Original research article / Originalus mokslinis straipsnis
Vol. 1 No. 9 (2025): Spalis

AI-Mediated English: How Generative Systems Reinforce English as a Global Lingua Franca

[1] Gerda Urbaite

Accepted: 10.15.2025
Published: 11.10.2025
https://doi.org/10.69760/portuni.0109004

Abstract. English has long been the world's lingua franca, dominating international communication, science, and media. The recent rise of generative Artificial Intelligence (AI) systems – from large language models like ChatGPT to AI-powered translation and writing tools – is poised to further entrench English's global dominance. This article explores how AI-mediated communication may be reinforcing English as the de facto global language. We synthesize current research and examples to examine biases in multilingual AI performance, the standardization of English via AI tools, and the ways AI adoption encourages even greater use of English worldwide. Results indicate that popular generative AI systems disproportionately favor English (and standard varieties of English), often at the expense of linguistic diversity. AI-generated content tends to homogenize toward mainstream English norms, marginalizing minority languages and non-standard dialects. Furthermore, non-native speakers increasingly rely on AI to produce English text, accelerating the spread of English in academia, business, and everyday communication. We discuss the implications of these findings for language learning and global linguistic equity. Finally, we highlight strategies – from diversifying training data to critical pedagogical practices – that could mitigate AI's English-centric biases and foster a more multilingual AI future.

Keywords: *Generative Artificial Intelligence, Global English, English as a Lingua Franca, Language Education*

Introduction

English has achieved an unprecedented global status as a lingua franca, serving as the primary medium for international communication in business, science, education, and diplomacy (Zeng & Yang, 2024). Over 1.2 billion people worldwide speak English either as a first or additional language (Zeng & Yang, 2024), and English is the predominant language of the internet and digital media (Zeng & Yang, 2024). This dominance did not occur overnight; historical processes of colonization, economic power, and cultural influence established English as a global language over centuries (Zeng & Yang, 2024; Lee et al., 2025). In the current era of the Fourth Industrial Revolution—marked by rapid advances in digital technology and artificial intelligence—English's global dominance not only endures but is expected to extend its reach even further (Zeng & Yang, 2024). English remains the primary language of international discourse and technological innovation, reinforced by the prominence of Silicon Valley and the outsized influence of American and British institutions in science and technology (Rajesh &

[1] Urbaite, G. Editor, Euro-Global Journal of Linguistics and Language Education. Email: urbaitee0013@gmail.com. ORCID: https://orcid.org/0009-0001-5471-6210

Padma, 2025). As a result, English is anticipated to become increasingly dominant as the language for programming and interacting with advanced technologies, including AI and machine-learning systems (Rajesh & Padma, 2025; OpenAI, 2023).

Against this backdrop, the rise of generative AI systems represents a new factor that could amplify English's role as a global lingua franca. Generative AI refers to algorithms—most notably large-language models (LLMs)—capable of producing human-like text, translations, images, and other content. Systems such as OpenAI's GPT-3/GPT-4 (the engines behind ChatGPT), Google's Bard, Meta's LLaMA, and others have rapidly been adopted for tasks ranging from answering questions and writing essays to real-time translation and personal tutoring. Crucially, these AI systems are largely products of an English-centric digital ecosystem, having been trained on vast troves of internet text that are predominantly in English (Vashee, 2023). It is estimated that more than 90 percent of GPT-3's training data was in English (Vashee, 2023), and OpenAI acknowledges that GPT-4's training and alignment processes were designed and tested mostly in English and from a U.S.-centric perspective (OpenAI, 2023). While state-of-the-art language models do ingest multilingual data, their capabilities and performance tend to be strongest in English, reflecting the unequal representation of languages online—often termed the *resourcedness gap*—(Lai et al., 2023; Vashee, 2023). In practical terms, English is the first language of AI: it is both the primary source of training information and the default language in which these systems excel (Crotty, 2024).

This article investigates how generative AI may be mediating and reinforcing the use of English globally, potentially at the expense of other languages. We define *AI-mediated English* as the evolving form and function of English as influenced by AI systems—from how AI biases elevate certain norms of English to how users worldwide leverage AI to communicate in English. While some observers hope AI translation and communication tools could level the linguistic playing field by making all languages equally accessible, current evidence suggests a more paradoxical outcome: AI often ends up privileging English even more. Popular large-language models like ChatGPT perform best in English and other high-resource languages but struggle or produce lower-quality output in less-resourced tongues (Sharma et al., 2024; Lai et al., 2023). These disparities risk deepening the digital-language divide rather than eliminating it (Patterson, 2025; Lee et al., 2025). Moreover, by making English content creation and translation easier, AI may further incentivize the use of English in international contexts, thereby reinforcing its dominance.

Research Questions. In this study, we ask: In what ways do generative AI systems reinforce the position of English as the global lingua franca? We examine three interrelated dimensions: (1) biases in multilingual AI systems that favor English (and other dominant languages) in information access and quality of output; (2) the tendency of AI tools to standardize English-language usage around certain norms (potentially narrowing the range of World Englishes); and (3) the impact of AI adoption on language behavior, particularly whether it encourages increased use of English by non-native speakers. Our goal is to synthesize emerging research on these issues and discuss the implications for linguistic diversity, language learning, and policy.

