JOURNAL OF E-LEARNING AND KNOWLEDGE SOCIETY
Vol. 22, No. 1 (2026), pp. 37-48

Politically motivated hate speech in the Albanian language social media

Rafaela Marteta[1]

Academy of Sciences of Albania, Institute of Linguistics and Literature, Dept. of Lexicology, Terminology and Applied Linguistics – Tiranë (Albania)

(submitted: 4/3/2026; accepted: 23/5/2026; published: 4/6/2026)

Abstract

The research addresses politically motivated hate speech in Albanian-language triggered by recent political events. An evident reactive hate dominates the comment section of political news outlet in social media, so that one is moved to detect the underlying partisan motivations of this hostility. Some comments provide answers and others suggest reasons that reflect a spreading animosity towards politics and its traditional or emerging actors. Robert Entman's frame paradigm offers a theoretical foundation to detect hate speech embedded in mediated interaction in digital platforms. This research combines framing with interpretative insights derived from critical discourse analysis to capture deeper causal logics embedded in frames as a communicating text. Findings reveal that politically motivated hate speech is often normalized and legitimized in digital interaction when it is frames around some thematic issues such as national identity, moral values, economic development. Hate speech targets include political actors and institutions, collective outer groups in some cases. Recurring narrative patterns highlight responsibility attribution and negative moral evaluation, which reinforce shared perceptions about political legitimacy, and underlying legacy. The research suggests that a context-sensitive approach when addressing hate speech detection helps to understand mechanisms of its construction in order to design preventive initiatives.

KEYWORDS: Albanian Language, Social Media, Hate Speech, Frame, Politics.

DOI
https://doi.org/10.20368/1971-8829/1136315

CITE AS
Marteta, R. (2026). Politically motivated hate speech in the Albanian language social media. *Journal of e-Learning and Knowledge Society*, *22*(1), 37-48. https://doi.org/10.20368/1971-8829/1136315

1. Introduction

Many studies have addressed the problem of growing hostility in online environment. It reflects the need to examine forms and dimensions of broader rise in hateful discourse across the public sphere. Recognizable systematic features have emerged. This indicates that many forms of hostility can no longer be regarded as accidental excesses of freedom of expression. Studies confirm that speech manifesting hostility is part of a system of meanings that construct social practices. The digital environment offers significant benefits, but also exposes users to negative, harmful and dangerous dynamics. Under these circumstances individuals may become either victims or perpetrators.

Digital communication environments are frequently permeated and polluted by hostility and prejudice. Stefano Pasta (2018) argues that this prejudiced thinking online relies on the banalization of hate pedagogies and the abdication of responsibility. This perspective is highlighted by Pier Cesare Rivoltella while introducing Pasta's study (2018): the sense of responsibility enables individuals to become wiser co-constructors of the digital space, conceived as a reality that fosters freedom and collective intelligence. Many social practices previously carried out offline have gradually "migrated", adopting to the characteristics of digital environments. It is now evident that the extent of human activity within digital spaces requires responsible forms of engagement. Such an approach is essential to ensuring the quality of the "onlife" (Floridi, 2015) experience.

Pier Cesare Rivoltella (2018) emphasizes also the importance of competences required for the exercise of

[1] corresponding author – rafaela.marteta@akad.gov.al – address: Sheshi "Nënë Tereza", nr. 3, Tiranë, Albania.

© Italian e-Learning Association

full and responsible digital citizenship. The author highlights the political dimension of a commitment to safeguarding and preserving the common good and a sustainable future. This scenario deteriorates further when users fail to assume responsibility, thereby undermining the fundamental principles of digital coexistence. Scholars have proposed literacy and identification strategies that extend beyond the mere recognition of hostile expressions in addressing the phenomenon of hate speech. Conceptualizing hate speech as part of "alphabets of hate" enables a deeper understanding of it as a structured repertoire of symbols, metaphors, and linguistic patterns.

These elements function collectively to convey, normalize and reproduce hate. When manifestations of hate assume multiple forms, they constitute alphabets capable of constructing and reinforcing collective narratives. Within these narratives, hate becomes recognizable through the mechanisms of animosity directed against specific individuals or social groups. Given the plurality of forms, performative practices, and actors involved, hate speech represents a contested and evolving issue. For the purpose of this research, one of the definitions adopted is that provided by the Council of Europe.

> *Hate speech is a complex and multidimensional phenomenon that has far-reaching consequences in democratic societies. It not only affects the dignity and human rights of the individuals directly targeted, but also of persons belonging to the same minority or group as those directly targeted. […] as all kinds of expression, which spread, incite to, promote or justify violence, hatred, discrimination or prejudice against a person, or a group of persons, that is based on presumed or real personal characteristics or status including ['race'/race], colour, language, religion, citizenship, national or ethnic origin, age, disability, sex, gender, gender identity and sexual orientation. (Council of Europe, 2022, p. 17).*

This definition concerns speech targeting vulnerable groups on the basis of protected identity characteristics. While this research focuses on hostility directed toward state actors and institutions that possess public authority and political agency. They are generally subject to broader limits of acceptable criticism. This form of expression may be more accurately understood as political hostility rather than hate speech strictly understood. Nevertheless, such discourse may still produce harmful democratic effects when it normalizes dehumanization, legitimizes violence, or intensifies political polarization. In this regard a conceptual clarification provide by the United Nations is useful. It defines hate speech as: "any kind of communication in speech, writing or behavior that attacks or uses pejorative or discriminatory language with references to a person or group on the basis of who they are" (United Nations, 2022).

Owing to the complexity of phenomenon and the lack of a universally accepted definition, existing research has adopted a range of methodological approaches. These include fully automated, manual, and hybrid frameworks. Findings have highlighted the need for an interdisciplinary approach to this topic, given that relevant contributions emerge from multiple academic disciplines. Alexandra Siegel (2020) argues that research on the effects of methodological approaches used to study online hate is relatively recent and remain limited in scope. She further notes that the boundaries of the field are often unclear, given that "definitions of hate speech can be extremely broad or fairly narrow" (Siegel, 2020).

