

An Ethical and Rhetorical Framework to Analyze LLM-generated Texts

Daniel Raffini

Sapienza, Università di Roma
daniel.raffini@uniroma1.it

Agnese Macori

Sapienza, Università di Roma
macori@diag.uniroma1.it

Marco Angelini

Link University of Rome
angelini@diag.uniroma1.it

Tiziana Catarci

Sapienza, Università di Roma ISCT-CNR
catarci@diag.uniroma1.it

Abstract

The rapid diffusion of Large Language Models (LLMs) is transforming the production and circulation of textual content, raising important linguistic, cognitive, and ethical questions. As AI-generated texts increasingly populate digital environments, understanding their rhetorical and stylistic characteristics becomes essential to assess their influence on readers and on the broader information ecosystem. This paper proposes an ethical and rhetorical framework for the analysis of LLM-generated texts based on a reader-oriented approach. The framework identifies a set of rhetorical functions (trustworthiness, apologetic, proximity, diversification, ambiguity, emphasis, explanatory, poetic, fairness, structure) and links them to specific rhetorical figures, providing a structured method to evaluate how linguistic strategies shape readers' perceptions and responses.

Methodologically, the study integrates theoretical modeling with qualitative close reading, combining deductive framework construction with inductive textual analysis. The framework is tested through the analysis of two corpora of AI-generated texts, consisting of argumentative essays and narrative short stories. The results highlight recurrent rhetorical patterns, including the overrepresentation of emphatic and structural functions in argumentative texts and the prevalence of figurative language with limited semantic depth in narrative outputs. These findings suggest that LLMs tend to reproduce recognizable rhetorical patterns while often relying on formal persuasion strategies rather than conceptual complexity. The proposed framework is designed as an open, scalable, and integrable model that can be progressively refined and expanded through further empirical applications across different textual typologies and communicative contexts.

Keywords: LLM; rhetoric; rhetorical figures; AI; ethics

La rapida diffusione dei Large Language Model (LLM) sta trasformando la produzione e la circolazione dei contenuti testuali, sollevando importanti questioni linguistiche, cognitive ed etiche. La crescente diffusione di testi generati dall'IA negli ambienti digitali rende necessario comprenderne le caratteristiche retoriche e stilistiche, per valutare l'influenza sui lettori e sull'ecosistema informativo. Questo articolo propone un framework etico e retorico per l'analisi dei testi generati da LLM, fondato su un approccio centrato sul lettore. Il framework identifica un insieme di funzioni retoriche (attendibilità, apologia, prossimità, diversificazione, ambiguità, enfasi, esplicazione, poetica, equità, struttura) e le collega a specifiche figure retoriche, fornendo un metodo strutturato per valutare come le strategie linguistiche plasmino le percezioni e le risposte dei lettori. Sul piano metodologico, lo studio integra la modellazione teorica con il close reading qualitativo, combinando la costruzione deduttiva del framework con l'analisi testuale induttiva. Il framework è testato attraverso l'analisi di due corpora di testi generati da LLM, composti rispettivamente da saggi argomentativi e racconti brevi. I risultati evidenziano schemi retorici ricorrenti, tra cui la sovrarappresentazione delle funzioni enfasi e struttura nei testi argomentativi e la prevalenza del linguaggio figurativo con limitata profondità semantica in quelli narrativi. Questi risultati suggeriscono che gli LLM tendono a riprodurre strutture retoriche ricorrenti, facendo affidamento su strategie di persuasione formale piuttosto che sulla complessità concettuale. Il framework proposto è concepito come un modello aperto, scalabile e integrabile, che può essere progressivamente affinato ed esteso attraverso ulteriori applicazioni empiriche su diverse tipologie testuali e contesti comunicativi.

Parole chiave: LLM; retorica; figure retoriche; IA; etica

1. Introduction

Beginning in 2022, several Large Language Models (LLMs) were released to the public and made available through online conversational interfaces, marking a turning point in the history of Artificial Intelligence (AI) and Human-Computer Interaction (HCI). By removing the barriers to communication between humans and machines, anyone – even with no specific technological knowledge – can engage in conversations with these systems using natural language. The spread of LLMs affects not only interaction methods and the definition of AI, but also information systems and expressive codes – i.e. the semiotic and communicative conventions through which meaning is produced and shared, including written language, narrative forms, and stylistic registers. We are witnessing a massive influx of artificially generated texts, audio, images, and videos into the information system, particularly on the web. Studies show that a significant amount of online content is already produced by AI systems, and projections indicate that this trend will grow exponentially (Muzumdar et al. 2025), (Sun et al. 2025), (Hanley and Durumeric 2024), (Liang et al. 2025). This development makes it essential to pay close attention to AI-generated content, including its linguistic and rhetorical aspects. As these texts become more

prevalent in communication, they will increasingly influence human language and the future development of AI models (Yakura et al. 2025).

The use of generative AI as a substitute for human writing leads to significant consequences from linguistic, societal and cognitive perspectives. As AI systems rely on vast datasets to generate content, they tend to reinforce the most frequent linguistic patterns while marginalizing the less common ones (Kendall and Teixeira da Silva 2024). Many studies have shown that LLMs still lack in linguistic diversity compared to humans (Shaib et al. 2024), (Martínez et al. 2024), (Giulianelli et al. 2023), and highlighted differences in style, grammar and rhetoric between human and AI-generated texts (Fröhling and Zubiaga 2021), (Muñoz-Ortiz et al. 2024), (Reinhart et al. 2025), (Carrasco-Farre 2024). Also relevant is the phenomenon of declining linguistic diversity in LLMs trained on AI-generated texts: LLM-generated texts are uploaded to the web, reducing linguistic diversity online; meanwhile, the practice of training new LLMs through web scraping leads to an exponential flattening of their performance (Guo et al. 2024). The reduction of linguistic diversity could have consequences such as impaired critical judgment, a reduced ability to evaluate situations (Alkhrisheh et al. 2021), manipulation and disinformation (Chen et al. 2025). Indeed, Yoo et al. (2025) show how the stylistic-rhetorical characterization of AI language can influence users' perception of truth and find that AI-generated language can create both trust and deception, depending on the rhetorical elements.

For these reasons, a deeper exploration of the rhetorical dimensions is necessary, along with an ethical reflection on AI's role in content generation. Studying rhetorical construction in generative models involves understanding how AI-generated texts can insert bias or persuasive techniques into language in ways not immediately apparent to users, especially in a context marked by increasing linguistic impoverishment and declining critical judgment. Therefore, we argue that it is crucial to examine LLMs' use of language and rhetoric both to enhance their linguistic competence and stylistic diversity, and to raise human awareness of the limitations inherent in AI-generated language. To do so, it is necessary to establish evaluation framework that consider qualitative aspects of textuality and human perception. In section 2 we address the problem of integrating rhetoric studies and computational methodologies, stating the importance of qualitative and human analysis of LLM's outputs; in section 3 we present the framework proposal based on textual functions and effects on readers, also suggesting how to link functions with specific rhetorical phenomena; in section 4 we analyze a corpus of LLM-generated narrative and argumentative texts to test the framework.

