

УДК 811.161.2'322:004.85:81'33

**ARCHITECTURES OF LEXICAL VOCABULARY RESTRICTION:
TRENDS AND CONTEMPORARY PRINCIPLES
IN ENGLISH AND UKRAINIAN**

Solomija Buk

*Ivan Franko National University of Lviv
Department of General Linguistic
1 Universytetska St., room 343, 79001, Lviv, Ukraine
phone: 032 239 47 56
e-mail: nouma: solomija@gmail.com
http://orcid.org/0000-0001-8026-3289*

Oksana Babelyuk

*Lviv State University of Life Safety
Lviv, Ukraine Kleparivska, 35, room 426, 79000, Lviv, Ukraine
phone: 032 239 47 56
e-mail: babelyuko@gmail.com
http://orcid.org/0000-0003-4837-1225*

This article systematically investigates the historical evolution, contemporary typology, and underlying construction principles of quantitatively restricted vocabulary systems in applied linguistics and lexicography. While conventional linguistic thought champions lexical richness, practical contexts – notably foreign language instruction and communication accessibility – necessitate the strategic implementation of limited lexicons. The study begins by tracing the foundational efforts to define essential vocabulary, rooted in the early 20th century. A historical distinction is established between the methodology employed by early simplified dictionaries, such as F. L. Thorndike's *The Teacher's Word Book* (1921), which primarily leveraged word frequency analysis across large corpora to aid readability assessment, and the contrasting approach of C. K. Ogden's *Basic English* (1925). Ogden's system deliberately reduced the English language to 850 words based on conceptual sufficiency – the capacity to convey essential ideas – rather than empirical frequency, giving rise to artificial international languages and resources like the Simple English Wikipedia.

The research then addresses the prevalent terminological diversity and conceptual overlaps inherent in this field by providing a comprehensive typology of related concepts. The analysis differentiates between: Controlled Natural Languages (CNLs), Plain Language (Plain English / Leichte Sprache), Survival Vocabulary, Vocabulary for Special Purposes etc. These systems differ in selection principles, target groups, and communicative purposes, reflecting the complex functional landscape of lexical reduction in contemporary society.

The final section focuses on the Lexical Base of Ukrainian Language (LBUL), presenting a detailed analysis of its structure, methodology, and cross-linguistic significance. The LBUL was constructed before the widespread adoption of international frequency dictionary series (such as the Routledge series, 2006–2018), relying on a balanced, 1.7-million-word multi-genre corpus representing five functional styles of modern Ukrainian. By selecting the most frequent 3,000 words from each style, the LBUL isolated a core of 1,386 lexical units. This chronological precedence highlights Ukraine's early and rigorous contribution to quantitative and applied lexicography. Crucially, a comparative analysis reveals that over 90% of the LBUL's vocabulary maps onto universal semantic fields (e.g., Nature, Human, Society, Abstract Relations) identified in other multilingual frequency dictionaries (German, Spanish, French, Turkish, Czech, and other languages). This strong structural convergence supports the hypothesis of a shared cognitive and semantic foundation across languages and positions the LBUL as an European model of corpus-based lexical universalism.

This high degree of cross-linguistic correlation demonstrates the universality of the foundational lexical layer across languages, offering strong pedagogical implications for utilizing the LBUL as a systematic “Common Framework Vocabulary” to facilitate cross-linguistic transfer and comprehension among language learners. The findings argue for a unified conceptual approach to restricted vocabularies, emphasizing the confluence of frequency, thematic coverage, and accessibility in constructing effective core lexicons.

Ultimately, the article proposes a refined conceptual taxonomy of restricted lexical systems and advanc-

es the notion of a Common Lexical Framework – a vocabulary-level complement to the CEFR – grounded in corpus-driven, cross-linguistic principles. The findings open pathways for future research involving quantitative alignment of semantic fields, integration into digital learning platforms, and broader comparative studies across diverse linguistic typologies.

Key words: Ukrainian as a foreign language, lexical base, dictionary reduction, frequency lists, language modeling, vocabulary optimization, applied linguistics, quantitative linguistics, lexicography.

DOI:

Formulation of the problem. The traditional linguistic perspective unequivocally upholds lexical richness as an essential positive characteristic of a language. However, the study of language, particularly in applied linguistics and foreign language pedagogy, reveals specific contexts where a highly controlled and quantitatively limited vocabulary offers significant practical and cognitive advantages. This principle is fundamental to the structured acquisition of a foreign language, where vocabulary is introduced progressively across established proficiency tiers (A1–C2) [11]. Yet, the utility of a simplified lexicon extends beyond the classroom, evidenced by a rising trend in employing quantitatively restricted vocabularies to enhance the accessibility and effective communication of information to diverse audiences, including language learners, non-specialists, and individuals with cognitive or language processing needs.

The historical quest to define the minimal essential lexicon for proficiency dates back to the early 20th century. This period saw the genesis of influential simplified dictionaries, driven primarily by the criterion of word frequency in large corpora. Landmark works such as F. L. Thorndike's *The Teacher's Word Book* (1921), which catalogued 10,000 common English words, provided a foundational framework for readability studies and pedagogical selection.