Significance. Understanding AI's role in global language dynamics is crucial for educators, policymakers, and technologists. If AI is inadvertently amplifying English linguistic hegemony, there are risks of further marginalizing speakers of less-supported languages and diminishing linguistic diversity (Amin et al., 2025; Lee et al., 2025). On the other hand, if approached critically, AI could also be harnessed to support multilingualism—for example, by improving translation or providing language-learning support in various tongues. By shedding light on the largely unprecedented phenomenon of AI-mediated English, this article offers insights into how we might steer technological development and pedagogical practice to promote equity and inclusivity in global communication.

Methodology

To address the research questions, we adopted a qualitative, integrative research design. First, we conducted a literature review of recent studies, reports, and theoretical papers at the intersection of AI and language use. Given the novelty of generative AI, much of the relevant research has emerged in the last 2–3 years (2022–2025), including peer-reviewed studies in computational linguistics and language education, preprint articles, and policy reports. We surveyed work on multilingual performance of large language models, AI in English language teaching (ELT), sociolinguistic analyses of AI output, and discussions of English as a lingua franca in the digital era. Key sources were drawn from journals in applied linguistics, AI conferences, and technical reports by AI developers. We prioritized studies with empirical evaluations of AI's multilingual capabilities or those offering data on how AI tools are used in various linguistic contexts (Lee et al., 2025; Sharma et al., 2024; Lai et al., 2023).

Second, we performed a content analysis of example cases and data points illustrating AI's language biases and effects. This involved examining documented examples such as AI-generated outputs in different languages, known issues with translation systems, and user behavior (e.g., researchers using ChatGPT to polish English manuscripts). We also included illustrative quotes and qualitative insights from educators and users about AI's influence on language use (Amin et al., 2025; Lepp & Smith, 2025).

Importantly, our analysis takes a *global Englishes* perspective, recognizing English not as a monolithic entity but as comprising diverse varieties and uses worldwide. We examine not only how much AI favors English, but which English it promotes—for instance, whether AI tools privilege Standard American/British norms over local English dialects (Amin et al., 2025). This approach is informed by *World Englishes* and *English as a Lingua Franca* (ELF) research paradigms, which emphasize pluralized English practices (Lee et al., 2025).

Finally, we synthesized findings to formulate an argument about AI-mediated English. Rather than a controlled experiment or survey, this study is primarily analytic and interpretive, drawing connections between disparate pieces of evidence to understand a broader sociolinguistic trend. We acknowledge that generative AI is a fast-moving field; therefore, our analysis represents a snapshot of the current state (as of late 2025) and an informed projection of possible trajectories.

Data Sources

Our sources include peer-reviewed articles from major journals such as the *Annual Review of Applied Linguistics* (Lee et al., 2025) and the *Qualitative Research Journal for Social Studies* (Amin et al., 2025), as well as technology-focused venues like the *Humanities & Social Sciences Communications* (Zeng & Yang, 2024) and *British Journal of Educational Technology*. Conference proceedings from computational-linguistics meetings—ACL, EMNLP, and NAACL—also provided empirical studies on multilingual model performance (Lai et al., 2023; Sharma et al., 2024). Whenever possible, we cited works that include DOI identifiers or stable URLs to ensure verifiability.

In total, more than thirty relevant publications were reviewed, and over twenty are cited directly in this article to substantiate key arguments. Additional contextual evidence was drawn from technical documentation and reports by major AI developers such as OpenAI (2023) and from professional commentaries and news features addressing policy debates around AI and language (Crotty, 2024; Patterson, 2025).

By combining empirical findings with sociolinguistic theory, this methodology enables a comprehensive understanding of how AI and global English interact. The following section presents the results of this review and analysis, organized around the principal themes identified.

Results

1. Biases in Multilingual AI Favoring English

Generative AI systems demonstrate a strong performance bias toward English, which reinforces English's dominance in information access. Numerous evaluations have shown that large-language models (LLMs) such as GPT-3 and GPT-4 perform substantially better in English than in most other languages (Sharma et al., 2024; Lai et al., 2023). For example, a comprehensive test of ChatGPT across 37 languages and multiple tasks found significantly worse performance for non-English inputs and outputs, especially on complex tasks (Lai et al., 2023). High-resource languages (e.g., Chinese, French, Spanish) fare reasonably well, though often still below English, while low-resource languages show dramatic drops in fluency, accuracy, and coherency of AI-generated text (Zeng & Yang, 2024; Vashee, 2023). This performance gap stems partly from training-data imbalances—English data dominate the corpora, and many less-common languages have sparse representation (Vashee, 2023). One analysis noted that English users get a far better experience with GPT-4 than users of other languages and that current models risk amplifying existing language inequities in global commerce and knowledge access (Crotty, 2024; Vashee, 2023). In short, the AI "brain" has effectively been wired with an English-centric knowledge base.

