AI vs Human News Headlines: A Comparative Study of Syntax and Emotional Language

Garibashvili Manana
https://orcid.org/0000-0002-5118-3690
Lazviashvili Shorena
https://orcid.org/0000-0002-4429-1044
Iakob Gogebashvili State University, Telavi
https://doi.org/10.52340/idw.2025.39

Abstract. This research examines the comparative linguistic features of newspaper headlines produced by artificial intelligence systems versus human journalists, concentrating particularly on grammatical structure and affective language use. Utilizing theoretical frameworks from Systemic Functional Linguistics, News Values Theory, and Critical Discourse Analysis, the study employs comparative content analysis methodology to evaluate paired headlines sourced from traditional news organizations (BBC, NBC) and AI news platforms (NewsGPT.ai). Results indicate that AI-produced headlines tend to be more concise, declarative, and sensationalized, utilizing active constructions and event-focused verbs. Conversely, human-crafted headlines exhibit greater length, explanatory depth, and analytical perspective, incorporating narrative contextualization, emotional engagement through stylistic devices, and direct quotations. Human-authored content additionally shows enhanced numerical specificity and grammatical intricacy, whereas AI-generated headlines prioritize compression and factual documentation. The research determines that although AI demonstrates capability in generating efficient and striking headlines, human journalists remain essential for maintaining precision, contextual richness, and empathetic authenticity. These results underscore the necessity for harmonious collaboration between AI and human contribution in digital journalism, as each approach offers distinctive yet mutually reinforcing capabilities to modern news media.

Keywords: AI-generated headlines; human-authored headlines; emotional language; comparative analysis; linguistic features; systemic functional linguistics; critical discourse analysis

Introduction

The role of newspaper language is to provide readers with information about various fields. Accordingly, newspaper language strives to be both informational and persuasive. In this regard, newspaper headlines play a leading role, as they serve as the titles of articles appearing above the news texts. The principal functions of headlines (Mallette, 1990, as cited in Bedřichová, 2006, p. 14) are the following: 1) to summarize the news; 2) to indicate the importance of stories; 3) to act as clear elements in page design; and 4) to persuade the looker to become a reader. Correspondingly, a headline should express the core content of the newspaper text. Salih and Abdulla (2012) state that "It is a 'miniature replica' of the story" (p. 291). Headlines, described as "little texts," have their own grammar, as they must convey much information in very few words (Halliday, 1985, p. 372). Most importantly, headlines are often scanned by readers rather than read in full (Dor, 2003). That is why they tend to be short, powerful, and influential over readers.

"Being a linguistic, journalistic, and cultural phenomenon, the news headline has

evolved with media technology and cultural contexts … of defining the headlines, those of informing and persuading remain constant" (Gherheş, Fărcaşiu, Cernicova-Buca, & Coman, 2025, p. 150). Recently, the process of AI-generated news has become increasingly active, producing newspaper articles that convey the main content of the news but often differ in form and emotional language. This difference is particularly noticeable in the headlines. Therefore, the observation and analysis of the language used in AI-generated news compared to human-written mainstream news articles is both highly interesting and important.

Theoretical Review

This study draws upon three complementary theoretical perspectives to analyze the linguistic differences between AI-generated and human-written news headlines: Systemic Functional Linguistics (SFL), News Values Theory, and Critical Discourse Analysis (CDA). Together, these frameworks provide a comprehensive lens for understanding how headlines function as both linguistic and journalistic artifacts.

Systemic Functional Linguistics (SFL), developed by Halliday (1985), serves as the fundamental linguistic framework for this study. SFL views language as a social semiotic system, where linguistic choices are shaped by context and serve specific communicative functions. From our perspective, SFL is particularly relevant to headline analysis because it highlights how grammatical structures convey meaning beyond the mere transfer of information, shaping how readers interpret and engage with news.

