# THE IMPACT

## OF GLOBALIZATION AND DIGITALIZATION ON THE AZERBAIJANI LANGUAGE

## EL IMPACTO DE LA GLOBALIZACIÓN Y LA DIGITALIZACIÓN EN LA LENGUA AZERBAIYANA

Lala Hasameddin Aliguliyeva[1*]
E-mail: lala.aliguliyeva@au.edu.az
ORCID: https://orcid.org/0000-0002-4312-8982
Aynura Rasim Khalilova[2]
E-mail: aynura.khalilova@au.edu.az
ORCID: https://orcid.org/0009-0004-3849-9158
Elzadeh Rahim Sadiqova[2]
E-mail: elzada.sadigova@au.edu.az
ORCID: https://orcid.org/0009-0007-3389-9277
Solmaz Suleyman Suleymanova[2]
E-mail : solmaz.suleymanova@au.edu.az
ORCID: https://orcid.org/0009-0003-4861-022X
Tamilla Yusif Maharramzade[2]
E-mail:  tamilla.maharramzada@au.edu.az
ORCID: https://orcid.org/0009-0009-3052-9224
[1]Baku Eurasian University. Azerbaijan.
[2]Azerbaijan University. Azerbaijan.
*Corresponding author

## ABSTRACT

The rapid advancement of digital technologies and globalization processes poses significant challenges for minority and regional languages worldwide, with the Azerbaijani language representing a critical case study in contemporary sociolinguistic transformation. This study addresses the multifaceted impact of digital communication platforms, social media integration, and global connectivity on Azerbaijani language use. It was found three critical challenges confronting modern Azerbaijani: (1) the proliferation of unregulated borrowings from English, Russian, and Turkish creates semantic overload and syntactic inconsistencies, particularly in technical and commercial domains; (2) the absence of standardized corpora and AI-powered language tools marginalizes Azerbaijani in natural language processing applications, limiting its digital infrastructure development; and (3) informal digital communication practices encourage non-standardized usage patterns that deviate from traditional grammatical structures. These findings demonstrate that unchecked globalization threatens to peripheralize Azerbaijani in professional, academic, and digital spheres, although simultaneously may create opportunities for linguistic innovation and cultural exchange. The implications of this research extend beyond descriptive linguistics to practical language policy and technological development. We advocate for strategic interventions including language-specific AI research investment, establishment of national terminology standardization councils, and development of publicly accessible intelligent language tools. We believe these measures are essential for ensuring the language's sustainable growth and cultural preservation in an increasingly digitalized/connected world, providing a framework applicable to other endangered and/or minority languages facing similar technological pressures.

Keywords: Language digitalization, Azerbaijani linguistics, Globalization impact, Minority language preservation.

## RESUMEN

## INTRODUCTION

Globalization refers to the process of economic, cultural, and communicative interdependence that expands the flow of information and goods beyond national borders. This has largely been one of the causes of the establishment of English (although also other hegemonic languages) as privileged vectors of prestige and modernity (Jacquemet, 2021). Furthermore, digitalization encompasses the transformation of discourses, practices, and resources into electronic formats, as well as the consolidation of online platforms that socialize language use in immediate and multimodal contexts (Lütge et al., 2019). In this context, digital sociolinguistics studies how new modes of interaction (instant messaging, social networks, forums) facilitate code-switching and code-mixing dynamics, incorporating elements from different languages within a single utterance and generating what is often called the "technolect" or "digital lingua franca" (Busch, 2021; Daniel et al., 2021). These forms of language contact not only introduce neologisms and acronyms—for example, terms like "applaşdırmaq" (to make applaable) or "like-lamaq" (to give a like)—but also foster morpho-syntactic adaptations that challenge traditional normative rules, opening fertile ground for the study of linguistic innovation and, simultaneously, for the analysis of semantic drift and contextual ambiguity.

