



<https://doi.org/10.70728/b.series.fil.v08.i08.101>

DEVIATIONS FROM ORTHOGRAPHIC NORMS IN THE LANGUAGE OF SOCIAL MEDIA AND THEIR CAUSES

Aslanov Akmal Subxanovich

Tashkent State Medical University

Candidate of Philological Sciences, Associate Professor

Tashkent, Uzbekistan

akmalaslanov66@gmail.com

ORCID: 0009-0009-7194-1346

ABSTRACT. This paper examines the systematic deviation of social media language from the codified orthographic norm and explains why such deviation occurs so consistently across languages and platforms. Orthography is treated not as a neutral technical code but, following Sebba, as a socially regulated system whose rules carry the weight of institutional authority; departure from that system is therefore always socially meaningful rather than accidental. The study surveys the principal categories of orthographic deviation attested in the international literature on computer mediated communication, including phonetic and abbreviated respelling, consonant and vowel deletion, letter and punctuation repetition, capitalization used for emphasis rather than grammatical marking, emoticon and emoji substitution for lexical or punctuational material, and deliberate anti normative respelling associated with particular online subcultures. It then classifies the causes of these deviations into four groups: technological and interface constraints, the economy and speed demanded by synchronous or near synchronous exchange, the hybrid oral written nature of computer mediated discourse, and the social and identity functions that non standard spelling performs for its users. The discussion draws on the established sociolinguistic and media linguistic scholarship of Crystal, Sebba, Androutsopoulos, Herring, Shaw, Squires, Tagliamonte and Denis, and Deumert and Masinyana, supported throughout by concrete examples in English, Russian and other languages. The paper concludes that orthographic deviation in social media should be understood as a regular, rule governed linguistic phenomenon rather than as simple carelessness or a decline in literacy, and that its systematic description has direct relevance for language norm policy and for the sociolinguistic study of digital communication generally.

Keywords: orthographic norm; social media language; computer mediated communication; non standard spelling; digital literacy; sociolinguistics; language variation.

ANNOTATSIYA. Ushbu maqola ijtimoiy tarmoqlar tilining kodifikatsiyalangan orfografik me'yordan tizimli chetlanishini o'rganadi va bunday chetlanish nima uchun tillar va platformalar bo'ylab bunchalik izchil sodir bo'lishini izohlaydi. Orfografiya neytral texnik kod sifatida emas, balki qoidalari institutsional hokimiyat vaznini o'zida mujassam etgan ijtimoiy jihatdan tartibga solinadigan tizim sifatida ko'rib chiqiladi. Shu sababli bu tizimdan chetlanish tasodifiy emas, balki har doim ijtimoiy ma'noga ega bo'ladi. Tadqiqot kompyuter vositachiligidagi kommunikatsiya bo'yicha xalqaro adabiyotda qayd etilgan orfografik chetlanishning asosiy toifalarini ko'rib chiqadi, jumladan fonetik va qisqartirilgan qayta yozish, undosh va unli tovushlarning tushirib qoldirilishi, harf va tinish belgilarining takrorlanishi, grammatik belgilash o'rniga urg'u berish uchun bosh harflardan foydalanish, leksik yoki punktuatsion materialni emotikon va emojiga almashtirish, hamda ma'lum onlayn submadaniyatlar bilan bog'liq bo'lgan atayin normaga zid qayta yozish. So'ngra u ushbu chetlanishlarning sabablarini to'rt guruhga tasniflaydi: texnologik va interfeys cheklovlari, sinxron yoki deyarli sinxron almashinuv talab qiladigan tejamkorlik va tezlik, kompyuter vositachiligidagi diskursning gibrid og'zaki-yozma tabiati, hamda nostandart imloning foydalanuvchilar uchun bajaradigan ijtimoiy va o'ziga xoslik funksiyalari. Muhokama Kristal, Sebba, Androutsopoulos, Herring, Shou, Skvayrs, Tagliamonte va Denis, Deumert va Masinyananing shakllangan sotsiolingvistik va media-lingvistik ilmiy asarlariga tayanadi, bu esa ingliz, rus va boshqa tillardagi aniq misollar bilan qo'llab-quvvatlanadi. Maqola shunday xulosaga keladiki, ijtimoiy tarmoqlardagi orfografik chetlanish oddiy beparvolik yoki savodxonlikning pasayishi sifatida emas, balki muntazam, qoidalarga bo'ysunuvchi lingvistik hodisa sifatida tushunilishi kerak, va uning tizimli tavsifi til me'yori siyosati hamda umuman raqamli kommunikatsiyaning sotsiolingvistik tadqiqi uchun bevosita ahamiyatga ega.