The second core framework, Critical Discourse Analysis (CDA), provides the basis for examining headlines as ideological and power-laden discourse. Developed by Fairclough (1995) and van Dijk (1988), CDA investigates how language both reflects and constructs social reality, power relations, and institutional practices. This perspective allows us to uncover how headlines are not only informative but also instrumental in reinforcing or challenging dominant ideologies.

Literature Review

There is a growing body of research comparing the linguistic features of AI-generated and human-authored newspaper headlines. Gherheş et al. (2025) explore the potential applications of AI technology in online journalism, noting predictions that speed and adaptability to new media will accelerate the automation of news production. Their study emphasizes that, while human supervision of the journalistic workflow remains vital, the nature of this workflow is changing, with the writing of micro-content increasingly entrusted to ChatGPT-3.5 as one of the most visible developments.

Similarly, Muñoz-Ortiz et al. (2024) compared English news texts produced by six large language models (LLMs) with human-written articles and identified measurable differences in sentence structure, vocabulary diversity, and emotional tone. Their findings show that human-authored texts tend to display more varied sentence lengths and stronger emotional expressions, whereas AI-generated texts are generally more syntactically uniform and objective. This underscores the tendency of AI-generated content to prioritize clarity and neutrality over stylistic richness.

Zamaraeva et al. (2025) present the first comprehensive comparison of New York Times–style text generated by six LLMs against authentic NYT writing, using a formal syntactic theory as the analytical basis. Their study highlights systematic structural differences between AI- and human-authored texts.

Finally, Durak (2025) compares human-written and AI-generated texts in educational discourse, focusing on features such as sentence length, vocabulary richness, and syntactic

complexity. The findings suggest that human-written texts generally exhibit longer sentences and more complex syntactic structures than those generated by AI.

Methodology

As the linguistic analysis of AI-generated text is still in its early stages, there remains limited research on direct comparisons between human- and AI-generated texts, particularly in the domain of newspaper headline analysis. To address this gap, we set out to conduct a study that identifies comparable newspaper articles from both human mainstream and AI-generated sources, contrasts them, and develops a comparative analysis of syntax and emotional language.

For this purpose, we applied comparative content analysis as the core methodological approach, enabling us to compare paired texts (AI vs. human headlines covering the same events) across multiple linguistic dimensions. Alongside this, we also employed a qualitative research method, through which headlines from mainstream media outlets and AI-driven news platforms were systematically examined. The study investigates linguistic similarities and differences between AI-generated and human-written English news reporting with a focus on syntactic structures, emotional language, lexical choice, and information hierarchy.

Results and Discussion

As we have already mentioned, the goal of our research is to compare and analyze news articles from mainstream media outlets and AI-driven news platforms in order to investigate linguistic similarities and differences between AI-generated and human-written English news reporting. The focus is on syntactic structures, emotional language, lexical choice, and information hierarchy. For in-depth analysis, we selected several articles covering the same events and examined both the headlines and the article texts themselves. The following findings illustrate these comparisons.

When comparing the following pair of titles:

a) *Bangkok Skyscraper Collapse: Arrest Warrants Issued* (AI-generated, NewsGPT.ai)

b) *Seventeen arrest warrants issued over Bangkok skyscraper collapse* (BBC News),

both emphasize the collapse and subsequent legal action, preserving the key terms "Bangkok," "skyscraper collapse," and "arrest warrants." Both are concise (6–8 words). However, differences are notable: the AI-generated title uses passive voice and presents information in a categorical, dramatic way, while the BBC title employs active voice, a fuller sentence structure, and includes the specific number "seventeen," enhancing factuality and precision. Thus, the AI headline spotlights the disaster dramatically, while the BBC version foregrounds the legal consequences with journalistic clarity.

A second pair demonstrates similar contrasts:

a) *1000-Year-Old Skeleton Moved to Gujarat Museum* (AI-generated, NewsGPT.ai)

b) *Ancient Indian skeleton gets a museum home six years after excavation* (BBC News).

