Registered in Germany Europa Revue Journal
ISSN: 1862-6564 Vol. 6 No. 11 (2026) Published November 17, 2026

The Impact of Artificial Intelligence on Language Evolution and Communication

Scientific supervisor: Nigora O‘rozova

English Teacher, Department of Language Teaching

Sabina Xoldorova

First-year student, Tourism and Hospitality Programme, Samarkand Institute of Economics and Service (SamIES), Uzbekistan

Abstract

Large language models (LLMs) are now everyday tools for drafting text, raising the question of whether they also influence how people write and speak. This narrative review synthesises recent evidence on the impact of artificial intelligence (AI) on lexical change and communication. Corpus studies show that LLM-preferred vocabulary has risen in scientific writing and spoken academic discourse since late 2022: at least 13.5% of 2024 biomedical abstracts showed signs of LLM processing, LLM-modified content in computer science abstracts rose from 2.4% to 22.5%, and “delve” increased significantly in speech. Implications for language education are outlined.

Keywords

artificial intelligence; large language models; language change; AI-mediated communication; lexical convergence; language education

Introduction. Language evolves through social contact, technology and usage frequency. Hancock et al. defined AI-mediated communication as communication in which an intelligent agent operates on behalf of a communicator by modifying, augmenting or generating messages [1]. Since ChatGPT’s release on 30 November 2022, such mediation has become routine. Because LLMs reproduce statistical regularities of their training data, their preferred wording may spread to users who read and adopt it; whether this occurs at measurable scale is the central question of this review.

Aim of the study. To synthesise recent evidence on how AI, and LLMs in particular, affects vocabulary change and communication, and to draw implications for language teaching.

Materials and methods. A narrative review was conducted in November 2026 using keyword searches ("large language models", "language change", "vocabulary", "AI-mediated communication", "homogenisation"). Sources published in 2020–2026 that analysed linguistic output in a corpus or offered an explicit conceptual framework were included; five were retained (three corpus studies, two conceptual works). Data source, sample size, time window and main indicator were compared descriptively; no pooling was attempted. No human participants were recruited.

Results. Written language. Kobak et al. analysed over 15 million PubMed abstracts (2010–2024): at least 13.5% of 2024 abstracts were estimated to have been processed with LLMs, and about 40% in some subcorpora [2]. Liang et al. examined 1,121,912 papers (2020–2024); the share of LLM-modified sentences in computer science abstracts rose from 2.4% (95% CI 2.1–2.7) in November 2022 to 22.5% (95% CI 21.7–23.3) in September 2024, against 2.5% to 7.7% in mathematics [3]. Spoken language. Yakura et al. analysed about 740,000 hours of speech (over 7.35 billion words). Using a synthetic control, they found that "delve" rose significantly after ChatGPT's release ($p = 0.010$), and many ChatGPT-preferred words grew by roughly 25–50% per year [4]; the effect was significant in science, business and education podcasts but not in religion or sports. Conceptual evidence. Sourati et al. argue that reliance on the same models may narrow linguistic and reasoning diversity [5].

Discussion. LLM influence is measurable, uneven and faster in technical fields. The spoken-language result is the most notable: machine-preferred words are apparently internalised and reused in speech [4], which fits the AI-mediated communication model [1]. Field differences [3] probably reflect writing norms and tool adoption. The homogenisation argument [5] suggests that lexical convergence may signal a wider loss of stylistic variety, relevant to learners of English as a foreign language. However, most data show association with a technological event; causal support comes only from the synthetic-control design in [4], and word-frequency shifts are not evidence of grammatical change.

Strengths and limitations. Strengths: very large corpora, both written and spoken data, and reported confidence intervals. Limitations: few sources, a narrative search,

reliance on English-language academic and preprint data, and no evidence yet for Uzbek or other less-resourced languages.

Practical significance. In tourism and hospitality education, where English is the language of service, students can be taught to draft with AI tools while recognising formulaic phrasing and keeping an appropriate register.

Conclusion. AI is already associated with measurable lexical shifts in written and spoken English, strongest in technical domains, and raises concerns about stylistic homogenisation. Longitudinal studies of learners and of languages other than English are the research priority.

References

1. Hancock JT, Naaman M, Levy K. AI-mediated communication: definition, research agenda, and ethical considerations. J Comput Mediat Commun. 2020;25(1):89-100. doi:10.1093/jcmc/zmz022

2. Kobak D, González-Márquez R, Horvát EÁ, Lause J. Delving into LLM-assisted writing in biomedical publications through excess vocabulary. Sci Adv. 2025;11(27):eadt3813. doi:10.1126/sciadv.adt3813

3. Liang W, Zhang Y, Wu Z, Lepp H, Ji W, Zhao X, et al. Quantifying large language model usage in scientific papers. Nat Hum Behav. 2025. doi:10.1038/s41562-025-02273-8

4. Yakura H, Lopez-Lopez E, Brinkmann L, Serna I, Gupta P, Soraperra I, et al. Empirical evidence of large language model's influence on human spoken communication. arXiv:2409.01754. 2024.

5. Sourati Z, Ziabari AS, Dehghani M. The homogenizing effect of large language models on human expression and thought. arXiv:2508.01491. 2025.