

DOI: <https://doi.org/10.15688/jvolsu2.2026.3.12>

UDC 811.111'33

LBC 81.432.1+81.11



Submitted: 28.11.2025

Accepted: 02.02.2026

**ARTIFICIAL INTELLIGENCE AND THE ENGLISH LANGUAGE:
THE IMPACT OF GENERATIVE MODELS ON WRITTEN DISCOURSE, STYLE,
AND LINGUISTIC COMPETENCE**

Yassien Shareef Saed Al-Bayati

Imam Ja'afar Al-Sadiq University, Kirkuk, Iraq;
Directorate of Religious Education and Islamic Studies, Baghdad, Iraq

Abstract. This article examines the multifaceted impact of artificial intelligence (AI) on contemporary English written communication, with particular attention to generative models such as ChatGPT and other large language models (LLMs). The analysis centers on written discourse, addressing stylistic formation, textual categories, genre conventions, and the redefinition of linguistic competence within digital contexts. Drawing on discourse analysis, stylistics, and sociolinguistics, the research demonstrates that AI technologies, functioning as both medium and mode of communication, are transforming textual conventions, challenging traditional notions of authorship, and blurring the distinction between human and machine discourse. The central argument asserts that AI operates not only as an autonomous agent in linguistic evolution – generating texts through specific algorithms – but also as a crucial driver of linguistic standardization. This process produces texts lacking human nuance, fosters stylistic and genre hybridization, and contributes to the emergence of a homogenized global form of English. Ultimately, LLMs signify a shift from individual to distributed cognition and from socially determined genre conventions to probabilistic constructions shaped by user input and computational recombination. The study concludes that heightened critical awareness of these developments is vital for preserving linguistic diversity and promoting interpretive literacy in an era increasingly shaped by algorithmic text production.

Key words: artificial intelligence, generative models, discourse, written communication, linguistic competence, stylistics.

Citation. Al-Bayati Y.Sh.S. Artificial Intelligence and the English Language: The Impact of Generative Models on Written Discourse, Style, and Linguistic Competence. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2. Yazykoznanie* [Science Journal of Volgograd State University. Linguistics], 2026, vol. 25, no. 3, pp. 137-147. DOI: <https://doi.org/10.15688/jvolsu2.2026.3.12>

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И АНГЛИЙСКИЙ ЯЗЫК: ВЛИЯНИЕ ГЕНЕРАТИВНЫХ МОДЕЛЕЙ НА ДИСКУРСИВНЫЕ ПРАКТИКИ, СТИЛЬ И ЯЗЫКОВУЮ КОМПЕТЕНЦИЮ

Яссин Шариф Саид Аль-Баяти

Университет Имама Джафара Аль-Садика, г. Киркук, Ирак;
Дирекция религиозного образования и исламоведения, г. Багдад, Ирак

Аннотация. В статье рассматривается многоаспектное влияние искусственного интеллекта (ИИ), особенно генеративных (как ChatGPT) и аналогичных языковых моделей (LLM), на английский язык. Анализ сосредоточен на письменном дискурсе, его стилистических нормах, текстовых категориях и жанровых конвенциях, а также меняющихся представлениях о языковой компетенции в цифровую эпоху. Опираясь на анализ дискурса, стилистику и социолингвистику, исследование показывает, как ИИ, выступая одновременно в качестве средства и способа коммуникации, трансформирует текстовые стандарты, ставя под вопрос авторство и стирая границы между человеческим и машинным дискурсом. ИИ также является не просто технологическим субъектом дискурса, порождающим текст по тем или иным алгоритмам, а непосредственным фактором языковой эволюции, способствующим лингвистической стандартизации и формирующим унифицированную глобальную форму английского языка, знаменующую переход от индивидуального к распределенному познанию, от социально обусловленной конвенциональности жанров к их вероятностному конструированию, определяемому пользовательским вводом и вычислительной рекомбинацией. В работе делается вывод, что понимание этих процессов имеет решающее значение для сохранения языкового разнообразия и формирования критической грамотности в эпоху алгоритмического авторства.

Ключевые слова: искусственный интеллект, генеративные модели, дискурс, письменная коммуникация, языковая компетенция, стилистика.

Цитирование. Аль-Баяти Я. Ш. С. Искусственный интеллект и английский язык: влияние генеративных моделей на дискурсивные практики, стиль и языковую компетенцию // Вестник Волгоградского государственного университета. Серия 2, Языкознание. – 2026. – Т. 25, № 3. – С. 137–147. – (На англ. яз.). – DOI: <https://doi.org/10.15688/jvolsu2.2026.3.12>

Introduction

The advent of generative artificial intelligence, following the introduction of OpenAI's ChatGPT in late 2022, has profoundly altered the domain of English written communication. Generative systems have started to perform a dual function, acting both as creators and participants in the writing process. Their capacity to reproduce human-like expression, adapt stylistic patterns, and generate coherent, contextually appropriate text has provoked significant debate about authorship, originality, and linguistic ability [Formosa et al., 2025]. Within this transformation, English, serving as the predominant language of digital interaction, plays a pivotal role.