A consequence of this bias is that the language one uses to query or interact with an AI can determine the quality and completeness of information received (Patterson, 2025; Lee et al., 2025). A recent study from Johns Hopkins introduced the term *information cocoons* to describe how multilingual LLMs give very different answers depending on the query language (Sharma et al., 2024). In their experiments, when the same question was posed in English, Arabic, Hindi, and Chinese, the answers diverged significantly in content and depth (Sharma et al., 2024). The LLM tended to retrieve and present information primarily from sources in the same language as the query, even when more comprehensive

information existed in another language (Sharma et al., 2024). This means that a user asking in a minority language might only get a narrow slice of information, missing perspectives that are well documented in English. Moreover, if information in the query language is scarce, the model often falls back to using English sources, thereby injecting the dominant English-language perspective into the answer (Sharma et al., 2024; Patterson, 2025). Sharma et al. (2024) conclude that, rather than truly breaking language barriers, current *multilingual* AI may reinforce dominant views and contribute to *linguistic imperialism*—the phenomenon of English overshadowing other languages in knowledge production. These findings highlight that, without intervention, LLMs could deepen the digital-knowledge divide between English and non-English information spheres.

Translation asymmetry is another manifestation of AI's English tilt. Generative models and AI translators are notably more adept at translating *into* English than translating English *into* other languages (Crotty, 2024). In evaluations, systems such as ChatGPT show strong ability to interpret foreign-language input and express it in English, thanks to abundant parallel data and English serving as a pivot language (Crotty, 2024). However, the reverse direction—producing, say, Arabic or Korean output from English input—often yields stilted or error-prone results (Lai et al., 2023). This asymmetry partly reflects the same training imbalance but also the fact that developers tend to fine-tune and evaluate models mostly on English outputs (Crotty, 2024). An internal OpenAI analysis acknowledged that the majority of GPT-4's training and alignment were conducted in English and that safety and quality tests were primarily U.S.-centric (OpenAI, 2023). Non-English capabilities are considered "a bonus" or side effect of the training (Crotty, 2024). The implication is that English enjoys first-class support, while other languages are effectively treated as lower-priority features.

This bias can have a self-reinforcing effect: users around the world may find that they get better results from AI systems by using English, which in turn encourages using English even when it is not their native language. Early evidence shows users strategically switching to English to interact with AI whenever possible. For instance, researchers have observed that ChatGPT can often handle tasks in other languages more accurately if the prompt is given in English—with a request to output in the target language—than if the prompt itself is in that language (Vashee, 2023; Crotty, 2024). This leads to the somewhat paradoxical situation that, to get a high-quality answer in language X, one might need to communicate with the AI *in English about language X*. Consequently, English proficiency becomes even more of an advantage in the age of AI: those who can formulate their queries and prompts in English can tap more of the AI's potential. In professional settings, it is often observed that high-quality English input tends to result in superior AI-generated output (Rajesh & Padma, 2025). Complex tasks—such as coding assistance or academic writing—frequently demand English interaction to yield the best results (Rajesh & Padma, 2025; Lee et al., 2025).

There is also a concern that AI's English dominance could exacerbate pressure on individuals and organizations to use English, further marginalizing other languages. If cutting-edge tools for business, education, or research "work best" in English, users might feel compelled to conduct more of their work in English to take full advantage of AI assistance. As Thien Huu Nguyen cautions, this dynamic might "exacerbate the bias for English and English speakers," leading people to conform to English norms and neglect their own languages or cultural-linguistic practices (Patterson, 2025; Zeng & Yang,

2024). Over time, this could contribute to a spiral in which English becomes even more entrenched—echoing past eras, from the British Empire to American globalization—but now supercharged by AI technology (Zeng & Yang, 2024).

In summary, the current generation of AI systems displays a marked English-centric bias in performance. Rather than democratizing linguistic access, they risk reinforcing English as the *digital default language*—the path of least resistance for reliable communication with AI. This sets the stage for English to maintain or even widen its lead as the global lingua franca unless conscious efforts are made to improve multilingual AI capabilities.

2. Standardization of English Through AI Tools

Beyond quantitative dominance, generative AI is influencing the qualitative aspects of global English use—effectively standardizing English conventions and narrowing variation. Studies indicate that AI-powered language tools tend to privilege certain "standard" forms of English (particularly American English) in their output, thereby reinforcing those norms worldwide (Amin et al., 2025; Lee et al., 2025). For instance, a 2025 sociolinguistic inquiry by Amin et al. examined AI-assisted writing and found that AI overwhelmingly preferred American spelling and grammar conventions, even when users from other English-dialect backgrounds interacted with the tool (Amin et al., 2025). Common differences like *color* vs. *colour* or *apartment* vs. *flat* were consistently resolved in favor of the American usage, reflecting an underlying bias of the training data toward U.S. English (Amin et al., 2025). Similarly, AI-based grammar checkers and text generators often enforce prescriptive grammar rules (e.g., avoiding split infinitives, using standard formal syntax) that align with traditional normative English teaching (Amin et al., 2025; Rajesh & Padma, 2025). While this can improve correctness and clarity, it also means that non-standard dialect features or local idioms are stripped away in favor of a more homogeneous style.

