THE USE OF THE UZBEK LANGUAGE ON SOCIAL MEDIA AND LINGUISTIC CHANGES

Ochilova Shahlo Shuxratjon qizi

1-st- year Master's student, Linguistics (Romance–Germanic Languages),

Renaissance Education University, Tashkent, Uzbekistan.

Email: ochilovashahlo@gmail.com

Doi :https://doi.org/10.5281/zenodo.20047041

Abstract: The present chapter examines the theoretical foundations for studying the use of the Uzbek language on social media platforms and the linguistic changes arising from digital communication. Drawing on sociolinguistics, computer-mediated communication theory, and code-switching research, the chapter argues that the emergence of social media as a primary communicative arena has introduced structural, lexical, and orthographic pressures on Uzbek that cannot be understood without reference to the language's post-Soviet sociolinguistic context. The chapter surveys the key theoretical frameworks — including David Crystal's concept of Netspeak, John Gumperz's approach to code-switching as a discourse strategy, Carol Myers-Scotton's Markedness Model, and research on language and identity in digital environments — and situates them within the specific conditions of the Uzbek linguistic landscape: the ongoing script transition from Cyrillic to Latin, the persistence of Russian as a prestige code, and the growing influence of English as the language of global digital culture. The findings of this theoretical survey suggest that Uzbek on social media is best understood as a site of active negotiation between national identity, globalization, and digital informality.

Keywords: Uzbek language, social media, linguistic change, code-switching, Netspeak, computer-mediated communication, sociolinguistics, language policy, digital identity, script mixing

Аннотация: Данная глава посвящена теоретическим основам изучения использования узбекского языка в социальных сетях и языковых изменений, возникающих в условиях цифровой коммуникации. Опираясь на социолингвистику, теорию компьютерно-опосредованной коммуникации и исследования в области переключения кодов, автор обосновывает тезис о том, что появление социальных сетей в качестве основного коммуникативного пространства создало структурное, лексическое и орфографическое давление на узбекский язык, которое невозможно понять без учёта постсоветского социолингвистического контекста. В главе рассматриваются ключевые теоретические концепции — нетспик Дэвида Кристала, подход Джона Гамперца к переключению кодов как дискурсивной стратегии, модель маркированности Кэрол Майерс-Скоттон, а также исследования языка и идентичности в цифровой среде — и помещаются в специфические условия узбекского языкового пространства: продолжающийся переход от кириллицы к латинице, сохранение русского языка в роли престижного кода и нарастающее влияние английского языка как глобального языка цифровой культуры.

Ключевые слова: узбекский язык, социальные сети, языковые изменения, переключение кодов, нетспик, компьютерно-опосредованная коммуникация, социолингвистика, языковая политика, цифровая идентичность, смешение графических систем

1. Introduction

The rise of social media as a dominant space for everyday communication has created conditions that no linguist of the twentieth century could have fully anticipated. Platforms such as Telegram,

Instagram, TikTok, and YouTube have become primary arenas for social interaction across the globe, and with them has emerged a set of communicative norms, lexical innovations, and structural conventions that differ in important ways from those governing face-to-face speech or formal written language. For speakers of Uzbek — a language navigating a historically complex sociolinguistic landscape shaped by Soviet language policy, post-independence reforms, an ongoing script transition, and the competing pressures of Russian and English — social media represents not just a new technology but a new kind of linguistic space in which old questions of identity, prestige, and norm are being renegotiated in real time.

Uzbek is spoken natively by an estimated 35 to 40 million people worldwide, making it the most widely spoken indigenous language in Central Asia and the second most widely spoken Turkic language after Turkish [Wikipedia: Uzbek language., 2024]. Despite this scale, the language finds itself at a critical juncture: the transition from Cyrillic to a Latin-based script, officially mandated since 1993, remains incomplete as of 2024, with different generations, institutions, and media channels using different writing systems. Social media has accelerated these tensions, producing a digital environment in which script choice, language mixing, and register informalization interact in ways that have no clear precedent in the pre-digital history of Uzbek.

