



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

CATEGORIZATION AND AUTHORIZATION OF EXTREMIST TEXT IN LINGUISTIC EXPERTISE

Kurbonmuratov Abror Abdurashidovich

PhD Researcher, Tashkent State University of Uzbek Language
and Literature named after Alisher Navoi

abrorbek1941@gmail.com

Abstract:

This article examines the linguistic features of extremist texts (ET) in the Uzbek language, focusing on their categorization by type (religious, political, social), authorship identification criteria, and the methodological foundations of forensic linguistic expertise. The study analyzes lexical, morphological, and syntactic properties of extremist discourse, as well as territorial, gender, and age-related markers of authorship. The findings contribute to the development of systematic criteria for linguistic expertise of extremist content in judicial and investigative practice.

Keywords: extremist text, linguistic expertise, authorship identification, forensic linguistics, Uzbek language, categorization, hate speech, extremist discourse.

The rapid spread of extremist content across digital platforms and social media has created an urgent need for systematic linguistic analysis of such materials in both academic and judicial contexts. Forensic linguistic expertise of extremist texts (ET) has become one of the most practically significant areas in applied linguistics, particularly in countries undergoing intensive socio-political transformation such as Uzbekistan.



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

While international scholarship has extensively examined extremist discourse in English, Russian, and other major languages, systematic research into the linguistic structure of Uzbek-language extremist texts remains underdeveloped. The agglutinative nature of the Uzbek language and its unique morpho-syntactic patterns require dedicated analytical frameworks that cannot simply be transferred from other linguistic traditions.

This article presents the key findings of a doctoral dissertation on the categorization and authorship authorization of extremist texts in linguistic expertise. The research addresses three core challenges: (1) developing a typology of extremist texts applicable to Uzbek-language materials; (2) identifying characteristic linguistic features at lexical, morphological, and syntactic levels; and (3) establishing criteria for determining the territorial, gender, and age-related characteristics of text authors.

Forensic linguistic analysis of extremist materials draws on two converging scholarly traditions. In the Anglo-American tradition, foundational contributions by Coulthard and Johnson (2007) and Grant established principles for the linguistic examination of anonymous threatening communications and social media content. Research by Nini and Olsson addressed methodological questions of organizing linguistic expertise in criminological contexts.

In Russian forensic linguistics, the theoretical and methodological foundations of linguistic expertise were developed by Baranov (2007), Belchikov, and Gorbanevsky, while Ablin (2016) specifically addressed expertise in extremism-related cases. These scholars emphasized the necessity of distinguishing between the linguistic and legal components of such analysis.

In Uzbek linguistics, contributions by Iskandarova, Shahobiddinova, Saidov, and Musulmonova have addressed applied and juridical linguistics. However, a



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

systematic framework for identifying and categorizing extremist texts specifically in the Uzbek language context had not yet been established prior to the current research.

A critical conceptual distinction fundamental to this research is the differentiation between "extreme language" and "extremist text." While extreme language encompasses sharp, radical expressions that may appear in various forms — images, gestures, insults — without necessarily constituting a criminal act, extremist text is characterized by specific communicative goals: direct threats, manipulation of target subjects, incitement to inter-ethnic conflict, or calls for undermining constitutional order.

Based on content, communicative strategy, and socio-political context, extremist texts can be categorized into three major types, each with distinctive linguistic properties:

3.1 Religious Extremist Texts

Religious ET are characterized by intensive use of specialized theological vocabulary with distorted semantic content (e.g., jihod, shahid, g'azot, xalifalik, tog'ut, murtad), citation of religious sources in support of violence, and a binary "believer vs. unbeliever" discourse structure. The communicative strategy typically combines doctrinal argumentation with emotional intensification. At the syntactic level, these texts frequently employ imperative constructions, conditional clauses, and introductory phrases of religious authority (alloh amriga binoan, shariatga ko'ra).