The study explores linguistic implications of AI-mediated language within established frameworks of discourse analysis and stylistics.

Artificial intelligence is interpreted not as a technical advancement alone but as a sociolinguistic force influencing contemporary textual conventions. The research seeks to determine (a) the extent to which generative models modify the stylistic configuration of English written discourse, (b) their influence on users' linguistic proficiency, and (c) their role in reshaping prevailing conceptions of textuality, authorship, and genre in the context of global English.

Theoretical framework and literature review

1. Discourse and technological mediation

Discourse, understood as both a linguistic and social construct, has consistently evolved in accordance with the media through which it is created and disseminated. Within this perspective,

technological mediation operates as an active agent reshaping communicative practices and linguistic conventions rather than serving as a neutral channel of transmission. Unlike previous instruments confined to conveying human-produced text, generative systems such as ChatGPT, Claude, and Gemini engage directly in the act of textual creation. Functioning as algorithmic collaborators instead of passive conduits, they generate original linguistic structures through probabilistic modeling.

Discourse is the way we use language. It is not just about words: it is deeply connected to society. According to Fairclough, discourse is a “social practice” that both reflects how society works and helps shape it [Fairclough, 2010]. Over time, as our ways of communicating have changed, from writing by hand, to printing, to radio and TV, and now to digital media, discourse has also changed. The researcher argues that these communication technologies do not simply carry our language: they actively change how we talk and what we talk about [Fairclough, 2010]. In other words, new media doesn’t just pass along human ideas: it helps create them. Technology becomes a kind of partner in making meaning.

This transformation gives rise to the notion of algorithmic co-authorship, a condition in which semantic meaning emerges from the interplay between human intention and computational inference. The discursive process, once an exclusively human social activity, now encompasses non-human agents capable of emulating and reproducing human communicative behavior.

Thus, the integration of AI into discourse production indeed challenges traditional linguistic categories like mode and medium. In systemic functional linguistics, mode refers to how language is organized and how texts are put together (e.g., how meaning is structured), while medium refers to the physical or technological channel that carries the message (spoken, written, digital) [Halliday, Matthiessen, 2014, pp. 33-34]. Generative AI models blur these distinctions: they act both as a medium (since they are part of the infrastructure through which communication happens) and as a mode (because they generate structured, cohesive discourse). Generative models function simultaneously as both infrastructural medium and discursive mode. Such convergence gives rise to a distinctive

configuration, in which human input, algorithmic generation, and digital interface collectively form a unified communicative act.

Technological mediation also entails a profound redistribution of linguistic authority on a global scale. English, already preeminent as an international language, attains enhanced dominance through its disproportionate presence in AI training corpora. Within this framework, generative models act simultaneously as instruments and embodiments of linguistic imperialism, perpetuating a standardized, global English that suppresses regional variants and stylistic diversity. Consequently, AI-driven mediation transforms not only the architecture of discourse but also reinforces existing linguistic hierarchies worldwide.

Previously, researchers in digital discourse analysis emphasized that the structures and affordances of communication technologies shape how people talk and write online. For instance, in his work on mediated discourse analysis, Jones argues that technological features are not neutral: they “enable and constrain” the way we communicate, influencing what kind of discursive practices are possible [Jones, 2020].

In the same vein, scholars in digital discourse theory explored how social media platforms (like Twitter, YouTube, or comment forums) give rise to new interaction patterns, dialogic exchange, and intertextual practices [Johansson, Tanskanen, Chovanec, 2021]. For example, Twitter’s tight character limits have been shown to affect users’ syntax and style: with fewer characters, people tend to use more abbreviations, contracted forms, and simpler sentence structures. Meanwhile, on platforms where commenting and discussion are encouraged, users engage in more back-and-forth, dialogic, intertextual exchange – a point made by discourse analysts studying digital practices [Jones, 2020].

When it comes to generative artificial intelligence, these same structural tendencies of online communication seem to be baked into the models. Because LLMs are trained on very large web-based corpora, they absorb common stylistic norms from digital texts (blogs, social media, comment threads). Recent work on model pre-training shows how register variation (i.e., differences in genre or style, such as opinion blogs vs. news) actually influences model behavior [Myntti et al., 2025].

As a result, AI-generated language often reflects a hybrid register: it may combine the clarity and precision of formal writing with elements of conversational, informal language – mimicking how people write in digital settings. Also, linguistic analysis shows that AI-generated texts differ from human texts in measurable ways (e.g., sentence length distribution, vocabulary variety) [Muñoz-Ortiz, Gymez-Rodríguez, Vilares, 2024]. Furthermore, technological mediation by AI affects global linguistic power. English, which already plays a dominant international role, is heavily overrepresented in the data used to train many large language models. Research indicates that LLMs are disproportionately trained on English texts, which can amplify the dominance of English in AI-generated discourse [Helm et al., 2023].

