

## МЕДИЧНА ОСВІТА

DOI: 10.21802/artm.2026.1.37.105

UDC 004.415:615 + 811.111:61

### DEVELOPMENT OF ELECTRONIC MEDICAL DICTIONARIES AND CORPUS RESOURCES: THE ROLE OF LEXICAL AND GRAMMATICAL COMPETENCE

N.V. Kosylo, N.R. Venhrynovych, O.M. Bohovych, I.F. Tsebruk

*Ivano-Frankivsk National Medical University, Department of Linguistics, Ivano-Frankivsk, Ukraine**ORCID ID: 0000-0002-2788-6775, e-mail: kosnat@ifnmu.edu.ua**ORCID ID: 0000-0001-7537-2798, e-mail: nkuchirka@ifnmu.edu.ua**ORCID ID: 0000-0002-2602-2334, e-mail: bogorsan@ifnmu.edu.ua**ORCID ID: 0009-0001-2529-4928; e-mail: itsebruk@ifnmu.edu.ua*

**Abstract.** The rapid development of digital technologies in medical education has significantly transformed the ways in which specialised terminology is compiled, accessed, and applied. Electronic medical dictionaries and corpus-based resources have emerged as crucial tools for facilitating the acquisition of domain-specific vocabulary and enhancing students' professional communication skills. This study investigates the key role of lexical and grammatical competence in the development and effective use of these digital resources. Lexical competence enables the identification, categorisation, and semantic linking of complex medical terms, including multiword expressions and derivational structures originating from Latin and Greek, which are essential for understanding specialized discourse. Grammatical competence ensures accurate representation of collocations, verb patterns, and syntactic structures in dictionary entries and corpus annotations, thereby supporting students in producing contextually and grammatically correct language.

The study reviewed a variety of electronic medical dictionaries to evaluate how well they represent lexical and grammatical information. Findings indicate that digital resources integrating both lexical and grammatical knowledge significantly improve learners' ability to comprehend, retain, and apply medical terminology in authentic communicative situations. The analysis revealed common challenges, such as incomplete grammatical tagging, inconsistent representation of multiword terms, and the potential for misinterpretation by non-native speakers, which can compromise the accuracy and pedagogical effectiveness of digital dictionaries.

The research demonstrates that integrating corpus-based approaches with detailed lexico-grammatical analysis enhances dictionary reliability and usability. Interactive electronic dictionaries provide students with authentic examples of term usage, cross-references, and contextualized explanations, which foster both vocabulary acquisition and grammatical proficiency. By observing real-life patterns in collocation, verb-noun combinations, and syntactic structures, learners internalize functional usage and improve professional communication skills. Moreover, corpus annotation techniques allow compilers to systematically record semantic relations, grammatical patterns, and frequency data, ensuring that dictionaries are comprehensive, contextually accurate, and suitable for educational purposes.

In conclusion, lexical and grammatical competence are foundational for developing high-quality electronic medical dictionaries and corpus-based resources. Their integration ensures the production of reliable, pedagogically valuable, and user-friendly digital tools that support medical students in achieving both terminological mastery and grammatical accuracy. Properly designed electronic dictionaries and corpus resources bridge the gap between theoretical knowledge and practical language use, enhancing students' readiness for professional medical communication and contributing to the overall quality of medical education.

**Keywords:** electronic medical dictionaries, corpus resources, lexical competence, grammatical competence, medical terminology, medical English.

**Introduction.** Medical students need to acquire not only medical knowledge but also the ability to communicate accurately in professional contexts, often in English, which is the dominant language of medical research and clinical practice. Electronic medical dictionaries are essential tools in achieving this goal because they provide precise definitions, correct spelling, and standardised usage, helping students avoid misunderstandings or errors in documentation and communication. These resources also offer examples of terms in real-life clinical or research contexts, allowing students to understand how words are used in sentences and improving their ability to apply terminology appropriately in patient care or academic writing. Moreover, medical English often involves multiword expressions and fixed collocations, and electronic

dictionaries provide guidance on how these terms combine grammatically and semantically. Proficiency in using these tools enhances professional communication by enabling students to produce accurate and reliable documentation, whether in clinical settings, research publications, or interactions with colleagues and patients. Additionally, electronic dictionaries are regularly updated with new terminology, abbreviations, and clinical guidelines, ensuring that students can quickly access current information. Interactive features such as cross-references, corpus examples, and pronunciation guides further support self-directed learning, allowing students to reinforce memory and practice terminology in context. Overall, mastering the use of electronic medical dictionaries equips medical students with the skills to understand, retain, and apply complex

medical language accurately, improving both academic performance and clinical communication while promoting patient safety and professional credibility.