Researchers have raised concerns that such AI-induced standardization narrows the space for *World Englishes*—the diverse regional and cultural varieties of English used globally (Lee et al., 2025). The very utility of AI in polishing language—helping writers fix errors or awkward phrasing—comes with the side effect of erasing distinctive fingerprints of a writer's dialect or idiolect. In educational contexts, teachers have noted that when students use tools like ChatGPT to draft or correct essays, the end result often reads in a blandly uniform register, lacking the colloquial expressions or local flavors the student might ordinarily use. In Amin et al.'s study, English learners and teachers appreciated AI's help in producing grammatically accurate text, but they simultaneously worried about "cultural erasure" and the loss of unique voice (Amin et al., 2025). The AI's preference for globally dominant vocabulary and phrasing can create a subtle pressure to conform to a monolithic English norm, potentially diminishing confidence in local expressions or second-language varieties. Over time, if AI tools are widely used to generate content—from work emails to published articles—we could see an increasing share of English text worldwide conforming to a similar tone and standard: an AI-mediated Global English that is efficient but less variegated.

Evidence of this trend is apparent in academia and science communication. As mentioned earlier, many non-native English researchers now use LLMs to refine their writing. A study of computer-

science papers' peer reviews in the wake of ChatGPT's emergence provides a telling insight: reviewers noticed that after late 2022 (when ChatGPT became publicly available), manuscripts—especially from authors in countries where English is not the first language—started to exhibit more uniform language quality, with fewer grammatical errors or idiosyncrasies (Lepp & Smith, 2025; Rajesh & Padma, 2025). At first glance, this suggests AI helped overcome some language barriers. However, reviewers then adapted by looking for other linguistic cues of non-nativity. They reported that certain turns of phrase and stylistic patterns common to LLM-generated text (for example, the frequent use of transitional words like *Moreover* or formal verbs like *delve into*) began appearing so often that they were taken as signs the text was AI-influenced and possibly authored by a non-native scientist (Lepp & Smith, 2025). In other words, a new kind of AI-shaped academic English is emerging—one that is grammatically impeccable and formulaic in style. Lepp and Smith (2025) refer to this as "ChatGPT style" and observed that reviewers sometimes infer an author's background from these stylistic tells. Crucially, the ideology equating "good English" with "good science" persisted; if anything, AI made it less about overt grammar mistakes and more about subtle stylistic conformity (Lepp & Smith, 2025; Rajesh & Padma, 2025). The study suggests that LLM use is reproducing existing language ideologies: to be taken seriously, one's English must sound like polished, standard academic prose—a bar that AI can help reach, but at the cost of homogenization (Amin et al., 2025).

From a global perspective, the risk is that AI acts as a sociolinguistic force imposing central norms. As Bender et al. (2021) famously argued, large-language models can behave like "stochastic parrots," regurgitating patterns from their training data without regard for context or diversity (Lee et al., 2025). If those training patterns over-represent standard American or British English, the model's outputs will by default carry those patterns forward (Lee et al., 2025; Zeng & Yang, 2024). This has been documented: Jeon et al. (2025) note that LLM-based tools often lack sensitivity to lingua-cultural diversity and tend to reflect mainstream norms unless explicitly guided otherwise (Lee et al., 2025; Amin et al., 2025). For example, in conversational AI, researchers have found that models might correct or ignore user input that uses non-standard English or code-mixed language (mixing English with another language), steering the dialogue back to standard English. This *algorithmic standardization* can inadvertently send the message that deviations from "proper" English are errors to be fixed rather than natural variations.

The marginalization of less-dominant English varieties is a serious concern for linguistic justice. English as used in India, Nigeria, Singapore, or among second-language communities has valid local norms—lexical, grammatical, and pragmatic. However, AI systems rarely reflect the breadth of these *World Englishes*. Instead, because of data imbalance and design choices, they lean into monolingual ideologies—the assumption that there is one correct form of English. The outputs thus perpetuate the notion that British or American English is the default English, and others are substandard. This dynamic was pointed out by Lee et al. (2025) in an ELT context: they warned that without intervention, LLMs may "perpetuate monolingual ideologies, reinforce societal inequities, and undermine linguistic justice" in language education (Lee et al., 2025). Biases in training data that marginalize minority voices or dialects can reinforce stereotypes and even limit intercultural understanding among learners (Amin et al., 2025; Lee et al., 2025).

To illustrate, consider spelling conventions taught by AI grammar tools. An ESL student writing *organisation* (an accepted spelling in many regions) might be prompted by an AI assistant to change it to *organization* (the U.S. spelling), even if the student is following British convention. The AI doesn't explain that both are acceptable; it simply flags the non-preferred one as a mistake. Such experiences, multiplied across millions of interactions, nudge everyone toward a uniform norm. Over time, linguistic features that are perfectly grammatical and common in non-U.S. dialects could be abandoned by users in favor of the AI-endorsed forms. Without explicit awareness, AI tools condition users to adopt a narrower form of English.

In summary, generative AI is acting as a global editor—smoothing out differences and promoting a standardized English. This has benefits—improved mutual intelligibility and fewer language errors—but also costs, including loss of diversity and potential cultural bias. English is a richer, more plural language than what AI currently represents. The concept of *AI-mediated English* captures this tension: English as filtered through AI becomes more uniform and possibly more bland, reflecting the biases of its training. This phenomenon is novel and largely absent from traditional discussions of English as a lingua franca, which mostly considered human factors. Now, AI is an active mediator of how English is written and spoken globally. The next section examines how the adoption of AI tools is influencing language choices and usage patterns, particularly the growing ubiquity of English in various domains.

