

*До 100-річчя
Національної академії наук України*

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
УКРАЇНСЬКИЙ МОВНО-ІНФОРМАЦІЙНИЙ ФОНД

ЛІНГВІСТИЧНО-ІНФОРМАЦІЙНІ СТУДІЇ

ПРАЦІ
УКРАЇНСЬКОГО МОВНО-ІНФОРМАЦІЙНОГО ФОНДУ
НАН УКРАЇНИ

У П'ЯТИ ТОМАХ

Дослідження 1991–2018 років

За загальною редакцією *В. А. Широкова*

КИЇВ 2018

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
УКРАЇНСЬКИЙ МОВНО-ІНФОРМАЦІЙНИЙ ФОНД

ТОМ III

ТЛУМАЧНА ЛЕКСИКОГРАФІЯ

КНИГА 2

**СИСТЕМНА СЕМАНТИКА
ТЛУМАЧНИХ СЛОВНИКІВ**

УКРАЇНСЬКИЙ МОВНО-ІНФОРМАЦІЙНИЙ ФОНД

*Рекомендовано до друку вченою радою
Українського мовно-інформаційного фонду НАН України
(протокол № 9 від 23 жовтня 2018 р.)*

Рецензенти:

академік НАН України *О. В. Палагін*,
член-кореспондент НАН України *А. О. Загнітко*

Автори другої книги третього тому:

*В. А. Широков, Н. М. Заїка, Є. В. Купріянов, О. О. Погрібна,
В. В. Чумак, І. В. Шевченко, К. В. Широков*

Лінгвістично-інформаційні студії: Праці Українського мовно-інформаційного фонду НАН України : у 5 т. Т. 3 : Тлумачна лексикографія. Кн. 2 : Системна семантика тлумачних словників / В. А. Широков та ін. Київ : Український мовно-інформаційний фонд НАН України, 2018. 250 с.

Другу книгу «Системна семантика тлумачних словників» третього тому «Тлумачна лексикографія» п'ятитомної монографії «Лінгвістично-інформаційні студії» присвячено викладу системних ефектів мови, які проявляються при аналізі великих тлумачних лексиконів.

Сформульовано вимоги лексикографічної повноти та замкненості, що висуваються до великих тлумачних словників, і проаналізовано логіко-лінгвістичні наслідки, які випливають з цих вимог. Уведено поняття автоморфізмів Л-систем, гіперланцюгів та гіперциклів на лексикографічних структурах і показано їх зв'язок з теоремою Геделя про неповноту формальних систем.

Із застосуванням ВЛЛ «Словник української мови» досліджено особливий клас дефініцій тлумачного словника – так звані формули квазісемантики, і сформульовано завдання про визначення ядра лексичної системи української мови.

Запропоновано уніфіковане представлення одиниць лексичного рівня та колокацій в єдиній формальній моделі на основі узагальнення поняття граматичного стану мовних одиниць.

Досліджено та викладено закони формування еквісемантичних рядів для українського дієслова та іменника. Для зазначених лексико-граматичних класів досліджено ряд семантичних властивостей та побудовано відповідні лексико-семантичні класифікації.

Системні засади теорії лексикографічних систем та теорії семантичних станів застосовано до побудови концептуальних моделей лексикографічних систем тлумачного типу для іспанської й турецької мов. На цій основі розроблено моделі відповідних віртуальних лексикографічних лабораторій (з комп'ютерною реалізацією для турецької мови).

Одержані результати надали шлях до формулювання ідеї так званого абсолютного словника, який в єдиному інтегрованому лексикографічному середовищі поєднує максимальний комплекс лінгвістичних об'єктів та відношень.

ЗМІСТ

ВСТУП

СИСТЕМНА ЛЕКСИКОГРАФІЯ ТА ІНТЕГРОВАНІЙ ОПИС МОВИ.....	6
---	---

Розділ 14. СИСТЕМНІ ЕФЕКТИ В ТЛУМАЧНИХ СЛОВНИКАХ УКРАЇНСЬКОЇ МОВИ.....