From a sociolinguistic perspective, this echoes ideas of linguistic imperialism. Pennycook's concept of metrolingual practices (from Metrolingualism) explains how globalization reshapes not just which languages are used, but how they are used: local varieties mix, but dominant forms (like standard/global English) often prevail. [Pennycook, Otsuji, 2016].

In this way, generative AI does more than mediate discourse: it can reproduce and reinforce a standardized global English, marginalizing regional or stylistically diverse variants. The technology role becomes not just a passive conduit but an active agent in upholding existing linguistic hierarchies.

2. Generative models and stylistic imitation

Stylistics has traditionally examined the expressive and aesthetic dimensions of language within its contextual environment [Leech, Short, 2007]. Operating through statistical modeling rather than intentional articulation, generative systems do not originate style but reconstruct it from pre-existing textual patterns. In this respect, AI authorship enacts Turing's "imitation game" within the sphere of discourse itself [Turing, 1950]. The output exemplifies what Herring and Androutsopoulos identify as "computational stylization" – a simulation of stylistic features generated through patterned frequency rather than conscious selection [Herring, Androutsopoulos, 2015]. Beyond surface regularity, algorithmic style provokes reconsideration of creativity, imitation, and authorship in linguistic performance.

Generative models reproduce stylistic features through probabilistic mechanisms, producing what may be defined as algorithmic style, marked by structural coherence, syntactic symmetry, and organized discourse flow, yet it lacks the capacity for intentional deviation that characterizes human stylistic creativity – what Carter and Nash describe as the "foregrounding" of linguistic norms to achieve expressive effect [Carter, Nash, 1990]. Although such imitation can yield rhetorically sophisticated expression, it remains devoid of the spontaneous variability characteristic of human creativity, where stylistic shifts convey identity, emotion, and attitude.

The resulting fluency of AI-generated language therefore carries both functional and ideological significance, promoting linguistic standardization and advancing a uniformized global form of English. Consequently, AI-mediated style occupies a contradictory space: it broadens access to refined linguistic forms while simultaneously diminishing the heterogeneity and individuality that endow discourse with its cultural and social depth.

3. Linguistic competence and the role of AI

Chomsky's concept of linguistic competence refers to an idealized, internalized system of grammatical knowledge that enables native speakers to produce and understand syntactically well-formed sentences [Chomsky, 1965]. However, this model deliberately excludes pragmatic, sociolinguistic, and discursive dimensions of language use – domains later captured by Hymes' notion of communicative competence, which encompasses not only grammaticality but also appropriateness, context sensitivity, and interactional effectiveness [Hymes, 1972].

In the contemporary context of artificial intelligence, many higher-order linguistic tasks, such as lexical selection, syntactic variation, and discourse structuring are increasingly mediated by algorithmic systems. Heavy reliance on AI-driven writing tools may lead users to offload aspects of communicative competence to these systems, potentially reducing their active engagement with idiomatic expression, pragmatic inference, and stylistic nuance.

Conversely, some users develop what scholars describe as distributive language competence [Van Lier, 2004] – a dynamic, co-constituted

capacity in which human judgment and algorithmic output jointly shape textual meaning. This aligns with observation that AI does not replace human skill so much as reconfigure it, fostering new meta-skills such as prompt formulation, output criticality, and iterative refinement [Floridi, Chiriatti, 2020]. In this model, linguistic agency becomes distributed across human and machine, reflecting broader theories of extended cognition [Clark, Chalmers, 1998] and sociotechnical co-constructivism [Krakowski, 2025].

From a sociolinguistic standpoint, this evolution signifies a transition from individual to distributed cognition, in which linguistic capability is co-enacted by human users and technological systems. Clark, building on the extended mind hypothesis, conceptualizes this integration as “extended cognition” [Clark, 2008], wherein digital tools become constitutive, not merely auxiliary, elements of cognitive activity. Within AI-supported composition, linguistic performance unfolds through layered mediation: grammar checkers, predictive text, and generative models interact to scaffold and shape textual production. This reconfiguration blurs the classical dichotomy between Chomskyan competence (internal knowledge) and performance (actual use), recasting language mastery as a socio-technical assemblage.

Yet this augmentation, while enhancing fluency and accessibility, risks fostering what Selwyn identifies as “instrumental literacy” – a mode of engagement focused on efficient system operation rather than critical or reflective linguistic understanding [Selwyn, 2016]. Similarly, Bender et al. warn that large language models function as “stochastic parrots,” reproducing patterns without grounding in meaning or intentionality [Bender et al., 2021]. Consequently, the pedagogical imperative lies in cultivating a critical-creative competence [Krakowski, 2025]: one that preserves human interpretive autonomy while ethically and effectively leveraging the generative affordances of algorithmic collaboration.

