REPRESENTATION OF GENDERED IDENTITY IN ENGLISH–UZBEK INFLUENCER DISCOURSE: AN AI-ENHANCED CORPUS STUDY

Qudratullayeva Muniraxon Abrorjon kizi

Teacher at the Department World Languages of Kokand University

Gmail: muniraqudratullayeva@gmail.com

Abstract. This article investigates the representation of gendered identity in English–Uzbek influencer discourse through an AI-enhanced corpus-based approach. Drawing on a large bilingual corpus of influencer-generated content from Instagram, YouTube, and Telegram, the study explores how masculinity and femininity are discursively constructed across cultures. Using natural language processing (NLP) tools combined with qualitative discourse analysis, the research identifies lexical patterns, evaluative strategies, and pragmatic markers associated with gendered identity. The findings demonstrate clear cross-cultural differences: English influencer discourse emphasizes individualism, self-expression, and gender fluidity, while Uzbek influencer discourse foregrounds collectivist values, moral responsibility, and culturally grounded gender norms. The study confirms the effectiveness of AI-assisted corpus methods for uncovering ideological tendencies in digital discourse and contributes to gender studies, intercultural communication, and applied linguistics.

Keywords: gendered identity; influencer discourse; AI-assisted corpus analysis; English–Uzbek comparison; digital communication.

Annotatsiya. Mazkur maqolada ingliz va o'zbek tillaridagi influencer diskursida gender identifikatsiyasining ifodalanishi sun'iy intellekt yordamida korpus asosida tadqiq etiladi. Instagram, YouTube va Telegram platformalaridan olingan keng ko'lamli ikki tilli korpus asosida erkaklik va ayollik tushunchalarining lingvistik ifodalanish usullari tahlil qilinadi. NLP texnologiyalari va sifatli diskurs tahlili natijalari shuni ko'rsatadiki, ingliz tilidagi diskursda gender moslashuvchan va individual xarakterga ega, o'zbek tilidagi diskursda esa an'anaviy qadriyatlar va ijtimoiy mas'uliyat ustuvor hisoblanadi.

Kalit so'zlar: gender identifikatsiyasi; influencer diskursi; sun'iy intellekt; ingliz–o'zbek tahlili; raqamli muloqot.

Аннотация. В статье рассматривается репрезентация гендерной идентичности в англо- и узбекоязычном дискурсе инфлюенсеров с применением корпусных методов анализа, усиленных технологиями искусственного интеллекта. Анализ цифровых текстов показывает культурно обусловленные различия в репрезентации маскулинности и фемининности.

Ключевые слова: гендерная идентичность; дискурс инфлюенсеров; искусственный интеллект; корпусный анализ; цифровая коммуникация**.**

Introduction.

Gendered identity is increasingly shaped, performed, and negotiated within digital communication spaces, where social media platforms function as powerful arenas for identity construction. Within these environments, influencers occupy a central role as cultural mediators who disseminate and normalize particular lifestyle ideals, behavioral expectations, and gendered identity frameworks. Through carefully curated content, influencers not only reflect prevailing social norms but also actively participate in redefining femininity and masculinity in ways that resonate with their audiences.

In multilingual and multicultural contexts such as English–Uzbek digital communication, influencer discourse becomes a particularly rich site for examining the interaction between globalized gender ideologies and locally grounded cultural values. Influencers operating across these linguistic spaces often engage in strategic code-switching, borrowing lexical items, discursive styles, and ideological positions from global English-language media while adapting them to Uzbek sociocultural norms. This hybrid discourse reveals ongoing negotiations between traditional gender expectations and more globalized, individualistic models of identity promoted through transnational digital culture.

Recent advances in artificial intelligence and natural language processing (NLP) have enabled large-scale, systematic analysis of such digital discourse, moving beyond small qualitative samples toward more comprehensive empirical investigations. Techniques such as sentiment analysis, topic modeling, stance detection, and gendered language analysis allow researchers to identify recurring patterns, ideological markers, and discursive strategies used by influencers when constructing gendered identities. These computational approaches make it possible to trace how specific gender norms are reinforced, challenged, or recontextualized across languages and platforms.

Literature Review.

Research in gender linguistics has long challenged essentialist views of gender, emphasizing that gender is not biologically fixed but socially and discursively constructed. Foundational scholars such as Butler conceptualize gender as performative, produced through repeated linguistic and social practices. Similarly, interactional and sociolinguistic approaches argue that gender emerges through discourse choices, interactional roles, and evaluative positioning rather than being a stable individual attribute. Language thus functions as a primary medium through which gendered identities are negotiated, legitimized, or resisted within specific sociocultural contexts.

