ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL

GENERATIVE AI AND THE EVOLUTION OF HUMAN DISCOURSE: THE EMERGENCE OF 'SYNTHETIC DIALECTS' IN ONLINE COMMUNICATION

Aqsa Shereen
Assistant Professor (English), Women University Swabi, Swabi.
aqsaharoon333@gmail.com
Atika Rub Nawaz
Lecturer in English, University of Wah, Wah Cantt.
ms.aatika@hotmail.com
Mahnoor Chaudhry
Lecturer, FAST University, Lahore.
mahnoor.chaudhary@nu.edu.pk
Rafia Dastgir
Lecturer, FAST University, Lahore.
rafiadastgir752@gmail.com

Abstract
In recent years, the rise of generative artificial intelligence (GenAI) especially large language models (LLMs) has altered online communication in subtle but profound ways. This paper explores how GenAI contributes to the evolution of human discourse by enabling and multiplying what we term synthetic dialects emergent patterns, registers or "ways of speaking" that blend human and machine-generated language, and which are increasingly present in online forums, social media, chatbots and collaborative writing tools. Drawing on theoretical perspectives from sociolinguistics, discourse theory and human-machine interaction, the present study analyzes how synthetic dialects emerge, how they propagate, and what implications they have for meaning-making, community formation, and human agency in digital spaces. Using a mixed-methods methodology; qualitative discourse analysis of online threads where GenAI outputs are edited or adopted, plus a survey of users' attitudes toward machine-augmented communication, the present paper identifies three key phenomena: (1) the standardisation of GenAI-influenced phrasing across diverse users, (2) the blending of human and machine registers into hybrid linguistic forms, and (3) the rising difficulty of distinguishing purely human discourse from hybrid or machine-influenced discourse. The discussion reflects on how synthetic dialects may reshape norms of participation, authority, and authenticity in online discourse, and concludes with implications for digital literacy, platform design and future linguistic research.

Keywords: Generative AI, Large Language Models, Synthetic Dialects, Digital Discourse, Human–Machine Interaction, Hybrid Authorship, Sociolinguistics, Online Communication, Language Standardization, Digital Literacy

1. Introduction
Digital communication has never been a static or uniform phenomenon. Since the emergence of early computer-mediated communication, linguistic practices in online environments have continually evolved in response to technological affordances, platform architectures, and shifting social norms. Early email exchanges, for instance, introduced new conventions of formality and turn-taking, while instant messaging and social media platforms later fostered abbreviated forms, multimodal expression, emoji usage, and meme-based discourse (Crystal, 2001; Herring, 2007). These changes illustrate that digital spaces function not merely as neutral channels for language use, but as active sites of linguistic innovation and transformation. In recent years, the rapid proliferation of generative artificial intelligence (GenAI), particularly large language models (LLMs) such as GPT-based systems, represents a qualitative departure from earlier stages of digital mediation. Unlike spell-checkers, predictive text, or grammar correction tools, GenAI systems do not simply support human expression; rather, they actively generate extended discourse, recommend stylistic structures, and optimize language for clarity, persuasion, or perceived professionalism. In doing so, these systems increasingly participate in the communicative process itself, shaping not only

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

what is said but also how it is said. This paper argues that sustained interaction with GenAI systems is giving rise to what can be conceptualized as synthetic dialects—emergent patterns of language that are neither fully human-authored nor entirely machine-generated. Instead, these dialects arise through iterative human–AI collaboration, in which users adopt, edit, and normalize AI-generated phrasing over time. Synthetic dialects are observable across a range of digital contexts, including social media captions, professional correspondence, academic writing, customer service interactions, and collaborative documents. Their defining feature lies not in overt markers of artificiality, but in subtle regularities of phrasing, register, and rhetorical structure that reflect the statistical tendencies of GenAI models. The emergence of synthetic dialects has significant implications for how authority, authenticity, and voice are constructed in digital communication. As GenAI-generated discourse becomes increasingly polished and socially acceptable, it may reshape normative expectations of what counts as "good" writing or appropriate expression. This shift raises critical questions about linguistic diversity, human agency, and the future trajectory of language change in algorithmically mediated environments. While existing scholarship has examined AI-generated language primarily through ethical, technical, and epistemological lenses focusing on issues such as bias, misinformation, authorship, and automation (Floridi et al., 2018; Bender et al., 2021) comparatively little attention has been paid to the linguistic consequences of routine GenAI use. Specifically, there is a lack of empirical and theoretical work examining how repeated exposure to machine-generated discourse may influence human writing practices and contribute to the stabilization of new hybrid registers. Addressing this gap, the present study investigates how synthetic dialects emerge, how they circulate across digital platforms, and how users perceive and negotiate machine-augmented communication in everyday discourse.

