

СТРУКТУРА СЕМАНТИКИ ПРИКМЕТНИКІВ УКРАЇНСЬКОЇ МОВИ В ЛЕКСИКОГРАФІЧНІЙ СИСТЕМІ

*І. В. Олександрук, доктор філософії, молодший науковий співробітник
Інституту мовознавства ім. О. О. Потебні НАН України
(Київ, Україна)*

e-mail: ira_oleksandruk@ukr.net

ORCID: <https://orcid.org/0000-0003-2701-1714>

У статті розглянуто структуру семантики прикметників української мови в лексикографічній системі тлумачного типу. Дослідження проведено на матеріалі електронної форми Словника української мови у 20 томах, який створюють в Українському мовно-інформаційному фонді Національної академії наук України. Працю представлено у вигляді Віртуальної лексикографічної лабораторії «Словник української мови».

Актуальність дослідження зумовлена застосуванням комп'ютерних технологій у дослідженнях із семантики, орієнтованих на створення автоматизованих систем опрацювання мовної інформації. Основним інструментом семантичного аналізу в цих системах є електронні тлумачні словники. Мета статті – проаналізувати структуру семантики відносних прикметників на основі інформаційної теорії лексикографічних систем для можливості подальшого опрацювання програмними засобами. Об'єктом дослідження є словникові статті прикметників української мови.

У розвідці проаналізовано структурне представлення параметрів лівої частини словникової статті, що репрезентує систему граматичної семантики, та правої – лексичної семантики прикметників української мови на засадах інформаційної теорії лексикографічних систем. У межах дослідження розроблено лексикографічну базу даних «Структура семантики прикметників української мови», яка є зручним інструментом для подальшого опрацювання структури прикметникових словникових статей, зокрема моделювання граматичної і лексичної семантики.

***Ключові слова:** електронний тлумачний словник, комп'ютерна лексикографія, семантика, відносні прикметники української мови.*

Актуальність. Актуальність обраної теми визначена недостатньою вивченістю застосування комп'ютерних технологій у дослідженнях, орієнтованих на розроблення автоматизованих систем опрацювання інформації. Важливе значення мають праці, спрямовані на створення систем семантичного автоматизованого аналізу тексту; розроблення методів представлення лексикографічної інформації у структурованій формі, доступній для оперування в комп'ютерах та мережах, а також способів програмного екстрагування інформації з текстів та інших словників з метою подальшого опрацювання.

Аналіз останніх досліджень і публікацій. Дослідженню та різнобічному вивченню українського прикметника, зокрема специфіки семантики, присвячено праці багатьох мовознавців: Д. Х. Баранника, В. М. Васильченка, І. Р. Вихованця, М. П. Кочергана, О. П. Штонь, М. І. Степаненка, І. Ю. Кірічек та ін. Однак ці студії зорієнтовані на теоретичне мовознавство. У працях В. А. Широкова розроблено інформаційну теорію лексикографічних систем, яка надає універсальну теоретичну основу як для традиційної, так і для цифрової лексикографії, а також розглянуто проблеми формалізованого опису структури прикметників української мови в межах інформаційного моделювання проблемно-орієнтованих структур українськомовних лексикографічних систем [3, 4, 5, 8 та ін.].

Теорія лексикографічних систем є узагальненням понять граматичних і лексичних значень, що слугує ефективним інструментальним засобом для провадження семантичних досліджень із використанням комп'ютерних технологій. Як зауважує автор, теорія надала концептуальний апарат і формальні засоби для розгляду лексикографічних структур як

самостійних мовно-інформаційних об'єктів, що допускають абстрагування від конкретних текстових реалізацій у словникових статтях і, представлені у вигляді певних формальних об'єктів, набувають самостійної лінгвістичної інтерпретації та можуть бути застосованими для нових мовознавчих досліджень [3, с. 98].

Лексикографічна система (Л-система) – це інформаційний об'єкт, який поєднує в собі риси моделі даних, моделі знань, логіко-лінгвістичного числення певного типу та є спеціальним інформаційним середовищем, у якому розвивається лексикографічний ефект (або певна їх сукупність) [5, с. 35]. Лексикографічний ефект постає основним системотвірним феноменологічним принципом мови, він відповідає за виявлення в певній ділянці мови комплексів дискретних одиниць, що виконують функцію структурних елементів лексикографування у відповідному словнику. Тому будь-який феномен мови може бути «лексикографований» і, отже, стати об'єктом певного словника [7, с. 5].

Основними системотвірними об'єктами Л-системи є комплекс елементарних інформаційних одиниць системи, а основними системотвірними відношеннями – відношення «суб'єкт – об'єкт» і «форма – зміст».

