. DOI: 10.1177/01461672211019394
- Pagliaro, S., Sacchi, S., Pacilli, M. G., & Brambilla, M. (2021). Trust and emotion in political rhetoric: How emotional language boosts perceived trustworthiness. *Personality and Social Psychology Bulletin*, 47(3), 385–398.
<https://doi.org/10.1177/0146167220969431>
- Pennycook, G., & Rand, D. G. (2020). Fighting misinformation on social media using cognitive science. *Nature Human Behaviour*, 4(5), 409–418. <https://doi.org/10.1038/s41562-020-0839-4>
- Soroka, S., Fournier, P., & Nir, L. (2022). Emotion and engagement: How emotional appeals affect polarization. *American Political Science Review*, 116(3), 721–738. DOI: 10.1017/S0003055422000224
- Stan, L., & Turcescu, L. (2017). *Church, state, and democracy in expanding Europe*. Oxford University Press.
- Stoica, E., & Florea, M. (2022). Historical narratives in Romanian political rhetoric: Legitimizing agendas through cultural and historical resonance. *Romanian Journal of Political Science*, 22(3), 45–68
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.
- Todorov, A., & Uleman, J. S. (2022). *Social perception and political communication: How automatic processes shape opinion*. Cambridge University Press.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The Spread of true and false news online. *Science*, 359 (6380), 1146-1151.
- Windahl, S., Signitzer, B., & Olson, J. T. (2009). *Using communication theory: An introduction to planned communication*. SAGE Publications.
- Wojcieszak, M., de Zavala, A. G., & Dvir-Gvirsman, S. (2021). The effects of commensurate and cross-cutting exposure on polarization and intergroup animosity in the context of COVID-19. *Group Processes & Intergroup Relations*, 24(5), 719-740. <https://doi.org/10.1177/136843022>