

ФОРМАЛЬНА МОДЕЛЬ ВІДОБРАЖЕННЯ СЕМАНТИЧНИХ ХАРАКТЕРИСТИК У СЛОВНИКАХ

© Кульчицький І. М., Лукач М. О., 2014

Розглянуто питання важливості інформаційних технологій для розв'язання практичних задач лінгвістики. Обґрунтовано доцільність представлення семантичних характеристик дієслів у словниках тлумачного типу. На основі інформаційної теорії лексикографічних систем побудовано формальну модель відображення семантичних характеристик дієслів у словниках.

Ключові слова: моделювання, формальна модель, семантична розмітка, інформаційна теорія лексикографічних систем.

The article deals with the importance of information technologies for solving practical problems of linguistics. The expediency of presenting the semantic characteristics of verbs in explanatory dictionaries is proven. On the basis of the information theory of lexicographic systems the formal model of representation of semantic characteristics of verbs in dictionaries is built.

Key words: modeling, formal model, semantic markup, information theory of lexicographical systems.

Вступ. Загальна постановка проблеми

В останні десятиліття лінгвістику як науку неможливо собі уявити без застосування інформаційних технологій. У світі активно розвивається лінгвістична технологія як галузь інформатики, орієнтована на розроблення у комп'ютерних системах різноманітних застосувань природномовних механізмів. Це зумовлено потребою у суспільстві в програмно-апаратних засобах, які ефективно маніпулюють природномовною інформацією. До найпопулярніших таких програм належать програмні засоби, які автоматично виправляють граматичні помилки, здійснюють автоматичне редагування, індексування та пошук інформації природної мови, а також лінгвістичні та текстові редактори і процесори, програми оптичного розпізнавання текстів та машинного перекладу. Це задає технологічний статус мови, що визначається функціональною повнотою, якісним рівнем та розповсюдженням засобів, які уможливають застосування природномовних об'єктів у штучно створених інформаційних системах [13].

Протягом останнього часу для більшості світових мов (наприклад, англійської, французької, німецької, іспанської, російської) було створено основні засоби опрацювання природномовних текстів (natural language processing), що ґрунтуються на потужних комп'ютерних словниках. В українському мовознавстві практично відсутні промислові програмні засоби укладання словників за допомогою комп'ютера. Тому сьогодні в Україні відчувається велика потреба появи нових словників різного типу, складовим елементом яких є українська лексика та відповідні лексикографічні програми [6].

Особливого значення для лінгвістики набули питання формалізації та моделювання об'єкта її дослідження – мови. Одним з аспектів моделювання мовних явищ є моделювання лексичної семантики, зокрема дієслівної. Для української мови є потреба у створенні семантичного словника. Щоб автоматизувати процес створення українського семантичного словника, потрібно спочатку розробити його формальну модель.

Аналіз останніх досліджень та публікацій

Протягом останніх 20-ти років в університеті м. Ланкастер, Великобританія розроблялася семантична система аналізу USAS (UCREL semantic analysis system, UCREL = [Lancaster] University Centre for Computer Corpus Research on Language). В основу системи семантичної розмітки тексту USAS покладено систему класифікації лексики відповідно до Лонгманівського лексикону сучасної англійської мови Т. МакАртура. В Лонгманівському лексиконі МакАртура 15 тис. слів, що належать до основної лексики, було згруповано за 14 семантичними полями / темами, які поділено на 127 груп і 2441 підгрупу. Зараз USAS охоплює 21 семантичну категорію верхнього рівня та 232 підкатегорії. Ця система активно використовується для семантичного маркування лексики не тільки англійської мови, а й італійської, китайської та голландської мов [20]. Семантичні теги позначають семантичні поля, які групують разом значення слів, які на якомусь загальному рівні пов'язані з таким самим ментальним концептом [17].

Подібні дослідження проводили й інші зарубіжні вчені, зокрема У. Чейф, Ф. Денеш, Ю. Д. Шведова, Л. М. Васильев та ін. В українському мовознавстві семантичною класифікацією дієслівної лексики займалися В. М. Русанівський, І. Р. Вихованець, С. М. Дишлева О. І. Леута та інші. І. Р. Вихованець поділив всі предикати на [3, с. 93–111]: 1) предикати дії (вказують на діяльність, яку активно і безпосередньо породжує і стимулює суб'єкт); 2) предикати процесу (позначають певну динамічну ситуацію, яка не потребує активного виконавця); 3) предикати стану (передбачають “пасивного” суб'єкта, реалізуються у тому відрізьку часу, з яким співвідносяться, отже, відображають тимчасову ознаку предмета); 4) предикати якості (позначають деяку внутрішню і постійну ознаку, яка в певному сенсі є невід'ємною щодо предмета); 5) локативні предикати (диференційною однакою вихідної локативності є статичність, позначають локативний стан); 5) предикати кількості (позначають точну або неозначену кількість відповідних предметів). В. М. Русанівський поділив дієслова на 41 семантичну групу за такими ознаками [12, с. 326]: 1) чи поєднується дієслово в особовій формі з дієсловом в інфінітивній формі; 2) чи поєднуються дієслова у різних відмінках з іменниками, які мають або не мають прийменників; 3) чи можна утворити пасивний дієприкметник від даного дієслова. Використовуючи логіко-граматичний підхід О. І. Леута виокремив 11 лексико-семантичних груп українських дієслів [8, с. 9]. Своєю чергою, С. М. Дишлева визначила 12 лексико-семантичних груп дієслів на основі спільних сем та відповідності дієслів із реальними фактами [4, с. 8].