Kalit so'zlar: orfografik me'yor, ijtimoiy tarmoq tili, kompyuter vositachiligidagi kommunikatsiya, nostandart imlo, raqamli savodxonlik, sotsiolingvistika, til varianti.

АННОТАЦИЯ. В данной статье рассматривается систематическое отклонение языка социальных сетей от кодифицированной орфографической нормы и объясняется, почему такое отклонение столь последовательно проявляется в разных языках и на разных платформах. Орфография рассматривается не как нейтральный технический код, а как социально регулируемая система, правила которой несут на себе вес институциональной власти. Поэтому отклонение от этой системы никогда не бывает случайным — оно всегда социально значимо. В исследовании рассматриваются основные категории орфографических отклонений, зафиксированные в международной литературе по компьютерно-опосредованной коммуникации, включая фонетическое и сокращённое переписывание слов, выпадение согласных и гласных, повторение букв и знаков препинания, использование заглавных букв для выделения, а не для грамматической маркировки, замену лексического или пунктуационного материала эмодзи и эмодзи, а также намеренное антинормативное написание, связанное с определёнными онлайн-субкультурами. Далее причины этих отклонений

классифицируются по четырём группам: технологические и интерфейсные ограничения, экономия и скорость, требуемые синхронным или почти синхронным обменом сообщениями, гибридная устно-письменная природа компьютерно-опосредованного дискурса, а также социальные функции и функции идентичности, которые нестандартное написание выполняет для своих пользователей. Обсуждение опирается на устоявшиеся социолингвистические и медиалингвистические труды Кристала, Себбы, Андруцопулоса, Херринг, Шоу, Сквайрс, Тальямонте и Дениса, Деумерт и Масиньяны, и на протяжении всей работы подкрепляется конкретными примерами на английском, русском и других языках. В статье делается вывод, что орфографическое отклонение в социальных сетях следует понимать как регулярное, подчиняющееся правилам языковое явление, а не как простую небрежность или снижение грамотности, и что его систематическое описание имеет непосредственное значение для языковой нормативной политики и для социолингвистического изучения цифровой коммуникации в целом.

Ключевые слова: орфографическая норма, язык социальных сетей, компьютерно-опосредованная коммуникация, нестандартное написание, цифровая грамотность, социолингвистика, языковая вариативность.

INTRODUCTION

The written norm of a national language is, historically, the product of institutions: schools, dictionaries, printing houses, editorial boards and, in many countries, an official language authority. Social media communication is produced almost entirely outside the reach of these institutions. A post, a comment or a private message is composed by an ordinary user, published without an intervening editor, and read within seconds by an audience that judges it primarily by its communicative effect rather than by its conformity to the codified spelling system. It is under precisely these conditions that deviation from the orthographic norm becomes not an occasional lapse but a pervasive and recurrent feature of the written language circulating online.

The theoretical starting point of the present study is the observation, developed at length by Sebba, that orthography is never a purely technical matter of representing sound by letter. A writing system is a social institution: its rules are taught, sanctioned and enforced, and a speaker's choice to follow or to depart from them is itself a social act, capable of signaling education, regional origin, generational belonging or deliberate resistance to institutional authority. Sebba's central claim, that spelling operates as a system of social meaning in addition to a system of graphic representation, is what makes possible a serious linguistic account of internet spelling rather than a purely normative complaint about falling standards.