3.2 Political Extremist Texts

Political ET are defined by opposition to democratic values and governmental legitimacy. The lexical layer is dominated by terms such as nepotizm,



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

manipulyatsiya, muxolif, moliyaviy piramida, and various pejorative references to state institutions. A distinctive pragmatic feature is the implicit communicative goal: texts may appear as ordinary political criticism while pursuing destabilizing aims. The presence of coordinated posting patterns (e.g., similar content published by multiple recently-created profiles) serves as an additional indicator of organized political ET rather than spontaneous criticism.

3.3 Social Extremist Texts

Social ET include content promoting racial, ethnic, or religious discrimination and inter-group violence. This category is lexically characterized by dehumanizing vocabulary targeting specific groups, including racist and xenophobic terminology. At the pragmatic level, these texts construct an "enemy image" using negative mythological, folkloric, or religious imagery to mobilize in-group members against out-groups.

4. Linguistic Features of Extremist Texts in Uzbek

4.1 Lexical Features

Lexical analysis reveals domain-specific vocabulary clusters for each ET type. Religious ET employ a dense network of Arabic and Persian religious borrowings. Political ET rely on Western political terminology and bureaucratic lexemes with negative connotations. Social ET make use of ethno-pejorative vocabulary and dehumanizing comparisons. Across all types, propagandistic lexemes (kurash, dushman, birlik, hamjihatlik) serve as high-frequency markers of extremist intent.



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

4.2 Morphological Features

The agglutinative nature of Uzbek creates specific morphological patterns in ET. The imperative mood (*buyruq mayli*) is predominant, reflecting the directive function of extremist communication (*Tezda xavfsiz joyga boring!*). Present tense verb forms emphasize the urgency of claimed threats (*Xavf kengaymoqda*). The plural suffix *-lar* is used extensively to construct collective in-group and out-group identities (*Fuqarolar, Mujohidlar, Murtadlar*). The conditional suffix *-sa/-se* frequently appears in threat constructions and calls to action.

4.3 Syntactic Features

At the syntactic level, extremist texts in Uzbek show a marked preference for simple sentences over complex constructions, enabling maximum communicative directness. Among complex sentence types, adversative compound sentences and conditional subordinate clauses appear with higher frequency. Seven distinct rhetorical syntactic strategies are identified in the corpus: threat rhetoric, temporal rhetoric (projecting future catastrophe), resistance rhetoric, romantic rhetoric (invoking solidarity), social justice rhetoric, mythological rhetoric, and lexical-syntactic repetition.

Parallel syntactic constructions (*parallelizm*) serve as a particularly effective intensification device. Structures such as "*Biz erkinlik uchun kurashamiz, tenglik uchun kurashamiz, adolat uchun kurashamiz*" achieve rhetorical impact through rhythmic repetition, conceptual reinforcement, and mnemonic effect — all functions well documented in persuasive political discourse.

5. Authorship Identification Criteria

Authorship identification (authorization) in forensic linguistic expertise of ET requires systematic criteria across three dimensions:



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

- Territorial markers: The presence, absence, or density of dialectal features can indicate the regional origin of the text author. For anonymous ET, dialectal analysis helps estimate the geographic distribution of the text and the specific extremist group affiliation.
- Gender markers: Female-authored ET display greater use of introductory lexical elements, tag questions, precise color terminology, intensifiers, and euphemisms. Male-authored ET more frequently employ direct imperatives, aggressive nominal constructions, and military/combat lexicon. Both may use coarse vocabulary, though the functional distribution differs.
- Age markers: Young authors tend toward shorter sentences, jargon, abbreviations, and emotionally expressive language reflecting social identity formation. Middle-aged authors show patterns of radical radicalization, instructional tone, and doctrinal argumentation. Older authors use more complex syntactic structures, traditional vocabulary, and authoritative-advisory communication strategies.

The case analysis demonstrates these distinctions through two contrasting examples: a message by a 2005-born young defendant employing persuasion-based extremist content without strong morpho-syntactic markers, and a text by a 1986-born defendant using explicit instruction, religious prohibition, and direct behavioral commands — reflecting divergent age-related communicative strategies within extremist discourse.

