

В. А. ШИРОКОВ, І. В. ОЛЕКСАНДРУК

ВІРТУАЛЬНА ЛЕКСИКОГРАФІЧНА ЛАБОРАТОРІЯ «СЛОВНИК УКРАЇНСЬКОЇ МОВИ» ЯК ІНСТРУМЕНТ СЕМАНТИЧНИХ ДОСЛІДЖЕНЬ

У статті йдеться про застосування теорії лексикографічних систем як методології організації та проведення різнопланових досліджень мовної системи. Інструментальним засобом для цього служить Віртуальна лексикографічна лабораторія «Словник української мови» (ВЛЛ СУМ). Подано приклади закономірностей лексичної системи української мови, встановлені з використанням ВЛЛ СУМ. Запропоновано семантичну та інструментальну структури електронного тлумачного словника, інтегрованого з ВЛЛ СУМ, для поглибленого дослідження семантичної структури українського прикметника.

Ключові слова: лексикографічна система, лексикографічний ефект, лексична семантика, лінгвістична база даних, віртуальна лексикографічна лабораторія.

Вступ. Загальновідомо, що великі словники є носіями глибоких системних закономірностей мови, які, будучи неочевидними, «прихованими» від читача, відіграють важливу роль у виявленні її когнітивно-комунікативного потенціалу. На нашу думку, такі системні ефекти найповніше представлені в одномовних тлумачних словниках, а саме — у великих багатотомних лексиконах, які містять основну частину національної лексики і фразеології й характеризуються докладним описом лексико-граматичної і лексико-семантичної системи відповідної мови. Завдяки своєму обсягові, складності будови та повноті лексикографічного опису такі словники є носіями великої кількості імпліцитно заданих лінгвістичних, когнітивних, логічних та інших зв'язків і відношень, переважно неконтрольованих, що перетворює ці лексикографічні об'єкти на певного роду «речі в собі»¹, чий глибини здебільшого недосяжні для користувачів. Проте тут варто зазначити, про яку категорію користувачів йдеться, адже когнітивний, так би мовити, діапазон зацікавленостей в інформації, що міститься в таких словниках, є надзвичайно широким — від елементарних відомостей про конкретне слово аж до узагальненої граматико-семантичної інформації, що стосується цілих класів мовних одиниць та різноманітних лінгвістичних відношень, що розвиваються і функціонують у мовній системі.

Стосовно такої узагальненої лінгвістичної інформації слід зазначити, що її «видобування» зі словника перетворюється на справжню проблему через згадану вище імпліцитність. Адже навіть лінгвістично освічений і кваліфіко-

¹ Широков В. А. Системна семантика тлумачних словників // Акцентологія. Етимологія. Семантика : До 75-річчя академіка НАН України В. Г. Скляренка.— К., 2012.— С. 487–510.

ваний користувач великого лексикона насправді лише «ковзається» по його поверхні, не маючи ефективних засобів для проникнення в глибину його змісту. Проте таке проникнення не лише доцільне або бажане, але й спроможне стати ефективним джерелом нових лінгвістичних фактів, закономірностей і навіть законів мовної системи. Ми виходимо тут з переконання, що великий тлумачний словник як лексикографічна система може справді містити у неявному вигляді досі не відкриті закони мови. Обґрунтуванням цього твердження служать принципи побудови та способи експлікування лінгвістичної інформації у словниках згаданого типу, а також особистий досвід авторів.

Мета статті — окреслити основні засади теорії лексикографічних систем як методу організації лінгвістичних досліджень та описати застосування ВЛЛ «Словник української мови» як інструменту для проведення семантичних досліджень і створення семантичного словника.

Тлумачний словник української мови як лексикографічна система

Узагальненням згаданого досвіду стала сформульована одним з авторів цієї статті (В. Ш.) теорія лексикографічних систем, згідно з якою будь-який словник (у тому числі й тлумачний) є лексикографічною системою певного типу. Сама лексикографічна система (*Л-система*) є інформаційною системою, у якій діють один або декілька лексикографічних ефектів². Основними системотвірними відношеннями Л-системи виступають відношення «суб'єкт – об'єкт» і «форма – зміст», розгортання яких веде до поняття структури Л-системи. Докладно теорію Л-систем, включаючи й теорію Л-системи академічного тлумачного «Словника української мови», викладено в монографіях³.

Відповідно до викладеного в будь-якій Л-системі визначаються такі сутності:

D — фрагмент дійсності, який виступає об'єктом лексикографічного опису;

S — суб'єкт, що здійснює аналіз та лексикографічний опис *D*;

Q — лексикографічний ефект, що його суб'єкт *S* спостерігає в *D*, у результаті чого виявляється і формується комплекс елементарних інформаційних одиниць $I^Q(D)$;

$V(I^Q(D))$ — множина лінгвістичних описів $I^Q(D)$; $S: I^Q(D) \rightarrow V(I^Q(D))$. При цьому для будь-якої елементарної інформаційної (ресстрової) одиниці *x* справедливе таке представлення:

$$I^Q(D) = \{x\}; \forall x \in I^Q(D) S: x \rightarrow V(x). \bigcup_{x \in I^Q(D)} V(x) = V(I^Q(D)). \quad (1)$$

² Широков В. А. Інформаційна теорія лексикографічних систем. — К., 1998. — С. 50–56. Лексикографічний ефект, за нашими уявленнями, є основним системотвірним феноменологічним принципом мови. Згідно з теорією лексикографічних систем він відповідає за виявлення у певній ділянці мови комплексів дискретних одиниць (за термінологією теорії лексикографічного ефекту — елементарних інформаційних одиниць), які відіграють роль одиниць лексикографування відповідного словника. Таким чином, будь-який феномен мови може бути «лексикографований» і, отже, стати об'єктом певного словника.

³ Див.: Широков В. А. Інформаційна теорія лексикографічних систем. — К., 1998. — 331 с.; Широков В. А. Елементи лексикографії. — К., 2005. — 304 с.; Широков В. А., Білоноженко В. М., Бугаков О. В. та ін. Лінгвістичні та технологічні основи тлумачної лексикографії. — К., 2010. — 296 с.; Широков В. А. Комп'ютерна лексикографія. — К., 2011. — 352 с.

Елементарною називається така Л-система, де діє (проявляється) лише один лексикографічний ефект. Лексикографічний опис $V(x)$ елементарної інформаційної (реєстрової) одиниці x формально задається множиною «слів» (текстів), представлених лінійними ланцюжками символів, які взято з певного скінченного алфавіту A . У словнику лексикографічний опис $V(x)$ фактично є текстом словникової статті з реєстровою одиницею x . Тоді $V(I^Q(D))$ є сукупністю всіх словникових статей. На множині описів $V(I^Q(D))$ (і, отже, на кожному $V(x)$) визначаються дві структури — β і $\sigma[\beta]$. Вони, по суті, задають семантику лінгвістичних фактів і закономірностей, які містить побудована в такий спосіб Л-система. При цьому β — множина «найпростіших» структурних елементів словника (слова, аббревіатури, поміти, ремарки, нумерації, елементи граматичного та семантичного описів і т. ін.). Структури β , як правило, задані в словнику експліцитно. Формально для кожного $x \in I^Q(D)$ визначається множина $\beta(x)$ «підслів» з $V(x)$ з такими властивостями:

1. $x \in \beta(x)$;
2. Будь-який фрагмент словникової статті $V(x)$ можна побудувати з елементів, що належать до $\beta(x)$;
3. Правило формування елементів $\beta(x)$ є єдиним для всіх $V(x)$, тобто для всіх $x \in I^Q(D)$.