2. Mixed methodology and qualitative analysis

Analyzing the language of LLMs raises important methodological issues, which are discussed in this section. The first point concerns the distinction between deductive and inductive research orientations. A deductive orientation begins with a theoretical framework or set of hypotheses and applies them to empirical material to test or refine them. An inductive orientation, by contrast, begins with the observation of empirical material (texts, data, interactions) and builds theoretical generalizations emerging patterns. The analysis of LLM-generated texts requires both: a deductive component to formalize the evaluative criteria, and an inductive component to test and expand them through engagement with real texts. This paper proposes a double movement: firstly, we introduce a theoretical framework based on functions to analyze LLM-generated texts; secondly, we analyze a corpus of AI-generated texts through rhetorical analysis

and by applying the framework. The two movements are linked and redefine each other in an iterative, adaptive, and scalable process.

Statistical and quantitative approaches to text analysis – such as stylometry, classifiers, and topic modeling – are prevalent in the study of LLM-generated texts (Soundararajan and Delany 2024; Lucy and Bamman 2021). Although these approaches can process large corpora, they are not designed to capture phenomena such as narrative relationships, plot development, character agency, rhetorical structure, and the construction of meaning. As several recent studies have argued, phenomena such as communication, rhetoric, and the production of meaning are not measurable but evaluable and require the integration of qualitative methods (Reinhart et al. 2025; Chandra et al. 2024; Tong et al. 2024; Ichien et al. 2024; Yao et al. 2024). In line with these positions, and consistent with our previous work (Raffini et al. 2025a; Raffini et al. 2025b), we maintain text analysis as a purely human activity. This choice also carries an ethical dimension that we consider worth making explicit. The use of LLMs as judges for evaluating LLM-generated texts removes the human from the evaluative process, creating an automation loop in which a text produced by a computational system is assessed exclusively by computational systems. This is ethically problematic because it undermines the principle of human-in-the-loop that is central to responsible AI development, and because it forecloses the interpretive accountability that evaluation requires. A recent Alan Turing Institute white paper identifies precisely this gap – the absence of interpretive frameworks from the humanities and qualitative social sciences in current AI systems – and advocates for their systematic integration (Hemment et al. 2025). Our methodological choice aligns with this call.

Qualitative human analysis requires a shift of gaze in LLM studies: while quantitative approaches demand experimental evidence and reproducible procedures, qualitative analysis rests on the authority of the interpreter, that is, the human expert who performs the analysis, whether a literary scholar, a rhetorician, or a communication researcher. This criterion of validity is well established in humanistic research traditions and has a precise theoretical basis. Gadamer's hermeneutics argues that understanding in the human sciences cannot be modeled on the experimental method of the natural sciences: meaning emerges not from the neutral application of a procedure, but from the historically situated judgment of the interpreter, shaped by tradition and what Gadamer calls *phronesis*, i.e. the practical wisdom that enables the expert to grasp the significance of a text in context (Gadamer 1975). Applied to LLM evaluation, this means that assessing whether a generated text communicates effectively, constructs meaning coherently, or deploys rhetoric appropriately cannot be reduced to a procedure: it requires a competent human reader exercising situated judgment. Fish's theory of interpretive communities adds a further dimension: meaning is produced not by isolated readers but by communities of practice, groups of experts who share methodological conventions, disciplinary assumptions, and interpretive standards consolidated over time (Fish 1980). The authority of the expert evaluator is therefore not arbitrary or merely subjective but derives from membership in such a community. In LLM assessment, this implies that qualitative evaluation gains its validity not from reproducibility in the experimental sense, but from the rigor of the interpretive tradition within which it is conducted and from the evaluator's accountability to a community of peers. We therefore argue that the evaluation of LLM-generated texts must integrate the interpretive component and the authority of the humanist expert, not necessarily as an alternative to quantitative methods, but as the indispensable framework within which their results become meaningful.

3. The Framework

In this section, we outline an ethical and rhetorical framework for analyzing LLM-generated texts, based on functions and adopting a reader-oriented approach. Given the specific focus on the language of LLMs examined from an ethical perspective, it is necessary to understand how stylistic and rhetorical devices shape textual features and, consequently, influence readers' perceptions and experiences. Subsection 3.1 provides the state of the art regarding reader-oriented approaches in rhetoric studies; subsection 3.2 provides the description of the functions proposed in our framework; subsection 3.3 proposes to operationalize the framework by linking functions to rhetorical devices, starting with the examples of rhetorical figures.

3.1 A reader-oriented approach

This subsection outlines the theoretical foundations of the proposed framework and the gaps in existing studies that it seeks to address. Rhetoric has traditionally been studied in relation to the speaker's intention: the classical tradition conceives of rhetoric as the art of persuasion, in which a human agent deliberately deploys linguistic strategies to achieve a communicative goal. Cognitive and affective effects on readers, by contrast, have remained comparatively less explored. Nevertheless, relevant contributions can be identified in interdisciplinary fields such as cognitive science and neuro-rhetoric (Jack and Appelbaum 2010), which provide insights into the perceptual and emotional processing of rhetorical forms. The choice to ground the framework in these reader-oriented approaches is motivated by the displacement of intentionality that characterizes LLM-generated discourse: if the ethical and analytical interest lies in the effects that AI-generated texts produce on readers, the most relevant theoretical contributions are those that study rhetoric from the reader's perspective. An important contribution in this direction can be found in Utsumi (2005), which examines the capacity of rhetorical figures to engage and move the reader. His research identifies the «resolution of incongruity» as a key cognitive mechanism: figurative language introduces an element of surprise or dissonance that requires the reader to reconcile this incongruity through active interpretation. The argument is further developed by Calabrese (2012), who situates rhetorical experience within the domains of embodied cognition and neural processing. Drawing on the principles of neuro-rhetoric, Calabrese argues that the persuasive power of rhetorical figures lies in their ability to activate neural circuits associated with schema reinforcement or disruption. When rhetorical expressions conform to or subtly modify established cognitive patterns, they facilitate comprehension and strengthen memory; conversely, when they disrupt these patterns, they require cognitive reconfiguration, often leading to deeper engagement and improved retention. Pilkington argues that rhetorical figures generate «weak implicatures» that prompt readers to invest cognitive effort in extracting indeterminate contextual meanings, producing aesthetic and emotional payoffs that make rhetoric effective (Pilkington 2000).

These contributions establish that rhetoric produces cognitive and affective effects on readers through mechanisms that are, in principle, independent of the speaker's identity or intention. However, as Hamilton and Schneider (2002) noted, empirical research on how rhetorical devices influence readers remains scarce. More importantly for the purposes of this paper, no systematic framework has been developed that applies this reader-oriented perspective specifically to LLM-generated texts – a type of discourse in which the absence of human intentionality is not an analytical complication but a constitutive feature. In human communication, the deployment of rhetorical strategies is embedded in a communicative context shared by sender and receiver, and subject to the norms and responsibilities that govern human interaction. When rhetorical effects are produced by a system that operates without shared context, communicative accountability,

or the constraints of human intentionality, those effects become ethically significant precisely because they are unregulated by any of the social and communicative mechanisms that ordinarily govern persuasion.

