НЕОЛОГІЗМИ В СУЧАСНІЙ АНГЛІЙСЬКІЙ ІТ-ТЕРМІНОЛОГІЇ: КОРПУСОБАЗОВАНИЙ АНАЛІЗ

NEOLOGISMS IN CONTEMPORARY ENGLISH IT TERMINOLOGY: A CORPUS-BASED ANALYSIS

Druzhbiak S.V.,
orcid.org/0000-0001-8971-9852
PhD,
Associate Professor at the Department of Applied Linguistics
Lviv Polytechnic National University

Byblyk A.V.,
orcid.org/0009-0003-5580-3310
Second-year Master's student at the Faculty of Linguistics
Lviv Polytechnic National University

This article develops an integrated framework for analysing the short-term diachronic evolution of English IT terminology and associated graphematic innovations. By bringing together established theories of lexical productivity with challenges specific to digital communication, the study highlights the interplay of multiple mechanisms that drive change. These include creative word-formation processes such as compounding, affixation, blending, and borrowing, alongside semantic extension and metaphorical reinterpretation. Orthographic practices such as camelCase, hashtags, re-hyphenation, and concatenated forms illustrate how digital platforms encourage experimentation with visual representation. In parallel, pragmatic and register shifts emerge as a consequence of platform-specific affordances, while multimodal resources such as emoji or interface symbols become increasingly relevant for meaning-making. The analysis also foregrounds genre diversification, showing how technical vocabulary circulates across forums, manuals, documentation, and mainstream media. A particular emphasis is placed on the interaction between bottom-up diffusion in communities of practice–such as developer forums–and top-down stabilization through institutional mechanisms, including termbases, standards, translation memories, and localization workflows. From these premises, the article derives methodological priorities for corpus research. It recommends systematic preprocessing of graphematic variants, joint use of normalized frequency and dispersion to track distribution, and the application of collocational and network measures for detecting semantic restructuring. The resulting framework provides not only a reproducible roadmap for corpus-based diachronic studies but also a set of practice-oriented guidelines for lexicographers, terminology managers, and localization specialists. Ultimately, the study demonstrates that IT terminology constitutes a dynamic system shaped by technological innovation, and it outlines a coherent agenda for documenting, testing, and understanding its ongoing transformation.

Key words: IT terminology, neologisms, corpus linguistics, orthographic innovation, multimodality, collocations, short-term diachronic analysis.

У статті розроблено інтегровану модель для аналізу короткодіахронної еволюції англомовної ІТ-термінології та пов'язаних із нею графематичних інновацій. Поєднання класичних теорій словотвірної продуктивності з викликами, що виникають у цифровому середовищі, дає змогу показати взаємодію різних механізмів мовних змін. До них належать новаторські словотвірні процеси, зокрема композиція, афіксація, блендинг і запозичення, а також семантичне розширення та метафорична переінтерпретація. Орфографічні практики на кшталт camelCase, хештегів, повторного дефісування чи конкатенації ілюструють, як цифрові платформи стимулюють експерименти з візуальною репрезентацією. Паралельно відбуваються прагматичні та реєстрові зрушення, зумовлені особливостями платформ, а також зростає роль мультимодальних ресурсів, зокрема емодзі чи інтерфейсних символів, у конструюванні значення. Увагу зосереджено й на жанровій диверсифікації, яка показує, як технічна лексика поширюється між форумами, документацією, посібниками та масмедійними ресурсами. Окремий акцент зроблено на взаємодії «знизу-вгору», що реалізується у спільнотах практики, наприклад на розробницьких форумах, і «згори-донизу», яку забезпечують інституційні механізми, зокрема термінологічні бази, стандарти, пам'яті перекладу та локалізаційні процеси. З цих позицій сформульовано методологічні пріоритети корпусних досліджень. Запропоновано систематичну попередню обробку графематичних варіантів, спільне використання нормалізованої частоти та дисперсії для відстеження поширення, а також застосування колокаційних і мережевих показників для виявлення семантичних перебудов. Запропонована модель формує не лише відтворюваний план для корпусних діахронних досліджень, але й практичні орієнтири для лексикографів, термінологів і фахівців з локалізації. У підсумку робота показує, що ІТ-термінологія є динамічною системою, зміненою під впливом технологічних інновацій, і пропонує цілісну дослідницьку програму для її документування та осмислення.