Правила 1–3 формування структур β відіграють важливу роль у лексикографії. Так, правило 2 фактично є вимогою універсальності метамови словника, настановою на те, що будь-який факт з розглядуваного універсуму лінгвістичних явищ, що підлягають лексикографуванню в конкретному словнику, повинен відбиватися в його метамові. Із правила 3 випливає вимога, щоб усі однотипні мовні факти і феномени, що розглядаються, мали уніфіковане представлення в лексикографічному описі. Зазначені правила становлять об'єктивні передумови для формалізованого визначення процесу лінгвістичного дослідження за допомогою лексикографічної системи.

Із елементів β будуються агреговані лексикографічні структури $\sigma[\beta]$, які вже містять змістові твердження про лінгвістичні властивості розглядуваної сфери мовних явищ. Структури $\sigma[\beta]$ можна наочно проілюструвати, використовуючи одну з найуніверсальніших моделей даних, а саме — реляційну⁴. Здійснюється це так. Припустимо, що в певному словнику (або Л-системі) визначено структуру β , яка складається з певного скінченного числа q різних структурних елементів, а саме:

$$\beta = \{\beta_1, \beta_2, \dots, \beta_q\}. \quad (2)$$

При цьому кожний з елементів $\beta_i, i=1, 2, \dots, q$ набуває певного конкретного значення на кожному з елементів реєстру x , які належать до $I^Q(D)$. Так визначається текстове представлення $\beta_i(x)$ для всіх $x \in I^Q(D)$ та $i=1, 2, \dots, q$. При цьому не виключені випадки, коли деякі з $\beta_i(x)$ для певних x можуть бути порожніми. Побудуємо таблицю:

β_i	β_2	...	β_q
$\beta_1(x_1)$	$\beta_2(x_1)$...	$\beta_q(x_1)$
$\beta_1(x_2)$	$\beta_2(x_2)$...	$\beta_q(x_2)$
$\beta_1(x_M)$	$\beta_2(x_M)$...	$\beta_q(x_M)$

⁴ Цикритзис Д., Лоховски Ф. Модели данных.— М., 1985.— 168 с.

Елементи $\beta_1, \beta_2, \dots, \beta_q$ можна інтерпретувати як імена атрибутів певної реляційної моделі даних. Значення цих атрибутів, що їх останні приймають на реєстрових *одиницях* x_1, x_2, \dots, x_M , тобто величини: $\beta_{ij} \equiv \beta_i(x_j), i=1, \dots, q; j=1, \dots, M$, інтерпретуємо як домени атрибутів $\beta_1, \beta_2, \dots, \beta_q$. Лексикографічна модель даних у такий спосіб інтерпретується через реляційну модель даних, причому структури, що належать до $\sigma[\beta]$, відіграють роль відношень цієї моделі, або елементів відповідної алгебри Кодда⁵. Таким чином, будь-яку реляційну модель можна вважати частинним випадком лексикографічної моделі даних. Зокрема, через певні елементи з $\sigma[\beta]$ визначаються ліва і права частини словникових статей, які відображають відношення «форма — зміст» для цього словника; відповідні елементи з $\sigma[\beta]$ (позначатимемо їх символом LP) є обов'язковими, нефакультивними елементами лексикографічної моделі даних. Крім цього, структура $\sigma[\beta]$ містить велику кількість спеціальних відношень, що визначають семантику предметної галузі, яка є об'єктом словника. Багато (якщо не більшість) таких елементів структури $\sigma[\beta]$ у словниках, як правило, задано імпліцитно і їх установлення та змістова лінгвістична інтерпретація є справою вміння, досвіду та мистецтва дослідника.

Отже, загальну форму елементарної лексикографічної моделі даних (Л-системи) тепер можемо записати так:

$$LM = \{D; S; Q; I^Q(D), V(I^Q(D)), \beta, \sigma[\beta]\}, \quad (3)$$

причому необхідність експліцитного подання відношення «форма — зміст» веде до виокремлення в структурі $\sigma[\beta]$ відношень LP та, як наслідок, до індукування комутативної діаграми:

$$\begin{array}{ccc}
 & V(I^Q(D)) & \\
 F \swarrow & & \searrow C \\
 \Lambda(I^Q(D)) & H & P(I^Q(D))
 \end{array} \quad (4)$$

де $HF = C$, причому ця діаграма справедлива для $\forall x \in I^Q(D)$, тобто таку діаграму можна створити для кожної словникової статті.

Застосовуючи викладене до академічного тлумачного Словника української мови, приходимо до розбудови структури його Л-системи. Для цього необхідно передусім побудувати для нього основну діаграму, тобто визначити об'єкти $\Lambda(W^{SUM}), P(W^{SUM}), F, C$ і H . Це означає, що потрібно визначити будову об'єктів $\Lambda(x)$ і $P(x)$ для всіх $x \in W^{SUM}$, де символом W^{SUM} позначено клас лексем української мови. У структурі $\Lambda(x)$ виділяються такі елементи:

$R(x)$ — Реєстровий ряд. Він складається із заголовкового слова x в канонічній формі, а також ряду інших слів, що мають із заголовковим словом загальний комплекс лексичних значень. Такі послідовності слів називаємо *еквісемантичними рядами*. Отже, мінімальний реєстровий ряд складається з одного лише заголовкового слова, яке є реєстровою одиницею словникової статті $V(x)$.

Елементи реєстрового ряду об'єднуються в групи, що належать до одного словозмінного класу (для повнозначних змінюваних частин мови). Після

⁵ Codd E. F. A Relational Model of Data for Large Shared Data Banks // Communications of the ACM. — New York, 1970. — Vol. 13. N. 6. — P. 377–387.

кожної такої групи в тексті словникової статті подається набір парадигматичних показників, що ідентифікують словозмінний клас, до якого належить ця група. Крім того, після кожного елемента реєстрового ряду наводиться його фонетичний варіант, якщо такий є. При потребі зазначаються показники синтаксичних зв'язків і функцій, граматичні параметри, стилістичні та інші ремарки. Зокрема, для однослівного реєстрового ряду загальна форма $\Lambda(x)$ має вигляд:

$$x, \text{Пар}(x), \text{Грам}(x), \text{Рем}(x), \text{Синт}(x), \quad (5)$$

де x — реєстрова одиниця; $\text{Пар}(x)$ — парадигматичні показники для змінюваних слів; $\text{Грам}(x)$ — граматичні параметри (рід для іменників, вид для дієслова, показник частини мови, що подається для займенників, назв числівників, прислівників, службових частин мови і вигуків); $\text{Рем}(x)$ — ремарки різних типів; $\text{Синт}(x)$ — показники синтаксичних зв'язків і функцій.