2. Literature Review

2.1 Digital Discourse and Language Change

Research on digital discourse has consistently emphasized that online communication environments accelerate processes of linguistic change. Computer-mediated communication (CMC) facilitates rapid diffusion of new forms, enables widespread imitation, and reduces institutional gatekeeping over language use (Herring, 2007). As a result, digital platforms often serve as incubators for novel registers, characterized by informality, hybridization, and multimodality (Tagg, 2015). Recent work has highlighted the role of online micro-communities or "linguistic micro-tribes" in shaping niche linguistic identities. Khan, Jameel, Parveen, and Sajjad (2025) demonstrate that digital communication allows small online communities to develop highly specialized registers that reinforce group identity while excluding outsiders. These findings emphasize that digital platforms are not neutral spaces; they actively mediate how language and identity are co-constructed. Androutsopoulos (2014) highlights that language change in digital spaces is closely tied to platform-specific affordances, such as character limits, algorithmic visibility, and interactional norms. Page et al. (2014) argue that digital discourse is shaped by circulation and repetition, whereby linguistic patterns gain legitimacy through visibility and uptake rather than institutional authority. Extending these ideas, research on bilingual youth in Pakistan shows that online communication often involves strategic code-switching, balancing cognitive load and identity negotiation in digital interactions (Khan, Parveen, Qadir, & Rashid, 2025). Such findings underscore the complex social and cognitive dimensions of digital language evolution.

2.2 Generative AI as a Linguistic Actor

Large language models are trained on massive datasets composed of human-authored text, enabling them to generate linguistically coherent and contextually appropriate discourse across genres (Brown et al., 2020). While these models lack intentionality or semantic understanding in the human sense, their outputs are often perceived as authoritative, fluent,

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL

and persuasive (Shanahan, 2022). Crucially, GenAI systems tend to favor statistically dominant patterns within their training data. Bender et al. (2021) warn that this tendency can reinforce hegemonic linguistic norms, marginalizing non-standard varieties and reducing stylistic diversity. Similarly, research on organizational leadership suggests that algorithmic or structured linguistic guidance such as that provided by AI can influence the development of professional discourse norms and perceived credibility (Mehinood & Lawa, 2025; Mehmood, Siddiqui, & Rashid, 2025). When deployed at scale, GenAI-generated discourse may contribute to linguistic standardization, reinforcing certain forms of expression across contexts.

2.3 Hybrid Authorship and Human–Machine Interaction

Emerging research on human–AI collaboration suggests that users rarely treat GenAI outputs as final products. Instead, they engage in iterative processes of editing, refinement, and contextual adaptation (Lee et al., 2022). Over time, this interaction blurs the boundary between original and assisted writing, giving rise to forms of hybrid authorship in which linguistic agency is distributed across human and machine contributors. From a sociolinguistic perspective, these processes mirror dynamics observed in online micro-tribes, where hybridized or borrowed linguistic elements are adapted to fit community norms (Khan, Jameel, Parveen, & Sajjad, 2025). Hybrid authorship challenges traditional notions of ownership and authenticity (Köbis & Mossink, 2021), highlighting that both human and algorithmic influences can become internalized as part of an individual's communicative repertoire. This underscores the conceptual importance of synthetic dialects, which emerge from the intersection of human creativity, AI assistance, and social interaction.