Ключові слова: ІТ-термінологія, неологізми, корпусна лінгвістика, орфографічні новації, мультимодальність, колокації, діахронічний аналіз.

Statement of the problem. The rapid expansion of information technologies has generated an unprecedented wave of lexical innovations in English. These innovations manifest not only in the constant emergence of new terms and abbreviations, but also in orthographic patterns such as camelCase, hashtags, and concatenated forms. Scholars such as David Crystal [7] emphasize that technological change has become a powerful driver of linguistic evolution, while Herring highlights the multimodal nature of computer-mediated discourse. Despite this recognition, systematic corpus-based studies of IT-related terminology remain limited, especially in terms of diachronic comparison across periods of rapid technological change. Moreover, the interaction between verbal innovations and multimodal markers such as emojis or interface symbols has not been sufficiently explored. Addressing these gaps is essential for understanding the mechanisms of lexical diffusion, stabilization of meaning, and the development of methodological tools for monitoring fast-evolving technical vocabularies[10].

Review of recent studies. The influence of technology on language development has been extensively discussed in linguistic scholarship. David Crystal in *Language and the Internet* (2006) argues that the Internet has become not only a new medium of communication but also a powerful catalyst of lexical change. He demonstrates how digital environments foster the rapid emergence of neologisms and non-standard forms, often bypassing traditional mechanisms of lexical codification [7].

Susan C. Herring in *Discourse in Web 2.0: Familiar, Reconfigured, and Emergent* (2013) stresses the multimodal nature of computer-mediated discourse, where textual, visual, and symbolic resources operate together [10]. Similarly, Crispin Thurlow and Kristine Mroczek in *Digital Discourse: Language in the New Media* (2011) highlight how online genres stimulate new linguistic practices and encourage the spread of hybrid multimodal signs. These insights are crucial for the study of IT-related neologisms, as many of them originate in digital platforms and are closely tied to multimodal user interaction [17].

From the perspective of word-formation and lexical innovation, Laurie Bauer in *Morphological Productivity [3]* and Jean Aitchison in *Words in the Mind[2]: An Introduction to the Mental Lexicon* (2013) underline the role of productivity, analogy, and cognitive factors in the creation and adoption of new lexical items. Kyo Kageura in *The Dynamics of Terminology: A Descriptive Theory of Term Formation and Terminological Growth* (2002) specifically examines how terminology evolves in sci-

entific and technical fields, drawing attention to the accelerating pace of terminological expansion. [11]

Corpus linguistics provides the methodological basis for tracing such changes. Michael Stubbs in *Words and Phrases: Corpus Studies of Lexical Semantics* (2001) demonstrates how frequency, collocation, and concordance analysis can be applied to identify semantic shifts and track lexical diffusion over time. [16]

Although these studies provide valuable theoretical and methodological frameworks, relatively few works have directly addressed the short-term diachronic, corpus-based analysis of IT-related terminology. Most research focuses either on sociolinguistic aspects of digital discourse or on individual lexical innovations without systematic comparison across time periods. Furthermore, the interaction of verbal neologisms with multimodal markers such as emojis or interface symbols remains largely unexplored. This gap underscores the importance of combining corpus methods with attention to multimodality in order to better understand how technological innovation reshapes the English lexicon.

Identification of previously unresolved parts of the overall problem addressed in this article

Despite the considerable scholarly attention to language change in the digital era, several aspects remain insufficiently investigated. First, existing research lacks systematic diachronic corpus analyses of IT-related terminology that would make it possible to trace the emergence, stabilization, and diffusion of new lexical items across different stages of technological development. Second, most studies concentrate either on the sociolinguistic features of online communication or on the description of individual neologisms, whereas comprehensive approaches that integrate quantitative indicators such as frequency, dispersion, and collocational patterns with qualitative concordance analysis are still rare. Third, the interaction of verbal neologisms with multimodal markers, including emojis and interface symbols, has so far received little attention, even though such elements play an important role in the stabilization of meaning and genre identification. Addressing these gaps constitutes the main focus of the present article.

The primary objective of this article is to examine the development of IT-related terminology and graphematic innovations in contemporary English through a corpus-based short-term diachronic approach. To achieve this aim, the following tasks are set.

1. To trace the frequency dynamics and dispersion of selected IT-related lexical items across two comparative time periods.

2. To identify and analyse collocational patterns

that indicate processes of semantic narrowing, broadening, or specialization.

3. To describe orthographic innovations, such as camelCase forms, hashtags, and concatenations, and to assess their role in lexical diffusion.