3. AI Adoption and the Proliferation of English Usage

The rapid adoption of generative AI tools around the world is not only a technical trend but also a linguistic one—it is shaping how and when people use English. One clear impact is that AI enables many non-native English speakers to produce content in English with greater ease and confidence, which in turn increases the volume of English communication globally. In professional and academic fields, where English is often the gatekeeper language, AI tools act as enablers for those who previously struggled with English writing. A striking example comes from scientific publishing: since 2023, there has been a surge in researchers using LLMs such as ChatGPT to help write or edit manuscripts in English (Lepp & Smith, 2025). Liang et al. (2024) estimated that by early 2024, about 17.5 percent of sentences in new computer-science preprint papers had been substantially edited or generated by ChatGPT (Lepp & Smith, 2025). The uptake was especially pronounced among authors in countries such as China and non-English-speaking parts of Europe, who historically might have faced language barriers in writing for top conferences (Rajesh & Padma, 2025). This suggests that AI is lowering the barrier to writing in English: researchers can draft in their native language and use AI to translate, or can write in English and rely on AI to refine grammar and style. Consequently, more scholars are contributing to English-language literature than would have done so unaided, simply because the task has become easier.

A similar pattern is observed in business and everyday communication. International companies report that employees who are not fluent in English are using AI translation and email-drafting tools to correspond with clients or colleagues in English rather than defaulting to local languages. By mediating these interactions, AI effectively allows English to penetrate domains where it might not have been

used due to human language-skill limitations. For instance, a small business owner in a non-English-speaking country can now create an English version of their product website or draft marketing emails in English with AI assistance, potentially reaching a broader audience. The availability of AI as a "universal translator" or writing coach means that using English has become more accessible to billions of people. Importantly, this does not mean those people have learned English in the traditional sense; rather, they can *perform* English through AI mediation. It blurs the line between who is an English user and who is not, since AI can fill in linguistic gaps.

In the field of education—especially English-language learning—generative AI is double-edged. On one hand, AI tutors and chatbots provide personalized English practice to learners worldwide, potentially accelerating English acquisition. Adaptive learning platforms can converse with students in English, correct their mistakes, and adjust to their proficiency level (Lee et al., 2025; Amin et al., 2025). This has led some to hail AI as a democratizer of English education—bringing quality English exposure to areas lacking human teachers. On the other hand, as AI translation improves, some educators wonder whether future students will bother to deeply learn foreign languages (English included) when instant translation is at their fingertips. However, current evidence suggests that English proficiency remains crucial even with AI translation. AI may translate basic meanings, but nuances and idioms often get lost or distorted (Crotty, 2024). Thus, being truly effective in using AI tools (for example, prompt engineering or interpreting AI outputs) often requires a solid grasp of English (Rajesh & Padma, 2025). As Rajesh and Padma (2025) note, many emerging tech roles—from AI prompt engineers to data analysts—demand strong English skills because interacting with AI systems and writing effective prompts hinge on mastery of English vocabulary and pragmatics. In short, English remains a key to unlocking AI's full potential.

The net effect of these trends is that English continues to solidify its position in global domains, with AI accelerating the process. We can already see AI's hand in the linguistic landscape: the majority of content generated by AI—articles, social-media posts by bots, and so on—is in English unless otherwise specified. Many AI-content platforms default to English output. For example, an AI image generator might require English prompts to produce the best results, or a coding assistant might only understand instructions given in English, since programming languages and documentation are predominantly English. This creates a virtuous cycle for English: the more English data and usage, the better AI performs in English, which further encourages using English to interface with AI.

However, it is important to consider who might be left out or disadvantaged by this shift. People with limited English proficiency may not benefit equally from the AI revolution. They might rely on AI translation, which is still imperfect and can introduce errors or misunderstandings (Patterson, 2025; Vashee, 2023). There is also the issue of tokenization and cost: some languages are more verbose or use writing systems that result in more tokens for the same content, meaning that using an AI service in those languages can be more expensive or hit usage limits faster (OpenAI, 2023). For instance, early analyses showed that GPT-4 could require significantly more tokens to express a given message in certain non-English languages compared with English, due to how it tokenizes text (OpenAI, 2023). This effectively makes English the most "cost-efficient" language to use with AI models in terms of token economy. It is another subtle incentive to use English over other languages on these platforms.

In summary, through a combination of enabling non-native speakers to produce English content and the systemic biases that make English the path of least resistance, AI is contributing to the wider spread and entrenchment of English. We term this phenomenon *AI-mediated English proliferation.* It is visible in academia (more papers in English), online content (machine-translated English articles, AI-written English text), and even interpersonal communication (people using AI to chat in English). English was already dominant, but AI is acting as a multiplier.

The results we have outlined depict a landscape in which generative AI currently favors English in both quality and quantity of output, encourages standard forms of English, and provides new avenues for English usage by non-traditional users. In the next section, we discuss what these findings imply for the future of global communication and linguistic diversity. We also consider what interventions or changes could ensure that AI develops in a way that supports multilingualism rather than undermining it.

Discussion

Our analysis reveals a paradox at the heart of AI-mediated language: the very technologies that have the potential to bridge language gaps are, in practice, reinforcing the supremacy of English. This has multifaceted implications—educational, cultural, and ethical. In this discussion, we unpack these implications and explore how stakeholders might respond to ensure a more linguistically equitable AI future.