Досліджуючи будову $P(x)$, виводимо, що вона складається з кількох вкладених одна в одну підструктур, ідентифікація яких пов'язана з певними лінгвістичними феноменами.

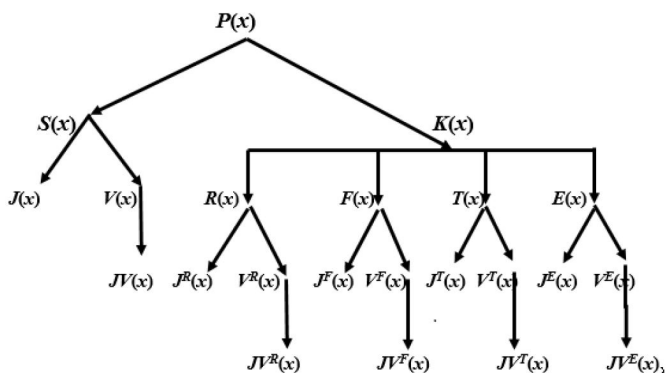
У структурі правої частині $P(x)$ словникової статті СУМ відображаються багатозначність реєстрової одиниці, а також її колокації (словосполучення, в яких представлена реєстрова одиниця). Багатозначність має двоступеневий характер і складається з рубрик, які представляють лексичні значення відповідних лексем та їх відтінки. Усі структурні ефекти, які ми тут описуємо, формально визначені в тексті словника за допомогою спеціальних роздільників, позиційно, шрифтовими виділеннями і т. ін. Ці елементи метамови словника також належать до структури β . Тому побудова формального репрезентанта структури може бути повністю алгоритмізована.

Найбільшим з елементів структури правої частини є той, що відтворює ефект багатозначності реєстрової одиниці (полісемія). Отже, полісемія: S_i — це частина $P(x)$, що містить семантичний опис i -того лексичного значення реєстрової одиниці x . Звідси довільну $P(x)$ можна представити як об'єднання:

$$P(x) \supseteq \bigcup_i S_i, S_i \cap S_j = \emptyset \text{ при } i \neq j. \quad (6)$$

У кожній рубриці S_i розміщуються дефініції та ілюстрації i -того лексичного значення (тобто приклади його вживання в літературному тексті — вони позначаються символом $J(i, q)$), а також фрагменти словникової статті, що відбивають відтінки значення. У поліграфічному оформленні вони представлені двома скісними лініями «//»; відтінки значення S_i позначаємо через $V(i, j)$. До кожного відтінку також подається одна або кілька ілюстрацій.

Аналогічною структурою володіє система словосполучень (колокацій). У СУМ-20 представлено словосполучення чотирьох типів: стійкі словосполучення, власне фразеологізми, термінологічні словосполучення та еквіваленти слова. Кожний з цих чотирьох типів у тексті СУМ-20 забезпечено ідентифікувальним маркером. Формально структуру $P(x)$ представлено у вигляді графа:



де символом $P(x)$ позначено праву частину словникової статті з реєстровою одиницею x ; відповідно: S — комплекс рубрик лексичного значення; K — комплекс колокацій; J — комплекс ілюстрацій всіх категорій; V — комплекс відтінків значень; R — комплекс стійких словосполучень; F — комплекс фразеологізмів; T — комплекс термінологічних словосполучень; E — комплекс еквівалентів слова. Стрілками на графі позначено відношення належності: $A \rightarrow B \Rightarrow A \supseteq B$.

Зауважимо, що всі структурні елементи та зв'язки в Л-системі «Словника української мови» є певними елементами його структури (β , $\sigma[\beta]$).

Використання ВЛЛ «Словник української мови» у лінгвістичних дослідженнях

В Українському мовно-інформаційному фонді НАНУ модель Л-системи «Словника української мови» реалізована у вигляді так званої Віртуальної лексикографічної лабораторії, яка використовується як інструментальна система для укладання «Словника української мови» у 20-ти томах, а також для проведення лінгвістичних досліджень. Адже ВЛЛ є інтегрованою лексикографічною системою, до складу якої входять, крім тлумачного словника як головного джерела семантичної інформації, електронний граматичний словник (ЕГС), що виконує в ній роль інтегративного компонента (за допомогою ЕГС здійснюється перехід від текстових слів до відповідних вихідних форм у реєстрі тлумачного словника), а також інші лексикографічні продукти — електронний словник паронімів, електронні синонімічний та антонімічний словники, електронний словник дієслівного керування, Український лінгвістичний корпус обсягом близько 200 млн слововживань.

На сьогодні тут одержано цілу низку нових наукових результатів, які навряд чи були б досяжними у традиційному дослідницькому модусі. Наведемо деякі з них. У підрозділах 1–3 використовуємо матеріал з монографій⁶.

1. Закони будови дієслівних та іменникових еквісемантичних рядів. У результаті дослідження будови тлумачного «Словника української мови» було встановлено, що структура дієслівних словникових статей підкоряється певній формальній закономірності. Як з'ясувалося, існує формальний

⁶ Широков В. А. Комп'ютерна лексикографія.— К., 2011.— 352 с.; Рабулець О. Г., Сухарина Н. М., Широков В. А., Якименко К. М. Дієслово в лексикографічній системі.— К., 2004.— 259 с.

механізм, або закон, який «породжує» структури граматичної семантики українського дієслова, цьому закону підкоряються всі без винятку структури словникових статей СУМ.

Якщо розглядати тексти лівих частин $\Lambda(x)$ словникових статей із заголовковими словами x , де x пробігає множину дієслів української мови, як лінійні ланцюжки символів, можемо помітити, що в структурі цих ланцюжків виділяються певні підланцюжки, які мають характер вкладень, причому ці вкладення виділяються певним інваріантним способом, коли різні підланцюжки не перетинаються один з одним.

Відзначена закономірність будови ланцюжків $L(x)$ є такою регулярною, що дозволяє сформулювати її навіть в аксіоматичній формі у вигляді таких тверджень:

1. Кожне українське дієслово реалізується в мові лексемою (з визначеною й зафіксованою лексичною семантикою), якій властиві одне або два значення категорії виду.

Лексикографічне представлення належності лексеми до певного значення атрибута «ВИД» (як воно подано в СУМ) назвемо *видовим комплексом*, або просто *комплексом*, і позначатимемо його символами C_1 та C_2 . За наявності одного видового комплексу маємо вкладення першого рівня $C_1 \subset \Lambda(x)$, яке графічно можна представити так:



де відношення вкладення « \subset » показано стрілкою. Аналогічно двокомплексна $L(x)$ графічно має такий вигляд:



Так, лексичне значення «Погоджуючись зі співрозмовником або підтверджуючи його слова, говорити часто “ага”»⁷ реалізується дієсловом **АГА#КАТИ** тільки у формі недоконаного виду; значення «Довго кличучи, змусити кого-небудь відізватися або прийти; догукатися»⁸ реалізується дієсловом **ДОКЛИ#КАТИСЯ** лише доконаного виду; значення «Прибудовувати зверху чого-небудь»⁹ може реалізовуватися і дієсловом **НАДБУДО#ВУВАТИ** недоконаного виду і **НАДБУДУВА#ТИ** доконаного виду; у значенні «Набувати носового відтінку (про звуки мови)»¹⁰, дієслово **НАЗАЛІЗУВА#ТИСЯ**, залежно від контексту, може бути як доконаного, так і недоконаного виду.