3. Theoretical Framework

To examine the emergence of synthetic dialects in online communication, this study draws on three complementary theoretical perspectives: sociolinguistics and register theory, critical discourse theory, and posthuman approaches to human–machine interaction. Together, these frameworks allow for an integrated analysis of linguistic form, social power, and distributed agency in AI-mediated discourse.

3.1 Sociolinguistics and Register Theory

In sociolinguistics, registers are understood as context-dependent varieties of language shaped by communicative purpose, audience, and situational norms (Halliday, 1978). Registers are not fixed linguistic systems but dynamic configurations of lexical choices, syntactic patterns, and pragmatic conventions that emerge through repeated social use. Traditionally, registers develop within identifiable social groups or professional communities, such as academic, legal, or journalistic discourse. In the context of generative AI, however, register formation is no longer driven solely by human social interaction. Synthetic dialects can be conceptualized as algorithmically mediated registers, emerging from the interaction between human users and GenAI systems that recommend and reproduce statistically dominant linguistic forms. Unlike conventional registers, which evolve through communal negotiation over time, synthetic dialects may stabilize rapidly due to the scale, speed, and repetition inherent in AI-generated language. This perspective highlights a crucial shift in sociolinguistic dynamics: language norms are increasingly shaped not only by peer interaction but also by algorithmic suggestion. As users across disparate contexts adopt similar GenAI-generated phrasing, stylistic conventions may converge, contributing to linguistic homogenization across platforms. Register theory thus provides a valuable lens for analyzing how synthetic dialects emerge, circulate, and acquire social legitimacy in digitally mediated environments.

3.2 Discourse Theory

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL

Critical discourse theory conceptualizes language as a site where power, ideology, and social meaning are produced and contested (Fairclough, 1995). Discourses do not merely reflect reality; they actively shape how knowledge, authority, and identity are constructed. From this perspective, linguistic patterns are inseparable from the social structures and power relations that sustain them. GenAI systems are trained on large-scale corpora that reflect historically dominant discourses, institutional language practices, and culturally privileged forms of expression. As a result, the outputs generated by these systems may implicitly reproduce and legitimize particular ways of speaking often privileging clarity, neutrality, and professionalism as defined by dominant socio-cultural standards. This raises the possibility that synthetic dialects function as vectors of discursive power, normalizing certain rhetorical styles while marginalizing alternative or non-standard forms. By applying discourse theory, this study examines how synthetic dialects participate in broader processes of normalization and authority-building in online communication. Specifically, it considers how AI-influenced discourse may shape perceptions of credibility, expertise, and authenticity, as well as how users negotiate these discursive effects through editing, adaptation, or resistance. This framework allows for a critical interrogation of GenAI not merely as a neutral tool, but as an active participant in the production and circulation of socially consequential discourse.

3.3 Posthuman and Human–Machine Interaction Theory

Posthuman scholarship challenges the human-centered conception of agency by emphasizing distributed cognition across humans, technologies, and material systems (Hayles, 1999). From this viewpoint, meaning-making is not the exclusive domain of individual human subjects, but emerges through complex assemblages involving both human and non-human actors. Synthetic dialects exemplify this posthuman condition of language production. Rather than being authored by a single speaker, AI-influenced discourse is co-produced through iterative interaction between users and GenAI systems. Human intention, algorithmic pattern recognition, and platform affordances jointly shape linguistic outcomes. Authorship thus becomes distributed, and agency is shared across human–machine assemblages. Human–machine interaction theory further emphasizes that users do not passively accept AI outputs; they actively negotiate, modify, and internalize them over time. This process leads to the gradual naturalization of AI-generated linguistic patterns as part of one's own communicative repertoire. The posthuman framework is therefore essential for understanding synthetic dialects as emergent phenomena that cannot be reduced to either human creativity or machine automation alone.