Despite these foundational efforts, the field is characterized by a significant terminological diversity and a lack of unified conceptual clarity among related concepts such as *Controlled Natural Language*, *Plain Language*, *Survival Vocabulary*, and *Vocabulary for Special Purposes*. C. K. Ogden's *Basic English* (1925), which prioritized conceptual importance over mere frequency, introduced a contrasting principle to word selection, leading to artificial languages and resources like the Simple English Wikipedia. This distinction highlights an existing gap in understanding the full spectrum of criteria – from frequency and communicability to thematic relevance – that drive the construction of effective restricted vocabularies. Furthermore, while numerous cross-linguistic studies on lexical cores exist (e.g., the Routledge Frequency Dictionaries), a detailed, historically contextualized analysis of the methodology and cross-linguistic consistency of the Lexical Base of Ukrainian Language (LBUL), a significant early contribution to quantitative lexicography, remains underexplored.

The geopolitical and cultural transformations of recent years have made Ukrainian not only a national symbol but also a strategic educational and cultural asset in Europe [1, 2, 13]. Teaching Ukrainian in a multilingual classroom – whether to Erasmus students, heritage learners, or foreign professionals – requires identifying a lexical common ground that unites speakers of diverse linguistic backgrounds. This common ground is represented by the core vocabulary, or lexical base, which covers the essential semantic fields of human life: nature, body, home, society, communication, cognition, values etc. Lexical base of Ukrainian language (LBUL) as a systematized model of this universal layer of vocabulary was compiled in 2004 (Buk, dyser) and it is important to analyzed it in the context of the achievements of the lexicographies of other countries, in particular in the cross-linguistic similarity to the Routledge *Frequency Dictionaries* (German, Spanish, Turkish, Czech, French, etc.).

This article, therefore, aims to:

1. Systematically review and typologize the historical background and current terminological landscape of quantitatively restricted vocabulary systems in applied linguistics.
2. Delineate and analyze the distinct principles (frequency, conceptual importance, thematic relevance) underlying the creation of various simplified lexicons, including Controlled Natural Language and Plain Language initiatives (e.g., *leichte Sprache*).
3. Provide a detailed analysis of the methodology, structure, and chronological precedence of the Lexical Base of Ukrainian Language (LBUL) and its significant cross-linguistic correlation with universal semantic fields found in contemporary international frequency dictionaries.

The ultimate goal is to establish a clearer conceptual taxonomy for these vocabulary systems and to position the LBUL within the international context of corpus-based lexicography.

1. Historical background of shortened and simplified dictionaries and related concepts.

In linguistics, it is generally accepted that a rich vocabulary is an unconditional positive characteristic of language. However, there are conditions under which richness may not be considered as such, and, conversely, the use of a limited number of words may be considered an advantage. This primarily concerns the study of a particular language as a foreign language in accordance with levels A1, A2, B1, B2, C1, and C2, which require the gradual introduction of vocabulary from a smaller to a larger number of words. But that is not all: there is a growing trend towards using a simplified, quantitatively limited vocabulary to effectively convey information to native speakers with low language proficiency: students, non-professionals in a particular field, and people with special needs.

The question of how many words one needs to know to be fluent in a language has long been of interest to humanity. At the beginning of the 20th century, a number of simplified dictionaries were created (F. Thorndike, C. Ogden, M. West, T. Corlett, etc.), where the main principle of word selection was the frequency of words in specially selected literature. F. L. Thorndike's *The Teacher's Word Book* (1921) published 10,000 of the **most commonly used words** in the English language, selected on the basis of an analysis of various texts containing a total of 4 million word uses. This list made it easier for teachers to choose books that matched the reading skills of their class and also became the basis for future research on readability.

In 1925, English writer, philosopher, and linguist C. K. Ogden identified 850 words that were "sufficient" for basic communication in his book *Basic English: A General Introduction with Rules and Grammar* (1925). He identified the most important ideas and concepts, and then selected the simplest words for these concepts. Thus, the principle for selecting this simplified version of the lexical composition of the language was the importance of a word for conveying a thought, rather than the importance of a word for understanding a text, i.e., not frequency (unlike lexical minimums). In this way, C. Ogden created 'Basic English' – an international artificial language based on English. 'Basic' stands for British, American, Scientific, International, Commercial. This term is used mainly in western lexicography. Basic English has similarities with simplified English, plain English, simple English, etc.

The idea of "Basic English" is the basis of **simple English Wikipedia** – a section

of Wikipedia created for readers who find it difficult to understand English Wikipedia and who have a basic level of English: schoolchildren, children, foreigners learning English, etc. It contains simplified and abridged versions of articles from English and other Wikipedias, using far fewer words and simpler grammar. It is effective for use in classrooms and for educational reading.