The phenomenon of lexical borrowings and morpho-syntactical adaptations thus becomes an indicator of prolonged contact between languages, including Azerbaijani, and global languages such as English or Russian, while simultaneously reconfiguring the domains of use of the national language (from informal to academic) (Mammadova, 2022). These interrelations strain the duality between linguistic modernization—necessary for integration into global spaces—and the preservation of identity integrity (Jorroch & Prawdzic-Jankowska, 2024). In this sense, a terrain has been defined in which standardization policies, digitalization, and the implementation of educational tools, possibly with the inclusion of artificial intelligence, emerge as indispensable strategic responses. Thus, this conceptual framework offers a solid basis for analyzing the vectors of change in the Azerbaijani language to propose a balance between innovation and cultural legacy, ensuring its resilience in the digital age.

Thus, as is the case with other languages, the Azerbaijani language is experiencing monumental transformations in this age of digital technology. As a result of the widespread use of digital communication technology, social media, and increased global connectedness, people's understanding of language and how they use it are undergoing significant transformations. Not only has the expanding influence of digitalization brought about new vocabulary items, such as online slang, acronyms, and borrowed foreign terminology, but it has also brought about changes in communication patterns and cultural expressions within the Azerbaijani linguistic landscape. A whole new language has emerged on the Internet as a consequence of the widespread use of mobile phones and social media platforms. This language is characterized by code-switching, informality, and abbreviations. In commercial and technical contexts, speakers of Azerbaijani often make use of loanwords from English and Russian. This linguistic growth is consistent with broader global patterns such as those seen in other national languages and languages spoken by minority groups. However, despite the fact that these developments bring about globalization and modernity, they also make it more difficult to preserve the integrity of national identities and languages. A case of special attention/interest is the Azerbaijani alphabet (Figure 1).

Fig 1. Azerbaijani alphabet.

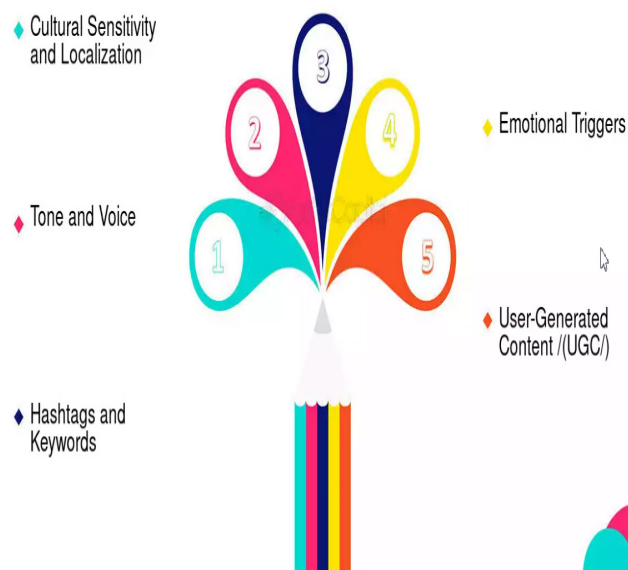| A a | Ä ä | B b | C c | Ç ç | D d | E e | F f | G g |
|---|---|---|---|---|---|---|---|---|
| a | ä | be | ce | çe | de | e | fe | ge |
| [ɑ] | [æ] | [b] | [dʒ] | [tʃ] | [d] | [e] | [f] | [ɟ] |
| Ğ ğ | H h | X x | I ı | İ i | J j | K k | Q q | L l |
| ğe | he | xe | ı | i | je | ke, ka | qe | el |
| [ɣ] | [h] | [x] | [ɯ] | [i] | [ʒ] | [c/k] | [g] | [l] |
| M m | N n | O o | Ö ö | P p | R r | S s | Ş ş | T t |
| em | en | o | ö | pe | er | se | şe | te |
| [m] | [n~ŋ] | [o] | [ø] | [p] | [r] | [s] | [ʃ] | [t] |
| U u | Ü ü | V v | Y y | Z z | | | | |
| u | ü | ve | ye | ze | | | | |
| [u] | [y] | [v~w] | [j] | [z] | | | | |

Source. Ager (2025).