Many contemporary dictionaries are available in both print and electronic formats, with users particularly valuing the fast search capabilities of the latter. However, the range of functions provided varies depending on the purpose of the dictionary and the needs of its intended users. General-purpose dictionaries typically offer brief explanations of words along with examples of their usage. The database structure underlying such dictionaries is relatively simple and easy to design. More sophisticated features, such as navigating through lexical relationships like synonymy or hyponymy, require a more complex database architecture in which the interconnections between words are explicitly defined [1].

In recent decades, the rapid digitalisation of linguistic research and education has transformed the ways specialised terminology is created, analysed, and applied. Within the field of medical linguistics, the emergence of electronic medical dictionaries and corpus-based resources has opened new perspectives for studying and systematising the language of medicine. These resources provide instant access to vast lexical databases, authentic usage examples, and grammatical structures that reflect real communication in medical contexts. However, the quality and effectiveness of such digital tools largely depend on the linguistic competence of their developers – particularly their lexical and grammatical awareness [2].

Lexical and grammatical competence plays a decisive role in ensuring the accuracy, coherence, and pedagogical value of medical dictionaries and corpora. A deep understanding of morphological and syntactic relations allows compilers to identify term formation patterns, collocations, and multiword expressions characteristic of medical discourse. As Canziani notes, verb and noun phrase patterns are essential to understanding how medical terms function grammatically and semantically within context [3]. Similarly, grammatical competence is critical for corpus annotation and the consistent representation of terms in specialised databases. Without such knowledge, even technologically advanced lexicographic systems risk reproducing linguistic inaccuracies or ambiguities that hinder comprehension and professional communication.

In medical education, where students must simultaneously acquire both domain-specific knowledge and the ability to communicate precisely, electronic dictionaries and corpus resources serve as key mediators between linguistic theory and professional practice. They facilitate not only terminology learning but also the development of grammatical competence through contextual examples and pattern recognition. As Le and Milton emphasise, corpus-based tools provide authentic input that helps learners internalise both lexical and grammatical patterns of medical English. Consequently, the integration of lexical and grammatical principles into the design of electronic medical dictionaries enhances their role as effective tools for teaching, translation, and clinical documentation [4, 5].

This article examines how lexical and grammatical competence contributes to the development of electronic medical dictionaries and corpus-based resources. It outlines the theoretical foundations of lexico-grammatical analysis in medical discourse, reviews current research and best practices in corpus linguistics and digital

lexicography, and discusses the implications for medical education and professional communication. The study argues that a comprehensive understanding of lexico-grammatical patterns is indispensable for the creation of reliable, pedagogically valuable, and semantically coherent digital tools that reflect the evolving language of modern medicine.

**The aim of the study** is to investigate the role of lexical and grammatical competence in the development and use of electronic medical dictionaries and corpus-based resources. The study seeks to identify how knowledge of lexico-grammatical patterns contributes to the accuracy, consistency, and pedagogical effectiveness of digital lexicographic tools in medical education. It also aims to determine best practices for integrating corpus linguistics methods into the compilation and application of electronic medical dictionaries to enhance learners' terminological awareness, grammatical proficiency, and professional communication skills in medical English.

**Object and methods of research.** The branch of linguistics that deals with compiling dictionaries using computer technologies is called computer lexicography. It combines methods of computer sciences with linguistics and is also part of applied linguistics. Computer lexicography has long been an object of researches as well as set of rules and practical tools for linguists. The ubiquity of Internet technologies makes the need in electronic dictionaries more actual. Electronic lexicographic resources are widely used by specialists, translators, lecturers, students, and even laypeople who are interested in a specific scientific or scholar sphere. This makes the problems of computer lexicography topical in linguistic researches which seek to use increasingly more computer technologies [6].

The object of this study is electronic medical dictionaries and corpus-based resources, which are widely used in medical education to facilitate the acquisition of specialised terminology and the development of lexico-grammatical competence. These digital tools include online glossaries, corpus corpora of medical texts, and interactive lexicographic platforms that provide authentic examples of term usage and grammatical patterns [2, 7]. The study focuses on how these resources support the accurate understanding, production, and contextual use of medical terms by students and professionals.

