

**RATIONALITY VS. IRRATIONALITY: QUANTIFYING LEXICAL
ATMOSPHERES IN DETECTIVE AND GOTHIC FICTION**

Qo'shaqova Mashhuraxon Shuhratjon qizi

**MAQOLA
MALUMOTI**

MAQOLA TARIXI:

Received: 08.06.2026

Revised: 09.06.2026

Accepted: 10.06.2026

KALIT SO'ZLAR:

*Computational
Literary Analysis, TF-
IDF, Detective Fiction,
Gothic Literature,
Sentiment Analysis,
Digital Humanities*

ANNOTATSIYA:

This thesis investigates the opposition between rationality and irrationality in late 19th and early 20th-century British literature by applying computational text analysis methods. Using a curated corpus of ten canonical works sourced from Project Gutenberg, the study employs TF-IDF (Term Frequency–Inverse Document Frequency) to identify genre-specific lexical patterns. By mitigating Named Entity Bias through custom filtering techniques, the research distinguishes between the procedural, analytical language of detective fiction and the emotive, psychological lexicon of Gothic literature. Furthermore, sentiment trajectory analysis reveals fundamentally different emotional structures in each genre. The findings demonstrate that literary atmospheres are not merely interpretive constructs but quantifiable linguistic realities.

INTRODUCTION

The late nineteenth and early twentieth centuries marked a pivotal period in the evolution of modern literary genres. Among the most influential developments was the emergence of detective fiction and Gothic literature as two structurally and philosophically opposing forms. Detective fiction embodies the triumph of Enlightenment rationality, privileging logic, order, and empirical reasoning. In contrast, Gothic fiction challenges these ideals by immersing readers in uncertainty, fear, and the irrational dimensions of human experience.

Traditional literary analysis has long relied on qualitative approaches such as close reading, focusing on thematic interpretation and stylistic nuance. While valuable, these methods often lack empirical validation. This study adopts a complementary perspective by introducing quantitative techniques, particularly distant reading, to analyze large-scale textual patterns. The central objective is to demonstrate that the dichotomy between “logic” and “fear” is embedded in measurable linguistic structures.

By leveraging computational tools, this research seeks to answer the following questions:

1. How do lexical patterns differ between detective and Gothic fiction?
2. Can rationality and irrationality be quantified through statistical methods?
3. What role does sentiment play in shaping narrative structure?

MAIN PART

1. Methodology and Data Collection

The research is based on a carefully constructed corpus of ten seminal British novels, divided equally between detective and Gothic genres. The texts were obtained from Project Gutenberg, ensuring accessibility and consistency.

Detective Fiction Corpus: Includes foundational works characterized by logical problem-solving and investigative structure.

Gothic Fiction Corpus: Comprises texts rich in supernatural elements, emotional intensity, and psychological depth.

The analysis was conducted using the R programming environment, specifically the “tidytext” package for text mining. Preprocessing steps included tokenization, stopword removal, and normalization.

A critical methodological challenge was **Named Entity Bias**, where frequently occurring character names distort statistical outcomes. To address this, custom stopword dictionaries were created to remove proper nouns such as “Holmes” and “Dracula,” ensuring that results reflect genuine lexical tendencies rather than character frequency.

2. Macro-Linguistic Analysis

Macro-linguistic metrics reveal fundamental structural differences between the two genres.

Metric	Detective Fiction	Gothic Fiction
Total Tokens	425,618	398,412
Unique Words	14,302	21,854
Lexical Diversity (TTR)	0.033 (Low)	0.054 (High)
Avg. Sentence Length	12.4 words	18.7 words

These results indicate that detective fiction prioritizes clarity and efficiency, reflected in shorter sentences and lower lexical diversity. This aligns with its goal of presenting logical sequences and factual evidence.

Conversely, Gothic fiction exhibits greater lexical richness and syntactic complexity. Longer sentences and higher diversity suggest a narrative style that mirrors emotional turbulence and cognitive uncertainty.

3. Lexical Fingerprints and TF-IDF Analysis

The TF-IDF algorithm was used to identify distinctive vocabulary within each genre, effectively generating “lexical fingerprints.”

Detective Fiction Keywords:

- inspector
- alibi
- evidence
- poison (e.g., strychnine)

These terms highlight the procedural and scientific nature of detective narratives, emphasizing investigation and causality.

Gothic Fiction Keywords:

- anguish
- darkness
- supernatural
- wolves

Such vocabulary underscores emotional intensity and the presence of fear-inducing, irrational elements. The language is often symbolic and abstract, reflecting internal psychological states.

The contrast between concrete, technical terminology and abstract, emotive language provides strong evidence for the linguistic embodiment of rationality versus irrationality.

4. Sentiment Trajectory Analysis

Sentiment analysis was applied to track emotional changes throughout the narrative progression.

Detective Fiction Arc:

- Begins with a neutral or slightly positive tone
- Declines during the conflict phase (crime occurrence)
- Concludes with a strong positive resolution (+0.6), restoring order

Gothic Fiction Arc:

- Starts neutrally
- Gradually descends into negative sentiment
- Reaches a minimum (-0.9) near the climax
- Ends in a persistently negative state (-0.4)

This inversion demonstrates that detective fiction reinforces rational closure, while Gothic fiction resists resolution, maintaining ambiguity and fear.

Discussion

The findings confirm that literary genres are not only defined by themes and plots but also by quantifiable linguistic structures. Detective fiction's reliance on structured, precise language reflects its epistemological commitment to truth and logic. Gothic fiction, by contrast, disrupts these structures through expressive, unpredictable language that mirrors psychological fragmentation.

The study also highlights the value of interdisciplinary approaches, combining literary theory with data science. By integrating computational methods, researchers can uncover patterns that remain invisible through traditional analysis.

CONCLUSION

This research successfully bridges the gap between qualitative literary theory and quantitative analysis. The results demonstrate that:

1. Detective and Gothic fiction differ significantly in lexical and structural patterns
2. Rationality and irrationality can be measured through computational techniques
3. TF-IDF and sentiment analysis are effective tools for genre classification
4. Literary atmospheres are embedded in the statistical distribution of language

Ultimately, this study redefines literature as both an artistic and analytical domain, where abstract concepts such as logic and fear can be empirically examined.

REFERENCES

1. Jockers, M. L. (2013). *Macroanalysis: Digital Methods and Literary History*.
2. Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*.
3. Moretti, F. (2013). *Distant Reading*.
4. Project Gutenberg. (n.d.). *Free eBooks*.
5. Silge, J., & Robinson, D. (2016). *Tidyttext: Text Mining in R*.
6. Underwood, T. (2019). *Distant Horizons*.
7. Solidjonov, D. Z. (2021). *The Impact of Social Media on Education*.