As stated before, because of globalization, there is greater interaction between different cultures and languages, which results in both the enrichment and the dilution of local linguistic systems. Code-mixing and language borrowing have the potential to be important forces that contribute to the advancement and creativity of language. Unless individuals actively try to stop this tendency and apply strategic measures, the expansion of global languages, notably English, threatens to drive national languages like Azerbaijani to the edges of professional, academic, and digital contexts. This is especially true because English is the dominant language in the world, and holds a notorious power (Figure 2). Therefore, the government of Azerbaijan has made the use of local terminology, the digitization of linguistic resources, and the enhancement of educational programs a priority in order to mitigate the consequences of these factors. Institutions such as the

Nasimi Institute of Linguistics are very important when it comes to studying the sociolinguistic consequences of technology and the standardization of terminology. The importance of having a comprehensive understanding of the ways in which digital culture shapes language patterns has never been higher than it is right now.

Fig 2. The power of language.



Source. FasterCapital (2025).

With this in mind, the main objective of this research is to investigate the changing nature of the Azerbaijani language in relation to technological progression, globalization, and the rise of social media. It is addressed historical evolution of terminology, communication techniques, and word usage in order to propose proposals for the preservation of language over the long term. It is also examined current trends and makes comparisons with other countries throughout the globe in order to establish the framework for understanding how the Azerbaijani language could thrive in a society that is more globalized and digital.

## DEVELOPMENT

The modernization of communication networks throughout the globe, as well as the expanding influence of social media and technology, have had a significant impact on minority and regional languages, such as Azerbaijani. Azerbaijani, like many other low-resource languages, has challenges in producing CodeMix texts (See Figure 3) owing to the frequency of hybrid word formations, informal syntactic structures, and multilingual interference, particularly on social media platforms (Sreelakshmi et al., 2024). This is especially true on platforms that allow for the use of two or more languages.

Recent developments in artificial intelligence-based linguistic tools and linguistic modeling have resulted in the development of new techniques that may assist with the preservation of language and the standardization of its structure. In this regard, Chen et al. (2014) developed an automated method with the purpose of highlighting the significance of reference aids in enhancing the productive language use of English as a Foreign Language (EFL) learners. To assist younger digital natives in acquiring traditional vocabulary and grammar, it may be beneficial to repurpose similar tools for the Azerbaijani language.

