МІЖНАРОДНИЙ ФІЛОЛОГІЧНИЙ ЧАСОПИС

Засновник:
Національний університет біоресурсів і природокористування України

Рік заснування: 1997
Періодичність випуску: щоквартально

Рекомендовано до друку та поширення
через мережу Інтернет Вченою радою
Національного університету біоресурсів і природокористування України
(протокол № 2 від 23 вересня 2025 р.)

Ідентифікатор медіа: R30-02294
(Рішення Національної ради України з питань телебачення
і радіомовлення № 1795, протокол № 31 від 21.12.2023)

Журнал входить до переліку
наукових фахових видань України

Категорія «Б». Галузь знань: 011 – Освітні, педагогічні науки; 035 – Філологія
(Наказ Міністерства освіти і науки України від 17 березня 2020 р. № 409)

Журнал представлено у міжнародних наукометричних базах даних, репозитаріях та пошукових системах: Google Scholar,
Національна бібліотека України імені В. І. Вернадського,
BASE, Ulrichsweb, Dimensions, UCSB Library,
OUCI (Open Ukrainian Citation Index), Litmaps

Адреса редакції:
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
E-mail: ph@philologicalscience.com.ua
https://philologicalscience.com.ua/uk

INTERNATIONAL JOURNAL OF PHILOLOGY

Founder:
National University of Life and Environmental Sciences of Ukraine

Year of foundation: 1997
Frequency: quarterly

Recommended for printing and distribution
via the Internet by the Academic Council
of National University of Life and Environmental Sciences of Ukraine
(Minutes No. 2 of September 23, 2025)

Media identifier: R30-02294
(Decision of the National Council of Television
and Radio Broadcasting of Ukraine No. 1795, Minutes No. 31, dated 21.12.2023)

The journal is included in the list
of Professional Scientific Publications of Ukraine
Category "B". Field of knowledge: 0111 – Education science;
0231 – Language acquisition
(Order of the Ministry of Education and Science of Ukraine of March 17, 2020, No. 409)

The journal is presented in the international scientometric databases, repositories and scientific systems: Google Scholar,
Vernadsky National Library of Ukraine,
BASE, Ulrichsweb, Dimensions, UCSB Library,
OUCI (Open Ukrainian Citation Index), Litmaps

Editors office address:
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
E-mail: ph@philologicalscience.com.ua
https://philologicalscience.com.ua/en

Редакційна колегія

Editorial Board

Зміст

Contents

International Journal of Philology, 29(3), 2025
Journal homepage: https://philologicalscience.com.ua/en
Received 18.05.2025 Revised 21.08.2025 Accepted 23.09.2025

UDC 81'373+563.7 DOI: 10.31548/philolog/3.2025.09

Morphological analysis of English-language biotechnology terms of the genomic era

Olena Syrotina*
PhD in Pedagogical Sciences, Associate Professor
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
https://orcid.org/0000-0002-4802-4891

Abstract. The relevance of this study is determined by the growing importance of English-language biotechnology terminology in the context of the rapid development of genomic technologies, which necessitates a systematic analysis of its morphological structure to enhance scientific communication and deepen the understanding of cognitive processes in term formation. The aim of the study was to analyse the morphological structure of English-language biotechnology terms of the genomic era through the lens of the cognitive approach. To achieve this goal, the following methods were used: lexicographic, morphological, and statistical analysis, elements of the cognitive approach to terminology research, as well as methods of systematisation and generalisation. It was found that English-language biotechnology terminology actively employs term formation models involving nouns, adjectives, and participles. The vast majority of terms are represented by nouns (N), with the most common being suffixal and prefix-suffixal formations, which together account for approximately 46% of the entire term base. The distribution of terms by conceptual categories within the N morphological model shows that the largest portion is made up of objects (54%)-key elements of biotechnology research such as genomes, proteins, and cells-followed by tools (19%)-modern methods, devices, and technologies-processes (17.2%)-dynamic actions such as sequencing or editing-and goals/results (9.8%)-final products or intended outcomes. Each of the analysed structures (N, N + N, A + N, A + A + N, A + A + A + N, V-ed + N, V-ing + N) was found to be associated with specific aspects of the conceptual domain: noun-based models predominantly represent the object of study (e.g., gene expression, protein synthesis), adjective-noun constructions typically refer to tools and characteristics of processes (e.g., synthetic genomic technologies), while participial forms

Suggested Citation:
Syrotina, O. (2025) Morphological analysis of English-language biotechnology terms of the genomic era. *International Journal of Philology*, 29(3), 9-28. doi: 10.31548/philolog/3.2025.09.

*Corresponding author

verbalise the results of actions, altered states, or acquired properties resulting from biotechnological interventions (e.g., engineered organism, modified gene), as well as functional features of objects (e.g., binding protein, binding site, coding sequence). Morphological analysis revealed not only the structural features of terms, but also the underlying conceptual models that reflect the patterns of scientific thinking in biotechnology and the cognitive-linguistic specificity of this dynamic field. The practical significance of the study lies in the possibility of applying the obtained results in teaching English for specific purposes to students majoring in "Biotechnology and Bioengineering"

Keywords: biotechnology terminology system; derivational structure; morphemic organisation; concept; cognitive approach

Introduction

Biotechnology is one of the most dynamic and interdisciplinary fields of knowledge, integrating advances in biology, chemistry, medicine, computer science, and engineering. This nature determines the constant updating of the terminological base, the emergence of new concepts, and the active development of the terminosystem. Despite the active development of biotechnological terminology, in particular English-language terminology, its morphological structure remains insufficiently investigated from the standpoint of a cognitive approach. Awareness of the lack of research on this issue justifies the relevance of analysing biotechnological terminology within the cognitive paradigm. Focusing on the morphological analysis of biotechnological terms of the genomic era is quite reasonable, given a number of key factors. During this period, there was a rapid increase in the number of terms in biotechnological discourse, due to revolutionary discoveries in the fields of genomics, genetic engineering, synthetic biology, and bioinformatics.

During 2000-2020, English-language biotechnological terminology became the object of increased interest from linguists. Ye. Striuk (2020) reviewed the cognitive model of the concept of biotechnology in contemporary English, analysing in detail its structural components and ways of representation in scientific discourse. The researcher stressed the importance of combining conventional terminology with cognitive interpretations, which contributed to a deeper understanding of the role of terms in the conceptual system of science. It was concluded that biotechnology is a flexible, prototype concept that does not lend itself to an adequate description due to classical definitions with fixed features. Therefore, an effective description of such concepts, according to the researcher's observations, requires alternative cognitive approaches that consider contextuality, interdisciplinarity, and variability of meaning. Consequently, cognitive terminography opens up new possibilities for describing complex scientific concepts, which allows considering both formal and content-cultural features of categorisation in scientific discourse.

O. Syrotina (2020) focused on the linguistic and cognitive aspects of metaphorisation in the English-language terminology of biotechnology, emphasising its role as a cognitive mechanism for creating new terms. The researcher came to the conclusion that one of the most productive types of metaphors is the anthropomorphic metaphor, when biotechnological terms are formed by analogy with a person or human activity, reflecting an active rethinking of scientific concepts through the prism of metaphorical images. The researcher also stressed that metaphorisation performs not only an explanatory, but also an innovative function, providing flexibility and adaptability of terminology to the rapid development of scientific

knowledge, technologies and interdisciplinary communication. Notably, the cognitive-discursive approach to the analysis of the biotechnological term system was deepened in the study by S. Molina-Plaza & A. Martínez-Sáez (2024), who investigated the functioning of metaphors in scientific texts on biotechnology. The researchers showed that metaphors not only perform a stylistic or heuristic function, but also actively participate in the conceptualisation of knowledge, contributing to the understanding of complex biotechnological processes through more understandable images. Metaphorical structures act as a tool of cognition, are fixed in scientific discourse and eventually become components of the term system, which confirms the expediency of a cognitive approach to analysing the morphological organisation of terms.

T. Almashadni & A. Jawad (2025) investigated the semantic content of the latest terms in biotechnology using methods of collocation analysis, discursive analysis, and semantic field theory. Discursive analysis has shown that these terms function in a variety of discourses – scientific, public, ethical, and political – often accompanied by conflicting interpretations. Through the application of semantic field theory, it was possible to group the latest terms into broader thematic domains, such as genetics, bioengineering, and medical biotechnology, reflecting the interdisciplinary nature and multifactorial nature of contemporary biotechnological terminology. The study confirmed and expanded on previous scientific developments in this area, while offering a new approach to the analysis of terms in both scientific and public speech. However, the study had certain limitations: the volume of the case was relatively small, and the selection of terms partly depended on the availability of sources. F.R. Adambaeva & N.A. Sadullayeva (2023) analysed eponym terms as part of biotechnological terminology, highlighting the historical and functional aspects of such terms and showing how personal names and discoveries become lexicalised in scientific discourse. In a follow-up study, F.R. Adambaeva (2025) investigated semantic hierarchies in biotechnology terminology in English and Uzbek, in particular, the process-object and property-object relationships. Both papers emphasised the importance of systematising terms for the accuracy of scientific communication, facilitating translation, and developing bilingual dictionaries, which contributes to a better understanding of specialised concepts in biotechnology.

All these studies show that terms in biotechnology are not only tools of nomination, but also means of conceptualisation, social impact, and interdisciplinary communication. In the context of the active development of scientific communication and the expansion of international terminological interaction, there is an increasing need to analyse terms not only from the standpoint of their formal organisation, but also considering cognitive function. Morphological analysis, which covers part-language affiliation, word-forming mechanisms, types of morphemes and models of creation, allows identifying how language represents scientific knowledge, how the structure of the term correlates with mental processes and categorical schemes of scientific thinking. It is in this context that morphological analysis of the terms of the genomic era, which was a turning point in the development of the biotechnological industry, becomes particularly important.

The purpose of the study was to identify morphological features of English-language biotechnological terms of the genomic era, considering their cognitive representation, and to analyse the part-language structure of terms and conceptualisation mechanisms that reflect the specifics of contemporary scientific thinking in the field of biotechnology. Achieving this goal involved solving the following tasks: perform morphological analysis of English-language biotechnological terms of the genomic era with the definition of their part-language affiliation; identify productive morphological models and

describe the typical structural organisation of terms; interpret established morphological models through the prism of a cognitive approach to identify the concepts that they represent.

Materials and Methods

In this study, the terms that characterise genomic era were selected according to a number of well-defined criteria. This category includes terms that directly relate to technologies and approaches that have emerged or are widely used, in particularб CRISPR-Cas technologies, new generation sequencing, synthetic biology, etc. The sample also included tokens related to post-genomic areas of modern science – genomics, transcriptomics, proteomics, metabolomics – and advanced practices in genetic engineering, genetic diagnostics, and gene therapy. Considerable attention was paid to the availability of terms in leading sources on biotechnology – dictionaries and glossaries. In particular, the Glossary of Biotechnology and Genetic Engineering (Zaid *et al.,* 2001) and Glossary of Biotechnology Terms (Kimball, 2002). The paper also considered the materials of the Encyclopaedia of Biotechnology (n.d.), Merriam-Webster (n.d.), and the American Heritage Dictionary of the English Language (2022). In addition, the Glossary of Genomics Terms (n.d.), National Human Genome Research Institute (n.d.a; n.d.b), and Genomics Glossary (n.d.) were analysed. Modern online resources dedicated to the creation of genetically modified organisms and biotechnological products were also considered. By drawing on a wide range of sources, from classic reference works to up-to-date glossaries of the genomic era, the authors compiled a representative sample of over 400 English-language terminology units. Thus, the terminology included in the analysis reflected the current state and dynamics of the development of biotechnological language in the context of genomic and post-genomic research.

In the process of studying the morphological features of English-language biotechnological terms of the genomic era, considering their cognitive representation, lexicographic, morphological and statistical analysis, elements of a cognitive approach to the study of terms and methods of systematisation and generalisation were used. At the first stage of the study, a corpus of English-language biotechnological terms representing key concepts and processes of the genomic era (in particular, CRISPR technologies, new generation sequencing, gene therapy, proteomics, genome editing, etc.) was formed. For the selection of lexical units, the lexicographic method was used, which performed a double function: it served as a tool for the initial selection of terms from authoritative industry dictionaries and glossaries and ensured the recording of definitions and verification of the terminological status of units. The latter involved checking whether the token is an established term in the biotechnological term system, or whether it functions only in certain contexts of specialised scientific use. The final sample included only those terms that were certified in at least two independent professional sources with a clear interpretation and stable subject reference. At the second stage, morphological encoding of terms was performed. The use of the morphological method at this stage involved determining the part-language affiliation of each component of the term, establishing the morphological structure, and distributing terms according to standard models that reflect their internal grammatical and semantic organisation. Based on the data obtained, terms were classified according to morphological models such as N + N (for example, gene therapy), A + N (genetic material), V-ed + N (modified organism) and V-ing + N (binding site), and which grammatical categories are most productive in the process of term formation in the English-language biotechnological term system of the genomic era. At the third stage of the study, a cognitive approach was applied, which allowed interpreting the results of morphological analysis from the standpoint

of cognitive linguistics. This approach considered language as a tool for representing, conceptualising and categorising knowledge about the world, and allowing the study to identify how term elements and their morphological structures relate to conceptual categories relevant to a particular scientific field. Within the framework of cognitive interpretation, each of the identified morphological models was compared with the conceptual categories that form the structure of the concept of "genomic era biotechnology". Methods of systematisation and generalisation at the final stage provided systematisation of the obtained data, identification of repeated structural models, and formulation of typical patterns of term formation in the English-language biotechnological term system. This synthesis of morphological, generalising and cognitive analysis helped to identify deep connections between the formal organisation of terms and the conceptual structure of knowledge that they represent.

Results and Discussion

Although biotechnology as a science was formed at the end of the 20th century, its terminological and practical foundations have a deep historical background, originating from the research of the heredity of Gregor Mendel in the 19th century and the use of microorganisms in fermentation processes, known since the 4th millennium BCE. This was highlighted in the paper by S. Rajak (2018), who noted that there are three main stages in the development of biotechnology: ancient, classical, and modern. It was modern biotechnology that became the basis for the development of the genomic era, which began in the second half of the 20th century and gained rapid development at the beginning of the 21st century. The genomic era became a defining stage in the development of scientific terminology, since its term system was not only replenished with new words, but also underwent a significant internal transformation: the principles of word formation changed, elements from different disciplines were combined, and the set of word-forming models was expanded.

To identify the conceptual features of the concept of "biotechnology of the genomic era", it is advisable to refer to the concept of concept. The term "concept", which has many scientific definitions, is widely used in the social sciences and humanities – such as literary studies, philosophy, linguistics, psychology, and cultural studies – and is also the basic unit of linguoconceptology (Yuldashev & Haydarov, 2022). O. Selivanova (2006) considered the concept as "an information structure of consciousness, an organised unit of memory that contains a set of knowledge about the object of cognition, verbal and nonverbal, acquired through the interaction of the functions of consciousness and the extra-conscious". This definition focuses on the cognitive nature of the concept, its connection with the conscious and subconscious experience of a person, and the multi-layered semantic content.

Since the concept of "biotechnology" functions mainly in the field of scientific discourse, it was appropriate to turn to the category of scientific (professional) concept, which helps to more deeply reveal the features of the term system and the knowledge content behind the corresponding term. The development of a scientific concept is carried out in the process of accumulation and development of special knowledge in a particular field, and the linguistic representation of concepts is terms. According to this opinion, terms are the result of cognitive activity, a reflection of professional concepts, and simultaneously – a link in communication between experts in a particular field. Thus, the term acts as a linguistic form of the concept, its sign expression, which ensures accurate and effective transfer of knowledge within the framework of scientific discourse. In this approach, the scientific concept is

considered not only as a set of knowledge, but as a metasense-value – that is, an important, meaningful term concept that is relevant for the subject of knowledge in specific circumstances. The scientific concept is the most important means of forming and developing a particular scientific field. Accordingly, the main function of a scientific concept is to represent knowledge, experience, meanings, associations, and scientific concepts that are most relevant to science or the scientific paradigm.

To determine the conceptual features of the concept of "biotechnology of the genomic era", the definitions of the term genomics in leading sources were analysed, which helped to outline the boundaries and content of this concept. According to the Merriam-Webster (n.d.), genomics is a branch of biotechnology concerned with applying the techniques of genetics and molecular biology to the genetic mapping and DNA sequencing of sets of genes or the complete genomes of selected organisms, with organising the results in databases, and with applications of the data (as in medicine or biology). In the Genomics Glossary (n.d.) emphasises that genomics covers the identification and characterisation of all genes and functional elements of the genome, and the analysis of their interaction. In the glossary of the National Human Genome Research Institute (n.d.a), genomics is defined as the branch of biology that studies the entire DNA set of an organism – its genome. Key aspects include the identification and characterisation of all genes and functional elements of the genome, and the study of their interaction.

Based on a definitive analysis of the term "genomics" in leading sources, genomic era biotechnology is defined as an interdisciplinary field of biotechnology that emerged at the intersection of molecular biology, genetics, genomics, and bioinformatics. Within this field, genomic information is considered as a key object of research and practical application. It is based on technologies of high-performance sequencing, genome editing (in particular CRISPR/Cas), synthetic biology and digital modelling and is aimed at developing innovative solutions in the field of medicine, agriculture, industry, and ecology. As a result of the definitive analysis of the concept "biotechnology of the genomic era", it became possible to determine its cognitive features: the use of genomic information as an object of research and practical application; the use of innovative technologies for sequencing, editing, and synthetic biology; integration of biological and digital approaches; orientation to the creation of new products and solutions in applied areas. The identified cognitive features allow structuring the concept of "biotechnology of the genomic era" by identifying relevant concepts in its concept sphere: the object is the genome and related information; the process is research, sequencing, editing and use of genomes; the goal/result is the creation of new biotechnological products and solutions; the tool is modern methods of genomics, bioinformatics, and synthetic biology.

Analysis of the research material proved that certain concepts are verbalised by most of the identified term-forming models, which include in their morphological composition the same parts of speech – noun, adjective, and verb (in participial form). Based on the results of morphological analysis, it was found that the structure of the English-language biotechnological terminological system is represented by the following models: N, N, N + N, A + N, A + A + N, A + N + N, A + A + A + N, and P + N, where N – noun, A – adjective, P – participle (verb in the form of a participle). Nouns play a key role in building the term system of the English biotechnological language, since they verbalise most industry concepts and are part of the structure of almost all terminological models. K. Deroey (2007) showed that noun compounds in scientific and technical texts in English occur significantly more often than in texts of a general thematic nature, which determines their

special functional load. This phenomenon has attracted the attention of researchers who have studied noun compounds from various scientific approaches, in particular: semantics (Bogachyk & Bihunov, 2020) and cognitive studies (Matvieieva & Torgovets, 2022), corpus linguistics (Bauer & Renouf, 2001), morphology and syntax (Junya, 2017).

Nouns fall into various lexical and grammatical categories: concrete, abstract, and proper nouns. In linguistic research, specific nouns are considered as nouns denoting physical objects or creatures that can be directly perceived through the senses. According to R. Matthews (2014), a specific noun is "a noun denoting a physical entity, an object that can be perceived by the senses". Specific nouns refer to various objects or beings that are key to human understanding. They contain in their semantics information about the properties, appearance, functions, and relationships of these objects with others, which is important for the categorisation process. In the biotechnological terminology of the genomic era, terms expressed by specific nouns verbalise the concept of "object". Examples of such terms: allele, codon, exon, gene, genome, histone, homolog, kinase, intron, neuron, vector, pathway, plasmid, promoter, transposon (Kimball, 2002). Specific nouns allow scientists to uniquely identify objects and results in biotechnological activities, ensuring accuracy and logical consistency in the development of a terminological system. To verbalise the concept of "goal/result" in biotechnology, terms expressed by specific nouns are also actively used. They usually denote the final effect, achieved goal or product of a biotechnological process, giving the terms clarity, concreteness and completeness, which is extremely important for the accuracy of the scientific term system. Examples of such nouns are the following terms: construct, mutant, strain, clone, biomarker, isolate, metabolite, knockout.

Specific nouns in the N + N model are widely used to verbalise the concepts of "object" and "process" in genomic era biotechnology. This model ensures accuracy and unambiguity of terminology, allowing scientists to clearly identify objects, phenomena, and processes. Examples of terms for objects: protein complex, plasmid vector, antibody molecule, lipid vesicle (Kimball, 2002), cell line, DNA sequence, membrane protein, virus particle (Zaid *et al*., 2001). The concept of "process" is verbalised in biotechnological terminology through terms such as gene expression, protein synthesis, virus replication, cell division. A special role in this is played by the gerund, which, acting as a noun, is one of the most common morphological means of expressing processality. It provides: nominativity – naming processes as independent objects (cloning, sequencing, culturing) and consistency – including complex terms in the structure, for example: genome editing, genetic engineering, gene cloning, molecular cloning, protein cloning, cell culture, tissue culture (Zaid *et al*., 2001). Thus, the use of specific nouns in the N + N model is an effective morphological means of nominating key concepts of biotechnological terminology, which allows clearly reflecting both static objects and dynamic processes, and accurately reflecting the relationships between objects and their characteristics or processes in the scientific discourse of biotechnology.

Abstract nouns are names of immaterial entities that denote an abstract feature, action, or general concept. Most often, these are nouns that refer to phenomena or processes that do not have a real physical embodiment and are perceived by the imagination. In the English terminology of biotechnology, abstract nouns are used to verbalise the concept of "process". Abstract nouns are used to represent concepts related to processes, functions, properties, or states of biological systems, such as replication, transcription, expression, regulation, interaction, transformation, differentiation, and mutation. The vast majority of nouns with abstract meaning are words with derived bases

of adjective (active → activation) or verb origin (replicate → replication). These examples indicate the abstract nature of the phenomenon reflected in the term. These phenomena cannot be counted, they cannot be measured or their limits determined, etc., but the characteristic features of these phenomena can be distinguished. Abstract nouns are symbiotic formations that combine in their categorical semantics the meaning of a phenomenon as a process and its qualities and form a special zone in the class of nouns. They allow describing dynamic and functional aspects of objects that cannot be directly observed, and thus play a key role in scientific discourse and categorisation of knowledge in the field of biotechnology.

Proper nouns that denote the individual name of a particular object and distinguish it from a number of objects of the same type are widely used in the verbalisation of the concept of "object". They allow describing concepts related to specific biological or technological objects. The use of proper nouns helps to accurately identify specific objects and their unique properties, which is important for scientific description and communication in the field of biotechnology. Proper names serve as a constant source for the formation of eponyms. In terminology, eponyms are terminological units used to name concepts that reflect unique objects, discoveries, processes, or technologies associated with specific individuals, organisations, or projects. Proper nouns are part of models of multicomponent eponymic terms and can be classified into several groups. These include the names of specific organisms or strains that uniquely identify biological objects, such as Escherichia coli, Saccharomyces cerevisiae, Watson-Crick; the names of cell lines, molecular tools, or enzymes that denote laboratory materials or technologies, such as HeLa cells, CRISPR-Cas9, Taq polymerase, Sanger sequencing method; and the names of large scientific projects, databases, or organisations that have become eponyms for large-scale research and initiatives, for example: Human Genome Project, ENCODE project, BioBricks Foundation (Kimball, 2002). The use of proper names in terminology allows accurately identifying objects and emphasising their unique properties, which is important for scientific description, communication, and standardisation of knowledge in biotechnology.

Thus, words denoting objects of reality, play an active role in the nomination process in the biotechnological sphere, since they provide accuracy and clarity in defining concepts, serve as the basis for the formation of terminological units and reflect key objects, phenomena, and processes around which scientific knowledge is built. By performing a quantitative analysis of the terms of the selected term system, it was revealed that the absolute majority are terms represented by a noun, which is quite natural for the scientific and technical industry. The total number of such terms was 206 units out of a total sample of 400 terms, which corresponded to 51.5% and indicated their dominant presence in the material under study. Similar dynamics were recorded by A.A. Tyutyunyk (2021), who noted that although the leading position is still occupied by simple terms (45.5%), phrase terms (41.34%) also make up a significant share, which indicates the relevance of multicomponent structures in modern term-making. This trend was also observed by G. Bondarenko (2020), who in the study of terminological vocabulary noted the dominance of nouns.

The morphological method of term formation is the basis for the development of English-language terms – nouns in biotechnology of the genomic era, since it is the foundations and affixes that carry the main word-forming load, since the morphological structure of the term formed by affixation organically lays the foundations of structural systematisation, which are of primary importance for terminology, because they allow tracing changes in the semantics of the word associated with the addition of so-called formal indicators – suffixes and affixes (Bondarenko, 2020). Using word-forming

analysis to distinguish between productive and unproductive ways of forming English-language terms represented by nouns in the field of biotechnology of the genomic era, several main groups were identified depending on the nature of their word-forming structure. This classification includes suffix formations – terms created by adding suffixes to bases that denote processes, phenomena, properties, or carriers of action. Such units as docking, transcription, replication, modification, stability, sequencer, activator (47 units – 22.8%) demonstrate high performance in expressing dynamic and static characteristics.

Prefix formations are formed by adding a prefix to the generating base without involving suffixes; the prefix changes or clarifies the meaning of the base word. In biotechnological terminology, such structures are less common, but they perform an important function of semantic modification, as illustrated by the examples of antiticodon, antigen, remark, pretest, and postgenome (29 units, 14.1%). A separate group consists of prefix-suffix formations, in which prefixes and suffixes simultaneously modify the meaning of the base, indicating the processality, state or action characteristic of scientific terminology. These include reprogramming, demethylation, misfolding, deactivation, reexpression, uncoating, rereplication, and misregulation (47 units – 22.8%). An important place among word-forming structures is also occupied by composite (root-forming) formations created by combining two or more bases, which can be either independent words or root morphemes. This type of creation is productive in biotechnological terminology, as it allows forming capacious, accurate, and semantically rich names of complex objects, processes and phenomena, for example: mutagen, bacteriophage, bacteriostat, immunosensor, ribosome, carcinogen, biotechnology, pharmacogenomics (41 units – 19.9%). A separate group consists of terms with Latin and Greek roots, which are formed on the basis of classical languages, which ensures their international clarity and stability in scientific use. These include epigenome, transcriptome, cytokine, mutagen, proteome, and metagenomics (41 units, 19.9%).

Terms represented by nouns in the English-language biotechnology terminology of the genomic era are often formed using neoclassical components such as bio-, geno-, -omics. Some researchers consider them the same status as ordinary words, while others give them a special status and consider them as a separate category. According to the rethinking of the status of these morphemes in the English derivation system, neoclassical components should be considered as a separate class of combining forms. They occupy an intermediate position between roots and affixes, since they have their own lexical meaning, but do not function as independent words. These forms are productive elements of the word-forming system, which allows creating new terms using regular morphological rules. This approach allows more accurately describing the structure and mechanisms of creating modern biotechnological terminology, which is especially relevant for the analysis of language units of the genomic era. This was also confirmed by a number of researchers. I. Plag (2018) emphasised the semi-autonomous nature of such forms, emphasising their ability to function as independent elements in word formation. A similar opinion was expressed by G. Booij (2005), who drew attention to their ability to form new words in regular, though somewhat limited, patterns. Similar observations were expressed by O. Garmash (2014; 2016), who analysed the verbalisation of English-language biocentric and technocentric concepts of affixal type and emphasised the active use of neoclassical morphemes as productive word-forming components in the development of complex scientific terms that reflect the latest concepts of genomics, bioinformatics, proteomics, etc. The results of quantitative analysis of the formulation of biotechnological nouns in the genomic era are presented in Table 1.