Implications for Language Learning and Education

From a language education perspective, the findings highlight both a boon and a challenge. Learners of English worldwide now have AI-powered tools at their disposal to practice and perfect their skills. This can accelerate learning outcomes; for example, AI tutoring systems can provide instant feedback on pronunciation and grammar, simulating immersive English environments in regions where competent teachers are scarce (Amin et al., 2025; Alisoy & Sadiqzade, 2024). However, the ease of relying on AI might also encourage surface-level competence—students may lean on AI to correct their English without fully internalizing the rules or nuances. More critically, English educators must grapple with the fact that AI's feedback is rooted in standard norms. Teachers should therefore integrate *critical AI literacy* into curricula, helping students question why a correction is made, whose English is being privileged, and what counts as acceptable variation (Lee et al., 2025). Pedagogical frameworks such as *Global Englishes Language Teaching* (GELT) become even more relevant: they emphasize exposure to diverse English varieties and communication over adherence to one "correct" form. Integrating AI into GELT could involve deliberately challenging AI—inputting non-standard dialect sentences and discussing why the AI flags them and whether that affects intelligibility.

Educators can also leverage AI's multilingual affordances to counterbalance English's pull. Generative AI can produce translation exercises or conversational practice in lesser-taught languages (though with varying quality). Schools might adopt a bilingual AI-assistant model: encouraging students to use AI in both English and their native language, thus developing parallel literacies. This approach would mitigate the risk of English overshadowing local languages in education. However, for this to be

effective, AI's capacity for local languages must be strengthened—a responsibility shared by developers, educational institutions, and governments through investment in localized NLP research and corpora (OpenAI, 2023; Vashee, 2023).

Cultural and Sociolinguistic Implications

At a societal level, AI's reinforcement of English extends the historical narrative of linguistic hegemony into the digital age. English linguistic hegemony has long been critiqued for its role in diminishing smaller languages and imposing a particular worldview (Zeng & Yang, 2024; Lee et al., 2025). AI now risks becoming an unwitting agent of that process under the banner of technological progress. As our results indicate, people adapt their linguistic behavior to suit AI—for instance, writing in standard English to obtain better responses—a subtle form of accommodation in which human linguistic choices are shaped by machine preferences. Over time, this could dilute cultural identity encoded in language. Because language reflects worldview, if AI interactions condition users to phrase ideas within English-dominant frameworks, valuable conceptual diversity may be lost (Amin et al., 2025; Rajesh & Padma, 2025).

There is also a risk of linguistic complacency among native English speakers. If non-native users rely on AI to communicate in English, Anglophone users may feel even less incentive to learn other languages or engage with linguistic diversity. This could exacerbate existing imbalances in intercultural communication, where English speakers expect others to accommodate them. AI might unintentionally promote *Anglophone insularity*—the belief that monolingualism is sufficient because technology will handle translation (Zeng & Yang, 2024). Such a dynamic undermines mutual multilingual exchange and empathy for the challenges faced by non-native English users.

Policy, Development, and Technological Interventions

The trajectory of AI-mediated language is not predetermined. To avoid perpetuating English dominance, AI developers must prioritize multilingual optimization. This entails curating large, high-quality datasets in a wide range of languages and dialects, and treating multilingual performance as a *core benchmark* of AI success rather than an afterthought (OpenAI, 2023; Lai et al., 2023). Some encouraging progress exists—initiatives to create open parallel corpora for low-resource languages and research challenges dedicated to multilingual NLP (Sharma et al., 2024). Policymakers and international bodies should view such initiatives as critical infrastructure for linguistic equity, akin to preserving cultural biodiversity. As several AI ethics panels have emphasized, we must strive to "break English dominance" by designing AI systems that inherently support linguistic diversity (Patterson, 2025; Vashee, 2023).

A related priority is addressing algorithmic fairness and inclusivity. Just as gender or racial bias is now recognized as an ethical concern, *linguistic bias* should also be foregrounded. AI tools should undergo linguistic-bias audits to examine disparities in output quality across languages and dialects. For example, if an AI essay evaluator consistently scores Indian English lower than American English despite similar content, that bias must be corrected (Lee et al., 2025). Training models to recognize and respect contextual variations—rather than treating deviations as "errors"—is a step toward

linguistic justice. Some researchers advocate incorporating *translanguaging* and *code-switching* practices into AI design, reflecting how multilingual speakers naturally use hybrid linguistic repertoires (Amin et al., 2025). Future systems should handle multilingual input fluidly rather than enforcing rigid monolingual boundaries.

Rights, Preservation, and the Future of Linguistic Diversity

From a digital-rights standpoint, access to AI in one's own language can be viewed as a *linguistic right*. Just as accessibility standards ensure equal participation for people with disabilities, linguistic accessibility ensures equitable participation in digital society. Regulations and public funding could require major AI companies to expand language support and open-source local-language models. Indeed, community-driven initiatives are already emerging for languages such as Bengali, Swahili, and Basque, though they remain underfunded (Crotty, 2024).