The screenshot shows the 'About Wikipedia' page of the Simple English Wikipedia. The page title is 'About Wikipedia' with a small icon of a person. The text on the page explains that it is the front page of the Simple English Wikipedia, which is a place where people work together to write encyclopedias in different languages. It uses simple English words and grammar. The Simple English Wikipedia is for everyone, including children and adults learning English. There are 256,596 articles. The pages are free to use and published under the Creative Commons Attribution/Share-Alike License 4.0 International License and the GNU Free Documentation License. It encourages users to help here, change pages, and make new pages. It also links to help pages and good pages. The page ends with a note about simple talk.

When writing articles here:

- Use **Basic English words and shorter sentences**. This allows people to understand complex terms or phrases.
- Write **good pages**. The best encyclopedia pages have useful, well-written information.
- Use the **pages to learn and teach**. These pages can help people learn English. You can also use them to **make a new Wikipedia** to help other people.
- **Simple does not mean short**. Writing in Simple English means that simple words are used. It does not mean readers want basic information. Articles do not have to be short to be simple; expand articles, add details, but use basic **vocabulary**.
- **Be bold!** Your article does not have to be perfect, because other editors will fix it and make it better. And most importantly, do not be afraid to start and make articles better yourself.

Special English or Learning English is a controlled, simplified version of English with about 1,500 words (it also avoids idioms), which was first used in 1959 and is still used today by the US government-funded television and radio organization Voice of America. Offering news, information, and cultural programs at a pace one-third slower than usual, it has a huge weekly international audience of hundreds of millions, to whom it effectively conveys information. Simplified English is aimed at foreign listeners who have an intermediate or advanced level of English. In 1962, Voice of America published Word Book (a dictionary of simplified English). In addition to definitions of selected words, it also contains additional lists: common prefixes, common expressions, numbers, days, months, chemical elements, body parts, computer terms, business terms, government structure, a list of presidents and a map of the United States, and a map of the world [8].

2. Terminological diversity of related concepts.

In this field of applied linguistics and language teaching, various other terms are also used, which differ slightly in meaning, so let us consider them in more detail.

Controlled natural language, CNL, is a general term for a simplified version of natural language. It has rules that limit the use of grammar, terminology, and idioms in order to avoid ambiguity and complexity. The purpose of restricted languages is to make them easier for humans (e.g., foreigners) to understand and to use language that is simple for automatic semantic analysis. For example, the American electronics corporation IBM,

one of the world's largest manufacturers of computers and software, uses simplified/technical English to improve the quality of technical documentation and, where possible, to simplify its (semi-)automatic translation. In controlled language, sentences are short and grammatically simple, with a preference for nouns over pronouns, active syntactic constructions over passive ones, etc.

Plain language is language that is easy to understand, clear and concise, avoids overly complex words, does not contain clichés or unnecessary technical or professional jargon, and corresponds to the level of development and education of the audience and their familiarity with the topic. It is a way of communicating so that the target audience understands what they read or hear the first time. Examples include “plain English” and “leichte Sprache” (German for “easy language”). The term “plain English” comes from the idiom “in plain English,” which means “in simple and understandable language.”

Plain English refers to the language requirements for government and business communication. It is used for official and business communication to make government, financial, business, legal, and other documents and publications understandable to ordinary people. Flowery language, verbosity, and imagery are not only inappropriate in formal business communication, but they can also be confusing and misleading. In the UK, these issues were criticized in George Orwell's essay “Politics and the English Language” (1946), which was followed by a series of books with recommendations for officials on how to avoid pompous and overly complex writing. Now all government, official, emergency, and other organizations are required to use plain/simple English, as complicated expressions can lead to misunderstandings and even deaths.

Similar trends have emerged in Ireland, the US, and other countries, affecting not only the government and legal sectors, but also transportation, public services, and health-care. For example, the official website of the US government contains information about the Plain Language Action and Information Network, an unfunded working group of federal employees from various agencies and specialties who support the use of clear communication in government documents [6]. The group emphasizes that plain language helps to avoid unnecessary responses to phone calls, writing explanatory and clarifying letters and documents, court proceedings, etc. [8]. The plain language initiative is also supported by the National Institutes of Health (NIH).

German ‘Leichte Sprache’ (easy language) is a special (usually written) variant of German. It is aimed at people who have a low level of German or reading skills in general. The rules were published by the German association Netzwerk Leichte Sprache [5] and cover sentence structure and word choice, as well as recommendations for typography and media use. A similar concept, German ‘Einfache Sprache’ (simple language), is aimed at people with cognitive impairments and those whose native language is not German.

Similar trends can be observed in many other countries: in France, in 2002, the Constitutional Council adopted a decision that defines “clarity and comprehensibility” of French legislation as one of the tasks of the constitution; in 2013, Israel's legislative body, the Knesset, adopted rules on the accessibility of services, which require the use of natural simplified language (https://www.gov.il/he/departments/ministry_of_justice/govil-landing-page); The International Plain Language Association was established (<https://plainlanguagenetwork.org/plain-language/what-is-plain-language/>). The Department of Public Works and Government Services of Canada has similar recommendations. In order to make information easy to read and understand, representatives of eight European

countries have developed European Union standards aimed at people with intellectual disabilities (in accordance with Article 9 of the EU Convention) and those for whom English is not their native language [4]. Nowadays, the ideas of simple, understandable, and accessible language are also being fruitfully applied in marketing to effectively convey information about a product, for example, in product descriptions on websites, advertising texts, calls to action, etc.