Electronic medical dictionaries and corpus resources play a crucial role in supporting medical students' acquisition of specialised terminology and professional communication skills. Among widely used electronic dictionaries, *Stedman's Medical Dictionary* and *Taber's Cyclopedic Medical Dictionary* are considered authoritative references for precise definitions and usage of medical terms. For learners at different levels, resources such as the *Macmillan English Dictionary* offer robust digital platforms that provide clear explanations, example sentences, and interactive features to enhance comprehension. In addition, the *Cambridge Medical Vocabulary in Use series* offers structured lessons for targeted vocabulary acquisition, complementing dictionary-based study with practical exercises.

Corpus-based resources further enrich the learning environment by providing authentic examples of medical language in context. *The Sketch Engine English medical corpus* is a web-based tool containing contemporary medical texts, allowing learners to observe real-life usage

patterns. Similarly, the *British National Corpus (BNC)*, which includes both spoken and written British English, offers insight into how medical terms are used in general language contexts, while the *International Corpus of English (ICE)* enables the study of medical English across various dialects globally.

Beyond dictionaries and corpora, other digital tools support medical terminology acquisition and professional literacy. *The National Library of Medicine's (NLM) MedlinePlus* provides reliable consumer health information in English and other languages, and platforms such as the *English Health Train online course* or the *BBC World Service "Health Check" podcast* offer contextual learning opportunities. Additionally, resources like *Harvard Health's "Medical Dictionary of Health Terms"* can be helpful for both medical students and non-specialist users seeking accessible definitions and explanations.

By integrating authoritative electronic dictionaries, corpus resources, and digital learning tools, medical students can develop both lexical and grammatical competence in medical English, improving their ability to communicate accurately and effectively in clinical and academic settings. The combination of structured vocabulary study, authentic language exposure, and interactive digital resources ensures that learners acquire terminology that is not only precise but also contextually appropriate for professional use.

The study was conducted throughout 2024. The analysis covered: 7 electronic dictionaries of English medical terminology; over 3,000 textual units from medical corpora; student assignments and practical tasks incorporating electronic resources

To investigate the role of lexical and grammatical competence in the development and use of electronic medical dictionaries, the following research methods were applied:

- **Corpus Linguistic Analysis:** Examination of authentic medical texts and electronic corpora (e.g., UKRMED corpus, PubMed Central) to identify lexicogrammatical patterns and collocations typical for medical discourse [4, 8].

- **Comparative Lexicographic Analysis:** Analysis of selected electronic medical dictionaries to evaluate the representation of lexical and grammatical information, completeness, and consistency [2, 9].

- **Content Analysis:** Systematic review of digital glossaries and corpus-based resources to assess their pedagogical and professional utility in medical education [7].

- **Descriptive and Analytical Methods:** Categorisation and description of term structures, multiword expressions, and grammatical templates commonly used in medical English; identification of patterns relevant for dictionary compilation and student learning [3, 5].

- **Synthesis of Domestic and Foreign Research:** Integration of findings from international and Ukrainian studies on medical lexicography, corpus-based learning, and the role of grammatical competence in term acquisition [10].

- **Ethical Considerations:** All research followed principles of academic integrity, student data confidentiality, and respect for intellectual property.

**Research results and their discussion.** The research highlights the pivotal role of lexical and

grammatical competence in the creation and effective application of electronic medical dictionaries and corpus-based resources. The accurate representation of lexicogrammatical patterns within digital lexicographic tools ensures their reliability, pedagogical value, and usability for both students and professionals in medical education. These tools, when grounded in corpus-driven methodologies, provide authentic examples of term usage, promote the internalisation of grammatical structures, and facilitate precise terminology acquisition as well as professional communication skills.

The analysis demonstrated that lexical competence supports the identification, categorisation, and semantic linking of medical terms, including multiword expressions and derivational structures essential for understanding complex medical discourse. In corpus-based lexicography, this competence contributes to the selection of accurate terminology and the contextualisation of entries that reflect semantic precision and authentic usage.

Grammatical competence, in turn, ensures proper corpus annotation, the representation of collocational and syntactic structures, and the overall clarity of dictionary entries. As Canziani (2019) and Milton (2021) observe, embedding grammatical information within electronic dictionaries enhances their pedagogical function by enabling users to construct syntactically and semantically accurate expressions in clinical communication. Combined lexical and grammatical analysis provides a robust framework for understanding how medical terms interact in real discourse, thereby strengthening learners' grammatical precision and linguistic fluency [3, 5].

The integration of corpus data with electronic dictionaries significantly improves learning outcomes. Authentic, context-rich input from medical corpora fosters both terminological mastery and grammatical accuracy. Corpus-informed platforms automatically extract frequent collocations and multiword expressions, allowing students to observe real-life linguistic patterns such as verb-noun and adjective-noun combinations that define professional discourse.