Both highlight an archaeological discovery but differ in style and detail. The AI title, using passive voice, stresses the age of the skeleton ("1000-Year-Old"), signaling historical significance. In contrast, the BBC headline employs active voice, narrative style, and a temporal marker ("six years after excavation"), which contextualizes the event and engages readers in a more process-oriented way. Again, the AI text is shorter and declarative, while the BBC headline appears more explanatory and reader-friendly.

Humanization is especially visible in the following example:

a) *Jessie J Announces Breast Cancer Diagnosis* (AI-generated, NewsGPT.ai)

b) *'I need a hug' — Singer Jessie J says she has breast cancer* (BBC News).

The BBC headline employs direct speech, personal pronouns, and emotionally charged language, immediately conveying vulnerability and human experience. The AI headline, while factual and concise, lacks emotional depth, underscoring the tendency of human-written headlines to "humanize" the story, whereas AI tends to strip it down to its informational essence.

Sentence length further illustrates differences:

a) *Crowd Crush Kills 11 in Victory Parade Tragedy* (AI-generated, NewsGPT.ai)

b) *Eleven die in India crush as fans gather for IPL victory parade* (BBC News).

The BBC headline is more neutral, using factual verbs like "die" and "gather." The AI headline, however, adopts a dramatic tone through lexical choices like "kills," "crush," and "tragedy." Both use active voice, but their subject treatment differs: BBC centers the victims ("Eleven die"), while AI foregrounds the event as the agent of action ("Crowd Crush Kills").

Another comparison makes the differences even clearer:

a) *Lisbon Glória Funicular Derailment Kills 15, Injures 18* (AI-generated, NewsGPT.ai)

b) *At least 15 people killed and 18 injured in crash of funicular railway in Lisbon* (NBC News).

Both highlight the accident and its casualties, but the NBC headline is longer (15 words), more explanatory, and follows traditional journalistic sentence structure. The AI headline is shorter (8 words), categorical, and formatted like a bulletin. Lexical choice also varies: NBC uses softer, descriptive terms ("crash of funicular railway," "people killed and injured"), whereas AI prefers technical and impactful language ("derailment," "kills," "injures"), producing a sharper, more dramatic tone.

A final pair illustrates the contrast between factual documentation and interpretive framing:

a) *Trump Signs Executive Order for Qatar Security Pact* (AI-generated, NewsGPT.ai)

b) *Trump's pact to defend Qatar could shake up or shore up a fragile region* (BBC News).

Both headlines highlight the Trump–Qatar agreement and include the keywords "Trump" and "Qatar." However, the AI title is straightforward, active-voice, and limited to 7 words, reflecting a legal/administrative register. The BBC headline is longer (14 words), complex, and speculative, using modal verbs and metaphorical language ("shake up or shore up") to suggest broader geopolitical consequences. This demonstrates that AI headlines typically document events as facts, whereas human-written headlines interpret events as stories with context, causality, and potential impact.

Discussion

Seven matched headlines from human-written (BBC, NBC) and AI-generated (NewsGPT.ai) news sources were compared, revealing systematic linguistic variations in four main areas: information hierarchy, lexical choice, syntactic structure, and emotive language. The most notable finding is that, while human-written headlines emphasize narrative contextualization and interpretive analysis, AI-generated headlines consistently prioritize categorical documentation of events.

In terms of sentence length, AI-generated headlines averaged 6–8 words, whereas human-written headlines averaged 10–15 words, nearly double the length. AI headlines tend to pack maximum factual content into minimal space using categorical, bulletin-style formats. In contrast, human-generated headlines incorporate additional words to provide contextual

anchoring and interpretive framing. Furthermore, human-written headlines consistently employ metaphorical and rhetorically sophisticated language, whereas AI headlines rely on literal, denotative language without rhetorical flourish.