Building on these studies, and in response to the gaps identified above, this paper proposes a reader-oriented approach to the analysis of LLM-generated discourse. By shifting the analytical focus from the act of textual production to the dynamics of reception, it becomes possible to more effectively assess the ethical and aesthetic implications of LLMs' rhetorical strategies, as well as their influence on user perception and interpretive engagement. The approach refrains from attributing intentionality or agency to LLMs in the sense of conscious communicative intent: what the framework analyzes is not what LLMs mean to do, but what their language does.

3.2 Description of the functions

Ethics of AI has developed several frameworks to evaluate LLM outputs, which are mostly based on principles, such as the UNESCO Recommendation on the Ethics of AI (2021) and the EU AI Act (2024). On the other hand, studies in computer science have developed benchmarks to assess LLMs based on quantitative and computational approaches. Integrating benchmarks and principles with rhetorical analysis can improve the qualitative side of LLMs' evaluation and help address ethical problems. According to McKee and Porter (2020), in fact, in human communication ethics is rooted in rhetoric and context, meaning that ethical evaluation cannot be separated from how language is used. We acknowledge that moral philosophy encompasses diverse traditions concerning the relationship between rhetoric and ethical behavior, a topic that would deserve a broader and more comprehensive discussion. Within the scope of this study, however, we examine the connection between rhetoric and ethics specifically through the lens of AI ethics. Our objective is to develop a rhetorical assessment framework that incorporates the core principles of AI ethics and operationalizes them within the domain of language, thereby enabling their systematic evaluation in communicative and textual practices.

To address AI ethics beyond the formalization of principles, it is necessary to analyze how AI communicates. In light of the approach outlined in subsection 3.1, we propose a structured classification of rhetorical functions especially conceived for analyzing LLM-generated texts from a reader-oriented perspective. The functions are designed to help detect rhetorical patterns that may carry ethical implications and should be considered as the combination of the purpose of rhetorical devices within the text and the effects they may produce on readers. At the core of our idea, in fact, lies the fundamental issue of empathy: producing an effect on the reader implies establishing an empathic relationship, which may take different forms. In the case of LLM-generated texts, this relationship is simulated rather than genuine: it exists only at the rhetorical surface, without a conscious human sender who shares the reader's communicative context. Rhetoric is a linguistic tool that humans employ within real contexts, whether these reflect reality or not; LLMs, by contrast, deploy it without such grounding, making it potentially more manipulative and misleading because they operate in a context-independent manner. Rhetoric detached from both context and sender risks losing its original function and can easily drift into forms of manipulation, miscommunication, and cognitive flattening.

We propose ten functions: trustworthiness, apologetic, proximity, diversification, ambiguity, emphasis, explanatory, poetic, fairness, and structure. For analytical clarity, these functions can be tentatively grouped into three macro-categories: (1) relational functions, which govern the positioning of the text in relation to the reader (trustworthiness, apologetic, proximity, fairness); (2) semantic-cognitive functions, which concern the production and organization of meaning (diversification, ambiguity, explanatory); and (3) formal-aesthetic functions, which operate at the level of expression and textual organization (emphasis, poetic, structure). These groupings are

heuristic rather than fixed, and functions may overlap across categories depending on context. This set of functions must be considered as a proposal and not as a closed and deterministic system. The definition of these functions is not based on a pre-existing theoretical model, but comes from the combination of:

1. Studies on the functions of communication (Jakobson 1960), (Werlich 1976);
2. principles of AI ethics (UNESCO 2021);
3. studies on AI rhetoric and language along with empirical observation (see Introduction for a review of relevant studies on LLM language and rhetoric).

We believe that this combination may be particularly useful to outline a model for analyzing LLMs' rhetorics. A more detailed description of the functions follows. For each function, the number in parentheses indicates whether it is inspired by communication studies (1), AI Ethics (2), or empirical studies and observation (3).

1. **Trustworthiness:** The effect of constructing the image of the speaker or text as credible, authoritative, or ethically trustworthy. It evokes classical ethos, generating trust in the speaker's competence, sincerity, or moral integrity, and invites the reader to receive the message with confidence. (2)
2. **Apologetic:** This function is characterized by a stance that anticipates potential objections or criticism, employing hesitation, mitigation, or deferential tones to soften the enunciation. It engenders a sense of discursive caution, thereby helping to disarm resistance and reduce perceptions of aggressiveness. (1, 3)
3. **Proximity:** This is achieved by reducing the distance between the speaker and the audience, thereby fostering a sense of familiarity and a shared perspective. This is reached through rhetorical strategies, cultural references, an inclusive tone, or direct address, thereby inviting the reader to feel personally engaged. (1, 3)
4. **Diversification:** A strategy that expands the semantic or argumentative scope of discourse. Diversification counters simplification and enriches message complexity by offering the reader a broader interpretative framework through variation or juxtaposition. (1, 3)
5. **Ambiguity:** It is characterized by the presence of multiple meanings, uncertainties, or semantic openness. Ambiguity invites interpretation and encourages the reader to explore layered or indeterminate meanings. (1, 3)
6. **Emphasis:** Highlights specific elements of the discourse. Achieved through repetition, phonetic accentuation, or syntactic prominence, with the aim of drawing the reader's attention to what is presented as central or significant. (1, 3)
7. **Explanatory:** The discourse's ability to clarify, make comprehensible, or systematically break down complex or abstract content. This function facilitates cognitive accessibility and improves reader understanding. (1, 3)
8. **Poetic:** This function is produced when the aesthetic or formal qualities of language are foregrounded. The poetic function stimulates the reader's sensitivity and imagination, creating resonances beyond the referential content. (1)

9. Fairness: This function is characterized by a rhetorical orientation marked by balance, moderation, or impartiality. Fairness is conveyed through the acknowledgement of alternative viewpoints, expressive caution, or an ethically measured tone that encourages the reader to perceive the discourse as reasoned and fair. (2)
10. Structure: This function is tied to coherent, logical, or hierarchical organization of discourse. Structure confers clarity, progression, and intelligibility, thereby enabling the reader to follow the development of ideas and perceive internal cohesion. (1, 3)

The set of functions serves both as a descriptive taxonomy and as an analytical device: it systematizes formal discursive mechanisms and makes hypotheses about their effects. To be applicable to text analysis and relevant for ethical purposes, we propose to operationalize these functions by linking them to rhetorical devices, thereby clarifying how ethical issues are rooted in language and how working on language may help address some of them. In the next subsection, we propose a possible way to apply the function framework starting from rhetorical figures.