4. To investigate the interaction between verbal neologisms and multimodal markers (e.g., emojis, interface symbols) in the stabilization of meaning.

5. To develop and test a methodological workflow that combines quantitative corpus techniques with qualitative concordance analysis, ensuring reproducibility for future research in the field.

Technological innovation as a driver of lexical and orthographic change

Technological innovation has consistently acted as one of the most powerful engines of linguistic change. Every significant technological advance produces novel referents, communicative needs, and interactional formats that language must encode. This cycle is visible in the vocabulary of information technologies, where new devices and services open conceptual gaps that require immediate lexical filling. The coinage of items such as *smartphone*, *app*, or *blockchain* illustrates how fresh technical realities demand linguistic representation. As Laurie Bauer notes in *An Introduction to International Varieties of English* (2002), languages draw on their word-formation resources – compounding, affixation, blending, and borrowing – to meet such demands [4]. In this way, technological innovation supplies the *what* (new realities and referents), while language supplies the *how* (morphological creativity and lexical productivity) to encode them.

Semantic extension and metaphorical mapping

A second pathway of technological lexical innovation is semantic extension, often supported by metaphorical mapping. Jean Aitchison in *Language Change: Progress or Decay?* (2013) demonstrates how words frequently acquire new senses through figurative use that later becomes conventionalized [1]. In the technological sphere, this mechanism is particularly visible: the term *cloud* originally belonged to meteorological vocabulary but now denotes large-scale data storage; the noun *mouse* has shifted from a physical animal to a standard computer peripheral. Such developments create layers of polysemy in which older senses coexist with emerging technological senses. Over time, the new sense stabilizes and may eventually dominate in relevant registers. The process confirms Aitchison's observation that semantic broadening and narrowing often coexist as transitional phases, especially when driven by rapid external pressures such as technological adoption [1].

Orthographic and graphematic innovation

Technological environments also reshape orthography and graphic practice. The need to create distinctive brand names, programming identifiers, or social media tokens encourages non-standard grafematic forms. Examples include camelCase (*iPhone*, *eBay*), hashtags (*#hashtag*), and concatenated compounds (*email* replacing *e-mail*). Mark Davies in *The Corpus of Contemporary American English (COCA)* (2008) provides evidence that such forms migrate from niche channels – technical forums, code repositories, marketing campaigns – into mainstream genres, gradually altering orthographic norms [8]. Their diffusion demonstrates the pressure of technology on writing systems, producing observable changes that lexicographers and corpus linguists must account for during text preprocessing. The normalization of camelCase or hashtag forms in corpus data, for instance, has become an unavoidable methodological step, since these tokens directly affect frequency counts and dispersion measures [8].

Interactional formats and pragmatic tendencies

The rise of new communication platforms also modifies grammar and pragmatics. Search engines, messaging applications, and social networks encourage linguistic forms optimized for speed and brevity. John Sinclair in *Corpus, Concordance, Collocation* (1991) emphasized that usage patterns reflect recurrent communicative needs [14]. In digital environments, such needs favor imperative, truncated, or fragmentary expressions, as well as the verbing of nouns (e.g., *to google*, *to tweet*). These tendencies are not merely stylistic but are symptomatic of register and genre shifts that can be empirically observed in corpus studies. They indicate how technological affordances – short message constraints, character limits, or real-time interaction – drive pragmatic economy and reshape syntactic preferences [14].

Diffusion through new communication channels

Technological innovation accelerates the diffusion of neologisms by providing faster channels of communication. While earlier lexical items often spread gradually through print media and technical reports, the Internet and social networks enable rapid, global circulation. Arja Piirainen-Marsh in *Conversation Analytic Research on Learning-in-Action* (2019) shows how new lexical items are negotiated and routinized in situated interactions such as developer chats, gamer talk, or online tutorials [13].

These micro-interactional processes precede macro-level visibility in corpora and highlight how community-based uptake acts as the initial stage of lexical diffusion. The shortening of the interval

between coinage and mainstream adoption under-scores the unique temporal dynamics of technological neologisms [13].