⁷ Віртуальна лексикографічна лабораторія «Словник української мови». Український мовно-інформаційний фонд НАН України, 2001–2018. — <http://lcorp.ulif.org.ua/LSlist/>.

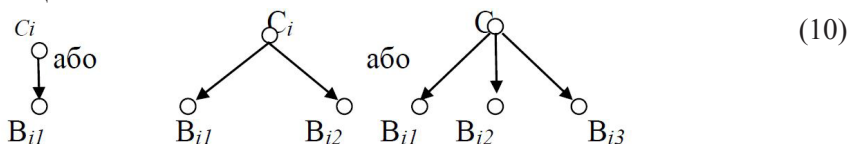
⁸ Там же.

⁹ Там же.

¹⁰ Там же.

2. Кожна дієслівна лексема з визначеною і зафіксованою лексичною семантикою та конкретним значенням виду може реалізуватися дієсловами, що належать не більше як до трьох різних словозмінних (парадигматичних) класів.

Ідентифікаторами словозмінних класів у СУМ виступають парадигматичні показники, якими, як правило, є набори квазіфлексій першої та другої особи однини теперішнього часу; якщо їх виявляється недостатньо для визначення системи словозміни, наводяться деякі інші форми. Так чи інакше, але набір парадигматичних показників має бути достатнім для ідентифікації словозмінної парадигми конкретної лексеми. Для лексикографічного представлення цих фактів уживатимемо термін *парадигматичний блок*, або просто *блок*, який позначатимемо символами B_{ij} , де індекс «i» відповідає номеру видового комплексу, до якого належить даний блок, і набуває значення 1, 2, а індекс «j» нумерує блоки всередині видового комплексу і набуває значення від 1 до 3. Таким чином, кожному видовому комплексу належить один, два або три словозмінні блоки. Графічно можемо зобразити це так:



3. Кожна дієслівна лексема з конкретним значенням виду, виявленою належністю до конкретного словозмінного класу та визначеною й зафіксованою лексичною семантикою може реалізуватися дієсловами, які мають не більше чотирьох фонетичних варіантів (як правило, це префіксальна та коренева варіація, але трапляються й суфіксальні).

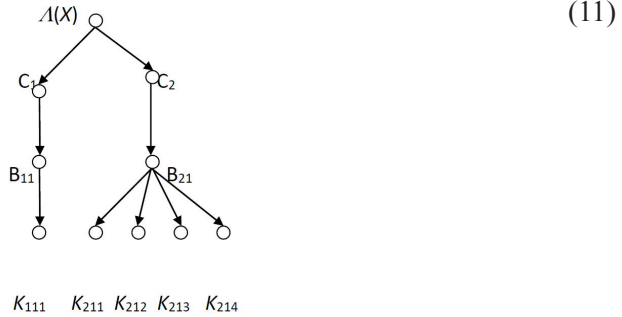
Для лексикографічного представлення цього факту вживатиметься термін *компонент*. Компоненти позначаються символами K_{ijr} , де індекс «i» відповідає номеру видового комплексу, до якого належить блок з номером «j», а індекс «r» пробігає значення від 1 до 4 залежно від того, скільки компонентів має розглядуваний блок.

Щоб проілюструвати ефект компонентності, розглянемо таку словникову статтю:

$A(x) =$ **ЗВИКА#ТИ**, рідко **ІЗВИКА#ТИ**, -аю, -аєш, *недок.*, **ЗВИ#КНУТИ**, рідко **ІЗВИ#КНУТИ**, **ЗВИ#КТИ**, рідко **ІЗВИ#КТИ**, -кну, -кнеш; *мин. ч.* звик, ізв'ик, -ла, -ло і зви#кнув, ізви#кнув, -нула, -ло; *док*¹¹. У цій статті є два видові комплекси. У першому комплексі: $C_1 =$ **ЗВИКА#ТИ** і рідко **ІЗВИКА#ТИ**, -аю, -аєш є всього один блок: $B_{11} =$ **ЗВИКА#ТИ** і **ІЗВИКА#ТИ** рідко, до якого, у свою чергу, належать два компоненти: $K_{111} =$ **ЗВИКА#ТИ**; $K_{112} =$ **ІЗВИКА#ТИ**. У другому комплексі: $C_2 =$ **ЗВІКНУТИ** і рідко **ІЗВІКНУТИ**, **ЗВІКТИ** і рідко **ІЗВІКТИ**, -кну, -кнеш; *мин. ч.* звик, ізв'ик, -ла, -ло і зви#кнув, ізви#кнув, -нула, -ло також є всього один блок: $B_{21} =$ **ЗВИКНУТИ**, і рідко **ІЗВІКНУТИ**, **ЗВІКТИ** і рідко **ІЗВІКТИ**, але в ньому наявна максимальна кількість компонентів — чотири, а саме: $K_{211} =$ **ЗВІКНУТИ**; $K_{212} =$ **ЗІКНУТИ**; $K_{213} =$ **ЗВІКТИ**; $K_{214} =$ **ІЗВІКТИ**.

¹¹ <http://lcorp.ulif.org.ua/LSlist/>.

Графічно структуру цієї словникової статті представлено так:



З описаних властивостей композиції дієслівних $L(x)$ випливає висновок про існування особливої системної лінгвістичної закономірності в їх структурі, яку формально подано послідовністю вкладень:

$$L(x) \supset C \supset B \supset K, \quad (12)$$

де символом C позначено сукупність видових комплексів, що належать до даного $L(x)$ (не більше двох комплексів), символом B — сукупність парадигматичних блоків (не більше трьох блоків у кожному комплексі), K — сукупність компонентів (не більше чотирьох компонентів у кожному блоці)¹².

Під час дослідження будови еквісемантичних рядів¹³ виокремлено кілька прикладів цих рядів різної кратності із «Словника української мови». Як зазначалося у згаданій праці, кількість членів рядів у СУМ варіювала від 1 до 8, причому рядів із кратністю 7 у Словнику тоді не було виявлено. Цей факт на час проведення дослідження видавався дещо незрозумілим. Але при написанні цієї статті було здійснено ретельніший аналіз структури словникових статей за допомогою ВЛЛ, причому було виявлено реєстрові ряди з кратністю 7. Наприклад:

ВЛИВА#ТИ1 (УЛИВАТИ), аю, аєш, *недок.*, **ВЛІТИ (УЛИ#ТИ),**
ВВІЛЛЯТИ (УВІЛЛЯТИ, УЛЛЯ#ТИ), увілляю, уві#ллєш і улляю, уллєш,
док., що, у що. 1. Лити що-небудь кудись, у середину чогось.
ВЛИВА#ТИСЯ1 (УЛИВАТИСЯ), аюся, аєшся, *недок.*, **ВЛІТИСЯ**
(УЛИ#ТИСЯ), **ВВІЛЛЯТИСЯ (УВІЛЛЯТИСЯ, УЛЛЯ#ТИСЯ),**
 увілляюся, уві#ллєшся і улляюся, уллєшся, *док.* 1. Литися в що-небудь,
 куди-небудь.