Fig 3. CodeMix example.
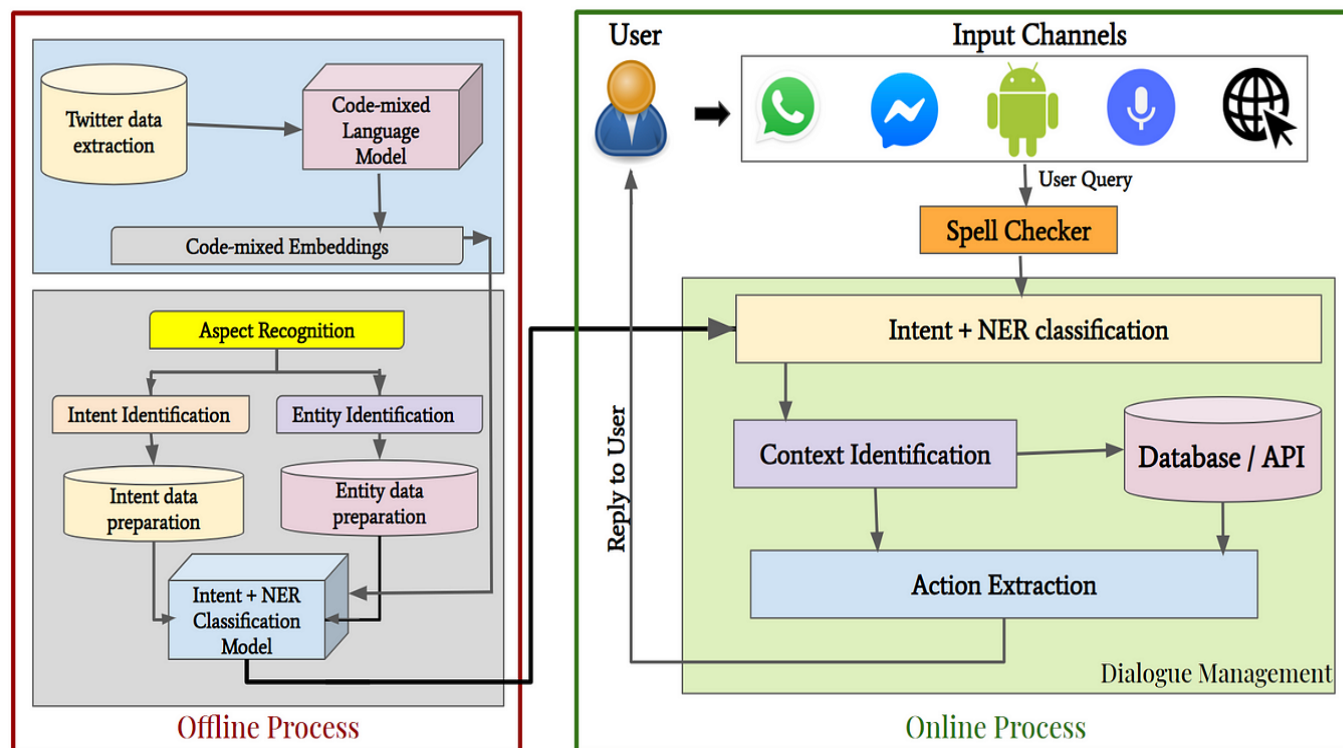


Source : Chatterjee (2022).

Through the application of GPT-based models to structured data visualizations (Maddigan & Susnjak, 2023) and multilingual programming frameworks (Katzy, 2024), it may be possible to get a deeper understanding of the potential complications and solutions that artificial intelligence may provide in multilingual environments. For example, Arnau-González et al. (2023) state that intelligent tutoring systems for arithmetic problems have the potential to increase students' ability to solve problems while they are in the classroom.  In light of this, it seems that there are equivalent options for acquiring a language and correcting terminology (Maimaiti et al., 2022).

Research on neural machine translation has brought to light the fact that lexical ambiguities and a dearth of corpus-based resources are factors that contribute to the escalation of language challenges in groups with limited resources. For the purpose of strengthening Azerbaijani-language artificial intelligence systems, tools that integrate domain-specific lexicons hold promise for ensuring terminological accuracy even in fields that are constantly evolving, such as law, medicine, or information and communication technology. Considering the above, the following are some of the most pressing issues that contemporary Azerbaijanis are facing.

Globalization and the advent of digital communication are directly responsible for the fact that there is a proliferation of borrowed terms and phrases on the Internet. The presence of such a wide range of languages makes it more difficult to develop efficient Azerbaijani search engines and learning platforms (Jeong et al., 2024) for educational purposes, just as it is in other language ecosystems. The Azerbaijani language continues to be one of the least used languages in the fields of artificial intelligence and natural language processing, leading to its low ranking. When compared to languages that have better resources, it is more challenging to create AI-powered learning systems for languages that do not have standardized corpora and lexicons (Maimaiti et al., 2022).

There is a lack of artificial intelligence in the classroom. Although AI has the potential to enhance tutoring and grammar correction (Arnau-González et al., 2023), there are presently no solutions that are suited for Azerbaijani. Despite ongoing

efforts, the local educational infrastructure continues to struggle with the challenge of incorporating such technologies. Because of the informal character of Azerbaijani's online presence, it tends to encourage non-standardized usage, especially on websites that neglect appropriate spelling, grammar, and code-switching. Sreelakshmi et al. (2024) discovered that multilingual contexts, particularly those including Dravidian languages, exhibited comparable trends.

With the rise of the Internet, the Azerbaijani language and its use have seen a sea change. Platforms such as social networking, instant messaging apps, and content production websites have introduced significant language variances. Some examples of these differences include the use of informal abbreviations and transliterations and the heavy use of codes from English, Russian, and Turkish. According to Sreelakshmi et al. (2024), these tendencies mirror the language mixing that occurs in CodeMix environments. These scholars bring attention to the risks of contextual ambiguity and semantic drift in digital communication systems that support many languages.

A result of unchecked language that confuses both native speakers and learners is the rapid influx of tech-related jargon, which is often and directly borrowed from English or Russian. This is significant for the development of modern Azerbaijani technolect, according to Katzy (2024), who argues that unregulated language situations in multilingual environments often lead to semantic overload and incorrect syntax. In the field of education, the Azerbaijani people face significant obstacles. There is a lack of comprehensive online learning platforms that include resources like corpora, automatic feedback systems, and AI-powered grammar checkers.

