

СТРУКТУРНА, ПРИКЛАДНА ТА МАТЕМАТИЧНА ЛІНГВІСТИКА

UDC 81'33

DOI <https://doi.org/10.32782/2710-4656/2024.3.1/24>**Harbera I. V.**

Vasyl' Stus Donetsk National University

Непывода І. В.

Vasyl' Stus Donetsk National University

MACRO- AND MICROSTRUCTURE OF THE DICTIONARY OF APPLIED LINGUISTICS TERMINOLOGY

The article is devoted to the creation of applied linguistics terms dictionary consisting of ten sections (applied phonetics, applied lexicography, natural language processing, applied morphology, corpus linguistics, applied syntax, quantitative linguistics, linguistic expertise, computational linguistics and applied linguistics itself). In the era of rapid development of linguistic sciences, changes in linguistic paradigms, research of practical aspects and a significant increase in theoretical developments, the number of terminology has also increased, which has led to a rethinking of the existing terminological apparatus. In this regard, we were faced with the task of unifying terms, and accordingly, our main goal was to collect and systematise the terms of the fields of applied linguistics. As a part of the research, the scientific literature of applied linguistics was studied. The compositional components of the dictionary, namely its macro- and microstructure, were determined. The article considers the concept of «dictionary article», which is a source of information that reveals the essential characteristics of a word, its contextual, paradigmatic and syntagmatic relations, stylistic and functional features, as well as derivational potential. The types and samples of the compiled dictionary entries are presented, the peculiarities of synonymy and homonymy representation within the dictionary articles are emphasised, and the dictionary entries of variant terms are demonstrated. The requirements for compiling an electronic terminology dictionary are presented. A general model of micro-level organisation of an electronic explanatory and translation dictionary of applied linguistics terms is proposed. Prospects for further scientific research, which are seen in the expansion and systematisation of the electronic dictionary of applied linguistics terms, are indicated.

Key words: applied linguistics, lexicography, macrostructure, microstructure, dictionary, dictionary article.

Statement of the problem and its relevance.

Under condition of intensive development of applied linguistics, independent directions were formed within it, forming their own conceptual and terminological apparatuses. Regarding this matter, there is an urgent need for the formation of an interpretative and translational dictionary of applied linguistics not only for the wide use of newly formed concepts, but also for specifying their semantics and avoiding scientific contradictions in the future. The purpose of the dictionary is to accurately and comprehensively reflect the conceptual and terminological apparatus of ten branches of applied linguistics (applied phonetics, applied lexicography, natural

language processing (NLP), applied morphology, corpus linguistics, applied syntax, quantitative linguistics, linguistic examination, computational linguistics and applied linguistics), as well as to ensure the educational, scientific and production activities of linguists, since the terminological dictionary is, actually, the most important way of organizing and presentation the terminology of the particular field. This, in fact, determines the urgency of developing outlined problem.

The purpose of this study is to collect and systematize terminology of applied linguistics' branches.

To achieve this goal, the following tasks were completed:

1) scientific literature on the research problem was processed;

2) a register of ten sections (applied phonetics, applied lexicography, natural language processing (NLP), applied morphology, corpus linguistics, applied syntax, quantitative linguistics, linguistic examination, computational linguistics and applied linguistics) was compiled;

3) the macrostructure of the dictionary was developed;

4) the structure of the dictionary article was proposed;

5) dictionary articles were concluded.

Analysis of recent research and publications.

Many modern researchers (V.V. Dubichynskyi, L.O. Symonenko, E.A. Karpilovska, T.R. Kiyak) understand the concept of “terminography” not only as a theoretical science but also as a practical activity of compiling dictionaries. Accordingly, lexicographers are trying to develop a certain methodology for unifying dictionary entries, including a set of additional characteristics to fully reflect the linguistic picture. However, despite the fact that well-known linguists (E.A. Karpilovskaya, L.S. Palamarchuk, P.Y. Goretsky, N.F. Klimenko) have already described the theory of dictionary classification in their works, a unified typology has not yet been developed in terminology.