Despite the difficulty of reaching consensus on a shared definition, scholars have continued to develop approaches that address various dimensions of explicitly harmful or offensive language, as well as the pedagogies of hate (Pasta, 2018). Pasta highlights the challenge and conceptual complexity involved in engaging with this field, thereby evidencing methodological and interpretative difficulties inherent in studying the phenomenon. Definitions, causes, and targets have all been central to research on the detection of online hate speech. Whether based on fully automated or hybrid approaches, this field is characterized by a persistent tension (Pasta, 2023) between purely technological studies and those that are integrate human and computational perspectives. Pasta (2023) highlights this ongoing methodological divide between technology driven approaches and those grounded in socio-humanistic and technology-informed frameworks.

Alexandra Siegel argues that fully automated approaches can be particularly effective in identifying subtle forms of hate speech; however, she also stresses the importance of theory-driven methodologies. According to her, the latter are crucial to ensuring that detection frameworks remain conceptually grounded and analytically robust (Siegel, 2020). There is no doubt about the advantages of automation enabled by technological progress. However, Siegel highlights the limitations inherent in these technologically driven methods. In particular, she notes that they have not yet been able to systematically capture forms of hate speech in a transversal manner (Siegel, 2020).

An additional limitation to be noted is that the majority of existing studies are based primarily on English-language content. Attention should also be drawn to contents expressed in languages that pose greater challenges for the automatic detection of hate speech. The development of approaches capable of addressing detection challenges in low-resource languages represents both a significant methodological challenge and a sustained area of scholarly interest for scholars

© Italian e-Learning Association

(Siegel, 2020). An effective methodological approach is crucial for examining new dimensions of hate speech dissemination within Web 2.0 environments, particularly with regard to its structural and formal characteristics. Above all, it is essential for analyzing its meanings and underlying motivations. (Forzinetti et al., 2024)

2. Structure of the manuscript

Following the general introduction to the research, this paper is structured as follows:

- A short review of studies on hate speech in the Albanian language
- Methodological design
- Results
- Discussion and Conclusions.

It begins with a review of existing studies on hate speech in the Albanian language. These contributions situate the research within the relevant scholarly context and identify key gaps in the scientific literature. This section is followed by an accurate presentation of the methodological design adopted for analysis. The subsequent section reports the main findings of the research. The final part offers a critical discussion of the results and draws the main conclusions, highlighting implications for future work.

3. Hate speech studies in the Albanian language

Studies on hate speech in Albanian language remain limited in scale compared to research conducted in other languages. The majority of these studies adopt fully automated approaches, primarily focusing on the analysis of comments sections across various online platforms. Hersi Kopani and Rovena Llapushi (2025) demonstrate that fine-tuning the GPT-3.4 API can yield significant improvements in the detection of hate speech on Albanian social media platforms. Fetahi et al. (2025) examine the methodological challenges involved in detecting hate speech in a low-resource language such as Albanian. The Albanian language is characterized by extensive dialectal variation and scarcity of recourses and tools for advancing natural language processing (NPL). Bogdani et al. (2021) have conducted a comprehensive study on the phenomenon of hate speech, with funding provided by the Council of Europe.

The findings of Muhamet Kaçuri's master thesis focus on the detection hate speech int the context of the Albanian language rule-based system. Kaçuri (2023) argues that, while deep learning models demonstrate the highest performance according to recent publications, they remain limited by their lack of explainability. In his research he introduces *Shaj* (English: curse), an annotated Albanian dataset for hate and offensive speech derived from user-generated content on social media platforms. Nevertheless, methodologies are often sporadic due to technical and technological constraints. The Albanian language continues to be an under-resourced area for morphological analysis in natural language processing (Çepani & Çerpja, 2026). Other contributing factors may be strategic in nature, including limited institutional attention to raising informational awareness, and addressing educational dimensions of the issue.

The present paper focuses on online hate speech driven by political events. Politics has been identified as a particularly potent trigger for hate speech (Petersen, Osmundsen & Arceneaux, 2023). This pattern is equally observable in the Albanian context. Online political discourse has emerged as a space for the production, circulation and even contestation of political content. Digital platforms are not merely communication channels. They constitute environments in which social and political realities are constructed through linguistic and symbolic practices.

The outcomes of the 2016 judicial reform continue to resonate within social media discourse, shaping public deliberation, and in many instances exacerbating the proliferation of hate speech. Following a period of intensified investigations, the SPAK (Special Structure Anti-Corruption) reported the arrest of dozens of high-ranking Albanian politicians (SPAK, 2024). At the societal level, the reform challenged prevailing perceptions of systemic impunity, an issue deeply politicized and socially salient. While this paper does not examine the activities of SPAK, the political context surrounding its operations provides essential background for the analysis of politically motivated hate speech in Albanian-language social media.

Developments at the institutional level are rapidly translated into the discursive arena of social media. In this context, processes of political accountability often trigger hate-driven commentary targeting politicians. There appears to be a widespread perception of an intensification in hate speech directed at politicians. The Opposition reinforces this perception through the organization of weekly protests. During these events, aggressive rhetoric frequently emerges. It is often characterized by denigration, humiliation and the criminalization of the government. Studies suggest that the circulation of hostile political rumors may benefit those who propagate them. In particular, it may enable interested parts to gain status through this process. (Petersen, Osmundsen & Arceneaux, 2023).

For the purpose of this research, politically motivated hate speech refers to any form of expression of hostile or denigrating discourse targeting individuals or group int their capacity as political actors or supporters of a political orientation. Its primary aim is the delegitimizing of these actors within the online public sphere.