Genre diversification and textual ecology

Another consequence of technological innovation is the diversification of discourse genres. The emergence of user manuals, API documentation, developer blogs, app store descriptions, and social-media microtexts has expanded the textual habitats where IT terminology appears. Douglas Biber, Susan Conrad, and Randi Reppen in *Corpus Linguistics: Investigating Language Structure and Use* (1998) stress that genre and register distinctions are crucial for interpreting frequency distributions [5]. A term may be highly common in developer forums but nearly absent from newspapers, and such disparities reveal different levels of register penetration. Therefore, dispersion across genres is as important as raw frequency in understanding the diffusion of technological vocabulary [5].

Institutional and terminological control

Technological neologisms also encounter pressures from institutions and standardization bodies. While grassroots usage favors creativity and informality, translation agencies, localization departments, and standards committees attempt to stabilize terminology for interoperability. Lynne Bowker in *Computer-Aided Translation Technology: A Practical Introduction* (2002) explains how tools such as translation memories and termbases contribute both to the acceleration of standardization (when updated promptly) and to the preservation of outdated variants (when not updated) [6]. This dual effect creates tension between bottom-up lexical change and top-down terminological control. The interplay illustrates that linguistic innovation in IT domains is shaped not only by user practices but also by institutional inertia [6].

Multimodality and semiotic extension

Technological communication is inherently multimodal. Emojis, icons, and interface symbols carry pragmatic and sometimes semantic meaning. Michael Stubbs in *Text and Corpus Analysis* (1996) notes that language must be studied in context, and in digital settings, context often includes non-verbal resources [15]. Emojis, for instance, can index affect, stance, or even replace words, functioning as lexical items in their own right. Their co-occurrence with technical terms adds a layer of semiotic complexity to corpus analysis, requiring either multimodal annotation or at least explicit recognition of such markers [15]. Ignoring these tokens risks underestimating the ways in which meaning is distributed across text and visual signs in online discourse.

International diffusion of IT terminology

English acts as the lingua franca of information technology, and this status accelerates the global spread of IT terminology. Kyo Kageura in *The Dynamics of Terminology: A Descriptive Theory of Term Formation and Terminological Growth* (2002) and *The Quantitative Analysis of the Dynamics and Structure of Terminologies* (2012) demonstrates how technical terms evolve rapidly and cross linguistic boundaries [11; 12]. Many new IT terms are borrowed directly into other languages in their English form or are slightly adapted phonologically and orthographically. This international diffusion complicates the task of terminologists and translators, since the boundaries between "loanword," "calque," and "brand name" are often blurred [12]. The global reach of English IT vocabulary illustrates the interplay between linguistic change, technological innovation, and sociolinguistic power relations [12].

Toward an integrated framework

Synthesizing these observations, the literature suggests a set of mechanisms through which technology shapes vocabulary and discourse [4; 1; 8; 14; 13; 5; 6; 15; 11, 12]:

1. **Lexical creation through morphological productivity** (compounding, affixation, borrowing).

2. **Semantic extension and metaphorization** leading to layered 1. polysemy.

3. **Orthographic and graphematic innovation** such as camelCase, hashtags, and hyphenation shifts.

4. **Pragmatic economy and register change** influenced by interactional affordances.

5. **Accelerated diffusion through online channels** that shorten the path from coinage to adoption.

6. **Genre diversification** that multiplies the textual habitats of new terms.

7. **Institutional control vs. grassroots creativity** creating tension in terminological standardization.

8. **Multimodal enrichment** where emojis and icons supplement or substitute verbal items.

9. **Internationalization of IT vocabulary** facilitated by English as a global lingua franca.

These mechanisms form a theoretical framework for studying the short-term diachronic evolution of IT terminology in English. While previous scholarship has described individual aspects – for example, Bauer on morphological productivity, Aitchison on semantic change, or Bowker on terminology management – a systematic integration of these mechanisms within a single analytic model has been lacking. The present article addresses this gap by synthesizing insights across linguistic subfields and aligning them with corpus-based methodology.

Implications for future empirical research

Although this article focuses on theoretical synthesis, the framework generates concrete methodological priorities for future corpus analysis [15; 8; 5]:

• **Preprocessing orthographic innovations** (e.g., normalize camelCase and hashtags) before calculating frequency and dispersion.

• **Measuring dispersion across genres** to capture the register penetration of IT terms.

• **Tracking collocational shifts** to detect semantic specialization or metaphorization.

• **Inspecting concordances** for early metaphorical or figurative uses that precede stabilization.

• **Considering multimodal tokens** in annotation schemes to represent the full semiotic environment.

• **Comparing international variants** to assess borrowing and adaptation patterns.