Traditionally, the structure of a dictionary comprises two main parts: macrostructure and microstructure. Each of them has a certain set of attributes and characteristics that are taken into account when compiling a dictionary, but each dictionary differs in its arrangement at both the macro and microstructure levels according to the author’s concept.

The concept of “macrostructure” is most often interpreted as the principle of organisation of dictionary entries in the main body of a dictionary, its external structure. Sergei Grinev adds to this definition the composition and interaction of all the composite parts of the dictionary. Thus, the term “macrostructure” is usually understood as the general ordering of a dictionary, its content and arrangement of dictionary entries [27: 41].

Orysia Demska notes that the dictionary macrostructure can be complete or truncated. According to her research, a complete macrostructure includes the dictionary itself; an introduction, preface; a theoretical section that briefly introduces the reader to the theory of lexicographic description of vocabulary items; rules for using the dictionary; grammatical reference, which is inherent in primarily explanatory and translation dictionaries; a list of abbreviations

and sources. Orysia Demska notes that a complete macrostructure is typical for large academic dictionaries. At the same time, the truncated (partial) macrostructure includes an incomplete set of these structural parts and is usually inherent in small reference books, translation dictionaries, and school editions [3: 102].

The term “microstructure” refers to the internal structure of the lexicographic description of a dictionary unit (article). In fact, the macrostructure is the dictionary itself, and the microstructure, or dictionary entry, is the main structural component of the dictionary [3: 102].

Outline of the main material of the study. One of the most important stages in the compilation of any dictionary is the elaboration of the concept, which involves defining the purpose, target audience, and principles of the dictionary. Our dictionary is intended for linguists, IT specialists, and anyone interested in applied linguistics.

The dictionary of applied linguistics terms is a work of explanatory and translation type that contains the terms themselves with their interpretation and an indication of the source from which the term and interpretation were derived. It also contains English equivalents and synonyms if they are available.

The sources of the search were Ukrainian and English-language scientific works devoted to the problems of applied linguistics.

The register actually defines the field of description of the dictionary, and in accordance with this statement, we have identified the register and the principles of its compilation as the defining components of the dictionary being created.

The register of the dictionary will include 10 thematic sections (applied phonetics, applied lexicography, natural language processing, applied morphology, corpus linguistics, applied syntax, quantitative linguistics, linguistic expertise, computational linguistics and applied linguistics), each of which will contain from 150 to 200 dictionary entries. As a result, the register of the dictionary of applied linguistics will contain about 2000 vocabulary items. The register units, arranged according to the traditional alphabetical principle, are single-word (simple) terms and phrase terms (compound, multi-word).

The vast majority of the register is composed of compound terms. The number of single-word terms is much smaller.

Here are samples of the concluded dictionary entries for each section, in the form and languages in which the dictionary is created:

ПРИКЛАДНА ЛІНГВІСТИКА –

- галузь мовознавства, в якій розробляють технології опрацювання мови (традиційні, тобто неавтоматизовані, й автоматизовані) для різних галузей суспільного життя [20];
- напрям у мовознавстві, який опрацьовує методи розв'язання практичних завдань, пов'язаних із використанням мови, і зорієнтований на задоволення суспільних потреб [15];
- це теоретичні та емпіричні дослідження проблем реального світу, в центрі яких знаходиться мова [32].

Відповідник: applied linguistics.

ПРИКЛАДНА ФОНЕТИКА – галузь загальної фонетики, в якій вивчення звуків та інших мовленнєвих одиниць, явищ і процесів проводять експериментальним шляхом [18].

Відповідник: applied phonetics.

ПРИКЛАДНА ЛЕКСИКОГРАФІЯ – це область прикладної лінгвістики, яка фокусується на укладенні словників (практична лексикографія), а також на описі різних типів зв'язків, наявних у лексиці (теоретична лексикографія) [36].

Відповідник: applied lexicology.

Синонім: комп'ютерна лексикографія.

КОМП'ЮТЕРНА ЛЕКСИКОГРАФІЯ – це галузь лексикографії, яка використовує комп'ютери в лексикографічній практиці [22].