With the rise of digital communication, scholars have increasingly turned their attention to online environments as key sites for gender construction. Digital discourse studies highlight social media platforms as spaces characterized by hybridity, multimodality, and rapid circulation of ideologies. Research suggests that digital media simultaneously reproduce traditional gender norms and enable alternative gender performances that challenge hegemonic models.

Social media influencers have been examined as influential actors in the production and dissemination of gendered meanings. Studies in predominantly English-language contexts demonstrate that influencer discourse often promotes normative ideals of femininity and masculinity linked to appearance, lifestyle, consumption, and self-optimization. At the same time, some research points to the emergence of counter-discourses that challenge traditional gender expectations, particularly in relation to body image, professional identity, and emotional expression. Influencers thus operate as cultural intermediaries who mediate between commercial interests, audience expectations, and broader gender ideologies.

Taken together, existing literature highlights the importance of discourse in constructing gender and recognizes social media influencers as key agents in this process. Nevertheless, there is a notable lack of comparative, multilingual research that integrates AI-based corpus methodologies to examine influencer discourse beyond English-dominant contexts. Addressing this gap, the present study situates itself at the intersection of gender linguistics, digital discourse analysis, and computational methods, offering a systematic investigation of gendered identity construction in English–Uzbek influencer communication.

Methodology.

A bilingual corpus of approximately 200,000 words was compiled from publicly available social media posts produced by English- and Uzbek-speaking influencers. The data were collected from verified or high-engagement influencer accounts to ensure relevance and discursive impact. Only publicly accessible content was included, in accordance with ethical standards for digital research.

The corpus was balanced across the two languages to allow for meaningful comparative analysis and consisted primarily of textual captions and accompanying comments authored by influencers.

AI-based natural language processing (NLP) tools were employed to process and analyze the corpus. The data underwent automated tokenization and normalization to ensure consistency across languages. Frequency analysis was conducted to identify salient lexical items and recurring gender-related terms, while collocation extraction was used to examine co-occurring patterns that signal evaluative stances and ideological positioning. These quantitative procedures facilitated the detection of dominant themes and discursive tendencies that would be difficult to observe through manual analysis alone.

To enhance analytical reliability, corpus-based findings were cross-checked through iterative qualitative interpretation. The integration of quantitative and qualitative methods strengthens the validity of the study by ensuring that observed patterns are both statistically salient and discursively meaningful. This methodological combination supports robust and replicable insights into gendered discourse across English–Uzbek digital communication.

Results and Discussion.

The analysis revealed that English influencer discourse frequently frames gender identity in terms of personal freedom, self-branding, and emotional openness. Uzbek influencer discourse, by contrast, emphasizes moral responsibility, social harmony, and culturally prescribed gender roles. These findings reflect broader sociocultural orientations and demonstrate the usefulness of AI-assisted corpus analysis for comparative discourse studies.

In contrast, Uzbek influencer discourse tends to construct gender identity within a framework of moral responsibility, social cohesion, and culturally prescribed gender roles. Frequency and collocation analyses reveal recurrent associations between gendered terms and values such as respect, family, modesty, and social duty. Rather than foregrounding individual self-realization, this discourse emphasizes relational identity, positioning gender as embedded within collective norms and ethical expectations. Influencers frequently invoke cultural traditions and social harmony, reinforcing established gender roles while framing them as morally grounded and socially stabilizing.

These findings demonstrate the effectiveness of AI-assisted corpus analysis in uncovering large-scale discursive patterns in multilingual influencer communication. The integration of NLP-based frequency and collocation analysis with critical discourse interpretation enables a nuanced understanding of how gender ideologies are linguistically encoded across languages and cultures. By combining computational precision with qualitative depth, the study highlights the potential of AI-driven methodologies for advancing comparative discourse studies, particularly in under-researched linguistic contexts.

Conclusion.

The study confirms that influencer discourse plays a crucial role in shaping gendered identity in digital environments. By integrating AI-enhanced corpus analysis with qualitative discourse interpretation, the research provides new insights into English–Uzbek gender representation and offers a replicable methodological framework for future studies.

By integrating AI-enhanced corpus linguistics with qualitative critical discourse interpretation, the study offers novel insights into how gendered meanings are linguistically constructed across distinct cultural and linguistic contexts. The findings reveal contrasting patterns of individualistic self-branding and emotional expressiveness in English discourse versus morally grounded, community-oriented gender roles in Uzbek discourse. These differences highlight the dynamic interaction between global digital norms and local cultural values in multilingual online communication.

Methodologically, the research advances digital discourse studies by proposing a replicable and scalable framework for analyzing large datasets of influencer content. The combined use of NLP-

based frequency and collocation analysis with interpretive discourse analysis allows for both empirical rigor and contextual sensitivity. This approach is particularly valuable for examining under-researched language pairs and contributes to expanding the methodological toolkit of gender linguistics in the age of artificial intelligence.