Survival vocabulary refers to the most basic words and expressions that are necessary to know in a foreign language environment, such as “good day,” “thank you,” “entrance,” etc. The term is used mainly in Western lexicography. The idea of creating such a list of words belongs to T. Corlett, who in 1963 published a list of “essential vocabulary” words. A distinctive feature of this type of vocabulary is that it is used by people to overcome communication difficulties in a foreign language environment, rather than for the purpose of learning the language. From this point of view, it is similar to a phrasebook, which simulates verbal situations, for example, at the airport, at the train station, in a restaurant, at a gas station, in a hotel, on an excursion, etc.

Vocabulary for special purposes is a reduced dictionary that presents the vocabulary of a specific field of knowledge, such as accounting or economics. The term is used mainly in Western lexicography. It differs in the degree of generalization of the lexical material: the vocabulary for special purposes is based on one of the functional styles, and the survival vocabulary is based on an analysis of words used in the most typical communication situations.

The **lexical base of language** is a compressed dictionary created by extracting vocabulary that is highly frequent in the main functional styles of language. As a rule, this is accompanied by a systemic structural analysis and thematic classification of this vocabulary. It has also been noted that the most frequent words are function words or general abstract concepts. On the other hand, words with specific meanings (necessary for conversation in everyday situations) are low-frequency. According to French linguist R. Michea, although they are rarely used, they are always present in the speaker’s mind. To enrich the vocabulary with this layer of lexicon, J. Gugeneim suggested that primary school students write down words associated with sixteen specific topics (parts of the body, clothing, home, food, etc.) in order to enrich their vocabulary with specific lexicon. In other words, the frequency criterion is supplemented by the thematic criterion.

3. The **Lexical base of Ukrainian Language (LBUL)** [12] was built on a 1.7-million-word balanced corpus representing five functional styles of modern Ukrainian: colloquial, belles-lettres, journalistic, scientific, and official [9, 10, 14, 15, 17]. From each, 3,000 most frequent words were compared – covering from 77.9% to 86.1% of text – yielding 1,386 lexical units constituting the Ukrainian core. The LBUL units were grouped into 16 semantic fields, including *Nature, Human, Society, Object, Abstract Relations, Values etc.* When compared with multilingual Routledge frequency dictionaries, these fields overlap significantly, showing that the same conceptual domains form the core lexicon across languages. Here, it is worth mentioning the interesting study by O. Maksymiv on a partial comparison of the vocabulary of these dictionaries [16].

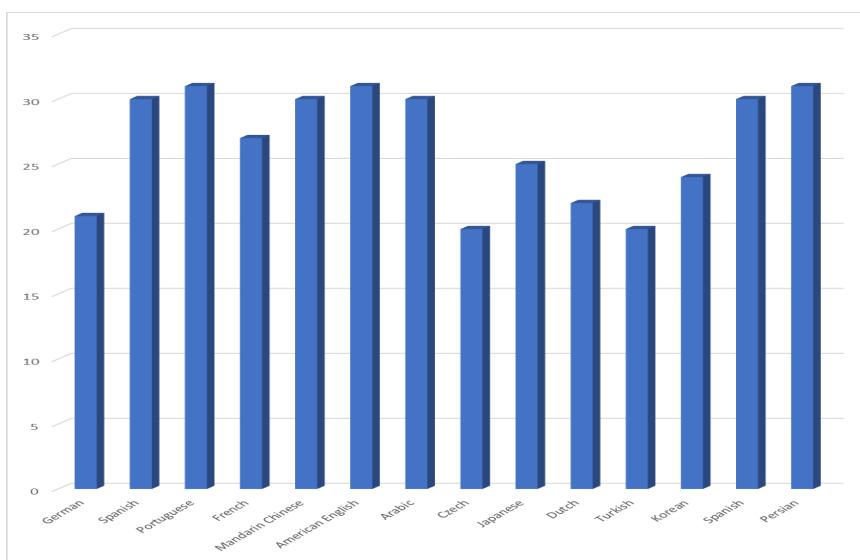


Chart 1. Amount of topics covered in the frequency dictionaries, Routledge 2006-2018
(see Appendix)

LBUL was compiled well before both the *Routledge Frequency Dictionary* series (2006–2018) (see Appendix) and the large-scale *Frequency Dictionaries Collection* of the Leipzig University (2011–2020) [1]. Despite being developed independently in 2004, the LBUL employed nearly identical principles of corpus selection and frequency analysis, relying on a balanced, multi-genre corpus and rigorous lemmatization and statistical procedures – methods that would later become standard in international corpus linguistics. This chronological precedence highlights Ukraine's early contribution to the evolution of quantitative and applied lexicography in Europe.