11

14.1. Лексикографічна замкненість і повнота. Автоморфізми L -систем, гіперланцюги і гіперцикли	11
14.2. Формули квазісемантики і ядро лексичної системи	19
14.3. Колокації та узагальнення граматичних станів	23
14.4. Дієслово в лексикографічній системі СУМа	29
14.4.1. Граматична семантика дієслова в лексикографічній системі	29
14.4.1.1. Еквісемантичні ряди дієслів у СУМі: структура та властивості	29
14.4.1.2. Словник сигнатурних класів дієслів сучасної української мови	55
14.4.2. Лексична семантика дієслова в лексикографічній системі	56
14.4.2.1. Репрезентація лексичної семантики у структурі правих частин дієслівних словникових статей	56
14.4.2.2. Формули тлумачення та їх рекурсивна редукція	61
14.4.2.3. Лексикографічна система «Словника дієслівних семантичних тем».....	73
14.4.2.4. Класи дієслівних лексико-семантичних варіантів із співвідносними суб'єктно-відображальними класами імен.....	80
14.4.2.5. Лексико-семантичні групи дієслів та їх диференційні семи.....	85
14.5. Іменник у лексикографічній системі СУМа	96

Розділ 15. ЛЕКСИКОГРАФІЧНА СИСТЕМА СЛОВНИКА ІСПАНСЬКОЇ МОВИ

128

15.1. Склад словникового реєстру	129
15.2. Метамова словника	131
15.3. Лексикографічні параметри та їх зміст	135
15.4. Семантичні стани та їх лексикографічна репрезентація	147
15.4.1. Семантичні стани як об'єкт тлумачного словника іспанської мови	148
15.4.2. Репрезентація граматичного і лексичного компонентів семантичних станів	161
15.5. Суперпозиції семантичних станів іспанських мовних одиниць	177
15.6. Побудова L -системи тлумачного словника іспанської мови.....	189
15.6.1. Загальна структура L -системи словника іспанської мови	191
15.7. Розкладання L -системи словника на β -структури	200
15.8. Встановлення відношень σ на структурах β	209

Розділ 16. ЛЕКСИКОГРАФІЧНА СИСТЕМА СЛОВНИКА ТУРЕЦЬКОЇ МОВИ

16.1. Застосування концепції лексикографічних систем у турецькомовній лексикографії	216
16.2. Концептуальне моделювання лексикографічної системи тлумачного словника турецької мови та його метамова.....	217
16.3. Структура бази даних тлумачного словника турецької мови	225
ЗАКЛЮЧНІ ЗАУВАЖЕННЯ	
НА ШЛЯХУ ДО «АБСОЛЮТНОГО» СЛОВНИКА	236
ЛІТЕРАТУРА.....	248

ВСТУП

СИСТЕМНА ЛЕКСИКОГРАФІЯ ТА ІНТЕГРОВАНІЙ ОПИС МОВИ

«Μέγα βιβλίον μέγα κακόν»
Καλλίμαχος
(«Велика книга – велике зло»
Калімахос)

Ідеї інтегрованого представлення різноаспектної лінгвістичної інформації в одному словниковому виданні, які знайшли своє формальне відображення та узагальнення у теорії лексикографічних середовищ та інтегрованих лексикографічних систем, на сьогодні представляють найактуальніші прагнення лексикографів. Ідеалом теорії лексикографії залишається універсальний, «абсолютний» словник, який з точністю і повнотою міг би відобразити мовну стихію, усі розряди одиниць лексикону зі всіма властивими їм ознаками: «Для теорії лексикографії необхідне поняття максимальної лексичної системи і поняття універсального словника, в якому зафіксовано максимальну лексичну систему. Користуючись таким словником, можна спілкуватися цією мовою в усіх ситуаціях, розмовляти на будь-яку тему з будь-яким носієм цієї мови. Легко зрозуміти, що це – усвідомлена ідеалізація, теоретичний світ лексикографії»¹.

Останнє зауваження, на нашу думку, вже не є таким категоричним, коли ми говоримо про комп'ютерні лексикографічні системи, спроможні до рецесії практично необмеженого інтеграційного потенціалу. Втім і досягнення традиційної лексикографії мають внутрішні ресурси до узагальнень та застосування у нових інформаційних середовищах. Так, наприклад, П. Н. Денисов² цілком слушно висуває питання про своєрідну «канонізацію» (визнання стандарту) тих схем словникових статей, які сформувалися в кращих словниках кожного із типів. І справді, найкращі словники мають настільки

¹ Денисов П. Н. Системность и связанность в лексике и система словарей // Проблематика определенных терминов в словарях разных типов / Ред. кол. С. Г. Бархударов и др. Ленинград : Наука, 1976. С. 68–69.

² Денисов П. Н. Лексика русского языка и принципы ее описания. М.: Рус. яз., 1980. С. 217.

досконалі структури словникових статей, що з них вдумливому дослідникові вдається досить легко абстрагувати елементи концептуальних моделей Л-систем, які складають субстрат цих словників, і потім збудувати відповідні внутрішні та зовнішні моделі, тобто усі стандартні елементи їхньої архітектури, передбачені теорією лексикографічних систем.