© Italian e-Learning Association

To provide a comprehensive framework of the political context that triggers broader hostility, the selected news topics focus on three key political actors: the head of government, the leader of the opposition, and a deputy minister. The latter is currently at the center of legal and institutional disputes having been accused of corruption-related misconduct while remaining in office. Given the political salience of these cases and the competing interpretations they generate, Robert Entman's framing offers an appropriate analytical lens for examining how media selectively emphasizes certain aspects of political reality, thereby structuring public discourse (Entman, 1993). At the same time, framing shapes the interpretative context of public debate.

Scholars have argued that audiences are not necessarily passively influenced by such frames (Entman, 1989; Graber, 1988). Rather as Pasta (2018) suggests, mediated communication extends the mere transmission of messages. It also serves to construct a sense of belonging among communities organized around shared beliefs. In this sense, politically driven frames non only structure interpretation but also foster collective identification. They reinforce discursive environments in which hostility may circulate and intensify. This paper specifically examines hate speech in comments responding to political news reels, treating these as communicative texts by a distinct semiotics of hate speech. The analysis focuses on the ways in which hate speech is produced. Specifically, hate frames were identified, and underlying cognitive schemata (Van Dijk, 1988: Fairclough, 2001) were examined through the discursive strategies they manifest (Van Dijk, 2011; Wodak, 2015).

4. Methodological design

A qualitative methodology was adopted to analyze users' comments on Facebook Reels related to political news. The comments were manually collected from Reels published on the official pages of Albanian-language news outlet between November 2025 and February 2026. The corpus was collected within this clearly delimited time period and corresponds to a phase of particular heightened political and public debate, driven by ongoing political crisis. Six official Facebook pages were selected on the basis of their relevance to national politics, visibility and levels of users' engagement. News items were selected according to their political relevance, the presence of a clearly identifiable target, and the salience of hate speech in the associated comment section. These criteria of selection and salience were employed to support frame identification (Entman, 1993).

Only publicly accessible comments were included, and user identities were anonymized to ensure confidentiality. Reels were selected over standard Facebook posts due to their higher levels of users' engagement and algorithmic visibility. From each reel, the 100 most relevant comments, including associated threads where present, were extracted, constituting a single dataset. Each dataset includes metadata specifying the identity of the official page identity, the date of publication, the news topic, the political orientation of the news outlet, number of comments per reel, and selection criteria applied. The overall corpus (approximately 600 comments) was determined according to the principle of theoretical saturation; reached when no new relevant information emerged from the data.

Comments were treated as autonomous discursive units capable of conveying coherent meaning. Each unit underwent manual annotation. Frames were assigned to comments that met Robert Entman's operational definition and minimal inclusion criteria, whereby at least two of the following conditions had to be present:

- problem definition: the target presented as threat or social problem etc.
- causal attribution: the target blamed for negative outcomes etc.
- moral evaluation: explicit moral condemnation, etc.
- solution recommended: calls for punishment, exclusion, removal etc.

Ambiguous cases were excluded from the analysis, including comments with no identifiable target, isolated insults, or those meeting only one for Entman's criteria. Finally, frames were interpreted using a theory-driven codebook developed for research purposes. Discursive strategies as outlined by Ruth Wodak (2015) and Teun A. van Dijk (2011), were integrated with socio-cognitive accounts of dehumanization proposed by Nick Haslam (2006) and Albert Bandura (1999). The codebook includes operational definition for hate speech, framing, discourse analysis, and discursive strategies. This instrument was constitutive for the dataset, structuring the coding of online comments according to: (1) textual content, (2) frame criteria, and (3) discursive features. The latter was assigned on the basis of the discursive strategies proposed by Wodak (2015) and Van Dijk (2011).

This procedure produced a structured and analyzable dataset represented in a grid format. While the codebook helped to limit subjective interpretation in this single-coder research, it also provided a consistent analytical framework for data classification. Table 1 below illustrates the structure of the dataset analysis. Some comments were translated into English for illustrative purposes only; otherwise, all comments in the dataset are in Albanian. The original comments have not been subjected to any linguistic correction and are reproduced as they appear in the comments section. This choice includes dialectal variations, phonetic features, systematic orthographic errors, and instances of language hybridization.

© Italian e-Learning Association

Hate speech classification followed a non-binary coding approach. Only comments previously identified as hate speech have been analyzed. Representative anonymized posts form the dataset are reproduced in Table 1, together with their assigned categories, illustrating how the coding scheme was applied in practice. The codebook is published as an addendum to the main paper. The codebook is divided into two sections. The first section presents the table of the operational definitions employed for the following concepts: hate speech (Council of Europe, 2021; United Nations, 2022); frame (Entman, 1993); Critical Discourse & Discourse Historical Approach (Van Dijk, 2001; Wodak, 2011, 2013, 2015; Reisigl & Wodak, 2001, 2009, 2015), as well as dehumanization (Bandura, 1999; Haslam, 2006). The second section contains a table outlining the strategies used in the dataset coding process. The first column contains the strategies, while the second specifies the associated discursive features, including linguistic indicators, psychologic & socio-cognitive accounts related to each strategy.

5. Results

Hate frames were conceptualized as discursive configuration that select specific aspects of reality and render them salient in order to construct the target as a political and social problem (Entman, 1993; Van Dijk, 2011). Certain frames were characterized by discursive features that reinforced polarization between in-group and out-group (Van Dijk, 2011). By identifying systematic patterns across discursive, socio-cognitive and social levels (Van Dijk, 2011; Haslam, 2006; Bandura, 1999), the analysis moves beyond a purely quantitative understanding of hate speech detection. For this reason, the research does not rely on statistical representation of data.

Although the corpus was coded manually a structured coding grid (Table 1) was employed to ensure transparency and analytical rigor. The decision not to translate the comments was motivated by the need to preserve their linguistic authenticity, as the original language contains structures that are essential for a deeper level of analysis. Translation could obscure linguistic patterns that express implicit forms of hate speech even without explicit hostile lexicon. It may also flatten certain cultural dimensions, as well as underlying mental models and attitudes embedded in the original discourse. Out of a total of approximately 600 comments around 440 met the criteria for classification as hate frames. The emerging targets of hate speech were categorized as follows:

1. individual politician;
2. political group (e.g., supporters of X Party);

Table 1 - Example of coded dataset based on the codebook.

