

A Morphological and Syntactic Study of Neologisms and Structures in Arabic Social Media

(Twitter, Facebook, Instagram as Case Studies)

Soudabeh mozaafari, Associate Professor in Arabic Language and Literature,
Faculty of Literature and Humanities, Kharazmi University, Tehran, Iran.

Ramin sedaghat, (corresponding author): master student of Arabic language and literature, Faculty of Literature and Humanities, Kharazmi University, Tehran, Iran.
Email: ramin.sedaghat@khu.ac.ir.

Introduction

The digital revolution has created a hybrid linguistic environment on social media, where Modern Standard Arabic, regional dialects, and foreign languages converge. This study adopts a descriptive perspective to examine systematic morphological and syntactic transformations in Arabic digital discourse, framing these changes as adaptive responses to new communicative constraints rather than as random errors. It also explores the challenges these hybrid forms pose to Arabic Natural Language Processing (NLP).

Methodology

The research employs a mixed descriptive-analytical and quantitative design. A purposive sample of 500 public posts and comments was collected from X (50%), Facebook (30%), and Instagram (20%) during 2025. After eliminating repetitions, 420 unique linguistic phenomena were identified and analyzed using qualitative interpretational alongside frequency-based quantitative measures.

Results and Discussion

Four major morphological patterns were identified: (1) *Arabizi*—the use of Latin letters and numerals for Arabic sounds—appeared in 42% of the sample; (2) morphological simplification, notably the unified colloquial relative pronoun "*elli*", replaced all standard gender- and number-distinctive forms in approximately 85% of relevant clauses; (3) hybrid derivation, where foreign roots were integrated into common Arabic verb templates, occurred in 30% of posts, with two specific patterns accounting for over 90% of these cases; and (4) blending and abbreviation, driven by linguistic economy, were especially frequent on X, where 25% of posts contained shortened expressions. Syntactically, a clear preference for Subject-Verb-Object (SVO) word order was observed, appearing in 65% of declarative sentences, alongside widespread use of colloquial substitutes for standard negation, questioning, and linking particles.

These transformations are systematic and motivated primarily by linguistic economy, communicative efficiency, platform constraints, and the expression of digital identity. However, the growing hybridity significantly challenges Arabic NLP systems, as many forms fall outside traditional linguistic resources.

Conclusion

he study strongly recommends the development of annotated hybrid corpora capable of capturing the full spectrum of contemporary digital Arabic usage, in order to support more robust NLP applications and better reflect the linguistic reality of Arabic-speaking online communities.

References

- Abdelali, A., Darwish, K., Durrani, N., & Mubarak, H. (2016). Farasa: A new fast and accurate Arabic tokenizer. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 219–223. [in English]
- Abdul-Raof, T. A. (2021). The use of Arabic neologisms in social media applications. *International Journal of Arabic-English Studies*, 21(1), 1–20. [in English]
- Al-Ajmi, S. (2014). Al-Arabitini: Arabic writing in Latin letters. In *The Arabic Youth Language in Modern Communication Media*, 7–26. Riyadh: King Abdullah International Center for Arabic Language Service. [in Arabic]
- Algahtani, M., & Albawardi, A. (2025). Levels of Arabic language in Saudi tweets. *Academic Journal of Interdisciplinary Studies*, 14(2). [in English]
- Al-Jarf, R. (2021). The impact of social media on the depletion of the Arabic language. ERIC, ED615197. [in Arabic]
- Alkaabi, H. A., Jasim, A. K., & Darroudi, A. (2025). Arabic NLP: A survey of pre-processing and representation techniques. *Journal of Computer Science and Information Technology*, 14(1). [in English]
- Almaoui, P., Bouillon, P., & Hengchen, S. (2025). Arabizi vs. LLMs: Can the Genie understand the language of Aladdin? In *Proceedings of Machine Translation Summit XX*, 28–41. Geneva, Switzerland. [in English]
- Al-Oufi, A. A. (2014). Arabizi (hybrid language) on social networking sites and its impact on Modern Standard Arabic. Riyadh: King Abdullah International Center for Arabic Language Service. [in Arabic]
- AlQadi, L., Alzubaidi, A., & Hacid, H. (2026). QIMMA: A quality-first approach to Arabic LLM evaluation. In *Proceedings of COLING 2026*. [in English]
- Al-Sayadi, S. (2024). Arabizi in Saudi youth online communication: Identity and frequency. *Journal of Saudi Linguistics*, 12(2), 45–67. [in English]

- Androutsopoulos, J. (2014). Languageing when online: Language and social media. In *The Routledge Handbook of Language and Digital Communication*, 231–245. Routledge. [in English]
- Benadla, D. (2024). Code choice in instant messaging among Algerian youth: A case study of Messenger. *Algerian Journal of Sociolinguistics*, 7(1), 22–41. [in English]
- Boudiar, A. (2018). Fus'ha in schools and colloquial Arabic on social networking sites. *Midad Al-Adab Journal*, Special Conference Issue, 648–655. [in Arabic]
- Bougrine, Z., et al. (2025). Benchmarking sentiment analysis of Algerian Arabic dialect on X (Twitter). *Journal of King Saud University – Computer and Information Sciences*, 37(5). [in English]
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5th ed. SAGE Publications. [in English]
- Darwish, A. H. (2019). The use of bilingualism, diglossia, and Latin script by social media users and their effects on the Arabic language. *Journal of Media Research*, Al-Azhar University, 51, 533–580. [in Arabic]
- Ferguson, C. A. (1959). Diglossia. *Word*, 15(2), 325–340. [in English]
- Fitrianto, I., & Hamid, R. (2024). Morphosemantic changes in the Arabic language in the social media era: A study of neologisms and their impact on youth communication. *International Journal of Arabic Studies*, 1(1). [in English]
- Gugliotta, E. (2025). *Tunisian Arabizi: Linguistic Analyses and Corpus Building Using Natural Language Processing*. DOAB. [in English]
- Kassas, S. (2019). Language use on social media among Arab youth: Reality, causes, and effects. *Journal of Issues in Language and Literature*, 8(3), 462–478. [in Arabic]
- Khaled, N., & Al-Mubarak, A. (2025). Arabic natural language processing (NLP): A comprehensive review of challenges, techniques, and emerging trends. *Computers*, 14(11), 497. [in English]
- Krippendorff, K. (2018). *Content Analysis: An Introduction to Its Methodology*. 4th ed. SAGE Publications. [in English]
- Neuman, W. L. (2014). *Social Research Methods: Qualitative and Quantitative Approaches*. 7th ed. Pearson. [in English]
- Nwesri, A. F. A., Shinbir, N. A. S., & Sharif, A. B. (2025). Sentiment analysis on Arabic dialects: A multi-dialect benchmark. In *Proceedings of the Shared Task on Sentiment Analysis for Arabic Dialects (RANLP 2025)*, 86–91. INCOMA Ltd. [in English]