Відповідник: computational lexicology.

Синонім: прикладна лексикографія.

АВТОМАТИЧНЕ ОПРАЦЮВАННЯ ПРИРОДНОЇ МОВИ (АОПМ) – це окремий технологічний напрямок в частині штучного інтелекту, який присвячений тому, як машина розуміє, аналізує та відтворює природну мову, якою людина звертається до неї [29].

Відповідник: Natural Language Processing (NLP).

МОРФОЛОГІЯ – це наука, що займається внутрішньою структурою та зовнішньою формою слів мови, включаючи частини мови та їх категорії [2].

Відповідник: morphology.

КОРПУСНА ЛІНГВІСТИКА –

- технологія добору, опрацювання, дослідження й збереження текстів (усних і писемних, як правило, сучасних) на комп'ютерних носіях інформації, яка дає змогу отримувати лінгвістичні характеристики вжитих у цих текстах слів, словосполучень, речень тощо [20];

- дисципліна в межах комп'ютерної лінгвістики, предметом досліджень якої є розроблення поняттєвого та процедурного апарату для форму-

вання корпусів текстів (текстозорієнтованих баз даних) та їх аналізу [11].

Відповідник: corpus linguistics.

СИНТАКСИС –

- це область аналізу мови, відповідальна за аналіз утворення речень з урахуванням слів мови [29];

- сфера граматичного ладу мови, що охоплює характерні для конкретних мов засоби і правила побудови речень, висловлень, текстів [8];

- наука, що вивчає правила структурування речень, висловлень, текстів [8].

Відповідник: syntax.

КВАНТИТАТИВНА ЛІНГВІСТИКА – напрям лінгвістики, що вивчає частотні характеристики мови – слів, їх комбінацій, синтаксичних конструкцій за допомогою методів «кількісної» математики (використовуються комбінаторика, теорія ймовірностей, математична статистика, теорія інформації) [21].

Відповідник: quantitative linguistics.

ЛІНГВІСТИЧНА ЕКСПЕРТИЗА –

- це процесуальна дія, призначена для встановлення фактів з метою вирішення судової справи (таким чином, це вид діяльності, що регулюється процесуальними нормами, предметом регулювання яких є суспільні відносини) [1];

- це дослідження, що дозволяє або не дозволяє встановити ці факти (таким чином, у ньому немає нічого юридичного, у ньому використовуються наукові теорії і методики, що описують фрагменти реальності) [1].

Відповідник: linguistic examination.

КОМП'ЮТЕРНА ЛІНГВІСТИКА –

- ділянка прикладної та експериментальної лінгвістики, в якій одиниці мови та лінгвістичні процеси моделюють на комп'ютері з метою створення комп'ютерних лінгвістичних систем [20];

- це широка царина використання комп'ютерних інструментів – програм, комп'ютерних технологій організації та оброблення даних – для моделювання функціонування мови в тих чи інших умовах, ситуаціях, проблемних галузях [11];

- сфера застосування комп'ютерних моделей мови не лише в лінгвістиці, а й у суміжних з нею дисциплінах [11];

- лінгвістична дисципліна, яка розв'язує теоретичні й прикладні завдання мовознавства за допомогою комп'ютера [11].

Відповідник: computational linguistics.

The structure of each dictionary entry includes the heading in its original form, its definition, references

to scientific literature, the English equivalent and synonyms.

In the register of our explanatory and translation dictionary, the heading of each article is printed on a new line in capital letters in bold. The English equivalent is given on a new line and in italics.

АВТОМАТИЧНЕ РОЗПІЗНАВАННЯ МОВЛЕННЯ – процес транскрипції з мови в текст: перетворення акустичного сигналу в послідовність слів, що не обов'язково передбачає розуміння значення тексту або намірів сказаного [29].

Відповідник: automatic speech recognition.

Синонім: автоматичний аналіз мовлення.

АВТОМАТИЧНЕ РОЗПІЗНАВАННЯ ТЕРМІНІВ – інструмент, що виділяє з тексту та автоматичного пошуку у базі всі терміни, що містяться у вихідному текстовому сегменті, над яким працює перекладач у цей час [25].