This chapter establishes the theoretical foundations for the empirical investigation that follows. It surveys the major frameworks relevant to the study of Uzbek on social media — computer-mediated communication theory, code-switching research, the sociolinguistics of language policy in post-Soviet Central Asia, and research on language and digital identity — and shows how each bears on the specific questions the dissertation addresses. No empirical data is analyzed here; the chapter's purpose is to build the conceptual architecture within which the analysis of the following chapters will operate.

2. Computer-Mediated Communication and the Concept of Netspeak

The foundational theoretical contribution to the study of language on the internet is David Crystal's concept of Netspeak. In *Language and the Internet*, Crystal defined Netspeak as a type of language displaying features that are unique to the internet, arising from its character as a medium that is electronic, global, and interactive [D. Crystal., 2006., p. 20]. Crystal's central argument was that the internet is not merely a technological revolution but a social one, and that language is central to that revolution [D. Crystal., 2006., p. 6]. Crucially, he insisted that Netspeak should not be understood as a corruption of existing language but as a genuinely new language dimension — one that selectively draws on properties of both spoken and written modes while doing things that neither can do alone.

Crystal identified several levels at which Netspeak innovations consistently appear: graphology and orthography (abbreviations, acronyms, emoticons, non-standard capitalization, and punctuation); the lexicon (new words coined for digital contexts, blends, borrowings); grammar (shortened syntactic structures, omitted function words); and discourse (the management of register, turn-taking, and topic across asynchronous and synchronous exchanges) [D. Crystal., 2006., p. 17]. He argued that graphology and the lexicon are the levels at which Netspeak distinctiveness is most immediately visible — the levels at which innovation is easiest to introduce and most readily noticed by participants [D. Crystal., 2006., p. 91].

For the study of Uzbek on social media, Crystal's framework is productive in two ways. First, it provides a descriptive vocabulary for the linguistic innovations Uzbek users introduce in digital contexts — the abbreviated forms, the mixed scripts, the borrowed slang that characterize Uzbek texting and social media posts. Second, it raises the theoretical question of whether Netspeak features in Uzbek are merely local adaptations of globally circulating digital norms, or whether they are shaped in specific ways by the linguistic history and current sociolinguistic situation of Uzbek. Research on the linguistic features of Uzbek internet texting has confirmed that digital Uzbek communication shows all the Netspeak features Crystal describes — abbreviations, slang, code-switching, informal syntax, and the

extensive use of emojis — but also exhibits the specific pressures of script instability and multilingualism that are unique to the Uzbek context [A. B. Abdullaeva & Z. T. Tukhtakhodjayeva., 2025].

A further theoretical contribution of relevance is Susan Herring's work on computer-mediated discourse analysis. Herring emphasized that internet language is not a single uniform variety but a set of practices shaped by the specific medium, interface, and communicative situation in which they occur [S. C. Herring., 2008., p. 2]. The language of a Telegram channel differs from that of an Instagram comment or a TikTok caption. This diversity of digital registers means that any study of Uzbek on social media must attend to the specific platform contexts in which data is collected, rather than treating all social media language as equivalent.

3. The Sociolinguistic Context: Uzbek Language Policy and Script Reform

No theoretical account of Uzbek on social media can proceed without engaging the historical forces that have shaped the current sociolinguistic landscape. The Uzbek language has undergone four script transitions in less than a century: from an Arabic-based script to Latin in 1929, from Latin to Cyrillic in 1940, and from Cyrillic back to a modified Latin alphabet following independence in 1993 [The Diplomat., 2023]. Each transition left its mark on the language's relationship to literacy, identity, and authority — and the most recent transition remains unfinished, making Uzbek the only major Turkic language currently caught between two active orthographic systems.

The 1993 script reform was intended as both a symbolic and practical break from the Soviet past. As William Fierman documented in his analysis of language policy and identity politics in Central Asia, the shift from the Cyrillic alphabet was not merely a technical orthographic decision but a deliberate repositioning of Uzbek identity — away from Soviet heritage and toward a Turkic and globally oriented future [W. Fierman., 2009., pp. 1207–1228]. However, the reform has been incomplete in practice. As of 2024, most institutions still use both alphabets simultaneously, and a full transition, originally planned for January 2023, has been repeatedly postponed [The Diplomat., 2023].