6. Methodological Implications for Forensic Linguistic Practice

The findings carry direct implications for the organization of forensic linguistic expertise in Uzbekistan. First, linguistic and legal components must be rigorously separated in expert conclusions: the linguist analyzes phonetic, lexical,



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

morphological, and syntactic features and identifies implicit pragmatic content, while legal qualification of danger and criminal liability assessment falls within the jurist's competence.

Second, the analysis of texts containing ideological-religious content should not rely solely on general reference dictionaries, as these typically provide approximate and superficial definitions of specialized religious terms. Such analysis requires interdisciplinary collaboration with subject-matter experts in religious studies.

Third, the research results have been applied in NLP-related projects developing morphological analyzers for the Uzbek language, demonstrating the practical value of this work for computational linguistics alongside forensic applications. This research has established a systematic typological and linguistic framework for the analysis of extremist texts in the Uzbek language. The study demonstrates that ET can be reliably classified into religious, political, and social types based on distinguishable lexical, morphological, syntactic, and pragmatic features. Authorship identification is made possible through the application of territorial, gender, and age-related linguistic criteria developed specifically for Uzbek-language materials.

The agglutinative morphology of Uzbek plays a distinctive role in the communicative effectiveness of extremist texts, particularly through the imperative mood, plural markers of collective identity, and conditional constructions. Syntactically, simple sentence structures dominate, while rhetorical devices such as repetition and parallelism function to intensify persuasive impact.

The findings contribute to Uzbek applied linguistics, forensic linguistics, and NLP development for the Uzbek language, and provide methodological



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

foundations for practitioners conducting linguistic expertise in judicial and law enforcement contexts.

References

1. Ablin M.V. (2016). Theoretical and methodological foundations of linguistic expertise in extremism cases [Теоретико-методологическое обоснование лингвистической экспертизы по делам об экстремизме]. Avtoref. diss. ... kand. filol. nauk. Ufa. 62 p.
2. Baranov A.N. (2007). Linguistic expertise of text: theory and practice [Лингвистическая экспертиза текста: теория и практика]. Moscow: Flinta: Nauka. 592 p.
3. Coulthard M. & Johnson A. (2007). An Introduction to Forensic Linguistics: Language in Evidence. New York: Routledge. 250 p.
4. Kurbonmuratov A.A. (2023). Category of 'extremist text' in linguistic expertise [Lingvistik ekspertizada «ekstremistik matn» kategoriyasi]. O‘zbekiston Milliy universiteti xabarlari. No. 1/1. P. 246–249.
5. Kurbonmuratov A.A. (2024). Language characteristics of extremist text [Ekstremistik matnning til xususiyatlari]. Xorazm Ma'mun akademiyasi axborotnomasi. No. 2/4. P. 91–94.
6. Kurbonmuratov A.A. (2024). The Concept of Extremist Text and its Types. Journal of Ethics and Diversity in International Communication. Vol. 3, Issue 3. P. 20–24.
7. Kurbonmuratov A.A. (2024). Author Marks in Extremist Text. Journal of Advanced Linguistic Studies. Vol. 11, No. 1. P. 471–479.



International Conference on Computing, Artificial Intelligence and Information Systems

Hosted Online from Warsaw, Poland

Date: 11th June, 2026

Website: <https://econferencia.com>

-
8. Musulmonova K.X. (2022). Process, stages and methods of linguistic expertise of written texts in the Uzbek language [O‘zbek tili yozma matnlarini lingvistik ekspertiza qilish jarayoni, bosqichlari va metodlari]. PhD diss. Qo‘qon. 132 p.
 9. Olsson J. (2008). Forensic Linguistics. London: Continuum. 218 p.
 10. Spohr D. (2017). Fake news and ideological polarization: filter bubbles and selective exposure on social media. Business Information Review. 34(3). P. 150–160.