ID - Text	Target	Frame criteria	Discursive features (linguistic indicators, psychologic and socio-cognitive accounts)
qetsohu se na e shpifen PD dhe PS un si dua as njeren un dua shqiptaret me zemer e puntore. Jo goj hale si puna jote	political group (supporters of Party X)	Problem Moral judgement	Perspectivation/Ingroup & outgroup membership: position of the commenter in terms of distance toward the target “I don’t like someone filthy mouthy like you” Moral disqualification (insults), Content of attributed difference between perceiver and target Polarization between in-group and out-group; Different deegrees of power, solidarity, intimacy: “I love Albanians who have a heart, and work hard, but not you filthy mouth”; Dehumanization (relational cognition): categorical distinction between them and us
Si mundet kjo hajdute me cizme te flase ne parlament pa u skuqur e as zverdhur per vjedhjet qe ka bere jo vetem per tunelin e Llogarase por lista eshte aq e gjate sa do ti rrethonte kete dhe ALIBABEN DISA HEREKETA RILINDASIT NUK KANE TURP ASHTU SI PARAARDHESIT E TYRE MASAKRUAN KETE POPULL Moj Laje hajdute mos dil me ne faqe te dheut se na e shpiiiiif.	individual politician	Problem Cause Moral judgement	Denomination/nomination: constructing the target as a thief “this thief, cat with the boots”; membership categorization devices, deictics: “this thief”; Moral disqualification of the target (metaphors fort theft, insults); Showing contempt to the target Participant description as different: Identifying the target with communist legacy (the target is heir to one of most important exponents of communist establishment) gender-based labeling, Predication (actor qualification) negative predicate and propositions, implications, presuppositions Intensification/Focus: capitalized words

© Italian e-Learning Association

3. Collective (e.g., “all Albanian people”);
4. Institutions (e.g., media, government);
5. multiple targets, when more than one of the identified targets co-occurred.

With a dominant presence, the primary target of politically motivated hate speech is the individual politician. This might be related to the way news is framed in social media Reels, which tend to highlight certain aspects while downplaying or omitting others. In such cases, the identification of the target is partly conditioned by the structural affordances of the medium.

Linguistic indicators reveal recurring patterns across the three columns of the analytical grid.

1) Particularly, at the discursive level (textual features), metaphors of dehumanization are frequently employed. All categories are illustrated below through specific example:

Metaphors of dehumanization: (Albanian) Ooooooo pyke,na ke ca trapin,kancer,ec vdis,se na je bo mikrob vdekje prures lugat dreqi gerr/ verr,ke 40- vite rrace e keqe, (English) ooooh wooow you scum, cancer, die, you’ve become a deadly microb devil’s servant damn you, you’ve been a bad breed for 40 years.

Intensification of the illocutionary mode (intensifiers, emphasize, quotation marks as in: (original) HA MUT SALI ME GJITH ATA QE MERR PAS TI SE ERE MUTI TE VJEN NGA SHKON E Nga vjen more Qen. (English) YOU SHITTY SALI NAD THOSE WHO FOLLOW YOU SHITTY STINKY WHEREVER YOU COME AND GO, you Dog.

Negative focus on the target, imperatives, call for action as in: (original) Ik or komunist mutit se ty vetem ish komunistat e spijunat te vijn per mbrapa se njerzit e ndershem nuk te degjoj as ste pierdhin ku rrin ti duhet te shkosh ne burg per krimet qe ke ber, (English) Go away you communist shit because only former communist spies come after you honest people don’t listen to you or even care where you are you should go to prison for the crimes you have committed.

“Us” vs. “Them” polarization (the target is constructed as different) as in: (original) Hahahaha ky vazhdon akoma aty sap dot nje dritare ne shtepin tende private pa vajtur ne 100 arkitetet e ketij per te vajtur ne 100 hajdutrit e ketij me zyrat me leket e popullit flas per ne popullin se ata vet ngren 100 kate pa e car trapin fare po populli pse te iken ty o godos, (English) Hahaha he still there while you her you can’t even open a window in your private house without going through 100 architects yet he goes through 100 thieves with flats funded with peoples’ money I speak for the people because they themselves built 100 floors without giving it a shit while your people is leaving the country you idiot. These examples construct a narrative that suggest a moral rupture through an “as versus them” opposition. Such discourse which reflects a polarized division between in-groups and out-groups, constitutes a central mechanism of conflictual political communication (van Dijk, 2006; Wodak et al., 2009). This discoursive configuration distinguishes from mere offensive language, because it operates through systematic processes of stigmatization na ddeligitimation (Faloppa, 2011).

Sarcastic or derogatory nicknaming, alteration of proper names to ridicule the target as in: (0riginal) Si ke ftyrēdh gēnjen. O KAMELEON I NDYR.Tregoi popullit sa MILJARDA I KE VJEDH. Ti dhe sekserēt e tu. Ju KRIMINELA. Qē nuk i jepni buk. Popullit pēr tē jetuar. Si ke ftyrē e del. Na tregon prallat e EDVIN NASTRADINIT. Ti duhet tē varesh nē LITAR O KRIMINEL.O buk shkelēs i POPULLIT, (English) What a lying face you have. YOU FILTHY CHAMELEON. Tell people how many BILLIONS YOU HAVE STOLEN. You and your intermediaries. YOU CRIMINALS. That don’t give bread to the people for living. How can you show your face. You tell us the tales of EDVIN NOSTRADAMUS. You should be HANGED YOU CRIMINAL. You OPPRESSOR OF THE PEOPLE.