Table 1. *Distribution of nouns by word-forming structure*

No.	Type of word-forming structure	Number of terms	Share
1	Suffix formations	47	22.8%
2	Prefix formations	29	14.1%
3	Prefix-suffix formations	47	22.8%
4	Composite (basic) formations	41	19.9%
5	Terms with Latin and Greek roots	41	19.9%
	Total	206	100%

Source: *compiled by the author*

The word-forming analysis of English-language noun terms in the field of biotechnology of the genomic era showed different degrees of productivity of word-forming models. Suffix and prefix-suffix formations were the most commonly used, which together cover almost half of the entire term base (about 46%). This indicates the high productivity of suffixing as the main method of creating scientific terminology, which allows denoting processes, phenomena, and actions characteristic of the biotechnological sphere. Composite formations and terms with Latin and Greek roots also occupy a significant share – 19.9% each. This is conditioned by the need to create succinct and precise terms that ensure interethnic clarity and consistency in scientific discourse. The relative low productivity of prefix nouns in term systems, which is 14.1%, is confirmed by studies of other researchers. In this context, Yu. Togano (2024) also noted that the prefix in English basically does not change the grammatical category of a word, but only expands its lexical and semantic meaning. Prefixes, according to the researcher, function more as lexemes or full-meaning units than as classical morphemes, which also limits their role in the creation of nouns in scientific terminology.

However, it should be noted that the study by M. Bogachyk & D. Bihunov (2020), devoted to the structural and semantic features of computer terminology in English, prefix formations account for 14.3% of the total number of analysed units, while prefix-suffix formations account for only 1.8%. This distribution demonstrates that the performance of word-forming models largely depends on the industry specifics of the term system. In the terminology of biotechnology, models related to the description of biological processes and structures predominate, where suffixation and basic assembly play a more important role, while in the computer sphere, the active use of prefixes is more often recorded due to the need to quickly create new functional names based on existing terms. Prefixes mostly only change the meaning of a word without changing its grammatical category, which narrows their word-forming use. Thus, the results of the analysis confirm the dominance of suffix and combined (prefix-suffix) models in the structure of English-language biotechnological terms, and the importance of composite formations and classical language elements in ensuring the scientific accuracy and universality of biotechnological terms – nouns of the genomic era.

Categorisation of the world at the level of features involves structuring knowledge about various characteristics of the surrounding reality, primarily qualities and properties, and finds its linguistic reflection at the level of the adjective category, which is an integral part of the terms that implement this concept. The adjective is an important part of speech, often found in syllabic models of English terms that verbalise concepts of genomic era biotechnology. In terminological phrases of attribute models, the noun fixes an object or object, and the adjective specifies its properties, origin, or functional specifics, for example: genomic sequence,

cellular pathway, recombinant protein, molecular marker, transgenic mouse, regulatory element, viral vector, bacterial strain, synthetic pathway, specific promoter (Zaid *et al.*, 2001). The presence of adjectives in models of English-language biotechnological terms indicates that during the nomination process, experts emphasise the essential features of objects necessary for their accurate scientific description and categorisation. Adjectives in terms of the genomic era mostly reflect properties that can be perceived or conceptualised through parameters of size (long-read sequencing), structure (structural genomics), shape (circular DNA), colour (green fluorescent protein), functional characteristics (functional genomics), spatial orientation (spatial transcriptomics), and temporal dynamics *(*real-time PCR*,* time-dependent expression) (Genomics Glossary, n.d.). Based on these features, adjectives play a key cognitive role: they highlight the most important characteristics of objects and processes, systematise knowledge, and clarify the meaning of terms. Thus, adjectives become not only a grammatical, but also a conceptual tool for term formation.

In genomic era biotechnology terminology, the A + N model is used to verbalise various concepts. The concept of "object" is represented by an adjective that specifies the type, kind, or characteristic of an object (for example: human genome, viral vector, or bacterial strain). The concept of "goal/result" is implemented through an adjective that emphasises the properties or functions of the final product (for example: recombinant protein, therapeutic molecule, therapeutic protein, transgenic organism). The A + N model provides high consistency and logic of terms, since it allows combining the basic concept (noun) with its refinement (adjective), which is especially important in genomic and molecular biotechnology. It allows accurately identifying both the objects of research and the final results of experimental activities, emphasising their properties, origin, and functions. This makes terms more specific, analytical, and unambiguously understandable in a scientific context.

In the analysed terminological system, two main types of adjectives can be distinguished: qualitative and relative. In the field of biotechnology, where accuracy, quality and efficiency play an important role, the language of science should be not only informative, but also as clear as possible. That is why qualitative adjectives and phrases constructed with them are an important tool for scientific description. They help to convey specific properties of biological processes, materials, and technologies. As part of multicomponent models such as A + A + N, A + N + N, or A + A + A + N, qualitative adjectives detail the properties of an object, for example: comprehensive functional genomic study, sensitive diagnostic test, stable gene expression, stable protein structure, strong promoter activity, pure DNA sample, robust experimental design, efficient delivery system, or advanced genomics methods: accurate measurement technique, efficient fermentation process. Phrases with high-quality adjectives are not only a way to decorate the language, but also a tool for scientific accuracy, which allows communicating effectively in an area where every characteristic matters. Knowledge and active use of such constructions increases the level of scientific literacy and professionalism of the specialist. Qualitative adjectives reflect the qualitative characteristics of an object or process.

Relative adjectives are adjectives that indicate a feature of an object because of its relation to another object or phenomenon. In a biotechnological context, such adjectives are widely used to accurately describe processes, materials, technologies, and equipment. In the field of biotechnology, relative adjectives in the A + N model indicate the relationship of an object to a particular industry, process, environment, or property. In biotechnology, this model is widely used for scientific accuracy, time creation, and classification of phenomena. Relative

adjectives in the A + N model allow verbalising concepts formally and clearly: specify a method, material, property, or industry, for example: genetic engineering, genetic variation, genetic material, recombinant DNA, recombinant protein, microbiological analysis, microbiological safety, immunological response, immunological assay, viral vector, viral infection, chromosomal abnormality, chromosomal mapping, biotechnological innovation, biotechnological production, antigenic determinant, metabolic pathway, enzymatic reaction, enzymatic degradation (Zaid *et al*., 2001).

In the biotechnological term system of the genomic era, the adjective+noun (A + N) model is one of the most common key concepts for verbalisation: 89 (45.9%) term units out of 194 were formed using this structural model, which allows simultaneously reflecting the main object of research and its characteristic, property or function, which significantly improves the accuracy of terminology. The noun + noun (N + N) model also plays an important role and makes up 57 phrases, which is 29.4% of the total number of terminological phrases. Model A + N is used with greater frequency because adjective-noun constructions provide a higher level of terminological accuracy, structural flexibility, and semantic transparency, which is critical for scientific discourse in biotechnology. It is these characteristics that allow the A + N model to act as a basic tool for nominating new concepts in the industry.

This conclusion was also reached by H. Syrotina (2022), who, as part of its research on the categories of concepts expressed in English-language terms of the biotechnological sphere, analysed grammatical models of term combinations and found that, despite the lower performance compared to the A + N model, the N + N structure remains relevant for denoting substantive terms with an attributive meaning, where the main component is located at the end of the phrase. O.A. Brona (2023) found that in English-language scientific texts, the frequency of using attributes-nouns exceeds the frequency of attributes-adjectives, although in the linguistic tradition it is common to use relative adjectives as terminological attributes that form complex terminological phrases, carrying a feature meaning and providing the synthetics of knowledge necessary to reveal the semantic potential of terms. The composition of terminology depends on the field of sciences – humanities, social sciences, natural sciences, or technical sciences. Common to them is the naming of beings, things, and phenomena, but in the terminology of the sciences of the humanities and the social cycle, there is an attraction to more nouns. But in the natural and technical sciences, the quantitative meaning of nouns.

The study by K. Pitkänen-Heikkilä (2015), dedicated to the analysis of adjectives as terms in the field of plant morphology, made similar observations, but with a focus on a specific subject area. The researcher noted that almost half of the terms in plant morphology are adjectives that function not only as modifiers, but also as full-fledged terminological units with a stable meaning. The researcher stressed that the traditional practice of terminography, which included the nominalisation of adjectives, was unsuitable for the field of plant morphology. This is conditioned by the fact that such nouns are not used in authentic scientific texts, and the transformation itself can lead to semantic distortions. Both researchers emphasised the conceptual significance of attributive adjectives, which do not act as secondary grammatical elements, but, on the contrary, play a key role in the formation of terminological meaning. If O.A. Brona (2023) considered this issue in a broader interdisciplinary context, K. Pitkänen-Heikkilä (2015) argued for the expediency of preserving adjectives as independent terms in specialised dictionaries.

Multicomponent adjective models, in particular A + A + N and A + A + A + N, play an important role in the classification of objects, since they allow reflecting in detail several features

of the concept under study simultaneously, for example: recombinant viral vector, robust experimental design, functional genomic analysis, experimental genetic manipulation, comprehensive functional genomic study. Despite the relatively low frequency, they remain significant in the structure of the terminological system: the model A + A + N is 11 units (5.7%), and A + A + A + N – 7 units (3.6%) of the total number of analysed term combinations. Their use indicates the need for a more accurate, multi-level characterisation of concepts in scientific discourse. They allow simultaneously conveying qualitative traits, attitudes, and affiliation, making the term more informative and accurate. This approach contributes to the effective systematisation of terms and facilitates communication between specialists in the field of biotechnology of the genomic era. The tendency to increase multicomponent terms in biotechnological terminology was noted by E. Myshak (2017) and A. Syrotin (2017), who emphasised their importance for accurately conveying complex concepts that have emerged in the development of scientific knowledge. E. Myshak (2017) noted that multicomponent terms often occur in biotech discourse, and their number increases with the complexity of the subject area. The researcher noted that "the more components in a term, the lower its polysemy", that is, the highest degree of accuracy and unambiguity, which is especially important for professional scientific communication. Similar observations were made by researchers in related fields. Arguments in favour of multicomponence were also supported by F. Remy *et al*. (2023), who, when analysing clinical terminology, emphasised the importance of accurate recognition of verbose (idiomatic) terms, since they provided adequate encoding of complex concepts in specialised subsections. Thus, in the context of the rapid development of scientific disciplines, multicomponent terms are not only a natural consequence of the complexity of knowledge, but also an effective tool for avoiding ambiguity, improving accuracy, and forming new nominations.

As a result of morphological analysis of terms of the English terminological system, a term formation model was identified that includes such a language unit as the participle. By its origin, the participle has a noun basis, but in the process of language evolution, it has acquired verbal characteristics, in particular, signs of state and time. From the standpoint of cognitive linguistics, a participle is considered as an intermediate language category that differs from both a verb and an adjective, but reflects the cognitive processes of adjectivation of an action based on verbal semantics. The participle retains procedural features in its meaning, that is, it reflects a new way of perceiving an object in human consciousness. Its formation is based on the heterogeneity of cognitive processes, which is manifested in varying degrees of expression of verbal and adjective properties. A participle is a verb form in which an action, process, or state, while remaining procedural in nature, acquires the meaning of a feature or property.

In the structure of English terms built on the model (participle + noun), verbalising the concepts of biotechnology, the participle is presented in two main forms – present (present participle, -ing) and past (past participle, -ed). Both forms perform an important cognitive and term-forming function, but differ in semantics and pragmatics of use. The past participle in the V-ed + N model performs the function of indicating the properties of objects or materials obtained as a result of biotechnological processes, for example: engineered organisation, cultured cells, isolated DNA, modified vector; denatured enzyme, activated pathway, inhibited expression, targeted delivery; integrated system, controlled release, immobilised biocatalyst, encapsulated microorganism (Zaid *et al.*, 2001). The V-ed + N model is represented by 16 units, which is 8.2% of the total number of analysed term combinations. The past participle in terms of the biotechnological field

not only performs a nominative function, but also serves as a cognitive marker that reflects the result, method or nature of an action, fixing important aspects of complex biotechnological processes in a linguistic form.

Present participles in the V-ing + N model structure serve as markers of the active action, process, or functional ability of an object. Such participles, unlike forms with the suffix -ed, do not focus on the result of the action, but reflect the processality, dynamics, incompleteness or continuing nature of the event, for example: binding protein, coding region, splicing site, signalling molecule, targeting system, processing enzyme, emerging pathogen (Genomics Glossary, n.d.). This is especially true in the genomic era biotechnology concept sphere, where constantly updated knowledge about the functioning of living systems, genomic processes, and molecular interactions is key. The V-ing + N model was recorded in 14 term combinations (7.2% of the total). Although the proportion of conjugations with participial forms – models V-ed + N and V-ing + N – is relatively small (8.2% and 7.2%, respectively), their use in biotechnological terminology is significant. Such constructions allow nominating complex scientific concepts compactly, accurately and cognitively motivated, which, in turn, contributes to effective communication within the framework of research, applied and educational activities. The results of quantitative analysis of morphological models of biotechnological terms of the genomic era are presented in Table 2.

Table 2. *Frequency of morphological models of biotechnological terms of the genomic era*

No.	Morphological model	Number of cases	Percentage of total (%)
1	N + N	57	29.4
2	A + N	89	45.9
3	A + A + N	11	5.7
4	A + A + A + N	7	3.6
5	V-ed + N and V-ing + N	30	15.5
Total		194	100

Source: *compiled by the author*

The analysis of the data presented in Table 2 allowed identifying the most productive morphological models of terms, identifying the relationship between noun and adjective structures, and trace trends in the dominance of individual types in the formation of terms of the genomic era. The most productive model is A + N (45.9%), which is actively used to indicate tools, methods and characteristics of processes in the field of genomic research. Noun structures N + N (29.4%) reflect objects of research and phenomena, while participial forms (V-ed + N, V-ing + N, 15.5%) verbalise the dynamic aspect – the process or result of an action. Models with two and three adjective components (A + A + N, A + A + A + N) are less common (9.3% together) and perform the function of detailing complex concepts. Thus, the predominance of adjective-noun constructions indicates the trend of biotechnological terminology of the genomic era towards analyticity, compactness, and accuracy of scientific nomination. To clarify the relationship between the morphological model N and the conceptual categories "object", "process", "goal/result" and "tool", an integrative analysis was carried out, which helped to identify which aspects of scientific knowledge are most often represented through a certain morphological structure. Table 3 shows the results of the distribution of model N terms by main conceptual categories.

Table 3. *Distribution of conceptual categories by model N*

Conceptual category	Number of terms	Percentage of total (%)
OBJECT	110	54.0
PROCESS	35	17.2
GOAL/RESULT	20	9.8
TOOL	39	19.0
Total	204	100

Source: *compiled by the author*

The distribution of terms by conceptual categories within morphological model N shows that the largest share is occupied by objects – 54%, which indicates that in the terminology of the genomic era, considerable attention is paid to the name of key objects of research, such as genomes, proteins, cells, and other biological structures. This is logical, because objects are the basis for further analysis, experiments, and development. The second most frequent category is "instrument" (19%). They include terms that refer to advanced methods, devices, and technologies that are actively used to study genomes and modify them. The high proportion of this category highlights the role of technological progress in biotechnology and genomics. The category "process" is 17.2%, which indicates the active use of terms that describe various dynamic phenomena – research, sequencing, editing, etc. This reflects the importance of the procedural component in scientific research and practical applications. A smaller share falls on the "goal/result" category (9.8%), which includes terms that denote the final products or goals of biotechnological activities. This may indicate that compared to objects, tools, and processes, the concepts of results or goals are still less formalised in terminology.

Thus, the analysis demonstrates that model N mainly verbalises static, object concepts, which is the basis for building further descriptions of processes, tools, and results in the field of genomic biotechnology. This division reflects the nature of the development of scientific knowledge in this dynamic field. In order to summarise the results of the study and identify the relationship between morphological models and conceptual categories, an integrative analysis of the terms of the genomic era was carried out. Table 4 shows a comparison of the main conceptual categories with typical models of their morphological implementation and frequency indicators.

Table 4. *Correlation of morphological models and conceptual categories in genomic era terminology*

Conceptual category	Morphological models	Number of terms	Share of the total number
OBJECT	N, N + N, A + N	108	55.7%
PROCESS	V-ing + N, V-ed + N, A + N	33	17.0%
GOAL/RESULT	A + A + N, A + A + A + N	20	10.3%
TOOL	N + N, A + N	33	17.0%
Total	-	194	100%

Source: *compiled by the author*

The above table shows the distribution of terms by main conceptual categories and their morphological models, which helps to better understand the structure and functionality of terminology in the field of genomic biotechnology. The largest share is made up of terms in the

"object" category (55.7%). This dominance reflects a focus on naming specific research objects – genomes, genetic structures, molecules, and other elements that form the basis for further scientific action and technological applications. The most common morphological models in this category are simple noun constructions (N) and phrases like N + N and A + N, which ensures compactness and accuracy in names. The second largest category is "process" (17.0%), which includes participle models (V-ing + N, V-ed + N) and adjective-noun constructions (A + N). This indicates the active use of terms that describe actions, operations, methods, and technological processes within genomic era biotechnology. This morphology contributes to a concise and understandable reflection of the dynamic aspects of scientific activity.

The tool category (17.0%) covers terms that refer to tools, methods, and technologies used in genomics and related disciplines. Here, too, noun models predominate, which emphasises the importance of clear identifiers for naming technical means and techniques. The smallest category – "goal/result" (10.3%) – reflects terms related to the final products of research and their meaning. They are represented by more complex adjective models (A + A + N, A + A + A + N), which is explained by the need for a detailed description and classification of new scientific concepts and products. Thus, the analysis confirmed that genomic era terminology is dominated by simple noun constructs for designating objects, and more complex models predominate in categories that require more accurate characterisation and description of processes and results. This division reflects the general tendency to conciseness in the names of objects and to expand in the designation of dynamic and functional aspects of scientific knowledge.

Conclusions

Morphological analysis of the terms of the English-language biotechnological terminology of the genomic era showed that its structure is represented by a number of productive models: N, N + N, A + A + N, A + N + N, A + A + A + N, and V-ing + N, V-ed + N, which include nouns, adjectives, and participles. The noun conceptualises phenomena of objective reality, such as processes, materials, structures, and technologies, while the adjective and participle refine the characteristics of an object by adding information about qualitative, quantitative, functional, spatial, or temporal parameters. Among the terms, the noun model N dominates (206 out of 400 terms, 51.5%), which indicates its leading role in the term system. The most productive model is A+N (45.9%), which denotes objects, tools, methods, and process characteristics. A smaller share is occupied by N + N structures (29.4%), which reflect the objects of research, and participial forms (15.5%), which verbalise the dynamics of processes or the results of actions. Models with multiple adjectives (9.3%) perform the function of detailing complex concepts, while simultaneously conveying qualitative features, attitudes, and belonging to the object. Comparing linguistic models with conceptual categories – "object", "process", "goal/result", "tool" – allows tracing how language structures scientific knowledge. Noun models mostly represent the object of research, adjective-noun constructions represent tools and characteristics of processes, and participial forms verbalise actions or target processes. The term system of the genomic era is dominated by the concept of "object" (55.7%), which reflects the priority of accurate naming of fundamental elements of research – genomes, molecules, structures. The "process" category (17%) includes participle and adjective-noun constructions, while the "tool" category (17%) covers tools and methods, mostly noun constructions. The smallest category – "goal/result" (10.3%) – is represented by more complex adjective models, which reflects the need for a detailed description of the final products of research. Thus, morphological

and cognitive analysis of terminological models demonstrated the relationship of formal language organisation with the conceptual content of the biotechnological terminosphere. Multicomponent models provide the accuracy, compactness, and detail of terms needed to describe complex biological processes and structures. A promising area of further research is the investigation of the dynamics of term formation in the context of the rapid development of the latest technologies, in particular, CRISPR, artificial intelligence in bioinformatics, and synthetic biology.

Acknowledgements

None.

Funding

None.

Conflict of Interest

None.

References

[1] Adambaeva, F.R. (2025). Semantic hierarchies in biotechnology terminology: A study of hyper hyponymic relations in English and Uzbek. *Modern American Journal of Linguistics, Education, and Pedagogy*, 1(3), 826-833.

[2] Adambaeva, F.R., & Sadullayeva, N.A. (2023). Terms-eponyms as a part of biotechnological terminology. *European Journal of Literature and Linguistics*, 4, 40-44. doi: 10.29013/EJLL-23-4-40-44.

[3] Almashadni, T.M.S., & Jawad, A.S. (2025). Semantics analysis of emerging terms in biotechnology: In English linguistics, send me a perfect structure for it. *Saudi Journal of Educational and Psychological Studies*, 6(1), 1-20.

[4] Bauer, L., & Renouf, A. (2001). A corpus-based study of compounding in English. *Journal of English Linguistics*, 29(2), 101-123. doi: 10.1177/00754240122005251.

[5] Bogachyk, M., & Bihunov, D. (2020). The structural-semantic features of computer terms in English. *Cognitive Studies|Études Cognitives*, 20, article number 2262. doi: 10.11649/cs.2262.

[6] Bondarenko, G.P. (2020). Morphological methods of creating branch terms to denote linguistic concepts. *Scientific Notes of the V.I. Vernadsky TNU. Series: Philology. Social Communications*, 31(70(4(1))), 12-17. doi: 10.32838/2663-6069/2020.4-1/03.

[7] Booij, G. (2005). *The grammar of words: An introduction to linguistic morphology*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199226245.001.0001.

[8] Brona, O.A. (2023). *Structure and semantics of term concepts in English-language geological discourse*. (PhD thesis, Ivan Franko National University of Lviv, Lviv, Ukraine).

[9] Deroey, K. (2007). Douglas Biber: University language: A corpus-based study of spoken and written registers. *Applied Linguistics*, 28(4), 624-627. doi: 10.1093/applin/amm039.

[10] Encyclopedia of Biotechnology. (n.d.). Retrieved from https://www.eolss.net/toc/c17-browsecontents.aspx.

[11] Garmash, O.L. (2014). Verbalization of English biocentric lexicalized concepts of affixal type. *Science and Education a New Dimension*. *Philolgy*, 2(7(34)), 47-50.

[12] Garmash, O.L. (2016). Verbalization of English technocentric lexicalized concepts formed by the metacognitive mechanism of affixation. *Scientific Notes of Mykola Gogol State University of Nizhyn. Philological Sciences*, 1, 158-164.

[13] Genomics Glossary. (n.d.) Retrieved from https://discoveringthegenome.org/glossary.

[14] Glossary of Genomics Terms. (n.d.) Retrieved from https://www.futurelearn.com/info/courses/the-genomicsera/0/steps/4864.

[15] Junya, M. (2017). A syntactic-morphological analysis of nominal compounds. *Florida Linguistic Papers*, 3, 51-62.
[16] Kimball, N. (2002). *Glossary of biotechnology terms*. Boca Raton: CRC Press.
[17] Matthews, P.H. (2014). *Oxford concise dictionary of linguistics* (3rd ed.). Oxford: Oxford University Press.
[18] Matvieieva, S.A., & Torgovets, M.A. (2022). Cognitive terminological structure as a unit of the professional worldview. *Scientific Journal of Mykhailo Dragomanov State University of Ukraine. Series 9. Current Trends in Language Development*, 23, 41-51. doi: 10.31392/NPU-nc.series9.2022.23.04.
[19] Merriam-Webster. (n.d.). *Genomics*. Retreived from https://surl.li/dgizuu.
[20] Molina-Plaza, S., & Martínez-Sáez, A. (2024). Exploring metaphors in biotechnology journals. In S. Molina-Plaza & N. Maroto (Eds.), *Aspects of cognitive terminology studies theoretical considerations and the role of metaphor in terminology* (pp. 209-236). Berlin: De Gruyter Mouton. doi: 10.1515/9783111073149-010.
[21] Myshak, E. (2017). The main means of formation of biotechnological terms. *American Journal of Research*, 4(4), 84-97.
[22] National Human Genome Research Institute. (n.d.a). *Genomics*. Retrieved from https://www.genome.gov/genetics-glossary/genomics.
[23] National Human Genome Research Institute. (n.d.b). *Talking glossary of genomic & genetic terms*. Retrieved from https://www.genome.gov/genetics-glossary.
[24] Pitkänen-Heikkilä, K. (2015). Adjectives as terms. *Terminology: International Journal of Theoretical and Applied Issues in Specialized Communication*, 21(1), 76-101. doi: 10.1075/term.21.1.04pit.
[25] Plag, I. (2018). *Word-formation in English*. Cambridge: Cambridge University Press.
[26] Rajak, S. (2018) Review on different stages of biotechnology and its application. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 5(9), 1044-1051.
[27] Remy, F., Khabibullina, A., & Demeester, T. (2023). Detecting idiomatic multiword expressions in clinical terminology using definition-based representation learning. In A. Bhatia, K. Evang, M. Garcia, V. Giouli, L. Han & S. Taslimipoor (Eds.), *Proceedings of the 19th workshop on multiword expressions (MWE 2023)* (pp. 73-80). Dubrovnik: Association for Computational Linguistics. doi: 10.18653/v1/2023.mwe-1.11.
[28] Selivanova, O. (2006). *Modern linguistics: A terminological encyclopaedia*. Poltava: Enviroment.
[29] Striuk, Ye. (2020). Analysis of the concept biotechnology in modern English. *Scientific Notes of V.I. Vernadsky TNU. Series: Philology. Social Communications*, 31(70(1(2))), 204-208. doi: 10.32838/2663-6069/2020.1-2/40.
[30] Syrotin, A. (2017). Multi-component English terms of biotechnology sphere. *Cogito. Multidisciplinary Research Journal*, 9(3), 78-86.
[31] Syrotina, H. (2022). Concept categories expressed in English terms of the biotechnology sphere. *Cogito. Multidisciplinary Research Journal*, 14(1),174-187.
[32] Syrotina, O. (2020). Anthropomorphic metaphor in the English terminological system of biotechnology. *The Journal of V.N. Karazin Kharkiv National University. Series "Philology"*, 84, 62-66. doi: 10.26565/2227-1864-2020-84-09.
[33] The American Heritage Dictionary of the English Language. (2022). *Biotechnology*. Retrieved from https://www.ahdictionary.com/word/search.html?q=Biotechnology.

[34] Togano, Yu. (2024). A word-based approach to the so-called category-changing usage of the English derivational prefix out-. *Languages*, 9(5), article number 172. doi: 10.3390/languages9050172.

[35] Tyutyunyk, A.A. (2021). Structural and semantic features of French terms in the field of biotechnology. In *Collection of materials of the V Round Table with international participation "Modern trends in phonetic research"* (pp. 99-102). Kyiv: Igor Sikorsky Kyiv Polytechnic Institute.

[36] Yuldashev, M., & Haydarov, S. (2022). The terms of concept and its interpretations in linguistics. *Internation Journal on Integrated Education*, 5(6), 455-461. doi: 10.31149/ijie.v5i6.3263.

[37] Zaid, A., Hughes, H.G., Porceddu, E., & Nicholas, F. (2001). *Glossary of biotechnology and genetic engineering*. Rome: Food and Agriculture Organization of the United Nations.

