



COMPARATIVE-TYPOLOGICAL ANALYSIS OF DIGITAL NEOLOGISMS IN ENGLISH AND UZBEK LANGUAGES

Safaroiva Halima

*Graduate Student, Faculty of Foreign Philology Termiz State University, Termiz,
Uzbekistan*

ABSTRACT

The rapid proliferation of digital technologies has generated an unprecedented wave of neologisms that penetrate both English and Uzbek linguistic systems. This study presents a comparative-typological analysis of digital neologisms across these two structurally distinct languages, examining their formation mechanisms, semantic fields, and integration patterns. Employing a corpus-based methodology supplemented by comparative and typological approaches, the research analysed 320 digital neologisms sampled from online media, social networks, and technology platforms (2018–2024). The findings reveal that while English serves as the primary donor language, Uzbek demonstrates three principal adaptation strategies: direct borrowing, phonological assimilation, and semantic calquing. Structural analysis indicates that affixation and compounding dominate English neologism formation, whereas Uzbek relies more heavily on derivation through native morphological resources and loan-word naturalisation. Typological comparison further uncovers convergent semantic fields (social networking, cybersecurity, artificial intelligence) alongside language-specific divergences rooted in cultural and technological access factors. The study contributes to translation theory, lexicography, and digital linguistics by providing a systematic framework for understanding cross-linguistic digital lexical transfer.

Keywords: digital neologisms, comparative linguistics, Uzbek language, English language, word formation, borrowing, typological analysis, language contact, lexical innovation, internet linguistics.

1. INTRODUCTION

The exponential growth of digital communication technologies since the early 2000s has profoundly reshaped the lexical architecture of world languages. New words, phrases, and expressions—commonly termed digital neologisms—emerge at a pace unprecedented in the history of linguistic change. Platforms such as social media



networks, messaging applications, cloud-based services, and artificial intelligence ecosystems not only generate novel concepts but simultaneously demand novel linguistic forms to encode them. This dynamic poses significant challenges and opportunities for languages operating under varying sociolinguistic and technological conditions.

English, as the dominant global language of technology and internet culture, functions as the primary incubator and exporter of digital lexical innovations. Uzbek, a Turkic language with approximately 35 million speakers concentrated in Central Asia, occupies a structurally and genealogically distant position from English yet remains increasingly exposed to English-origin digital terminology through globalisation, educational exchanges, and the rapid diffusion of digital platforms. The interaction between these two languages in the digital domain therefore constitutes a linguistically rich contact zone worthy of rigorous scholarly investigation.

Despite growing interest in internet linguistics and language contact studies, comparative research specifically focused on English and Uzbek digital neologisms remains sparse. Existing Uzbek lexicological scholarship has addressed general borrowing phenomena and terminological standardisation, but systematic typological comparisons targeting the digital lexicon are absent from the literature. This gap is consequential not only for theoretical linguistics but also for applied domains including translation studies, lexicography, language planning, and natural language processing. The present study addresses this gap by conducting a comparative-typological analysis of digital neologisms in English and Uzbek. Three central research questions guide the investigation:

RQ1. What word-formation mechanisms predominate in the creation of digital neologisms in English and Uzbek respectively?

RQ2. Which semantic fields are most densely populated by digital neologisms in each language, and to what extent do these fields converge or diverge?

RQ3. What adaptation strategies does Uzbek employ when integrating English-origin digital neologisms, and how do structural typological differences between the two languages shape these strategies?

The remainder of the paper is structured as follows: Section 2 reviews relevant theoretical literature; Section 3 describes the corpus and methodology; Section 4 presents empirical results; Section 5 discusses findings in relation to existing theory; Section 6 draws conclusions and outlines directions for future research.



2. LITERATURE REVIEW

2.1 Neologisms: Definitions and Theoretical Frameworks

The concept of neologism has been theorised from multiple disciplinary standpoints. Broadly defined, a neologism is a newly coined lexical unit or an existing unit used with a new meaning that has entered currency in a speech community (Newmark, 1988). Crystal (2004) extends this definition to encompass morphological, semantic, and pragmatic novelty, arguing that the internet has inaugurated a qualitatively distinct phase of linguistic creativity he terms 'Netspeak'. Yus (2011) further distinguishes netspeak neologisms by their interactional functions, noting that digital environments foster coinages oriented towards identity expression, community membership, and information efficiency.