3.3 Linking functions to rhetoric: the example of rhetorical figures

Linguistic and rhetorical analysis encompasses a wide range of approaches and areas of interest, and the definition of rhetoric itself brings together multiple facets (Zhao and Cui 2024). Rhetoric analysis of a text includes different elements, e.g. the organization of discourse, macro-structure, tone selection (from formal to informal, serious to ironic), lexical choices (words and expressions that have a specific impact on the audience), style (the set of linguistic and stylistic choices that define the speaker's voice), and rhetorical figures. This variety makes it difficult to formalize all these elements in a single framework. Nevertheless, for the framework to be operationalized, it is important to link the functions described in subsection 3.2 with specific rhetorical devices. We decided to start with rhetorical figures, also known as figures of speech. These figures involve shifts in meaning and creative uses of language, making them among the most dynamic and distinctive features of human communication. We acknowledge that rhetorical figures may seem outdated in contemporary rhetorical studies; however, we consider them productive for our purposes for several reasons. First, they are prone to formalization, as demonstrated by various ontologies of rhetorical figures (Kelly et al 2010), (Harris et al. 2017), (Mladenović and Mitrović 2013), (Wang et al. 2021), (Kühn et al. 2022), (Kühn et al. 2023). Second, as discussed in subsection 3.1, rhetorical figures have been widely studied in relation to their effects on the reader, showing that they operate both linguistically and cognitively by creating mental images, reinforcing meanings, and directing attention to key points in discourse. Finally, it must be considered that the scalable and open nature of the framework, together with the double movement outlined in section 2, allows us to extend the analysis to other rhetorical aspects, making rhetorical figures only a starting point.

In line with the aim of linking functions with specific rhetorical devices, Table 1 proposes interpretive associations between functions and a set of rhetorical figures. Table 1 indicates (1) the name of the figure and (2) the set of functions marked with an "X" to indicate the association. Criteria that we used to establish possible associations are the following: the definition of each rhetorical figure, the function in context, the way the figure positions the speaker in relation to

the reader, and the cognitive or emotional effects it appears to generate¹. It is important to emphasize that these associations are not deterministic: the presence of rhetorical figures does not necessarily cause or exclude a particular function in the text. The associations are interpretative, suggesting potential tendencies rather than fixed rules. Their validity must be substantiated through qualitative textual analysis, since context may strongly influence the interpretation and effects of rhetorical figures (see Section 3).

Table 1. Associations between functions and rhetorical figures

	Trustworthiness	Apologetic	Proximity	Diversification	Ambiguity	Emphasis	Explanatory	Poetic	Fairness	Structure
Allegory				X	X		X	X		
Alliteration						X		X		
Allusion			X	X	X				X	
Antanacsis					X	X		X		
Antithesis						X	X		X	X
Aporia		X	X		X				X	
Chiasmus						X		X		X
Climax	X					X		X		X
Commoratio	X					X	X			
Decrementum	X					X		X		X
Enumeration	X					X	X			X
Euphemism		X	X						X	

¹ For the definitions, we drew on authoritative sources such as the *RhetFig Project: Computational Rhetorics and Models of Persuasion* (Di Marco and Gladkova 2011), Richard A. Lanham’s *A Handlist of Rhetorical Terms*, and the *Oxford English Dictionary Online*.

Excusatio	X	X	X						X	
Hyperbole						X		X		
Irony				X	X					
Litotes		X			X				X	
Metaphor				X	X		X	X		
Metonymy				X	X			X		
Oxymoron					X	X		X		
Paradox					X	X				
Parison						X		X		X
Periphrasis				X			X	X		
Personification			X					X		
Polyptoton						X		X		
Preterition		X			X	X				
Repetition						X		X		X
Reticence					X				X	
Sententia	X					X	X			
Proverb	X		X							
Simile							X	X		
Symbol				X	X			X		
Taxis	X						X			X

Although the framework includes a substantial number of rhetorical figures, many others have been excluded. The primary criterion for inclusion was the prevalence of a figure in contemporary communication, with priority given to those most employed in discourse. A second criterion concerned the relevance of each figure to both the functions of our framework and LLM-generated discourse. In this regard, tropes are more extensively represented than schemes, as they are more likely to produce specific interpretative and persuasive effects on

readers and are particularly significant for the analysis of LLM-generated texts. This is especially true when LLMs are understood not merely as technical tools, but as cultural objects and meaning-making systems characterized by distinctive modes of generating and organizing discourse.

In the case of LLM-generated texts, it is particularly relevant to analyze the rhetorical figures that shape the perceived relationality of the speaker. Figures such as allusion, aporia, and excusatio, for example, are associated with proximity, simulating a closer dialogic relationship between speaker and reader: allusion presupposes shared knowledge, aporia stages hesitation or doubt, excusatio anticipates and responds to possible objections. These strategies frequently emerge in LLM-generated texts, contributing to the impression of a voice that engages empathetically with the user. Such figures can be regarded as strategies that, more obliquely, produce the effect of anthropomorphizing LLMs, demonstrating how ethical questions can be closely related to rhetorical devices.

Equally significant are figures associated with poetic function, in their ability to evoke, move, or aesthetically structure experience. Rhetorical figures such as metaphor, simile or oxymoron are associated with this function. It is common for LLMs to use figurative language, even when not asked for poetic content, to produce expressions that are perceived as creative or expressive, and to create a sense of emotional depth (see subsection 4.1). The structure and explanatory functions are also important for understanding the rhetorical patterns of LLMs. Figures such as enumeration, climax, taxis and chiasmus offer not only semantic balance but also logical progression or conceptual articulation. These rhetorical devices assist LLMs in simulating argumentative development or hierarchical organization of discourse, thereby rendering output more coherent or persuasive. Furthermore, explanatory functions may emerge through analogies or repetition, suggesting a didactic posture even in the absence of true intentionality.

The framework is not intended to be closed, either in terms of the number and definition of functions or the correlations between functions and figures, which may vary depending on the context. Functions and correlations are not fixed but arise from a human, dialogic classification that can be modified and expanded according to specific communicative contexts.

4. LLM-generated texts analysis

Analysis of real texts is essential for testing and refining the theoretical framework. Accordingly, we apply and attempt to integrate the framework by analyzing a corpus of AI-generated texts. The analysis should be regarded as a limited application of the proposed framework, intended at this stage primarily to illustrate the methodology rather than to provide extensive results for evaluating LLMs. The aim is to understand how the framework can interact with, adapt to, and expand through the analysis of real cases. Rather than creating a new textual corpus from scratch, we rely on two corpora of LLM-generated texts from our previous studies (Raffini et al. 2025a), (Raffini et al. 2025b). These corpora have already proven valuable, as they allowed us to identify complex manipulative discourse and gender bias through close reading and rhetorical analysis. Moreover, they represent two common textual types, namely argumentative and narrative. Full details of corpus composition, generation methodology, and data distribution are provided in the original publications. The distinctive feature of the analysis presented here, compared with previous analyses conducted on the same corpora, is that the rhetorical analysis is now carried out starting from functions specifically designed to define an ethical process for evaluating LLM-generated texts, as shown in Section 3.

Each text was analyzed by human experts (specifically, the first and second authors of this paper, both of whom have a background in literary, linguistic, and rhetorical studies) with the aim of identifying and annotating the rhetorical figures. Rather than working independently and subsequently measuring inter-annotator agreement, we decided to conduct the analysis together through joint working sessions, ensuring continuous discussion and consensus on the identification and classification of rhetorical devices, according to the principles of human accountability and shared responsibility described in Section 2. The counts of rhetorical figures identified, which will be presented in subsections 4.1 and 4.2, are therefore the result of a qualitative human analysis. The quantitative dimension helps make the phenomenon more readable and measurable, but numbers are the result of a human, interpretative annotation process. Furthermore, building on these quantitative findings, we offer interpretations grounded in the contextual and semantic dimensions of the texts.