Морфологічний аналіз англомовних біотехнологічних термінів геномної ери

Олена Сиротіна
Кандидат педагогічних наук, доцент
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
https://orcid.org/0000-0002-4802-4891

Анотація. Актуальність дослідження зумовлена зростанням значення англомовної біотехнологічної термінології в контексті інтенсивного розвитку геномних технологій, що потребує системного аналізу її морфологічної структури для покращення наукової комунікації та глибшого розуміння когнітивних процесів термінотворення. Мета роботи полягала в аналізі морфологічної структури англомовних біотехнологічних термінів геномної ери крізь призму когнітивного підходу. Для досягнення поставленої мети використано такі методи: лексикографічний, морфологічний та статистичний аналіз, елементи когнітивного підходу до вивчення термінів, а також методи систематизації та узагальнення. Встановлено, що в англомовній біотехнологічній терміносистемі активно використовуються термінологічні моделі, до структури яких входять іменники, прикметники та дієприкметники. Абсолютну більшість становлять терміни, представлені іменниками (N), причому найбільш уживаними виявилися суфіксальні та префіксально-суфіксальні утворення, які разом охоплюють близько 46 % усієї термінної бази. Розподіл термінів за концептуальними категоріями у межах морфологічної моделі N показав, що найбільшу частку займають об'єкти (54 %) – ключові елементи біотехнологічного дослідження, такі як геноми, білки, клітини, за якими слідують інструменти (19 %) – сучасні методи, прилади й технології, процеси (17,2 %) – динамічні дії на кшталт секвенування чи редагування, та мета/результат (9,8 %) – кінцеві продукти або цілі. Кожна із проаналізованих структур (N, N + N, A + N, A + A + N, A + A + A + N, V-ed + N, V-ing + N) виявилась пов'язаною з окремими аспектами концептуальної сфери: іменникові моделі здебільшого репрезентують об'єкт дослідження (наприклад, gene expression, protein synthesis), прикметниково-іменникові конструкції – інструменти та характеристики процесів (наприклад, synthetic genomic technologies), тоді як дієприкметникові форми вербалізують результати дії, змінені стани або властивості, набуті внаслідок біотехнологічного втручання (наприклад, engineered organism, modified gene) або функціональні властивості об'єктів (наприклад, binding protein, binding site, coding sequence). Морфологічний аналіз дозволив виявити не лише структурні особливості термінів, а й ті концептуальні моделі, що лежать в основі наукового мислення у сфері біотехнологій, відображаючи динаміку розвитку галузі та її лінгвокогнітивну специфіку. Практичне значення дослідження зумовлене можливістю використання отриманих результатів у процесі викладання іноземної мови за професійним спрямуванням для студентів спеціальності «Біотехнології та біоінженерія»

Ключові слова: терміносистема біотехнології; словотвірна структура; морфемна організація; концепт; когнітивний підхід

International Journal of Philology, 29(3), 2025
Journal homepage: https://philologicalscience.com.ua/en
Received 26.04.2025 Revised 09.08.2025 Accepted 23.09.2025

UDC 811.161.2'42:030:001.4 DOI: 10.31548/philolog/3.2025.29

Composition of the encyclopaedic article in projection onto the biblionymic cross-section: Cognition, structure, specifics of determinative relations

Mykola Stepanenko*
Doctor of Philological Sciences, Professor
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
https://orcid.org/0000-0002-6727-1265
Nina Stepanenko
PhD in Philological Siences, Associate Professor
National University of Life and Environmental Sciences of Ukraine
03041, 15, Heroiv Oborony Str., Kyiv, Ukraine
https://orcid.org/0000-0002-6727-1265

Abstract. The relevance of the proposed scholarly study was conditioned by the orientation of modern linguistics towards the study of encyclopaedics as a poly-paradigmatic science in projection onto its modern content, technical equipment and the process of encyclopaedographing – the creation of encyclopaedic dictionaries and the development of the theoretical foundations. The aim of the work was a comprehensive investigation of the specifics of the interaction between the pre-textual (title complex) and the proper textual structure of the encyclopaedic article as the basic unit of encyclopaedic discourse. The reliability of the results obtained was ensured by the use of general scientific and linguistic methods: induction and deduction, the descriptive, descriptive-analytical, comparative and biographical methods, dictionary definitions, contextual and component analysis. On the basis of an analysis of folkloristic, linguistic, literary-studies and historical encyclopaedic publications, specialised regional-biographical, sectoral regional-biographical and other sources, relevant features of the encyclopaedic text were interpreted, and the classical composition of the encyclopaedic article as the principal structural unit of universal or sectoral encyclopaedic publications was elucidated. The main focus of the study was the autosemantic-pragmatic pre-textual unit – the title – which systematically and consistently correlated in content and in

Suggested Citation:
Stepanenko, M., & Stepanenko, N. (2025). Composition of the encyclopaedic article in projection onto the biblionymic cross-section: Cognition, structure, specifics of determinative relations. *International Journal of Philology*, 29(3), 29-47. doi: 10.31548/philolog/3.2025.29.

*Corresponding author

the means of language architectonics with the definitive-textual part. The immanent features (consistently actualised informativeness, rigid logicality, maximal concreteness, structural-semantic concision, non-disputatiousness) and functions (integrative, distinctive, nominative, representative, conceptual-content, aconnotative) of the biblionym-anthroponym as an obligatory constituent of the encyclopaedic article – the most important actualiser of the post-message – were ascertained; it has a specific propositional-dictum motivational base, a strictly regulated structural organisation, and a one-way cognitive-discursive trajectory – from title to text. The fundamental criteria for identifying anthroponymic biblionyms with an identical verbal set were established. The anthroponymic-familial cross-reference was disclosed according to the identifying features "degree of kinship" ["full kinship", "kinship of varying degrees between grandparents and grandchildren", "collateral kinship of varying degrees", etc.] and "degree of affinity". The practical significance of the study conducted was determined primarily by the fact that its results can be applied to systematise, deepen and refine recommendations that directly concern the qualitative potential of preparing an encyclopaedic article

Keywords: structural unit of the encyclopaedic text; encyclopaedic pre-text and text; classical constituents of the encyclopaedic text; anthroponymic pre-textual component; multi-component anthroponymic formula; anthroponymic-familial cross-reference

Introduction

Encyclopaedic discourse organically correlated with scholarly and popular-scholarly discourse, was united with these discourses by various parameters at the level of form and content, and at the same time appeared as a specific genre of the discursive paradigm. The relevant characteristics of encyclopaedic discourse were considered to be its systemacity, hierarchical nature, laconicism, compactness of form, abstractness, clear content determinacy grounded in strictly regulated causal relations, terminologicality, absence of an authorial conception, intertextuality, hypertextuality, connotative indifference, and unified structure. The relevance of the proposed research was motivated by the need to expand the theory of encyclopaedic discourse with new partial data concerning cognition, reception, structure, the nature of determinative relations, the extralingual context of the encyclopaedic article; and by the need for the more active introduction into encyclopaedographical practice of technologies for its high-quality preparation in view of the experience developed and the concrete pragmatic demands of the present. An important component of such preparation was the set of issues that directly concerned the process of nominating entries, with the prospect that titles should organically harmonise with the definitive part and its semantically relevant components.

Encyclopaedic articles, arranged in a certain sequence – most often alphabetically, less often according to another principle – formed the corpus of entries of the corresponding compendium of knowledge (encyclopaedia, encyclopaedic dictionary, encyclopaedic handbook, etc.). In the opinion of R. Serdeha (2020), it was appropriate to conditionally distinguish two key components in an encyclopaedic article: the first – the article title or lexical unit fixed as the heading; and the second – the explanatory part, which contained a coherent and systematic disclosure of the content. The corpus of articles was formed purposefully and systematically – on the basis of a previously determined list of topics selected in accordance with the overall concept of the publication. The term "encyclopaedia", despite its entrenched dynamic

history and wide – worldwide – area of use, did not have an unambiguous, or more precisely, an all-embracing interpretation. Most often, this term was treated as a collection of encyclopaedic articles of a scholarly or popular-scholarly nature created by specialists in accordance with defined standards and clear criteria. Such texts were usually distinguished by informativeness, objectivity, scholarly reliability, accuracy and propaedeuticity. This corpus – a universally recognised and convincing marker of human progress in its various dimensions: scientific and technological progress, intellectual and spiritual advancement; a register of national achievements, the wealth of a country; a systemic fixer of its rises and declines; ultimately, the source into which every generation invests all that it wishes to pass on to its descendants.

Encyclopaedic publications were classified by various criteria, conditioned primarily by the multifunctionality of this important bearer of information about everything that surrounds a person, affects people, that people cognise or transform, as well as by its projection onto the collective recipient, i.e., its being addressed to an audience not limited by any boundaries or to particular categories of people, for example those united by type of activity, profession, ethnic affiliation, confession, political convictions, age, tastes, preferences, etc. According to the work of I. Dziuba (2021), it was appropriate to classify encyclopaedias by four criteria, because this integrative model fully embraces the encyclopaedic matrix. The first criterion concerned the purpose of the continuum described (scholarly, popular-scholarly, popular publications), and the second – the readership (publications for mass readers, for specialists in particular branches of knowledge, for young people and children). The third criterion envisaged such an important characteristic of encyclopaedic publications as the structure (alphabetical, systematic, alphabetical-systematic publications), and the fourth took into account the nature of the information such publications represent (universal, sectoral, specialised, regional publications. Each of the distinguished encyclopaedic subtypes can be differentiated by more specific criteria into narrower genre modifications. An important place in this context was assigned to electronic encyclopaedias as a landmark achievement of modern encyclopaedistics. The encyclopaedias were reasonably considered an element of open science, the essence of which is defined by the fact that research results, according to A.V. Yatsyshyn *et al.* (2021), become accessible to all, since such results do not presuppose any technical or financial barriers. The emergence of such publications, in the opinion of S.-K. Lin (2020), was also conditioned by the fact that in the Internet era people wish to learn quickly about all that is topical; however, owing to the absence of reliable sources, sensational information spreads in social networks faster than verified information and creates so-called fake news. Online encyclopaedias must become reliable, verified information that counteracts this negative and dangerous process. In these modern compendia of knowledge, the structural unit that represents such compendia acquires a new quality. The advantages and popularity of modern carriers of information lie not only in free access, but also in the fact that the verbally presented content can, where necessary, be accompanied by multimedia elements, be constantly updated, changed and supplemented without the involvement of financial resources. Ultimately, it is more convenient to use, since websites have "quick search (simple and advanced)", and also "various interesting sections for users (forums, competitions, news, quizzes, surveys, etc.)" (Yatsyshyn *et al.*, 2021).

The aim of the article was to analyse systematically the biblionymic cross-section of the encyclopaedic entry against the background of its compositional structure, above all those determinative relations that arise between the anthroponymic title, represented by one-, two- and three-component

anthroponymic formulae, and the definitive-textual part with its regular programming of anthroponymic components as important informative centres of the semantic structure.

Materials and Methods

The source base consisted of encyclopaedic works, whose foundation was "an all-embracing corpus of reliable information intended to satisfy the informational needs of society", in particular "an integral scientifically integrated system of knowledge about Ukraine in the full compass of its spatio-temporal existence, development and prospects, a science of historical memory, culture, experience, a peculiar philosophy and politics of statehood", and so on (Chernysh, 2018): folkloristic, linguistic, literary-studies, historical encyclopaedic publications, specialised regional-biographical, sectoral regional-biographical and other publications. The basis of the author's card index was formed by contextual uses of official and unofficial anthroponymic formulae, carefully selected from V. Kubiiovych (2003), M. Zhulynskyi *et al.* (2015). This repertoire was supplemented by factual material obtained from V. Rusanivskyi *et al.* (2007), M. Dmytrenko (2020). Such specialised regional-biographical publications as O. Bilousko (2009), P. Rotach (2013) were examined in detail. As a kind of model for interpreting anthroponyms as titles and organic constituents of the definitive part of the article of the compendium of knowledge under examination, I. Dziuba *et al.* (2022) was chosen, whose specificity was determined by the fact that anthroponymic entries predominated in it, and this practice was consciously motivated: for various reasons of an objective and subjective nature, personalities attested in history had not entered the encyclopaedic sphere. There was no obvious need to expand the source spectrum, since the phenomenon under analysis did not display any specificity in encyclopaedic studies of different genres, and the mechanical accumulation of factual material did not materially affect the projected results of the research being carried out.

The following general-scientific and special linguistic methods were used in the work: the descriptive method – for the selection and ordering of the material, the cataloguing of the phenomenon under study; the descriptive-analytical method – for classifying anthroponymic formulae according to different criteria; induction and deduction – for consistently reflecting the determinative relations between the title and the definitive part with an anthroponymic resource; the biographical method – for establishing relations of blood and non-blood kinship in the anthroponymic space of an individual entry and of different articles with a shared set of one-, two- or three-component anthroponymic formulae, as well as for identifying the specificity of the anthroponymic formulae of particular social groups, and also for the study of national and partly global traditions of the formation and legitimisation of the anthroponymic formula, its possible ellipsisation and expansion caused partly by intra-, but mainly by extra-lingual factors; contextual analysis – for characterising the conditions of microtextual functioning within the structure of the entry, macrotextual functioning within the space of the encyclopaedic publication of anthroponyms and for establishing the direct and reverse connection with the biblionym; component analysis – for interpreting the phenomenon of variability within the structure of the anthroponymic formula and for differentiating titles by this feature; the comparative method – for clarifying the common and the different in the sphere of the anthroponymic formula in different encyclopaedic publications; dictionary definitions – for a comprehensive analysis of the specificity of encyclopaedic entries with an anthroponymic biblionym.

Results and Discussion

Scholars T. Vakaliuk *et al.* (2021) noted that what brings encyclopaedic discourse closer to other

genres of the discursive paradigm is, on the one hand, that it is inseparable from the discursive practice customary in society and is pre-programmed by an emphasised goal as well as by the social mechanisms that ensure its functioning, and, on the other, that it is distinguished by a set of communicative events regulated by certain features of communicative intentions, and is also defined by content and structural similarity, by the particularities of syntactic, lexical and stylistic organisation, and by its orientation towards a given audience. The principal concrete representative of the discourse under study, according to T. Vakaliuk *et al.* (2021), is the encyclopaedic article. It is appropriate to treat it as an organised set of lexicographical – more precisely, encyclopaedographical – parameters, the core among which is formed by the object and scope of lexicographising (encyclopaedographising). The former "determines the composition and boundaries of the dictionary", whereas the latter "conditions the selection of relevant linguistic information about the language unit; the dictionary's addressee, which presupposes taking into account the age and educational determinants of the average user". The analysed discourse and the content formed on its basis in the form of articles have been studied from different angles. There arises a need to characterise encyclopaedic discourse from a diachronic-synchronic perspective. This issue is fragmentarily covered in the work of N. Kovalchuk (2023).

Equally, within the constant field of attention are such important aspects of the issue under consideration as the dissemination of universal (sectoral) knowledge and the revival of the national memory of the Ukrainian people through the prism of the encyclopaedic domain. The conceptual contours of the problem in question are set out in the work of M. Zhelezniak *et al.* (2023). These, however, require the expansion of the theoretical base and factual filling within a systematic presentation – chronological, thematic, generic, functional-stylistic, etc. No less important are works that provide valuable information on high-quality editing of encyclopaedic content and academic integrity in scholarly communication. Of particular note in this context are the works of A. Petrushka (2021) and N. Berezovska-Savchuk (2023). The arguments advanced therein may serve as the canvas for future scholarly studies.

As for the encyclopaedic article as a concrete representative of encyclopaedic discourse in general and of a particular encyclopaedic publication in particular, this problem has not yet found an unambiguous solution. This concerns above all such research segments as the differential spectrum, the preparation technology oriented towards compositional structure, language-and-style features, the author's academic integrity, and non-verbal accompaniment. In particular, universal characteristics have been singled out that are inherent in any encyclopaedic entry and in an encyclopaedic publication as a whole. These include primarily conformity to the propaedeutic aim, which "consists in presenting, in an accessible form, knowledge sufficient to introduce the topic", objectivity of material, verifiability, reliability of information, reliance on scientific results and established facts, and the use of the method of generalisation. As regards the proposed methodology, its spectrum requires expansion. Against this background, researchers present a typology of the main errors, the emergence of which is caused by failure to observe, violation of, or disregard for one or more or even all the universal characteristics, namely: the encyclopaedic article is too superficial owing to the intention to include as much additional information as possible; very expansive due to the desire to cram all information into the exposition; rather one-sided because of detailed commentary on one thing and the failure to take account of other relevant information.

In conformity with O. Ishchenko & O. Savchenko (2023), those who compose Wikipedia entries should avoid the analysed and

other errors. The authors reached the reasoned conclusion that guidelines of the proposed kind do not guarantee that "the material of the article will be detailed, comprehensive, of quality, something that can be fully trusted", yet such guidelines will still perform the main function – facilitating the creation of full-fledged content. Each of these aspects requires theoretical adjustment and pragmatic concretisation. As for the composition of the encyclopaedic entry, its components – particularly the title – are specific, since, as the analysis conducted has confirmed, such components are based on rich factual material, are indifferent with respect to connotation, and do not correlate with the axiological potential inherent, for example, in artistic or publicistic discourse.

The quality formation of an encyclopaedic article begins with the title, which must vividly mirror its essence and serve as an important transmitter of the pragmatic nature of the text, functioning either autosemantically, i.e., as an independent text or message, or synsemantically, i.e., as its part (Yuldasheva *et al.*, 2022). Kh. Li *et al.* (2021) have convincingly proved another essential point: the biblionym is an integral component of the informational architectonics of encyclopaedic discourse that fits organically into the research landscape. The highlighted propositions served as one of the methodological premises of the present article. In view of the analysis above, no less significant are the informative links of the title with the key components of the encyclopaedic article which, by specifying, clarifying and expanding its semantic vectors, at the same time take an active part in the constitution of the overall content. Exceptional attention was paid to this issue in the present study. The most essential aspect of the consideration was anthroponymic titles in projection onto the specifics of the use as an anthroponymic formula, as well as onto the hierarchical links of this formula with textual anthroponyms. The material at the disposal allows stating that the fundamental propositions of the anthropocentric issue described systematically and consistently correlate with the methodological principles of a new linguistic field – linguopersonology, whose definitional boundaries have been clearly outlined by O. Mykytiuk (2025): "a syncretic science that demands study in social, linguo-cultural, linguo-philosophical, and linguo-political dimensions".

Works do not lose the acuteness whose object of research is the canonical attributes, structural organisation, and classification parameters of the basic structural unit of an encyclopaedic publication. In terms of content-type characteristics, one distinguishes review articles, reference articles, explanatory articles, cross-reference articles, and biographical articles, and, by nature of preparation, authorial and editorial articles (Kyrydon, 2020). The proposed classification is not universal, yet it may be considered exhaustive given the biblionymic matrix of encyclopaedias. What is highly essential in this cognitive space is that all the highlighted structural units function not separately, but in a certain interconnection, to a greater or lesser extent organically and harmoniously complementing one another, so that the publication which these units jointly and objectively reflect is "coherent both in content and in the manner of presenting the material". As for biblionyms, the biblionyms do not perform a demarcation function at the level of the genre typology of encyclopaedic articles, which is conditioned by the purely nominative nature, the attachment to a specific denotatum with its individualised status. The title, like the main exposition (the definitive part and the proper textual part as components of a single whole), bibliography, and author, are classical components of an encyclopaedic article.

The title should be interpreted as the first sign that provides a whole complex of notions about the content and form of what is communicated (Ruban, 2024). It is appropriate to treat it not simply as a component of the text,

but as a symbol of the meaning embedded in it, as the most important step of the receptive mechanism. There are every grounds to assign to the biblionym the status of a unique unit of the text and the most eloquent actualiser of textual categories, a highly important component of a publicistic, scholarly, religious or any other work or its part; ultimately, an integral attribute of a newspaper, journal, almanac, article, message, post, review – that is, of any verbal whole (Lebed & Bilohrats, 2020; Monakhova *et al.*, 2021).

The title as a compositional attribute of the text. The title stands with the main text in various prospective-retrospective relations. The peculiarity of this determinative phenomenon depends primarily on the functional-stylistic and genre nature of the text. The role the title plays in the artistic style differs significantly from its purpose in the scholarly or official-business style. It is important to consider, in addition to asymmetrical, also symmetrical characteristics. Among the latter is the name given by a micro- or macro-text, which has its own regular conditioning, since it is intended to reflect the content dominant in its strategic presentation. A specifically actualised title always appears as an independent – often visualised – component with a fixed initial placement above the text, which makes it possible to perceive it as the bearer of a certain code, a marker of informative relevance, an original phenomenon that focuses the semantic features of the author's narration and, as O.A. Rarytskyi (2017) wrote, clearly directs the recipient towards a concrete interpretative model of what is communicated, acting as a conventional signal for its decoding. Practically every title is an explicit or implicit secondary denomination that participates in the structural-semantic organisation of the text and, as L. Yuldasheva (2018) emphasised, as an important component of the text, significantly influences the formation of the main textual properties. According to F. Moncomble (2017), the most important task is the optimisation of the relevance of titles, which opens for the reader a direct, motivated access to the content.

The title as an obligatory compositional attribute of the encyclopaedic article. The communicative complex of the encyclopaedic text "title ↔ text" has pronounced specific features. These are demonstrated by the cognitive-discursive trajectory of the components of this complex. It proceeds from the title (pre-text) to the text, and not vice versa, as happens, for example, in an artistic work, when, after grasping the main exposition, recipients decode and justify its polysemous, metaphorised, emotively or emotionally connoted title, and so on, which often focuses on subtexts, prompts reflection, and requires additional commentary, hence an expansion of the range of background knowledge. In encyclopaedic entries the title is an obligatory unit, rarely expressed by a single lexeme but more often by a grammatically ordered structure which, in laconic form and without additional comments or unnecessary explanations, actualises the main content. Non-verbal means of representing the component of the article under description – graphic, numerical, symbolic, silent titles, i.e., graphic-interval lines that separate textual fragments, and the like – are not characteristic of encyclopaedic discourse in view of its clearly and consistently implemented informativeness, maximal logicality and concreteness and, which is particularly important for the genre under analysis, its non-disputatiousness.

Encyclopaedic titles have a specific motivational basis. In accordance with the status of the motivator and the model of the structure of knowledge about the world that surrounds the person and the world of the person, O. Selivanova (2011) distinguished the following types of motivation: propositional-dictum, which presupposes the use of motivators with a direct meaning, i.e., those in whose onomasiological structure real information about reality is embedded; associative, whose essence is

represented by metaphorical meaning; modus, which is based on evaluative reception; mixed, which operates on motivational criteria of different status. The encyclopaedic title clearly and unambiguously identifies the content of the article; outside encyclopaedic discourse it realises the same semantics as in this textual space; it contains concrete information about a concrete denotatum – a person (people), phenomenon(s), event(s), epoch(s), etc.; therefore, propositional-dictum motivation is immanent to it. A double structure, a complicated build, hypertrophied nominative character are not typical of the title under description. It functions as an autosemantic-pragmatic pre-textual unit that explicates the most relevant information in the paradigms of the formal, natural and social-humanitarian sciences, more rarely of practical skills, abilities, individual and collective experience, and is oriented "towards the subsequent text (prospection), to which text-centricity, self-centricity (the realisation of the author's intentions) and anthropocentricity (orientation towards the addressee) are adjacent. Taken together, these transform the title into a means of pragmatic influence and facilitate the process of interpreting the text" (Yuldasheva, 2018) and its content and formal-grammatical organisation. The basis of this complex mechanism is formed by vertical cohesion, which arises between the title as a pre-textual semantic code and the text as a systemic set of content differential features of the verbal whole realised by the corresponding verbal repertoire, and by horizontal cohesion, which exists between the title and its functioning in the text in full if it is monolexemic, or in the form of separate fragments if it is multi-component. According to the observations of T. Yeshchenko (2021), "the repetition of lexemes from the title in the structure of the text creates the so-called "semantic radius of the title", which binds the entire text". Vertical and horizontal cohesion are consistently actualised in lengthy, mostly multi-plot, encyclopaedic articles, chiefly those that represent the social-humanitarian sciences.

All biblionyms are characterised by integrative (generalising), distinctive and nominative functions. The latter may or may not be accompanied by connotation. For encyclopaedic titles, the following functions are also important: representative – the designation of the world of things – and conceptual-content – the establishment of systemic, hierarchical, determinative, paritarian, causal and other links and relations that pervade the world of things, reflecting evolutionary processes in nature or human society. As already noted, the quality of the names of encyclopaedic entries is determined primarily by the maximal consonance of the content with the article, by pre-textual explicitness, i.e., the real capacity to reveal, in general outlines, the denotative essence of relevant information. Moreover, such names must be structurally and semantically laconic, unburdened by unnecessary details, maximally correct, and indifferent with respect to positive or negative evaluation. The type of denomination described is not inherent in such functions as evaluative-expressive, emotive, imperative, aesthetic, or advertising; therefore, such titles occupy a separate place in the paradigm of titles, just as the encyclopaedic text itself does in the multifunctional space of the discursive paradigm.

Researchers classify biblionyms in different ways. The names of encyclopaedic articles are distinguished primarily by the fact that such names a priori confirm, argue for, and prove the reliability, correctness, truth of something; indicate the place and role in the world of everything that surrounds people, of a particular society or even of civilisation as a whole, of individual personalities or associations of people, of individual nations or ethno-communities, and thereby activate or inhibit receptive engagement, shape the level of attention to what is communicated, or lower the degree of interest in it. Such names belong to the corpus of constative titles, according to the classification of

D. Wunderlich (1976). The biblionyms thus distinguished do not lose this capacity even when formally acquiring other classificatory features. For example, in the definitive parts of encyclopaedic entries that represent the sphere of the humanities, a distinctly declared connotative orientation serves as an auxiliary factor in the conceptualisation of fragments of surrounding reality on the basis of cognitive experience, in revealing the conceptual-content dominant embedded in such fragments. The established regularity applies, in particular, to personal encyclopaedias (Zhulynskyi *et al.*, 2015) of an artist's legacy, and the role of titles is regularly performed by intertextual structures, symbols, cryptonyms, authorial names in the form of a metaphor, oxymoron, antithesis, etc., for articles devoted to an in-depth analysis of artistic, pictorial or any other works with the same name: "The Wind Converses with the Grove", "O, Boisterous Wind!", "P.S." → poems included in "Kobzar" (Zhulynskyi *et al.*, 2015); "Parable of the Labourers in the Vineyard", "Lost at Cards" → drawings by Taras Shevchenko (Zhulynskyi *et al.*, 2015). In such a case, not only the nominative character characteristic of this factual situation is actualised, but also implicitly or explicitly declared predicativity.

According to content filling, encyclopaedic biblionyms are integrated into thematic groups. The strategic identifying factor is the archisemes "human", "nature", "human and nature", which undergo internal sememe-seme differentiation. The titles of each group that directly nominate a concrete denotatum are united in subgroups, which can likewise be subdivided into narrower lexico-semantic formations, by this semasiological procedure manifesting maximally specified relevant information. The type of differentiation under consideration is conditioned by the generic essence of encyclopaedic discourse.

The anthroponymic title as an obligatory compositional attribute of the encyclopaedic article. The regular filler of the thematic group "human" is the anthroponymic formula of a one-, two- or three-component structure. In the biblionymic paradigm, alongside official anthroponyms there exist unofficial denominations – pseudonyms, nicknames, cryptonyms, gyneconyms, patronyms, etc. – which have not become a separate subject of this study. The main reason for unofficial naming is the desire of precedent-setting, extraordinary personalities to be identified by other anthroponymic indicators. In particular, this is the conceptualisation of unofficial names in various spheres of being, most often in the literary, cultural and artistic. V. Soprykina & O. Malenko (2021) noted that "an ordinary person may not know who Pavlo Hubenko is, but through the pseudonym Ostap Vyshnia that person will realise that it is a Ukrainian humourist writer, prose author, author of "smiles", etc. Who Larysa Kosach is will be confidently said by philologists or pupils, whereas for the majority this figure is emblematic under the pseudonym Lesia Ukrainka. The same happens with the anthroponyms Mykola Fitylov – Mykola Khvylovy; Mariia Vilinska – Marko Vovchok; Panas Rudchenko – Panas Myrnyi. In these situations, the pseudonym performs the function of concealing a person's real proper name". De-actualisation, concealment, non-publicising of the official name in its broad sense has various character, depending on the degree of changeability of the anthroponymic formula (only the first name and surname, or the entire component composition); the degree of anthroponymic correlation of the official name, in its various parametric transformations, with unofficial ones; and so forth. The designation "absolutely independent titles" is fixed for official and unofficial biblionyms.

The one-component anthroponymic model is represented by two sub-models: "surname" and "first name". The surname is "of great weight for every bearer of a surname"; it appears as "the principal accent of each individual, which was not the case in the Middle Ages. At that time, the given name of an

individual was still emphasised at this or that point of contact in society" (Irklivskyi, 1987). The peculiarity of the "surname" sub-model lies in the fact that its filler usually has the plural form. The content of articles whose definitive component is the sub-model under description is most often milestones in the biography of a particular family from the point of view of its history or present. The entries described are of two types: such entries contain systematic information about people who descend from one ancestor, and such entries provide only general information which is specified in articles with other models of the anthroponymic formula – "the same surname + first name" or "the same surname + first name + patronymic". Cf.: 1) Halahany (Halahany) – a Cossack-noble family originating from Hnat (see), a colonel, and the brother Semen, the acting colonel of Myrhorod in the Crimean campaign of 1736; the most famous – Hryhorii (1716-77), colonel of Pryluky (1739-63), participant in campaigns against the Turks and in Prussia, and, above all, public figure Hryhorii (see); after the death of the son Pavlo, the family died out (Kubiiovych, 2003); 2) Muraviov-Apostols, a noble family in Poltava region, which on the female line descended from Hetman D. Apostol. → Muraviov-Apostol Ivan (1770-1851), Russian statesman (ambassador in Hamburg, Copenhagen, and Madrid; senator) and writer (principal work "A Journey through Taurida in 1820", 1823)...; Muraviov-Apostol Ippolit (1806-26), a Decembrist, son of Ivan M.-A. (see), Russian staff officer...; Muraviov-Apostol Matvii (1793-1886), son of Ivan M.-A. (see), Russian lieutenant-colonel...; Muraviov-Apostol Serhii (1796-1826), a Decembrist, son of Ivan M.-A. (see), Russian lieutenant-colonel... (Kubiiovych, 2003). A separate type is the defective sub-model in which the titular first name and patronymic are not established due to temporal remoteness, lack of documentation or absence of an adequate factual base, the insignificance of the personage, sporadic presence in a given sphere of social existence, and other weighty reasons. This is usually reported in the entry: Akerman (first name, patronymic, years of birth and death unknown) – a landowner from the Myrhorod area, a passionate admirer of Shevchenko's poetry (Rotach, 2013). The phenomenon of defectiveness is also characteristic of two-component anthroponymic formulas in which the non-verbalised member could not be established for the serious reasons presented above.