This unresolved script coexistence has produced one of the most distinctive features of Uzbek digital communication: the simultaneous use of multiple orthographic conventions within a single text or even a single message. Research on the linguistic and cultural aspects of Uzbek internet communication has confirmed that formal versus informal register distinctions, the influence of Russian and English, and the coexistence of Latin and Cyrillic are among the defining characteristics of online Uzbek — features that the internet has not created but has amplified and accelerated [T. S. Vohidova., 2024., p. 301].

The persistence of Russian as a prestige code adds a further layer of complexity. Since 1991, Uzbekistan has promoted Uzbek as the state language, yet Russian continues to play a significant role in education, business, and media. Fierman's analysis of post-Soviet language policy traced this tension between the symbolic promotion of Uzbek and the practical dominance of Russian in professional and technical domains [W. Fierman., 2009., pp. 1215–1219]. On social media, this tension is visible in the widespread Uzbek-Russian code-switching that characterizes the digital communication of urban, educated users, particularly in Tashkent and other major cities.

4. Code-Switching Theory and Its Relevance to Uzbek Digital Communication

The central theoretical framework for analyzing the multilingual practices of Uzbek social media users is code-switching theory. Code-switching — the use of two or more languages or language varieties within a single communicative event — has been studied from several theoretical perspectives, each illuminating a different dimension of the phenomenon.

The foundational contribution is John Gumperz's interactional sociolinguistic approach. In *Discourse Strategies*, Gumperz defined conversational code-switching as the juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems [J. J. Gumperz., 1982., p. 59]. His key insight was that code-switching is not a sign of

linguistic deficiency but a communicative resource — a discourse strategy through which speakers signal shifts in topic, interpersonal stance, and social identity. Gumperz identified several conversational functions of code-switching, including quotation, addressee specification, interjection, reiteration for emphasis, message qualification, and personalization versus objectivization — functions that are well attested across bilingual speech communities [J. J. Gumperz., 1982., pp. 75–84]. All of these functions are observable in Uzbek social media discourse, where speakers switch codes to signal expertise, align themselves with particular social identities, or navigate the expectations of mixed audiences.

Carol Myers-Scotton's Markedness Model offers a complementary and more explicitly sociological account of code-switching motivation. Myers-Scotton proposes that bilingual speakers select their language code rationally, as a means of indexing the set of rights and obligations they wish to establish in a given interaction [C. Myers-Scotton., 1993]. Every speech community has an unmarked code — the expected, default choice for a given interaction type — and speakers who deviate from this unmarked choice make a marked choice that carries additional social meaning. On Uzbek social media, standard Uzbek may function as the unmarked code for nationally oriented or formal content, while switching to Russian or English constitutes a marked choice signaling cosmopolitanism, modernity, or professional status.

The Markedness Model is particularly relevant for understanding language choice in Uzbek digital discourse because the social indexicality of Uzbek, Russian, and English is historically weighted in ways that differ from many other multilingual contexts. Russian retains associations with Soviet-era education and urban sophistication; Uzbek is associated with national identity and cultural authenticity; English increasingly indexes global digital culture and professional aspiration. A post that mixes all three languages is not a product of incomplete competence — it is a socially meaningful performance of a hybrid identity, as research on Uzbek-English code-switching on Instagram has confirmed [S. Karimova., 2025., pp. 1066–1067].

Research specifically on Uzbek-Russian code-switching in contemporary Uzbekistan provides empirical grounding for these theoretical claims. Mizuki Sakurama Nakamura's discourse analysis of code-switching practices among Uzbek-Russian bilinguals in Tashkent found that intra-sentential switching — switching within a clause — accounts for 58% of all code-switching instances, and that contextual factors such as topic, setting, and the age of participants significantly shape which code is chosen [M. Sakurama Nakamura., 2024., p. 118]. The study also identified the main functions of code-switching as lexical gap-filling, emphasis, quotation, and expression of group identity — findings that align directly with Gumperz's functional taxonomy and confirm its applicability to the Uzbek context.