Задля отримання такого результату необхідне позитивне вирішення двох завдань: по-перше, володіти необхідною лінгвістичною інформацією, а по-друге – розробити формат, адаптований для її представлення. Щодо вирішення першого питання, то, на наше бачення, одномовна тлумачна лексикографія найближче підійшла до системного опису мови. Що ж до забезпечення прийняттого формату для представлення такого опису мови, то це вже є прерогативою комп'ютерної лінгвістики та лексикографії, про що буде йтися далі.

Основи системної лексикографії ґрунтовно опрацював Ю. Д. Апресян³. Він же увів поняття лексикографічного типу та лексикографічного портрета.

Поняття лексикографічного типу відрізняється від більш традиційних понять (семантичне поле, семантичний клас, лексико-семантична група слів), оскільки основою для виділення лексикографічного типу є будь-які спільні властивості лексем, а не лише семантичні. Поняття лексикографічного типу має сенс лише в рамках інтегрального опису мови, тобто узгодженого опису граматики і словника. В інтегральному словнику лексемам приписуються ті властивості, звернення до яких вимагають якісь лінгвістичні правила. Членування ж на семантичні класи визначається лише ідеографічними міркуваннями. Лексикографічний тип, на відміну від семантичних класів (як вони представлені, наприклад, в ідеографічному словнику) утворюють не строгую ієрархію, а багаторазово перехрещувані класи, оскільки одна і та ж лексема може потрапляти в різні класи за кожною із своїх властивостей і за будь-якою сукупністю властивостей.

³ Апресян Ю. Д. Избранные труды. Т. 1: Лексическая семантика: синонимические средства языка. 2 изд., испр. и доп. Москва : Школа «Языки рус. культуры», 1995. 472 с.; Апресян Ю. Д. Избранные труды. Т. 2: Интегральное описание языка и системная лексикография. Москва : Школа «Языки рус. культуры», 1995. 767 с.; Апресян Ю. Д. Интегральное описание языка и толковый словарь. Вопросы языкознания. 1986. № 2. С. 57–70.

Другою опорою системної лексикографії виступає поняття лексикографічного портрета – вичерпної і ненадлишкової характеристики усіх лінгвістично суттєвих властивостей даної лексеми в рамках інтегрального опису мови.

У такий спосіб системна лексикографія прагне адекватно розв'язати два головні завдання – уніфікація (лексикографічні типи) та індивідуалізація (лексикографічні портрети): «...лексикографічний портрет повинен відповідати одному чи декільком лексикографічним типам. Тільки тоді можна вважати, що лексикографічний опис задовольняє вимогу системності»⁴.

Кінець ХХ – початок ХХІ ст. ознаменувалися значним поступом у розкритті законів організації мовної системи, що пов'язано, безумовно, і з новими інструментальними можливостями вивчення мовних одиниць, які надають комп'ютерні технології. Комп'ютеризація лінгвістичних досліджень, стимульована набуттям природною мовою технологічного статусу в сучасному мережецентричному суспільстві, спонукала до інтенсифікації досліджень та розробок у галузі системної лінгвістики і лексикографії. Це, у свою чергу, стимулювало й намагання уточнити й поглибити засади системного підходу в лінгвістичній науці. У першому томі нашого багатотомника нами здійснено уточнення поняття системи в лінгвістиці через виділення системної тріади «структура – субстанція – суб'єкт», властивості якої яскраво проявляються практично у всіх ділянках теоретичного опису мови та численних практичних застосуваннях. Таке тріадичне визначення поняття «система» надає конструктивний шлях до узагальнення і водночас конкретизації методик й засобів системної лексикографії. Зокрема, в Українському мовно-інформаційному фонді системний підхід було застосовано до таких фундаментальних лексикографічних об'єктів, якими є тлумачні словники національних мов.

Це не випадково.

Системні властивості мови, взагалі, розглядаються в багатьох наукових працях. Але ми вважаємо, що системні ефекти найбільш повно представлено в одномовних тлумачних словниках. Йдеться про великі багатотомні лексикони, які

⁴ Дубичинский В. В. Теоретическая и практическая лексикография. Вена, Харьков: Wienar Slawistischer Almanach Sonderband 45, 1998. С. 57.