2) Recurring patterns at a socio-cognitive level concerned frame assignment through mental models and social attitudes. These patterns involve the definition of the problem, whereby the target is constructed as a threat, or as a broader social and political issue, as illustrated in the examples below:

(original) Po mos be shume moral o te korruptuar qe i moret meqaf ate popull, (English) Don’t act all moral you corrupted ones who ruined those people’s lives. Moral evaluation/judgement implicating, presuming, alleging as in: (original) Je njeriu me i korruptuar ne bote nuk ka te dyte mbas teje, (English) You are the most corrupted person in the world, there is no one second to you.

Causal attribution of blame/responsibility for societal issues as in: (original) Ti sje ne rregull nga trute e kokes;ke te gjithe njerzit qe ke qeveris deri sot ne burg;dhe ti je i paster ipa korruptuar;nuk i ha njeri mo kto perralla ;pij longun 83 mandateve qe ke mar se i di gjith bota si i more ;do fusesh Shqiprin ne europ me pensjoniste 100mileksha ;kurr mos i hegshi ato leke Zoti ju ndershkofte, (English) You are out of your mind; all the people you have governed until today are in prison and you are “clean” and not corrupted;nobody believes these stories anymore ; now you can drink out of those 83 mandates you’ve had because everybody knows how you got them ; you want to push Albania into Europe with 100-lek pensioners; may God punish you for touching pensioners’ money. Stereotyping, invoking generalized negative traits about a target group as in: (original) Te tere hajduta pisa gjakpirsa asnje nga ju plera qe ulni bythen te ato karrige .vetem per te vjedhur, (English) All of you are thieves and bloodsuckers none of you scum who sit on those chairs are able to do else but steal.

© Italian e-Learning Association

3) Recurring patterns at the social level relate to broader societal and political implications, particularly when the focus is placed on the target, often constructed as individual politician: (original) *Korupsjoni ini ju bashk me stafin tuaj dhe ju ini anti shqiptar,* (English) *You are the corruption you and your staff, you anti-Albanians.* Exclusion, marginalization of the target: call for silencing the, disappear, delegitimation as in: (original) *Ku shkuan dy urat e projektuara ? Ne drejtim te pa ditur ...ne tynelin e llogarase? je e vetmja qe nuk mund te besh moral shqiptareve..,* (English) *Where did the money for two planned bridges go? Towards the unknown… into the llogara tunnel? You are the only one who cannot lecture Albanians.*

Individual targeting based on arguments, implications, presumptions for societal harm, crisis or threat to the community as in: (original) *Ky duhet te arestohet dhe denohet per tradheti kombetare. Pa u lodhur shume per te gjetur edhe provat se nuk ja vlen te shpenzohet me shume per tradhetaret e atedheut. Ne emer te popullit si dikur ; vetem keshtu e pastron ate vend nga shushunjat, (English) He should be arrested and punished for national treason. No need to spent much time gathering evidence as it is not worth for traitors of the homeland. In the name of the people as in the past; only in this way the country will be cleansed of the parasites*.

Collective versus individual targeting refers to whether hostility is directed toward a group or toward a single figure, as illustrated in the following example: (original) *Po normalisht do sulmojne sos kan pune te tjera per te bere. Rri dhe merr mend ktu ne komente ku pjesa me e madhe hane buken kot. Ku ne kohen e PD kta skishin goje te ndiheshin qe po u shtypnin kokat e spo e merrnin vesh e po jetonin ne mizerje pa drita pa uje,po ne kto kohe ku JU keni sjell zhvillim e keni rritur rrogat e keni shtruar rruget e keni kthyer Tiranen ne Kryeqytetin me te bukur te Europes dhe Veriun me nje zhvillim te papare e Jugu gjithashtu, tani u vjen goja ktyre se Ai i pari sikur u kish q* plakat kush tia mbante i pari kandilin e vët ua mbante nga koka*
(English) *Well, of course they will attack, they have nothing else to do. They sit here in the comments while most of them are living off nothing. Back in the PD era, they didn't even have the words to realize they were being crushed, living in misery without electricity or water, but now in these times when YOU have brought development, increased salaries, paved roads, turned Tirana into the most beautiful capital in Europe, developed the North like never before, and the South as well, now they complain – because, first of all, it's as if he f***ed the old ladies, and who held the candle for him It stayed on their heads.*

6. Discussions and conclusions

Rather than generalizing, this research conceptualizes framing as a mechanism that defines and delimits hate speech within contemporary digital environments. While automated approaches offer notable advantages, they may come at the cost of overlooking key discursive features in favor of computational efficiency. The automatic detection of online hate speech presents several methodological and practical challenges that must be addressed. Some of these challenges have been outlined in reports by the European Online Hate Lab (2023), in particular the limited capacity of automated systems to capture meaning during data processing (European Commission, 2023).

For this reason, there is a need for studies adopting context-sensitive and interpretative approaches that take into account the socio-cultural variables contributing to the production of hate speech. In the reports by the European Online Hate Lab various terms are used to categorize hateful language. The report identifies a scale of severity and the intensity of hateful language. It distinguishes six levels of escalation, ranging from disagreement as the least intense to death as the most intense. This typology is particularly useful for this research, as the findings indicate a relevant presence of high-intensity forms of hateful language, including expressions that move beyond criticism or hostility and approach calls for harm and elimination. High-intensity hateful speech tends to generate morally charged frames in which emotive, polarized commentary reinforces symbolic "us versus them" boundaries rather than persuasive argumentation (Matthes & Kohring, 2008).