- Shared Lexical Core: Over 90% of LBUL vocabulary corresponds to universal semantic fields found in other languages. For example, see the amount of topics covered in the frequency dictionaries, Routledge 2006-2018 (Chart 1).
- Limited Cultural Specificity: Only 5–7% of lexemes are culturally bound or lack direct equivalents (e.g., *watch/clock*, *arm/hand* etc.).
- Pedagogical Implication: Teaching Ukrainian through this Common Framework Vocabulary allows learners to map Ukrainian words onto familiar cognitive and semantic structures.
- Cross-Linguistic Transfer: The universality of lexical domains facilitates transfer and comprehension among multilingual learners.

The LBUL provides a systematized model of this universal layer of vocabulary and demonstrates remarkable cross-linguistic similarity to the Routledge *Frequency Dictionaries* (German, Spanish, Turkish, Czech, French, etc.).

The LBUL represents a scientifically grounded core of universal Ukrainian vocabulary, suitable for multilingual education. The comparison with Routledge frequency dictionaries reveals a shared cross-linguistic structure of human conceptualization. This Common

Lexical Framework promotes understanding, intercultural dialogue, and inclusive learning in multilingual environments. The universality of lexical domains offers a practical and theoretical foundation for teaching Ukrainian as part of Europe's multilingual mosaic.

Integrating the Ukrainian lexical base into multilingual curricula and digital teaching tools (such as frequency-based apps or AI-driven vocabulary builders) supports the development of a European Common Lexical Framework—a step beyond the Common European Framework of Reference (CEFR), oriented toward vocabulary-level interoperability.

Conclusions and future research. This study has systematically reviewed the historical development and diverse contemporary typology of quantitatively restricted vocabulary systems, demonstrating that the purposeful limitation of lexicon serves crucial functions in both foreign language pedagogy and enhanced domestic communication accessibility. Key Findings were obtained in several dimensions: methodological, taxonomical, and cross-linguistic universality.

Diverse Methodologies: The historical analysis highlights a fundamental divergence in principles: early simplified dictionaries (Thorndike) were based primarily on frequency, while systems like C.K. Ogden's Basic English were constructed on conceptual and communicational sufficiency.

Terminological Taxonomy: Concepts such as Controlled Natural Language (CNL), Plain Language (Plain English, Leichte Sprache), Survival Vocabulary, and Vocabulary for Special Purposes each represent distinct applications of lexical restriction, varying in their target audience (learners vs. native speakers with low proficiency), purpose (ambiguity reduction vs. accessibility), and underlying rules (grammatical constraints vs. word choice). The global move towards Plain Language underscores its critical role in government and business communication to prevent misunderstandings and promote clarity, as evidenced by initiatives across the UK, US, and the EU.

Role of LBUL: The analysis of LBUL confirms its methodological validity and chronological precedence in European corpus linguistics. Developed independently in 2004, the LBUL employed nearly identical principles of balanced corpus selection and rigorous statistical analysis that later became standard in international collections, such as the Routledge Frequency Dictionaries (2006–2018). In the multilingual classroom, highlighting differences among languages is necessary for cultural depth, yet building on lexical universals is crucial for comprehension, motivation, and confidence. The LBUL provides a functional bridge between Ukrainian and other European languages, reinforcing the idea that mutual understanding starts with shared meaning fields, not with contrastive exceptions.

Cross-Linguistic Universality: The LBUL exhibits remarkable cross-linguistic consistency, with over 90% of its vocabulary corresponding to universal semantic fields (e.g., Nature, Human, Society) found in other major language corpora. This reinforces the hypothesis that a shared cognitive and conceptual framework underlies the core lexicon across disparate languages.

Future research should focus on a direct, quantitative comparison of the LBUL's semantic field distribution with the most recent large-scale frequency dictionaries (e.g., Leipzig University Collection) to further solidify its cross-linguistic validity. For a broader perspective, other similar series should be included in the study: Oxford Word Lists (Oxford University Press), Collins COBUILD Word Frequency Series, Cambridge English Vocabulary Wordlists, Le Monde Frequency Lists, SUBTLEX Series, Berlitz Phrasebooks, Langenscheidt Sprachführer / Grundwortschatz, etc. Additionally, an empirical

study assessing the pedagogical efficacy of using the LBUL as a “Common Framework Vocabulary” for teaching Ukrainian to multilingual students would provide valuable insights into optimizing cross-linguistic transfer and learning outcomes.

For languages with limited global presence, including Ukrainian, the question of defining an evidence-based, functionally universal lexical basis has both scientific and practical urgency. On the one hand, the growth of digital communication, multilingual information environments, and adaptive educational technologies requires clear and empirically grounded lexical standards. On the other hand, the increasing prevalence of simplified language models – such as plain language, easy language, and controlled language – creates new challenges for maintaining linguistic richness while ensuring communicative accessibility. Thus, the problem of determining a corpus-based, pedagogically validated lexical foundation for teaching UFL intersects with broader tasks of digital linguodidactics, cognitive optimization, and human–AI interaction.