4.1 Analysis of the argumentative-text corpus

The first corpus consists of 62 argumentative texts generated by ChatGPT during a controlled user study, with each participant producing one text using a fixed prompt requiring a ~3000-character text supporting a clear opinion on one of two ethically nuanced topics. The generation followed a standardized methodology, where all participants used the same prompt and the free version of ChatGPT-4 (April 2025) to ensure consistency across outputs. In the previous study on this corpus, the texts were embedded in a pre-post experimental design, where users' opinions were measured before and after exposure to the generated content to assess persuasive impact (Raffini et al. 2025b). Along with the user study, we also analyzed the texts using rhetorical analysis. In the present paper, we will build on that rhetorical analysis with special attention to the functions described in the proposed framework.

The corpus shows a certain degree of uniformity, both in terms of content and textual structure. For this reason, only 10 texts were included in the analysis, with five texts for each topic, specifically, texts 1–5 (topic 1: robot care for the elderly) and texts 40–44 (topic 2: the productivity of the 4-day workweek). Given the uniformity of the texts, the sample can be considered representative, particularly since the primary aim here is to test the methodology. Table 2 shows the occurrence of rhetorical figures in the corpus. The number is only indicative, since the definition of some rhetorical figures may be subject to slightly different interpretations.

Table 2. Occurrences of rhetorical figures in argumentative corpus.

Rhetorical figure	Occurrences
Enumeration	57
Antithesis	49
Climax	45
Excusatio	18
Parison	17
Hyperbole	16
Sententia	12
Metaphor	11
Personification	10

Periphrasis	8
Alliteration	4
Euphemism	4
Aporia	3
Chiasmus	1

Results show a predominance of rhetorical figures indicating accumulation (enumeration, climax, parison) or opposition (antithesis, excusatio). Enumeration and climax often take the form of a tricolon, i.e., the juxtaposition of three elements, while adjectival and syntagmatic formations often follow the juxtaposition of two terms, coordinated either by asyndeton or polysyndeton. The presence of these figures is disproportionate, resulting in excessive structural and formal repetitiveness, combined with redundancy in terms of content. Although there are no specific cases of commoratio, the texts present what we might call a structural or diluted commoratio, i.e., the tendency to repeat concepts throughout the text, in many cases also with the same lexical choices. Similarly, we could not find examples of taxis, but the principle underlying this rhetorical figure – which is the division of a concept into two or more parts – can be noticed in the overall text structure.

Table 3 shows functions' values, which are obtained by summing up the rhetorical figures related to each function, according to the correlations established in Table 1.

Table 3. Functions in argumentative corpus.

Function	Value
Emphasis	201
Structure	169
Explanatory	137
Trustworthiness	132
Poetic	112
Fairness	74
Proximity	35
Apologetic	25
Diversification	19
Ambiguity	14

The functions typically expected in argumentative texts are represented, namely structure, explanatory, and trustworthiness. An intermediate presence of functions such as proximity and apologetic also aligns with expectations, as does the relatively low presence of diversification and ambiguity. Argumentative texts typically aim at clarity and persuasion, leaving little room for doubt or polysemy. The correspondence between textual type and the results obtained confirms

the overall validity of the framework. By examining the differences from expectations and focusing on details, we can identify the distinctive characteristics of the corpus under analysis with respect to the textual type it belongs to, and thus some characteristics of argumentative texts generated by LLMs compared to those written by humans.

A significant finding is the predominance of the emphatic function. While emphasis is a useful rhetorical tool for highlighting key aspects of an argument and making a discourse more convincing, its overrepresentation indicates a rhetorical strategy that focuses more on persuasion itself than on conveying a message. Emphasis should appear, together with proximity and apologetic, among the intermediate values, as these functions tend to persuade through empathy and emotion rather than through rigorous and truthful argumentation. This is also confirmed by the relatively lower values of the functions of trustworthiness and fairness, which appear to be affected by the overrepresentation of the emphasis function. The value of the poetic function also diverges from what would be expected in argumentative writing. Although the poetic dimension is often associated with cultivated or refined writing, such a high value would not normally be expected in argumentative texts. This suggests that the texts, at a rhetorical level, aim to elicit a positive response through a more refined formal style, relying on formal devices that may obscure the centrality of truth and argumentation.

The structure function rightly appears among the highest values, considering the high degree of formalization typical of argumentative texts. This function is further reinforced when other rhetorical elements beyond rhetorical figures are considered, such as the macrotextual dimension, which is highly structured in the texts analyzed. Most texts, in fact, follow a six-paragraph structure consisting of an introduction, four body paragraphs, and a conclusion. This reveals a tendency toward standardized discourse templates, where variation occurs mainly at the level of wording rather than argument organization.

Among the functions, trustworthiness is the one that requires the greatest integration with qualitative and interpretative analysis. It is important to assess whether the effect of trustworthiness is based on real or false elements. In the first case, the argumentation is valid; in the latter, it may constitute rhetorical manipulation. In our corpus, trustworthiness results from the following rhetorical figures: enumeration, climax, excusatio, and sententia. When examining enumeration and climax, the accumulation of terms often appears to be purely formal, without leading to any genuine semantic development. Excusatio and sententia prove particularly useful in defining the boundary between truthful argumentation and rhetorical manipulation. According to Lanham's definition, excusatio means «offering an excuse of such weight that it overcomes all objections» (Lanham, 1991). In the cases analyzed, the responses to criticism do not appear convincing and often consist merely of the repetition of the problem itself. Excusatio is often built on a tautological argument, since the criticism is addressed by invoking the very cause that triggered it, resulting in circular reasoning in which the initial problem corresponds to the reason given for overcoming the objection. In other cases, the critical issues are merely minimized or excusatio tends to be merely formal.

In general, excusatio appears to be used as a method to neutralize criticism without real depth. The rhetorical device creates an effect of trustworthiness that is often not supported by conceptual solidity. A similar case is the use of sententia, which often comes close to clichés. The lapidary and gnomic nature of the sententia is used by LLMs to convey concepts that are not genuine maxims of wisdom, but rather common and in some cases stereotypical ideas. The fact that excusatio and sententia are deprived of their capacity to establish knowledge and reduced to simple persuasive devices shows that the function of trustworthiness, in the cases analyzed, should be understood in a manipulative sense. In other words, the trustworthiness produced through rhetoric does not correspond to actual trustworthiness. This is also an

example that human interpretation remains a fundamental step in the evaluation of LLMs and that a framework is needed which, while starting from the occurrence of rhetorical phenomena, subsequently moves to the interpretative level through qualitative analysis.