It is also noteworthy that, although AI headlines occasionally generalize or omit numerical details, human-written headlines always retain specific figures (e.g., "Seventeen arrest warrants," "At least 15 people killed and 18 injured"). Given expectations that AI systems could be highly proficient in precise numerical reporting, this outcome was somewhat surprising. The human emphasis on numerical accuracy reflects the professional journalism criteria of verifiability and credibility, enabling readers to assess the scope and gravity of events. This demonstrates that human journalists successfully merge factual rigor with narrative engagement, even while incorporating contextual and interpretive elements.

Finally, the analysis reveals a contrasting pattern in emotive language. Human-written headlines typically use softer, neutral lexical choices such as "die," "gather," and "crash." AI-generated headlines, in contrast, often employ more dramatic and emotionally charged terms like "kills," "tragedy," and "crush." However, the AI's dramatic wording does not involve genuine emotional humanization. Human journalists, by contrast, use techniques such as direct speech quotations ("I need a hug"), personal pronouns (I, she), vulnerability markers, and sympathetic framing to establish a genuine emotional connection with readers, reflecting a level of emotional sophistication absent in AI-generated headlines.

Conclusion

Systematic disparities in syntax, lexical choice, emotional language, and information hierarchy have been observed in the comparison of newspaper headlines written by humans versus AI. While both types of headlines fulfill the fundamental journalistic tasks of describing events and capturing readers' attention, they do so in markedly different ways. AI-generated headlines are typically shorter, more categorical, and often employ dramatic or action-oriented verbs, yet they lack the subtle humanizing techniques that professional journalists use. In contrast, human-authored headlines are generally longer, more informative, and incorporate emotional resonance, numerical specificity, and narrative framing to better engage readers.

These findings suggest that AI-generated headlines prioritize efficiency and conciseness, reflecting the technological trend toward automated, rapid news production. However, this comes at the cost of journalism's contextual richness and empathetic dimension. Human journalists remain essential for ensuring accuracy, depth of interpretation, and emotional genuineness in reporting.

Future research should examine how these linguistic differences influence reader trust, engagement, and information processing as AI tools are increasingly integrated into newsrooms. Investigating hybrid models, where AI supports headline creation under human editorial supervision, is also crucial to maintain ethical standards and contextual awareness. Ultimately, the interplay between AI and human-authored journalism will shape not only the form and style of headlines but also the cultural role and credibility of the news media in society.

References:

Bedrichova, L. (2006). Headlines and subheadlines in newspaper reporting (MA thesis, Masaryk University). Masaryk University.

Dor, D. (2003). On newspaper headlines as relevance optimizers. Journal of Pragmatics, 35(5), 695–721 retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0378216602001340?via%3Dihub

Gherheş, V., Fărcaşiu, M. A., Cernicova‑Buca, M., & Coman, C. (2025). AI vs. human-authored headlines: Evaluating the effectiveness, trust, and linguistic features of ChatGPT-generated clickbait and informative headlines in digital news. Information, 16(2), Article 150 retrieved from https://www.mdpi.com/2078-2489/16/2/150

Halliday, M. A. K. (1985). An introduction to functional grammar (2nd ed.). London: Edward Arnold.

Muñoz-Ortiz, A., Gómez-Rodríguez, C., & Vilares, D. (2024). Contrasting linguistic patterns in human and LLM-generated news text. Artificial Intelligence Review, 57(2), 265–296 retrieved from https://link.springer.com/article/10.1007/s10462-024-10903-2

Salih, Y. M., & Abdulla, Q. A. M. (2012). Linguistic features of newspaper headlines. Journal of Al-Anbar University for Language and Literature, (7), 291–310. Retrieved from //efaidnbmnnnibpcajpcglclefindmkaj/https://iasj.rdd.edu.iq/journals/uploads/2025/01/05/f2bd558d801258af29f0264c2342421a.pdf

Zamaraeva, O., Flickinger, D., Bond, F., & Gómez-Rodríguez, C. (2025). Comparing LLM-generated and human-authored news text using formal syntactic theory. Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 9041–9060 retrieved from https://aclanthology.org/2025.acl-long.443/