The lexico-semantic range of the "first name" sub-model is quite wide. Its core is constituted by precedent names of pagan [Biloboh, Vlos, Yarylo] and Christian religions (God the Father, God the Son), of Greek and Roman mythology (Athena, Hera, Faun), biblical personages [Abel, Adam, David, Eve, Cain, Petro, Pavlo, Yahweh (Yahveh, Jehovah)], mythical beings representing various spheres of human activity (poetry, eloquence, art, etc.) [Abeona, Herma, Orpheus], thinkers of different times (Aristotle, Democritus, Socrates), founders of different forms of social consciousness, social movements, trends, etc. [Augustine, Buddha, Muhammad (Magommed)], prophets [Veleda, Jeremiah, Moses (Moshe)] (Shevchenko, 2004): Veles, Volos – the Slavic god of plenty and gold, patron of wizards, seers, soothsayers, domestic cattle and "flocks of clouds", trade, commerce; God the Father ... – in Christianity – the first person (hypostasis) of the Divine Trinity; Aristotle (384-322) – an ancient Greek philosopher and scholar-encyclopaedist; Jeremiah – the second great prophet of the Old Testament; Orpheus ... – in Greek myth. – the son of the muse Calliope, a Thracian singer and musician whose art's magical power was obeyed by people, gods, beasts and birds, and nature itself. It is worth distinguishing separately publications whose register is formed on the basis of one-component anthroponymic formulas. These are studies whose object of treatment is the already mentioned Byzantine, ancient Greek, ancient Roman, ancient Hebrew, pagan, Slavic and other names (Shevchenko, 2004), the names of

Ukrainian collectors and bearers-performers of folklore (singers, lyre-players, kobzars, storytellers, narrators, etc.) (Dmytrenko, 2020), as well as individual (Iskra, Kim, Lenera, Oktyabryna, Vladlen) (Triiniak, 2005), official (full) or unofficial (incomplete) male and female names (Nadiia, Nadiienka) (Skrypnyk & Dziatkivska, 2005), surnames of various origins and structures (Symonenko, Ivanchuk, Khomyshyna) (Irklivskyi, 1987).

The two-component anthroponymic formula is constituted by the first name given to a person after birth or consciously changed in adulthood and the hereditary surname borne by other members of the family and which, at the bearer's wish, may also be officially altered in adulthood. Official titles "surname + first name" are traditional in foreign encyclopaedography: Naumenko Volodymyr (1852-1919), public and educational figure, literary scholar, ethnographer, philologist, a native of Novhorod-Siverskyi (Kubiiovych, 2003). Ukrainian compilers and authors of encyclopaedias also use this way of presenting the definitive part.

The three-component anthroponymic formula is represented by the first name given to a person after birth or consciously changed in adulthood, the hereditary surname and the patronymic: Naumenko Volodymyr Pavlovych [7(19) VII 1852, the town of Novhorod-Siverskyi, now Chernihiv Oblast – 3 VII 1919, Kyiv] – a Ukrainian teacher, philologist, journalist, ethnographer, public-political figure, corresponding member of the South-Western Department of the Russian Geographical Society, of the Shevchenko Scientific Society (Lviv) (Rusanivskyi *et al.*, 2007). This model was actively used in Soviet times and continues to be consistently employed in Ukrainian encyclopaedic studies. The type of anthroponymic nomination analysed has advantages: it is more informatively relevant, more convenient to use, since by means of the patronymic it more clearly identifies personalities – primarily those who have the same first names and surnames – who moreover belong to the same or related spheres of activity and intersect in earthly time-space.

A special category singles out anthroponymic formulae-headings of encyclopaedic entries with an identical verbalisation set, which are arranged in the register usually by date of birth. In such a case, there arises a need to identify a person by those criteria which either represent the idea of self-perception or indicate the relation to other people. These include, in particular, "individual features of a person, that is, unique physical and biological properties, facts of life", and social identity – "the typification of a person by the attributes of the social group to which the person belongs and whose members the person perceives as the own, may denote the relations with people as "we" (by age, sex, education, profession, nationality, place of residence, etc.)" (Yermolenko *et al.*, 2023). The isolated criteria are differentiated into two types: family personal formations – a conscious, predictable, desired way of naming caused by family aspirations, and non-family – an accidental coincidence of first name, surname, and patronymic. The distinguished propriative formations require additional criteria of identification outside the encyclopaedic discourse. The most typical of these is "type of activity": Lysenko Oleksandr Yakymovych (06.05.1920, village of Bohoslovske, now Balakhonovske, Stavropol Krai, Russian Federation – 08.01.1944, buried in the village of Nakhov, reburied in the city of Kalinkavichy, both in Homel Oblast, Belarus) – a military serviceman; and Lysenko Oleksandr Yakymovych (05.05.1925, village of Hannivka, now in Petrove District, Kirovohrad Oblast – 19.04.1990, Kyiv) – a philosopher (Dziuba *et al.*, 2022).

If this marker coincides, others are used – "dates of life", "place of residence", "ethnic affiliation", "special achievements", etc.: Il'yin Ivan Petrovych (1888, village of Myrtsiv, now the village of Cherneta, Orzhytsia District, Poltava Oblast – ?) – a clergyman; and Il'yin Ivan Petrovych (1900, village of Mykhnivtsi, now

Lubny District, Poltava Oblast – ?) – a clergyman (Bilousko, 2009). Some encyclopaedic publications also record distinctions such as "first, 1st / second, 2nd", "elder / younger", etc.: Saharda Mykola Ivanovych, 1st (01.12.1870, township of Zolotonosha, Poltava Governorate, now a city, district centre of Cherkasy Oblast – ?) – a teacher; and Saharda Mykola Ivanovych, 2nd (?, township of Hlobyno, Poltava Governorate, now a city, district centre of Poltava Oblast – 26.04.1911) – a philanthropist (Bilousko, 2009); Kholodnyi Petro, the elder (18.12.1876, Pereiaslav – 7.06.1930, Warsaw), a prominent impressionist painter with a penchant for lyricism and neo-Byzantinism, by profession a physico-mechanic; and Kholodnyi Petro, the younger (22.7.1902, Kyiv), a painter and graphic artist, son of Petro the elder (Kubiiovych, 2003). Superscript Arabic numerals are also used as identifiers. These numerals help to highlight personalities in other articles of the same publication where such personalities are mentioned: Abel1 (from Heb., possibly from Akkad. – "son") – according to the Bible, the son of Adam and Eve, who was killed by the elder brother Cain; Abel2 (secular name Vasyliev Vasyl; 1757-1841) – a monk, a seer (Shevchenko, 2004).

For encyclopaedic articles with an anthroponymic heading, anthroponymic-family cross-reference is typical, with or without a hyperlink, according to such identification criteria as "degree of kinship" and "degree of affinity". The degree of kinship presupposes the grouping of blood relatives – those who descend from a common ancestor. Full kinship is distinguished ["father ↔ son (daughter)", "mother ↔ daughter (son)"]; kinship of varying degrees between grandparents and grandchildren ["grandfather ↔ grandson (granddaughter)", "great-great-grandmother ↔ great-great-granddaughter (great-great-grandson)"...], collateral kinship of varying degrees ["brother ↔ brother (sister)", "aunt ↔ second niece (second nephew)"...], "age" ("elder ↔ younger"). These are recorded in all entries or in the one that is second in the register and contains an anthroponymic component with the semantics "a person of male/female sex in relation to the parents, blood or non-blood relatives". Active cross-references are not always used in such cases, since there is no particular need: the entries are placed one after another. The advantages of a technology that displays information about kinship and affinity in all articles lie in the fact that it activates receptive engagement not only in the process of continuous study of an encyclopaedic publication, but also of selective acquaintance with it, conditioned by the extraction of specifically segmented information.

As for the marker "degree of kinship, affinity", it is consistently manifested in articles with anthroponymic formulae-titles that have a partial onymic coincidence, usually of the surname, or even a completely different structure. The technology of through-fixation is applied to it in the encyclopaedic publication, since the entries may be placed contiguously or at a distance. In the proprial field "Liashenko", represented in the "Encyclopaedia of Modern Ukraine", the following same-surname and different-surname verbal models of active cross-reference are recorded: "brother1 ↔ brother2" / "brother2 ↔ brother1", "father ↔ daughter" / "daughter ↔ father": Liashenko Borys Mykolaiovych (11.10.1948, city of Bohuslav, Kyiv Oblast – 04.12.2015, Zhytomyr) – a specialist in applied mathematics and informatics. Brother of I. Liashenko ↔ Liashenko Ihor Mykolaiovych (06.09.1939, Kyiv – 04.05.2012) – a specialist in applied mathematics and informatics. Father of Olena, brother of Borys Liashenkos ↔ Liashenko Olena Ihorivna (15.10.1968, Kyiv) – a specialist in economic cybernetics. Daughter of I. Liashenko; Liashenko Hennadii Ivanovych (01.06.1938, town of Nohaisk, now the city of Prymorsk, Zaporizhzhia Oblast) – a composer, teacher, musicologist, and musical public figure. Father of M. Denysenko ↔ Denysenko Maryna Hennadiivna (29.01.1962, Lviv) – a composer,

teacher. Daughter of H. Liashenko; "father ↔ son" / "son ↔ father": Liashenko Ivan Fedorovych (14.02.1927, village of Mykolaivka, now Stanytsia-Luhanska District, Luhansk Oblast – 12.12.1998, Kyiv) – a musicologist, teacher, musical public figure. Father of O. Liashenko ↔ Liashenko Oleksandr Ivanovych (16.03.1950, Lviv) – a teacher. Son of I. Liashenko; "husband ↔ wife" / "wife ↔ husband", "father ↔ son" / "son ↔ father": Liashenko Tetiana Vasylivna (29.10.1944, city of Petropavlovsk-Kamchatsky, Khabarovsk Krai, Russian Federation) – a specialist in computer-based building materials science and experimental-statistical modelling. Wife of V. Voznesensky ↔ Voznesensky Vitalii Anatoliiovych (08.01.1934, city of Rostov-on-Don, Russian Federation – 07.04.2012, Odesa, buried in Kyiv) – a specialist in building materials science and experimental-statistical modelling. Son of A. Voznesensky, husband of T. Liashenko ↔ Voznesensky Anatolii Oleksandrovych (19.04(02.05).1901, city of Alatyr, Simbirsk Governorate, now Chuvashia, Russian Federation – 13.04.1968, Kyiv) – a specialist in heat engineering. Father of V. Voznesensky (Dziuba *et al.*, 2022).

Within the structure of a single family entry whose principal figure is the oldest representative of the lineage, all models of degrees of blood kinship are provided via hyperlinks. The analysed phenomenon is clearly demonstrated by the proprial field "Paton", in which two figures are iconic: Paton Yevhen Oskarovych and Paton Borys Yevhenovych. In the "Encyclopaedia of Ukrainian Studies" a single model – "son ↔ father" – is recorded in the form of a verbalised structure "son of Yevhen P.": Paton Borys (*1918), son of Yevhen P., a scholar and political figure, metallurgist, and researcher of electric welding… In the article "Paton Yevhen", which is placed in the register immediately after the entry "Paton Borys", the active cross-reference "father ↔ son" is absent: Paton Yevhen (1870-1953), a distinguished scholar in the field of bridge construction and electric welding… (Kubiiovych, 2003). In the electronic version of the "Encyclopaedia of Modern Ukraine", which appeared later than the "Encyclopaedia of Ukrainian Studies", new genealogical data are attested; in addition to the model already analysed, four more are distinguished, represented by eight specific markers: Paton Yevhen Oskarovych (04.03.1870, city of Nice, France – 12.08.1953, Kyiv) – a specialist in electric welding and bridge construction. Father of Borys and Volodymyr, grandfather of Yevheniia Patons. = 1) "father ↔ son1" / "son1 ↔ father": Paton Yevhen Oskarovych ↔ Paton Borys Yevhenovych (27.11.1918, Kyiv – 19.08.2020, ibid.) – a specialist in metallurgy and metal technology. Son of Yevhen, brother of Volodymyr, father of Yevheniia Patons; 2) "father ↔ son2" / "son2 ↔ father": Paton Yevhen Oskarovych ↔ Paton Volodymyr Yevhenovych (18.03.1917, Kyiv – 28.02.1987, ibid.) – a specialist in welding. Son of Yevhen, brother of Borys Patons; 3) "brother1 ↔ brother2" / "brother2 ↔ brother1": Paton Borys Yevhenovych ↔ Paton Volodymyr Yevhenovych; 4) "grandfather ↔ granddaughter" / "granddaughter ↔ grandfather": Paton Yevhen Oskarovych ↔ Paton Yevheniia Borysivna (12.03.1956, Kyiv – 23.12.2009, ibid.) – a biologist. Daughter of Borys, granddaughter of Yevhen Patons; 5) "father ↔ daughter" / "daughter ↔ father": Paton Borys Yevhenovych ↔ Paton Yevheniia Borysivna. The spectrum of highlighted active cross-references can be complemented by a sixth model indicating a more distant type of kinship – "uncle ↔ niece" / "niece ↔ uncle": Paton Volodymyr Yevhenovych ↔ Paton Yevheniia Borysivna (Dziuba *et al.*, 2022).

Markers of kinship such as "grandniece" ("second niece"), "third cousin", "distant relative of X", "descendant of X", "great-grandson through X's brother", "great-great-grandson through the sister's line", "from the family of X" function regularly in personal encyclopaedias whose object of treatment is a unique figure in the multifaceted dimension, one component

of which is a talented, extraordinary family environment: Vidomenko Oleksandr Andronykovych (7.07.1931, village of Moryntsi, Zvenyhorod District, Cherkasy Oblast) – a Ukrainian literary scholar, publicist, local historian, and public figure. A distant descendant of Shevchenko (the poet's great-grandnephew through the brother Y. Shevchenko)... (Zhulynskyi *et al.*, 2015); Korsun Mariia Davydivna (from the Horlenko family; 1812-1886) – a landowner of the village of Velyka Krucha, Pyriatyn County (now in the same district), an acquaintance of T.H. Shevchenko... (Rotach, 2013).

In the family hierarchy, discrepancies and inaccuracies sometimes occur, which violate the logic of the thematised encyclopaedic presentation and impoverish or distort a specific plot line or the real picture of the genealogy. Such undesirable moments arise mainly when the entries united by a single conceptual-genealogical dominant are written by different authors. In M. Zhulynskyi *et al.* (2015) P. Lobas, in the articles "Alchevska Khrystyna Danylivna", "Alchevska Khrystia Oleksiivna", distinguishes three verbal models of the degree of blood kinship: "mother ↔ daughter" / "daughter ↔ mother", "mother ↔ son" / "son ↔ mother", "sister ↔ brother" / "brother ↔ sister": Alchevska Khrystyna Danylivna (4/16.04.1841, Borzna, now a district centre of Chernihiv Oblast – 15.03.1920, Kharkiv) – a Ukrainian public figure, teacher, publicist, member of the Kharkiv "Hromada". Mother of Kh. O. Alchevska and I. Alchevskyi; Alchevska Khrystia Oleksiivna (Khrystyna; 4/16.03.1882, Kharkiv – 27.10.1931, ibid.) – a Ukrainian writer, publicist, teacher. Daughter of Kh.D. Alchevska. In the same publication, the author of the articles "Alchevskyi Hryhorii Oleksiiovych" and "Alchevskyi Ivan Oleksiiovych", I. Lysenko, operates with a broader repertoire of the described models: "mother ↔ daughter" / "daughter ↔ mother", "mother ↔ son1" / "son1 ↔ mother", "mother ↔ son2" / "son2 ↔ mother", "sister ↔ brother1" / "brother1 ↔ sister", "sister ↔ brother2" / "brother2 ↔ sister", "brother1 ↔ brother2" / "brother2 ↔ brother1", providing information that is more relevant from the viewpoint of family relationships and the social role of the individuals: Alchevskyi Hryhorii Oleksiiovych (1866, Kharkiv – 1920, Moscow) – a Ukrainian composer, vocal pedagogue, pianist. Son of Kh. D. Alchevska, brother of I. Alchevskyi and Kh.O. Alchevska; Alchevskyi Ivan Oleksiiovych (15/27.12.1876, Kharkiv – 27.04/10.05.1917, Baku; buried in Kharkiv) – a Ukrainian and Russian opera and chamber singer (lyric-dramatic tenor). Son of Kh.D. Alchevska, brother of H. Alchevskyi and Kh.O. Alchevska.

All components of a two- or three-component anthroponymic formula must be reproduced exactly in the encyclopaedia. Such feedback-remarks were expressed during a survey conducted by the Institute of Encyclopaedic Research of the National Academy of Sciences of Ukraine on the occasion of the 20th anniversary of the appearance of the first volume of the "Encyclopaedia of Modern Ukraine" by users of the e-ESU (Dziuba, 2021). In real life, cases of the parallel existence of an official/unofficial surname, first name, or patronymic are not uncommon. Sometimes such a naming practice is recorded in writing, including in authoritative sources. The politician and human rights activist Stepan Khmara is called Stepan Ilkovych and – occasionally – Stepan Illich (corresponding to Ilko / Illia; according to the passport – Illich): ... Without background knowledge, it is difficult to establish causal links in the anthroponymic formula "wife of Vasyl Verkhovynets", represented in R. Rotach (2013) by two entries, with the principle of cross-referencing applied from the first: 1) Verkhovynets-Kosteva Yevdokiia Ivanivna (1885-1987) – wife of V. Verkhovynets, an actress. See the article "Dolia Yevdokiia Ivanivna"; 2) Dolia-Verkhovynets (Verkhovynets-Kostiv) Yevdokiia Ivanivna (née Pohorila; 1.03.1885, Poltava – 2.11.1988, ibid.) – a Ukrainian dramatic actress, wife of the composer and conductor V. Verkhovynets.

The content of the nomination is partially decoded by the article about the outstanding artist placed in this publication next to the analysed one (1): Verkhovynets Vasyl (real surname – Kostiv Vasyl Mykolaiovych; 5.01.1880, village of Staryi Mizun, now Ivano-Frankivsk Oblast – 11.04.1938, Kyiv, executed) – a Ukrainian composer, conductor, choreographer, folklorist, ethnographer (Rotach, 2013). The issue of variation at the level of the biblionym-anthroponym is the subject of another serious scholarly study. In this context, it is important to emphasise that official titular nominations explicitly harmonise in the encyclopaedographic essence with such methodological principles of the encyclopaedic article as objectivity, accuracy, argumentation, and maximal informative correctness.

Conclusions

The concrete representative of encyclopaedic discourse and the basic structural unit of an encyclopaedic publication is the encyclopaedic article, which has a coherent compositional structure (title, definitive-textual part, bibliography, author), whose strategic component is the communicative complex "title ↔ text". In terms of content-type characteristics, articles differ from one another and at the same time have universal features in the form of such textual codes as informative relevance, systemacity, logicality, causal regularity, conciseness, density, internal or external intertextuality, connotative indifference.

The introductive link in the algorithm of the reception of any text is the pre-textual unit, or title, which stands in hierarchically explicated prospective-retrospective relations with the main text. The components of the communicative complex "title ↔ text" are characterised by a two-plane cognitive-discursive trajectory: "from the title to the text" and "from the text to the title", the manifestation of which is conditioned by the genre nature of the concrete representative of encyclopaedic discourse. The autosemantic-pragmatic pre-textual units of encyclopaedic articles, which, being arranged according to different principles, form a universal or sectoral encyclopaedic publication, have the own specificity. The spectrum of functions assigned to such units is limited: integrative, distinctive, nominative, representative, conceptual-content, aconnotative. The motivational basis is strictly regulated for biblionymic titles: by the onomasiological structure, the titles convey real information about a specific denotatum. As for the cognitive-discursive trajectory, it is characterised by a one-way movement – "from the title to the text".

The thematic paradigm of encyclopaedic titles is represented by three thematic groups: "human", "nature", "human and nature". A special place in it is given to anthroponymic biblionyms, realised in the form of one-, two- and three-component anthroponymic formulae. The first and second formulae are differentiated into structures narrower in terms of component composition: "first name", "surname". Each of these formulae has its own features, not only at the structural level, but also at the level of informative and even lexico-semantic content, which has been consistently revealed in the sub-paradigm of Byzantine, pagan, Slavic, and other widely known, as well as individual, personal names. According to the criterion of two-/three-component structure, encyclopaedic publications issued abroad (surname + first name) contrast with those in the former Soviet Union and independent Ukraine (surname + first name + patronymic). The practice of two-component nomination has also been actively applied in Ukraine since 1990. The advantages of three-component formations lie in the fact that such formations provide more specific information about the person. The identification of persons with the same verbalisation set, who represent family and non-family personal formations, is facilitated by such markers of the definitive part as "dates of life", "place of birth (residence)", "ethnic affiliation", "social status", "special

achievements", as well as verbalised distinctions such as "first / second", "elder / younger" or the substitutes – superscript Arabic numerals.

The specificity of encyclopaedic articles with an anthroponymic heading is determined by the anthroponymic-family cross-reference typical of such articles, which is based on the identifying features "degree of kinship" and "degree of affinity". Separately and together, these features convey more specific plot information about a given person and the family and genealogy. Real biblionyms explicitly reflect the content of the main part and directly correlate with the spectrum of its cross-references. Unofficial names require the introduction of additional comments to ensure informative completeness and correctness. An exception in this context is constituted by the so-called precedent, unofficial substitutes for a real anthroponym. The prospect for further research lies in the use of the developed technology to study pseudonyms, nicknames, cryptonyms, gyneconyms, patronymics and other unofficial anthroponymic biblionyms, as well as the titles of such thematic groups as "nature", "human and nature", primarily from the perspective of cognition, reception, structure, and the specifics of the manifestation of external and internal determinative relations.

Acknowledgements

None.

Funding

None.

Conflict of Interest

None.

References

[1] Berezovska-Savchuk, N. (2023). Editing reference works. In Zh. Koloiz (Ed.), *Editorial practice: Current state and development prospects* (pp 159-192). Kryvyi Rih: KSPU.

[2] Bilousko, O.A. (Ed.). (2009). *Poltavika: The Poltava encyclopedia*. Poltava: Poltavskiy Literator.

[3] Chernysh, N. (2018). *Studies in the history of Ukrainian encyclopedic tradition*. Lviv: LA "Piramida".

[4] Dmytrenko, M. (Ed.). (2020) *Ukrainian folklore encyclopedia*. Kyiv: Institute of Art History, Folklore and Ethnology of the National Academy of Sciences of Ukraine (NASU).

[5] Dziuba, I. (Ed.). (2021). *Ukrainian electronic based encyclopedias in light of socio-humanities*. Kyiv: Institute of Encyclopaedic Research of the National Academy of Sciences of Ukraine (NASU). doi: 10.37068%2Fb%2F9789660296770.

[6] Dziuba, I.M., Zhukovskyi, A.I., Zhelezniak, M.H., Andrieiev, V.L., Boriak, O.V., Haran, O.M., Zadyraka, V.A., Kyselov, O.I., & Yatskiv, Ya.S. (Eds.). (2022). *Encyclopedia of modern Ukraine*. Kyiv: Institute of Encyclopaedic Research of the National Academy of Sciences of Ukraine (NASU).

[7] Irklivskyi, V. (1987). *Etymological dictionary of Ukrainian surnames*. Munich: Ukrainian Fraternal Union.

[8] Ishchenko, O., & Savchenko, O. (2023). How to write a Wikipedia article. *The Encyclopedia Herald of Ukraine*, 15, 78-86. doi: 10.37068/evu.15.7.

[9] Kovalchuk, H. (2023). Encyclopedias and lexicons of the 15-18th centuries in the collections from the Old and Rare Books Department of the Vernadsky National Library of Ukraine. *The Encyclopedia Herald of Ukraine*, 15, 9-24. doi: 10.37068/evu.15.2.

[10] Kubiiovych, V. (Ed.). (2003). *Encyclopedia of Ukrainian studies*. Lviv: Taras Shevchenko Scientific Society in Lviv.

[11] Kyrydon, A.M. (Ed.). (2020). *Encyclopedias as glocal media.* Kyiv: State Scientific Institution "Encyclopaedic Publishing House".

[12] Lebed, I., & Bilohrats, Kh. (2020). Headings of publications in the media on the example of the newspaper "Den'" and magazine "Krayina". *Bulletin of Lviv Polytechnic National University. Series: Journalistic Sciences*, 4, 63-67. doi: 10.23939/sjs2020.01.063.

[13] Li, X., Thelwall, M., & Mohammadi, E. (2021). How are encyclopedias cited in academic research? Wikipedia, Britannica, Baidu Baike, and Scholarpedia. *Profesional De La Información*, 30(5). doi: 10.3145/epi.2021.sep.08.

[14] Lin, S.-K. (2020). Encyclopedia – a peer-reviewed scholarly community project. *Encyclopedia*, 1(1), 1-2. doi: 10.3390/encyclopedia1010001.

[15] Monakhova, T.V., Tuluzakova, O.H., & Monakhov, S.D. (2021). Effective headlines as the problem in media industry and an object of linguistics. *Scientific notes of V.I. Vernadsky Taurida National University. Series: Philology. Journalism*, 32(2(2)), 258-262. doi: 10.32838/2710-4656/2021.2-2/45.

[16] Moncomble, F. (2017). Beauty and the Tweet: How traditional media seduce in the digital age. *E-Rea*, 15(1). doi: 10.4000/erea.5917.

[17] Mykytiuk, O.R. (2025). *Political linguistic personology: Dmytro Dontsov's political lingual personality as a phenomenon*. (Doctoral dissertation, Odesa I.I. Mechnikov National University, Odesa, Ukraine).

[18] Petrushka, A.I. (2021). Social media: Academic integrity in the system of scientific communication. *State and Regions. Series: Social Communications*, 3(47), 73-79. doi: 10.32840/cpu2219-8741/2021.3(47).10.

[19] Rarytskyi, O.A. (2017). *Non-fiction prose by the Ukrainian writers of the 1960s: Genre specificity and poetics*. (Doctoral dissertation, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine).

[20] Rotach, P. (Ed.). (2013). *Poltava Shevchenkiana: An attempt at a regional Shevchenko encyclopedia*. Poltava: Dyvosvit.

[21] Ruban, A. (2024). The title as a constituent and structural component of T.G. Shevchenko's poetic works: General characteristics, specifics. *The Journal of V.N. Karazin Kharkiv National University. Series "Philology"*, 94, 13-18. doi: 10.26565/2227-1864-2024-94-02.

[22] Rusanivskyi, V.M., *et al*. (Eds.). (2007). *Ukrainian language. Encyclopedia*. Kyiv: Mykola Bazhan All-Ukrainian State Specialised Publisher "Ukrainian Encyclopedia".

[23] Selivanova, O.O. (2011). *Linguistic encyclopedia*. Poltava: Dovkillia-K.

[24] Serdega, R. (2020). The structure of dictionary entries in the dictionary of agricultural vocabulary of central Slobozhanshchyna (Kharkiv Region). *Linguistics*, 1(42), 58-70. doi: 10.12958/2227-2631-2020-1-42-58-70.

[25] Shevchenko, V.M. (2004). *Dictionary-reference guide to religious studies*. Kyiv: Naukova Dumka.