References

1. Babelyuk O. Linguistic Evolution In The Digital Age: English Cyber Neologisms Word-Formation Processes. *Nova Filologia*. № 97 (2025). P. 5–10.
2. Buk S., Babelyuk O. Linguistic myths in Russian-Ukrainian war: Changes and challenges. *Baltic Rim Economies*. Special issue on Ukraine’s integration to west. February 2024. No. 1. P. 63–64.
3. Collection of frequency dictionaries of various languages from the University of Leipzig Leipziger Universitätsverlag und AVA - Akademische Verlagsanstalt 2011–2020. Rezhym dostupu : https://www.univerlag-leipzig.de/catalog/bookstore/category/142-Frequency_Dictionaries_Haeufigkeitswoerterbuecher. (Date of access 01.12.2025).
4. European standards for making information easy to read and understand. Rezhym dostupu : https://www.inclusion-europe.eu/wp-content/uploads/2017/06/EN_Information_for_all.pdf. (Date of access 01.12.2025).
5. Netzwerk Leichte Sprache. Online verfügbar unter. Rezhym dostupu : www.leichtesprache.org. (Date of access 01.12.2025).
6. Plain Language Action and Information Network. Rezhym dostupu : <https://www.nih.gov/institutes-nih/nih-office-director/office-communications-public-liaison/clear-communication/plain-language>. (Date of access 01.12.2025).
7. Plain Language. URL : <https://www.plainlanguage.gov>. (Date of access 01.12.2025).
8. Voice of America news special English. Rezhym dostupu : <https://web.archive.org/web/20010902230430/http://www.voanews.com/specialenglish/> (Date of access 01.12.2025).
9. Buk S. 3 000 naichastotnishykh sliv rozmovno-pobutovoho styliu suchasnoi ukrainskoi movy. Lviv : LNU imeni Ivana Franka, 2006. 180 s.
10. Buk S. 3 000 naichastotnishykh sliv naukovoho styliu suchasnoi ukrainskoi movy. Lviv : LNU imeni Ivana Franka, 2006. 192 s.
11. Buk S. Lehka nimetska, prosta anhliiska, leksychna osnova ukrainskoi movy ta inshi typy sproshchenykh slovnykiv. *Jahrbuch der XV. Internationalen virtuellen Konferenz der Ukrainistik “Dialog der Sprachen - Dialog der Kulturen. Die Ukraine aus globaler Sicht”*. Reihe: *Internationale virtuelle Konferenz der Ukrainistik*. Bd. 2024. (Hgg.) Novikova, O., Schweier, U. Verlag Buchschmiede von Dataform Media GmbH Wien, UB LMU München. 2025. S. 47–57.
12. Buk S. N. Leksychna osnova ukrainskoi movy: vydilennia ta systemno-strukturna orhanizatsiia. Rukopys. Dys... kand. filol. nauk. Lvivskyi natsionalnyi universytet imeni Ivana Franka. Lviv, 2004. 281 s. Rezhym dostupu : <https://www.academia.edu/102016180> (Data zvernennia 10.12.2025).
13. Buk S. Transformatsiia motyvatsii vyvchennia ukrainskoi movy yak inozemnoi u zviaz-

ku z viinoiu. *Teoriia i praktyka vykladannia ukrainskoi movy yak inozemnoi*. 2025. № 19. S. 63–77.

14. Buk S. Chastotnyi slovnyk ofitsiino-dilovoho stylu: pryntsypy ukladannia ta statystychni kharakterystyky. *Linhvistichni studii*. 2006. Vyp. 14. S. 184–188.

15. Darchuk N. Chastotnyi slovnyk suchasnoi ukrainskoi publitsystyky / Ukl. Darchuk N. P. ta in. 1996. Rezhym dostupu : <http://www.mova.info/freqcard.aspx?l1=98&sl=publicist>. (Data os-tannoho dostupu 20.08.2024).

16. Maksymiv O. Sproba poshyrennia zastosuvannia metateoretychnykh pryntsypiv zistav-lemnia danykh chastotnykh slovnykiv na material 14 mov. *Materialy pershoi mizhnarodnoi naukovoi konferentsii pamyati Tyschenka*. Kyiv. 2024. S. 31–36.

17. Chastotnyi slovnyk suchasnoi ukrainskoi khudozhnoi prozy / Za red. Perebyinis V. C. K. : Naukova dumka, 1981. T. 1, 2.