Цей факт підтверджує правильність закону «1-2-3-4», а також думку про те, що структура Л-системи «Словника української мови» цілком відповідає лінгвістичним законам українського дієслова, дозволяє простежити різні типи варіативності слів української мови і задовольняє загальну вимогу лексикографічної повноти і замкненості великих тлумачних лексиконів.

Теорія Л-системи «Словника української мови» є також універсальним інструментом у виявленні закономірностей морфологічної (граматичної) структури лексем іншої частини мови — іменника. Дослідження корпусу іменникових словникових статей «Словника української мови» дозволяють зробити висновок про те, що структура $\Lambda(x)$ формується на основі

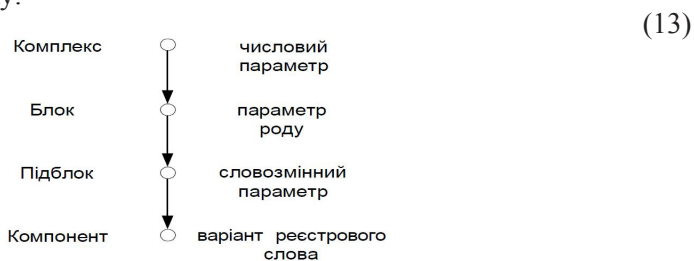
¹² Широков В. А. Системна семантика тлумачних словників...

¹³ Див.: Рабулець О. Г., Сухарина Н. М., Широков В. А., Якименко К. М. *Знач. праця.*

морфосемантичних властивостей елементарних інформаційних одиниць. Ними є заголовне слово, його лексемні варіанти та лексикографічні параметри (позначення роду, словозмінного класу, числа та функціонального стилю), за допомогою яких здійснюється опис граматичної семантики іменників.

Як виявилось під час дослідження, функціонування елементарних структуротвірних компонентів $\Lambda(x)$ не є випадковим і підпорядковується загальним граматичним законам мови, на основі яких творяться типи структур граматичної семантики українських іменників. Закономірне породження таких структур виявляється в ієрархічній організації структуротвірних компонентів $\Lambda(x)$ ¹⁴.

Згідно із засадами теорії Л-системи «Словника української мови», ієрархізацію параметрів граматичної семантики іменників подаємо у вигляді схеми, на якій показано співвідношення формальних граматичних параметрів та умовно названих структурних елементів, розташованих в ієрархічному порядку:



Як бачимо, існує закономірність для всього іменникового корпусу, за якою текст $\Lambda(x)$ для будь-якого x з множини українських іменників однозначно подано у вигляді системи вкладень:

$$\Lambda(x) \supset \mathcal{C} \supset \mathcal{P} \supset \mathcal{I} \supset \mathcal{K}, \quad (14)$$

де символом \supset позначено відношення належності, а кожний з чотирьох елементів \mathcal{C} , \mathcal{P} , \mathcal{I} , \mathcal{K} являє собою певну частину $\Lambda(x)$, яка має чітко визначену лінгвістичну кваліфікацію: \mathcal{C} — відображення параметра на позначення числа, \mathcal{P} — відображення параметра на позначення роду, \mathcal{I} — відображення параметра на позначення парадигматичного (словозмінного) класу, \mathcal{K} — відображення параметра на позначення реєстрового слова (та його варіанта)¹⁵.

2. Дослідження формул квазісемантики та визначення ядра лексичної системи української мови. З-поміж типів словникових тлумачень у «Словнику української мови» виокремлюється група дефініцій, в яких тлумачення слова здійснюється через посилання на словникову статтю з близьким за семантикою реєстровим словом, що перебуває з означуваним у певному дериваційному відношенні. Характерним для дефініцій цього типу є набір формул тлумачення — своєрідних семантичних кліше, які маркують тип відношення між означуваним словом та його відповідником у дефініції; такі формули не містять семантичних складників (архі- та диференційні

¹⁴ Відповідні результати наводимо за працею: Погрібна О. О., Чумак В. В., Широков В. А., Шевченко І. В. Лінгвістична класифікація українського іменника у світлі теорії лексикографічних систем // Мовознавство. — 2004. — № 5–6. — С. 62–82.

¹⁵ Докладніше про дослідження властивостей іменника в дослідницькій парадигмі ВЛЛ СУМ див. у кн.: Погрібна О. О. Лексична та граматична семантика іменників у «Словнику української мови» (назви знярядь праці та споруд) : Дис. ... канд. філол. наук. — К., 2005. — 244 с.

семи), у них часто використовується спосіб тлумачення лексичної семантики через граматичну (іменник до ..., прикметник до ..., прислівник до ...). За допомогою сервісу ВЛЛ СУМ таких формул квазісемантики¹⁶ нами виявлено 35 (див. рис.).

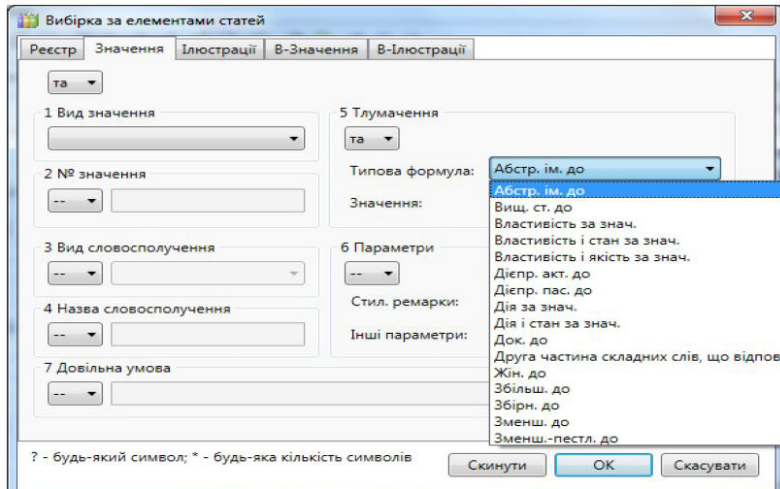


Рис. Вікно з ВЛЛ СУМ для формування файлів лексем, тлумачення яких здійснено за формулами квазісемантики

Аналіз формул квазісемантики засобами ВЛЛ СУМ дозволяє простежити системність, що пов'язано з визначенням семантично і граматично первинної та похідної лексики в мові. Наприклад, у формулах тлумачення:

«Абстр. ім. до ...» **АГІТАЦІЙНІСТЬ**, ності, жс. Абстр. ім. до **агітаційний**, «Властивість за знач. ...» **ЕКОЛОГІЧНІСТЬ**, ності, жс. Властивість за знач. **екологічний**, «Властивість і якість за знач. ...» **НАДПОТУЖНІСТЬ**, ності, жс. Властивість і якість за знач. **надпотужний**, «Стан за знач. ...» **БАДЬОРІСТЬ**, рості, жс. Стан за знач. **бадьорий**, «Стан і властивість за знач. ...» **ХАОТИЧНІСТЬ**, ності, жс. Стан і властивість за знач. **хаотичний**, «Якість за знач. ...» **КОМФОРТНІСТЬ**, ності, жс. Якість за знач. **комфортний**, «Якість і властивість за знач. ...» **БРЯЗКІТЛИВІСТЬ**, вості, жс. Якість і властивість за знач. **брязкітливий**¹⁷ домінує словотвірна модель «X + ість», де X — представляє базову частину деривата.