Мета дослідження – проаналізувати структуру семантики прикметників української мови на засадах інформаційної теорії лексикографічних систем для можливості подальшого опрацювання програмними засобами.

Матеріали і методи дослідження. Матеріал дослідження – словникові статті прикметників, автоматично екстраговані з електронної форми Словника української мови у 20 томах [2], обсягом більше ніж 200 тисяч реєстрових слів. Словник функціонує у двох формах: електронній та паперовій. Електронна форма словника представлена Віртуальною лексикографічною лабораторією «Словник української мови» (ВЛЛ «СУМ»), яка служить інструментальною системою для укладання Словника української мови у 20-ти томах, а також для проведення лінгвістичних досліджень. ВЛЛ «СУМ» є відкритою системою, тому постійно поповнюється новими лексемами.

Для досягнення поставленої мети було застосовано такі **методи дослідження**: описовий, дистрибутивний аналіз (для дослідження особливостей уживання відносних прикметників у сполученні з іменниками), інформаційна теорія лексикографічних систем (для поглибленого дослідження семантичної структури українського прикметника).

Результати дослідження та їх обговорення. Застосування інформаційної теорії лексикографічних систем до інтерпретації прикметникових словникових статей вимагає встановлення формальних репрезентантів їхньої структури. Згідно з теорією, Словник розглядаємо як лексикографічну систему, що має свою структуру та складається з множини словникових статей, виокремлення структурних елементів яких є необхідною умовою для автоматизованого опрацювання інформації. Як зазначає В. А. Широков, для проведення автоматизованого аналізу граматичних, лексичних, семантичних та інших ефектів, відтворених у СУМ, необхідно в явному вигляді мати такі формальні об'єкти, які репрезентують структуру кожної його словникової статті, структуру міжсловникових відображень і словника в цілому [4, с. 183].

Аналіз структурних елементів семантики відносних прикметників зроблено на основі їх опису в Словнику української мови у 20 томах відповідно до теоретичних засад інформаційної теорії лексикографічних систем.

У тлумачному словнику відображено системний опис лінгвістичної інформації, яку впорядковано та структуровано відповідно до лексикографічних правил і традицій. Як лексикографічна система словник містить мінімум два необхідні складники: реєстрову (ліву) та інтерпретаційну (праву), що є виявом відношення «форма – зміст». Саме наявність інтерпретаційної частини – носія змістового компонента відношення «форма – зміст» відрізняє словник від звичайного списку слів [6, с. 118].

Структурні елементи лівої частини словникової статті в Словнику мають певний набір параметрів, які використано для опису граматичної семантики, що відображено за відповідними лексикографічними ремарками.

Реєстрове слово виділене жирним шрифтом та обов'язково з наголосом. Узагальнено структуру прикметникових словникових статей можна представити в такій формі:

X	V	F	S
---	---	---	---

X – реєстрове слово;

V – лексемні варіанти реєстрового слова;

F – флексія реєстрового слова (жіночого і середнього роду);

S – функціонально-стильова характеристика прикметника.

Лексемні варіанти, подані у лівій частині словникової статті, можуть бути утворені на основі суфіксальної, префіксальної, акцентуаційної чи фонетичної варіації реєстрового слова. За кількістю лексичних компонентів реєстрового ряду словникова стаття може бути однокомпонентною (K^1), двокомпонентною (K^2), трикомпонентною (K^3) і т. ін.:

АГЛИЦЬКИЙ, а, е, *заст.*: K^1 = аглицький.

АВАНТЮРНИЦЬКИЙ, *рідко* **АВАНТУ́РНИЦЬКИЙ**, а, е.: K^1 = авантюрницький; K^2 = авантурницький; фонетична варіація.

УКРАЇНСЬКИЙ (ВКРАЇНСЬКИЙ), а, е.: K^1 = український; K^2 = вкраїнський; префіксальна варіація.

УСМІХНЕНИЙ (ВСМІХНЕНИЙ), УСМІХНУТИЙ (ВСМІХНУТИЙ), а, е.: K^1 = усміхнений; K^2 = всміхнений; K^3 = усміхнутий; K^4 = всміхнутий; префіксальна варіація.

ВСЕВИДЮЩИЙ (УСЕВИДЮЩИЙ), рідше ВСЕВИДЮЧИЙ (УСЕВИДЮЧИЙ), ВСЕВІДЯЩИЙ (УСЕВІДЯЩИЙ), ВСЕВІДЯЧИЙ (УСЕВІДЯЧИЙ), а, е.: K^1 = всевидющий; K^2 = усевидющий; K^3 = всевидючий; K^4 = усевидючий; K^5 = всевидящий; K^6 = усевидячий; K^7 = всевидячий; K^8 = усевидячий; префіксальна варіація.