4. Authorship, originality, and the ethics of AI discourse

Foucault’s concept of the “author function” [Foucault, 1984] is particularly relevant to AI-generated discourse. When a machine produces text, who – or what – is the

author? The concept of authorship becomes decentered, and linguistic style becomes detached from individual identity. This decentering has ethical implications for academic writing, journalism, and creative industries. But some scholars have argued that AI-generated writing challenges traditional notions of intellectual property and linguistic authenticity [Floridi, Chiriatti, 2020].

The decentering of authorship brought about by generative AI extends beyond questions of textual ownership to challenge the very ontology of originality. Traditionally, originality has been linked to individual creativity, intentionality, and affect – qualities that machines, by definition, do not possess. Yet, the linguistic outputs of large language models often exhibit syntactic sophistication and stylistic coherence that mimic the hallmarks of creativity. This paradox destabilizes the humanist assumption that authorship is an expression of a singular consciousness. As Barthes famously proclaimed, “the author is dead” [Barthes, 1977]; in the context of AI-generated discourse, this assertion acquires renewed relevance. What emerges is a form of distributed authorship, where agency is shared among human prompt design, algorithmic patterning, and data-driven intertextuality. Each generated text becomes a recombination of pre-existing linguistic material, recontextualized through probabilistic synthesis rather than original invention.

Ethically, this redistribution of authorship raises critical issues of accountability, transparency, and moral responsibility. When AI systems produce biased, inaccurate, or plagiarized content, determining liability becomes complex. The opacity of algorithmic processes – what Pasquale calls the “black box society” [Pasquale, 2015] – means that the origins of linguistic choices are often hidden from both users and audiences. Furthermore, the normalization of machine-generated text in academic and journalistic domains risks eroding public trust in textual authenticity. Scholars such as Floridi [Floridi, 2023] and Bender et al. warn that without clear frameworks for attribution and ethical disclosure, AI-mediated discourse may blur the boundaries between fabrication and fact [Bender et al., 2021]. Thus, the ethics of AI authorship must move beyond legal ownership to encompass epistemic integrity – the responsibility to make visible the technological mediation behind the words we read and write.

Research methodology

The research adopts a mixed-methods discourse-analytic framework. The corpus comprises both AI-generated and human-produced English texts representing diverse genres, including academic abstracts, journalistic writing, essays, and creative compositions. Through stylistic analysis instruments such as Coh-Metrix and corpus-based concordance, the study measures key textual indicators: lexical variety, syntactic sophistication, coherence devices, and rhetorical organization.

Quantitative outcomes are further contextualized through qualitative discourse analysis. The interpretive component examines modality, stance, and cohesion within the theoretical parameters of systemic functional linguistics [Halliday, Matthiessen, 2014]. Each sample is evaluated for distinctive traits of AI-generated discourse, including tonal uniformity, structural equilibrium, and constrained lexical range. Data interpretation adheres to Fairclough's tripartite model encompassing textual, discursive, and social dimensions of analysis [Fairclough, 2010].

Discussion

1. Stylistic uniformity and the erosion of idiolect

One of the defining characteristics of AI-produced English is stylistic homogeneity. Generated texts display refined syntax, formulaic phrasing, and structurally balanced organization. Although such features promote precision and readability, they simultaneously suppress authorial distinctiveness. The idiolect – the individual linguistic signature of a writer – is replaced by algorithmic regularity. This development aligns with what Bourdieu terms 'the regulation of the linguistic market', wherein adherence to prevailing norms generates symbolic value while constraining creative divergence [Bourdieu, 1991].

The absence of an idiolect in AI-mediated texts signals a broader shift from expressive individuality to standardized textual production. The homogeneity intrinsic to generative systems stems from probabilistic optimization, which diminishes deviations of all kinds to maintain grammatical and stylistic cohesion. Consequently,

the distinctive features that traditionally signal creativity, displayed in the choice of stylistic resources, syntactic deviation, and inventive imagery, are progressively eliminated. This reduction of linguistic individuality not only narrows the aesthetic range but also transforms the notion of authorship into an algorithmically shared act rather than an expression of an individual style. In this environment, linguistic creativity is subject to computational rationality, producing a homogeneous, cohesive discourse that parallels broader mechanisms of cultural homogeneity within the digital sphere.

2. Register hybridization

Register hybridization refers to the linguistic phenomenon where distinct language registers – varieties of language used in different social contexts or communicative situations – blend or overlap, creating hybrid registers that combine features of multiple traditional registers. Scientific research in linguistics and corpus studies has examined how hybrid registers exhibit both the core characteristics of their influencing registers and unique lexico-grammatical nuances that emerge from this blending. Studies such as [Degaetano-Ortlieb, 2019] analyze hybridization effects in literary texts, showing that hybrid registers maintain resemblance to traditional registers while also reflecting features from more influential registers, analyzed through dimensions like lexico-grammatical features and process types, under frameworks including register theory [Biber, Conrad, 2019] and Systemic Functional Linguistics [Halliday, Matthiessen, 2014]. This helps explain inter- and intra-textual variation in hybrid registers [Degaetano-Ortlieb, 2019].