ROPH (Research on online political hostility, 2025) is a research project aimed at examining and addressing extreme forms of online political hostility. The research has identified offline sources of frustration that are transferred to online contexts, mediated by psychological, sociological, and political predispositions (Petersen, Osmundsen & Arceneaux, 2023). "A need for chaos" (Petersen, Osmundsen and Arceneaux, 2023) has also emerged from our dataset. Commenters issue call for uprising, violence, boycott, capital punishment, divine punishment against individual politicians or political groups. A "burn-it-all-to-the ground" social attitude identified by Petersen, Osmundsen and Arceneaux (2023) is also present int this research. It appears as an explicit pattern associated with a disintegrative social orientation. Hannah Arendt (1970) associated processes of social and political disintegration to the growing demands of the masses and the weakening of political parties. In contemporary settings, this dynamic unfolds within what Rivoltella describes "as another place, a parallel space" (Rivoltella, 2017). Within this "other space" experiences are redefined, while they are "empowered by the suppression of the body" (Rivoltella, 2017). They may also generate frustration due to limited opportunities for concrete action. In line with Arendt's (1970) argument such frustration may become a source of violence.

© Italian e-Learning Association

Indeed, the dataset reveals evidence of recurring narratives that construct present-day politics as fundamentally immutable in its nature. It is presented as intrinsically corrupted, hypocritical, and ultimately a threat to national interest. This essentialization, frequently reinforced through copular constructions with the verb "to be", has a legitimizing function in the case of hate speech. It generalizes blame and responsibility, and indiscriminately delegitimize political processes and political actors. In this respect, the Albanian language displays certain linguistic features that may facilitate processes of essentialization and collectivization. The frequent use of the verb "to be" for ontological attribution and the use of absolute quantifiers – all, always, never – supports the generalization of guilt. Within this framework, essentialization is intertwined with recurrent socio-cultural and anthropological patterns. Political dissent is reframed in logics of collective survival and generalized distrust.

According to Albert Nikolla (2011), the lack of trust among Albanians is a social attitude toward elites who continue to hold power despite having lost citizen's trust. The findings of this research suggest that hostility is directed primarily toward public authorities which are portrayed as incompetent, corrupt and oppressive. This form of discourse overlaps with processes of institutional delegitimization, carrying significant implications for democratic legitimacy and the quality of democratic discourse. Manuel Castells (2017) argues that distrust towards particular candidates may extend beyond individual to institutions and, ultimately, to the political system as a whole. In this sense hate against individual politician functions as a delegitimizing discourse that contributes to the erosion of institutional trust. By portraying them as enemies of the people who operate within corrupt structures, such narratives can foster antagonistic political identities and normalize hostility toward institutional authority.

Therefore, it is important to distinguish hostility toward institutions from legitimate political criticism. Criticism challenges policies, decisions within the framework of democratic debate. Hate is expressed through discursive strategies that delegitimize individual politicians and the institutions they represent as inherently corrupt or criminal. Such discourse often relies on accusation, inflammatory language and calls for punishment or rejection of institutional authority.

Many comments attribute responsibility for political decline to the target, frequently invoking death, violent agency, or divine punishment. This discursive pattern activates what Bandura (1999) defines moral disengagement; through which harmful judgements are morally justified and normalized. Both physical distance and the inability to directly experience others' suffering further contributes to moral disengagement (Bandura, 1999). As argued by Pasta (2018), new media have reshaped the ways in which emotions are experienced and expressed. Value-based labeling, vertical comparisons, and the deployment of the disgust and contempt functions to delegitimize the target's right to speak, in extreme cases, their very right to exist. Explicit calls for death or capital punishment exemplify this attitude.

Notably, some social discourses legitimize the use of communist-era punitive practices as a justified response to the target's presumed corruption and criminality. Through presuppositions, predications, implications, and constructions of negative agency the targets or their actions are criminalized (Van Dijk, 2011). Many frames operate through dehumanization by animalizing and objectifying the target on behalf of non-human characteristics (Haslam, 2006). A recurrent labeling of targets as "communists" evokes a totalitarian legacy, constructing them as dangerous, destructive and threatening. Violent action may be perceived as necessary when a group is framed as threatening (Pollack et al., 2023). Hate speech grounded in a totalitarian legacy is highly prevalent in the dataset, especially when the target is a deputy minister and the descendant to one of the most prominent representatives of the communist establishment. Hate frames imply the incrimination of her political legacy by associating her current political activity with that of her grandfather. This association serves to delegitimize both figures as responsible for crimes and theft, while constructing them as enemies of the country. Studies suggest that enemies are constructed on the basis of perceived inherent features rather than temporary one (Van Dijk, 2011).

In conclusion, it can be argued that social media provide a space for the propagation and normalization of hate speech. Hateful messages, often encountering little resistance and limited moderation, contribute to the normalization of hostility towards targeted individuals. This occurs when targets are constructed as e problem, a cause of social harm, or as morally despicable actors. Understanding the deeper motivations behind hatred expressed online may help individuals and institutions to identify it as a driver of conflictual digital citizenship, and social conflict. Groups and institutions addressing hate speech may respond through awareness-raising and prevention programs, or by developing counter-narratives.

A key finding of this research is the transversality of hate speech: hostile and exclusionary expressions are not confined to specific actors or settings. This suggests that hate speech is not the product of isolated individuals, but part broader discursive patterns circulating in everyday online communication. Its transversal spread blurs the boundary between explicit hostility and normalized stigmatizing language. This highlights the need to address this behavior and the wider communicative environments that sustain it. This concern is echoed in the Joint Communication of the European Commission to the European Parliament and

© Italian e-Learning Association

the Council of Europe. It emphasizes that language which belittles or dehumanizes others contributes to a sense of impunity against attacks on specific communities or individuals. While freedom of expression is a fundamental right and a democratic value, human rights law recognizes that it cannot be used to incite hatred and violence. And there should be no impunity for those who cross the line. (European Commission, 2023).