[26] Skrypnyk, L.H., & Dziatkivska, N.P. (2005). *Proper names of people: Slavic-reference*. Kyiv: Naukova Dumka.

[27] Soprykina, V., & Malenko, O. (2021). *Pseudo-onymikon of the Ukrainian cultural and artistic sphere of the early 21st century: A linguistic interpretation*. Kharkiv: Kharkiv Linguistic School.

[28] Triiniak, I.I. (2005). *Dictionary of Ukrainian names*. Kyiv: Dovira.

[29] Vakaliuk, T.A., Chernysh, O.A., & Syvak, O.B. (2021). Dictionary entry genre. *New Philology*, 1(81), 60-65. doi: 10.26661/2414-1135-2021-81-1-9.

[30] Wunderlich, D. (1976). *Studies on speech act theory*. Frankfurt am Main: Suhrkamp.

[31] Yatsyshyn, A.V., Burov, O.Iu., & Nosenko, Yu.H. (2021). Online encyclopedias: Modern state and development prospects. In *Academic youth-2021*: *Proceedings of the 9th all-Ukrainian scientific and practical conference of young scholars* (pp. 148-156). Kyiv: Institute for Digitalisation of Education of the NAES of Ukraine.

[32] Yermolenko, S.Ya., Bybyk, S.P., Hanzha, A.Yu., Kots, T.A., & Siuta, H.M. (2023). *Linguosophy of Ukrainian texts of the 21st century*. Kyiv: Institute for the Ukrainian Language of the NAS of Ukraine.

[33] Yeshchenko, T.A. (2021). *Literary text phenomenon: Communicative, semantic and pragmatic aspects*. Retrieved from https://surl.li/hrmlfh.

[34] Yuldasheva, L.P. (2018). *Structural-semantic and pragmatic features of the headings of the works of the late XX-early XXI centuries*. (PhD dissertation, Bogdan Khmelnitsky Cherkasy National University, Cherkasy, Ukraine).

[35] Yuldasheva, L.P., Kobylina, Yu.M., & Kolomynska, T.B. (2022). Pragmatic features of the titles of modern Ukrainian literature. *Scientific notes of V.I. Vernadsky Taurida National University. Series: Philology. Journalism*, 33(1(1)), 83-90. doi: 10.32838/2710-4656/2022.1-1/15.

[36] Zhelezniak, M., Ishchenko, O., & Bortnyk, C. (2023). Ukrainian encyclopedias and dictionaries in view of spreading geological knowledge. *Visnyk of Taras Shevchenko National University of Kyiv. Geology*, 4(103), 19-25. doi: 10.17721/1728-2713.103.03.

[37] Zhulynskyi, M.H. *et al*. (Eds.). (2015). *Shevchenko encyclopedia*. Kyiv: National Academy of Sciences of Ukraine.

Композиція енциклопедійної статті в проєкції на бібліонімний зріз: когніція, структура, специфіка детермінативних відношень

Микола Степаненко

Доктор філологічних наук, професор
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
https://orcid.org/0000-0002-6727-1265

Ніна Степаненко

Кандидат філологічних наук, доцент
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
https://orcid.org/0000-0002-0934-784X

Анотація. Актуальність пропонованої наукової студії зумовлена спрямованістю сучасної лінгвістики на вивчення енциклопедознавства як поліпарадигмальної науки в проєкції її модерного наповнення, технічного оснащення та процесу енциклопедографування – створення енциклопедійних словників, опрацювання їхніх теоретичних засад. Метою роботи було комплексне дослідження специфіки взаємодії претекстової (заголовкового комплексу) та власне текстової структури енциклопедійної статті як основної одиниці енциклопедійного дискурсу. Достовірність отриманих результатів забезпечило використання загальнонаукових і лінгвістичних методів: індукції та дедукції, описового, описово-аналітичного, порівняльного, біографійного, словникових дефініцій, контекстного й компонентного аналізу. На основі аналізу фольклористичних, лінгвістичних, літературознавчих, історичних енциклопедійнних видань, спеціалізовано-регіонально-біографічних, галузево-регіонально-біографічних та інших джерел проінтерпретовано релевантні ознаки енциклопедійного тексту, витлумачено класичну композицію енциклопедійної статті як основної структурної одиниці універсальних або галузевих енциклопедійних видань. Основний фокус дослідження сформувала автосемантично-прагматична претекстова одиниця – заголовок, яка системно й послідовно корелює за змістом та засобами мовної архітектоніки з дефінітивно-текстовою частиною. З'ясовано іманентні ознаки (послідовно актуалізована інформативність, жорстка логічність, максимальна конкретність, структурно-семантична лаконічність, недискусійність) та функції (інтегративна, видільна, номінативна, репрезентативна, концептуально-змістова, аконотативна) бібліоніма-антропоніма як обов'язкового конституента енциклопедійної статті – найважливішого актуалізатора постповідомлення, що має специфічну пропозиційно-диктумну мотиваційну базу, жорстко регламентовану структурну організацію, односпрямовану когнітивно-дискурсивну траєкторію – від заголовка до тексту. Установлено засадничі критерії ідентифікування антропонімних бібліонімів з ідентичним вербальним набором. Розкрито антропонімно-родинне перехресне посилання за ідентифікувальними ознаками «ступінь спорідненості» [«повна спорідненість», «спорідненість різного ступеня між дідами (бабами) та онуками», «бокова спорідненість різного ступеня» і под.] та «ступінь свояцтва». Практичне значення проведеного дослідження визначило передовсім те, що його результати можна застосовувати для систематизування, поглиблення, уточнення рекомендацій, які безпосередньо торкнулися якісного потенціалу підготовки енциклопедійної статті

Ключові слова: структурна одиниця енциклопедійного тексту; енциклопедійний претекст і текст; класичні конституенти енциклопедійного тексту; антропонімний претекстовий компонент; різнокомпонентна антропонімна формула; антропонімно-родинне перехресне посилання

International Journal of Philology, 29(3), 2025
Journal homepage: https://philologicalscience.com.ua/en
Received 04.04.2025 Revised 11.08.2025 Accepted 23.09.2025

UDC 373.46 DOI: 10.31548/philolog/3.2025.48

Word-formation mechanisms in Azerbaijani and world languages: A comparative morphological study

Ramila Farajova*
PhD in Philology
Azerbaijan State Marine Academy
AZ1000, 18 Zarifa Aliyeva Str., Baku, Azerbaijan
https://orcid.org/0009-0000-2006-1706

Abstract. This research paper presented an extensive and in-depth comparative examination of word-formation mechanisms in the Azerbaijani language in relation to several major world languages, specifically English, Russian, German, Turkish, and Persian. By situating Azerbaijani within a broader cross-linguistic framework, the study aimed to elucidate both typological particularities and universal patterns in morphological processes. The methodology integrated cognitive morphological analysis, corpus-based investigation, computational and digital resource modelling, and cross-linguistic typological comparison to investigate Azerbaijani morphological structures. The research demonstrated that Azerbaijani preserved its core agglutinative structure while developing hybrid formations through loan-affix integration, showing increased frequency of mixed morphological chains in digital corpora and expanding productive affixation patterns in response to contact-driven lexical influx. Empirical analysis showed that Azerbaijani morphology was both flexible and resilient, capable of generating novel lexical items and accommodating semantic shifts in response to social, technological, and intercultural developments. These findings underscored the dual character of morphological evolution: it revealed itself as universal in its structural tendencies while being uniquely shaped by the cultural and linguistic context of Azerbaijani speakers. By situating Azerbaijani morphology within the comparative landscape of world languages, this study contributed to a deeper understanding of cross-linguistic creativity, typological variation, and the interplay between morphology and sociolinguistic dynamics, offering insights relevant to theoretical linguistics, language teaching, and applied lexicography. The practical value of this research lies in providing linguists, lexicographers, educators, and digital language-technology developers with empirically grounded models of Azerbaijani word-

Suggested Citation:
Farajova, R. (2025). Word-formation mechanisms in Azerbaijani and world languages: A comparative morphological study. *International Journal of Philology*, 29(3), 48-60. doi: 10.31548/philolog/3.2025.48.

*Corresponding author

formation that can be directly applied to dictionary compilation, curriculum design, automated morphology processing, and the development of NLP tools such as morphological analysers and spell-checkers

Keywords: agglutinative structure; typological contrast; lexical innovation; contact-induced change; affix productivity; digital corpora; cross-linguistic creativity

Introduction

Word formation represents the most dynamic layer of linguistic structure, serving as the mechanism through which languages expand conceptual possibilities and encode shifting realities. The study of morphological creativity provides insight into how linguistic systems maintain internal coherence while remaining open to innovation, particularly in contexts where multilingual interaction accelerates the emergence of new forms. The subject matter of this research is justified by the need to understand how Azerbaijani morphology functions as a productive system capable of both preserving structural identity and integrating external influences through affixation, hybridisation, and semantic extension.

Researcher G.Y. Rakhimova (2023) examined derivational productivity in Azerbaijani through corpus-tagged datasets and found that suffixation remained the dominant word-formation mechanism, although hybrid forms constructed from English stems showed measurable growth in social media registers. This study highlighted that productivity could be quantified statistically, demonstrating that digital communication environments contribute strongly to the expansion of derivational categories. Linguist V. Abdullayeva-Nebiyeva (2025) analysed semantic shifts triggered by foreign lexical input and concluded that Azerbaijani speakers assimilate new stems using native derivational models rather than adopting foreign morphology wholesale. According to results, morphological adaptation occurs through cognitive schema integration rather than structural replacement. In a study conducted by B. Özenç *et al*. (2018) the authors explored reduplication as a multifunctional mechanism in Azerbaijani and demonstrated that reduplicative constructions served intensifying, iterative, and stylistic purposes across both colloquial and literary registers. Their findings confirmed that morphological repetition supported pragmatic expressivity and discourse-level nuance.

D. Beck (2017) investigated derivational patterns across academic and journalistic text corpora and established that affixes related to abstraction and agentive meaning, particularly *-lıq/-lik* and *-çı*, exhibited the highest productivity indices. The author concluded that morphological creativity was strongly motivated by social and occupational diversification after 2020. According to international researcher. Researcher M. Huseynova (2019) offered a computational typology model that compared word-formation networks across Kazakh, Turkish, German, and Russian. The study demonstrated that hybrid derivation increased most rapidly in languages with significant digital borrowing pressure, thereby supporting the relevance of computational contrastive analysis. Another contact-focused study was conducted by Finally, in a comparative acquisition-based study, I. Plag (2025) noted that learners of agglutinative languages produce derivational expansions earlier than learners of fusional ones, suggesting that morphological transparency facilitates lexical growth and frequency-driven generalisation.

The reviewed scholarship identified important features of Azerbaijani morphology, but none of the works simultaneously quantified

productivity, compared Azerbaijani with structurally diverse languages, and evaluated contact-induced derivation within a unified analytical model. The aim of this research was to establish a comprehensive cross-linguistic framework for examining how Azerbaijani word-formation systems function, evolve, and interact with contact influences by integrating cognitive, corpus-based, and computational methods.

Materials and Methods

The study employed a comparative, descriptive, typological, and corpus-based methodology to examine Azerbaijani word-formation in a cross-linguistic perspective. Qualitative morphological analysis was combined with quantitative corpus procedures, including morphological segmentation, derivational pattern classification, and frequency-based statistical modelling using contemporary digital corpora. A dedicated historical-philological approach was applied to manuscripts and classical texts through the use of authoritative editions, normalisation of variant forms, and chronological interpretation of interfixes and derivational structures, ensuring that diachronic claims were empirically grounded. The research integrated qualitative description with quantitative corpus metrics, combining morphological segmentation, historical interpretation, and frequency-based statistical modelling.

The analysis was grounded in five interconnected theoretical principles (Booij, 2018). First, a comparative-typological perspective was applied, contrasting Azerbaijani with fusional languages such as Russian and German, analytic structures such as English, and other synthetic systems including Persian and Turkish, in order to determine points of convergence and divergence in affix behaviour. Second, productivity was assessed through established quantitative measures – primarily frequency-per-million indices and type-token ratios – to evaluate the output of derivation, compounding, and reduplication. Third, diachronic development was examined by consulting historical corpora and etymological dictionaries, enabling the tracing of morphological change across four chronological layers spanning 1900 to 2024. (Durkin, 2015; López-Couso, 2016; Tsanko, 2025). Fourth, sociolinguistic dynamics were considered through the analysis of loan-driven innovations documented in digital media corpora, including conversational registers and youth-oriented discourse. Finally, the study adopted a cognitive-functional perspective in which derivational patterns were interpreted as reflections of conceptual categorisation processes and indicators of relative processing efficiency. In response to concerns regarding the representation of hybrid morphological forms such as bio-i-təhlükəsizlik (bio-safety) and other interfix compounds, the present study adopted a multi-step methodology to ensure that analysed forms reflect genuine usage rather than occasionalisms, neologistic experiments, or errors arising from machine translation.

Corpus-based data selection: All morphological forms were extracted from a combination of contemporary corpora, including The Azerbaijani National Corpus (National Corpus of Azerbaijani Language, n.d.) (spanning written, journalistic, and academic texts); Social media and digital communication corpora (Twitter, Instagram, blogs, online news portals); Official administrative and scientific documents from 2000-2025 (Milli Arxiv İdarəsi, n.d.).

Validation criteria required that forms be included only if they appeared in at least three independent sources or texts, demonstrated clear semantic transparency according to Azerbaijani morphological rules, and were consistent with the language's phonological and interfixation patterns, while excluding forms found solely in machine-translated content or arising from individual user idiosyncrasies. Frequency Analysis: Each candidate form underwent quantitative frequency analysis. This confirms functional productivity by demonstrating repeated use across registers (formal, academic, digital) and age groups.

Comparative analysis across formal, informal, and digital registers ensures that identified forms are not artifacts of a single medium but represent active, productive morphology in real communicative contexts. By incorporating these methodological steps, the study substantiates claims regarding hybrid morphology and interfixation in Azerbaijani, providing empirical grounding rather than relying solely on isolated examples. The study drew on four major corpora, including national, multilingual, and media datasets, as summarised in Table 1.

Table 1. *Corpus data*

Source	Token Count	Function	Notes
Azerbaijani National Corpus (n.d.)	50M+	diachronic + literary	tagged, lemmatised
SketchEngine (EN/RU/DE/TR)	6 corpora	cross-language comparison	concordance pulled 2024-2025
Google Books Dataset	10M sample	diachronic expansion	query-filtered by author
Author-compiled Media Corpus (2020-2024)	11.2M	neologism tracking	Telegram, news, Twitter-X

Source: *compiled by the author based on Russian National Corpus (n.d.), British National Corpus (n.d.), Turkish National Corpus (n.d.), Humboldt-Universität zu Berlin, Institute of Linguistics (n.d.), National Corpus of Azerbaijani Language, (n.d.)*

To ensure data reliability and eliminate noise, strict selection criteria were applied to all lexical items retrieved from the corpora, as shown in Table 2.

Table 2. *Selection criteria applied to avoid noise, occasionalisms and MT-generated anomalies*

Criterion	Required Threshold
Min. attested occurrences	⩾50 tokens across ⩾3 sources
Exclusion of errors	MT-detected orthographic patterns removed
Concordance verification	Human-validated across two corpora
Hybrid forms counted only if productive	must form ⩾3 independent derivatives

Notes: *examples such as bio-i-təhlükəsizlik (bio-safety) were excluded from productivity claims because they did not meet the attestation or derivative-branch thresholds*
Source: *compiled by the author*

The data presented in the tables are interpreted and discussed in the running text, as tables serve only as a means of structured presentation and do not, by themselves, constitute analysis. Accordingly, each table in this section is accompanied by an explicit analytical commentary that explains the relevance of the selected criteria, thresholds, and exclusions for the research objectives. This approach ensures that the tabulated data are fully integrated into the methodological and analytical argument of the study rather than remaining descriptive. The analysis drew on four major corpora for Azerbaijani (National Corpus of Azerbaijani Language, n.d.) and comparable corpora for English, Russian, German, and Turkish, as summarized in Table 3. All lexical items included in the analysis were filtered according to the selection criteria outlined in Table 2, ensuring that only forms attested in at least 50 tokens across three independent sources and producing at least three independent derivatives were considered. Forms failing these thresholds, such as occasional or machine-generated items (e.g., bio-i-təhlükəsizlik; bio-safety), were excluded from productivity counts.

Table 3. *Corpora used in the study for diachronic and cross-linguistic analysis*

Language	Corpus	Token Count	Function	Notes
Azerbaijani	ANC + Media	61.2M	Diachronic, neologism tracking	Tagged, lemmatised
English	SketchEngine	25M	Cross-language comparison	Concordances extracted 2024-2025
Russian	RNC	30M	Diachronic + derivational analysis	Tagged, lemmatised
German	DWDS	28M	Diachronic + derivational analysis	Tagged, lemmatised
Turkish	TNC	22M	Diachronic + derivational analysis	Tagged, lemmatised

Source: *compiled by the author based on Azerbaijani National Corpus (n.d.), Milli Arxiv İdarəsi. (n.d.) Russian National Corpus (n.d.), Turkish National Corpus (n.d.), British National Corpus (n.d.), HumboldtUniversität zu Berlin, Institute of Linguistics, (n.d.)*

The study drew on both national and multilingual corpora, including Azerbaijani, Turkish, Russian, and English language datasets, to provide a robust foundation for cross-linguistic analysis and to examine patterns of morphological, phonological, and lexical variation across different languages.

Results and Discussion

Suffixation dominates Azerbaijani word-formation, accounting for 82.4% of all tokens (n = 43,912 types) in the 2020-2024 corpora. As shown in Table 4, the suffixes -lıq/-lik, -çı/-çi, -sız/-siz, and -lı/-li contribute the majority of derivational output. The high concentration of derivational productivity in these four suffixes indicates a core agglutinative engine, where speakers preferentially use these morphemes to generate new lexical items. The prominence of -lıq/-lik, for example, aligns with previous claims regarding nominalisation productivity in Azerbaijani (Huseynova, 2019), but compiled corpus-based data provide measurable confirmation. The diminishing frequency of less productive suffixes such as -varı/-vari demonstrates that, while the language remains agglutinative, productivity is highly skewed toward a limited set of high-yield morphemes, reflecting both historical continuity and contemporary lexical stability.

Table 4. *Frequency of top suffixes in 2020-2024 corpora*

Suffix	Tokens
-lıq/-lik	17,504
-çı/-çi	9,612
-sız/-siz	7,921
-lı/-li	5,332
-varı/-vari	1,543

Source: *compiled by the author*

The table confirmed derivational density concentrated in four high-yield morphemes, supporting M. Huseynova (2019), but replacing citation-only claims with original data. Hybrid derivations – where borrowed stems combine with native affixes – demonstrate measurable productivity. SketchEngine concordances from 2024 (Table 5) show that forms like follow-çu (follower) and blok-lamaq (to block) generate multiple derivatives (e.g., follow-çular, follow-çuluq, bloklanma, bloklayıcı; follower, following, block, blocker), confirming

productive integration into Azerbaijani morphology. By contrast, items such as like-ləmək (to like) fail to reach the three-branch derivational threshold and are therefore non-productive. These results indicate that productive hybridisation depends not only on stem familiarity but also on compatibility with the native morphological system. High fpm values (e.g., 112.6 for blok-lamaq; to block) suggest widespread usage and acceptance in both digital media and everyday speech, supporting the hypothesis that contact-driven innovation can expand the lexicon while respecting agglutinative constraints.

Table 5. *SketchEngine concordance output (2024 scrape)*

Term	fpm	Confirmed Derivatives	Status
follow-çu (follower)	78.3	Followçular (followers), followçuluq (following)	PRODUCTIVE
blok-lamaq (block)	112.6	Bloklanma (blocking), bloklayıcı (blocker)	PRODUCTIVE
koment-çi (comment)	61.1	Komentçilik (commenting)	PRODUCTIVE
like-ləmək (like)	32.4	- no 3-branch derivation	NON-PRODUCTIVE

Source: *compiled by the author*

Although prefixation remains rare, representing only 11.3% of new derivations, its domain-specific usage shows systematic patterns. Table 6 reveals that qeyri- dominates in academic and news registers, anti- in medical and political contexts, and super- and ekstra- primarily in youth discourse and marketing. This distribution demonstrates that prefixation in Azerbaijani is register-sensitive and semantically specialised, often emerging in response to new lexical needs (e.g., technological, sociopolitical, or marketing contexts). While the total number of prefixal tokens is smaller than suffixal derivations, the increasing frequency in specific domains suggests a gradual acceptance and productive adaptation of prefixal strategies, complementing the dominant suffix-based morphology.

Table 6. *Corpus attested prefix counts (2020-24)*

Prefix	Tokens	Register Peak
qeyri-	4,480	academic + news
anti-	2,931	medical + political
super-	1,282	youth + social media
ekstra-	731	marketing

Source: *compiled by the author*

Azerbaijani morphology represents a hybrid and adaptable system, combining traditional agglutinative structures, inherited literary influences, and modern digital innovations. This hybrid system is typologically unique, as it integrates concatenative processes (classic suffixation and interfixation), non-concatenative processes (apophony/internal stem alternation), and zero-derivation/conversion, enabling speakers to generate complex words efficiently while maintaining semantic transparency (Kornfilt, 2002). Interfixation in Azerbaijani serves multiple linguistic functions: phonological linking, semantic integration, stylistic marking, and cognitive facilitation. Historically, interfixes entered Azerbaijani through contact with Persian and Arabic during the development of the literary language. Classical texts and administrative manuscripts frequently show interfixes connecting roots from different languages, e.g., elm-i-fən (science of knowledge), ədəb-i-nəzər (literary theory),

tarix-i-dövlət (history of the state), and qanun-i-mədəniyyət (law of culture). These compounds demonstrate precision, clarity, and aesthetic balance, confirming J. Kornfilt (2002) observations that interfixes preserve formal and semantic integrity.

In the XX century, interfixation expanded to scientific and administrative neologisms, such as bio-i-təhlükəsizlik (biosecurity), texnologiya-i-inkişaf (technological development), virtual-i-muzey (virtual museum), and onlayn-i-təhsil (online education). These examples confirm M. Huseynova (2019) assertion that interfixes are not mere borrowings but partially grammaticalised morphemes, contributing to formal integration and semantic transparency. In digital communication, young speakers further innovate by hybridizing Azerbaijani roots with English or Turkish, e.g., selfi-çubuğu, zoom-məşq, blog-yazı, and e-poçt-məktub (selfie stick, zoom workout, blog post, and email) (Huseynova, 2019). For example, compounds like virtual-muzey (virtual museum) and onlayn-təhsil (online education) show consistent presence in educational and digital corpora, while younger digital users frequently employ hybrid forms such as selfi-çubuğu (selfie stick) or zoom-məşq (zoom workout), validating their integration into contemporary spoken and written Azerbaijani. These forms demonstrate morphological creativity and cognitive efficiency: multi-root compounds are processed as single lexical units, aiding comprehension in fast-paced communication.

Typologically, Azerbaijani interfixation exhibits systematic linking, often guided by vowel harmony, similar to Turkish, and shows functional parallels with German linking elements (e.g., -e-, -s-) in technical compounds. English generally lacks native interfixation in core vocabulary; however, in words of Latin or Greek origin (e.g., speed-o-meter, psych-o-path), connecting vowels perform an interfix-like function, linking stems within complex compounds. This nuance indicates that Azerbaijani's interfixation system is both productive and regular, facilitating cognitive parsing in native compounds, while English relies on historical borrowings and orthographic markers to achieve similar linking effects (Kornfilt, 2002). Sociolinguistic variation is also notable: urban youth employ hybrid digital forms, whereas older or rural speakers retain classical interfixes, showing register-sensitive morphological adaptation. Historical and contemporary corpora confirm that interfixation preserves hierarchical and formal semantic relationships across registers (Nasirova *et al.*, 2023). Reduplication is a highly productive process, serving plurality, intensity, repetition, stylistic nuance, and sociolinguistic marking. Traditional folk poetry, oral storytelling, and proverbs exploit reduplication to enhance rhythm, memorability, and cognitive salience, e.g., şirin-şirin, çox-çox, yavaş-yavaş, göz-göz, çırp-çırp (sweetly, very much, slowly, eye for eye, flutter by flutter). Cognitive research supports that repeated phonological elements reinforce mental representation, facilitating processing and memory retention in agglutinative systems (Nasirova *et al.*, 2023).

In digital communication, reduplication adapts to stylistic conventions, e.g., tez-tez yenilənən, paylaşıla-paylaşıla, sürətli-sürətli (frequently updated, shared over and over, quickly-quickly), often combined with interfixes and conversion, producing complex hybrid compounds like tez-tez-yenilənən-onlayn-məlumat (frequently updated online information) as stated by O.B. Jalilbayli (2022). These findings confirm A. Mirzayeva (2021) and G.Y. Rakhimova (2023), showing that Azerbaijani reduplication preserves traditional cognitive and expressive functions while adapting to technological contexts. Typologically, Azerbaijani reduplication shares features with Japanese, Indonesian, and Malay, where it signals plurality, iteration, or emphasis, whereas English employs it primarily for onomatopoeia or colloquial intensifiers, illustrating Azerbaijani's exceptional expressive flexibility.

Sociolinguistic variation indicates younger, digitally literate speakers innovate reduplications for humor and style, whereas older or rural speakers preserve canonical patterns.

Apophony, or internal stem alternation, allows roots to signal tense, aspect, causativity, plurality, or derivation, e.g., al → aldı, yaz → yazdır, otur → oturt, gözlə → gözlət (buy → bought, write → wrote, sit → sat, wait → waited) (Rakhimova, 2023). This non-concatenative process differs from English (sing → sang) or German (sprechen → sprichst), allowing semantic nuance without external affixation. Historically, apophony was influenced by Persian and Arabic and appears in literary and administrative texts, e.g., müdür → müdürləşdir, təhsil → təhsilləndir (director → direct, education → educate). In modern contexts, apophony continues in formal and digital registers, adapting to technological verbs (oxu → oxut, yüklə → yüklət; read → read, load → load), demonstrating morphological resilience and continuity. Cognitively, internal alternation reduces processing load, supporting faster comprehension and production. Typologically, Azerbaijani uniquely combines agglutinative suffixation and non-concatenative alternation, contrasting Turkish (suffix-only) and English (analytic).

Conversion enables lexical category shifts without overt affixation, e.g., oyun → oyun oynamaq (game → to play a game), film → film çəkmək (to make a film), zoom → zoom etmək (to zoom). Conversion often interacts with affixation, interfixation, and reduplication, creating morphologically complex, semantically transparent forms like tez-tez-yenilənən-onlayn-məlumat (frequently updated online information). Historically, zero-derivation maximised lexical productivity in classical literature (dərs → dərs etmək, məktub → məktub yazmaq; lesson → to give a lesson, letter → to write a letter). Cognitive efficiency is enhanced as speakers generate novel verbs without memorising additional affixed forms. Sociolinguistically, younger urban speakers innovate with playful conversions (selfi-çək-çək; take a selfie), whereas older speakers maintain classical forms, illustrating register- and context-sensitive morphological adaptation. The digital revolution accelerates hybrid morphological processes. Affixation, interfixation, reduplication, apophony, and conversion now operate in digital, English, and Turkish borrowing contexts, e.g., onlayn-dərs (online lesson), virtual-muzey (virtual museum), bio-i-təhlükəsizlik (bio-security), selfi-çubuğu (selfie-stick), video-görüş (video call), chat-başlamaq (start a chat). Hybrid compounds carry phonological and semantic cues, supporting rapid parsing and comprehension in digital communication. Historical continuity is evident: interfixes and reduplication transfer classical patterns into modern, technologically mediated registers.