More broadly, register variation research reveals systematic linguistic differences across registers, with multidimensional approaches to studying register identifying distinct patterns in lexical and grammatical usage depending on context, medium, and audience. Registers exhibit overlapping linguistic features, contributing to hybridization observed in both spoken and written language varieties, including academic, conversational, and digital communication styles [Biber, Conrad, 2019].

Empirical work also connects register variation to cognitive and social dimensions of language processing, noting that people adapt their language use based on social setting, interlocutor,

and purpose, which naturally fosters hybridized registers especially in multilingual or digital contexts [Masullo et al., 2025]. In the domain of AI and computational linguistics, although specific scientific articles on “register hybridization” in AI-generated language are emerging, it is recognized that large language models generate text blending formal and informal registers based on training data mixtures and prompt specifications. This leads to outputs that defy rigid genre boundaries and represent a hybrid linguistic style, reflecting the flexibility and context sensitivity of language use [Myntti et al., 2025].

The phenomenon of register hybridization further demonstrates how algorithmic composition mediates between divergent stylistic frameworks. By synthesizing heterogeneous linguistic inputs, generative systems construct a stratified textual voice that unites institutional authority with conversational fluidity. The result may be described as a “synthetic register,” governed less by situational context than by statistical convergence among genres. Such convergence operates ambivalently: it broadens access to academic discourse while simultaneously diminishing its disciplinary specificity and stylistic rigor. As AI systems continue to navigate between formal and informal linguistic codes, the concept of register as a fixed, context-bound category transforms into a fluid continuum shaped primarily by algorithmic processes rather than social convention or communicative circumstance.

3. Genre transformation in the digital sphere

Traditional genre theory [Swales, 1990] conceptualizes genres as socially sanctioned communicative frameworks that organize meaning and purpose within specific contexts. Generative artificial intelligence disrupts established genre distinctions by creating texts that extend beyond traditional communicative frameworks. AI-produced academic prose often integrates journalistic narration, while machine-generated news content may adopt argumentative patterns characteristic of scholarly discourse. Such fluidity reflects the multimodal and cross-linguistic dynamics inherent to digital communication and corresponds with Crystal’s concept of Internet linguistics, which underscores the textual instability of online expression [Crystal, 2019].

The transformation of genre in AI-mediated contexts illustrates how algorithmic composition restructures communicative organization within the digital domain. Functioning outside conventional genre boundaries, generative systems amalgamate varied textual models to produce hybrid constructions that defy fixed classification, thereby redefining the parameters of genre as an adaptable and evolving communicative form. This intergeneric synthesis results in what may be termed “genre convergence,” wherein stylistic, rhetorical, and structural traits from multiple genres coalesce within a single composition. As algorithmic systems assume an active role in meaning production, genre evolves from a stable convention into a negotiable, probabilistic construct determined by user input and computational recombination. The emergent digital discourse environment thus embodies a post-generic condition, in which textual formation is guided by adaptive algorithmic logic rather than fixed social convention.

4. Coherence and cohesion:

mechanical precision versus human subtlety

Coherence and cohesion, while closely related, reflect distinct aspects of discourse quality. Cohesion is the mechanical connection of sentences and ideas through linguistic devices such as conjunctions, pronouns, and lexical ties that create surface-level flow. Coherence, by contrast, involves a deeper, meaningful connection: the logical and semantic unity that makes the text make sense as a whole and aligns with the reader’s expectations and background knowledge.

Mechanical cohesion can be precisely controlled by algorithms – AI language models excel at producing syntactically smooth and grammatically correct transitions between sentences – what can be called “mechanical precision.” However, human discourse subtly navigates coherence through pragmatics, shared context, and interpretative flexibility, intertwining emotional and experiential understanding. This is “human subtlety,” where meaning unfolds not just through linguistic links but through inferred intentions, cultural cues, and adaptive, creative thinking.

AI-generated text often achieves cohesion with apparent ease, but it may lack the nuanced coherence characteristic of human communication,

leading to predictable or superficially logical connections without deep semantic integration or pragmatic engagement. Ultimately, coherence in human discourse exceeds mechanical precision by incorporating subtle, context-sensitive judgments that computers currently cannot fully replicate [Matson, 2025; Gao, Wang, 2023].

This tension between mechanical cohesion and human coherence underscores ongoing challenges for AI discourse systems striving to emulate not only fluent but meaningfully connected human language.