4.2 Analysis of the narrative-text corpus

The corpus of narrative texts consists of 15 short stories generated using a zero-shot prompting strategy designed to avoid bias induced by examples. The texts were generated using three LLMs (ChatGPT, Gemini, and Claude – five stories each) in separate sessions and conversational windows to prevent context retention effects. The prompt required a ~500-word narrative structured according to Freytag’s pyramid and including five characters adapted from Propp’s roles (main character, villain, helper, desired character, dispatcher). The methodology followed an open-ended generation approach, allowing variability while maintaining comparability through a standardized prompt. The resulting corpus was intentionally small to enable a close reading and interpretative analysis of narrative and gender biases (Raffini et al. 2025b).

To analyze this corpus using the proposed framework, we follow the same strategy used for the argumentative-texts corpus. Table 4 shows the number of rhetorical figures found in the entire corpus.

Table 4. Occurrences of rhetorical figures in narrative corpus.

Rhetorical Figure	Occurrences
Metaphor	119
Simile	54
Personification	40
Metonymy	36
Hyperbole	16
Symbol	14
Parison	12
Sententia	11
Oxymoron	9
Antithesis	9
Enumeration	7
Paradox	7
Allusion	6
Reticence	4
Repetition	3
Litote	3
Irony	3

Polyptoton	2
Alliteration	1
Euphemism	1
Climax	1

In the case of narrative texts compared to argumentative ones, greater effort was required in interpreting rhetorical figures. This is due, as we shall see, to a rather superficial and unconscious use of rhetorical figures in the corpus, but also by the fact that tropes are often open to interpretation. For example, personification and metonymy may vary depending on the context. Metaphor is another rhetorical figure subject to interpretation, since metaphorical language has entered everyday communication, making it difficult to distinguish between a poetical intention and common use. Given the absence of authorial intention in LLM-generated texts, in our analysis, we decided to use a broad inclusion criterion for metaphor. Table 5 shows the value of each function.

Table 5. Functions in narrative corpus.

Function	Value
Poetic	307
Explanatory	200
Ambiguity	199
Diversification	178
Emphasis	78
Proximity	47
Structure	32
Fairness	23
Trustworthiness	19
Apologetic	4

As in the case of argumentative texts, the dominant functions in narrative texts coincide with the expectations for this textual type. In a narrative text, we expected the predominance of the poetic function, diversification, and ambiguity, which indeed rank among the highest values. Once these values are set aside, as they coincide with the expectations of the genre, the values of the other functions reveal the peculiarities of the narrative generated by LLMs. Particularly relevant are the values of the explanatory and emphatic functions, which do not directly coincide with expectations. The explanatory function, in fact, is more appropriate for informative or argumentative texts. In narrative texts, it is mainly linked to the rhetorical figure of simile, but it would not normally be expected to reach levels close to those of ambiguity, which should predominate given the polysemic nature of creative works. Contemporary literary tradition, in fact, favors rhetorical devices that stimulate processes of meaning construction on the part of the reader, such as symbols and metaphors, whereas simile requires less interpretative

participation. Dissolving ambiguity goes against fundamental features of literature, namely doubt and polysemy. The excessive use of sententia is also related to the explanatory function. Sententia gives the impression of moralizing and overly explanatory writing, dissolving the ambiguity typical of literary texts into clearly stated morals.

The high value of emphasis reflects a rather stereotypical conception of creativity, in which artistic effect is associated with stylistic exaggeration and pretentiousness. This is also linked to an element that emerges from the qualitative analysis of rhetorical figures. As mentioned, in the corpus analyzed, the construction of figures is often quite vague, making their identification difficult in many cases. This is particularly evident in the case of metaphors, which, although prevalent, rarely appear highly meaningful; rather, they often take the form of lemmatized metaphors (i.e., expressions that have entered idiomatic usage) or empty metaphors that do not offer strong or convincing interpretative possibilities. This leads many metaphors to verge on hypotyposis, a rhetorical figure consisting in the vivid description of a scene in such detail that it appears to unfold before the reader's eyes. Hypotyposis is linked to an impressionistic perspective: while it enriches complex rhetorical contexts, in semantically weak contexts the convergence of metaphors into hypotyposis increases the emphatic component of the text. This figure, which was not included in our initial framework, illustrates how qualitative analysis can help refine the framework, while also highlighting how the evaluation of a rhetorical device depends on the context in which it is used.

Another tendency is the scarcity of sound figures, or schemes, which would otherwise be expected in narrative texts. The poetic effect is pursued almost exclusively through content rather than form, confirming the poor literary quality of the texts and the exclusion of the craft dimension of the creative act. Rhetorical figures empty of meaning, the overrepresentation of the emphatic function, the overly explanatory tendency, and the lack of formal attention create a rhetorical dimension that is inadequate for narrative texts and that, although replicating some of their more superficial aspects, proves insufficient when subjected to careful critical scrutiny. The tendency toward trivialization and stereotyping is also reflected in the narrative structures and character development, as we demonstrated in our previous analysis of this corpus (Raffini et al. 2025b). In this case, the framework based on functions productively engages with interpretative approaches from literary studies.

5. Conclusions

This paper has proposed an ethical and rhetorical framework for the analysis of LLM-generated texts, grounded in a reader-oriented perspective and combining deductive theoretical modeling with inductive close reading. The framework identifies ten communicative functions and links them systematically to rhetorical figures, offering a structured instrument for evaluating how LLM-generated language can shape readers' perceptions and interpretations. A defining feature of the framework is its integration of rhetorical analysis with ethical evaluation. By anchoring ethical concerns in specific linguistic phenomena, the framework makes visible how questions of manipulation, anthropomorphization, and cognitive flattening are not external to language but constitutive of it. The analysis of argumentative and narrative corpora has confirmed both the validity of the framework and its capacity to reveal the distinctive rhetorical patterns of LLM-generated discourse. A significant finding is the strong dependence between rhetorical choices and textual type: the coincidence between expected functions and genre conventions serves as a validation criterion, while deviations from expectation constitute the most analytically

productive results. In argumentative texts, the overrepresentation of emphatic and structural functions indicates a rhetorical orientation that prioritizes persuasive form over conceptual depth, often generating an effect of trustworthiness that is not supported by genuine argumentative solidity. In narrative texts, the predominance of metaphor, simile, and personification produces a strong poetic effect, but one that is frequently superficial or conventional, relying on figurative language that lacks semantic density and interpretive complexity.

The framework is not intended as a closed or definitive system. Its functions, associations, and analytical categories are open to revision and expansion. As the analysis has shown, empirical engagement with real texts can refine the framework productively: the emergence of hypotypis as a relevant but initially unlisted figure illustrates how qualitative close reading contributes to the iterative development of the framework. At the same time, the analysis confirmed the effectiveness of using rhetorical figures as a starting point: while the framework can and should be extended to other rhetorical dimensions, figures proved to be a productive entry point both for identifying recurrent patterns and for prompting interpretive refinement. At the methodological level, the paper has also made a broader argument about the necessity of human, qualitative evaluation in LLM assessment. Consistent with the hermeneutic tradition and with recent calls from the humanities and interpretive social sciences, we maintain that evaluating whether a generated text communicates effectively, constructs meaning coherently, or deploys rhetoric responsibly requires the situated judgment of a competent reader, which cannot be replaced by automated procedures.