4. Methodology

4.1 Research Design

This study adopts a mixed-methods research design to capture both the structural characteristics of synthetic dialects and users' subjective experiences of AI-mediated communication. Combining qualitative discourse analysis with quantitative survey data allows for a comprehensive examination of how GenAI influences linguistic practices and how these influences are perceived and interpreted by users. The mixed-methods approach is particularly appropriate for studying emergent linguistic phenomena, as it enables triangulation between observed discourse patterns and reported user attitudes. This design enhances the validity of the findings by situating linguistic analysis within broader social and perceptual contexts.

4.2 Qualitative Discourse Analysis

The qualitative component of the study consists of a corpus-based discourse analysis of approximately 150 text samples collected from multiple digital platforms. These include:

1. Reddit threads in which users explicitly reference editing or relying on AI-generated text

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL

2. LinkedIn posts exhibiting stylistic and rhetorical features commonly associated with GenAI outputs
3. Collaborative writing environments where AI-assisted drafting and revision are evident

The selected texts were analyzed using critical discourse analysis (Fairclough, 1995), with particular attention to recurring lexical items, syntactic constructions, and pragmatic markers indicative of GenAI influence. Analytical focus was placed on identifying patterns of standardization, hybridization, and rhetorical regularity that suggest the emergence of synthetic dialects.

Rather than seeking to identify individual authorship, the analysis prioritized pattern recognition across texts, emphasizing how shared linguistic features circulate and stabilize across different users and platforms.

4.3 Survey of User Attitudes

To complement the discourse analysis, an online survey was conducted with 312 participants who reported varying degrees of GenAI use in their writing practices. The survey examined:

1. Frequency and context of GenAI use in everyday communication
2. Perceived influence of GenAI on personal writing style and linguistic choices
3. Self-reported ability to distinguish between human-authored and AI-influenced text
4. Attitudes toward authenticity, originality, and authorship in AI-assisted communication

Quantitative data were analyzed using descriptive statistical methods to identify general trends, while open-ended responses were subjected to thematic coding. This dual analytical approach enabled the identification of recurring attitudinal patterns and interpretive strategies employed by users when engaging with GenAI-assisted discourse.

5. Findings

5.1 Standardization of GenAI-Influenced Phrasing

Analysis of the corpus revealed that users across different platforms frequently adopted AI-generated phrasing. Common examples included:

"It is important to note that…"

1. "This highlights the significance of…"
2. "In today's rapidly evolving landscape…"

These expressions, originally suggested by GenAI tools, appeared consistently in posts and messages authored by diverse users. The repeated use of such phrases indicates a process of linguistic convergence, in which algorithmically generated patterns begin to influence broader communication norms. This aligns with sociolinguistic theory suggesting that registers stabilize through repetition and social uptake (Halliday, 1978). In this case, however, repetition is accelerated by the predictive and prescriptive nature of GenAI systems, which standardize certain linguistic forms across contexts and user groups. Survey data corroborated these observations: 72% of participants reported using AI-suggested phrases "sometimes" or "often" in their writing. Notably, respondents highlighted that these phrases were perceived as more polished, authoritative, or professional, demonstrating GenAI's role in shaping users' perception of acceptable discourse norms.

5.2 Hybrid Linguistic Forms

While AI-generated language influenced users' choices, verbatim adoption was rare. Instead, participants typically edited or adapted outputs to fit personal style, context, or audience. These adaptations created hybrid registers, where human creativity and AI optimization coexist. For example, users often combined AI-suggested sentence structures with colloquial expressions, emojis, or context-specific terminology, producing a linguistic form that is both polished and distinctively human. This finding supports previous research on human–AI collaboration, which emphasizes that iterative editing blurs authorship boundaries and fosters co-created discourse (Lee et al., 2022; Köbis & Mossink, 2021). Hybrid forms were

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

particularly evident in professional platforms such as LinkedIn, where users sought to balance the clarity and authority of AI-generated language with authentic, personalized expression.

