УДК 811.111’373:004.77

ЕМЕРДЖЕНТНІ АБРЕВІАТУРИ В АНГЛОМОВНОМУ TWITTER/X ДИСКУРСІ: ЛІНГВІСТИЧНІ ПАТЕРНИ ТА ФУНКЦІЇ

Басараб Вікторія Сергіївна, Київ, Україна
Студентка Факультету романо-германської філології
Київський столичний університет імені Бориса Грінченка
vsbasarab.frgf24m@kubg.edu.ua

Serhiivna, Kyiv, Ukraine
Student of the Faculty of Romance and Germanic Philology Borys Grinchenko Kyiv Metropolitan University
vsbasarab.frgf24m@kubg.edu.ua

У статті представлено корпусне дослідження структурних та функціональних особливостей абревіатур в англійському дискурсі Twitter/X як проявів інноваційних трансформацій у мові соціальних медіа. Актуальність дослідження визначається необхідністю систематичного вивчення мовних інновацій у цифровій комунікації, які формують нові конвенції грамотності та відображають адаптацію мови до технологічних обмежень платформ. На основі корпусу з 500 англомовних твітів, проаналізованих за допомогою Voyant Tools, досліджуються механізми формування, закономірності поширення та прагматичні функції абревіатур у типових для цієї платформи контекстах. Методологія дослідження включає аналіз частотності, розподілу та контексту, що дозволяє виявити систематичні закономірності у використанні абревіатур. Функціональний аналіз показує, що абревіатури виконують три основні прагматичні функції: маркерів епістемічної модальності, засобів хеджування, що пом'якшують категоричні висловлювання, та індикаторів внутрішньогрупової ідентичності. Результати дослідження вказують на систематичний характер скорочень у дискурсі Twitter і підтверджують гіпотезу щодо формування стабільних конвенцій цифрової грамотності, що визначаються взаємодією між технологічними обмеженнями платформи та комунікативними потребами користувачів. Дослідження сприяє розумінню процесів мовних інновацій у комп'ютерно-опосередкованій комунікації та демонструє ефективність корпусної методології у дослідженні явищ цифрового дискурсу.

Ключові слова: *дискурс Twitter/X, скорочення, цифрова комунікація, прагматичні функції, маркери дискурсу, соціальні медіа*

This article presents a corpus study of the structural and functional features of abbreviations in English Twitter/X discourse as manifestations of innovative transformations in social media language. The relevance of the study is determined by the need for systematic study of linguistic innovations in digital communication, which form new conventions of literacy and reflect the adaptation of language to the technological limitations of platforms. Based on a corpus of 500 English-language tweets analyzed using VoyantTools, the mechanisms of formation, patterns of distribution, and pragmatic functions of abbreviations in contexts typical for this platform are investigated. The research methodology includes analysis of frequency, distribution, and context, which allows us to identify systematic patterns in the use of abbreviations. Functional analysis shows that abbreviations perform three main pragmatic functions: markers of epistemic modality, hedging devices that soften categorical statements, and indicators of in-group identity. The results of the study indicate the systematic nature of abbreviations in Twitter discourse and confirm the hypothesis of the formation of stable conventions of digital literacy, which are determined by the interaction between the technological limitations of the platform and the communicative needs of users. The study contributes to the understanding of language innovation processes in computer-mediated

communication and demonstrates the effectiveness of corpus methodology in the study of digital discourse phenomena.

Keywords: *Twitter/X discourse, abbreviations, digital communication, pragmatic functions, discourse markers, social media.*

Introduction. The digitisation of the communicative environment has led to the emergence of new forms of speech interaction, characterised by specific linguistic practices determined by the technological capabilities and limitations of social networks. *Twitter* (now renamed *X*), as one of the most influential microblogging platforms, demonstrates intense processes of linguistic innovation, among which abbreviation occupies a prominent place. Unlike traditional written genres, where abbreviations are used primarily to save space in technical or reference texts, in Twitter discourse, abbreviations take on complex pragmatic and sociolinguistic functions.

Despite the growing interest in studying the language of social networks, systematic analysis of abbreviations in Twitter discourse remains underdeveloped. Most existing studies focus on descriptive cataloguing of forms without detailed analysis of their distribution, functional variability, and role in structuring discourse. This necessitates the use of corpus methodologies that allow statistically significant patterns to be identified in a representative data set.