These technologies are very critical in today's educational system, especially for languages that do not have enough resources. Chen et al. (2014) demonstrated the efficacy of intelligent formulaic expression aid for English language learners. As a concept, this is rather new to Azerbaijani speakers. Arnau-González et al. (2023) introduced intelligent tutoring systems for structured learning, which is insult to injury. In order to make these techniques work for Azerbaijani, we may change them so that words are easier to remember and utilize consistently. However, machine translation, voice recognition, and intelligent assistants are all negatively impacted by the persistent underrepresentation of Azerbaijani in the training data of large language models (LLMs). Improving search operations in pre-trained models and effectively managing data diversity are crucial for better support of low-resource languages (Jeong et al., 2024).

Similarly, Maimaiti et al. (2022) suggested that pre-trained lexicon embeddings might improve machine translation transfer learning in low-resource scenarios. As far as Azerbaijani NLP is concerned, this is a promising development. In addition, Maddigan and Susnjak (2023) introduced Chat2VIS, a language-based data visualization tool, which is at the forefront of human-artificial intelligence interaction. This technology shows great promise as a tool for Azerbaijani language training and research.

Finally, it is worth emphasizing that the tension between modernization and identity preservation is not unique to Azerbaijani, but rather reflects a global challenge for minority and emerging languages: the challenge is to design adaptive models that allow for the progressive incorporation of lexical innovations and technological regulations without sacrificing the internal cohesion of the linguistic system or its symbolic and cultural value. In our opinion, this balance will be achieved through a holistic approach that combines: (1) academic research in digital sociolinguistics and NLP; (2) the creation of inclusive digital infrastructures; (3) flexible yet firm regulations in the field of terminology; and (4) ongoing training for speakers and language professionals. Only in this way can Azerbaijani establish itself as a vibrant language, capable of engaging on an equal footing in the global arena while simultaneously preserving its historical richness and central role in national identity.

## CONCLUSIONS

There is a convergence of issues in the fields of linguistics, technology, and education that are having an impact on the Azerbaijani language. There is a dearth of variety in the datasets that are used to train natural language processing (NLS), there is a rising discrepancy in the availability of standardized current vocabulary, and there is a paucity of educational materials that are driven by artificial intelligence. To begin, the language is not as helpful in the business or in the classroom since there is not enough digital infrastructure for things like intelligent conversation systems, predictive typing, and grammatical correction software. Both of these things are essential for effective communication. The development of natural language processing tools for low-resource languages is not only feasible but also essential, as Sreelakshmi et al. (2024) and Maimaiti et al. (2022) have pointed out. This is necessary in order for these languages to continue to exist and adapt to the modern day.

The second issue is that unrestrained lexical borrowing is causing a complex network of terminology to emerge, which is making it more difficult to use language. Compounding the issue is the absence of a regulatory

organization that has the capacity to update and regulate the Azerbaijani language. Because of those authors like Chen et al. (2014) and Katzy (2024) both emphasize the need of systematic vocabulary increase as a means of preserving language in modern contexts. In addition, the absence of digital dictionaries, corpus databases, interactive applications, and intelligent tutoring systems that are available in the Azerbaijani language makes the educational limits even more severe. In contrast to other languages, which have powerful learning environments assisted by artificial intelligence, those who study Azerbaijani continue to employ older methods that are based on print. One of the last repercussions of the global neglect of artificial intelligence is that Azerbaijani citizens are unable to take part in the growing ecosystem of communication platforms driven by AI. A consequence of this digital marginalization that may result in a generational gap is the possibility that younger users would renounce the use of their native language in favor of English or Russian for reasons related to their professional and technical responsibilities.

In order to address these issues, a number of long-term strategies are required:

1. Allocating funds to research on artificial intelligence and natural language processing that is particular to a language.

2. The establishment of a national council with the purpose of standardizing speech in the fields of science, technology, and social media.

3. In the process of developing language visualization systems, chatbots, and intelligent dictionaries that are publicly accessible and powered by artificial intelligence.

4. Azerbaijani will be added to databases all over the globe in order to guarantee that it will be included into AI systems of the future generation.

We believe that the future of the Azerbaijani language in the digital world is contingent on a collaborative effort on the part of linguists, technologists, and politicians working together inside the same direction.