Відповідник: automatic recognition of terms.

Синонім: автоматичний аналіз термінів.

Homonyms are superscripted and have separate dictionary entries.

МОВА¹ – будь-яка підмножина множини рядків латинського алфавіту [29].

Відповідник: language.

МОВА² – це набір рядків, які є результатами похідних дерев із S як кореневою міткою та термінальними символами як мітками листкових вузлів [29].

Відповідник: language.

Variants of terms have a common core word and capture a set of meanings within a single vocabulary unit.

ДИНАМІЧНА СТРУКТУРА ДАНИХ –

- послідовність даних, зв'язана відносинами лінійного списку [26];

- послідовність комірок в пам'яті, де кожний наступний елемент розташований в пам'яті в наступній комірці [26].

Відповідник: dynamic data structure.

РЯДОК –

- це скінченна послідовність букв [29];

- це будь-яка послідовність буквено-цифрових символів (букви, цифри, пробіли, табуляції та розділові знаки) [31].

Відповідник: string.

Conclusions and prospects for further research.

Thus, the organisation of the macro- and microstructure of a terminological dictionary is the result of in-depth theoretical analysis and comprehension. Compiling a dictionary entry requires a thorough study of the peculiarities of using the relevant terminological unit, as well as repeated verification, clarification and supplementation of its content. A prerequisite is full compliance with the principles of building a dictionary entry, which makes the dictionary user-friendly.

On the basis of all these principles and features of construction, which will contribute to the perfect presentation of information and significantly increase the information content and value of a lexicographic publication, the macro- and microstructure of a terminological dictionary of applied linguistics has been developed.

The prospect of further research will be the compilation of an electronic dictionary of applied linguistics terms.

Bibliography:

1. Гарбера І.В. Основи лінгвістичної експертизи тексту. Вінниця, 2020. 220 с.
2. Горпинич В.О. Морфологія української мови. Київ, 2004. 336 с.
3. Демська О. Вступ до лексикографії. Київ, 2010. 266 с.
4. Дубічинський В.В. Лексикографія. Харків, 2011. 66 с.
5. Жуковська В.В. Вступ до корпусної лінгвістики. Житомир, 2013. 142 с.
6. Загнітко А. та ін. Парадигмально-категорійні основи прикладної лінгвістики. Вінниця, 2015. 472 с.
7. Загнітко А. Теорія лінгвоперсонології. Вінниця, 2017. 136 с.
8. Загнітко А.П. Теорія сучасного синтаксису. Донецьк, 2007. 294 с.
9. Ільман В.М. та ін. Алгоритми, дані і структури. Дніпро, 2019. 134 с.
10. Карпенко Ю.О. Вступ до мовознавства. Київ, 2006. 336 с.
11. Карпіловська Є.А. Вступ до прикладної лінгвістики: комп'ютерна лінгвістика. Донецьк, 2006. 188 с.
12. Козуб Г.О. Програмування. Старобільськ, 2020. 108 с.
13. Копей В.Б. Мова програмування Python для інженерів і науковців. Івано-Франківськ, 2019. 272 с.
14. Кормен Т. Вступ до алгоритмів. Київ, 2019. 1288 с.
15. Кочерган М.П. Загальне мовознавство. Київ, 2006. 464 с.
16. Кренич А.П. Алгоритми і структури даних. Київ, 2021. 200 с.
17. Мелешко Є.В. та ін. Алгоритми та структури даних. Кропивницький, 2019. 156 с.
18. Мойсієнко А.К. та ін. Сучасна українська літературна мова: Лексикологія. Фонетика. Київ, 2010. 270 с.