By articulating these priorities, the theoretical framework not only consolidates existing knowledge but also prepares the ground for systematic empirical investigation.

Disappearance and semantic narrowing / widening of terms

Words do not only appear: they also decline in use or shift their senses, and these two phenomena – lexical disappearance (attrition) and semantic narrowing or widening – are central to understanding the life cycle of technical vocabulary.

Jean Aitchison in *Language Change: Progress or Decay?* (2013) emphasises that language change is an adaptive, socially driven process in which items fall out of favour as the referents they name cease to be relevant or as alternative labels supplant them [1]. She cautions that disappearance should not be interpreted as linguistic "decay" but as a normal consequence of cultural and technological shifts. In the IT domain, disappearance is often rapid and clearly traceable: when a technology is replaced, the lexical items closely tied to it (e.g., *floppy disk*, *dial-up*, *VCR*) decline in frequency because the artefacts and practices that licensed their use are no longer part of everyday life [1].

Several causal mechanisms underlie lexical disappearance in technological vocabularies. The most direct is **technological obsolescence**: when an object or procedure ceases to be produced or widely used, references to it decline correspondingly. A second mechanism is **lexical replacement**, whereby new concepts or more economical labels supplant older ones. For instance, *USB drive*, *external hard drive*, and ultimately *cloud storage* have successively displaced *floppy disk* as dominant referents for portable storage. A third mechanism is **institutional and market influence**: corporate branding, standardiza-

tion bodies, and dominant platforms can accelerate adoption of particular terms while marginalising competing alternatives; conversely, preservationist institutions such as dictionaries, translation memories, and standards committees can slow disappearance by continuing to record or circulate older forms [6;9]. Finally, **sociolinguistic prestige and reach** matter: terms that remain confined to small communities of practice (e.g., specialised developer forums) are more likely to remain marginal and eventually vanish from broader registers.

Semantic narrowing and widening represent complementary processes that reshape the conceptual scope of surviving items. **Narrowing (specialization)** occurs when a general word acquires a more restricted technical sense. A canonical historical case is *computer*, which shifted from denoting a person who computes to denoting a machine that computes [1]. Similar processes affect IT nouns such as *server*, *router*, and *tablet*, where common words are recruited to designate specific devices or roles. **Widening (generalization)**, by contrast, expands a term's application: *window* broadened from an architectural opening to include graphical user interface (GUI) elements, and now any framed digital display can be described using the term. Metaphor and metonymy frequently catalyse these shifts: for example, *cloud* transferred from meteorological imagery to remote storage, then generalized to a wide range of services (*cloud storage, cloud provider, cloud migration*) [1].

Corpus linguistics provides concrete indicators for these semantic processes. Narrowing tends to be accompanied by a contraction of the collocational profile: the target word increasingly co-occurs with a small set of technical partners (e.g., *server* with *host, http, cloud, container*) while earlier, more general associations decline.

Widening, by contrast, is signalled by a broadening of collocates and dispersion across registers: a word that once appeared mainly with domestic or physical environment terms begins to appear with computing, corporate, or media vocabulary.

Michael Stubbs in *Text and Corpus Analysis* (1996) argues that such collocation patterns are among the most reliable corpus signals of semantic change [15]. His emphasis on examining authentic concordance lines is crucial: statistical tendencies must be confirmed qualitatively by reviewing actual usage contexts.

Quantitative perspectives complement these insights. Kyo Kageura in *The Quantitative Analysis of the Dynamics and Structure of Terminologies* (2012) shows how network centrality and clustering measures can capture semantic reconfiguration, while

growth and decline curves quantify the tempo of lexical change [12]. In his framework, the replacement of one item by another can be visualised as the transfer of centrality in a co-occurrence network, while logistic S-curves fitted to normalized frequency trajectories model the dynamics of adoption and attrition.

On this basis, a practical typology of corpus indicators can be outlined for future empirical research:

• **Frequency decline / disappearance:** A sustained proportional drop in normalized frequency (tokens per million) across periods, often >90%, accompanied by low dispersion (concentration in very few documents) and contexts marked as historical or retrospective (e.g., "used to be stored on a floppy disk"). These criteria align with lexicographic practice for declaring items obsolete [9].

• **Lexical replacement signatures:** Parallel timelines showing the rise of a candidate replacement while the older term falls; network analysis demonstrates the new item occupying the functional neighbourhood of the obsolete one. Logistic diffusion curves can be used to quantify these dynamics [12].