5. Language and Identity in Digital Spaces

A third theoretical strand necessary for the analysis of Uzbek on social media is the relationship between language use and identity construction in digital environments. Sociolinguistic research has moved away from essentialist conceptions of identity — identity as a fixed property of individuals or groups — toward performative accounts, in which identity is understood as something produced through social practice rather than simply expressed by it. Language choices actively construct, rather than merely reflect, the speaker's social positioning.

In digital contexts, this performative dimension is particularly salient. Social media users curate their online personas through deliberate choices of language, code, script, register, and genre. Research on Uzbek-English bilinguals on Instagram has shown that code-switching on social media serves identity-constructing functions: users deploy code-switching to perform hybrid identities, demonstrate language proficiency, signal affiliation with global digital culture, and engage with diverse audiences whose expectations differ [S. Karimova., 2025., pp. 1066–1067]. An Instagram post, a Telegram message, and a TikTok caption are not neutral communicative acts — they are identity performances directed at imagined audiences with specific social expectations.

For Uzbek speakers, the identity stakes of language choice on social media are particularly high because of the language's role in post-independence national identity construction. Since 1991, the Uzbek language has been positioned as the primary symbol of Uzbek national identity — the linguistic embodiment of post-Soviet cultural sovereignty. Yet the daily reality of digital communication places this national language in constant contact with Russian and English, creating conditions in which every language choice carries symbolic weight beyond its immediate communicative function [W. Fierman., 2009., pp. 1219–1228].

Research on code-switching among Uzbek youth in digital communication has documented this dynamic directly. The study published in the *International Journal of Artificial Intelligence* (2025) found that bilingual and multilingual practices have become the norm rather than the exception in the online communication of young Uzbek speakers, driven by the influence of social media, messaging platforms, and globalized digital culture. The study describes linguistic, social, and cultural motivations behind code-switching and notes that social identity expression is among the most consistent motivators — a finding that aligns with both Gumperz's functional account and Myers-Scotton's markedness framework [International Journal of Artificial Intelligence., 2025].

6. Categories of Linguistic Change Induced by Digital Communication

The theoretical literature on language change in digital environments identifies several categories of linguistic innovation that are directly relevant to Uzbek. These categories will organize the empirical analysis of the subsequent chapters.

The first category is *lexical borrowing and neologism*. Digital communication introduces new referential needs — platform names, digital actions, technology terms — that languages must meet through borrowing, calquing, or native word formation. Uzbek on social media shows extensive borrowing from English (*like qilmoq*, *post solmoq*, *story*) and from Russian (particularly in technical and professional domains), alongside the emergence of hybrid forms that combine Uzbek morphology with borrowed lexical stems. Research on the linguistic features of Uzbek internet texting has catalogued these borrowings and noted that they spread through social media use faster than any language planning body can regulate [A. B. Abdullaeva & Z. T. Tukhtakhodjayeva., 2025].

The second category is *orthographic and graphological innovation*. As noted in section 3, script mixing — the simultaneous use of Latin and Cyrillic characters within a single text — is a defining feature of Uzbek digital communication. Beyond script mixing, Uzbek social media language shows the standard Netspeak orthographic features Crystal describes: abbreviations, phonetic spellings, capitalization for emphasis, and the use of numbers and symbols as letter substitutes. Research published in the *Central Asian Journal of Literature, Philosophy and Culture* identifies these features as constituting hybrid communicative codes in Uzbek social media texts — systematic patterns of norm variation that function not as simple errors but as markers of the digital register [CAJLPC., 2025].

The third category is *register informalization*. Social media platforms encourage informal, colloquial language use regardless of topic — a tendency that Crystal noted as a general feature of Netspeak [D. Crystal., 2006., p. 6]. In Uzbek, where the distinction between formal and informal written registers has historically been maintained carefully, this informalization represents a meaningful shift. Research on online Uzbek communication confirms that internet discourse consistently blurs formal and informal boundaries, producing a new register that is neither the careful written standard of official communication nor the intimate colloquial of private speech [T. S. Vohidova., 2024., p. 302].

The fourth category is *discourse structural innovation*. Social media platforms impose specific constraints on how discourse is organized — character limits, threading conventions, visual formats, and emoji systems shape what can be said and how. In Uzbek, these platform constraints interact with existing discourse conventions in ways that are not yet fully documented. The use of emoji as pragmatic markers, the shortening of explanatory structures, the use of hashtags as metalinguistic commentary —

all of these represent structural innovations in Uzbek written discourse that have no direct precedent in pre-digital Uzbek writing.