Future research can extend the framework by applying it to larger and more diverse corpora, enabling systematic comparisons across textual typologies, language models, and communicative contexts. The framework may also be expanded to incorporate additional rhetorical dimensions, such as discourse organization, tone, lexical strategies, narrative structure, as well as cognitive-linguistic approaches such as conceptual blending, image schemata, and emotional evocation, which can deepen the analysis of the poetic and persuasive effects produced by AI-generated language. Through iterative empirical refinement, the framework is designed to evolve into a broader interdisciplinary methodology for evaluating the linguistic, rhetorical, and ethical dimensions of LLM-generated discourse, with a focus on the ethical dimension of linguistic phenomena and on the role of human evaluation in the assessment of AI systems.

References

- Alkhrisheh, Hazim Taisir Dayij, Failasofah Failasofah, Taisir Alkhrisheh. 2021. “The Interrelationship among Language Use, Linguistic Competence, and Higher Order Skills.” *Indonesian Research Journal in Education* 5, no. (april 2021): 6–22. <https://doi.org/10.22437/irje.v5i1.11333>
- Calabrese, Stefano. 2012. *Retorica e scienze neurocognitive*. Roma: Carocci.
- Carrasco-Farre, Carlos. 2024. “Large Language Models are as persuasive as humans, but how? About the cognitive effort and moral-emotional language of LLM arguments.” [arXiv:2404.09329](https://arxiv.org/abs/2404.09329). <https://doi.org/10.48550/arXiv.2404.09329>
- Chandra, Rohitash, Abhishek Tiwari, Naman Jain, and Sushrut Badhe. 2024. “Large Language Models for Metaphor Detection: Bhagavad Gita and Sermon on the Mount.” *IEEE Access* 12: 84452–84469. <https://doi.org/10.1109/ACCESS.2024.3411060>

- Chen, Zhiyi, Jinyi Ye, Emilio Ferrara, and Luca Luceri. 2025. "Prevalence, Sharing Patterns, and Spreaders of Multimodal AI-Generated Content on X during the 2024 U.S. Presidential Election." [arXiv:2502.11248](https://arxiv.org/abs/2502.11248). <https://doi.org/10.48550/arXiv.2502.11248>
- Di Marco, Chrysanne, and Olga Gladkova. 2011. "The RhetFig Project: Computational Rhetorics and Models of Persuasion." CMNA (Computational Models of Natural Argument) Workshop, San Francisco, CA, August 7–11.
- Fish, Stanley. 1980. *Is There a Text in This Class? The Authority of Interpretive Communities*. Cambridge: Harvard University Press.
- Fröhling, Leon, and Arkaitz Zubiaga. 2021. "Feature-Based Detection of Automated Language Models: Tackling GPT-2, GPT-3 and Grover." *PeerJ Computer Science* 7. <https://doi.org/10.7717/peerj-cs.443>
- Gadamer, Hans-Georg. 1975. *Truth and Method*. London: Sheed and Ward.
- Giulianelli, Mario, Joris Baan, Wilker Aziz, Raquel Fernández, and Barbara Plank. 2023. "What Comes Next? Evaluating Uncertainty in Neural Text Generators against Human Production Variability." In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 14349–371.
- Guo, Yanzhu, Guokan Shang, Michalis Vazirgiannis, and Chloé Clavel. 2024. "The Curious Decline of Linguistic Diversity: Training Language Models on Synthetic Text." In *Findings of the Association for Computational Linguistics: NAACL 2024*, 3589–3604. Mexico City, Mexico: Association for Computational Linguistics. [10.18653/v1/2024.findings-naacl.228](https://arxiv.org/abs/2406.18653)
- Hamilton, Craig A., and Ralf Schneider. 2002. "From Iser to Turner and Beyond: Reception Theory Meets Cognitive Criticism." *Style* 36, no. 4: 640–58. <https://www.jstor.org/stable/10.5325/style.36.4.640>
- Hanley, Hans W. A., and Zakir Durumeric. 2024. "Machine-Made Media: Monitoring the Mobilization of Machine-Generated Articles on Misinformation and Mainstream News Websites." arXiv preprint arXiv:2305.09820. <https://arxiv.org/abs/2305.09820>
- Harris, Randy Allen, Chrysanne Di Marco, Ashley Rose Mehlenbacher et al. 2017. "A cognitive ontology of rhetorical figures." *Cognition and Ontologies*: 18–21.
- Hemment, Drew, Cody Kommers, et al. 2025. *Doing AI Differently: Rethinking the Foundations of AI via the Humanities*. White Paper. London: The Alan Turing Institute. <https://www.turing.ac.uk/news/publications/doing-ai-differently>
- Ichien, Nicholas, Dušan Stamenković, and Keith J. Holyoak. 2024. "Large Language Model Displays Emergent Ability to Interpret Novel Literary Metaphors." *Metaphor and Symbol* 39, no. 4 (2024): 296–309. <https://doi.org/10.1080/10926488.2024.2380348>
- Jack, Jordynn, and L. Gregory Appelbaum. 2010. "'This Is Your Brain on Rhetoric': Research Directions for Neurorhetorics." *Rhetoric Society Quarterly* 40 (5): 411–37. [doi:10.1080/02773945.2010.516303](https://doi.org/10.1080/02773945.2010.516303).
- Jakobson, Roman. 1960. "Linguistics and Poetics." In *Style in Language*, edited by Thomas A. Sebeok, 350–77. Cambridge, MA: MIT Press.
- Kelly, Ashley R., Nike A. Abbott, Randy Allen Harris, Chrysanne DiMarco, and David R. Cheriton. 2010. "Toward an Ontology of Rhetorical Figures." In *Proceedings of the 28th ACM*