ხელოვნური ინტელექტისა და ადამიანის მიერ შექმნილი საინფორმაციო სათაურები: სინტაქსისა და ემოციური ენის შედარებითი ანალიზი

ლარიბაშვილი მანანა
ლაზვიაშვილიშორენა
იაკობ გოგებაშვილის სახელობის სახელმწიფო უნივერსიტეტი, თელავი

აბსტრაქტი

კვლევა შეისწავლის ხელოვნური ინტელექტის სისტემებისა და ჟურნალისტების მიერ შექმნილი საინფორმაციო სათაურების შედარებით ლინგვისტურ მახასიათებლებს, განსაკუთრებული ყურადღება ექცევა გრამატიკულ სტრუქტურას და აფექტური ენის გამოყენებას. სისტემურ-ფუნქციური ლინგვისტიკის, საინფორმაციო ღირებულებების თეორიისა და კრიტიკული დისკურსის ანალიზის თეორიული ჩარჩოების გამოყენებით, კვლევა იყენებს შედარებითი ანალიზის მეთოდს ტრადიციული საინფორმაციო ონლაინ მედიიდან (BBC, NBC) და ხელოვნური ინტელექტის საინფორმაციო პლატფორმებიდან (NewsGPT.ai) აღებული იდენტური შინაარსის მატარებელი სათაურების შესაფასებლად.

მიღებული შედეგები მიუთითებს, რომ ხელოვნური ინტელექტის მიერ შექმნილი სათაურები უფრო ლაკონური, მტკიცებითი და სენსაციური ხასიათისაა; გამოყენებულია მოქმედებითი გავრის კონსტრუქციები და მოვლენაზე ორიენტირებულ ზმნები. ხოლო, ადამიანის მიერ შექმნილი სათაურები უფრო გრძელია სიტყვების რაოდენობის მიხედვით, განმარტებითი სიღრმითა და ანალიტიკური პერსპექტივით. მათში გაერთიანებულია ნარატიული კონტექსტუალიზაცია. აგრეთვე, ემოციური ჩართულობა სტილისტური საშუალებებითა და პირდაპირი ციტირებით.

ადამიანის ავტორობით შექმნილ სათაურებში შეიმჩნევა რიცხვითი სახელების სიხშუსტე და გრამატიკულად კომპლექსური კონსტრუქციების სიჭარბე, მაშინ როცა ხელოვნური ინტელექტის მიერ გენერირებული სათაურები პრიორიტეტს ანიჭებს შემჭიდროვებულ ინფორმაციასა და ფაქტობრივ დოკუმენტირებას.

კვლევის მხიედვით დგინდება რომ, მაშინ როდესაც ხელოვნური ინტელექტი აგენერირებს ეფექტურ და თავლშისახემ სათაურებს, ჟურნალისტების მიერ შექმნილი სათაურები გამოირჩევა სიხუსტით, რათა უზრუნველყოფილ იქნეს კონტექსტუალური სიმდიდრე და ემპათიური ავთენტურობა.   ეს შედეგები ხაზს უსვამს ჰარმონიული თანამშრომლობის აუცილებლობაზე ხელოვნურ ინტელექტსა და ცოცხალ ჟურნალისტიკას

შორის, რადგან თითოეული მიდგომა უზრუნველყოფს გამორჩეულ, მაგრამ ურთიერთგამაძლიერებელ შესაძლებლობებს თანამედროვე საინფორმაციო მედიისთვის.

საკვანძო სიტყვები: AI-ს მიერ შექმნილი სათაურები; ადამიანის მიერ შექმნილისათაურები; ემოციური ენა; შედარებითი ანალიზი; ლინგვისტური მახასიათებლები; სისტემური ფუნქციონალური ლინგვისტიკა; კრიტიკული დისკურსის ანალიზი