До цих моделей варто додати й інші лінгвістичні механізми, зокрема, моделі типу «Префікс + X», де X — певні іменники, дієслова, прикметники, до яких можуть додаватися префікси **а-, ви-, від-, до-, з-, за-, на-, од-, по-, про-, роз-, с-, у-** тощо.

Навіть побіжний аналіз окреслених моделей дозволяє поставити питання про визначення ядра лексичної системи української мови і кваліфікацію й кодифікацію породжувальних механізмів, що діють у цій системі. Зокрема, доцільно дослідити, якими афіксальними моделями забезпечується функціонування формул квазісемантики, а також навпаки: які афіксальні словотвірні моделі беруть участь у формуванні лексики, що належить до певного квазісемантичного типу. При цьому «Словник української мови» у формі ВЛЛ СУМ, безперечно, слугуватиме ефективним інструментом для

¹⁶ Широков В. А. Системна семантика тлумачних словників... — С. 487–510.

¹⁷ <http://lcorp.ulif.org.ua/LSlist/>.

визначення лексичного ядра української мови, означення й формалізації породжувальних механізмів на класі українських лексичних одиниць.

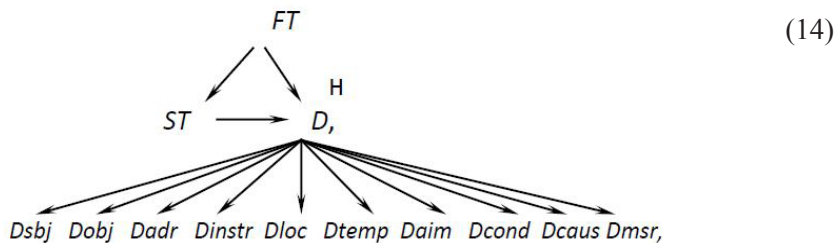
Фактично йдеться про встановлення певної формалізованої системи на класі українських слів зі структурою $Я + ПГ$, де символом $Я$ позначено, власне, лексичне ядро української мови, а $ПГ$ означає породжувальну граматику, дія елементів якої на $Я$ визначає весь клас українських слів.

Застосування такої структури в лінгвістичних технологіях дозволило б значно ущільнити обсяг лінгвістичних даних і, отже, збільшити ефективність алгоритмів опрацювання лінгвістичної інформації. На наш погляд, розв'язання цього важливого завдання практично неможливе без застосування потужних інструментальних засобів, які надає ВЛІ СУМ.

3. Дослідження структури та виведення лінійного формату для словникових дефініцій українського дієслова. Під час дослідження структури дефініцій шляхом аналізу тлумачень дієслів Н. М. Сухариною було виділено типи диференційних сем, що характерні для дієслівних формул тлумачення та встановлено типові форми їх представлення¹⁸. Виявлено та ідентифіковано набір з 10-ти диференційних сем: $Dsbj$ — «суб'єкт»; $Dobj$ — «об'єкт»; $Dadr$ — «адресат»; $Dinstr$ — «спосіб, засіб»; $Dloc$ — «середовище, місце»; $Dtemp$ — «час»; $Daim$ — «мета»; $Dcond$ — «умова»; $Dcaus$ — «причина»; $Dmsr$ — «міра, інтенсивність». Зазначений набір є інваріантним, унаслідок чого будь-яка дієслівна формула тлумачення розкладається у лінійний (табличний) формат з певною структурою:

Лінійний формат дефініцій дієслова										
семантична сема	суб'єкт	об'єкт	адресат	спосіб, засіб	середовище, місце	час	мета	умова	причина	міра, інтенсивність
1	2	3	4	5	6	7	8	9	10	11

Графічно це зображено у вигляді такої схеми:



де FT — множина формул тлумачення, що розглядається як спеціальна лексикографічна система;

ST — множина семантичних тем, що відіграє роль реєстрової (формальної) частини лексикографічної системи;

D — множина диференційних сем, що виступають інтерпретаційною (змістовою) частиною FT ;

H — оператор, що здійснює зв'язок між ST і D .

Отже, структуру узагальненої формули тлумачення для дієслова можна передати так:

$$FT = ST + D = ST + Dsbj + Dobj + Dadr + Dinstr + Dloc + Dtemp + Daim + Dcond + Dcaus + Dmsr. \quad (15)$$

¹⁸ Див.: Сухарина Н. М. Граматична та лексична семантика українського дієслова в лексикографічній системі : Автореф. дис. ... канд. філол. наук.— К., 2003.— 19 с.

Порядок розташування компонентів тут не відображено (як і синтаксичну структуру формули тлумачення).

Запропонований формат дієслівних формул тлумачення виступає простішим об'єктом, ніж словникова дефініція, що робить його доступним для автоматизованого дослідження. Переведення словникових тлумачень до ідеографічно-таксономічного формату дозволило виявити скінченні списки універсальних семантичних ознак різного походження, що співвідносяться з певним атрибутом формату (семантичною темою чи диференційними семами «суб'єкт», «об'єкт», «адресат», «спосіб, засіб», «середовище, місце», «час», «мета», «причина», «умова», «міра, інтенсивність») і служать елементами метамови лексикографічного опису лексичної семантики усіх лексико-семантичних варіантів (ЛСВ) дієслова.

Саме виділені атрибути формату дають можливість побудувати багатопараметричні лексико-граматичні класифікації дієслова. Загалом ці класифікації можна об'єднати у три групи:

1) за атрибутами (одним чи кількома) ідеографічно-таксономічного формату;

2) за конкретними значеннями, які може приймати атрибут (один чи кілька) ідеографічно-таксономічного формату;

3) за атрибутом (одним чи кількома) та значеннями іншого (чи інших) атрибутів ідеографічно-таксономічного формату.