- International Conference on Design of Communication (SIGDOC '10)*, 123–30. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/1878450.1878471>.
- Kendall, Graham, and Teixeira da Silva Jaime A. 2024. "Risks of Abuse of Large Language Models, Like ChatGPT, in Scientific Publishing: Authorship, Predatory Publishing, and Paper Mills." *Learned Publishing* 37 (1): 55–62. <https://doi.org/10.1002/leap.1578>
- Kühn, Ramona, Jelena Mitrović, and Michael Granitzer. 2022. "GRhOOT: Ontology of Rhetorical Figures in German." In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 4001–4010. Marseille, France: European Language Resources Association. <https://aclanthology.org/2022.lrec-1.426/>.
- Kühn, Ramona, Jelena Mitrović, and Michael Granitzer. 2023. "ESTHER: Ontology of Rhetorical Figures in English." In *Proceedings of the Joint Ontology Workshops 2023 Episode IX: The Quebec Summer of Ontology co-located with the 13th International Conference on Formal Ontology in Information Systems (FOIS 2023)*, edited by Fumiaki Toyoshima et al., Vol. 3637. CEUR Workshop Proceedings. Sherbrooke, Québec, Canada, July 19–20, 2023. Aachen: CEUR-WS.org.
- Lanham, Richard A. 1991. *A Handlist of Rhetorical Terms*. Berkeley: University of California Press. <https://doi.org/10.1525/9780520912045>.
- Liang, Weixin, Yaohui Zhang, Mihai Codreanu, Jiayu Wang, Hancheng Cao, and James Zou. 2025. "The Widespread Adoption of Large Language Model-Assisted Writing Across Society." <https://arxiv.org/abs/2502.09747>
- Lucy, Li, and David Bamman. 2021. "Gender and Representation Bias in GPT-3 Generated Stories." In *Proceedings of the Third Workshop on Narrative Understanding*, edited by Nader Akoury, Faeze Brahman, Snigdha Chaturvedi, Elizabeth Clark, Mohit Iyyer, and Lara J. Martin, 48–55. Virtual: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.nuse-1.5>.
- Martínez, Gonzalo, José Alberto Hernández, Javier Conde, Pedro Reviriego, and Elena Merino-Gómez. 2024. "Beware of Words: Evaluating the Lexical Diversity of Conversational LLMs Using ChatGPT as Case Study." *ACM Transactions on Intelligent Systems and Technology*. <https://doi.org/10.1145/3696459>
- McKee, Heidi, and James Porter. 2020. "Ethics for AI Writing: The Importance of Rhetorical Context." In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 110–116. <https://doi.org/10.1145/3375627.3375811>
- Mladenović, Miljana, and Jelena Mitrović. 2013. "Ontology of Rhetorical Figures for Serbian." In *Text, Speech, and Dialogue*, edited by Václav Matoušek and Ivan Habernal, 386–93. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer.
- Muñoz-Ortiz, Alberto, Carlos Gómez-Rodríguez, and David Vilares. 2024. "Contrasting Linguistic Patterns in Human and LLM-Generated News Text." *Artificial Intelligence Review* 57: 265. <https://doi.org/10.1007/s10462-024-10903-2>
- Muzumdar, Prathamesh, Sumanth Cheemalapati, Srikanth Reddy RamiReddy, et al. 2025. "The Dead Internet Theory: A Survey on Artificial Interactions and the Future of Social Media." *Asian Journal of Research in Computer Science* 18, no. 1 (January): 67–73. <https://doi.org/10.9734/ajrcos/2025/v18i1549>
- Pilkington, Adrian. 2000. *Poetic Effects: A relevance theory perspective*. John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.75>

- Raffini, Daniel, Agnese Macori, Lorenzo Porcaro, Tiziana Catarci, and Marco Angelini. 2025a. "How Persuasive Could LLMs Be? A First Study Combining Linguistic-Rhetorical Analysis and User Experiments." 20th International Conference on Artificial Intelligence and Law (ICAAIL) LCIC-CLAIRvoyantS Workshop, 2025. arXiv preprint arXiv:2508.09614. <https://arxiv.org/abs/2508.09614>
- Raffini, Daniel, Agnese Macori, Marco Angelini and Tiziana Catarci. 2025b. "A Close Reading Approach to Gender Narrative Biases in AI-Generated Stories." 2025 IEEE International Conference on Cyber Humanities (IEEE-CH), Florence, Italy, 2025, 1-9. <https://doi.org/10.48550/arXiv.2508.09651>.
- Reinhart, Alex, David West Brown, Ben Markey, et al. 2025. "Do LLMs Write like Humans? Variation in Grammatical and Rhetorical Styles." *Proceedings of the National Academy of Sciences* 122, no. 8 (February). <https://doi.org/10.1073/pnas.2422455122>
- Shaib, Chantal, Joe Barrow, Jiuding Sun Sun, Alexa Siu, Byron C. Wallace, and Ani Nenkova. 2024. "Standardizing the Measurement of Text Diversity: A Tool and a Comparative Analysis of Scores." [arXiv:2403.00553v2](https://arxiv.org/abs/2403.00553v2). <https://doi.org/10.48550/arXiv.2403.00553>
- Soundararajan, Shweta and Sarah Jane Delany. 2024. "Investigating Gender Bias in Large Language Models Through Text Generation". In *Proceedings of the 7th International Conference on Natural Language and Speech Processing (ICNLSP 2024)*, edited by Mourad Abbas e Abed Alhakim Freihat, 410–424. Trento: Association for Computational Linguistics.
- Sun, Zhen, Zongmin Zhang, Xinyue et al. 2025. "Are We in the AI-Generated Text World Already? Quantifying and Monitoring AIGT on Social Media." arXiv preprint arXiv:2412.18148. <https://arxiv.org/abs/2412.18148>
- Tong, Xiaoyu, Rochelle Choenni, Martha Lewis, and Ekaterina Shutova. 2024. "Metaphor Understanding Challenge Dataset for LLMs." In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 3517–3536. Bangkok, Thailand: Association for Computational Linguistics
- UNESCO. 2021. *Recommendation on the Ethics of Artificial Intelligence*. United Nations Educational, Scientific and Cultural Organization
- Utsumi, Akira. 2005. "The Role of Feature Emergence in Metaphor Appreciation," *Metaphor and Symbol* 20, no. 3 (May): 151–72. https://doi.org/10.1207/s15327868ms2003_1
- Wang, Yetian, Randy Allen Harris, and Daniel M. Berry. 2021. "An Ontology for Ploke: Rhetorical Figures of Lexical Repetitions." In *Proceedings of the Joint Ontology Workshops 2021 Episode VII: The Bolzano Summer of Knowledge co-located with the 12th International Conference on Formal Ontology in Information Systems (FOIS 2021), and the 12th International Conference on Biomedical Ontologies (ICBO 2021)*, edited by Emilio M. Sanfilippo et al., Vol. 2969. CEUR Workshop Proceedings. Bolzano, Italy, September 11-18, 2021. Aachen: CEUR-WS.org. <http://ceur-ws.org/Vol-2969/paper42-CAOS.pdf>.
- Werlich, Egon. 1976. *A Text Grammar of English*. Heidelberg: Quelle & Meyer.
- Yakura, Hiromu, Lopez-Lopez Ezequiel, Brinkmann Levin, et al. 2025. "Empirical evidence of Large Language Model's influence on human spoken communication". [arXiv:2409.01754](https://arxiv.org/abs/2409.01754)

- Yao, Ben, Yazhou Zhang, Qiuchi Li, and Jing Qin. 2024 “Is Sarcasm Detection A Step-by-Step Reasoning Process in Large Language Models?” [arXiv:2407.12725](https://arxiv.org/abs/2407.12725). <https://doi.org/10.48550/arXiv.2407.12725>
- Yoo, Dahey, Hyunmin Kang, and Changhoon Oh. 2025. “Deciphering Deception: How Different Rhetoric of AI Language Impacts Users’ Sense of Truth in LLMs.” *International Journal of Human–Computer Interaction* 41, no. 4: 2163–2183. <https://doi.org/10.1080/10447929.2025.2471111>
- Zhao, Xia, and Xiliang Cui. “Rhetorical Studies in the Context of Large Language Models (LLMs).” Apollo - University of Cambridge Repository, 2024. <https://doi.org/10.17863/CAM.108500>