Overall, the study underscores the importance of interdisciplinary approaches in understanding gendered identity construction in digital media and provides a foundation for future comparative research exploring the evolving relationship between language, culture, technology, and gender.

References.

1. Abdulxayeva, M. M. qizi . (2025). DIALECTAL INTERFERENCE IN LEARNING ENGLISH AMONG UZBEK-SPEAKING STUDENTS. GOLDEN BRAIN, 3(17), 144–149. https://doi.org/10.5281/zenodo.17686082
2. Abdulxayeva, M. M. qizi ., & Botirov, H. O. o'g'li . (2025). THE ROLE OF CULTURAL BACKGROUND IN ENGLISH–UZBEK FOREIGN LANGUAGE ACQUISITION. GOLDEN BRAIN, 3(17), 137–143. https://doi.org/10.5281/zenodo.17686076
3. Abidin, C. (2016). Visibility labour: Engaging with influencers' fashion brands and advertorial campaigns on Instagram. Media International Australia, 161(1), 86–100. https://doi.org/10.1177/1329878X16665177
4. Baker, P. (2006). Using corpora in discourse analysis. London: Continuum.
5. Biber, D., Conrad, S., & Reppen, R. (1998). Corpus linguistics: Investigating language structure and use. Cambridge: Cambridge University Press.
6. Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. Discourse Studies, 7(4–5), 585–614. https://doi.org/10.1177/1461445605054407
7. Butler, J. (1990). Gender trouble: Feminism and the subversion of identity. New York: Routledge.
8. Cameron, D. (2005). Language, gender, and sexuality: Current issues and new directions. Applied Linguistics, 26(4), 482–502. https://doi.org/10.1093/applin/ami027
9. Eckert, P., & McConnell-Ginet, S. (2013). Language and gender (2nd ed.). Cambridge: Cambridge University Press.
10. Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. Online Readings in Psychology and Culture, 2(1). https://doi.org/10.9707/2307-0919.1014
11. https://doi.org/10.1080/19392397.2016.1218292
12. Kasper, G., & Rose, K. R. (2002). Pragmatic development in a second language. Oxford: Blackwell.
13. Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, 'micro-celebrity' and the rise of social media influencers. Celebrity Studies, 8(2), 191–208.
14. Marwick, A. E. (2015). Instafame: Luxury selfies in the attention economy. Public Culture, 27(1), 137–160. https://doi.org/10.1215/08992363-2798379
15. McEnery, T., & Hardie, A. (2012). Corpus linguistics: Method, theory and practice. Cambridge: Cambridge University Press.
16. Numonova Muharramxon. (2025). BARQAROR TURIZM:ATROF-MUHITNI HIMOYA QILISH VA AHOLI TURMUSHINI YAXSHILASH BO'YICHA MUVAFFAQIYATLI XALQARO TAJRIBALAR. *University Research Base*, 604–607. Retrieved from https://scholar.kokanduni.uz/index.php/rb/article/view/943
17. Numonova, M. (2025). Zoonyms in children's poetic literature: A cross linguistic ecolinguistic and axiological study in English, Uzbek, and Russian. QO'QON UNIVERSITETI XABARNOMASI, 16, 171–175. https://doi.org/10.54613/ku.v16i1275
18. Qudratullayeva, M. A. qizi . (2025). GENDER BIAS IN ENGLISH–UZBEK MACHINE TRANSLATION: A COMPARATIVE STUDY OF AI AND HUMAN OUTPUTS. GOLDEN BRAIN, 3(17), 123–130. https://doi.org/10.5281/zenodo.17685835

19. Qudratullayeva, M. A. qizi ., & Mardonov, Q. B. o'g'li . (2025). AI-BASED ANALYSIS OF GENDERED LANGUAGE PATTERNS IN ENGLISH–UZBEK SOCIAL MEDIA DISCOURSE. GOLDEN BRAIN, 3(17), 150–158. https://doi.org/10.5281/zenodo.17686094

20. Spencer-Oatey, H. (2008). Culturally speaking: Culture, communication and politeness theory (2nd ed.). London: Continuum.

21. THE IMPACT OF UZBEK CULTURAL NORMS ON ENGLISH PRAGMATIC COMPETENCE. (2025). International Journal of Artificial Intelligence, 5(12), 2848-2851. https://www.academicpublishers.org/journals/index.php/ijai/article/view/9479

22. Ziyodaxon Qodirova: 4.Qodirova, Z. (2025, October). Cross-Linguistic Representation of National Identity through Proverbs (Insights from Abdulla Qahhor and O. Henry). In International Conference on Global Trends and Innovations in Multidisciplinary Research (Vol. 1, No. 4, pp. 22-24).