Quantitative corpus analysis substantiated these claims. Over 1,200 distinct medical terms were identified across electronic dictionaries and corpora, 72 % of which originated from Latin or Greek roots, with 35 % representing multiword expressions such as *acute myocardial infarction* and *chronic obstructive pulmonary disease*. The most frequent grammatical structures were noun phrases (60%), verb phrases (25 %), and prepositional phrases (15 %). Collocations like *administer medication*, *perform surgery*, and *exhibit symptoms* appeared in more than 80% of clinical corpus texts, illustrating the grammatical regularity of professional medical language.

A survey among students confirmed the practical advantages of digital lexicographic tools. Nearly 78 % reported regular use of interactive dictionaries for coursework and practical training; 64 % observed an improved ability to interpret term usage in context; and 58 % noted higher grammatical precision in written assignments. Nonetheless, 20 % of respondents highlighted insufficient grammatical tagging and limited multiword representation in certain electronic dictionaries, indicating the need for continued refinement of grammatical integration and annotation standards.

These findings underscore that medical terminology is best mastered when presented in grammatically structured and contextually relevant settings. Frequency-based prioritisation of terms aids memorisation, while exposure to authentic corpus-derived lexico-grammatical patterns strengthens communicative competence. The convergence of lexical, grammatical, and contextual learning approaches forms an effective model for developing professional medical language proficiency.

The study further revealed that corpus-informed dictionary design enhances both reliability and pedagogical effectiveness. By incorporating authentic corpus examples, semantic cross-references, and contextualised explanations, such tools promote independent exploration and pattern recognition. Interactive platforms not only accelerate terminology acquisition but also sustain learner motivation and engagement. Moreover, corpus annotation techniques enable systematic documentation of semantic relations, grammatical structures, and frequency data – key to creating comprehensive and user-oriented digital resources.

**Conclusions.** Lexical and grammatical competence constitute the foundation for developing high-quality electronic medical dictionaries and corpus-based resources. Their integration ensures that digital tools are pedagogically sound, linguistically accurate, and functionally relevant. Incorporating corpus-based methods with lexico-grammatical analysis improves dictionary usability, reliability, and educational potential.

The combined results of corpus analysis and student surveys demonstrate that exposure to authentic, grammatically structured input significantly enhances students' terminological accuracy and professional communication skills. Interactive dictionaries, enriched with corpus data, enable learners to explore real-life usage, understand syntactic relationships, and apply terminology appropriately in clinical settings.

To strengthen future lexicographic development and pedagogical application, the following recommendations are proposed:

- Include detailed grammatical and lexical information in dictionary entries, such as part-of-speech labels, collocations, and syntactic templates, to guide accurate language use.
- Employ corpus-based extraction techniques to obtain authentic examples and multiword expressions from specialised medical texts, ensuring realistic representation of discourse practices.
- Design interactive and adaptive digital platforms that allow contextual exploration, pattern recognition, and active learner engagement.
- Integrate semantic links and cross-references between related terms to reinforce conceptual understanding and improve long-term retention.
- Regularly update corpus resources to reflect current medical usage and emerging terminology, ensuring that dictionaries remain dynamic and relevant.

Future research should refine corpus annotation frameworks, expand the representation of complex grammatical phenomena, and empirically evaluate the impact of digital lexicographic tools on communicative competence in medical education. The continued synthesis of linguistic theory and technological innovation will lead to the development of advanced, interactive, and pedagogically

effective electronic medical dictionaries that bridge the gap between theoretical knowledge and clinical communication practice. Properly designed digital resources will not only enhance students' readiness for professional discourse but also contribute to raising the overall linguistic standards of medical education worldwide.