Crystal's description of the language of electronic communication as a distinct variety, which he termed Netspeak, situates this social account within a functional one. Crystal argued that computer mediated writing combines properties of speech, such as its rapid, unplanned, interactive character, with the permanence and visibility of writing, and that many of the

orthographic peculiarities of online text follow directly from this hybrid status rather than from ignorance of the standard system. A message composed in seconds, read once and then scrolled past, is simply not held to the same orthographic standard as a printed document intended for repeated, careful reading, and its author, in most cases, is perfectly capable of producing the standard spelling when the communicative situation calls for it.

Empirical studies conducted since the early years of mass internet use have confirmed that non standard spelling in digital communication is neither random nor evidence of declining literacy. Tagliamonte and Denis, examining a large corpus of instant messaging transcripts produced by Canadian teenagers, found that the great majority of forms in the corpus were, in fact, orthographically standard, and that the non standard forms which did occur followed regular, describable patterns rather than appearing at random; the authors explicitly rejected the popular characterization of instant messaging as linguistically ruinous. Herring's foundational work on computer mediated discourse likewise established that the linguistic features of online text, orthography included, are systematically shaped by the technical parameters of the medium: whether the exchange is synchronous or asynchronous, whether messages are public or private, and what character or bandwidth limitations the platform imposes.

Illustrative material from a range of languages and platforms makes the object of the present study concrete from the outset. English social media text regularly reduces function words and grammatical morphemes, as in u for you, ur for your or you are, and thx for thanks; it lengthens words through letter repetition to convey emphasis or prolonged vowel sound, as in soooo good or nooooo way; it multiplies punctuation marks well beyond their grammatical function, as in what?!?!?! or wait.....; and it uses full capitalization not to mark a proper noun but to signal raised volume or intensity, as in STOP DOING THAT. Runet, the Russian language segment of the internet, developed in the early two thousands an entire subcultural register, described in detail by Gorham, built on deliberately anti normative respellings such as preved for privet, kpacavчer for krasavchik, or aффтрап жжот for avtor zhzheth, in which the departure from the codified Cyrillic orthography is not accidental but is the very point of the utterance. South African bilingual texting studied by Deumert and Masinyana shows comparable phonetic respelling and language mixing across English and isiXhosa within a single message. These examples, drawn from typologically and culturally distant languages, converge on the same conclusion: orthographic deviation in digital writing is a structured, cross linguistically recurrent phenomenon and not a series of unconnected individual errors.

The aim of the present study is to classify the principal types of orthographic norm deviation documented in the scholarship on social media and computer mediated communication, and to relate each type systematically to the communicative and social conditions that produce it. Its contribution lies in bringing together, within a single explanatory scheme, the technological, economic, discourse structural and social identity based accounts of internet orthography that the literature has so far developed largely as separate lines of argument.

The material examined in this study consists of examples of orthographic deviation documented in the published scholarship on computer mediated communication together with illustrative instances drawn from publicly available social media text in English, Russian and other languages. Because the aim of the paper is classificatory and explanatory rather than corpus statistical, no new corpus was compiled; instead the study synthesizes findings already established through corpus based research, most notably the large scale instant messaging corpus analyzed by Tagliamonte and Denis and the bilingual SMS corpus analyzed by Deumert and Masinyana, alongside the qualitative and discourse analytic descriptions offered by Crystal, Sebba, Shaw, Squires and Heath.

The method applied is descriptive and classificatory, supplemented by functional analysis of the communicative purpose each type of deviation serves. Deviations are grouped first by the linguistic level at which they occur, distinguishing graphemic substitution, segmental deletion, iconic repetition, capitalization and punctuation, and non alphabetic substitution by emoji or emoticon. They are then related to four categories of cause identified in the literature: the technological and interface constraints of the device or platform used; the economy and speed required by real time or near real time exchange; the hybrid oral written character of computer mediated discourse described by Herring and by Crystal; and the social and identity signaling function of non standard spelling described by Sebba, Shaw and Squires. This dual classification, by linguistic type and by underlying cause, is what allows the study to move from a mere inventory of curiosities to an explanatory typology.