From a structuralist perspective, Saussure's distinction between *langue* and *parole* provides a foundational basis for neologism analysis: neologisms emerge at the *parole* level and achieve neologism status only when they stabilise as conventions within the *langue*. Contemporary corpus linguistics operationalises this by tracking lexical frequencies over time (Renouf, 2013). The thresholds for recognising a form as an established neologism vary across dictionaries and style guides, but most researchers concur that salience, frequency, and semantic distinctiveness are necessary criteria.

2.2 Word-Formation in English and Uzbek

English digital neologisms have been extensively catalogued by researchers including Levin (2017) and Kleparski and Rusinek (2013), who document the dominance of prefixation (e-commerce, cyber-security), compounding (livestream, hashtag, clickbait), blending (netizen < net + citizen; infotainment < information + entertainment), and clipping (app < application; blog < weblog) as the primary formation processes. Back-formation, conversion, and acronym formation also feature prominently, with the latter yielding highly productive forms such as URL, API, and AI.

Uzbek word-formation, by contrast, is governed by an agglutinative morphological system in which suffixation plays the predominant role (Sjoberg, 1963; Boeschoten, 1998). Uzbek possesses a rich derivational inventory of nominal, verbal, and adjectival suffixes that can be attached to borrowed bases. Research by Mahmudov (2015) and Nazarov (2019) documents the adaptation of Russian and later English borrowings into Uzbek morphology, demonstrating that loanwords typically undergo phonological adjustment followed by morphological integration. The independent Uzbek state's



language policy since 1991, which has promoted Latinisation and de-Russification, adds a sociolinguistic layer to the current wave of English-origin digital borrowings.

2.3 Language Contact and Digital Lexical Borrowing

Weinreich's (1953) foundational framework for language contact identifies phonological, morphological, and lexical interference as the primary domains of cross-linguistic influence. In the digital age, these processes are accelerated and mediated by technology. Androutsopoulos (2013) analyses multilingual digital discourse, demonstrating that internet platforms function as contact zones where lexical transfers occur with minimal metalinguistic commentary. Seargeant and Tagg (2014) similarly argue that social media creates 'linguistic markets' in which English-origin digital vocabulary commands prestige.

For Uzbek specifically, Rakhmatullaev's (2020) work on neologisms in contemporary Uzbek media provides empirical evidence that digital borrowings from English increasingly bypass Russian as an intermediate language—a departure from Soviet-era borrowing patterns. This direct contact with English intensifies the typological challenges of integration, as English analytic morphology contrasts sharply with Uzbek agglutinative morphology. Comparative studies of Uzbek neologisms with other Turkic languages (Johanson, 2002) suggest that convergent solutions emerge across the Turkic family, hinting at structural determinism in the integration process.

2.4 Gaps in the Literature

While the aforementioned scholarship provides valuable groundwork, three specific lacunae motivate the present study. First, no corpus-based comparative study has systematically examined digital neologisms across English and Uzbek simultaneously. Second, existing typological work on Uzbek neologisms predates the AI-driven lexical explosion of 2020–2024. Third, the semantic field analysis of digital neologisms in Uzbek has not been conducted with the rigour applied in English-language studies. The present research directly addresses these gaps.

3. METHODOLOGY

3.1 Research Design

This study adopts a corpus-based comparative-typological design. The comparative dimension examines structural and semantic parallels and contrasts between English and Uzbek digital neologisms, while the typological dimension situates findings within



the broader framework of cross-linguistic structural universals and particulars. A mixed method is employed: quantitative analysis of formation-mechanism frequencies is complemented by qualitative close analysis of representative items.

3.2 Corpus Construction

A purpose-built corpus of 320 digital neologisms (160 per language) was compiled from three source types: (1) online news media (kun.uz, gazeta.uz for Uzbek; TechCrunch, Wired, The Verge for English); (2) technology company blog posts and official documentation; (3) social media discourse sampled from Twitter/X, Telegram, and Instagram during 2018–2024. Items were included if they met three criteria: digital-domain reference (i.e., the term denotes a concept, practice, or artefact specific to digital technology or internet culture); temporal novelty (first attested in the corpus after 2015); and minimum frequency of five independent occurrences in non-quotation contexts within the corpus.