Структурне представлення параметрів лівої частини словникової статті, що репрезентує систему граматичної семантики прикметників української мови, сприятиме побудові алгоритмів для подальшого опрацювання.

Лексична семантика мовних одиниць у Словнику представлена в правій частині його статей, де відображена смислова структура слова, багатозначність реєстрової одиниці та її фразеологізми, зв'язок значень з іншими словами. Усі структурні ефекти формально визначено в тексті словника за допомогою спеціальних розділових позначок, позиційно, шрифтовими виділеннями тощо, тому побудову формального репрезентанта структури може бути повністю алгоритмізовано [4, с. 206].

На основі інформаційної теорії лексикографічних систем у дослідженні правої частини прикметникових словникових статей Словника, вирізняємо набір структурних елементів, за допомогою яких провадимо системний опис лексичного значення реєстрового слова.

Для прикметникових правих частин статті такими є елементи:

T – тлумачення лексичного значення прикметника;

S – один або кілька синонімів;

A – один або кілька антонімів;

I_1 – ілюстрація / ілюстрації до основного тлумачення;

V_1 – відтінок значення, який представлений у Словнику двома скісними рисками «//»;

I_2 – ілюстрація / ілюстрації до відтінку значення;

F – інформація про функціонування прикметника в складі усталених словосполучень [стійкі словосполучення, словосполучення – термінологічні словосполучення з позначкою трикутник (Δ), фразеологічні одиниці з позначкою ромб ($\diamond$)];

V_2 – відтінок значення до стійких словосполучень;

I_3 – ілюстрація / ілюстрації до стійких словосполучень.

Унаочнимо викладене на прикладі словникової статті:

«АМАТОРСЬКИЙ, а, е. Прикм. До **аматор**; любительський. *По містах і містечках почали організовуватися аматорські трупи (І. Франко); – Так тоді добре, слухайте і мотайте на вус. Чорнишенко хоче заснувати театр, перший на цьому клаптику землі, справжній, не аматорський театр (Б. Лепкий); // Виконуваний аматорами. На вакаціях учителі здебільшого робляться акторами і грають преподано аматорські спектаклі (І. Карпенко-Карий); В місті його знали як українофіла, бо носив під форменою тужуркою вишивану сорочку з стьожкою і був організатором аматорських малоросійських спектаклів (Б. Антоненко-Давидович); Від поверхового, малокомпетентного, аматорського вивчення та залучення випадкових сортів Л. П. Симиренко перейшов до науково обґрунтованої інтродукції та всебічного дослідження їхніх господарських і біологічних особливостей (з газ.)» [2].*

Зазначені елементи, за допомогою яких зроблено лексикографічний опис лексичної семантики мовних одиниць, набувають статусу формальних репрезентантів структури правої частини словникової статті.

На основі аналізу словникових статей з'ясовано, що лексична семантика прикметників представлена за допомогою основних способів тлумачення слів у тлумачному словнику: описовий, синонімічний, відсильний, трапляються також комбіновані способи тлумачень – описово-синонімічний, описово-антонімічний, описово-перелічувальний та ін.

Окремого розгляду потребують відносні прикметники, лексичне значення яких побудовано за відсильними формулами тлумачення «Прикм. до» та «Стос. до», які лише встановлюють факт зв'язку значення прикметників із лексичними значеннями іменників, від яких їх утворено, що внаслідок проведення аналізу їхніх семантичних властивостей за допомогою комп'ютерних програм. Зазначені формули тлумачення мають досить широку семантику й передають загальне значення відносних прикметників. За своїм семантичним обсягом такі формули повністю охоплюють усі смислові навантаження, зазначені в ілюстративній частині словникової статті. Більш докладно вони розглянуті в праці [1].

За своєю структурою словникові статті відносних прикметників із формулами тлумачення «Прикм. до» та «Стос. до» мають таке відображення: X див. Y, наприклад: **«АБІТУРІЄНТСЬКИЙ**, а, е. Прикм. До **абітурієнт**» [2].

Будова словникової статті може бути складнішою, оскільки довжина низки відсилянй може бути більшою за одиницю й мати цілий ланцюг відсилянй:

x ... див. y^1, y^2
x ... див. y^1, y^2, y^3
 $x^1, x^2, \dots$ див. y^1, y^2
 $x^1, x^2, x^3, \dots$ див. $y^1, y^2, y^3, \dots$

Наприклад:

«АРГОТІЧНИЙ, а, е, лінгв. Прикм. До **арго** і **арготізм**.