## REFERENCES

Ager, S. (2025). *Azerbaijani language, alphabets and pronunciation*. Omniglot. https://www.omniglot.com/writing/azeri.htm

Arnau-González, P., Serrano-Mamolar, A., Katsigiannis, S., Althobaiti, T., & Arevalillo-Herráez, M. (2023). Toward Automatic Tutoring of Math Word Problems in Intelligent Tutoring Systems. *IEEE Access*, *11*, 67030–67039. https://doi.org/10.1109/ACCESS.2023.3290478

Busch, F. (2021). Enregistered spellings in interaction: Social indexicality in digital written communication. *Zeitschrift Für Sprachwissenschaft*, *40*(3), 297–323. https://doi.org/10.1515/zfs-2021-2033

Chatterjee, A. (2022, February 25). A Tale of Two Languages: The Code-Mixing story. *Medium*. https://wiprotechblogs.medium.com/a-tale-of-two-languages-the-code-switch-mix-story-d88e105b07f4

Chen, M.-H., Huang, C.-C., Huang, S.-T., Chang, J. S., & Liou, H.-C. (2014). An Automatic Reference Aid for Improving EFL Learners' Formulaic Expressions in Productive Language Use. *IEEE Transactions on Learning Technologies*, *7*(1), 57–68. https://doi.org/10.1109/TLT.2013.34

Daniel, M., Chechuro, I., Verhees, S., & Dobrushina, N. (2021). Lingua francas as lexical donors: Evidence from Daghestan. *Language*, *97*(3), 520–560.

FasterCapital. (2025). *The power of language in social media campaigns*. https://fastercapital.com/topics/the-power-of-language-in-social-media-campaigns.html

Jacquemet, M. (2021). Transidioma and Territory: Language and Power in the Age of Globalization. *WORD*, *67*(1), 117–136. https://doi.org/10.1080/00437956.2021.1882651

Jeong, Y., Yang, J., Choi, I. H., & Lee, J. (2024). Feature-Based Text Search Engine Mitigating Data Diversity Problem Using Pre-Trained Large Language Model for Fast Deployment Services. *IEEE Access*, *12*, 48145–48157. https://doi.org/10.1109/ACCESS.2024.3373470

Jorroch, A., & Prawdzic-Jankowska, I. (2024). Corpus Linguistics meets Contact Linguistics and Sociolinguistics: Multimodal German–Polish digital linguistic corpus. *Corpora*, *19*(3), 375–394. https://doi.org/10.3366/cor.2024.0319

Katzy, J. (2024). Programming Language Models in Multilingual Settings. *Proceedings of the 2024 IEEE/ACM 46th International Conference on Software Engineering: Companion Proceedings*, 204–206. https://doi.org/10.1145/3639478.3639787

Lütge, C., Merse, T., Owczarek, C., & Stannard, M. (2019). Crossovers: Digitalization and literature in foreign language education. *Studies in Second Language Learning and Teaching*, *9*(3), Article 3. https://doi.org/10.14746/ssllt.2019.9.3.5

Maddigan, P., & Susnjak, T. (2023). Chat2VIS: Generating Data Visualizations via Natural Language Using ChatGPT, Codex and GPT-3 Large Language Models. *IEEE Access*, *11*, 45181–45193. https://doi.org/10.1109/ACCESS.2023.3274199

Maimaiti, M., Liu, Y., Luan, H., & Sun, M. (2022). Enriching the transfer learning with pre-trained lexicon embedding for low-resource neural machine translation. *Tsinghua Science and Technology*, *27*(1), 150–163. https://doi.org/10.26599/TST.2020.9010029

Mammadova, F. (2022). Analysis of the Borrowing Methods of Foreign Words in Modern Russian and Azerbaijani Languages. *Ad Alta-Journal of Interdisciplinary Research*, *12*(2), 57–60.

Sreelakshmi, K., Premjith, B., Chakravarthi, B. R., & Soman, K. P. (2024). Detection of Hate Speech and Offensive Language CodeMix Text in Dravidian Languages Using Cost-Sensitive Learning Approach. *IEEE Access*, *12*, 20064–20090. https://doi.org/10.1109/ACCESS.2024.3358811