Formulation of the problem. The study of abbreviations in Twitter discourse is relevant for several reasons. First, traditional approaches to studying abbreviations, which focus primarily on orthographic and morphological aspects, are insufficient for analysing dynamic platform-specific practices where abbreviations perform complex pragmatic and sociolinguistic functions. Second, previous studies have been limited mainly to descriptive cataloguing of forms without systematic analysis of their distribution, functional variability, and diachronic dynamics. Thirdly, the methodological challenge lies in the need to process large amounts of data to identify statistically significant patterns, which requires the use of computer analysis tools.

The central research question is formulated as follows: what systemic linguistic mechanisms determine the formation, spread and functioning of abbreviations in English-language Twitter discourse? The operationalisation of this question involves considering three interrelated aspects: (1) the structural typology of abbreviations, (2) their pragmatic functions in different communicative contexts, and (3) the sociolinguistic parameters of variability.

Analysis of recent studies and publications. The theoretical basis of the study consists of works on digital linguistics, corpus linguistics, and discourse analysis. D. Crystal (2011) made a fundamental contribution to understanding the specifics of ‘Internet language’ by defining digital discourse as a separate variety that combines features of spoken and written language. According to D. Crystal, Internet communication ‘is neither recorded speech nor spoken writing, but a new form of speech with its own characteristics’ (Crystal, 2011). Research specifically on Twitter discourse is presented in the works of M. Zappavigna (2012, 2021), who applied a systemic-functional approach to the analysis of microblogging. M. Zappavigna (2012) argues that Twitter is characterised by ‘ambient affiliation’, i.e., a form of social kinship realised through shared evaluative positions marked by hashtags and other semiotic resources. However, the issue of abbreviations remains peripheral in these studies. The issue of abbreviations in digital communication has been examined in the context of studying the characteristics of text message language. S. A. Tagliamonte & D. Denis (2008) analysed a corpus of instant messages sent by teenagers and found that, contrary to stereotypes, the language of online communication is not linguistically impoverished, and that abbreviations are used selectively and perform sociolinguistic functions. In the field of corpus linguistics, the works of D. Biber & S. Conrad (2019) are relevant, who developed a methodology for multidimensional register analysis based on statistical study of linguistic features. This study aims to fill this gap by applying corpus methodologies to the systematic study of Twitter abbreviations.

The article **aims** to identify and systematically describe the pragmatic functions of abbreviations in English-language Twitter/X discourse based on a corpus analysis of 500 tweets.

To achieve this goal, the following tasks have been set:

1) to identify the main types of abbreviations used in Twitter/X discourse and determine their frequency using the Voyant Tools tool;

2) to conduct a functional analysis of the identified abbreviations to determine their pragmatic functions in the contexts of digital communication;

3) to identify systematic patterns of abbreviation use as markers of epistemic modality, means of hedging, and indicators of in-group identity;

The provided analysis is focused on abbreviations in English-language Twitter/X discourse as a linguistic phenomenon of digital communication.

Research results. An analysis of a corpus of 500 English-language tweets using VoyantTools revealed systematic patterns in the use of abbreviations, characterized by predictable structural types, distribution patterns, and pragmatic functions. The corpus contains 7,353 word occurrences and 1,933 unique word forms, providing a sufficient basis for identifying frequency and contextual patterns in digital discourse. The basic indicators of the corpus demonstrate the specific characteristics of Twitter discourse. Vocabulary density is 0.263, which indicates a relatively high level of lexical diversity, despite the limited volume of individual tweets. The readability index is 5.853, confirming the accessibility of the content to a wide audience. The average sentence length of 31.2 words exceeds that typical for informal written speech, which can be explained by the combination of several thoughts within a single tweet due to platform limitations.

The corpus was analyzed by extracting all identified abbreviations, classifying them by structural types and functional categories, and studying their positional distribution within tweets. Each tweet was analyzed for the presence of abbreviations, their number, position in the statement, and pragmatic context of use.

This corpus has 1 document with 7,353 total words and 1,933 unique word forms. Created now.

Vocabulary Density: 0.263

Readability Index: 5.853

Average Words Per Sentence: 31.2

Figure 1.1 Summary tool.

Quantitative analysis of the corpus demonstrated the high prevalence of abbreviations in Twitter discourse. The vast majority of tweets contained at least one abbreviation, confirming the central role of these forms in the platform's digital communication.