In this context, media-education based pedagogical activism may offers a response. Practically, pedagogical activism can be implemented through structured classroom or community-based practices that directly engage with real digital discourse. A counter-discourse production is a possible method. Learners collaboratively create alternative narratives – videos, posts, memes – that reframe polarized or hostile discussion into empathetic discussion. Gagliardone et al. (2015) argue that, despite the variety of approaches and intended audiences, initiatives aimed at addressing hate speech should converge around three educational objectives: "to inform, to analyze and to confront hate speech" (Gagliardone et al., 2015)

The research acknowledges several limitations, including the absence of inter-coder reliability due to a single author coding. Although this limitation is partially mitigated by theoretical grounding and a documented codebook, internal interpretative processes may still introduce subjectivity. Moreover, the methodological approach itself may constitute a limitation due to difficulties in applying it to a large dataset. Subjectivity in framing comments may also function as a frame in its own right, shaping how issues are interpreted (Matthes & Kohring, 2008). Finally, while the reel format enables the observation of spontaneous public reactions, it also reflects the influence of certain groups in shaping structural aspects of news production.

Datasets and reproducibility

Datasets will be published as an addendum to the main paper.

References

Arendt, H. (1990). On violence. Harcourt, Brace & World.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. Personality and social Psychology Review, 3(3), 193-209. https://doi.org/10.1207/s15327957pspr0303_3

Bayer, J., & Bard, P. (2020). Hate speech and hate crime in the EU and the evaluation of online content regulation approaches (Study requested by the LIBE Committee, Directorate-General for Internal Policies, European Parliament). Publications Office of the European Union. http://doi.org/10.2861/28047

Bogdani, M., Faloppa, F., & Karaj, S. (2021). Mapping hate speech in Albania. Council of Europe.

Castells, M. (2009). Communication power. Oxford University Press.

Council of Europe. (2022). Recommendation CM/Rec (2022)16 of the Committee of Ministers to member States on combating hate speech. Council of Europe.

Çepani, A., Çerpja, A. (2026). Digital morphology of the Albanian language. Journal of Responsible Technology, 25, 100151. https://doi.org/10.1016/j.jrt.2026.100151

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. Journal of Communication, 43(4), 51-58. http://doi.org/10.1111/j.1460-2466.1993.tb01304.x

European Commission & High Representative of the Union for Foreign Affairs and Security Policy. (2023). Joint communication to the European Parliament and the Council: No place for hate: A Europe united against hatred https://commission.europa.eu/news-and-media/news/no-place-hate-europe-united-against-hatred-2023-12-06_en

European Commission. (2023). The European Online Hate Lab. Publications Office of the European Union. http://doi.org/10.2838/59442

Fetahi, E., Hamiti, M., Susuri, A., Ajdari, J., & Zenuni, X. (2024). AI-based hate speech detection in Albanian social media: New dataset and mobile web application integration. International Journal of Interactive Mobile Technologies (iJIM), 18(24), 190-208. https://doi.org/10.3991/ijim.v18i24.50851

Faloppa F. (2011). Razzisti a parole (per tacer dei fatti). Laterza.

Floridi, L. (Ed.) (2015). The Onlife Manifesto: Being human in a hyperconnected era. Springer. https://doi.org/10.1007/978-3-319-04093-6

Fetahi, M., Susuri E., Hamiti, M., Kastrati, Z., Canhasi, E., & Misin9, A. (2025). Enhancing social media hate speech detection in low-resource languages using transformers and explainable AI. Social Network Analysis and Mining, 15, Article 82. https://doi.org/10.1007/s13278-025-01497

Forzinetti, E., Della Vedova, M. L., Pasta, S., & Santerini, M. (2024). Indicators for characterising online hate speech and its automatic detection. arXiv. https://arxiv.org/abs/2402.08462

© Italian e-Learning Association

Haslam, N. (2006). Dehumanization: An Integrative review. Personality and Social Psychology Review, 10(3), 252-264. https://doi.org/10.1207/s15327957pspr1003_4

Kaçuri, M (2023). Explainability of hate speech classification for Albanian language using rule-based system and neural networks (Diploma thesis). Technische Universität Wien. https://doi.org/10.34726/hss.2023.105780

Kopani, H., & Llapushi, R. (2025). Fine-tuning GPT-3.5 for hate speech detection in Albanian YouTube comments: Challenges and analysis. In K., Dhoska & E. Spaho (Eds.) Bridging Horizons in Artificial Intelligence, Robotics, Cybersecurity, Smart cities and Digital Economy: Sustainable Economy and Ecotechnology (pp. 243-261). Springer. https://doi.org/10.1007/978-3-031-72029-1 20.

Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. Journal of Communication, 58(2), 258-279. https://doi.org/10.1111/j.1460-2466.2008.00384.x

Nikolla, A. (2011). L'uomo nuovo Albanese. Tra morale comunista e crisi della transizione. Bonanno Editore.

Pasta, S. (2018). Razzismi 2.0: Analisi socio-educativa dell'odio online. Scholé.

Petersen, M. B., Osmundsen, M., Arceneaux, K., et al. (2023). The need for chaos and the sharing of hostile political rumors in advanced democracies. American Political Science Review, 117(1), 1-18. https://doi.org/10.1017/S0003055422000759

Petersen, M. B., Osmundsen, M., Bechmann, A., Lindekilde, L., Sherson, J., Bor, A., et al. (2020-2025). Research on online political hostility (ROPH) Aarhus University, Department of Political Science. https://ps.au.dk/forskningsprojekter/research-on-online-politcal-hostility/about-the-roph-project

Pollack Sarnecki, H., Lioufas, A., & Jarlsbo, M. (2023). Rotten democracy – Conspiracy propaganda, racism and violence (Report No. FOI-R-5473-SE). Swedish Defence Research Agency (FOI)

Rivoltella, P. C. (2018). Prefazione in Razzismi 2.0: Analisi socio-educativa dell'odio online. Scholé.

Rivoltella, P. C. (2017). Tecnologie di comunità. ELS La Scuola.

Siegel, A. A. (2020). Online hate speech. Princeton University Press.

United Nations. (2022). United Nations strategy and plan of action on hate speech. United Nations. https://www.un.org/en/genocide-prevention/hate-speech/strategy-plan-action

Van Dijk, T. A. (2008). Toward a theory of context: The socio-cognitive approach. Cambridge University Press.