Paradoxically, AI could also become an ally in language preservation. The same generative models that now privilege English could be adapted to revitalize endangered languages through automatic transcription, text-to-speech systems, and interactive teaching tools (Amin et al., 2025). However, such efforts must involve native speakers and cultural experts to ensure authenticity and respect for linguistic identity.

Finally, the concept of *AI-mediated English* introduced here implies the emergence of a distinct variety—a technolect of English shaped by AI's preferences and constraints. Scholars should begin documenting its grammatical, lexical, and stylistic tendencies: Does it simplify syntax? Prefer certain discourse markers? Alter tone or register? Such inquiry would help educators and developers alike recognize where AI's influence is reshaping linguistic norms (Lepp & Smith, 2025). Teachers might, for instance, caution students not to imitate formulaic AI phrasing ("In addition," "Moreover") uncritically.

Ethical Reflections

A broader ethical question underlies these debates: Do we want a future in which all communication passes through one linguistic filter, even if technologically efficient, or one that sustains the world's linguistic plurality? The findings of this study sound an urgent warning: without conscious effort, AI may nudge humanity toward linguistic homogenization dominated by English. Yet, as numerous scholars argue, maintaining multilingual ecosystems is both possible and necessary (Lee et al., 2025; Zeng & Yang, 2024). Strategies such as multilingual education, local-language media promotion, and supportive language policy must accompany AI's rise. Ultimately, we should treat the current AI–English nexus as a call to pursue *inclusive innovation*—ensuring that technological progress empowers speakers of all languages rather than privileging only those fluent in English.

Conclusion

Generative AI is reshaping the linguistic balance of our globalized world—and in doing so, it appears to be tilting the scales further toward English. Our examination of contemporary research and examples has shown that *AI-mediated English* is a double-edged phenomenon. On one side, AI tools

have made English communication more accessible than ever, helping non-native speakers write and speak in English and breaking down some barriers in international exchange. On the other side, these same tools exhibit entrenched biases: they prefer English, perform best in English, and often encourage English-centric norms in subtle ways. The result is that English's position as the global lingua franca is not challenged by AI, but rather reinforced and possibly redefined by it.

Several notable findings emerged. First, current multilingual AI systems privilege high-resource languages—especially English—in both information retrieval and content generation (Sharma et al., 2024; Lai et al., 2023). This creates a risk that speakers of less-resourced languages will be trapped in "information cocoons" or feel compelled to use English to achieve better outcomes when interacting with AI (Sharma et al., 2024; Zeng & Yang, 2024). Second, AI outputs tend to standardize English, promoting mainstream dialects and grammatical norms while filtering out regional or non-standard variation (Amin et al., 2025; Lee et al., 2025). While this can improve clarity, it raises concerns about the diminishing space for *World Englishes* and the potential loss of linguistic diversity. Third, the widespread adoption of AI is accelerating the use of English across multiple domains—from academia, where AI helps scholars publish in English (Lepp & Smith, 2025), to business and media, where translation tools enable English-language outreach (Rajesh & Padma, 2025). These trends collectively underscore that AI is not yet the great language equalizer many hoped for. Instead, it reflects and amplifies the existing linguistic power structure, with English at the apex (Zeng & Yang, 2024; Rajesh & Padma, 2025).

However, the future trajectory of AI and global language use remains within our influence. To avoid a scenario in which AI becomes an agent of linguistic homogenization—or an instrument of global *Anglification*—deliberate and inclusive measures must be taken. Our review suggests several strategies:

- **Diversify AI Training and Evaluation.** Developers should expand language diversity in training datasets and treat multilingual performance as a central benchmark of AI success. Future large-language models (LLMs) should aim to produce equally high-quality output across a broad spectrum of world languages. Initiatives to collect data for low-resource languages and to develop benchmarks beyond English are crucial (Lai et al., 2023). Linguists and cultural experts must be directly involved to ensure that lesser-spoken languages and dialects are handled accurately and respectfully rather than reduced to English approximations.

- **Incorporate Linguistic Fairness in AI Ethics.** The AI community should regard linguistic bias as an ethical concern equal to racial or gender bias. Documentation such as system or model cards should transparently disclose language coverage and deficiencies (OpenAI, 2023). When imbalances—such as favoring one dialect—are identified, they should be addressed through iterative retraining and feedback from international users.

- **Critical AI Literacy in Education.** Educators should integrate discussions of AI's linguistic biases into language and media literacy curricula. Students and users must understand that AI outputs are not neutral; they reflect the data and perspectives that shaped them (Lee et al., 2025). By cultivating awareness, users can resist overreliance on AI's standardized English,

verify translations with native speakers, and preserve their unique linguistic expressions even when AI suggests changes.

- **Policy and Support for Linguistic Diversity.** Governments, universities, and cultural institutions have a key role to play in ensuring that national and minority languages thrive in the AI era. This includes funding NLP research, mandating multilingual accessibility in AI services, and promoting open-source projects for underrepresented languages. International organizations such as UNESCO and the European Union could coordinate global initiatives on *AI and multilingualism* (Patterson, 2025; Vashee, 2023). Just as biodiversity is preserved through global collaboration, linguistic diversity deserves similar protection in digital spaces.
- **Designing AI for Multilingual Interaction.** Future AI systems should move beyond monolingual design paradigms. Advanced models could support code-switching, translingual dialogue, and cross-language summarization—allowing participants to converse naturally in their preferred languages while AI mediates understanding. Rather than everyone defaulting to English, such systems could make multilingual interaction the norm, embodying the inclusive potential of AI.