Figure 1.2 Summary tool.

Frequency analysis revealed fifteen different types of abbreviations, demonstrating an uneven distribution of usage. The most common forms were *tbx* (thanks), *ngl* (not gonna lie), *idk* (I don't

know), *lol* (laughing out loud), and *btw* (by the way). These five abbreviations account for about half of all recorded instances, indicating their central role in shaping Twitter's discursive practices.

The abbreviations *smh* (shaking my head), *BRB* (be right back), and *THX* (thanks) showed average frequency, appearing regularly in the corpus but with lower intensity compared to the most common forms.

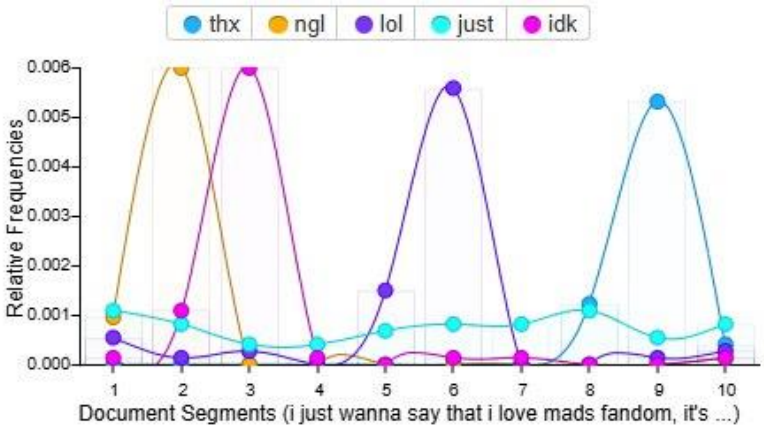

Figure 2.1 Frequency analysis.

Less frequent were community-specific abbreviations such as *IYKYK* (if you know you know), *IYKWIM* (if you know what I mean), *B4N* (bye for now), and *NVM* (never mind). Despite their lower frequency of use, these forms play an important role in constructing intra-group identity and demonstrating familiarity with platform conventions.

The frequency distribution shows a pattern characteristic of natural language: a small number of high-frequency forms coexist with a larger number of less frequently used variants. This pattern confirms the systematic, rather than random, nature of abbreviation use in Twitter discourse.

The functional classification was based on studying the contexts in which each abbreviation is used and identifying recurring communicative goals that are realized through these forms.

Markers of epistemic modality. It should be noted that abbreviations function as markers of epistemic modality, indicating the speaker's subjective assessment of the truthfulness and sincerity of their own statement. These forms do not add new propositional information to the statement, but rather modify the epistemic frame in which the message's content is interpreted.

A contextual analysis of *tbh* revealed that this abbreviation systematically accompanies the expression of personal opinions, assessments, and subjective judgments: "*Life without academic stress will be boring tbh*" (Twitter, 2025). The function of *tbh* is not so much to confirm the literal honesty of the statement as to mark it as a sincere personal opinion that may contradict the audience's generally accepted views or expectations.

Ngl demonstrates a similar but slightly different function. This abbreviation appears more often in contexts of acknowledging facts or emotional states that the author perceives as potentially uncomfortable or unexpected: *"The state of the world got me a little depressed ngl"* (Twitter, 2023). The function of *ngl* is to mark a statement as a sincere admission that the author is making despite some internal resistance or social conventions. Both abbreviations are systematically used to construct authenticity in digital discourse. In the context of social media, where users often present curated, idealized versions of their lives and opinions, *tbh* and *ngl* function as linguistic markers of the "real" self, signaling a temporary removal of social filters.

Hedging and softening categorical statements. The second group of abbreviations performs hedging functions — linguistically softening categorical statements and reducing the author's responsibility for the accuracy or universality of the statement. This category includes *idk*, *lol*, and *smh*.

Idk softens the categorical nature of statements: *"Idk why people hate Geminis so much, we're literally the best"* (Twitter, 2025). *Idk* functions as a rhetorical device to express surprise or

misunderstanding about a particular state of affairs: *"idk how i didn't realize fathers day is this weekend…… wow*"(Twitter, 2024). The form signals not a lack of knowledge as such, but cognitive disorientation or the unexpectedness of the situation for the author.