• **Narrowing (specialization) evidence:** Shift from diffuse collocations across everyday domains to dense collocations with technical lexis, reduced polysemy in concordance samples, and contraction in genre spread (concentration in manuals, standards, journals). Manual KWIC inspection confirms the disappearance of non-technical senses [15].

• **Widening (generalization) evidence:** Rising dispersion across genres, greater diversity of collocational partners, and conventionalisation of metaphorical or metonymic uses. Frequency growth in both unigrams and multiword expressions, together with

appearance in mainstream media and user-oriented documentation, indicates diffusion beyond specialist circles [1].

Together, disappearance, narrowing, and widening constitute essential dimensions of the lexical life cycle in technology-driven vocabularies. Integrating theoretical insights from Aitchison, Bowker, Gouws, Stubbs, and Kageura within a corpus-based framework provides a robust agenda for analysing how IT terminology emerges, shifts, and declines over time.

Conclusions This study shows that technological innovation fundamentally reshapes the vocabulary and writing practices of modern English. The proposed framework identifies how new terms emerge, stabilize, shift in meaning, and eventually decline under the influence of digital technologies. It also demonstrates that corpus-based methods, when combined with attention to orthographic variants and multimodal features, provide an effective basis for tracing these processes. The findings underline the importance of treating IT terminology as a dynamic system and point to practical applications for lexicography, terminology management, and language technology.

Prospects for further research. The conceptual framework points to several clear and high-impact avenues for future empirical work. Each suggested direction includes methodological recommendations and concrete research questions that would operationalize the framework into testable studies: short-term diachronic corpus validation of diffusion and attrition, preprocessing standards and reproducible pipelines, multimodal annotation and functional classification of emojis/UI tokens.

REFERENCES:

1. Aitchison, Jean. *Language Change: Progress or Decay?* 4th ed. Cambridge: Cambridge University Press, 2013. 368 p.

2. Aitchison, Jean. *Words in the Mind: An Introduction to the Mental Lexicon.* 4th ed. Oxford: Wiley-Blackwell, 2013.

3. Bauer, Laurie. *Morphological Productivity.* Cambridge: Cambridge University Press, 2001.

4. Bauer, Laurie. *An Introduction to International Varieties of English.* Edinburgh: Edinburgh University Press, 2002. 192 p.

5. Biber, Douglas; Conrad, Susan; Reppen, Randi. *Corpus Linguistics: Investigating Language Structure and Use.* Cambridge: Cambridge University Press, 1998. 320 p.

6. Bowker, Lynne. *Computer-Aided Translation Technology: A Practical Introduction.* Ottawa: University of Ottawa Press, 2002. 288 p.

7. Crystal, David. *Language and the Internet.* Cambridge: Cambridge University Press, 2006.

8. Davies, Mark. *The Corpus of Contemporary American English (COCA).* 2008. Available at: *https://www.english-corpora.org/coca/* (accessed: 17.09.2025).

9. Gouws, Rufus H.; Heid, Ulrich; Schweickard, Wolfgang (eds.). *Lexicography and Terminology: A Worldwide Perspective.* Amsterdam; Philadelphia: John Benjamins, 2013. 370 p.

10. Herring, Susan C. "Discourse in Web 2.0: Familiar, Reconfigured, and Emergent." In: Tannen, Deborah; Trester, Anna Marie (eds.) *Discourse 2.0: Language and New Media.* Washington, DC: Georgetown University Press, 2013.

11. Kageura, Kyo. *The Dynamics of Terminology: A Descriptive Theory of Term Formation and Terminological Growth.* Amsterdam; Philadelphia: John Benjamins, 2002. 254 p.

12. Kageura, Kyo. *The Quantitative Analysis of the Dynamics and Structure of Terminologies.* Amsterdam; Philadelphia: John Benjamins, 2012. 340 p.

13. Piirainen-Marsh, Arja. *Conversation Analytic Research on Learning-in-Action.* Helsinki: Finnish Literature Society, 2019.

14. Sinclair, John. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press, 1991. 179 p.

15. Stubbs, Michael. *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture.* Oxford: Blackwell, 1996. 272 p.

16. Stubbs, Michael. *Words and Phrases: Corpus Studies of Lexical Semantics.* Oxford: Blackwell, 2001.

17. Thurlow, Crispin; Mroczek, Kristine (eds.). *Digital Discourse: Language in the New Media.* Oxford: Oxford University Press, 2011.