Табличний формат представлення словникових дефініцій може служити ефективним засобом у лексикографічній практиці, адже сприятиме встановленню логічних хиб при побудові тлумачень, зокрема таких, як:

– цикли у тлумаченнях (ця хиба полягає в тому, що слово, яке виступає одним із компонентів формули тлумачення, саме може тлумачитися лише через це витлумачуване слово);

– несумірність обсягів понять (помилка полягає в тому, що обсяг тлумачення не дорівнює обсягу витлумачуваного);

– логічні протиріччя (помилка зумовлена тим, що значення слів-компонентів тлумачення даного слова суперечать один одному);

– «перерване» тлумачення (у випадку, коли серед слів, які використовуються у метамові тлумачень, є такі, що відсутні у словнику як реєстрові).

Запропонований лінійний формат для інтерпретаційної частини словникових статей, поширений на весь корпус дієслів, розроблених у «Словнику української мови», на наш погляд, є ефективним засобом не лише для мовознавчих студій або для вдосконалення самої лексикографічної продукції, але й для майбутніх автоматизованих систем розуміння природномовних текстів, у яких неодмінним складником є лінгвістичне забезпечення¹⁹. Зрозуміло, що теорія має бути поширена і на інші лексико-граматичні класи української мови. У наступному підрозділі це завдання ми частково розв'язуємо для класу українського прикметника.

4. Дослідження семантики відносних прикметників з похідною іменниковою основою засобами ВЛЛ СУМ. Ефективність використання ВЛЛ СУМ як інструменту семантичних досліджень залежить значною мірою від того, наскільки формули тлумачень забезпечують виконання

¹⁹ Рабулець О. Г., Сухарина Н. М., Широков В. А., Якименко К. М. Зазнач. праця.

загальної вимоги, висунутої до електронних словників щодо експліцитності представлення в них інформації²⁰.

Для електронного тлумачного словника ця вимога інтерпретується як реалізація формулою тлумачення такої форми представлення в тлумачній статті інформації про лексичне значення, яка б дозволяла за допомогою комп'ютерних програм встановлювати схожість / відмінність семантичних характеристик текстової словоформи і слова із словника, будувати в автоматизованому режимі певні семантичні класифікатори слів, передавати вилучену інформацію щодо семантичних властивостей слів на входження лінгвістичних аналізаторів.

Проведений з цього погляду попередній аналіз структури дефініцій у тлумачному словнику показав: у більшості з них застосовано спосіб опису семантичних характеристик слова через гіперонімічні зв'язки в лексичній системі мови, які є відображенням відношень логічного підпорядкування «вид — рід» або «частина — ціле», що існують між позначуваними референтами. Наприклад:

БАМБУ#К, а, ч. Висока гнучка тропічна і субтропічна деревоподібна злакова рослина з порожнистим колінчастим стеблом і видовженим листям²¹ (відношення «рід — вид» «рослина — бамбук»).

Із урахуванням гіперонімічних зв'язків будують семантичні класифікації, зорієнтовані на формування тезаурусів (родо-видові відношення належать до основного типу зв'язків, встановлюваних між тезаурусними одиницями). Встановлення гіперонімічних зв'язків передбачено в компонентному семантичному аналізі. За цим принципом будують і онтології предметних галузей.

Отже, спосіб виявлення лексичного значення слова на основі його гіперонімічних зв'язків можна віднести до типових формул експліцитного тлумачення слова. Цей спосіб варто запропонувати для опису лексичної семантики нових слів, що з'являються в мові, під час введення їх до тлумачних чи енциклопедичних словників, а також для уточнення тлумачень в електронних тлумачних словниках, заданих неекспліцитними формулами представлення лексичного значення у вихідному тлумачному словнику. До останніх за результатами попереднього аналізу можуть бути віднесені формули тлумачення, які передбачають виявлення значення слів опосередковано — шляхом встановлення у мовленні їхніх словотвірних зв'язків, синонімічних чи антонімічних і навіть синтагматичних.

Частина формул тлумачення, а саме відсильні формули «те саме, що...»:

АКО#РДОВИЙ, а, е. Те саме, що **акордний1.**, «див...» **АБХА#ЗЕЦЬ** див. **абха#зці**²²,

формули, за якими значення слова описують через встановлення синонімічних відношень, можуть уважатися умовно неекспліцитними, тому що сервіс інтегрованої лексикографічної системи дозволяє в лінгвістичній

²⁰ Грязнухина Т. А., Любченко Т. П., Якименко К. М. Эксплицитное представление семантической информации в Электронном толковом словаре украинского языка (синонимическая и паронимическая параметризация) // Прикладная лингвистика и лингвистические технологии: MegaLing-2012. — К., 2012. — С. 79–91.

²¹ <http://lcorp.ulif.org.ua/LSlist/>.

²² <http://lcorp.ulif.org.ua/LSlist/>.

базі електронного тлумачного словника автоматично продублювати тлумачення відповідних слів.

До формул тлумачення, для яких експліцитний опис семантики слова зводиться до завдання розроблення тлумачних статей нових слів, належать формули тлумачення відносних прикметників з похідною іменниковою основою «Прикметник до...» та «Стосовно до...». Зазначені формули встановлюють існування зв'язку значення прикметників з лексичним значенням іменників, від яких ці прикметники утворено. Саме цей зв'язок і повинен насамперед бути представленим в експліцитному вигляді в тлумаченнях відносних прикметників.

Другою семантичною властивістю цієї групи прикметників є реляційний характер позначуваної ними ознаки. Така властивість належить до іманентних властивостей відносних прикметників, вона розрізняє їх та якісні прикметники і, отже, також має бути обов'язково відображена експліцитно в лексичному значенні представників цієї групи.

Експліцитним представленням реляційної ознаки, вираженої відносними прикметниками, є семантичні відношення як відображення ситуативних зв'язків, у які можуть вступати референти, позначувані іменниками, від яких утворені прикметники, і референти, позначувані іменниками, синтаксично зв'язаними з цими прикметниками в сполученнях «відносний прикметник + іменник» (A + I).

Можливість і характер ситуативних зв'язків, у свою чергу, зумовлені іманентними властивостями предметів, що вступають у певні відношення, а також характером реальних ситуацій, у яких ці предмети перебувають чи функціонують.

Таким чином, через включення семантичних відношень, які можуть позначатися відносним прикметником у сполученнях A + I, до структури його лексичного значення експлікується певна співвіднесеність останнього з позамовною дійсністю.

Для експлікації семантичних відношень використаємо класифікацію, прийняту в інформаційно-пошукових системах (у предметних онтологіях), так як її подано, наприклад, у словнику Г. А. Золотової²³. Класифікація представлена групами семантичних відношень: *бути виготовленим з* (феритове осереддя), *відношення «ціле — частина»* (атомне ядро), *посесивні* (двірницьке пальто), *просторові* (далматинський острів), *темпоральні* (декабристська доба) та ін.

Кількість значень (відтінків у межах значення) конкретного прикметника перебуває в прямій залежності від кількості семантичних відношень, які ним можуть бути реалізовані в сполученні з іменником.

Визначальною для конкретних типів семантичних відношень є інформація про властивості предметів реальної дійсності, позначуваних прикметником та іменником, що зумовлюють їхню здатність вступати в певні ситуативні зв'язки — відповідники цих семантичних відношень. На мовленнєвому рівні ця зумовленість формально виявляється в лексичній сполучуваності відносних прикметників.