5.3 Erosion of Human–Machine Distinction

Survey results revealed that 61% of respondents were unable to reliably distinguish between AI-assisted and fully human-authored text. This ambiguity suggests that synthetic dialects increasingly occupy a middle ground between human and machine discourse, challenging traditional notions of authorship, originality, and voice. Qualitative responses indicated nuanced attitudes: some users viewed AI influence as enhancing communication efficiency, while others expressed concern that overreliance might diminish personal style or creativity. This highlights the dual role of GenAI as both an enabler of effective discourse and a potential homogenizer of expression.

6. Discussion

The findings indicate that synthetic dialects represent a novel stage in language evolution, with three interrelated characteristics:

1. **Algorithmic norm-setting:** GenAI models standardize certain expressions and rhetorical structures, influencing perceptions of what is professional, persuasive, or "correct" across online spaces. This supports earlier theoretical arguments regarding AI's role in reinforcing dominant linguistic norms (Bender et al., 2021).
2. **Distributed agency:** Linguistic choices increasingly emerge from collaborative human–machine processes, rather than individual intention alone. Users act as co-creators, selectively adopting, editing, and integrating AI-generated forms into their discourse. This aligns with posthuman perspectives emphasizing distributed cognition and hybrid authorship (Hayles, 1999).
3. **Shifts in authenticity:** The rise of synthetic dialects challenges traditional concepts of originality. Authorship becomes less about the source of a phrase and more about its functional effectiveness in communication. Users increasingly evaluate language based on clarity, persuasiveness, and contextual fit rather than purely human generation.

These findings suggest that synthetic dialects are not mere artifacts of AI technology, but active components of a socio-linguistic ecosystem. While concerns about homogenization are valid, the study also emphasizes users' agency in shaping and negotiating hybrid registers. This mirrors research on digital linguistic micro-tribes, where niche communities adapt shared language norms to express identity while engaging with broader technological influences (Androutsopoulos, 2014; Tagg, 2015).

7. Implications

7.1 Digital Literacy

The emergence of synthetic dialects highlights the pressing need for enhanced digital literacy, extending beyond basic technological proficiency to include critical linguistic awareness. Users of AI-assisted writing tools must understand how GenAI can subtly shape discourse, influence stylistic norms, and potentially homogenize language across communities. Awareness of these influences enables users to preserve personal voice and cultural-linguistic diversity while leveraging AI for efficiency and clarity (Androutsopoulos, 2014; Tagg, 2015). Educational programs and professional development initiatives could incorporate targeted exercises designed to Identify and analyze AI-generated phrasing in texts:

1. Evaluate the ethical implications of AI-mediated authorship
2. Develop strategies for integrating AI outputs without compromising originality or authenticity

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL

By fostering these competencies, individuals can engage with AI tools responsibly, mitigating risks of linguistic homogenization while retaining agency over communicative choices.

7.2 Platform Design

The study also underscores the responsibility of platform designers in shaping user engagement with AI-assisted writing. Transparency is essential: users benefit from understanding when and how AI contributes to text generation, without being stigmatized for employing AI support. Features such as subtle AI usage indicators, editable AI suggestions, and contextual guidance can enable users to make informed choices about adopting or modifying AI-generated language. Design interventions that support creative collaboration rather than mere automation are particularly important. Such features can:

1. Promote user agency, ensuring humans remain co-creators rather than passive recipients
2. Encourage the development of hybrid registers that integrate AI-assisted clarity with individual style
3. Provide educational scaffolding, supporting users in distinguishing between AI influence and their own linguistic contribution

Thoughtful platform design therefore functions not only as a technical enhancement but also as a sociolinguistic intervention, shaping how hybrid discourse norms emerge and circulate.

7.3 Future Linguistic Research

Synthetic dialects represent a new frontier for research in sociolinguistics, discourse studies, and posthuman communication. Potential directions for future inquiry include:

1. Cross-linguistic emergence of AI-influenced registers: Investigating whether synthetic dialects manifest similarly across languages and cultural contexts, or whether local linguistic norms mediate AI influence.
2. Genre-specific hybridization: Examining differences in AI-human collaboration across professional, academic, and casual communication, to determine how context shapes the adoption and adaptation of AI-generated forms.
3. Longitudinal effects on writing and social norms: Tracking how repeated exposure to GenAI outputs affects writing development, communicative competence, and shared norms of authorship over time.