19. Муковнін Є.В. Труднощі і роботі систем обробки природної мови та основні методи їх вирішення.
20. Партико З.В. Прикладна і комп'ютерна лінгвістика: Вступ до спеціальності. Львів, 2008. 224 с.
21. Пасічник В.В. та ін. Математична лінгвістика. Львів, 2012. 359 с.
22. Перебийніс В.І., Сорокін В.М. Традиційна та комп'ютерна лексикографія. Київ, 2009. 218 с.
23. Прийма С.М. Теорія алгоритмів. Мелітополь, 2018. 116 с.
24. Ситар Г.В. Синтаксичні фразеологізми в розрізі конструкційної граматики. Вінниця, 2017. 458 с.
25. Тарануха В.Ю. Інтелектуальна обробка текстів: навчальний посібник. Київ, 2014. 80 с.
26. Ткачук В.М. Алгоритми і структура даних. Івано-Франківськ, 2016. 286 с.
27. Туровська Л. В. Макро- та мікроструктура термінологічного словника тлумачного типу. Українська мова. 2018. № 3. С. 40–53.
28. Bot K. History of Applied Linguistics: From 1980 to the Present. 2015. 168 p.
29. Clark A. and oth. The handbook of computational linguistics and natural language processing. 2010. 775 p.
30. Gee J.P. An Introduction to Discourse Analysis. Routledge. 2010. 218 p.
31. Jurafsky D., Martin J. Speech and Language Processing. New Jersey, 1999. 950 p.
32. Littlemore J., Groom N. Doing Applied Linguistics: A Guide for Students. 2012. 205 p.
33. McCarthy M., O'Keeffe A. Routledge Handbook of Corpus Linguistics. 2021. 682 p.
34. McKinley J. and oth. Doing Research in Applied Linguistics. New York, 2016. 262 p.
35. Oakey D., Hunston S. Introducing Applied Linguistics: Concepts and Skills. 2009. 251 p.
36. Simpson J. Routledge Handbook of Applied Linguistics. 2011. 732 p.
37. Viana V. and oth. Perspectives on Corpus Linguistics. Amsterdam, 2011. 256 p.
38. Wei L. Applied Linguistics. 2013. 328 p.
39. Wei L. Contemporary applied linguistics. London, 2009. 277 p.
40. Woodrow L. Writing about Quantitative Research in Applied Linguistics. 2014. 220 p.

Гарбера І. В., Непийвода І. В. МАКРО- І МІКРОСТРУКТУРА СЛОВНИКА ТЕРМІНІВ ПРИКЛАДНОЇ ЛІНГВІСТИКИ

Статтю присвячено створенню словника термінів прикладної лінгвістики, що складається з десяти її розділів (прикладної фонетики, прикладної лексикографії, автоматичного опрацювання природної мови, прикладної морфології, корпусної лінгвістики, прикладного синтаксису, квантитативної лінгвістики, лінгвістичної експертизи, комп'ютерної лінгвістики та власне прикладної лінгвістики). В епоху стрімкого розвитку мовознавчих наук, зміни лінгвістичних парадигм, дослідження практичних сторін та значного збільшення теоретичних напрацювань зростає кількість термінології, що зумовило переосмислення наявного термінологічного апарату. У зв'язку з цим, перед нами постало завдання уніфікації термінів, відповідно нашою основною метою стало зібрання та систематизація термінів галузей прикладної лінгвістики. В межах наукового дослідження було опрацьовано наукову літературу розділів прикладної лінгвістики. Визначено композиційні складові словника, а саме його макро- і мікроструктуру. Розглянуто поняття "словникова стаття", під яким вбачається джерело інформації, що розкриває сутнісні характеристики слова, його контекстні, парадигматичні та синтагматичні зв'язки, стилістичні та функціональні особливості, а також дериваційний потенціал. Наведено типи та зразки укладених словникових статей, наголошено на особливостях представлення синонімії та омонімії в межах словникових статей, а також продемонстровано словникові статті варіантних термінів. Представлено вимоги до укладання електронного термінологічного словника. Запропоновано загальну модель мікрорівневої організації електронного тлумачно-перекладного словника термінів прикладної лінгвістики. Зазначено перспективи подальших наукових досліджень, які вбачаються в розширенні та систематизації електронного словника термінів прикладної лінгвістики.

Ключові слова: *прикладна лінгвістика, лексикографія, макроструктура, мікроструктура, словник, словникова стаття.*