RESULTS AND DISCUSSION

Table 1 sets out the principal types of orthographic deviation identified in the material, together with representative examples and the linguistic level at which each operates.

Table 1. Types of orthographic deviation in social media language

Type	Linguistic operation	Representative examples
Phonetic respelling	Letters replaced to reflect pronunciation or to shorten the word	kool, luv, thx, gr8, nite, m8
Segmental deletion	Omission of vowels, silent letters or grammatical morphemes	u, ur, plz, msg, dont, cant
Letter repetition	Iconic lengthening of a grapheme to mark emphasis or prolonged sound	soooo good, nooooo, heyyyy
Punctuation multiplication	Repetition of exclamation, question or ellipsis marks beyond grammatical need	what?!?!, wait....., NO!!!!

Emphatic capitalization	Full capitalization used prosodically rather than grammatically	STOP DOING THAT, I SAID NO
Anti normative respelling	Deliberate, community specific departure from the codified script	preved, кросачег, аффтар жжот
Emoji and emoticon substitution	Non alphabetic sign replacing a lexeme or a punctuation mark	see you soon 😊 instead of a closing sentence
Script and code mixing	Alternation between scripts or languages within one message	isiXhosa and English mixed within a single SMS

Phonetic and abbreviated respelling. The most frequently discussed type of orthographic deviation involves replacing the standard spelling of a word with a shorter or more phonetically direct sequence of letters and numerals: kool for cool, luv for love, thx for thanks, gr8 for great, m8 for mate, and 2moro for tomorrow. Crystal, in his study of text messaging, showed that this device long predates the mobile phone and the internet, tracing analogous abbreviatory spelling back to nineteenth century rebus writing and to telegraphic style, and argued that its recent proliferation is a matter of degree, driven by character limits and typing effort, rather than an entirely new linguistic phenomenon.

Segmental deletion. A closely related type is the deletion of vowels, of silent consonants, or of grammatical morphemes that carry comparatively low informational load: u for you, ur for your or you are, plz for please, msg for message, and the frequent dropping of the apostrophe in contracted forms such as dont, cant and im. Tagliamonte and Denis found that even in their instant messaging corpus, where such deletions were fully available to teenage writers, the standard spelling of these very words still predominated overall, which indicates that segmental deletion functions as an optional stylistic resource rather than as an obligatory replacement of the norm.

Letter and punctuation repetition. A third type exploits the graphic medium iconically rather than phonemically. Letters are repeated to represent prolonged pronunciation or heightened emotion, as in soooo tired, nooooo way, or heyyyyy, and punctuation marks are multiplied well beyond the single mark the grammar requires, as in what?!?!?! or I cant believe it..... Heath's analysis of capitalization in social media similarly shows that full capitalization is regularly read by users not as a grammatical device but as a prosodic one, functioning as the written equivalent of raised volume or stress, so that STOP typed in capitals communicates something a lower case stop cannot.

Emoji and emoticon substitution. A further, medium specific type of deviation from the written norm consists in replacing an alphabetic sequence, or a punctuation mark, with a non alphabetic pictographic sign. A message ending simply in a smiling face emoji performs the function a closing sentence or a terminal punctuation mark would otherwise perform, and a heart emoji frequently substitutes for the word love altogether. Because such signs fall

outside the alphabetic writing system entirely, their increasing frequency represents not merely a variant spelling of an existing word but a partial displacement of alphabetic writing itself within certain communicative functions, a development the codified orthographic norm, framed for an alphabetic system, has no means of accommodating.

Deliberate anti normative respelling. The most socially marked type of deviation is not accidental at all but consciously and deliberately anti normative. Gorham's account of the Russian language Padonki subculture of the early Runet documents an entire register built on systematic, rule governed misspelling of standard Russian: *privet* respelled as *preved*, *krasavchik* as *красавчер*, and the phrase *avtor zhzheth* as *аффтар жжот*. Squires' study of what she terms the enregisterment of internet language demonstrates that such deliberately non standard forms come, over time, to index a recognizable social persona, in her case that of the unsophisticated or uneducated internet user, a persona competent writers can then adopt playfully precisely because they are able to depart from, and therefore also to return to, the standard norm at will.