Через те, що тут ідеться про врахування властивостей предметів позамовної дійсності через їхні лексичні значення, доцільно лексичну

²³ Золотова Г. А. Синтаксический словарь : Репертуар элементар. единиц рус. синтаксиса.— М., 1998.— 440 с.

сполучуваність описувати в термінах семантичних класів слів, а не окремих слів. До одного семантичного класу належать слова, які мають спільні ознаки в їхньому сигніфікативному значенні.

Урахування параметрів, за якими повинен здійснюватися опис значення відносних прикметників з похідною іменниковою основою в експліцитному вигляді, а також значний обсяг досліджуваного матеріалу вимагають від дослідника застосування операційного способу вирішення поставленого завдання, узгодженого з експліцитним описом значення в електронному тлумачному словнику через гіперонімічні зв'язки. Для експлікації семантичних властивостей відносних прикметників тут ми обираємо формалізм теорії семантичних станів²⁴. Останній є похідним від феноменологічного підходу до опису мовних явищ²⁵, який передбачає реалізацію опису семантики через визначення загального семантичного стану мовної одиниці, детермінованого максимально повною сумою контекстів, в яких вона функціонує. Для одиниць лексичного рівня цей стан представлено сумою ознак граматичної та лексичної семантики, організованою через гіперонімічні (та інші) зв'язки з іншими одиницями.

Теорію семантичних станів побудовано як узагальнення аналізу системної організації представлення семантичної інформації в тлумачних словниках. Стисло її можна передати формулою:

$$Fg_i(S) = \{M(L), Z(S), X_i\}, \quad (16)$$

де гіперонімічні зв'язки представлені як семантична тема Z — родова ознака, X_i — видові ознаки Z . До оператора Z вводимо також і диференційні ознаки. Багатопараметричність формули семантичного стану, практично необмежена кількість її складників, зміст яких зумовлений семантичним контекстом, що є відображенням денотативного рівня значення, наділяє формалізм значними потенціями для узагальнень, робить його гнучким інструментом у дослідженнях лексичної семантики.

Так, для відображення специфіки семантики відносних прикметників з іменниковою похідною основою до загального оператора F , який інтерпретується як оператор сукупності значень певних семантичних категорій, вводимо в ролі репрезентантів реляційної ознаки, позначуваної відносним прикметником, новий оператор V (семантичне відношення), йому підпорядковується множина операторів $\{X_i(V)\}$, що представляють конкретні відношення (об'єктні, суб'єктні, інструментальні, просторові, ...), що характеризують семантичні стани конкретних ЛСВ, а також вводимо до оператора Z диференційні ознаки: X_1 (семантичний клас іменника, від якого утворено прикметник) і X_2 (семантичний клас означуваного цим прикметником іменника).

Для ілюстрації розглянемо приклад прикметника з формулою тлумачення, представленою в «Словнику української мови» одним лексичним значенням:

ДІВО#ЧИЙ, а, е. 1. Прикм. до **діва** і **дівчина**²⁶.

²⁴ Широков В. А. Феноменологія лексикографічних систем.— К., 2004.— 327 с

²⁵ Широков В. А., Любченко Т. П., Шевченко І. В., Широков К. В. Граматичні системи. Феноменологічний підхід.— К., 2018.— С. 16–20.

²⁶ <http://lcorp.ulif.org.ua/LSlist/>.

Це значення в електронному тлумачному словнику варто подати у формі «прикметник від **дiва** і **дiвчина** у значенні “молода незаміжня особа жіночої статі”» з відповідними відтінками. Ними можуть бути, наприклад, такі семантичні відношення:

– у сполученні з іменником на позначення зовнішньої характеристики (дівочі плечі) — семантичне відношення «атрибутивне», відтінок — «які є зовнішньою характеристикою дівчини, дівчат»;

– у сполученні з іменником на позначення групи людей (дівочий гурт) — семантичне відношення «бути елементом», відтінок — «який має у своєму складі дівчат»;

– у сполученні з іменником на позначення одягу (дівочий капелюшок) — семантичне відношення «посесивне», відтінок — «який належить дівчині».

Інформаційно-програмним інструментарієм для такого дослідження варто визначити певну проблемно-орієнтовану програмну надбудову над ВЛЛ СУМ, за допомогою якої формується лінгвістична база даних (ЛБД), зорієнтована на вирішення конкретного завдання — дослідження семантичних властивостей прикметників. У цьому підході ЛБД поєднуватиме в собі автоматизоване робоче місце дослідника, масив мовного матеріалу дослідження, засоби семантичної диференціації, проміжні результати певних етапів аналізу та кінцеві результати. Останні, власне, й набувають форми електронного тлумачного словника. У ЛБД повинні експлікуватися семантичні складники лексичного значення відносних прикметників, структуровані у вигляді окремих полів. У такий спосіб на ЛБД індукується структура певної Л-системи. Зрозуміло, що ця система мусить бути обладнана необхідними сервісними функціями. Сформована в такий спосіб лінгвістична база даних є репрезентативним матеріалом для проведення в автоматизованому режимі лінгвістичних досліджень з питань однозначності/багато-значності, словотвірної семантики, вона дозволяє отримати інформацію про семантичні класи слів, входження лексеми до стійких словосполучень тощо.

Експліцитне представлення значення відносних прикметників в автоматичних перекладних словниках дає змогу в системі машинного перекладу визначати перекладні еквіваленти багатозначних слів з урахуванням конкретних значень слів, які перекладаються, розрізняти синоніми і перекладні елементи. Викладена схема може і мусить бути поширена на інші лексико-семантичні групи та лексико-граматичні класи слів.

Проведені дослідження підтверджують думку про те, що великі словники є носіями глибоких системних закономірностей мови. Одержані наукові результати демонструють ефективність Віртуальної лексикографічної лабораторії «Словник української мови» для дослідження нових лінгвістичних фактів, закономірностей і законів мовної системи.

Автори висловлюють подяку Н. М. Заїці та О. О. Погрібній за люб'язний дозвіл на використання матеріалів їхніх дисертацій у нашій статті.

V. A. SHYROKOV, I. V. OLEKSANDRUK

THE VIRTUAL LEXICOGRAPHIC LABORATORY «DICTIONARY OF THE UKRAINIAN LANGUAGE» AS A TOOL FOR SEMANTIC RESEARCH

The article deals with the application of the theory of lexicographic systems as a methodology for organizing and conducting diverse studies of the language system. A tool for this purpose serves the Virtual lexicographic laboratory «Dictionary of Ukrainian Language» (VLL SUM). Examples of regularities of the lexical system of the Ukrainian language, established using the VLL SUM, are given. The semantic and instrumental structures of the electronic explanatory dictionary integrated with VLL SUM for the in-depth study of the semantic structure of the Ukrainian adjective are proposed.

Keywords: lexicographic system, lexicographic effect, lexical semantics, linguistic database, virtual lexicographic laboratory.