Typologically, Azerbaijani shares certain linguistic features with Turkish, Russian, English, Japanese, and Persian, while still retaining distinct cognitive and expressive advantages. For instance, in comparison with Turkish, Azerbaijani exhibits similar agglutinative structures and vowel harmony, but it uses fewer interfixes and internal alternations. Compared to Russian, which is heavily fusional, Azerbaijani preserves agglutinative roots with optional internal alternation, as in yaz → yazdır (write → print) versus the Russian писать (pisat; to write) → заставить писать (zastavit pisat; to make someone write). In relation to English, which is largely analytic and exhibits limited morphological productivity, Azerbaijani can compress complex semantic information morphologically. With respect to Japanese, which often employs reduplication for plurality or emphasis, Azerbaijani integrates similar semantic effects through the use of interfixes and stem alternations. Finally, in comparison with Persian, Azerbaijani has borrowed roots and interfixes that have been adapted to its agglutinative patterns, as seen in examples like elm-i-fən (science) and ədəb-i-nəzər (litetature). Sociolinguistically, urban digital users innovate,

whereas formal and classical registers preserve historical morphology. Cognitively, this hybrid system supports efficient parsing, recognition, and production, particularly in digital and multilingual environments (López-Couso, 2016).

The present study demonstrates that Azerbaijani retains its agglutinative core while generating productive hybrid derivations under contact-driven influence. This pattern aligns with recent findings in morphological typology, psycholinguistics, and contact linguistics, and situates Azerbaijani within broader global research on language adaptation. The present study demonstrates that Azerbaijani retains its agglutinative core while simultaneously expanding its derivational potential through contactdriven hybrid formations – a dynamic mirrored in recent crosslinguistic and typological research on Turkic and other languages. Recent semanticmorphological work by C. Ismoilova (2025) on Turkish has shown that in agglutinative languages, affixation is guided by semantic coherence rather than purely formal concatenation. This supports the find that many of the productive hybrid derivatives in Azerbaijani preserve semantic transparency, which helps them integrate smoothly into everyday usage rather than remain sporadic borrowings. In other words, semantic consistency appears to enable hybrid affixation to become part of the productive core, rather than marginal exceptions.

Further, studies on loan-verb adaptation in Russian reveal systematic patterns in suffix assignment: loan verbs often receive variant suffixes and coexist with native morphological patterns (Sokolova *et al*., 2025). This resonates with current observation that borrowed stems in Azerbaijani likewise take native affixes or hybrid stemaffix combinations systematically rather than haphazardly. The crosslinguistic similarity suggests that contactinduced morphological adaptation may follow universal tendencies shaped by phonological compatibility and morphological economy, regardless of the donor or recipient language's typology.

Notably, a study on prefix-like formation in Turkish by G. Booij (2018) – which documents emergent "prefixoids" derived from adjectival constituents (e.g., ana "main/primary," ön "pre") – argues that even languages without traditional prefixation may develop creative wordformation strategies to address lexical gaps. This parallels current finding of hybrid affixation in Azerbaijani: as contact and lexical influx introduce new stems, speakers exploit the flexibility of the agglutinative system to generate novel forms. In doing so, Azerbaijani expands its wordformation repertoire in a way reminiscent of emergent morphological innovation observed in related Turkic languages.

However, computationalmorphology research cautions about the consequences of such morphological diversity. A recent study on language modelling for morphologically rich languages by C. Toraman *et al*. (2022) showed that standard tokenisation and NLP tools often struggle with complex affixation and irregular morphology. This suggests that while hybrid derivations may enrich the lexicon and ease human processing, they pose challenges for automated processing, morphological tagging, and computational lexicography – a limitation that must be acknowledged when considering applications of this research (e.g., morphological analysers, spell-checkers). Moreover, historicalphilological investigation of ancient Turkic root morphemes revealed the long-standing productivity of affix-based derivation in the family, pointing to diachronic continuity in word-formation strategies, as proved by D. Beck (2017) in a recent study. This supports diachronic findings made in this work: hybrid and novel forms in Azerbaijani may not be aberrations but part of a deep-rooted morphological tradition, albeit one evolving under modern sociolinguistic pressures. From that perspective, the hybrid derivations documented in this study – especially those combining borrowed stems with native affixes – might blur the boundaries between derivation, compounding, and morphological construction. This

raises the question whether some of identified "words" in this work are better analysed as constructionlevel or compoundderived units, suggesting a fruitful direction for future research.

In light of these comparisons, the present study contributes to a growing body of evidence that morphological systems – even in agglutinative languages – are not static. They respond to sociolinguistic dynamics, contact, lexical influx, and speaker needs. For Azerbaijani, this means that hybrid derivation and loanaffix integration are not peripheral anomalies but active mechanisms enriching the lexicon and reflecting real language use. At the same time, current findings highlight a tension: the same processes that enhance expressive capacity and cognitive accessibility may complicate computational processing or formal standardsation. Thus, while supporting many of the generalisations found in contemporary studies – about agglutination, borrowing adaptation, semantic transparency, and cognitive processing – this research also pushes the boundary of typological and morphological theory by showing how language contact can dynamically reshape morphological productivity without erasing core structural traits. In sum, Azerbaijani emerges as a morphologically resilient, flexible, and adaptive system: rooted in agglutination, yet capable of accommodating hybrid innovation – a finding that bridges typological tradition with modern contact-induced evolution.

Conclusions

The present study demonstrated that Azerbaijani morphology is highly dynamic, adaptive, and cognitively optimised, integrating multiple word-formation mechanisms-including suffixation, interfixation, compounding, reduplication, apophony, and conversion-to generate a rich, expressive, and efficient lexicon. Corpus analysis of 2020-2024 data revealed that suffixation dominates derivational productivity, accounting for 82.4% of all tokens (n = 43,912 types), with the most productive suffixes being -lıq/-lik (17,504 tokens), -çı/-çi (9,612 tokens), -sız/-siz (7,921 tokens), and -lı/-li (5,332 tokens). Hybridisation with borrowed stems was also productive: forms such as follow-çu (follower) and blok-lamaq (blocker) generated multiple derivatives, achieving frequencies per million (fpm) of 78.3 and 112.6 respectively, whereas less integrated forms like like-ləmək (to like) remained non-productive (fpm = 32.4). Prefixation remains limited (11.3% of new derivations), with domain-specific prominence in technological, academic, and youth discourse (e.g., qeyri- 4,480 tokens; anti- 2,931 tokens). Interfixation and compounding enhance formal clarity and cognitive processing, while reduplication supports emphasis and memorability. Apophony allows precise derivational adjustments, and conversion facilitates flexibility and rapid lexical expansion. Historical layers of Persian, Arabic, Turkish, and Russian influence coexist with modern digital innovations, producing a morphological system that operates effectively across literary, formal, and online registers. Comparative analysis showed that Azerbaijani integrates features of agglutinative, concatenative, and non-concatenative languages while maintaining semantic transparency and processing efficiency, distinguishing it as a hybrid system capable of accommodating neologisms, hybrid compounds, and contact-induced borrowings. Future research should expand the integration of corpus linguistics, psycholinguistic experiments, and computational modelling. Large-scale spoken, written, and digital corpora can be used to track frequency, productivity, and register variation of derivational mechanisms. Cognitive experiments measuring reaction times and accuracy for parsing complex forms (e.g., tez-tez-yenilənən-onlayn-məlumat; frequently updated online information) can validate hypotheses about processing efficiency. Computational approaches, including NLP and AI-based morphological analysers, will benefit from comprehensive morphological databases to handle compounds, reduplications, and apophony. Dialectal

studies, longitudinal tracking, and interdisciplinary collaboration will further illuminate how classical morphology adapts to modern technological, educational, and multilingual demands. Overall, Azerbaijani morphology exemplifies a living system that remains resilient, adaptive, and socially functional while preserving clarity, expressiveness, and cognitive efficiency.

Acknowledgements

None.

Funding

None.

Conflict of Interest

None.

References

[1] Abdullayeva-Nebiyeva, V. (2025). Cultural determinants and grammatical outcomes in the Azerbaijani: An analysis of sociolinguistic influence on Turkic grammar. *Path of Science*, 11(4), 4028-4032. doi: 10.22178/pos.116-35.

[2] Beck, D. (2017). The typology of morphological processes: Form and function. In A.Y. Aikhenvald & R.M.W. Dixon (Eds.), *The Cambridge handbook of linguistic typology* (pp. 325-360). Cambridge: Cambridge University Press. doi: 10.1017/9781316135716.011.

[3] Booij, G. (2018). The construction of words. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics* (pp. 229-245). Cambridge: Cambridge University Press. doi: 10.1017/9781316339732.016.

[4] British National Corpus. (n.d.). Retrieved from https://www.english-corpora.org/bnc/.

[5] Durkin, P. (Ed.). (2015). *The Oxford handbook of lexicography*. Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199691630.001.0001.

[6] Humboldt-Universität zu Berlin, Institute of Linguistics. (n.d.). *Corpus linguistics: Corpora links*. Retrieved from https://www.linguistik.hu-berlin.de/en/institut-en/professuren-en/korpuslinguistik/links-en/korpora_links.

[7] Huseynova, M. (2019). Language contacts of Azerbaijani and Kazakh Turkic languages (on the basis of Azerbaijani dialectologist, academician Mammadaga Shiraliyev's creative works). *Turkic Studies Journal*, 1(2), 34-40. doi: 10.32523/tsj.02-2019/2-4.

[8] Ismoilova, C. (2025). The foundations of Turkish word formation: A semantic approach. *International Journal of Artificial Intelligence*, 1(5), 1581-1583.

[9] Jalilbayli, O.B. (2022). Philosophy of linguistic culture and new perspectives in modern Azerbaijani linguistics. *Futurity Philosophy*, 1(4), 53-65. doi: 10.57125/FP.2022.12.30.05.

[10] Kornfilt, J. (2002). The Turkic languages (Routledge language family descriptions). *Journal of Linguistics*, 38(2), 397-439. doi: 10.1017/S0022226702251622.

[11] LópezCouso, M.J. (2016). Corpora and online resources in English historical linguistics. In M. Kytö & P. Pahta (Eds.), *The Cambridge handbook of English historical linguistics* (pp. 127-145). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139600231.009.

[12] Milli Arxiv İdarəsi. (n.d.). *Publication and use of documents*. Retrieved from http://www.milliarxiv.gov.az/az/news/senedlerin-nesri-ve-istifadesi.

[13] Mirzayeva, A. (2021). A phytonymic phraseology and linguocultural features of paremiological units. *Langua: Journal of Linguistics, Literature, and Language Education*, 4(2), 39-44. doi: 10.5281/zenodo.5540285.

[14] Nasirova, U., Hasanova, S., & Dashdemirova, P. (2023). Phonetic rules in the Azerbaijani language. *ASES International Journal of Culture, Art and Literature*, 2(1), 33-39. doi: 10.5281/zenodo.10492377.

[15] National Corpus of Azerbaijani Language. (n.d.). Retrieved from https://korpus.azerbaycandili.az/.
[16] Özenç, B., Ehsani, R., & Solak, E. (2018). MorAz: An open-source morphological analyzer for Azerbaijani Turkish. In E. Blanco & W. Lu (Eds.), *Proceedings of the 2018 conference on empirical methods in natural language processing: System demonstrations* (pp. 25-29). Brussels: Association for Computational Linguistics. doi: 10.18653/v1/D18-2005.
[17] Plag, I. (2025). Productivity and the mental lexicon. In A.J. Merrison, P. Griffiths & A. Bloomer (Eds.), *Language in use* (pp. 106-115). London: Routledge.
[18] Rakhimova, G.Y. (2023). Functional semantics of English reduplicative elements. *International Journal of Applied Linguistics and Translation*, 9(3), 90-92. doi: 10.11648/j.ijalt.20230903.12.
[19] Russian National Corpus. (n.d.). Retrieved from https://ruscorpora.ru/en.
[20] Sokolova, S., Birzer, S., Ignateva, A., & Kibisova, E. (2025). Suffix diversity: Investigating the morphological landscape of Russian loan verbs. *Russian Linguistics*, 49, article number 17. doi: 10.1007/s11185-025-09324-8.
[21] Toraman, C., Yılmaz, D., & Şahinuç, A. (2022). Evaluating morphological compositional generalization in large language models for agglutinative languages. *ArXiv*. doi: 10.48550/arXiv.2410.12656.
[22] Tsanko, I. (2025). Corpora and the study of language evolution: A historical linguistics perspective. *SWorldJournal*, 4(30-04), 122-129. doi: 10.30888/2663-5712.2025-30-04-041.
[23] Turkish National Corpus. (n.d.). Retrieved from https://www.tnc.org.tr/.

Механізми словотворення в азербайджанській та світових мовах: порівняльне морфологічне дослідження

Раміла Фараджова
Доктор філологічних наук
Азербайджанська Державна Морська Академія
AZ1000, вул. Заріфи Алієвої, 18, м. Баку, Азербайджан
https://orcid.org/0009-0000-2006-1706

Анотація. У цій дослідницькій роботі представлено розширене та глибоке порівняльне дослідження механізмів словотворення в азербайджанській мові у зв'язку з кількома основними мовами світу, зокрема англійською, російською, німецькою, турецькою та перською. Розміщуючи азербайджанську мову в ширших міжмовних рамках, дослідження мало на меті з'ясувати як типологічні особливості, так і універсальні закономірності в морфологічних процесах. Методологія інтегрувала когнітивно-морфологічний аналіз, корпусне дослідження, обчислювальне та цифрове моделювання ресурсів, а також міжмовне типологічне порівняння для дослідження морфологічних структур азербайджанської мови. Дослідження показало, що азербайджанська мова зберегла свою основну аглютинативну структуру, розвиваючи гібридні утворення через інтеграцію запозичень та афіксів, демонструючи збільшення частоти змішаних морфологічних ланцюгів у цифрових корпусах та розширення продуктивних афіксаційних моделей у відповідь на лексичний приплив, зумовлений контактами. Емпіричний аналіз показав, що азербайджанська морфологія була одночасно гнучкою та стійкою, здатною генерувати нові лексичні одиниці та адаптуватися до семантичних зрушень у відповідь на соціальний, технологічний та міжкультурний розвиток. Ці висновки підкреслили подвійний характер морфологічної еволюції: вона виявилася універсальною у своїх структурних тенденціях, водночас унікально формуючись культурним та лінгвістичним контекстом азербайджанських носіїв. Розміщуючи азербайджанську морфологію в порівняльному ландшафті світових мов, це дослідження сприяло глибшому розумінню міжмовної креативності, типологічної варіації та взаємодії між морфологією та соціолінгвістичною динамікою, пропонуючи розуміння, актуальне для теоретичної лінгвістики, викладання мов та прикладної лексикографії. Практична цінність цього дослідження полягає в тому, що воно надає лінгвістам, лексикографам, викладачам та розробникам цифрових мовних технологій емпірично обґрунтовані моделі азербайджанського словотворення, які можна безпосередньо застосовувати для складання словників, розробки навчальних програм, автоматизованої обробки морфології та розробки інструментів вивчення природного мови, таких як морфологічні аналізатори та засоби перевірки орфографії

Ключові слова: аглютинативна структура; типологічний контраст; лексичні інновації; контактно-індуковані зміни; афіксальна продуктивність; цифрові корпуси; міжмовна креативність

International Journal of Philology, 29(3), 2025
Journal homepage: https://philologicalscience.com.ua/en
Received 15.03.2025 Revised 02.09.2025 Accepted 23.09.2025

UDC 81.27:316.35 DOI: 10.31548/philolog/3.2025.61

Features of language functioning and communication in social networks

Svitlana Heyko*
PhD in Philosophy, Associate Professor
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
https://orcid.org/0009-0008-0119-7673
Olena Lauta
PhD in Philosophy, Associate Professor
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
https://orcid.org/0000-0002-6118-0258

Abstract. The article analysed the features of communication in social networks, as well as the essential characteristics and existing classifications of hate speech and support speech. The need for a deeper understanding of these phenomena is due to their impact on society, communication norms and the need to develop effective mechanisms for countering aggression in the network. The purpose of the study was to identify, analyse and systematise communicative strategies for expressing hate speech and support in social networks. The methodology included content analysis of texts, structural and semantic modelling and sociolinguistic observation. The results of the study showed that the language of social networks is a hybrid, which is characterised by informality and brevity such as simple sentences, agrammatisms, abbreviations; multimodality such as the use of emojis, memes, images; hypertextuality and compression such as use of hashtags, links; impact of technical limitations: character limits stimulate creativity; adaptation to censorship: modification of words to bypass moderation algorithms. Hate speech is implemented through strategies of sarcasm, manipulation, the opposition "friend-stranger" and the use of modified vocabulary to bypass censorship. Supportive speech is mainly expressed through strategies of emotional support, informational assistance and the use of verbal and non-verbal markers (likes, retweets, hashtags). The interaction of polar discourses of hostility and support in the digital environment reflects the adaptation of communicative practices to the conditions of online communication and moderation algorithms. The study found that these

Suggested Citation:
Heyko, S., & Lauta, O. (2025). Features of language functioning and communication in social networks. *International Journal of Philology*, 29(3), 61-73. doi: 10.31548/philolog/3.2025.61.

*Corresponding author

discourses often exist in parallel spaces (for example, in replies to the same tweet). A common strategy for both is the use of the opposition "friend-stranger". Supportive speech often arises as a reaction to hate speech, forming a mechanism of collective resistance

Keywords: hate speech; supportive speech; Internet; online interaction; networking platforms; discourse strategies

Introduction

The rapid development of digital technologies and the global popularity of social networks have led to the emergence of new forms of communication that are significantly transforming traditional language practices. Platforms such as Facebook, Instagram, X (Twitter), TikTok and Telegram form a specific communicative space with their own norms, genres and linguistic innovations. In these environments, speech takes on a multimodal character, combining verbal and non-verbal means that influence the structure and functioning of linguistic discourse. Social networks are an important tool for shaping public opinion, but at the same time they are an environment for the spread of hate speech. Alongside it, supportive language is developing as an alternative communicative practice focused on solidarity and mutual assistance. The relevance of the study is due to the fact that social networks not only reflect linguistic changes, but also actively initiate them, setting linguistics the task of understanding the mechanisms of digital language formation and its interaction with traditional forms of speech. The study of social media language is an interdisciplinary field that involves linguists, philosophers, and digital technology analysts.

British linguist D. Crystal (2021) conducted a thorough analysis of the impact of digital technologies on language and became one of the founders of the field of internet linguistics. In his works, the scientist outlined the main prospects for the development of this field of research (sociolinguistic, educational, stylistic and applied), defining their significance for understanding language processes in the digital environment, which laid the foundation for the further development of Internet linguistics as a separate field. Internet linguist G. McCulloch (2020) analysed the evolutionary changes characteristic of modern Internet communication, in particular how the digital environment affects the structure and functioning of the English language. Based on the conducted analysis, the researcher concluded that the Internet as a global space for communication generates new language practices and changes the mechanisms of meaning transmission. In the context of studying language processes in the digital environment, the article by D. Strelēvica-Ošiņa (2019) is particularly noteworthy, in which the researcher analysed the terminology used to describe linguistic correctness, emphasising the Latin origin of a significant part of the key concepts. This approach made it possible to trace the historical factors that shaped normative ideas that influence the interpretation of linguistic usage, particularly in the context of social networks.

A new direction in the study of language practices on social networks has been the study of the phenomenon of “brainrot language”. This term is used to refer to internet content with low informational value or potentially negative psychological effects. The concept also covers the phenomenon of excessive consumption of digital media, primarily short-form videos and texts, which can have adverse effects on cognitive functioning. This phenomenon has been studied by K. Tiffany (2021). Research into the lexical-semantic aspects of language functioning is considered promising. Ukrainian

researchers L. Tereshchenko & L. Stanislavova (2024) concluded in their work that lexical-semantic categories constitute one of the most complex levels of the language system, as they combine the lexical fund of the language and the structure of meanings. The authors analysed in detail the lexical-semantic fields covering sets of lexemes united by common meaning; lexical-semantic groups formed on the basis of a common generalised seme; and the sememe as a unit reflecting the separate meaning of a polysemous word. This approach made it possible to systematically describe the patterns of lexical meaning organisation and the mechanisms of its variation within the linguistic structure of the Internet community.

A comprehensive comparative analysis of the strategic parameters of communicative interaction between Ukrainian-speaking and English-speaking social media users was conducted by L. Shvelidze (2021). In the scientific works, the researcher outlined the mechanisms of communication conflict and cooperation in network discourse, developed a typology of communication strategies, and defined the essence of communication tactics in Ukrainian and English. An important achievement in the research was the identification of specific features of the speech behaviour of participants in the Ukrainian- and English-speaking segments of social networks and the delineation of typological trends in the functioning of communication in the digital environment. The peculiarities of the Ukrainian and French languages in the context of network communication remain relatively understudied, which highlights the need for further interlinguistic and intercultural studies.

The psychological aspect of supportive language has been studied by T.I. Buialska (2016) and A.P. Romanchenko (2024), while its linguistic aspects have not been sufficiently researched. In her article, T.M. Geiko (2025) examined the influence of social networks on the modern French language as a special communicative environment. The author analysed how vocabulary and word meanings change in digital communication, how neologisms, borrowings and new language models appear, and how specific discursive practices of online interaction are formed. The conclusions emphasise that social media are significantly transforming the communicative system of the French language, creating dynamic, flexible and context-dependent forms of speech that differ from traditional norms. Despite a significant number of studies devoted to individual aspects of Internet communication, a comprehensive analysis of the functioning of language in the context of social networks remains relevant. It is necessary to investigate in more detail how the technological features of platforms influence the ways in which messages are constructed, how users adapt linguistic means to the requirements of the digital environment, and what pragmatic, social and cultural functions language performs in online interaction.

The aim of the study was to conduct a comprehensive analysis of language of support as a linguistic phenomenon and to investigate its interaction with language of hostility in the context of French-language digital communication. Objectives: to characterise the features of Internet communication and social media language; to refine the definitions and classifications of the concepts of hostility and support; to analyse the contexts of their use and interaction. The achievement of the stated aim and objectives made it possible not only to gain a deeper understanding of the patterns of digital communication, but also to outline prospects for further linguistic and interdisciplinary research in this field.

Materials and Methods

The methodological basis of the study was a set of general scientific and special linguistic methods, the application of which ensured the systematic and multifaceted nature of the analysis. In particular, general scientific methods of analysis and synthesis were used to process

theoretical sources, generalise scientific approaches to the study of network discourse, and systematise factual material. Descriptive analysis was aimed at systematically describing linguistic and communicative phenomena, which made it possible to focus on the process of linguistic interaction in social networks within the French-speaking communicative space, identify and classify the linguistic mechanisms of forming new units in social networks, describe linguistic innovations that arise as a result of interaction between standard language and digital slang, and systematise them according to certain criteria.

The main tool of analysis was content analysis, which made it possible to systematise empirical material, identify thematic groups of manifestations of hate speech (in particular, racist, sexist and other forms of discriminatory discourse) and supportive speech, as well as track their frequency and functional orientation. At the same time, linguistic and stylistic analysis was used to identify the linguistic means and rhetorical devices used to implement the relevant communication strategies. Given the multimodal nature of online communication, the study also relied on an analysis of the interaction between verbal and non-verbal components of messages, in particular text with emojis, memes, GIF animations, stickers and video clips, which made it possible to take into account the role of technological markers in the formation of meaning. The interpretation of supportive language was carried out taking into account a psychological approach, which involves the analysis of stress coping resources and forms of structural and functional support, as well as a functional approach, which allows to distinguish between emotional, informational and network support. The combination of linguistic, psychological, and multimodal methods provided a comprehensive understanding of communicative practices in social networks and their discursive and pragmatic characteristics.

Observation of the language practices of social network respondents showed that the language of social networks is a kind of linguistic hybrid, combining elements of standard language, slang, argot, neologisms, graphic abbreviations, emojis, etc. The study of this phenomenon made it possible to trace how the lexical system of language, in particular French, functions in new conditions and how it adapts to the challenges of digital communication. Quantitative analysis was used to establish the frequency of individual linguistic phenomena, as well as to determine the dynamics of their use in the Ukrainian- and French-speaking segments of social networks. Special linguistic methods play an important role in the study. Comparative analysis made it possible to study the communicative behaviour of Ukrainian- and French-speaking social network users, which allowed to identify common and distinctive features of linguistic representation of social interaction. Linguistic and stylistic analysis was used to determine the stylistic and emotional-expressive load of lexical units used in online discourse, as well as to clarify the role of evaluative and expressive means in shaping the communicative effect.

The typological method was used to identify typological differences between the linguistic means of the Ukrainian and French languages in the digital environment. Contextual-interpretative analysis made it possible to study the communicative intentions of participants in online interaction, their communicative roles and the peculiarities of meaning realisation in specific communication situations. Discursive analysis helped to identify the specifics of the realisation of speech genres and speech acts in the process of Internet communication, taking into account sociocultural and pragmatic factors. Linguistic-pragmatic analysis was used to develop a typology of communicative strategies and tactics of communicants in Internet communication, and component analysis was used to determine the semantic composition of

lexical units used in network discourse. The combined application of these methods provided a comprehensive approach to the study of the linguistic and communicative features of modern online discourse and made it possible to obtain well-founded scientific conclusion.

Results and Discussion

Digital technologies have become an integral part of everyday life around the world, reflecting a high level of digital development and the deep integration of online platforms into social communication. Social networks, along with other electronic media such as news sites, blogs, mobile applications and online forums, occupy an important place in the information environment of the modern world. They have not only changed the way news is received, but have also become an effective communication tool. Government officials and members of society actively use digital platforms to interact with citizens, disseminate messages, and so on. Digital media has also become a key resource for civil society activists, particularly volunteers, who use social networks to promote their ideas and attract new supporters to socio-political activities.

Large-scale digitalisation has a significant impact on linguistic culture. Social networks contribute to the rapid spread of new words, technical vocabulary and expressions, often of English origin. Through active use in online communication, these linguistic innovations quickly enter everyday use. Entertainment media products such as reality shows, TV series and popular bloggers also play an important role in changing the linguistic environment. Charismatic presenters and fictional characters become linguistic role models, especially for young audiences. Their expressions, style of speech and communication characteristics are often imitated by viewers and gradually become established in everyday language. The internet has had a radical impact on people's everyday lives, and one of the most noticeable areas of this impact has been the change in communication methods. Social networks have made it possible to maintain social contacts regardless of distance, create personal networks, and shape public image. Online communication has become a factor that has a significant and lasting impact on both language and social processes. The term "social media language" is used to refer to the specific vocabulary, grammatical structures and writing style that have emerged in the digital environment. This type of speech is characterised by informality, active use of slang and a tendency towards conciseness: many words and expressions are shortened or take on new forms. This language system is constantly changing in response to new trends and cultural processes shaped by the values of the younger generation.

To a large extent, the language of social platforms has become a kind of marker of identity for modern youth. It allows to convey their thoughts, emotions, and beliefs in a way that is natural and understandable to their peers. Moreover, this format of communication helps to strengthen the sense of community and belonging, primarily through hashtags, memes and various elements of internet slang, which form a common cultural and linguistic space for users from different regions. Social networks are forming their own unique style of communication. These are short and witty messages, often accompanied by irony, abbreviations and hashtags. It can be visual material around which interests and communities are formed. Video sequences with sound effects are common – popular expressions, music fragments or excerpts from speeches.

The style of digital language is largely determined by the technical limitations of social platforms. For example, Twitter still imposes a 140-character limit per post, which encourages users to shorten their thoughts as much as possible, often sacrificing grammatical or stylistic correctness. As a result, abbreviations, symbolic substitutions, and various shortcuts have

become commonplace not only for ordinary users but also for media representatives. The language of social networks is also significantly influenced by content censorship, especially when it comes to topics considered sensitive, such as violence, suicide, sexuality, or politics. To avoid automatic blocking, users deliberately distort the spelling of words, replacing individual letters with symbols that are similar in form or sound. Such practices have become not only a way to circumvent censorship, but also part of a new visual and linguistic aesthetic that is actively forming in the digital environment.

Internet communication, particularly on social media, has created a unique communicative environment with its own rules and norms. The language of social media (computer-mediated communication) is considered a separate sociolinguistic phenomenon characterised by a number of specific features (Geiko, 2025). At the syntactic level, simple short sentences, frequently used incomplete and one-clause sentences are most common. Such sentences do not overload the text and keep the recipients' attention. The phenomenon of agrammatism is also common. The text producer deliberately deviates from the norms of syntax and punctuation. Punctuation marks may be absent, incorrectly placed, or deliberately overused. In addition, the concept of language varieties is blurred in the social media environment. In addition to the usual styles and registers, the choice of font, writing system and punctuation also play a meaningful role. For example, there is a so-called Greeklish on the Internet, which is Greek written in the Latin alphabet.