Script and code mixing. In multilingual settings the deviation from a single codified orthography frequently takes the form of alternation between two writing systems or two languages within one message. Deumert and Masinyana's study of bilingual South African text messaging documents frequent alternation between English and isiXhosa spelling conventions within a single SMS, driven by the bilingual repertoire of the writer rather than by any single language's internal orthographic instability. Comparable script mixing is attested wherever two alphabets serve one speech community, and in each case the deviation is best described not as an error in either system taken separately but as a genuinely bilingual or bigraphic practice with its own regularities.

Table 2 relates each of the types described above to the underlying cause or causes the literature identifies for it, distinguishing technological, economic, discourse structural and social identity based explanations.

Table 2. Causes of orthographic deviation in social media language

Cause	Explanation	Associated deviation types
Technological and interface constraints	Small keyboards, character limits and predictive input encourage shorter, simplified spelling	Phonetic respelling, segmental deletion
Speed and economy of real time exchange	Rapid, unplanned composition favors the shortest adequate form over the codified one	Segmental deletion, abbreviation
Hybrid oral written character of the medium	Written text carries prosodic and paralinguistic information normally conveyed by speech	Letter repetition, punctuation multiplication, emphatic capitalization, emoji substitution

Social and identity signaling	Spelling choice indexes group membership, generational identity or a deliberately adopted persona	Anti normative respelling, script and code mixing
-------------------------------	---	---

Taken together, these four causal categories show that the deviations catalogued in Table 1 are not reducible to a single explanation. The technological and economic causes account well for segmental deletion and phonetic abbreviation, which are most frequent precisely where input is slowest or most constrained, on small mobile keyboards and within strict character limits. The discourse structural cause, the hybrid orality of computer mediated communication that Herring and Crystal both describe, accounts for the iconic and prosodic devices such as letter repetition, punctuation multiplication and emphatic capitalization, all of which supply, in written form, information that spoken language would otherwise convey through intonation, volume and pause. The social identity cause, developed most fully by Sebba and applied specifically to digital contexts by Shaw and by Squires, accounts for the type of deviation that cannot be explained by speed or by device constraint at all, namely the deliberately anti normative respelling documented in subcultures such as the Russian Padonki, where competent, literate writers choose the non standard form precisely because it is non standard, and where the resulting spelling becomes, in Squires' terms, an enregistered marker of a social persona rather than a mere shortcut.

This causal heterogeneity has an important consequence for how the phenomenon should be evaluated. A deviation caused by keyboard constraint, such as the dropped apostrophe in *dont*, carries little or no social meaning and is unlikely to be perceived by other users as a meaningful choice. A deviation caused by deliberate identity signaling, such as the Padonki respellings, carries a great deal of social meaning and is fully intended to be recognized as such. Treating the whole of internet orthography as a single undifferentiated departure from the norm, as much popular commentary continues to do, therefore obscures precisely the distinction that a linguistically informed description must draw. Milroy and Milroy's account of the standard language ideology is relevant here: the belief that any departure from the codified norm signals carelessness or a decline in literacy is itself a social construct maintained by institutions with authority over language, and it is this ideology, more than any linguistic fact about the deviations themselves, that produces the recurring public anxiety about the effect of digital communication on literacy.

The evidence surveyed here does not support that anxiety. Tagliamonte and Denis's corpus study found standard spelling to be the overwhelming majority pattern even among the demographic, teenage instant message users, most often assumed to write the most non standardly, and found that the non standard forms which did occur were used systematically and for identifiable communicative purposes rather than at random. Androutsopoulos's review of language change in digital media likewise concludes that the digital medium functions above all as an arena in which existing sociolinguistic variation becomes newly visible and newly discussable, rather than as a cause of linguistic decline in its own right. The

orthographic deviations catalogued in this study are, in this light, best understood as an expansion of the expressive resources available to the writer of digital text, achieved by selectively departing from a codified norm the writer generally continues to command.