Van Dijk, T. A. (1998). Ideology: A multidisciplinary approach. Sage.

Van Dijk, T. A. (2011). Discourse studies: A multidisciplinary introduction (2nd ed.). Sage Publications

Wodak, R. (2015). The historical-discursive approach. In R. Wodak & M. Meyer (Eds.), Methods of critical discourse studies (3rd ed., pp.23-61). Sage Publications.

© Italian e-Learning Association

Appendix - Minimal Codebook Structure

Part 1. Table of operational definitions for concepts.

Concepts	Operational Definition for
Hate speech	Hate speech is a complex and multidimensional phenomenon that has far-reaching and dangerous consequences in democratic societies. It not only affects the dignity and human rights of the individual directly targeted, but also of persons belonging to the same minority or group as those directly targeted. Hate speech leads to dangerous divisions in society as a whole, affects the participation and inclusion of all those targeted by it and threatens democracy. […] Hate speech is understood as all kinds of expressions, which spread, incite to, promote or justify violence, hatred, discrimination or prejudice against a person, or a group of persons, that is based on presumed or real personal characteristics or status including ['race'/race], colour, language, religion, citizenship, national or ethnic origin, age, disability, sex, gender, gender identity and sexual orientation (Council of Europe, 2021). In the context of this document, the term hate speech is understood as any kind of communication in speech, writing or behavior that attacks or uses pejorative or discriminatory language with references to a person or group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor. This is often rooted in, and generates intolerance and hatred, in certain contexts, can be demeaning and divisive (United Nations, 2022).
Frame	Framing essentially involves selection and salience. To frame is to select some aspects of a perceived reality and make them more salient in a communicating text in such way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described (Entman, 1993).
Critical Discourse & Discourse Historical Approach.	Critical Discourse Analysis (CDA) is a research enterprise which critically analyses the relation between language and society. More specifically, CDA is a type of discourse-analytical research that studies the way ideology, identity and inequality are (re)enacted through texts produced in social and political contexts. (Van Dijk, 2001; Wodak, 2011). The discourse-historical approach (DHA) belongs in the broadly defined field of critical discourse studies (CDS), or also critical discourse analysis (CDA) (Reisigl & Wodak, 2001, 2009; Wodak, 2011, 2013). CDS in general investigates language use beyond the sentence level, as well as other forms of meaning-making such as visual sounds, seeing them as irreducible elements in the (re)production of society via semiosis. CDS aims to denaturalize the role discourse splay in the (re)production of noninclusive and nonegalitarian structures and challenges the social conditions in which they are embedded. (Wodak, 2015; Reisigl & Wodak, 2015).
Dehumanization	Dehumanization is a phenomenon in interpersonal as well as in intergroup contexts, occurs outside the domains of violence and conflict, and has socio-cognitive dimensions in addition to the motivational determinant that are usually emphasized. Two distinct forms of dehumanization should occur when the respective properties are denied to others. In most theoretical accounts, dehumanization is seen primarily as a motivated phenomenon, enabling the release of aggression or removing the burden of moral qualms or vicarious distress (Bandura, 1999; Haslam, 2006).

© Italian e-Learning Association

Part 2. Table of strategies for dataset coding.

Strategies	**Discursive Features** **(linguistic indicators, psychologic & socio-cognitive accounts)**
Nomination (Wodak)	*Actor construction* Membership categorization devices (Deictics; Verbs and nouns for actions and processes; Metaphors)
Denomination (Van Dijk)	*Participant description as different* Strangers, immigrants, others, Opponents, enemies, negative agency (Pronouns, demonstratives)
Dehumanization (Haslam)	*Essentialism* (Nature of difference between perceiver and target, content of attributed difference between perceiver and target)
Predication (Wodak)	*Actor qualification* (The labeling of social actors, positively or negatively, appreciatively or depreciatorily. Stereotypical, evaluative attributions (metaphors, allusions, evocations, adjectives, implicatures)
Predication (Van Dijk)	*Meaning of sentences* (Negative predicate and propositions, implications, presuppositions, allegations, negative emphasize, negative properties)
Dehumanization (Haslam)	*Social categorization* (Mechanistic, a view of others as fungible, and lacking individuality, depersonalized, deindividuated outgroups, moral exclusion)
Argumentation (Wodak)	*Justification/question claims of truth and norms* (topoi, fallacies)
Argumentation (Van Dijk)	*Persuasion by arguments and fallacies* (Formats/schemas/alleged negative properties, rhetorical moves, deictics, adjectives, pronouns)
Dehumanization (Haslam)	*Behavior Explanation* (a more mechanistic view of the other, emphasizing factors that are deterministic, attenuate personal agency an de-emphasize intentional states)
Perspectivation (Wodak)	*Positioning the speaker in terms of involvement or distance* (Metaphors, quotation marks, deictics, indirect or free speech, other)
In-group and out-group membership (Van Dijk)	(as in) *Us* vs. *Them* Different degrees of power, solidarity, intimacy (pronouns, rhetorical moves)
Dehumanization Haslam	*Relational cognition* Communal sharing but categorical distinction between "us" and them"/asocial, Value-based dehumanization: an outgroup is perceived to have dissimilar values to the ingroup; lack of prosocial values and incongruent with one's ingroups' values.
Intensification (Wodak)	*Intensifying the illocutionary mode (*Diminutives, hesitations, verbs of saying, feeling, thinking, hyperboles, litotes, other)
Focus (Van Dijk)	*Special focus on participant (*Stress, volume, size, colour, enumeration, alliteration, visual structures)
Dehumanization (Haslam)	*Emotion* (disgust and revulsion, object of disgust and contempt, elicitors of disgust: death, excretion, sexuality. In its animalistic form the implicit contrast is animals, and it involves denial of uniquely human characteristics. In its mechanistic form it involves emotional distancing, the other is cold, robotic, passive, lacking depth.)

© Italian e-Learning Association