The language used in online communication is close to colloquial speech, as it has characteristics such as emotionality, expressiveness, informality, and evaluative reactions. Most often, such language is not formal in nature. Based on the example of the French language, the main aspects of the influence of social networks on the language system will be considered. Firstly, there is a noticeable emergence of hashtags for keywords or phrases preceded by the symbol '#' (octothorpe), used to structure text messages by topic or type. They create a parallel language where a common lexical unit can unite thousands of like-minded people. For example, cult hashtags that have been lexicographically recorded on the Internet include #ThrowbackThursday, #FoodPorn, and #Selfie.

Secondly, there is a significant spread of various abbreviations, primarily acronyms and abbreviations, which indicates the activity of the word-formation subsystem of the French language. Certain restrictions on the use of symbols in social networks contribute to the linguistic creativity of speakers-users, thanks to which a multitude of acronyms and abbreviations have appeared, enriching the vocabulary of everyday communication, such as lol (laughing out loud), mdr (mort de rire), OMG (Oh my God). Thirdly, an important aspect is the emergence of ideograms – emojis. These iconographic symbols of emotions and objects/items have become a universal means of communication, making them more functional than simple linguistic signs and essential elements of online conversations, as they convey certain emotional concepts more concisely than simple lexemes.

Next, it is worth mentioning the surge in neologisms caused by the active creativity of users. Many original words and expressions appear on social media, mostly occasionalisms, which are transient due to the specificity of their denotation. Such units are on the periphery of the lexical-semantic system of the French language. For example, during the World Cup, the term Mbappémania appeared to describe the enthusiasm for the play of young footballer Kylian Mbappé. Others enter the common vocabulary, such as "tweeter" and "liker," which refer to common practices on social media. Finally, there are new linguistic challenges related to communication through interactive game formats, where users have to follow certain rules. For example, lipograms are texts written without using one or more

letters of the alphabet. In this way, the potential of communication is developed and diversified, particularly in its linguistic and conceptual dimensions (Geiko, 2025). Thus, social media platforms create a favourable environment for borrowing, while activating native lexemes and word formation tools.

Other features of social media language are also emerging as alternatives. First and foremost is hypertextuality: in a network environment, linear text can be transformed into non-linear hypertext. In addition, online communication is characterised by a compressive language-creating process, where alphanumeric symbols are formed according to the principle of homophony: ok1- aucun, pl1- plein. Furthermore, there is a phenomenon of reterminology, i.e. the use of an existing term in a new meaning or in a different field. For example, the word virus began to be used to refer to an "internet virus" (Holubovska, 2015). Among other features of the language of social media, informality and brevity can also be highlighted. Communications are characterised by simple syntactic constructions, deliberate deviations from grammatical and punctuation norms (agrammaticalisms), as well as the active use of abbreviations, acronyms and initialisms, which is due to the desire for efficiency and often technical limitations of platforms (Antonyuk & Hoza, 2023).

Multimodality: The text component is combined with visual (emojis, memes, images, GIFs) and audio information. Emojis and emoticons perform a compensatory function, replacing non-verbal means of communication (facial expressions, gestures, intonation). Adaptation to censorship: To circumvent moderation algorithms, users deliberately distort the spelling of banned words by replacing letters with similar symbols (e.g., "n@zi"), which has become part of the visual and linguistic aesthetics of the digital environment (Määttä, 2023). Each social platform forms its own communicative style. Network X (formerly Twitter) is characterised by conciseness, a high concentration of irony and sarcasm, and the use of hashtags to participate in public discussions (Adawiyah, 2023).

Communication on the intern*et al*so has a number of features. It can be carried out using written text, images, video and audio files, as well as through sharing and liking. Despite this number of means of communication, the most common is the use of written text. The most common types of communication on social media include posts, stories, comments, reviews, and private chats (Antonyuk & Hoza, 2023). Since social networks allow for the rapid dissemination of content and responses to it, online messages, prioritising immediacy, may rely less on verified sources and neglect the principle of objectivity (Mabillard *et al*., 2024). It is also worth noting that due to the physical absence of a communication partner, the communicator can express their thoughts and feelings more freely (Udovichenko & Samoilenko, 2021).

A characteristic feature of communication on social networks is the expression of support and hostility. Hate speech is a complex and multifaceted phenomenon, for which there is no single definition in the scientific community. Within the scope of this study, hate speech is defined as derogatory, discriminatory statements based on social prejudices and directed at a person or group of persons because of their membership in a particular social, ethnic, religious, gender or other category. There are a number of approaches to classifying hate speech. By thematic focus: Spanish researchers distinguish between racist, sexist, ideological and violent hate speech (Brändle *et al*., 2024). By explicitness: the German school proposes a dichotomy between explicit (direct insults, threats) and implicit (negative stereotypes, misinformation, minimisation of historical facts). By social affiliation of the victim: manifestations of xenophobia, anti-Semitism, Romaphobia, Afrophobia, etc.

Discursive features of hate speech encompass a variety of lexical, pragmatic and stylistic features. These include the use of sexist

expressions directed against both women and men, the use of slurs and vulgarisms with derogatory connotations. Another characteristic feature is the use of controversial symbols, nicknames and hashtags, particularly those referring to groups involved in hate crimes. Such messages are often manipulative or deliberately distort facts, and may contain sarcasm, irony, ambiguous wording, personal attacks, biased attacks and accusations of lying, deception or ignorance (Papcunová *et al.*, 2023). An important discursive feature of hate speech is its clear focus on an object. Hate speech is a verbalised form of primary, undefined emotions (domestic affects) that do not have a specific object in themselves. When these emotions are expressed in speech, they become directed, i.e. they turn into hostility towards someone or something. Another characteristic feature is the use of so-called "magic words" or discursive "fetishes": repeated terms or expressions that, through excessive use, lose their original meaning and context, turning into emotional triggers and a kind of "advertising slogans" that function almost automatically (Voirol & Martini, 2023). Some researchers analyse direct hate speech as a manifestation of exclusion, denial and annihilation – processes aimed at depriving a particular person or group of their right to exist. Indirect hate speech, on the other hand, can take the form of denying or minimising hateful actions or historical facts, i.e. downplaying their significance (Moïse & Hugonnier, 2019). Since posts with overtly negative vocabulary are often removed by moderation algorithms, hate speech on social media rarely uses openly negative epithets. A much more common practice is to modify keywords to bypass algorithms, for example, replacing the letter "a" with "@" (Määttä, 2023).

The compositional features of the discursive characteristics of hate speech also include the actualisation of the binary opposition "in-group-out-group", which contributes to the polarisation of society. A clear antonymic position is created between "us", the members of a certain group, and "them", the opposing group, whose dissent causes dissatisfaction. Stylistic features lead to the use of so-called "magic words" or fetishes, which, through repetition, lose their original meaning and become emotional triggers, as well as the modification of vocabulary to circumvent moderation (Voirol & Martini, 2023). Thus, the discursive features of hate speech are: the use of pejorative, sexist vocabulary and vulgarisms; manipulation, sarcasm, irony, distortion of facts; actualisation of the "in-group-out-group" opposition; use of "magic words" and modification of vocabulary.

Supportive language is conceptualised in this study as multidimensional communication aimed at providing assistance, emotional encouragement, and expressing approval or solidarity with the views or actions of others. Contemporary scientific literature distinguishes three leading approaches to its analysis. The psychological approach considers social support as one of the key resources for overcoming stress, proposing a distinction between structural (quantitative characteristics of social contacts) and functional (qualitative parameters of interaction) support. The communicative approach focuses on speech strategies and tactics. Within this approach, a distinction is made between internal support (verbalisation of one's own beliefs) and external support (interaction with other communicators) (Buialska, 2016). Support tactics include support itself, praise, consolidation, gratitude, and other types of supportive speech (Romanchenko, 2024).

The functional approach classifies support according to its social functions. Emotional support includes expressions of comfort, empathy and encouragement. Informational support involves providing advice, instructions and relevant information. Self-esteem support is achieved through affirmation of a person's values and competencies. Network support provides a sense of belonging to a community. Practical and tangible support manifests itself in specific proposals for solving problematic

situations. It is worth noting that online communication is dominated by informational support, which is most often represented in the form of advice; emotional support is the second most common type (Millette & Boislard, 2023). Forms of expressing support on social networks are diverse. They are verbal, expressing direct words of support, appeals, comments. They are also non-verbal/technological, expressed through likes, retweets (with or without comments), the use of supportive hashtags, changing avatars or profile names to show solidarity (Aubanelle, 2022). The characteristics of supportive discourse are its dialogical nature, cooperative model of interaction, contextuality, meaningfulness, and interdiscursivity – the ability to integrate into other types of discourse (Buialska, 2016). One way to show support is by sharing posts and tweets. Reposting (retweeting) can take two forms: without comment or with an added opinion, which creates an additional metadiscursive level.

Other means of expressing support include the choice of nickname and profile photo. On social network X, it is important to distinguish between the username, which begins with @, and the account name. Most often, it is the account name that is changed to show support. This is for pragmatic reasons: the account name is more visible due to its location and bold font, and it also provides broader discursive possibilities – up to 50 characters compared to 15 for the username. The profile photo, which embodies the user's digital identity, can also be easily changed according to communicative intentions, in particular to express solidarity (Aubanelle, 2022). Studies have shown that users who post their own face on their avatar and use their real name receive more polite and personalised responses because their profiles are perceived as more socially "present" (Feng *et al.*, 2013).

Hashtags can also be a means of supporting a particular person or phenomenon. They perform a classificatory function, but at the same time can carry interpretative or metadiscursive meaning (Aubanelle, 2022). Research has found that these discourses often exist in parallel spaces (e.g., in responses to the same tweet). A common strategy for both is the use of the "in-group-out-group" opposition. Supportive language often arises as a reaction to hate speech, forming a mechanism of collective resistance. The results of the study show that social media language is shaped by specific communicative conditions that distinguish it from traditional forms of writing and oral interaction. One of the key features is the high degree of dynamism of linguistic means: new lexical units, memes, hashtags and abbreviations appear extremely quickly and spread thanks to the network effect. The results are consistent with the study by L. Shvelidze (2021), which points to the speed of linguistic transformations. This confirmed the assumption that digital discourse is a self-regulating, adaptive system that responds to changes in the social context almost instantly.

An important trend is the growing role of multimodal means of communication. The use of emojis, GIF animations, stickers, reactions, and video clips serves not only a decorative but also a meaningful function, complementing or even replacing the verbal component. The results confirm the conclusions of V. Sikorska *et al*. (2024). Participants in communicative interaction actively use these means to convey emotional nuances, irony, or pragmatic intentions, which in traditional written speech would require detailed descriptions. The data obtained demonstrate that emojis and other non-verbal elements act as markers of interpersonal closeness and social engagement. The phenomenon of fragmentation and non-linearity of communication attracts particular attention. O. Kudelska (2023). The presence of comments, threads, reposts, and instant reactions creates a polycentric discourse in which messages can be reinterpreted or broken away from their original context. This leads to a change in the logic of constructing statements: users often focus not on a consistent presentation of their thoughts, but on

maximum clarity within a short fragment. Thus, a communicative strategy of "minimal semantic blocks" is formed, characteristic of platforms where audience attention is a limited resource.

A trend towards increased personalised communication has also been identified (H. Matsiuk, 2021). Users are increasingly forming individual language styles, which are manifested in the choice of specific language markers, the use of emojis and special ways of formatting text. This practice contributes to the construction of digital identity, but can complicate mutual understanding between representatives of different online communities with their own norms and traditions of communication. In addition, a correlation has been found between linguistic behaviour and the linguistic rights of ethnolinguistic communities, as well as issues of linguistic identity, linguistic-national relations and language planning O. Mykhalchuk (2024).

Despite the convincing results, the study has certain limitations. First, the analysis covered mainly the most popular social networks, while niche platforms with specific audiences were left out. Second, the work was based on the analysis of public texts, while private messages may demonstrate different patterns of linguistic behaviour. Future research may focus on comparing linguistic strategies in different digital environments and identifying universal mechanisms of online communication formation. Overall, the results suggest that social media language is a unique communicative phenomenon that combines features of traditional writing, oral interaction, and multimodal digital content. Its development reflects social changes, technological innovations, and new forms of interaction between people, making this area promising for further linguistic and interdisciplinary research.

Conclusions

The study found that communication on social networks is a complex interdisciplinary phenomenon that combines linguistic, philosophical, and socio-analytical approaches. The language of social networks is shaped by rapid digitalisation, which determines its hybrid, dynamic, and rapidly changing nature. It is characterised by informality, multimodality, hypertextuality, content compression, as well as dependence on technical limitations and moderation algorithms that stimulate linguistic creativity and adaptation. Social networks are a space where polar communication strategies, in particular hate speech and supportive speech, function simultaneously. Hate speech is defined as a set of derogatory and discriminatory statements directed at individuals or groups, expressed through sarcasm, irony, negative framing, wordplay, and manipulative language techniques, often adapted to circumvent censorship. Supportive language is interpreted as a multidimensional communicative practice aimed at solidarity, providing emotional and informational support, expressing approval, and consolidating the community. It is implemented through verbal and non-verbal means, including positive connotation vocabulary, emojis, hashtags, and symbolic changes in user accounts. The most common strategy was the transmission of emotions, while the strategy of supporting self-esteem was less represented. Common to both types of discourse is the strategy of constructing the "in-group-out-group" opposition, which indicates a high level of polarisation in digital communication. The results confirmed the ability of users to adapt language strategies to the conditions of online interaction and demonstrate the importance of analysing supportive and hostile language for a deeper understanding of the mechanisms of modern Internet communication, as well as for further research in the field of digital linguistics, content moderation, and media literacy. Prospects for further research include a comparative study of the implementation of hate speech and supportive language in different types of online

communities, as well as research into the impact of moderation algorithms and artificial intelligence on the formation, transformation and masking of communication practices. Another promising area is the analysis of multimodal means (emojis, memes, visual symbols) in combination with verbal elements as tools of emotional influence and polarisation.

Acknowledgements

None.

Funding

None.

Conflict of Interest

None.

References

[1] Adawiyah, R. (2023). *Mini research: Language and social media*. Retrieved from https://www.academia.edu/106797085/LANGUAGE_AND_SOCIAL_MEDIA.

[2] Antonyuk, H.D., & Hoza, V.S. (2023). Internet speech and social media: An analysis of the use of linguistic means in users' comments on forums and social networks. *Scientific Notes of Ostroh Academy National University, Philology Series*, 19(87), 79-82. doi: 10.25264/2519-2558-2023-19(87)-79-82.

[3] Aubanelle, C. (2023). *Markers of alignment and opposition in political discourse on Twitter*. (Doctoral dissertation, Sorbonne Nouvelle University, Paris, France).

[4] Brändle, G., Cáceres-Zapatero, M.-D., & Paz-Rebollo, M.-A. (2024). Feeling hatred: An analysis of the perceived severity of hate speech in the Spanish population. *Spanish Journal of Sociology (RES)*, 33(2), article number a219. doi: 10.22325/fes/res.2024.219.

[5] Buialska, T.I. (2016). The interdiscursive potential of expressing communicative support in contemporary English. *Scientific Notes of Ostroh Academy National University, Philology Series*, 62, 52-55.

[6] Crystal, D. (2021). *Internet linguistics*. London: Routledge. doi: 10.4324/9780203830901.

[7] Feng, B., Li, S., & Li, N. (2013). Is a profile worth a thousand words? How online support-seekers' profile features may influence the quality of received support messages. *Communication Research*, 43(2). doi: 10.1177/0093650213510942.

[8] Geiko, T.M. (2025). Social media in the communicative dimension of the French language: Lexical-semantic and discursive aspects. *Bulletin of Science and Education*, 4(34), 357-370. doi: 10.52058/2786-6165-2025-4(34)-357-370.

[9] Holubovska, I.O. (2015). The language of social networks as a novel linguistic phenomenon. *Linguistic and Conceptual Worldviews*, 5, 49-56.

[10] Kudelska, O. (2023). The effectiveness of social media in teaching English. *Contemporary Studies in Foreign Philology*, 1(23), 264-272. doi: 10.32782/2617-3921.2023.23.264-271.

[11] Mabillard, V., Zumofen, R., & Dubois, P. (2024). *Public communication on social media: Challenges, risks, and perspectives*. Quebec: University of Quebec Press.

[12] Matsiuk, H. (Ed.). (2021). *Sociolinguistics: An integrative research approach*. Lviv: Ivan Franko National University of Lviv.

[13] Määttä, S.K. (2023). Linguistic and discursive properties of hate speech and speech facilitating the expression of hatred: Evidence from Finnish and French online discussion boards. *Internet Pragmatics*, 6(2), 156-172. doi: 10.1075/ip.00094.maa.

[14] McCulloch, G. (2020). *Because Internet: Understanding how language is changing*. New York, NY: Vintage Books.

[15] Millette, M., & Boislard, M.-A. (2023). "Come on, it's not that bad!": Social support and expertises in exchanges about sexual inexperience among young adults on Reddit. *Questions de Communication*, 43, 33-60.

[16] Moïse, C., & Hugonnier, C. (2019). Homophobic discourse. Testimony as an alternative discourse. *Semen: Revue de Sémio-Linguistique des Textes et Discours*, 47, 121-136. doi: 10.4000/semen.12795.

[17] Mykhalchuk, O. (2024). Language and territory: Sociolinguistic parameters of linguistic identity. In B. Azhniuk (Ed.), *Language and war: Dynamics of the language system and language policy* (pp. 411-459). Kyiv: Dmytro Buraho Publishing House.

[18] Papcunová, J., Martončik, M., Fedáková, D., Kentoš, M., Bozogáňová, M., Srba, I., Moro, R., Pikuliak, M., Šimko, M., & Adamkovič, M. (2023). Hate speech operationalization: A preliminary examination of hate speech indicators and their structure. *Complex & Intelligent Systems*, 9(3), 2827-2842. doi: 10.1007/s40747-021-00561-0.

[19] Romanchenko, A.P. (2024). Communicative support in media discourse during the wartime period: Specifics of implementation. *Scientific notes of V.I. Vernadsky Taurida National University Series Philology Journalism*, 1(1), 59-66. doi: 10.32782/2710-4656/2024.1.1/10.

[20] Sikorska, V., Nenov, I., Nogai, V., & Didur, L. (2024). Language discourse in social networks: An analysis of linguistic means and communicative strategies in the online environment. *Black Sea Philological Studies*, 5, 123-128. doi: 10.32782/bsps-2024.5.18.

[21] Shvelidze, L.D. (2021). *Linguistic means of implementation of communicative strategies in the social media discourse*. (PhD dissertation, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine).

[22] Strelēvica-Ošiņa, D. (2019). The language of correctness: Some terms of Latin origin. *Antiquitas Viva*, 5, 185-194. doi: 10.22364/av5.16.

[23] Tereshchenko, L., & Stanislavova, L. (2024). Lexical and semantic categories: A linguistic perspective. *International Humanitarian University Herald Philology*, 65, 128-131. doi: 10.32782/2409-1154.2024.65.27.

[24] Tiffany, K. (2021). *Maybe you missed it, but the Internet "died" five years ago*. Retrieved from https://www.theatlantic.com/technology/archive/2021/08/dead-internet-theory-wrong-but-feels-true/619937/.

[25] Udovichenko, H.M., & Samoilenko, H.A. (2021). Certain language means of emotions expression in superior internet communication. *Academic Studies. Series "Humanities"*, 3, 153-159. doi: 10.52726/as.humanities/2021.3.22.

[26] Voirol, O., & Martini, É. (2023). The discursive construction of hatred: Affects, fascist agitation, and the politics of resentment. *Réseaux*, 241(5), 39-77. doi: 10.3917/res.241.0039.

Особливості функціонування мови та здійснення комунікації у соціальних мережах

Світлана Гейко
Кадидат філософських наук, доцент
Національний університет біоресурсів та природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
https://orcid.org/0009-0008-0119-7673
Олена Лаута
Кадидат філософських наук, доцент
Національний університет біоресурсів та природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
https://orcid.org/0000-0002-6118-0258

Анотація. У статті проаналізовані особливості комунікації в соціальних мережах, а також сутнісні характеристики та існуючі класифікації мови ворожнечі та мови підтримки. Необхідність глибшого розуміння цих явищ зумовлена їхнім впливом на суспільство, комунікаційні норми та необхідністю розробки ефективних механізмів протидії агресії в мережі. Мета дослідження полягала в виявленні, аналізі та систематизації комунікативних стратегій для вираження мови ворожнечі та підтримки в соціальних мережах. Методологія включала контент-аналіз текстів, структурно-семантичне моделювання та соціолінгвістичне спостереження. Результати дослідження показали, що мова соціальних мереж є гібридом, для якого характерні неформальність та стислість, такі як прості речення, аграматизми, скорочення; мультимодальність, тобто використання емодзі, мемів, зображень; гіпертекстуальність та компресія, а саме використання гештегів, лінків; вплив технічних обмежень, таких як обмеження кількості символів стимулює креативність; адаптація до цензури, а саме модифікація слів для обходу алгоритмів модерації. Мова ворожнечі реалізується через стратегії сарказму, маніпуляції, опозиції «свій-чужий» та використання модифікованої лексики для обходу цензури. Мова підтримки переважно виражається через стратегії емоційної підтримки, інформаційної допомоги та використання вербальних і невербальних маркерів (вподобайки, ретвіти, гештеги). Взаємодія полярних дискурсів ворожнечі та підтримки у цифровому середовищі відображає адаптацію комунікативних практик до умов онлайн-спілкування та алгоритмів модерації. Дослідження виявило, що ці дискурси часто існують у паралельних просторах (наприклад, у відповідях на один і той же твіт). Спільною стратегією для обох є використання опозиції «свій-чужий». Мова підтримки часто виникає як реакція на мову ворожнечі, формуючи механізм колективного опору

Ключові слова: мова ворожнечі; мова підтримки; інтернет-комунікація; соціальні мережі; комунікативні стратегії

International Journal of Philology, 29(3), 2025
Journal homepage: https://philologicalscience.com.ua/en
Received 03.05.2025 Revised 17.08.2025 Accepted 23.09.2025

UDC 001.83:165.192 Doi: 10.31548/philolog/3.2025.74

Scientific discourse through the lens of the internal form of the term

Liudmyla Shanaieva-Tsymbal*
PhD in Public Administration, Associate Professor
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
https://orcid.org/0000-0002-4571-3827

Hamest Tamrazyan
PhD in Philology, Associate Professor
Swiss Federal Institute of Technology in Lausanne
1015, INN 116 Station 14 Str., Lausanne, Switzerland
https://orcid.org/0000-0001-9630-0195

Nataliia Yamnych
Senior Lecturer
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
https://orcid.org/0000-0003-4298-9075

Abstract. Discourse is a dynamic linguistic phenomenon that occupies an intermediate position between language and speech, and its study through the lens of the internal form of the term allows for the revelation of its semantic motivation and cognitive foundations. The research aimed to investigate the functioning of terms in scientific discourse and to analyse the importance of their internal form for precise interpretation and communication. The methodological framework included descriptive, analytical, component, definitional, comparative, contextual and discourse analysis. It was revealed that a great number of terms show clear evidence of internal form indicators. The internal form is substantiated to function as a cognitive mechanism, connecting linguistic form and conceptual content. Statistical analysis confirmed that the majority of terms possess at least partial motivation, either morphological, semantic, or metaphorical. The analysis noted that metaphors play a crucial role, extending common physical experiences into abstract

Suggested Citation:
Shanaieva-Tsymbal, L., Tamrazyan, H., & Yamnych, N. (2025). Scientific discourse through the lens of the internal form of the term. *International Journal of Philology*, 29(3), 74-85. doi: 10.31548/philolog/3.2025.74.

*Corresponding author

economic reasoning. It was further revealed that metaphors showcase the internal form as a link between empirical experience and theoretical abstraction. The cross-linguistic comparison showed that English-Ukrainian term pairs display full or partial motivation equivalence. It was identified that the internal form also preserves conceptual precision by anchoring the term to its original semantic motivation, which is particularly vital in written discourse. The findings confirmed that the internal form of a term is an active component in meaning construction and a mental schema that mediates between an abstract concept and its linguistic expression. Its cognitive, communicative, and epistemological functions were studied. The research concluded that preserving the internal form in translation and adhering to terminological standards is vital for maintaining semantic accuracy and international consistency in scientific communication

Keywords: linguistic expression; terminology; cognitive motivation; semantic accuracy; conceptual content; motivational equivalence; metaphor

Introduction

Discourse holds a key position in modern linguistics, extending beyond mere linguistic forms to encompass purposeful social action that bridges language, thought, and communication, integrating non-verbal elements and reflecting pragmatic, sociocultural and cognitive factors. The relevance of this study lies in its approach to scientific discourse through the internal form of the term, which is essential for understanding the semantic motivation and cognitive structure of specialised language units. This perspective is important for ensuring the accuracy and conceptual clarity necessary in scientific communication, particularly in interdisciplinary research and translation. Despite the dynamic evolution of discourse studies from text-centric analysis to a complex communicative and social phenomenon, a unified understanding of its essence and typology, particularly in scientific contexts, remains elusive, leading to conceptual ambiguity in classifying terminological systems.

Analysis of recent studies and publications (2019-2025) demonstrated a pronounced growing interest in the intersection of discourse research and terminology. In the theoretical domain of discourse studies, scholars continue to refine the field's foundations: P. Cap (2019) characterised discourse studies as a trans-disciplinary project rooted in two philosophical bases: social constructionism and linguistics and identified a gradual evolution toward models that consolidate social theory and linguistic analysis as well. The scholar discussed recent cognitive-pragmalinguistic models, such as Discourse Space Theory, Critical Metaphor Analysis and the Legitimisation-Proximisation Model, noting that these new theories represent a consolidation of the social-theoretical and linguistic bases. O. Gryshchenko (2024) emphasised that discourse functions as communication, dialogue and interaction between participants, involving both linguistic and extralinguistic factors.

In the specialised domain of terminology and scholarly language, contemporary research focuses on quantifying language barriers and defining precision. L. Li *et al*. (2022) developed an interpretable approach for measuring scholarly jargon, defining it as discipline-specific word types and senses. Their study highlighted the "gatekeeping effect" of specialised language, which, while facilitating in-group communication, can hinder understanding for out-groups. They demonstrated that jargon is nearly always negatively correlated with interdisciplinary impact. L. Shanaieva-Tsymbal & N. Yamnych (2023) investigated the enhancement

of professional communication in public administration by counteracting the increasing informality trend through the strategic study of phraseological units and metaphors in foreign language learning. They argued that metaphors serve as cognitive tools that, when used correctly, provide precision, conciseness and cultural authenticity, conveying complex ideas and potentially mitigating conflict.

Complementary research explored the cognitive dimension of terminological phenomena. R. Rehak (2023) specifically addressed the complexity within Artificial Intelligence discourse, noting the pervasive lack of clarity and standardisation in terminology. Author's central conclusion characterised this ambiguity as a "language labyrinth", underscoring how conceptual inconsistency impedes effective communication and shared understanding within the field. However, author's work focused primarily on the consequence of the problem – the resulting confusion – without deeply exploring the remedy offered by analysing a term's cognitive or structural properties, such as its internal form. In contrast, O. Lysychenko & K. Melko (2025) investigated the structural models and evolution of terms in Ukrainian IT discourse, concentrating on the dynamics of linguistic adaptation. Their key finding was the demonstration that translated Ukrainian equivalents of English verbs can actively serve as a source for further terminological derivation through semantic expansion. This highlights a critical mechanism of how specialised vocabulary expands and adapts under cross-linguistic influence, offering insights into the process of creating new terms. Nevertheless, this study focused on structural derivation and linguistic change and did not extensively analyse the role of the internal form in making these new or derived terms inherently more comprehensible or semantically motivated for the user.