5. The impact on linguistic competence

Exposure to AI-generated text influences human linguistic behavior. Students who rely on AI for writing support tend to adopt the model's preferred syntactic patterns and lexical collocations. While this may increase grammatical accuracy, it narrows expressive range. In contrast, skilled users engage in critical prompting – using AI as a dialogic partner to expand rather than replace linguistic competence. This kind of interaction embodies Vygotsky's concept of the “zone of proximal development” [Vygotsky, 1978], where cognitive growth occurs through mediated interaction.

AI's impact on linguistic competence, therefore, seems much more profound than mere stylistic imitations. Continuous interaction with generative systems fosters what is called “algorithmic alignment”, a tendency to imitate models' statistically more frequent lexical, syntactic, and rhetorical choices. Although these writing habits efficiently contribute to clarity, accuracy, and grammatical fluency, they risk diminishing linguistic creativity and adaptability to the contexts in which writing is produced. Over time, reliance on algorithmic scaffolding may lead to passive language skills when users, rather than developing an individual language repertoire, reproduce patterns optimized by the model. Conversely, when approached critically, AI-assisted writing can stimulate metalinguistic awareness, enhancing language performance by prompting reflection on register, genre, and rhetorical strategy.

The pedagogical challenge, therefore, lies in transforming AI from an instrument of linguistic conformity into a catalyst for reflective and adaptive language learning.

6. Ethical and educational implications

The pedagogical implications of AI writing tools are profound. Institutions must balance the benefits of linguistic support with the risks of dependency and plagiarism. Scholars such as Sayers argue for “AI literacy” – the ability to critically evaluate, adapt, and ethically integrate machine-generated language [Goro, 2025]. Such literacy requires awareness of stylistic homogenization, source opacity, and algorithmic bias. Educators are urged to promote reflective writing practices that combine human creativity with technological affordances.

The ethical implications of AI-assisted writing extend far beyond concerns of academic misconduct, encompassing fundamental issues of authorship, responsibility, and epistemological reliability. Dependence on generative systems introduces challenges of transparency, as users frequently remain unaware of the origins, training biases, and ideological assumptions underlying algorithmic output. The growing normalization of machine-generated language simultaneously perpetuates dominant linguistic standards, marginalizing peripheral varieties and diminishing creative deviation. Consequently, pedagogical policy must promote critical interaction with AI technologies, prioritizing ethical awareness and interpretive autonomy alongside technical competence. Through the development of reflective writing practices that integrate human judgment with algorithmic support, educational institutions can achieve a constructive equilibrium in which technological mediation enriches learning without compromising originality, intellectual agency, or cultural plurality.

Results

The comparative linguistic analysis of AI-generated English reveals several consistent tendencies distinguished from human-authored texts.

– AI-generated texts exhibit significantly reduced lexical diversity, with studies showing that vocabulary richness measured by standardized type-token ratio (STTR) and Measure of Textual Lexical Diversity (MTLD) is 15–30% lower compared to human texts. AI tends to overuse high-frequency collocations such as *in this context*, *it can be argued that*, and *plays a crucial role*,

contributing to lexical repetitiveness and predictability [Farooq et al., 2026];

- sentences generated by AI predominantly follow canonical Subject-Verb-Object structures with less subordination and fewer complex clauses, resulting in higher readability scores but lower rhetorical dynamism and syntactic variation than human writing. AI overuses straightforward syntactic templates, leading to a formulaic style [Liu W., Liu X., 2025; Tudino, Qin, 2024];

- drawing on the categories, namely reference, substitution, conjunction, and lexical cohesion proposed by [Halliday, Hasan, 2014], the findings revealed that the human-generated texts used richer and more varied cohesion devices, including personal and demonstrative references, substitutions, and diverse conjunctions, which enhanced readability and engagement. In contrast, while efficient and coherent, AI-generated texts often lack personal touch and variety, resulting in a more segmented narrative [Faridha et al., 2026];

- AI texts deploy epistemic markers at nearly double the frequency of human-written abstracts (AI: $n = 61$ vs Human: $n = 33$). The findings show that qualitatively, human writers favour the modal verb *may* to signal tentativeness and use direct assertions to express certainty, while AI-generated abstracts predominantly use *can* to project theoretical capability and mediate certainty through high-commitment reporting verbs [Mahmud, Hasan, Khan, 2026];

- variants of Global English converge within AI-generated texts, creating a transnational yet standardized English register. This register minimizes the use of regional idioms and idiomatic expressions, resulting in a homogenized style that emphasizes international comprehensibility over local identity [Lee, Ho, Gu, 2026].

In summary, AI-generated English achieves mechanical precision in cohesion and syntactic clarity but at the expense of human-like subtlety, linguistic variety, and pragmatic depth, leading to stylistic predictability and standardized discourse.

Conclusion

Generative artificial intelligence has emerged as a significant force in the transformation of English discourse. Its impact extends across stylistic conventions, genre configurations, linguistic competence, and the ethical dimensions of communication. Research indicates that AI-generated language can lead to stylistic homogeneity, as it often reflects the patterns and biases present in the training data used to develop large language models. This can result in a blending of registers and increased permeability between genres, which may influence how language is used in various contexts. Furthermore, the interaction between computational logic and sociocultural dynamics contributes to the evolution of global English, raising important ethical considerations regarding authorship, authenticity, and the potential for misinformation.