АБА́ТСЬКИЙ, а, е. 1. Прикм. До **абат**, **абатіса** і **абатство**.

АУДИТО́РНИЙ, рідко **АВДИТО́РНИЙ**, а, е. Прикм. До **аудито́рія**, **авдито́рія** 1.

А́НГЕЛЬСЬКИЙ, рідше **А́НГОЛЬСЬКИЙ**, **Я́НГОЛЬСЬКИЙ**, а, е. 1. бібл. Прикм. до **а́нгел**, **а́нгол**, **я́нгол** 1» [2].

На основі опрацьованого матеріалу в дослідженні створено лексикографічну базу даних «Структура семантики прикметників української мови», яка даних містить такі поля:

Поле 1 – відображення прикметникової реєстрової лексеми.

Поле 2 – відображення граматичних параметрів реєстрового слова.

Поле 3 – відображення тлумачення реєстрової лексеми.

Поле 4 – відображення синонімів / антонімів.

Поле 5 – відображення лексико-семантичного варіанта реєстрового слова.

Фрагмент лексикографічної бази даних

Поле 1	Поле 2	Поле 3	Поле 4	Поле 5
АДЕКВА́ТНИЙ	а, е	Цілком відповідний	тотожний	
ВИШНЕ́ВИЙ	а, е	1. Прикм. до ви́шня 2. Темно-червоний, кольору стиглої темної вишні		// Вигот. із вишень (у 2 знач.)
КВІТЧА́СТИЙ	а, е <i>перен.</i>	1. Укритий квітами 2. Різного кольору 3. Дуже прикрашений	різнобарвний примхливий, пишний (про мову, слова і т. ін.)	// Який зображує квіти, зелень // На якому зображені квіти // Пістрявий, строкатий
КОЛЕКТІ́ВНИЙ	а, е.	Прикм. до колектив	спільний	// Який належить колективу. // Який здійснює колектив. // Власт. колективу, який ґрунтується на спільності праці, інтересів та на принципах колективізму. // Признач. для колективу.

База даних є зручним інструментом для одержання різнотипної інформації за будь-яким інформативним елементом, що значно спрощує та оптимізує аналіз семантичної структури досліджуваних прикметникових лексем. Крім того, формат лексикографічної бази даних має такі переваги: слугує ефективним способом систематизації і впорядкування мовної інформації, дає можливість провадити дослідження за окремим лексикографічним параметром, у зручній формі структурувати інформацію, зробивши її придатною для подальшого опрацювання програмними засобами, що є необхідною умовою для автоматизованого оброблення інформації.

Висновки і перспективи. На засадах інформаційної теорії лексикографічних систем виокремлено й проаналізовано структурні елементи словникових статей українських прикметників, що є необхідною умовою для подальшого автоматизованого опрацювання інформації в лексикографічній системі, зокрема для моделювання граматичної і лексичної семантики.

Розроблена лексикографічна база даних «Структура семантики відносних прикметників» слугує зручним інструментом для одержання семантичної інформації способом фільтрування даних за будь-яким інформативним елементом, що значно спрощує й оптимізує аналіз семантичної структури досліджуваних прикметникових лексем, уможливорює провадження автоматичних вибірок слів за наявністю того чи того компонента в структурі їхнього значення.

Список використаних джерел

1. Олександрук І. В. Застосування комп'ютерних технологій під час дослідження семантики відносних прикметників української мови. *Pomiedzy. Polonistyczno-Ukrainoznawcze Studia Naukowe*. 2022, nr 4 (7), ss. 59–66.
2. Словник української мови у 20 томах online. URL: <https://services.ulif.org.ua/expl/Entry/index?wordid=25129&page=835> (дата звернення 10.03.2025).
3. Широков В. А. Елементи лексикографії. Київ : Довіра, 2005. 304 с.
4. Широков В. А. Інформаційна теорія лексикографічних систем. Київ : Довіра. 1998. 331 с.
5. Широков В. А. Комп'ютерна лексикографія. Київ : Наукова думка, 2011. 352 с.
6. Широков В. А. та ін. Лінгвістично-інформаційні студії: праці Українського мовно-інформаційного фонду НАН України : у 5 т. Київ : УМІФ, 2018. Т. 1 : Наукова парадигма та основні мовно-інформаційні структури. 271 с.
7. Широков В. А., Олександрук І. В. Віртуальна лексикографічна лабораторія «Словник української мови» як інструмент семантичних досліджень. *Мовознавство*. 2019. № 5. С. 3–20.
8. Широков В. А. Феноменологія лексикографічних систем. Київ : Наукова думка, 2004. 326 с.