Список використаної літератури

1. Babelyuk O. Linguistic Evolution In The Digital Age: English Cyber Neologisms Word-Formation Processes. *Nova Filologia*. № 97 (2025). P. 5–10.
2. Buk S., Babelyuk O. Linguistic myths in Russian-Ukrainian war: Changes and challenges. *Baltic Rim Economies. Special issue on Ukraine's integration to west*. February 2024. No. 1. P. 63–64.
3. Collection of frequency dictionaries of various languages from the University of Leipzig Leipziger Universitätsverlag und AVA - Academische Verlagsanstalt 2011–2020. Режим доступу : https://www.univerlag-leipzig.de/catalog/bookstore/category/142-Frequency_Dictionaries_Haeufigkeitswoerterbuecher. (Date of access 01.12.2025).
4. European standards for making information easy to read and understand. Режим доступу : https://www.inclusion-europe.eu/wp-content/uploads/2017/06/EN_Information_for_all.pdf. (Date of access 01.12.2025).
5. Netzwerk Leichte Sprache. Online verfügbar unter. Режим доступу : www.leichte-sprache.org. (Date of access 01.12.2025).
6. Plain Language Action and Information Network. Режим доступу : <https://www.nih.gov/institutes-nih/nih-office-director/office-communications-public-liaison/clear-communication/plain-language>. (Date of access 01.12.2025).
7. Plain Language. URL: <https://www.plainlanguage.gov>. (Date of access 01.12.2025).
8. Voice of America news special English. Режим доступу : <https://web.archive.org/web/20010902230430/http://www.voanews.com/specialenglish/> (Date of access 01.12.2025).
9. Buk S. 3 000 naichastotnishykh sliv rozmovno-pobutovoho stylu suchasnoi ukrainskoi movy. Lviv : LNU imeni Ivana Franka, 2006. 180 s.
10. Buk S. 3 000 naichastotnishykh sliv naukovoho stylu suchasnoi ukrainskoi movy. Lviv : LNU imeni Ivana Franka, 2006. 192 s.
11. Buk S. Lehka nimetska, prosta anhliiska, leksychna osnova ukrainskoi movy ta inshi typy sproshchenykh slovnykiv // Jahrbuch der XV. Internationalen virtuellen Konferenz der Ukrainistik "Dialog der Sprachen - Dialog der Kulturen. Die Ukraine aus globaler Sicht". Reihe: Internationale virtuelle Konferenz der Ukrainistik. Bd. 2024. (Hgg.) Novikova, O., Schweier, U. Verlag Buchschmiede von Dataform Media GmbH Wien, UB LMU München. 2025. S. 47–57.
12. Бук С. Н. Лексична основа української мови: виділення та системно-структурна організація. Рукопис. Дис... канд. фіол. наук. Львівський національний університет імені Івана Франка. Львів, 2004. 281 с. Режим доступу : <https://www.academia.edu/102016180> (Дата звернення 10.12.2025).
13. Бук С. Трансформація мотивації вивчення української мови як іноземної у зв'язку з війною. *Teoriia i praktyka vykladannia ukraїnської movy yak inozemnoi*. 2025. № 19. С. 63–77.
14. Бук С. Частотний словник офіційно-ділового стилю: принципи укладання та статистичні характеристики. *Linhvistichni studii*. 2006. Вип. 14. С. 184–188.

15. Дарчук Н. Частотний словник сучасної української публіцистики / Укл. Дарчук Н. П. та ін. 1996. Режим доступу : <http://www.mova.info/freqcard.aspx?l1=98&sl=publicist>. [Дата останнього доступу 20.08.2024].

16. Максимів О. Спроба поширення застосування метатеоретичних принципів зіставлення даних частотних словників на матеріал 14 мов. *Матеріали першої міжнародної наукової конференції пам'яті Тищенка*. Київ, 2024. С. 31–36.

17. Частотний словник сучасної української художньої прози. За ред. Перебийніс В. С. Київ : Наукова думка. 1981. Т. 1, 2.

Appendix: Routledge dictionary series / Додаток: Серія частотних словників Рутледж

1. Aksan Y., Aksan M., Mersinli Ü. and Demirhan U. U. A frequency dictionary of Turkish: Core vocabulary for learners. New York : Routledge, 2017.
2. Buckwalter T. and Parkinson D. L. A frequency dictionary of Arabic: Core vocabulary for learners. New York : Routledge, 2011.
3. Čermák F. and Kren M. A frequency dictionary of Czech: Core vocabulary for learners. New York : Routledge, 2011.
4. Davies M. and Davies K. A frequency dictionary of Spanish: Core vocabulary for learners. 2nd ed. New York : Routledge, 2018.
5. Davies M. and Gardner D. A Frequency Dictionary of Contemporary American English: Word Sketches, Collocates and Thematic Lists. New York : Routledge, 2010.
6. Davies M. and Preto-Bay A. M. R. A frequency dictionary of Portuguese: Core vocabulary for learners. New York : Routledge, 2008.
7. Jones R. L. and Tschorner E. A frequency dictionary of German: Core vocabulary for learners. New York : Routledge, 2006.
8. Lonsdale D. and Le Bras Y. A frequency dictionary of French: Core vocabulary for learners. New York : Routledge, 2009.
9. Miller C. and Aghajanian-Stewart K. A frequency dictionary of Persian: Core vocabulary for learners. New York : Routledge. 2018.
10. Rayson P., and Davies M. Series preface. In: A frequency dictionary of Spanish: Core vocabulary for learners. New York : Routledge, 2008.
11. Tiberius C. and Schoonheim T. A frequency dictionary of Dutch: Core vocabulary for learners. New York : Routledge, 2014.
12. Tono Y., Makawa K. and Yamazaki M. A frequency dictionary of Japanese: Core vocabulary for learners. New York : Routledge, 2013.
13. Xiao R., Rayson P. and McEnery T. A frequency dictionary of Mandarin Chinese: Core vocabulary for learners. New York : Routledge, 2009.
14. Yi S., Chang S., and Sō S. A frequency dictionary of Korean: Core vocabulary for learners. New York : Routledge, 2017.