Lol demonstrates a significant functional evolution from its original literal meaning (laughing out loud) to an abstract marker of informality and a reduction in the seriousness of a statement. Analysis of the contexts in which it is used shows that in the vast majority of cases, *lol* does not indicate actual laughter or even a humorous perception of the situation. Instead, *lol* functions as a universal means of softening the tone of a message: *"Twitter is so boring now lol"*(Twitter, 2024). In such contexts, the abbreviation signals a relaxed, detached attitude toward the content of the statement, reducing its emotional weight and potential for confrontation.

Metadiscursive markers and identity indicators. The third functional category includes abbreviations that regulate discourse structure or construct intragroup identity. This group includes *btw*, *BRB*, *IYKYK*, and *IYKWIM*.

IYKYK and *IYKWIM* stand out among other abbreviations for their specific function of constructing intra-group identity. These forms appeal to a shared cultural code between the author and a certain part of the audience, creating a sense of exclusive community.

IYKYK accompanies references to specific cultural phenomena, historical events, or shared experiences that need no explanation for "those in the know": *"This weather reminds me of the Grambling game in 1980. IYKYK"*(Twitter,2024). The function of the abbreviation is not to convey information, but to construct a social boundary between those who understand the context and those who do not.

IYKWIM performs a similar function, but with a greater emphasis on implicit meaning: *"I love when people are a bit dumb but hate when they are stupid iykwim"*(Twitter, 2024). The abbreviation signals that the statement contains additional, indirect meaning that should be obvious to those who share a certain perspective or experience with the author.

Both forms function as markers of commonality and mutual understanding, simultaneously including part of the audience (those who understand) and excluding another part (those who do not understand). This creates a sense of intimacy and in-group solidarity, which is an important aspect of social interaction on Twitter.

Conclusions and prospects for further research. This study employed a corpus-based methodology, analysing a data-set comprising 500 English-language tweets using Voyant Tools. The findings demonstrate that abbreviations in Twitter/X discourse are a systematic linguistic phenomenon characterized by predictable pragmatic functions and patterns of use. A subsequent functional analysis revealed three main categories of abbreviation use which reflect key aspects of digital communication on this platform.

It is shown that the theoretical significance of these findings extends to several areas of linguistic research. For theories of language change, Twitter abbreviations are an example of rapid pragmatization processes in computer-mediated communication, where technological constraints interact with communicative needs, generating stable innovations in extremely short periods of time. The transformation of forms from literal abbreviations to discursive markers with specific pragmatic functions illustrates the mechanisms of grammaticalization and pragmatization in real time.

It follows that Twitter abbreviations constitute a structured subsystem of digital English, characterized by systematic pragmatic functions and functional specialization. The analyzed forms do not indicate a decline in language or communication skills, but rather demonstrate the language's ability to adapt to new communication conditions, creating effective means for performing repetitive pragmatic functions within the constraints of a specific platform.

Consequently abbreviations function as multifunctional tools that allow users to construct authenticity, regulate interpersonal relationships, mark epistemic position, soften the categorical nature of statements, and signal intragroup membership. Understanding these functions and patterns of their use contributes to a broader awareness of the evolution of language in the digital context and the emergence of new literacy practices in computer-mediated communication that deserve serious linguistic research.

REFERENCES

1. Baron, N. S. (2010). *Always on: Language in an online and mobile world*. Oxford University Press.
2. Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.
3. Crystal, D. (2011). *Internet linguistics: A student guide*. Routledge. https://doi.org/10.4324/9780203830901
4. Eisenstein, J. (2013, June). What to do about bad language on the internet. In Proceedings of the 2013 conference of the North American Chapter of the association for computational linguistics: *Human language technologies* (pp. 359-369).
5. Herring, S. C. (2013). Discourse in Web 2.0: Familiar, reconfigured, and emergent. *Discourse, 2*(0), 1-26.
6. Tagliamonte, S. A., & Denis, D. (2008). Linguistic ruin? LOL! Instant messaging and teen language. *American speech, 83*(1), 3-34.
7. Vessey, R. (2015). Zappavigna, M.(2012). Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web. London: Bloomsbury. In *Yearbook of Corpus Linguistics and Pragmatics 2015: Current Approaches to Discourse and Translation Studies* (pp. 295-299). Cham: Springer International Publishing.
8. Zappavigna, M. I. C. H. E. L. E. (2021). Discourse and social media. *The Bloomsbury handbook of discourse analysis*, 295-309.