7. Code-Switching Among Uzbek Youth on Social Media: A Research Overview

Research specifically focused on Uzbek youth and digital code-switching has produced findings that help specify the theoretical picture outlined above. A study on code-switching among Uzbek youth in digital communication, published in the *International Journal of Artificial Intelligence*, found that bilingual and multilingual practices have become normalized in the online communication of young Uzbek speakers, driven by the growing influence of social media and globalized communication environments [International Journal of Artificial Intelligence., 2025]. The study examines linguistic, social, and cultural motivations behind code-switching and its structural patterns, drawing on the theoretical frameworks of Gumperz (1982) and Myers-Scotton (1993) — the same frameworks that the present dissertation applies.

On Instagram specifically, research on Uzbek-English bilinguals found that code-switching serves to perform what the author terms hybrid identities — identities that are simultaneously local and global, nationally rooted and digitally mobile [S. Karimova., 2025., p. 1066]. This hybrid identity performance is not available through Uzbek alone or through English alone — it requires the visible alternation between them, the simultaneous presence of both codes in a single post or caption. This finding aligns with the theoretical prediction of Myers-Scotton's Markedness Model: the choice to switch codes is rational and socially motivated, indexing a specific relational position and social persona.

Sakurama Nakamura's discourse analysis of Uzbek-Russian bilingual code-switching in Tashkent reaches a parallel conclusion for the Uzbek-Russian pair: code-switching is not random but serves well-defined communicative and identity functions, including lexical gap-filling (switching to Russian for technical or professional vocabulary unavailable or unfamiliar in Uzbek), emphasis, quotation, and expression of group identity [M. Sakurama Nakamura., 2024., p. 118]. Together, these studies establish a consistent empirical picture in which Uzbek digital code-switching is a purposeful sociolinguistic practice, governed by the same structural and social logics that Gumperz and Myers-Scotton identified in other multilingual communities.

8. The Tension Between Language Preservation and Digital Language Change

A recurring theme in the literature on Uzbek and social media is the tension between the preservation of linguistic norms and the dynamics of language change that digital communication accelerates. This tension is not unique to Uzbek — it characterizes the relationship between any standard language and its digital vernacular — but it takes on particular intensity in the Uzbek case because of the language's unfinished script reform and the active role of the state in promoting standard Uzbek as a national symbol.

Crystal argued that internet language, far from threatening established languages, creates opportunities for linguistic expansion and creativity, enabling forms of expression that the standard language alone cannot provide [D. Crystal., 2006., p. 6]. From this perspective, the innovations of Uzbek social media language are not violations of the standard but extensions of it — creative adaptations to a new communicative environment that enrich the expressive resources of the language.

A more cautious assessment is offered by research on internet communication and Uzbek language norms. A study published in the *Central Asian Journal of Literature, Philosophy and Culture* notes that social media texts in Uzbek display patterns of phonetic, lexical, syntactic, and stylistic norm violation that erode traditional literary standards and complicate the task of standardizing Uzbek across the Latin-Cyrillic divide [CAJLPC., 2025]. From this perspective, social media language represents a genuine challenge to language planning, particularly given that social media is now the primary writing environment for many young Uzbek speakers who receive most of their written language input through digital channels rather than formal education.

These two perspectives are not incompatible. The theoretical literature suggests that both are accurate simultaneously: digital communication does enrich the expressive range of Uzbek through new vocabulary, registers, and hybrid codes, and it does challenge the standard norms that language planning seeks to establish. Resolving this tension is a task for language policy, not for linguistic analysis. What the present study aims to do is document and analyze the linguistic changes themselves — to describe what is happening before prescribing what should happen.

9. Theoretical Synthesis and the Framework of the Present Study

The preceding survey of theoretical frameworks allows for a synthesis that will guide the empirical analysis of the dissertation's subsequent chapters. The present study operates within a sociolinguistic framework that draws on three interconnected theoretical pillars.