АРХІТЕКТУРИ ЛЕКСИЧНОГО ОБМЕЖЕННЯ СЛОВНИКА: ТЕНДЕНЦІЇ ТА СУЧASNІ ПРИНЦИПИ В АНГЛІЙСЬКІЙ ТА УКРАЇНСЬКІЙ МОВАХ

Соломія Бук

Львівський національний університет імені Івана Франка
кафедра загального мовознавства
бул. Університетська, 1, кімн. 343, 79001, Львів, Україна
тел.: 032 239 47 56
ел. пошта: solomija@gmail.com
<http://orcid.org/0000-0001-8026-3289>

Оксана Бабельюк

Львівський державний університет безпеки життєдіяльності
бул. Клепарівська, 35, кімн. 426, 79000, Львів, Україна
тел.: 032 239 47 56
ел. пошта: babelyuko@gmail.com
<http://orcid.org/0000-0003-4837-1225>

У цій статті системно досліджено історичну еволюцію, сучасну типологію та основні принципи побудови кількісно обмежених систем словникового запасу в прикладній лінгвістиці та лексикографії.Хоча традиційна лінгвістична думка відстоює лексичне багатство, практичні контексти, зокрема викладання іноземних мов та доступність комунікації, вимагають стратегічного впровадження обмеженої лексики. Дослідження починається з відстеження фундаментальних зусиль щодо визначення основної лексики, що сягають корінням на початок 20 століття. Встановлено історичну відмінність між методологією, що використовувалася ранніми спрощеними словниками, такими як «Словник для вчителя» Ф. Л. Торндейка (1921), який переважно використовував аналіз частоти слів у великих корпусах для покращення оцінки прочитності, та контрастним підходом «Базової англійської мови» К. К. Огдена (1925). Система Огдена навмисно скоротила англійську мову до 850 слів на основі концептуальної достатності — здатності передавати важливі ідеї — а не емпіричної частоти, що привело до появи штучних міжнародних мов та ресурсів, таких як Вікіпедія простою англійською. Далі дослідження розглядає поширене термінологічне розмайття та концептуальні перекріття, властиві цій галузі, надаючи комплексну типологію пов'язаних понять. Аналіз розрізняє: контрольовані природні мови (CNL), просту мову (Plain English / Leichte Sprache), словник для виживання, словник для спеціальних цілей тощо. Ці системи відрізняються принципами відбору, цільовими групами та комунікативними цілями, що відображає складний функціональний ландшафт лексичної редукції в сучасному суспільстві.

Також увагу зосереджено на лексичній основі української мови (ЛОУМ), представляючи детальний аналіз її структури, методології та міжмовного значення. ЛОУМ було створено ще до широкого впровадження міжнародних серій частотних словників (таких як, наприклад, серія Routledge, 2006–2018), опираючись на збалансований корпус обсягом 1,7 мільйона слів, що представляє п'ять функційних стилів сучасної української мови. Вибравши 3000 найчастіших слів з кожного стилю, ЛОУМ виділила ядро з 1386 лексичних одиниць. Цей хронологічний прецедент підкреслює ранній внесок України в кількісну та прикладну лексикографію. Найважливіше те, що порівняльний аналіз показує, що понад 90% словникового запасу ЛОУМ відповідає універсальним семантичним полям (наприклад, Природа, Людина, Суспільство, Абстрактні відносини), виявленим в інших багатомовних частотних словниках (німецькому, іспанському, французькому, турецькому, чеському та інших мовах). Ця сильна структурна конвергенція підтверджує гіпотезу про спільну когнітивну та семантичну основу між мовами та позиціонує ЛОУМ як європейську модель корпусного лексичного універсалізму.

Такий високий ступінь міжмовної кореляції демонструє універсальність базового лексичного шару між мовами, пропонуючи вагомі педагогічні висновки для використання ЛОУМ як системного «СпільногоРамкового словника» для сприяння міжмовному переносу та розумінню серед тих, хто вивчає мову. Результати дослідження свідчать про необхідність єдиного концептуального підходу до обмежено-го словникового запасу, підкреслюючи поєднання частоти, тематичного охоплення та доступності при відборі лексичної основи. Зрештою, у статті пропонується уточнена концептуальна таксономія обмежених лексичних систем та просувається поняття Загальної лексичної рамки (Common Lexical Framework) – доповнення до CEFR на рівні словникового запасу, що ґрунтуються на корпусно-орієнтованих, міжлінгвістичних принципах. Результати дослідження відкривають шляхи для майбутніх досліджень, що включають кількісне вирівнювання семантичних полів, інтеграцію в цифрові навчальні платформи та ширші порівняльні дослідження різних лінгвістичних типологій.

Ключові слова: українська мова як іноземна, лексична основа, редукція словника, частотні списки, моделювання мови, оптимізація словникового запасу, прикладна лінгвістика, квантитативна лінгвістика, лексикографія.

Стаття надійшла до редакції 17.01.2026
доопрацьована 25.01.2026
прийнята до друку 30.01.2026