The convergence of these two works reveals the primary justification for the current research. While R. Rehak (2023) established the difficulty of a "conceptual labyrinth" and O. Lysychenko & K. Melko (2025) detailed a mechanism of structural expansion, neither study sufficiently explored the internal form of the term as the fundamental cognitive link between a term's structure and its precise interpretation. The need to directly connect the analysis of a term's internal form to the goal of achieving conceptual clarity and accuracy in scientific discourse represents the crucial gap that this study aims to fill. Eastern European linguistics has made significant contributions to explaining the notion of discourse as a form of communicative activity that ties together speech, thought and social action. As F. Batsevych (2004) defined it, discourse is "a type of communicative activity, an interactive phenomenon, a flow of speech that occurs within a specific channel of communication and is regulated by the strategies and tactics of the participants".

The present study set out to examine how terms operate within scientific discourse and to analyse the significance of their internal form for accurate interpretation and clear communication. The main objectives encompassed recognising the semantic and cognitive motivation of selected terms, examining their definitional structure and defining the degree of transparency between the inner form and the conceptual content. Academic papers from the field of economics have been analysed, as this area is currently characterised by significant conceptual renewal and active terminological borrowing.

Materials and Methods

The materials selected for this study included a specialised corpus of texts from the field of economic sciences, which was chosen due to its current characteristics of significant conceptual renewal and active terminological borrowing. The approach is fundamentally rooted in the understanding that studying the term's internal form requires a blend of linguistic, cognitive, and communicative methods. The

internal form refers to the semantic motivation or the underlying figurative meaning of the term's component morphemes, connecting the literal structure of the term to the specialised concept it denotes. Terminology was classified into three categories based on the clarity of this connection: transparent, semi-transparent and opaque terms.

The methodological foundation was based on the works of O. Tsaruk (2016), P. Cap (2019) and L. Li *et al*. (2022). Further theoretical perspectives were informed by the research of N. Suzen *et al*. (2020). Additional grounding was found in the studies by O. Lysychenko & K. Melko (2025), S. Albota (2024), and T.A. van Dijk (2017). Further insights were drawn from M.V. Tomakhiv (2015), I. Synytsia (2021) and O. Stadnichenko *et al*. (2024). The research was additionally supported by the works of S. Shepitko (2010) and R. Rehak (2023). The framework incorporated the findings of I.A. Kolesnikova (2009), H.V. Holubova (2022) and M. Holianych (2008). Finally, the methodology was informed by the studies of K. Hokhman (2018), O. Gryshchenko (2024) and F. Batsevych (2004). The specialised nature of the corpus was consistent with general findings that scientific discourse is highly saturated with specialised vocabulary, affirming the term's role as the main carrier of professional knowledge and the primary tool for conveying specialised information.

The source base was meticulously compiled to ensure specificity and representativeness within the field of Economics. The corpus comprised over 1,500 terminological units that were were deemed sufficient because the selection process aimed for semantic saturation. The extraction of the linguistic material was conducted using a selective sampling method, which involved a combination of automated and manual processing. The initial extraction was semi-automated, utilising AntConc to generate word frequency lists and N-gram concordances. This assisted in quickly identifying high-frequency noun phrases, which form the structural core of most terminology.

A terminological unit was identified based on its function as a name for a specialised concept within the field of economics, requiring a definitional context that distinguished it from general language use. This included nouns and noun phrases, adjective-noun constructions, verbs and verb phrases. The boundaries of a terminological unit were determined based on the minimum required set of lexical items needed to fully realise the specialised. The initial automated lists were subject to manual verification and selection to ensure each included unit met the strict criterion of specialised conceptual reference. Descriptive and analytical methods were applied for the inventory and characterisation of terms, enabling the identification and classification of the terms found in the selected corpus. The descriptive method was used for collecting and systematising information regarding the problem of term definition. It was combined with structural and classification approaches to identify and describe the main structural models of term formation. The analytic method involved studying the linguistic means employed and determining the place of terminological vocabulary in the overall lexicon.

Component and definitional analysis were utilised to determine the internal form and semantic motivation of each term. This qualitative analysis was necessary because, as attested in terminology theory, terminology leans heavily on the structure and metaphorical power of common vocabulary. The method of definitional analysis relied on the principle of using dictionary definitions as a reliable source for identifying thematic concepts. Definitions were regarded as "mini-knowledge-representations", explaining the location of a concept in the specialised domain's conceptual structure. Their analysis determined the meaning of terms, along with linguistic characteristics such as part of speech, gender, or frequency. Component analysis was used to investigate

the lexical content and semantic structure. In this approach, term meanings were examined through the sequential comparison of terms grouped into one lexical microsystem based on a specific integral semantic feature.

The comparative method was employed to reveal semantic shifts and cross-linguistic equivalences between English and Ukrainian economic terms. The comparison of terms across languages was essential for terminological translation, especially for harmonising national terminologies. Determining equivalence relied on a detailed dual analysis of both knowledge structures and their linguistic representations. Contextual and discourse analysis were used to study the functioning of terms in the structure of scientific discourse and to trace their pragmatic roles. Contextual analysis was applied to describe the influence of the immediate environment on the understanding of terms. Terminology analysis within text involved treating terms as pivots (term pivot) to identify their lexical and predicative links with other units in a corpus. Local linguistic context was vital for term disambiguation and subsequent conceptual modelling.

The research was performed in three main stages: compilation of the corpus, extraction and analysis of terminological units based on structure, motivation, and internal form, and utilisation of findings to draw conclusions regarding terms' roles in creating cohesive scientific discourse. Following extraction and verification, term density was calculated by dividing the total number of terminological units in each text by the overall word count. Word counts were obtained using AntConc, and term lists were verified against economic glossaries to exclude non-specialised vocabulary. The resulting percentages represent the proportion of terms relative to total words in each genre, enabling quantitative comparison across research articles, textbooks, and methodological guides.

In addition to frequency, terms were categorised according to transparency of their internal form using a three-tiered scale. Fully transparent terms clearly reflect their conceptual meaning and can be interpreted without additional context. Semi-transparent terms partially convey meaning and require domain knowledge. Opaque terms do not reveal meaning through internal form and rely entirely on convention. Classification was based on morphological composition and semantic motivation, independently assessed by two researchers with disagreements resolved through discussion, ensuring consistent and replicable assignment linking quantitative term usage with qualitative cognitive and linguistic analysis.

Results and Discussion

The concept of internal form in terminology has long attracted scholarly attention, tracing back to Aristotle, who posited that names do not exist inherently in nature but acquire meaning when they become symbolic representations (Aristotle, 1963; Matthews, 1997). Even seemingly inarticulate sounds, such as those produced by animals, may convey meaning, illustrating that linguistic signs are not entirely arbitrary. Contemporary research in specialised language has extended this idea, suggesting that while general vocabulary may sometimes lack clear motivation, professional terms are generally cognitively grounded. Terminological units are not randomly assigned; rather, they emerge from a deliberate conceptual necessity. This principle applies even to borrowed terms: for example, many English-origin terms in economic discourse, such as soft currency, maintain their semantic motivation in Ukrainian equivalents, reflecting both structural and conceptual inheritance from the source language. Consequently, internal form can be regarded as a fundamental characteristic of professional vocabulary, linking linguistic expression with conceptual structure.

Previous research examined the patterns, cognitive functions, and cross-linguistic

correspondences of internal form in economic terminology. Based on analyses of English and Ukrainian economic texts, scholars identified morphological, semantic, and conventionalised types of motivation and highlighted the important role of metaphor in conceptual structuring. Earlier studies also addressed the extent to which internal form was preserved in translation and discussed its influence on the precision of scientific discourse. These findings contributed to a broader understanding of internal form as a mnemonic and epistemological component of terminology. Building on these typologies, researchers have also distinguished degrees of transparency that determine how clearly the internal form of a term reveals its underlying concept.

Transparent terms preserve internal form fully and directly or metaphorically convey the specialised concept without requiring extensive domain knowledge for initial comprehension (e.g., liquidity trap – which metaphorically evokes a situation in which money is "trapped" due to low interest rates). Semi-transparent terms provide only partial motivation: their constituent morphemes offer a general hint, while full interpretation requires professional knowledge or familiarity with the definitional context (e.g., hidden unemployment, where "unemployment" is clear but "hidden" demands understanding of statistical measurement and reporting practices). In contrast, opaque terms lack an accessible internal form: the meaning of the term is arbitrary from the perspective of its surface structure and must be learned entirely through convention (e.g., Giffen good, named after economist Robert Giffen, whose surname provides no semantic cue to the economic paradox described by the term). This classification complements existing motivational frameworks by illustrating how terminological transparency shapes ease of conceptual access.

The significance of internal form in terminology has been emphasised since classical philosophy. Aristotle (1963) noted that names do not inherently exist in nature but acquire meaning when they function as symbols; even non-verbal sounds, such as animal calls, can convey information. In modern studies, the question of whether all professional terms possess internal form remains debated. Specialised terms are rarely arbitrary: each term exhibits some degree of cognitive motivation. This applies even to borrowed terms: for example, many English-origin terms in economic discourse, particularly those formed morphologically, retain motivation in Ukrainian equivalents. Soft currency, for instance, illustrates how the internal form is inherited from the source language, reflecting both structure and meaning. These observations support the view that internal form is a pervasive characteristic of terminological systems, linking linguistic expression with conceptual content.

Furthermore, the analysis revealed that terms with strong metaphorical motivation tend to appear in argumentative or explanatory contexts, whereas morphologically transparent terms prevail in definitional statements. This observation is supported by adopting quantitative metrics developed by scholars such as N. Suzen *et al*. (2020), who utilised Relative Information Gain to determine the science-specific meaning and importance of a word, and L. Li *et al*. (2022), who employed normalised pointwise mutual information to measure scholarly jargon. Comparative analysis is further confirmed as a valuable tool by O. Lysychenko & K. Melko (2025), who demonstrated its utility for identifying structural transformations in terminological units.

A corpus analysis of 1,500 terminological units revealed that 72% of terms display indicators of internal form. Among these, 48% are morphologically or etymologically motivated, while 24% exhibit semantic or metaphorical motivation. The remaining 28% are conventionalised or opaque, with their meaning accessible only through convention rather than word structure. Morphologically motivated terms,

such as microeconomics or underemployment, demonstrate the productivity and transparency characteristic of scientific vocabulary. Semantic motivations often involve metaphoric or metonymic transfer, as in "bubble economy" or "invisible hand", reflecting the extension of concrete experience into abstract economic reasoning. Borrowed terms, such as "soft currency", highlight the donor-language motivation, while conventionalised terms like "bond" or "stock" illustrate opaque forms whose internal structure no longer transparently conveys meaning (Table 1).

Table 1. *Types of motivation in economic terms*

Type of Motivation	Number of Terms	Percentage	Example
Morphological (affixation, compounding)	720	48	microeconomics, underemployment
Semantic (metaphoric, metonymic)	360	24	bubble economy, invisible hand
Borrowed/Loan translation	270	18	soft currency, trade-off
Conventionalised (opaque)	150	10	bond, stock

Source: *compiled by the authors*

The findings suggest that motivation in terminology functions as a cognitive mechanism connecting linguistic form with conceptual understanding. Internal form serves not only as a mnemonic device but also as a reflection of the cognitive models underlying economic reasoning. Metaphorical structures in terminology contribute significantly to conceptual coherence. Terms such as "market elasticity", "financial bubble", or "liquidity trap" extend tangible experiences to abstract economic phenomena, linking perception and cognition. According to G. Lakoff & M. Johnson (1980), metaphor is not merely rhetorical but a fundamental cognitive mechanism. In the compiled corpus, 24% of terms rely on metaphorical transfer. For instance, "liquidity" originally denotes a physical property of substances but in economics signifies the ease of converting assets into cash. Similarly, "economic depression" borrows from the physical sense of downward motion to represent a decline in economic activity. These examples demonstrated that internal form preserves traces of embodied cognition even within highly abstract domains.

Cross-linguistic comparison reveals that internal form is not always maintained in translation. Among English-Ukrainian term pairs, 63% retain full or partial motivational equivalence, while 37% exhibit semantic shifts or neutralisation (Table 2). Translation strategies such as calquing preserve conceptual metaphors (floating exchange rate – плаваючий валютний курс; plavayuchyi valiutnyi kurs), whereas partial borrowings may reduce motivational transparency (hedging – хеджування; khedzhuvannia). Some terms, like bond – облігація (obligatsiya), enrich imagery in Ukrainian, while others lose structural motivation, reflecting differences in metaphorical and conceptual frameworks across languages.

Table 2. *Cross-linguistic motivation correspondence*

Motivation Equivalence	Percentage	Example
Full equivalence	41	floating exchange rate – плаваючий валютний курс (plavayuchyi valiutnyi kurs)
Partial equivalence	22	hedging – хеджування (khedzhuvannia)
Neutralisation	37	leverage – фінансовий важіль (finansovyi vazhil)

Source: *compiled by the authors*

Precision is crucial in scientific communication. It was shown that 95% of terms in written discourse are used consistently, compared to 83% in spoken contexts. In lectures or seminars, near-synonyms or paraphrases occasionally replace exact terms, risking terminological ambiguity. For example, "costs" and "expenses" may be interchanged orally, whereas written texts demand strict differentiation. Thus, internal form ensures conceptual precision by anchoring terms to their semantic motivation. Term frequency analysis indicates variation across genres (Table 3). Research articles display higher terminological density (30%), reflecting focused argumentation and data presentation, whereas textbooks and methodological guides show lower density (21-24%) due to explanatory content. These patterns align with previous observations that terminological saturation correlates with scientific abstraction and specialisation.

Table 3. *Term density across genres*

Genre	Average Words per Text	Number of Terms	Term Density (%)
Research article	5,000	1,500	30
Textbook chapter	7,000	1,700	24
Methodological guide	4,500	950	21

Source: *compiled by the authors*

Comparing the present findings with previous linguistic and cognitive research reveals both significant parallels and novel insights. Earlier studies estimated that approximately 30-40% of vocabulary in scientific discourse consists of specialised terminological units. This estimation aligns closely with corpus-based findings, which indicate an average terminological density of 31.8% across English and Ukrainian economic texts. While these figures suggest continuity in terminological usage patterns, the analytical scope was extended by not only quantifying the frequency of terminological units but also examining the nature of their internal form and the types of motivation underlying these terms. Unlike the descriptive and typological studies of the 20th century, which mainly catalogued terms without exploring their pragmatic or functional roles, the current research incorporates discourse analysis to examine how the internal form shapes the communicative structure of scientific texts. It was revealed that terms with strong metaphorical motivation, such as liquidity trap or financial bubble, frequently appear in argumentative, explanatory, or illustrative contexts. These metaphorically motivated terms serve as cognitive bridges, mapping abstract economic concepts onto more concrete or familiar experiential domains, thereby facilitating comprehension and reasoning. On the other hand, morphologically transparent terms, such as "microeconomics" or "underemployment", are most prevalent in definitional, descriptive, or formal statement contexts, where terminological precision and clarity are paramount. This functional stratification demonstrates that the type of motivation is closely linked to the discourse function of a term, reflecting the communicative purpose and context in which it occurs.

The study further identified a gradient of motivation transparency within economic terminology, revealing how cognitive salience and linguistic form interact. Using a three-tiered scale, terms were categorised as fully transparent, semi-transparent, or opaque. Fully transparent terms, such as "microloan" or "inflationary spiral", possess clear morphological or semantic motivation, allowing their meaning to be readily inferred from form. Semi-transparent terms, for example "credit crunch" or "liquidity trap", retain partial motivation, often

requiring contextual knowledge or cognitive mapping to be fully understood. Opaque terms, such as "bond" or "derivative", no longer display overt internal motivation and rely instead on convention, definition, or domain-specific expertise for comprehension.

Statistical analysis has shown that 43% of terms fall into the fully transparent category, 29% into the semi-transparent category, and 28% into the opaque category. This distribution indicates that, despite high specialisation, scientific discourse maintains a substantial cognitive connection with general language. Terminology is not an isolated or self-contained system; rather, it draws heavily on familiar linguistic structures and metaphorical frameworks to facilitate understanding. Morphological and metaphorical features provide the scaffolding that links abstract concepts to tangible representations, while opaque terms often emerge from historical development or borrowing processes, reflecting the evolution of knowledge within the discipline.

Understanding the internal form has important practical implications for terminological standardisation, particularly in multilingual and international contexts. International bodies, increasingly recognise the need for motivation-based equivalence to preserve semantic integrity across languages. This perspective is supported by evidence demonstrating that terminological stability is higher when internal form and cognitive motivation are preserved during translation. For instance, the Ukrainian calque "плаваючий валютний курс" (plavayuchyi valiutnyi kurs) fully reproduces the conceptual metaphor of instability inherent in "floating exchange rate", ensuring equivalence at both linguistic and cognitive levels. In contrast, partial borrowings, such as "хеджування" (khedzhuvannia) from "hedging", reduce interpretative clarity and may require supplementary explanation to achieve conceptual transparency. Maintaining motivation in cross-linguistic transfer not only facilitates comprehension but also strengthens the epistemological function of terminology as a vehicle for precise scientific reasoning.

Beyond practical considerations, the analysis revealed that internal form functions as a dynamic interface between thought and language. This aligns with theories in cognitive linguistics, particularly the concepts of frames and conceptual metaphor (Lakoff & Johnson, 1980). Each term embodies a conceptual frame that encapsulates knowledge structures and logical relationships within a field, allowing abstract concepts to be encoded in language. Morphologically transparent and metaphorically motivated terms exemplify this interaction, enabling cognitive processing and recall. The findings further support the hypothesis that terminological motivation is a universal phenomenon, though it is variably realised depending on the stage of conceptual development, degree of borrowing, and language-specific cognitive patterns. Such a perspective aligns with the works of R. Temmerman (2000), who emphasised the dynamic, context-dependent nature of terminology. Motivation serves as a cognitive anchor, connecting terms to mental models, historical usage, and cultural conceptualisation, thereby making scientific vocabulary both efficient and comprehensible. While the present study provided a comprehensive statistical, qualitative, and cross-linguistic analysis, certain limitations remain. The corpus was restricted to English and Ukrainian economic texts, which may not fully represent other languages, fields, or cultural contexts. Additionally, the study relies on manual analysis of motivation, which, despite being rigorous, may be complemented by computational tools to scale the investigation.

Conclusions

Discourse represents a dynamic form of linguistic activity operating within social and

cognitive contexts, and scientific discourse, in particular, stands out for its organisation, terminological precision, and functional coherence. Analysis of a corpus demonstrated that the internal form of a term is central to preserving conceptual integrity. Statistical findings indicate that the majority of terms possess some degree of motivation – morphological, semantic, or metaphorical. Morphologically motivated terms, such as "microeconomics" or "underemployment", demonstrate transparency and productivity, while semantically and metaphorically motivated terms, like "bubble economy" or "liquidity trap", reveal the role of conceptual metaphor in structuring abstract concepts.

Comparative analysis of English-Ukrainian term pairs showed that 63% of terms retain full or partial motivational equivalence, while 37% exhibit semantic shifts or neutralisation. Terms such as floating exchange rate – плаваючий валютний курс (plavayuchyi valiutnyi kurs) preserve metaphorical and morphological cues, whereas others, like hedging – хеджування (khedzhuvannia), show reduced transparency. These patterns highlight that motivation varies across languages and disciplines, reflecting both universal cognitive processes and culture-specific conceptualisations.

The internal form contributes to multiple dimensions of scientific discourse. Cognitively, it links abstract notions to mental representations, facilitating comprehension and retention. Communicatively, it supports mutual understanding among specialists through transparent and logically motivated terminology. Epistemologically, it serves as a vehicle for conceptualisation, reflecting the historical development of disciplinary knowledge. The coexistence of morphological and metaphorical motivation demonstrates the dual nature of scientific language, combining systematic regularity with creative elaboration.

Furthermore, corpus-based analyses indicated higher terminological density in research articles (30%) compared to textbooks and methodological guides (21-24%), suggesting that terminological saturation correlates with abstraction and argumentation. Consistency of term usage was also observed, with 95% of terms in written discourse appearing consistently versus 83% in spoken contexts. These quantitative results support the conclusion that internal form and cognitive motivation are crucial for coherent and precise scientific communication. Future research should expand to additional disciplines and languages, employing corpus-based, computational, and psycholinguistic methods to quantify the cognitive and communicative effects of internal form and motivation. Understanding these dynamics promises to deepen insight into the interplay between language, cognition, and scientific reasoning, demonstrating that terminology is not arbitrary but a reflection of conceptual sophistication in specialised fields.

Acknowledgements

None.

Funding

None.

Conflict of Interest

None.

References

[1] Albota, S. (2024). Linguistic representation of Ukraine in English-language media discourse: A corpus-assisted approach. In *CLW-2024: Computational linguistics workshop at 8th international conference on computational linguistics and intelligent systems (CoLInS-2024)* (pp. 1-28). Lviv: Lviv Polytechnic National University.

[2] Aristotle. (1963). *Categories and de interpretatione*. Oxford: Oxford University Press.

[3] Batsevych, F. (2004). *Fundamentals of communicative linguistics*. Kyiv: Academy.
[4] Cap, P. (2019). Discourse studies: Between social constructionism and linguistics. A critical overview. *Topics in Linguistics*, 20(2), 1-16. doi: 10.2478/topling-2019-0006.
[5] Gryshchenko, O. (2024). Discourse: Knowledge, news, and fake intertwined. *Philological Review*, 1(23), 4-11. doi: 10.31499/2415-8828.1.2024.306854.
[6] Hokhman, K. (2018). Academic discourse in the system of institutional discourses. *Scientific Bulletin of South Ukrainian National Pedagogical University named after K.D. Ushynsky*, 26, 37-47.
[7] Holianych, M. (2008). *Internal form of words and discourse*. Ivano-Frankivsk: Publishing and Design Department of the Information and Communication Technology Center of the Vasyl Stefanyk Precarpathian National University.
[8] Holubova, H.V. (2022). Variability of the concept "scientific discourse": Modern dimension. *Transcarpathian Philological Studies*, 1(25), 89-92. doi: 10.32782/tps2663-4880/2022.25.1.17.
[9] Kolesnikova, I.A. (2009). *Linguacognitive and communicative-pragmatic characteristics of the professional discourse*. (Doctoral thesis, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine).
[10] Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press. doi: 10.7208/chicago/9780226470993.001.0001.
[11] Li, L., Dodge, J., Bamman, D., & Keith, A.K. (2022). Words as gatekeepers: Measuring discipline-specific terms and meanings in scholarly publications. *ArXiv*. doi: 10.48550/arXiv.2212.09676.
[12] Lysychenko, O.B., & Melko, K.B. (2025). Structural models of term formation in Ukrainian IT discourse: Transformations and trends. *Messenger of Kyiv National Linguistic University. Series Philology*, 28(1), 122-130. doi: 10.32589/2311-0821.1.2025.335651.
[13] Matthews, P.H. (1997). *The concise Oxford dictionary of linguistics* (3rd ed.). Oxford: Oxford University Press. doi: 10.1093/acref/9780199675128.001.0001.
[14] Rehak, R. (2023). The language labyrinth: Constructive critique on the terminology used in the AI discourse. In P. Verdegem (Ed.), *AI for everyone?: Critical perspectives* (pp. 87-102). London: University of Westminster Press. doi: 10.16997/book55.f.
[15] Shanaieva-Tsymbal, L., & Yamnych, N. (2023). Enhancing effective communication in public administration: The role of phraseological units in foreign language learning. *Cogito – Multidisciplinary Research Journal*, 15(3), 127-135.
[16] Shepitko, S. (2010). Scientific discourse components. *Research Bulletin. Series: Philological Sciences*, 89(5), 164-167.
[17] Stadnichenko, O., Stasyk, M., Shulha, O., Nazarenko, I., & Chaikovska, U. (2024). Interaction of cultural and linguistic aspects of Ukrainian philology: Modern vectors of scientific discourse. *Synesis*, 16(3), 169-187.
[18] Suzen, N., Mirkes, E.M., & Gorban, A.N. (2020). Informational space of meaning for scientific texts. *ArXiv*. doi: 10.48550/arXiv.2004.13717.
[19] Synytsia, I. (2021). The discursive opposition and the "power of discourse" in the Ukrainian mass media. *SHS Web of Conferences*, 100, article number 02007. doi: 10.1051/shsconf/202110002007.
[20] Temmerman, R. (2000). Why traditional terminology theory impedes a realistic description of categories and terms in the life sciences. *Terminology: International Journal of Theoretical and Applied Issues in Specialised Communication*, 5(1), 77-92. doi: 10.1075/term.5.1.07tem.
[21] Tomakhiv, M.V. (2015). Scientific English discourse: The current state of development and future prospects of study. *Odesa Linguistic Journal*, 5, 154-157.

[22] Tsaruk, O. (2016). Semantic constituents of a scientific and technical term. *Scientific Notes of Ostroh Academy National University, Philology Series*, 60, 247-249.
[23] Van Dijk, T.A. (2017). *Discourse and power*. London: Bloomsbury Publishing.

Науковий дискурс крізь призму внутрішньої форми терміна

Людмила Шанаєва-Цимбал
Кандидат наук з державного управління, доцент
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
https://orcid.org/0000-0002-4571-3827
Амест Тамразян
Кандидат філологічних наук, доцент
Швейцарський федеральний політехнічний інститут в Лозанні
1015, вул. Станція 14, INN 116, м. Лозанна, Швейцарія
https://orcid.org/0000-0001-9630-0195
Наталія Ямнич
Старший викладач
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
https://orcid.org/0000-0003-4298-9075

Анотація. Дискурс є динамічним лінгвістичним феноменом, що займає проміжне місце між мовою та мовленням, і його вивчення через призму внутрішньої форми терміна дозволяє розкрити його семантичну мотивацію та когнітивні основи. Дослідження мало на меті дослідити функціонування термінів у науковому дискурсі та проаналізувати важливість їх внутрішньої форми для точного тлумачення та комунікації. Методологічна база включала описовий, аналітичний, компонентний, дефініційний, порівняльний, контекстуальний та дискурсивний аналіз. Було виявлено, що значна кількість термінів має чіткі ознаки внутрішньої форми як когнітивного механізму, що поєднує мовну форму та концептуальний зміст. Статистичний аналіз підтвердив, що більшість термінів мають принаймні часткову мотивацію, морфологічну, семантичну або метафоричну. Аналіз показав, що метафори відіграють вирішальну роль, поширюючи загальні фізичні враження на абстрактне економічне мислення. Було також виявлено, що метафори демонструють внутрішню форму як зв'язок між емпіричним досвідом та теоретичною абстракцією. Міжмовне порівняння виявило, що пари термінів англійської та української мов демонструють повну або часткову мотиваційну еквівалентність. Було встановлено, що внутрішня форма також зберігає концептуальну точність, прив'язуючи термін до його первісної семантичної мотивації, що є особливо важливим у письмовому дискурсі. Результати дослідження підтвердили, що внутрішня форма терміну є активним компонентом у побудові значення та ментальною схемою, яка пов'язує абстрактне поняття з його лінгвістичним вираженням. Були досліджені когнітивні, комунікативні та епістемологічні функції внутрішньої форми. Дослідження довело, що збереження внутрішньої форми в перекладі та дотримання термінологічних стандартів є надзвичайно важливим для підтримання семантичної точності та узгодженості наукового спілкування на міжнародному рівні

Ключові слова: лінгвістичне вираження; термінологія; когнітивна мотивація; семантична точність; концептуальний зміст; метафора

МІЖНАРОДНИЙ ФІЛОЛОГІЧНИЙ ЧАСОПИС

Науковий журнал

Том 29, № 3, 2025

Оригінал-макет видання виготовлено у відділі науково-технічної інформації
Національного університету біоресурсів
і природокористування України

Відповідальний редактор:
Н. Шевченко

Підписано до друку 23 вересня 2025 р.
Формат 70*100/16
Умов. друк. арк. 7
Наклад 100 прим.

Адреса видавництва:
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
E-mail: ph@philologicalscience.com.ua
https://philologicalscience.com.ua/uk

INTERNATIONAL JOURNAL OF PHILOLOGY

Scientific Journal

Volume 29, No. 3, 2025

The original layout of the publication is made in the Department of Scientific and Technical Information of National University of Life and Environmental Sciences of Ukraine

Managing Editor:
N. Shevchenko

Signed for print of September 23, 2025
Format 70*100/16
Conventional printed pages 7
Circulation 100 copies

Editors Office Address:
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
E-mail: ph@philologicalscience.com.ua
https://philologicalscience.com.ua/en