The first pillar is Crystal's computer-mediated communication theory and the concept of Netspeak [D. Crystal., 2006], which provides a descriptive vocabulary for the formal linguistic features of Uzbek digital language — the orthographic, lexical, and structural innovations that distinguish social media Uzbek from standard written Uzbek. These features will be identified and analyzed in the empirical chapters.

The second pillar is Gumperz's interactional sociolinguistics and Myers-Scotton's Markedness Model [J. J. Gumperz., 1982; C. Myers-Scotton., 1993], which provide the theoretical account of code-switching as a socially motivated communicative strategy. These frameworks will be used to interpret the patterns of language alternation observed in the data — to ask not only what switches occur but why they occur, what social meanings they carry, and how they relate to the specific sociolinguistic conditions of Uzbekistan.

The third pillar is research on language and identity in digital environments [S. Karimova., 2025; M. Sakurama Nakamura., 2024], which situates formal and sociolinguistic analysis within the broader question of what it means to perform a social identity through digital language choices. This pillar connects the micro-level analysis of individual language features to the macro-level dynamics of language policy, national identity, and globalization.

10. Conclusion

This chapter has established the theoretical foundations for the empirical investigation of Uzbek language use on social media. The review of the relevant literature has shown that the phenomenon requires the integration of at least three theoretical perspectives: computer-mediated communication theory, code-switching research, and the sociolinguistics of language policy and identity.

The Uzbek case is theoretically productive precisely because of its distinctive complexity. The ongoing script transition, the triglossic relationship between Uzbek, Russian, and English, the state's active promotion of Uzbek as a national symbol, and the rapidly growing engagement of Uzbek youth with global social media platforms together create a sociolinguistic situation in which the linguistic changes induced by digital communication intersect with long-standing processes of identity construction and post-Soviet cultural reconfiguration.

The theoretical frameworks surveyed in this chapter — Crystal's Netspeak, Gumperz's discourse strategies, Myers-Scotton's Markedness Model, and research on digital identity — provide the conceptual tools needed to approach this complexity systematically and rigorously. The following chapters will apply these tools to empirical data, with the goal of contributing to a fuller understanding of how Uzbek is changing in the digital age and what those changes mean for the language and its speakers.

References

1. Abdullaeva, A. B., & Tukhtakhodjayeva, Z. T. (2025). Linguistic features of internet texting in Uzbek language. European International Journal of Philological Sciences.
2. Crystal, D. (2006). Language and the Internet (2nd ed., pp. 6, 17, 20, 91). Cambridge University Press.
3. Fierman, W. (2009). Identity, symbolism, and the politics of language in Central Asia. Europe-Asia Studies, 61(7), 1207–1228. https://doi.org/10.1080/09668130903068731
4. Gumperz, J. J. (1982). Discourse Strategies (pp. 59, 75–84). Cambridge University Press.
5. Herring, S. C. (2008). Language and the Internet. In W. Donsbach (Ed.), International Encyclopedia of Communication (p. 2). Wiley-Blackwell. PDF
6. Internet linguistic features and the threat to Uzbek language norms. (2025). Central Asian Journal of Literature, Philosophy and Culture (CAJLPC), 6(4). https://cajlpc.casjournal.org/index.php/CAJLPC/article/view/1320
7. Karimova, S. (2025). The role of code-switching in identity construction among Uzbek-English bilinguals on Instagram. Journal of Applied Science and Social Science, 15(09), 1066–1072. https://www.internationaljournal.co.in/index.php/jasass/article/view/1945
8. Myers-Scotton, C. (1993). Social Motivations for Code-Switching: Evidence from Africa. Clarendon Press. Archive.org
9. Sakurama Nakamura, M. (2024). Beyond bilingualism: A discourse analysis of Uzbek-Russian code-switching in contemporary Uzbekistan (p. 118). Turkophone.
10. Uzbekistan's drawn-out journey from Cyrillic to Latin script. (2023). The Diplomat.
11. Vohidova, T. S. (2024). Linguistic and cultural aspects of internet communication in Uzbek (pp. 301–302). Kokand University Scientific Repository.
12. Code-switching among Uzbek youth in digital communication. (2025). International Journal of Artificial Intelligence.