Although such technology democratizes access to linguistic production, it simultaneously threatens to diminish individuality, creativity, and regional diversity. The sustainability of English writing therefore relies on cultivating critical AI literacy – the capacity to discern, interpret, and ethically regulate the stylistic and ideological influences embedded in algorithmic text generation. In this regard, artificial intelligence functions as both collaborator and reflector of collective communicative values. The essential task for linguists, educators, and authors is to preserve human expressiveness and interpretive depth within an increasingly automated linguistic environment.

REFERENCES

- Barthes R., 1977. The Death of the Author. Barthes R. *Image – Music – Text*. New York, Fontana, pp. 142-148.
- Bender E.M., Gabry T., McMillan-Major A., Shmitchell S., 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610-623. DOI: <https://doi.org/10.1145/3442188.3445922>
- Biber D., Conrad S., 2019. *Register, Genre, and Style*. Cambridge, Cambridge University Press. 420 p.
- Bourdieu P., 1991. *Language and Symbolic Power*. Cambridge, Harvard University Press. 302 p.
- Carter R., Nash W., 1990. *Seeing Through Language: A Guide to Styles of English Usage*. S.I., Basil Blackwell. 267 p.
- Chomsky N., 1965. *Aspects of the Theory of Syntax*. Cambridge, MIT Press. 251 p.

- Clark A., 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford, Oxford University Press. 317 p.
- Clark A., Chalmers D. 1998. The Extended Mind. *Analysis*, vol. 58, no 1, pp. 7-19.
- Crystal D., 2019. *The Cambridge Encyclopedia of the English Language*. Cambridge, Cambridge University Press. 506p.
- Degaetano-Ortlieb S., 2019. *Hybridization Effects in Literary Text*. URL: https://stefaniadegaetano.com/wp-content/uploads/2019/05/abstract_cl2019_hybridization_final.pdf
- Fairclough N., 2010. *Critical Discourse Analysis: The Critical Study of Language*. London, Longman. 591 p.
- Faridha H., Khoirunnisaa S., Maharani N., Nurjanah L., 2026. A Discourse Analysis of Cohesive Devices in Human and AI-Produced Personal Statements. *English Learning Innovation*, vol. 7, pp. 124-141. DOI: 10.22219/englie.v7i1.41321
- Farooq S., Naz T., Habib P., Sajjad K., 2026. A Comparative Linguistic Analysis of Lexical Diversity in AI-Generated and Human-Written Academic Research. *Advance Social Science Archive Journal*, vol. 5, no. 2, pp. 846-858. URL: <https://assajournal.com/index.php/36/article/view/1749>
- Floridi L., Chiriatti M., 2020. GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, vol. 30, no. 4, pp. 681-694.
- Floridi L., 2023. *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford, Oxford University Press. 279 p.
- Formosa P., Bankins S., Matulionyte R. et al., 2025. Can ChatGPT Be an Author? Generative AI Creative Writing Assistance and Perceptions of Authorship, Creatorship, Responsibility, and Disclosure. *AI & Soc*, vol. 40, pp. 3405-3417. DOI: <https://doi.org/10.1007/s00146-024-02081-0>
- Foucault M., 1984. What Is an Author? // *The Foucault Reader* / ed. by P. Rabinow. New York, Pantheon Books, pp. 101-120.
- Gao C., Wang R., 2023. *Distinguishing AI-Generated and Human-Written Text Through Psycholinguistic Analysis*. URL: <https://arxiv.org/html/2505.01800v1>
- Goro G., 2025. The New Literacy in the Age of AI: Educational Foundations, Pedagogical Practices, and Ethical Challenges. Vallejo R.B., ed. *New Literacies in the Age of AI: Ethics, Teaching, and Writing*. Montevideo, South American Publishing, pp. 2-18. DOI: <https://doi.org/10.62486/978-9915-9854-5-9.ch01>
- Halliday M.A.K., Hasan R., 2014. *Cohesion in English*. London, Routledge. 392 p.
- Halliday M.A.K., Matthiessen C., 2014. *Halliday's Introduction to Functional Grammar*. London, New York, Routledge. 786 p.
- Helm P., Bella B., Koch G., Giunchiglia F., 2023. *Diversity and Language Technology: How Techno-Linguistic Bias Can Cause Epistemic Injustice*. DOI: <https://doi.org/10.48550/arXiv.2307.13714>
- Herring S.C., Androutsopoulos J., 2015. Computer-Mediated Discourse 2.0. Tannen D., Hamilton H.E., Schiffrin D., eds. *The Handbook of Discourse Analysis*. Hoboken, John Wiley & Sons, pp. 1753-1764.
- Hymes D., 1972. On Communicative Competence. Pride J.B., Holmes J., eds. *Sociolinguistics: Selected Readings*. Harmondsworth, Penguin Books, pp. 269-293.
- Johansson M., Tanskanen S.-K., Chovanec J., 2021. *Analyzing Digital Discourses Between Convergence and Controversy*. Cham, Springer International Publishing AG. 374 p.
- Jones R., 2020. Mediated Discourse Analysis. Adolphs S., Knight D., eds. *The Routledge Handbook of English Language and Digital Humanities*. London, Routledge, pp. 202-219.
- Krakowski S., 2025. Human-AI Agency in the Age of Generative AI. *Information and Organization*, vol. 35, no. 1, art. 100560 DOI: <https://doi.org/10.1016/j.infoandorg.2025.100560>.
- Lee J.C.-K., Ho J.P.J., Gu M.M., 2026. Translanguaging and Artificial Intelligence in Education Development: Implications for Leadership in Language Education in a Changing World. *International Journal of Applied Linguistics*. DOI: <https://doi.org/10.1111/ijal.70182>
- Leech G., Short M., 2007. *Style in Fiction*. London, Pearson Education. 404 p.
- Liu W, Liu X., 2025. A Comparative Analysis of Syntactic Complexity in Argumentative Essays from Rhetorical Perspective: ChatGPT vs. English Native Speakers. *PLoS One*, vol. 20, no. 8, e0329410. DOI: 10.1371/journal.pone.0329410