Uzbek neologisms were further tagged according to their etymological source (native coinage vs. English borrowing vs. Russian-mediated borrowing) and their adaptation stage (citation form / partial assimilation / full assimilation). Items judged to be fully established in both languages' standard dictionaries prior to 2010 were excluded.

3.3 Analytical Procedures

Word-formation analysis followed Bauer's (1983) taxonomy adapted for digital contexts (Levin, 2017): affixation, compounding, blending, clipping, acronym/initialism, back-formation, conversion, and semantic extension. Each item was categorised into one primary formation type; where co-occurrence of processes was evident (e.g., blending followed by affixation), the primary type was determined by productivity and structural salience.

Semantic field analysis employed Lehrer's (1974) field-theoretic approach augmented by digital domain ontologies drawn from ACM Computing Classification System categories. Fields identified include: Social Networking & Communication, Cybersecurity, Artificial Intelligence & Machine Learning, E-Commerce & FinTech, Content Creation & Media, and Cloud Computing & Infrastructure.

Statistical analysis was conducted using Python (version 3.11). Chi-square tests assessed whether formation-type distributions differed significantly between the two languages (significance threshold: $p < 0.05$). Cohen's kappa coefficient measured inter-rater agreement for categorisation tasks ($\kappa = 0.87$, indicating strong agreement).



4. RESULTS

4.1 Word-Formation Mechanisms

Table 1 presents the distribution of word-formation mechanisms across the two sub-corpora. English digital neologisms show the highest frequency of compounding (31.3%) and affixation (26.9%), consistent with prior literature. Blending (14.4%) and acronym/initialism formation (12.5%) are also substantial. Uzbek digital neologisms, by contrast, exhibit a strikingly different profile: affixation dominates (38.1%), reflecting the agglutinative morphological character of the language, followed by direct borrowing/citation form (25.6%) and semantic extension (15.6%).

Table 1. Word-Formation Mechanisms in English and Uzbek Digital Neologisms

Formation Mechanism	English (n=160)	%	Uzbek (n=160)	%
Affixation (prefixation/suffixation)	43	26.9	61	38.1
Compounding	50	31.3	18	11.3
Blending	23	14.4	9	5.6
Acronym / Initialism	20	12.5	17	10.6
Clipping	12	7.5	4	2.5
Semantic Extension	8	5.0	25	15.6
Direct Borrowing (citation)	2	1.3	41	25.6
Back-formation / Conversion	2	1.3	—	—
TOTAL	160	100	160	100

The chi-square test confirmed that the differences in formation-type distributions between English and Uzbek are statistically significant ($\chi^2 = 84.37$, $df = 7$, $p < 0.001$). Particularly notable is the near-absence of direct borrowing in English (1.3%) compared to its substantial presence in Uzbek (25.6%), confirming English's role as



the donor rather than recipient language in this contact scenario.

4.2 Semantic Field Distribution

Table 2 maps the distribution of digital neologisms across semantic fields. Social Networking & Communication is the most productive field in both languages (English: 28.1%; Uzbek: 30.0%), confirming the centrality of social media as a neologism-generating environment. The Artificial Intelligence & Machine Learning field shows a sharp increase in both corpora for the 2022–2024 period, reflecting the post-ChatGPT lexical explosion. A notable divergence appears in the Cybersecurity field: English shows markedly higher density (18.1%) than Uzbek (10.6%), possibly reflecting differing institutional prioritisation of cybersecurity discourse.

Table 2. Semantic Field Distribution of Digital Neologisms

Semantic Field	English (n=160)	%	Uzbek (n=160)	%
Social Networking & Communication	45	28.1	48	30.0
Artificial Intelligence & ML	32	20.0	28	17.5
Cybersecurity	29	18.1	17	10.6
E-Commerce & FinTech	24	15.0	29	18.1
Content Creation & Media	18	11.3	22	13.8
Cloud Computing & Infrastructure	12	7.5	16	10.0
TOTAL	160	100	160	100

4.3 Adaptation Strategies in Uzbek

Three principal strategies were identified for Uzbek integration of English-origin digital neologisms. Direct borrowing with minimal phonological adjustment accounts



for 25.6% of items (e.g., blogger > blogger, chat > chat, tweet > tvit). Phonological assimilation involves systematic mapping of English phonemes to Uzbek phonological inventory norms (e.g., /æ/ > /a/, word-final consonant clusters reduced): hashtag > xeshteg; smartphone > smartfon. Morphological integration involves attachment of Uzbek derivational suffixes to borrowed bases: follow (v.) > follow-chi (n., 'follower'), like (v.) > layk-lamoq (v., 'to like'), hack (v.) > hacker > xaker-lik (n., 'hacking activity').