Such research will deepen our understanding of language evolution in AI-mediated environments, providing insight into both the opportunities and challenges posed by hybrid discourse ecosystems.

8. Conclusion

Generative AI is not merely a tool for writing; it functions as an active participant in the evolution of human discourse. The study demonstrates that synthetic dialects hybrid linguistic forms co-produced by humans and AI systems are increasingly present across digital platforms. These dialects contribute to the standardization of certain expressions, reshape perceptions of authorship and authenticity, and challenge conventional assumptions about originality and voice. The findings highlight a fundamental shift toward hybrid linguistic ecosystems, in which meaning, authority, and identity are collaboratively negotiated between human users and algorithmic systems. Recognizing this shift has important implications for:

1. **Digital literacy:** Equipping users with the knowledge and skills to navigate AI-mediated communication responsibly
2. **Platform design:** Creating features that foster transparency, user agency, and hybrid creativity
3. **Linguistic research:** Expanding understanding of register formation, language change, and posthuman authorship

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL

By critically engaging with the dynamics of AI-assisted communication, scholars, educators, and practitioners can better anticipate the future trajectory of human discourse in an increasingly AI-augmented digital landscape.

References

Androutsopoulos, J. (2011). Language change and digital media: A review of conceptions and evidence. *Standard Languages and Language Standards*, 145–161.

Androutsopoulos, J. (2014). Mediatization and sociolinguistic change. *Key Topics in Sociolinguistics*, 63–79.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, 610–623.

Brown, T. B., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.

Crystal, D. (2001). *Language and the Internet*. Cambridge University Press.

Fairclough, N. (1995). *Critical Discourse Analysis: The Critical Study of Language*. Longman.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., … & Schafer, B. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707.

Halliday, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. Edward Arnold.

Hayles, N. K. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press.

Herring, S. C. (2007). A faceted classification scheme for computer-mediated discourse. *Language@Internet*, 4.

Khan, H., Jameel, T., Parveen, N., & Sajjad, A. (2025, July). The sociolinguistics of "online linguistic micro-tribes": How digital communication shapes niche identity and exclusion. *Qualitative Research Journal for Social Studies, 2*(2), 542–552. https://doi.org/10.63878/qrjs81

Khan, H., Parveen, N., Qadir, A., & Rashid, S. (2025, September). Code-switching, cognitive load, and identity negotiation: A psycholinguistic study of bilingual youth in urban Pakistan. *Journal of Applied Linguistics and TESOL (JALT), 8*(3), 2533–2540. https://doi.org/10.63878/jalt1278

Köbis, N. C., & Mossink, L. (2021). Hybrid authorship in human–AI collaboration. *Journal of Human–Computer Interaction*, 37(12), 1121–1143.

Lee, M. K., et al. (2022). Human–AI collaboration in creative writing: Understanding user practices. *CHI Conference Proceedings*.

Mehinood, A., & Lawa, A. (2025). Ethical leadership as a catalyst for organizational resilience in times of crisis. *Review of Applied Management and Social Sciences, 8*(1), 407–420. https://doi.org/10.47067/ramss.vSi1.472

Mehmood, A., Siddiqui, S., & Rashid, R. (2025). Mentorship as a catalyst for leadership development: Strategies, outcomes, and implications for emerging leaders in organizational contexts. *Social Science Review Archives, 3*(1), 2516–2525. https://doi.org/10.70670/sra.v3i2.56l

Page, R., Barton, D., & Lee, D. (2014). Discourse and language change in digital media. *Journal of Sociolinguistics, 18*(2), 123–145.

Shanahan, M. (2022). Talking about large language models. *Communications of the ACM, 65*(10), 68–78.

Tagg, C. (2015). *Exploring Digital Communication: Language in Action*. Routledge.