In closing, the notion of *AI-mediated English* encapsulates a defining reality of our time: English is no longer spreading solely through human-to-human contact but through human–AI–human communication loops that transform and transmit language. This dynamic offers both a challenge and an opportunity—to observe linguistic evolution in real time and to ensure it unfolds equitably. As scholars, technologists, and educators, we must ask what kind of linguistic world we want AI to help create. Will it be one where a single, algorithmically polished English dominates all discourse—or one where technology empowers all languages to coexist and flourish? The current trajectory leans toward the former (Zeng & Yang, 2024), but with awareness and intentional action, we can alter course. Generative AI, a product of human ingenuity, should ultimately serve humanity's full linguistic richness—not merely reinforce the hegemony of the already powerful. By recognizing how AI reinforces English's global position, we take the first step toward ensuring that the next chapters of global communication are written in many voices, not just one.

References

Amin, M., Rasheed, H., Ali, S., & Ara, A. (2025). Artificial intelligence and the standardization of global English: A sociolinguistic inquiry. *Qualitative Research Journal for Social Studies, 2*(3), 365–384. https://doi.org/10.63878/qrjs3.10

Alisoy, H. (2022). The importance of listening in language acquisition. *Xarici Dillərin Tədrisi Və Tədqiqində Ənənəviliyin Və Müasirliyin Vəhdəti, 5*(5), 44–50.*

Alisoy, H. (2025). Structural and semantic taxonomy of English phraseological units: A theoretical perspective. *Porta Universorum, 1*(8), 61–82.*

Alisoy, H. (2025). From echo chambers to critical dialogue: A comparative case study of social media-based pedagogy for addressing scientific misinformation. *Global Spectrum of Research and Humanities, 2*(4), 35–47. https://doi.org/10.69760/gsrh.0250203002

Alisoy, H., Hajiyeva, B., & Sadiqzade, Z. (2024). *Connect with English A2–B1 Speaking Handbook. Journal of Azerbaijan Language and Education Studies, 1*(2), 1–115.*

Alisoy, H., Mammadova, I., Asadova, B., Ismayilli, F., & Aliyeva, T. (2024). The future of language education: Integrating smartphones into the curriculum. *Edelweiss Applied Science and Technology, 8*(6), 4539–4556.*

Alisoy, H., & Sadiqzade, Z. (2024). Mobile-assisted language learning (MALL): Revolutionizing language education. *Luminis Applied Science and Engineering, 1*(1), 60–72. https://doi.org/10.69760/lumin.202400002

Crotty, D. (2024, February 23). AI, translations, and the dominance of the English language. *The Scholarly Kitchen.* https://scholarlykitchen.sspnet.org

Lai, V. D., Ngo, N. T., Veyseh, A. P. B., et al. (2023). ChatGPT beyond English: Towards a comprehensive evaluation of large language models in multilingual learning. *Findings of EMNLP 2023 (Long Papers).* https://openreview.net

Lee, S., Jeon, J., & McKinley, J. (2025). Generative AI and English language teaching: A global Englishes perspective. *Annual Review of Applied Linguistics, 45*, 85–108. https://doi.org/10.1017/S0267190525100184

Lepp, H., & Smith, D. S. (2025). "You cannot sound like GPT": Signs of language discrimination and resistance in computer science publishing. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25).* https://doi.org/10.1145/3715275.3732202

Mammadova, I. (2024). The role of proverbs in language learning: A cognitive and cultural perspective. *EuroGlobal Journal of Linguistics and Language Education, 1*(1), 40–45. https://doi.org/10.69760/8qj8tr41

OpenAI. (2023). *GPT-4 System Card* [Technical report]. OpenAI.

Patterson, J. (2025, September 2). Multilingual artificial intelligence often reinforces bias. *Johns Hopkins Hub.* https://hub.jhu.edu

Rajesh, A., & Padma, P. (2025). Lingua Franca 2.0: The enduring power of English amidst artificial intelligence. *International Journal of Language, Literature and Culture, 5*(4), 60–65. https://doi.org/10.22161/ijllc.5.4.9

Sadiqzade, Z. (2024). The foundational role of auditory skills in language mastery. *Acta Globalis Humanitatis et Linguarum, 1*(1), 82–87.*

Sadiqzade, Z. (2024). The impact of music on language learning: A harmonious path to mastery. *EuroGlobal Journal of Linguistics and Language Education, 1*(1), 134–140.*

Sharma, N., Murray, K., & Xiao, Z. (2024). Faux polyglot: A study on information disparity in multilingual large language models (NAACL 2025 presentation). *arXiv Preprint, arXiv:2407.05502*. https://arxiv.org/abs/2407.05502

Vashee, K. (2023, June 28). Making generative AI effectively multilingual at scale. *ModernMT Blog*. https://www.modernmt.com/blog

Zeng, J., & Yang, J. (2024). English language hegemony: Retrospect and prospect. *Humanities and Social Sciences Communications, 11*(317). https://doi.org/10.1057/s41599-024-02821-z