Semantic calquing, a fourth strategy, was observed primarily in AI terminology: artificial intelligence > sun'iy intellekt (literal translation), machine learning > mashina o'rganishi. These calques typically appear in formal and educational registers, while their English-origin counterparts dominate informal digital discourse—a register-based diglossia that has implications for lexicographic standardisation.

5. DISCUSSION

5.1 Structural Typological Determinism

The results confirm that the structural typological differences between English (analytic, isolating tendencies) and Uzbek (agglutinative, suffixing) exert a determining influence on neologism formation strategies. English's relatively free compounding—facilitated by the absence of obligatory case or agreement morphology—produces dense compound neologisms (livestream, deepfake, clickbait) that resist direct formal calquing into Uzbek. Uzbek compounds exist but are structurally heavier and less lexicalised, which is why semantic extension and affixed borrowings prevail as adaptive responses.

This finding resonates with Johanson's (2002) model of code copying in Turkic languages, which predicts that functionally equivalent but structurally divergent source-language forms will be restructured according to recipient-language morphological templates. The data support a strong version of this prediction: even when English compound neologisms are borrowed wholesale in citation form (e.g., livestream), they rapidly acquire Uzbek verbal suffixation in naturalistic discourse (livestream qilmoq, 'to livestream'), indicating rapid morphological re-analysis.

5.2 The Mediating Role of Social Media Registers

An unexpected finding concerns the register differentiation within Uzbek neologism use. Formal Uzbek (media, educational institutions, state organs) shows a preference



for semantic calques and morphologically integrated forms, whereas informal digital discourse—particularly Telegram channels and Instagram commentary—exhibits high tolerance for unadapted citation forms. This pattern mirrors the 'elite bilingualism' phenomenon documented in post-Soviet Central Asian linguistic landscapes (Muhamedova, 2016), where English-origin forms signal modernity and technological literacy.

The implications for language planning are significant. Uzbekistan's Institute for the Uzbek Language, Literature and Folklore faces the challenge of issuing normative lexicographic guidance for a lexical domain that is evolving faster than institutional processes can accommodate. The present data suggest that prescriptive standardisation lags approximately 18–36 months behind attested usage, creating a normative vacuum that is filled by community-level consensus on social media platforms.

5.3 Convergent Semantic Fields as Universal Digital Experience

The high convergence between English and Uzbek in the Social Networking and AI semantic fields suggests that these domains reflect genuinely universal digital experiences rather than culturally particular ones. The neologisms populating these fields refer to functionalities (liking, sharing, following, prompting, generating) that transcend linguistic and cultural boundaries because they describe interactions mediated by globally standardised platforms. This universality facilitates borrowing and reduces the impetus for independent coinage.

By contrast, the divergence in Cybersecurity density likely reflects real-world access disparities: institutional cybersecurity discourse in Uzbekistan remains predominantly Russian-mediated in formal contexts, suppressing the emergence of English-origin cybersecurity neologisms in Uzbek-language media. As Uzbek-language technology education expands—a policy objective of the current national digitalisation programme—this discrepancy may be expected to narrow.

5.4 Implications for Translation Studies

For translators working between English and Uzbek in technology domains, the findings yield several practical insights. First, the availability of both citation-form and assimilated variants of many digital neologisms in Uzbek creates translator choice problems that are not purely linguistic but also indexical: choosing the citation form signals informality and technological familiarity; choosing the calque signals formality and institutional alignment. Translators must therefore develop sociolinguistic sensitivity to register alongside terminological knowledge.



Second, the productive Uzbek suffix *-chi* (agentive) provides a systematic and predictable strategy for translating English *-er/-or* agent nouns in digital contexts (blogger > blogger-*chi*; streamer > strimer-*chi*), which could be codified in domain-specific translation guidelines.