CONCLUSIONS

The study has classified the principal types of orthographic norm deviation attested in the language of social media, distinguishing phonetic and abbreviated respelling, segmental deletion, letter and punctuation repetition, emphatic capitalization, emoji and emoticon substitution, deliberate anti normative respelling and script or code mixing. Each type was shown to correspond to one or more of four underlying causes: the technological and interface constraints of the device or platform, the economy and speed demanded by rapid or synchronous exchange, the hybrid oral written character of computer mediated discourse, and the social and identity signaling function that non standard spelling performs for its users. The classification demonstrates that orthographic deviation online is neither uniform nor random. Deviations produced by device constraint and by communicative speed differ, in their social meaning and in their communicative intent, from deviations produced deliberately to signal group membership or a particular social persona, and a serious linguistic description of digital writing must keep these causal categories apart rather than treating every departure from the codified norm as evidence of the same underlying process.

These findings carry a direct implication for language norm policy in the digital era. Because a substantial share of the orthographic deviation documented in social media is functional rather than erroneous, in the sense that it performs a genuine expressive or social task the codified norm was never designed to perform, blanket corrective approaches that treat all such deviation as a literacy problem are unlikely to be either accurate or effective. A more productive approach, consistent with the descriptive tradition represented by Crystal, Sebba and Herring, is to teach the boundary between register appropriate deviation and genuine orthographic error, so that competent bidialectal or bigraphic writing in digital contexts is recognized as a skill rather than dismissed as its absence. Further research, building on the corpus based methods of Tagliamonte and Denis and of Deumert and Masinyana, might usefully extend this typology through systematic quantitative comparison across languages, platforms and age groups, and through longitudinal study of which deviations, over time, move from marked social signal toward unmarked, accepted usage.

REFERENCES

1. Androutsopoulos, J. 2011. Language change and digital media: a review of conceptions and evidence. In K. Kelly Holmes and T. Milani (eds.), *Standard Languages and Language Standards in a Changing Europe*, Novus Press, 145 to 160.
2. Baron, N. 2008. *Always On: Language in an Online and Mobile World*. Oxford: Oxford University Press.
3. Crystal, D. 2006. *Language and the Internet*. 2nd edition. Cambridge: Cambridge University Press.

4. Crystal, D. 2008. Txtng: The Gr8 Db8. Oxford: Oxford University Press.
5. Deumert, A. and Masinyana, S. O. 2008. Mobile language choices: the use of English and isiXhosa in text messages (SMS), evidence from a bilingual South African sample. *English World Wide*, 29(2), 117 to 147.
6. Gorham, M. S. 2014. *After Newspeak: Language, Culture, and Politics in Russia from Gorbachev to Putin*. Ithaca: Cornell University Press.
7. Heath, M. 2018. Orthography in social media: pragmatic and prosodic interpretations of caps lock. *Proceedings of the Linguistic Society of America*, 3, 55, 1 to 13.
8. Herring, S. C. (ed.) 1996. *Computer Mediated Communication: Linguistic, Social and Cross Cultural Perspectives*. Amsterdam: John Benjamins.
9. Milroy, J. and Milroy, L. 2012. *Authority in Language: Investigating Standard English*. 4th edition. London: Routledge.
10. Sebba, M. 2007. *Spelling and Society: The Culture and Politics of Orthography Around the World*. Cambridge: Cambridge University Press.
11. Shaw, P. 2008. Spelling, accent and identity in computer mediated communication. *English Today*, 24(2), 42 to 49.
12. Squires, L. 2010. Enregistering internet language. *Language in Society*, 39(4), 457 to 492.
13. Tagliamonte, S. A. and Denis, D. 2008. Linguistic ruin? LOL! Instant messaging and teen language. *American Speech*, 83(1), 3 to 34.
14. Thurlow, C. and Mroczek, K. (eds.) 2011. *Digital Discourse: Language in the New Media*. Oxford: Oxford University Press.