References

1. Oleksandruk, I. V. (2022). Zastosuvannya komp'yuternykh tekhnolohiy pid chas doslidzhennya semantyky vidnosnykh prykmetnykiv ukrayins'koyi movy [Application of computer technologies in the study of semantics of relative adjectives in the Ukrainian language]. In: *Pomiedzy. Polonistyczno-Ukrainoznawcze Studia Naukowe* [Between. Polish-Ukrainian Studies], 4 (7), 59–66 (in Ukr.).
2. Slovnyk ukrayins'koyi movy u 20 tomakh online [Dictionary of the Ukrainian language in 20 volumes online]. Available at: <https://services.ulif.org.ua/expl/Entry/index?wordid=25129&page=835> (in Ukr.).
3. Shyrokov, V. A. (2005). Elementy leksykohrafiyi [Elements of lexicography]. Kyiv: Dovira, 304 (in Ukr.).
4. Shyrokov, V. A. (1998). Informatsiyna teoriya leksykohrafichnykh system [Information theory of lexicographic systems]. Kyiv: Dovira, 331 (in Ukr.).
5. Shyrokov V. A. (2011). Komp'yuterna leksykohrafiya [Computer lexicography]. Kyiv: Naukova dumka, 352 (in Ukr.).
6. Shyrokov, V. A. ta in. (2018). Linhvistychno-informatsiyni studiyi: pratsi Ukrayins'koho movno-informatsiynoho fondu NAN Ukrayiny: u 5 t. [Linguistic and information studies: works of the Ukrainian Language and Information Foundation of the National Academy of Sciences of Ukraine: in 5 vols.]. Kyiv: UMIF, 1: Naukova paradyhma ta osnovni movno-informatsiyni struktury, 271 (in Ukr.).
7. Shyrokov, V. A. & Oleksandruk I. V. (2019). Virtual'na leksykohrafichna laboratoriya «Slovnyk ukrayins'koyi movy» yak instrument semantichnykh doslidzhen' [Virtual lexicographic laboratory “Dictionary of the Ukrainian Language” as a tool for semantic research]. *Movoznavstvo* [Linguistics], 5, 3–20 (in Ukr.).
8. Shyrokov, V. A. (2004). Fenomenolohiya leksykohrafichnykh system [Phenomenology of lexicographic systems]. Kyiv: Naukova dumka, 326 (in Ukr.).

THE STRUCTURE OF THE SEMANTICS OF UKRAINIAN ADJECTIVES IN THE LEXICOGRAPHIC SYSTEM

Iryna Oleksandruk, PhD, junior research fellow

*at the O. O. Potebnya Institute of the National Academy of Sciences of Ukraine
(Kyiv, Ukraine)*

e-mail: ira_oleksandruk@ukr.net

ORCID: <https://orcid.org/0000-0003-2701-1714>

Abstract. Introduction. The article deals with the structure of the semantics of Ukrainian adjectives in the lexicographic system of the explanatory type. The study was conducted on the basis of the electronic form of the Dictionary of the Ukrainian Language in 20 volumes, which is being created by the Ukrainian Language and Information Foundation of the National Academy of Sciences of Ukraine. The dictionary is presented in the form of the Virtual Lexicographic Laboratory “Dictionary of the Ukrainian Language”. The relevance of the study is due to the use of computer technologies in semantics

research aimed at creating automated systems for processing language information. The main tool for semantic analysis in these systems is electronic explanatory dictionaries.

The purpose of the article is to analyze the structure of the semantics of relative adjectives on the basis of the information theory of lexicographic systems for the possibility of further processing by software. The object of the study is the dictionary entries of adjectives in the Ukrainian language.

The methods. Research methods – lexicographic analysis, information theory of lexicographic systems.

Main results of the study. Originality. The article analyzes the structural representation of the parameters of the left part of the dictionary entry, which represents the system of grammatical semantics, and the right part, which represents the lexical semantics of adjectives in the Ukrainian language, on the basis of the information theory of lexicographic systems. As part of the study, the lexicographic database “Structure of Semantics of Ukrainian Adjectives” was developed, which is a convenient tool for further elaboration of the structure of adjectival dictionary entries.

Conclusion. On the basis of the information theory of lexicographic systems, the structural elements of dictionary entries of Ukrainian adjectives are identified and analyzed, which is a prerequisite for further automated processing of information in a lexicographic system, in particular, modeling of grammatical and lexical semantics.

Key words: electronic explanatory dictionary, computer lexicography, semantics, relative adjectives of the Ukrainian language.

Надійшла до редакції: 03.04.2025

Прийнято до друку: 21.05.2025