6. CONCLUSION

This study has presented a systematic comparative-typological analysis of 320 digital neologisms in English and Uzbek, addressing formation mechanisms, semantic field distribution, and lexical adaptation strategies. The principal findings can be summarised as follows:

1. English digital neologism formation is dominated by compounding and affixation, while Uzbek relies primarily on affixation and borrowing, reflecting the agglutinative typological character of the recipient language.
2. Social Networking and Artificial Intelligence constitute the most productive semantic fields in both languages, indicating universal digital experience as a driver of lexical convergence.
3. Uzbek employs four principal adaptation strategies for English-origin digital neologisms: direct borrowing, phonological assimilation, morphological integration, and semantic calquing, with register determining strategy selection.
4. Register-based diglossia between formal calqued forms and informal citation borrowings creates lexicographic standardisation challenges for Uzbek language planning institutions.

These findings contribute to comparative linguistics, translation studies, and sociolinguistics by providing both a descriptive account of the current digital lexical landscape in English and Uzbek and a theoretical framework—typologically-informed contact linguistics—for analysing future waves of digital vocabulary.

Future research should extend the corpus temporally to track the stabilisation or obsolescence of identified neologisms, and geographically to compare Uzbek patterns with those of other Turkic languages facing analogous contact pressures. Computational approaches, including diachronic word embedding analyses, may enable more fine-grained tracking of semantic drift in digital neologisms over time.



REFERENCES

- Androutsopoulos, J. (2013). Code-switching in computer-mediated communication. In S. Herring, D. Stein, & T. Virtanen (Eds.), *Pragmatics of Computer-Mediated Communication* (pp. 667–694). De Gruyter Mouton.
- Bauer, L. (1983). *English Word-Formation*. Cambridge University Press.
- Boeschoten, H. (1998). The speakers of Uzbek: A sociolinguistic profile. In L. Johanson & É. Csató (Eds.), *The Turkic Languages* (pp. 357–363). Routledge.
- Crystal, D. (2004). *Language and the Internet* (2nd ed.). Cambridge University Press.
- Johanson, L. (2002). *Structural Factors in Turkic Language Contacts*. Curzon Press.
- Kleparski, G. A., & Rusinek, A. (2013). The domain of word-formation: English neologisms at the turn of the 21st century. *Studies in Linguistics and Methodology*, 4(2), 45–67.
- Lehrer, A. (1974). *Semantic Fields and Lexical Structure*. North-Holland.
- Levin, B. (2017). The lexicon in the age of social media: Formation processes and digital neologisms. *Journal of English Linguistics*, 45(3), 219–244.
- Mahmudov, N. (2015). O'zbek tilida yangi so'z yasash jarayonlari [New word-formation processes in the Uzbek language]. *O'zbek tili va adabiyoti*, 3, 12–21.
- Muhamedova, R. (2016). Kazakh and Uzbek: Language policy, attitudes and use in Central Asia. *Language Policy*, 15(4), 463–480.
- Nazarov, A. (2019). Ingliz tilidan o'zbekchaga so'z o'zlashtirishning zamonaviy tendentsiyalari [Contemporary trends in borrowing from English into Uzbek]. *Filologiya masalalari*, 2, 34–49.
- Newmark, P. (1988). *A Textbook of Translation*. Prentice Hall.
- Rakhmatullaev, Sh. (2020). O'zbek tilining yangi so'zlari va ularning lug'atlarda aks etishi [New words of the Uzbek language and their reflection in dictionaries]. Navruz nashri.
- Renouf, A. (2013). Corpus linguistics: Progress and prospects. In S. Granger & S. Petch-Tyson (Eds.), *Extending the Scope of Corpus-Based Research* (pp. 3–21). Rodopi.
- Seargeant, P., & Tagg, C. (2014). *The Language of Social Media: Identity and Community on the Internet*. Palgrave Macmillan.
- Sjoberg, A. F. (1963). *Uzbek Structural Grammar*. Indiana University Publications.
- Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. Linguistic Circle of New York.
- Yus, F. (2011). *Cyberpragmatics: Internet-Mediated Communication in Context*. John Benjamins.