містять основну частину національної лексики і фразеології й характери-зуються докладним описом лексико-граматичної і лексико-семантичної системи мови. Завдяки великому обсягу, складності будови та повноті лексикографічного опису такі словники є носіями величезної кількості глибоких, імпліцитно заданих лінгвістичних, когнітивних, логічних та інших системних зв'язків і відношень, переважно неконтрольованих, які, будучи неочевидними, «прихованими» від читача, відіграють важливу роль у формуванні інформаційного потенціалу мови. Сказане перетворює ці великі лексикографічні системи на певного роду «речі в собі». Справді, пересічний користувач великого лексикону лише «ковзається» по його поверхні, не маючи ефективних засобів для проникнення в глибину його змісту. Це добре видно зі схеми: «Геометричний образ словника»:



Але що ще гірше, що така можливість не доступна навіть для укладачів словника, які змушені враховувати системні відношення мови, лише спираючись на власний досвід та лексикографічну інтуїцію. Викладене виправдовує наведене в епіграфі твердження, що «велика книга – це велике зло».

Ми підкреслюємо інноваційну роль великих словників у сучасну епоху як семантичних аналізаторів, тегерів, елементів концепто- і онтографічних систем, резидентних елементів мовної підтримки різного типу прикладних мовно-інформаційних систем і процесів (експертиза текстів, законодавство і правозастосування, інженерія знань тощо). Таким чином, значення словників у сучасних інформаційних технологіях постійно зростає, що робить необхідним

переведення їх до цифрових форматів. Але цей самий чинник висуває нові вимоги і до якості словників та засобів і способів експлікації їхньої внутрішньої системної семантичної і логічної структури. Отже, виникає проблема виявлення і аналізу системних відношень, сконцентрованих у тлумачних словниках. Деякі з них добре відомі й зрозумілі. Однак тут ми зосередимося на аналізі властивостей і відношень, які зазвичай не розглядаються у стандартній лексикографії.

Наше дослідження системних властивостей мови виконано на матеріалі Словника української мови в 11 томах (СУМ-11) і Словника української мови в 20 томах (СУМ-20); теорія і технологія, а також стан розроблення викладено у першій книзі третього тому. Підкреслимо, що йдеться саме про електронні версії цих словників, адже, як буде показано в ході викладу, виявити системні закономірності у, так би мовити, «ручному режимі» майже неможливо. Тому експериментальні дослідження було здійснено переважно за допомогою Віртуальної лексикографічної лабораторії «Словник української мови» (ВЛЛ_СУМ), яка не лише забезпечує спільні, скоординовані зусилля значної, територіально розподіленої команди лексикографів, але й надає необхідні інструменти для проведення лінгвістичного дослідження системних властивостей мови, втілених у Словнику.

РОЗДІЛ 14

СИСТЕМНІ ЕФЕКТИ В ТЛУМАЧНИХ СЛОВНИКАХ

14.1. Лексикографічна замкненість і повнота.

Автоморфізми Л-систем, гіперланцюги і гіперцикли

Перейдемо до визначення і опису системних відношень, що експлікуються із тлумачних словників. Перше з них індуковано так званими гіперланцюгами на лексикографічних структурах, які вводяться в такий спосіб (зауважимо, що наш подальший виклад базується на теорії лексикографічних систем, поняття та позначення якої ми використовуємо, слідуючи за змістом попередніх томів).

Нехай задано лексикографічну систему (надалі: Л-система) $V(U_0)$. На її структурі $\sigma[\beta]$ виокремимо певну підмножину \mathbf{A} , елементи якої A ($A \in \mathbf{A}$) назвемо *автоморфізмами* Л-системи $V(U_0)$. Смысл цих елементів полягає в тому, що вони визначають внутрішні відображення Л-системи

$$A: V(U_0) \rightarrow V(U_0) \quad (14.1)$$

спеціального типу – відображення між певними словниковими статтями $A: V(x) \rightarrow V(y)$ для різних заголовних слів x та y , що належать до множини U_0 . Довжина ланцюга посилань може бути довшою за 1, так що зрештою одержуємо ланцюг рекурсивного типу:

$$V(x) \rightarrow \{V(x')\} \rightarrow \dots \rightarrow \{V(x'')\} \rightarrow \dots$$

Таким чином, автоморфізми з множини \mathbf{A} (надалі називатимемо їх \mathbf{A} -автоморфізмами) індукують відображення множини заголовних слів у себе:

$$A: U_0 \rightarrow U_0. \quad (14.2)$$

Розглянемо один із алгоритмів такого відображення, який має прозору семантичну інтерпретацію.

Обов'язковим (нефакультативним) елементом словникової статті тлумачного словника є лексичне значення, яке представлено в лексикографічній структурі словниковою дефініцією. Ми постулюємо два основні системні