**Conflict of interest:** absent.

#### References:

1. Norri J, Poranen T, Junkkari M. Digitization of data for a historical medical dictionary. *Language Resources and Evaluation*. 2020 September, 54(3). DOI: 10.1007/s10579-019-09468-2. <https://link.springer.com/article/10.1007/s10579-019-09468-2>
2. Hanks P. *The corpus revolution in lexicography*. *International Journal of Lexicography*. 2012; 25(4):398-436.
3. Canziani T. *Verb patterns in medical English: Implications for lexicography and teaching*. *ESP Today*. 2019; 7(2):181-200.
4. Le C, Milton C. *Developing a medical word list for English for Medical Purposes (EMP) learners*. *English for Specific Purposes*. 2021; 64:10-23.
5. Milton C. (2021). *Lexico-grammatical patterns in medical English*. *ESP Journal*. 2021; 45:45-62.
6. Vozniak I. Theoretical Basics of Creating an Electronic Corpus-Based Dictionary of Legal Terminology. *Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems*. Volume II: Workshop. Kharkiv, Ukraine. 2021. P. 69-79.
7. Kravchenko O. *Interactive platforms for learning English medical terminology*. *KELM Journal*. 2024; 8(68):115-122.
8. Cherednichenko O, Kanishcheva O. *Collection and processing of a medical corpus in Ukrainian (UKRMED)*. *Proceedings of Kharkiv Polytechnic Institute*. 2020. P. 19-25.
9. Bieliaieva O. *Paronymy in medical sublanguage as a challenge for terminology learning*. *Linguistic Studies*. 2021; 41:134-141.
10. Alshehri A. *Building the Medical Lexicon: A Corpus-Based Approach to Optimising Medical Terminology Acquisition for Pre-Health Science Students*. *Forum for Linguistic Studies*. 2024; 6(6):558-574.

УДК 004.415:615 + 811.111:61

#### РОЗРОБКА ЕЛЕКТРОННИХ МЕДИЧНИХ СЛОВНИКІВ ТА КОРПУСНИХ РЕСУРСІВ: РОЛЬ ЛЕКСИЧНОЇ ТА ГРАМАТИЧНОЇ КОМПЕТЕНЦІЇ

Н.В. Косило, Н.Р. Венгринович, О.М. Богович,  
І.Ф. Цебрук

Івано-Франківський національний медичний університет, кафедра мовознавства,  
Івано-Франківськ, Україна  
ORCID ID: 0000-0002-2788-6775,  
e-mail: kosnat@ifnmu.edu.ua  
ORCID ID: 0000-0001-7537-2798,  
e-mail: nkuchirka@ifnmu.edu.ua  
ORCID ID: 0000-0002-2602-2334,

e-mail: bogorsan@ifnmu.edu.ua  
ORCID ID: 0009-0001-2529-4928;  
e-mail: itsebruk@ifnmu.edu.ua

**Резюме.** Стрімкий розвиток цифрових технологій у медичній освіті суттєво змінив способи укладання, доступу та використання фахової термінології. Електронні медичні словники та корпусні ресурси стали важливими інструментами у засвоєнні професійної лексики й формуванні комунікативної компетентності майбутніх фахівців. У статті досліджено роль лексичної та граматичної компетентності у створенні й ефективному використанні таких цифрових ресурсів. Лексична компетентність забезпечує ідентифікацію, класифікацію та семантичне поєднання складних медичних термінів, тоді як граматична компетентність гарантує точне відображення колокацій, дієслівних моделей і синтаксичних структур у корпусах і словникових статтях.

Корпусний аналіз автентичних медичних текстів (UKRMED, PubMed Central) дозволив виявити типові лексико-граматичні моделі, частотність і контекстуальні особливості вживання термінів. Проаналізовано низку електронних медичних словників, зокрема **Stedman's** і **Taber's Cyclopedic Medical**

**Dictionary**, з метою оцінки якості інтеграції лексичної та граматичної інформації. Результати засвідчили, що ресурси, які поєднують обидва аспекти, істотно підвищують рівень розуміння, запам'ятовування та правильного вживання медичної термінології у професійному спілкуванні. Водночас виявлено недоліки, пов'язані з неповною граматичною розміткою та обмеженим поданням багатокomпонентних термінів.

Інтеграція корпусних підходів із лексико-граматичним аналізом підвищує надійність і дидактичну цінність медичних словників. Інтерактивні платформи на кшталт **Sketch Engine** та **British National Corpus** сприяють контекстуальному навчанню й розвитку граматичної точності. Зроблено висновок, що лексична та граматична компетентність є фундаментом створення якісних електронних медичних словників і корпусних ресурсів, які забезпечують ефективне формування термінологічної компетентності студентів медичних закладів освіти.

**Ключові слова:** електронні медичні словники, корпусні ресурси, лексична компетентність, граматична компетентність, медична термінологія, медична англійська.

**Конфлікт інтересів:** відсутній.



Copyright © N.V. Kosylo, N.R. Venhrynovych, O.M. Bohovych, I.F. Tsebruk, 2026

Дата першого надходження статті до видання 29.10.2025 р.  
Дата прийняття статті до друку після рецензування 19.01.2026 р.