- Mahmud A., Hasan W., Khan A.A., 2026. Epistemic Modality in Human and AI-Generated Academic Abstracts: A Corpus-Based Analysis. *Qlantic Journal of Social Sciences*, vol. 7, no. 1, pp. 309-318. DOI: <https://doi.org/10.55737/qjss.viii.26494>
- Masullo C., Casado A., Leivada E., Sorace A., 2025. Register Variation and Linguistic Background Modulate Accuracy in Detecting Morphosyntactic Errors. *Open Journal of Romance Linguistics*, vol. 11, no. 1, pp.1-36. DOI: <https://doi.org/10.5565/rev/isogloss.527>
- Matson J.O., 2025. *How AI Has Disrupted the Meaning of Coherence*. URL: https://www.linkedin.com/posts/drjohnowenmatson_how-ai-has-made-coherence-incoherent-few-activity-7373943814014738433-s5E3
- Muñoz-Ortiz A., Gymez-Rodríguez C., Vilares D., 2024. Contrasting Linguistic Patterns in Human and LLM-Generated News Text. *ArtifIntell Rev.*, vol. 57, art. 265. DOI: 10.1007/s10462-024-10903-2
- Myntti A., Henriksson E., Laippala V., Pyysalo S., 2025. Register Always Matters: Analysis of LLM Pretraining Data Through the Lens of Language Variation. *Computer Science. Computation and Language*. DOI: <https://doi.org/10.48550/arXiv.2504.01542>
- Pasquale F., 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, Harvard University Press. 320 p.
- Pennycook A., Otsuji E., 2016. Lingoing, Language Labels and Metrolingual Practices. *Applied Linguistics Review*, vol. 7, no. 3, pp. 259-277. DOI: <https://doi.org/10.1515/applirev-2016-0011>
- Selwyn N., 2016. *Is Technology Good for Education?* Cambridge, Polity Press. 160 p.
- Swales J., 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge, Cambridge University Press. 260 p.
- Tudino G., Qin Y., 2024. A Corpus-Driven Comparative Analysis of AI in Academic Discourse: Investigating ChatGPT-Generated Academic Texts in Social Sciences. *Lingua*, vol. 312, art. 103838. DOI: <https://doi.org/10.1016/j.lingua.2024.103838>
- Turing A.M., 1950. Computing Machinery and Intelligence. *Mind*, vol. 59, no. 236, pp. 433-460.
- Van Lier L., 2004. *The Ecology and Semiotics of Language Learning: A Sociocultural Perspective*. Norwell, MA, Kluwer Academic Publishers. 264 p.
- Vygotsky L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, Harvard University Press. 159 p.

Information About the Author

Yassien Shareef Saed Al-Bayati, Candidate of Sciences (Philology), Education College, Department of English Language, Imam Ja'afar Al-Sadiq University, Kirkuk, Iraq; Department of Educational Research, Directorate of Religious Education and Islamic Studies, Baghdad, Iraq, Yassien_shareef@ijsu.edu.iq, <https://orcid.org/0000-0001-7219-0235>

Информация об авторе

Яссин Шариф Саид Аль-Баяти, кандидат филологических наук, Университет Имама Джафара Аль-Садики, г. Киркук, Ирак; Дирекция религиозного образования и исламоведения, г. Багдад, Ирак, Yassien_shareef@ijsu.edu.iq, <https://orcid.org/0000-0001-7219-0235>