Формулювання мети

Мета статті – побудувати формальну модель відображення семантичних характеристик дієслів у словниках за допомогою інформаційної теорії лексикографічних систем.

Аналіз отриманих наукових результатів

У лінгвістичній літературі термін «модель» вперше використав американець З. Херріс у 1944 р., характеризує різницю між методологічними прийомами двох лінгвістів С. Х'юмена та Е. Сепіра, а потім у 1951 р. він ще раз вжив цей термін для визначення результатів описової термінології Е. Сепіра. Проте в більш специфічному сенсі (а саме для застосування в граматиці) термін «модель» використав Ч. Хокетт у 1954 р., а також Н. Хомський у 1956 р. Останні визначають модель як узагальнену та формалізовану структуру або процес певних фактів мови. А. Еттінджер у 1957 р. вже говорить про моделі не тільки як про відтворення дійсності, але і як про вплив на цю дійсність. Ще один лінгвіст В. Інґве поширює поняття моделі з конкретних мов на механізм мови загалом [10, с. 14].

Поняття «модель» поширено у різноманітних галузях науки і, відповідно, відрізняються його визначення залежно від конкретного застосування. Розглянемо деякі підходи до розуміння цього терміна у мовознавчій науці.

Як зазначено в лінгвістичному енциклопедичному словнику [9, с. 304], модель у лінгвістиці: 1) штучно створений лінгвістом реальний або уявний пристрій, що відтворює, імітує своєю поведінкою (зазвичай у спрощеному вигляді) поведінку якогось іншого («справжнього») пристрою

(оригіналу) в лінгвістичних цілях; 2) зразок, що слугує стандартом (еталоном) для масового відтворення; те саме, що і «тип», «схема», «парадигма», «структура» тощо (наприклад, «модель дієвідміни або відміни», «словотвірна модель», «модель речення» і т. д.).

В іншому словнику [2, с. 238] подано подібне визначення. Модель (англ. pattern, model): 1) (схема) закономірне розташування послідовностей сегментних і надсегментних елементів у складних мовних одиницях, що визначається характеристиками цієї мовної структури і регулярно відтворюється в мовленні (наприклад, словотвірна модель, модель речення); 2) характерна для певної мови парадигматична схема елементів (складових) її структури; компактне символічне зображення цієї схеми (чи будь-яких її елементів), що є основою методу моделювання мови (наприклад, породжувальна модель).

Важливими напрямками застосування комп'ютерів у лінгвістиці є [1]:

- класифікація джерел;
- структурно-семантичне, функціонально-стилістичне виділення мінімальних відрізків мови;
- дослідження семантики та лексики текстів;
- розроблення методів семантичного представлення текстів та тлумачень у словниках;
- комп'ютерне проектування словників в заданих узагальнених структурах словникових статей;
- формування комп'ютерних морфемних довідників для індексування масивів текстів та слів тощо.

За допомогою автоматизації у лексикографії можна ефективно вирішувати такі завдання [6]:

1. Спостерігати зміни у вживанні лексики у текстах, породжувати тексти та текстові блоки для їх подальшого багатократного використання.

2. При перетворенні та моделюванні засобами комп'ютера традиційних словників можна здійснювати різноманітні лексикографічні дослідження в автоматичному режимі та узагальнювати отримані результати.

3. Здійснювати всебічний аналіз вже раніше опублікованих словників та створювати потрібні авторські словники. Це забезпечує повноту лексики загальнономовних тлумачних словників і підвищує якість виправлення під час перевидання словників.

4. Формувати базу з лексикою певної мови з визначенням її належності до певного типу словника, узгоджувати реєстри однотипних двомовних словників, а також здійснювати аналіз за багатьма критеріями для формування реєстру словника у перекладній лексикографії.

5. У галузі термінографії:

- створювати словники, що якомога повніше описують суть понять, цілком відповідають загальноприйнятним правилам та нормам української мови, забезпечують погодженість термінів, задовольняють вимоги лінгвістів, а також фахівців певної галузі;
- на основі вимог цілісності, повноти, експлікативності та несуперечливості уніфікувати системи термінів;
- укладати автоматизовані банки термінів.

6. Впорядкувати та уніфікувати наявні масиви фактологічних даних для створення різних енциклопедичних видань.

Отже, актуальними все ще є питання створення сучасних інформаційних технологій у галузі лексикографії. Такі технології мають бути орієнтовними як на комп'ютерну підтримку створення традиційних лексикографічних праць, так і на створення комп'ютерних лексикографічних систем, які належать до програмних комплексів автоматизованого чи автоматичного опрацювання природномовних об'єктів [6].

Характерною рисою сучасних програм, які працюють з об'єктами природної мови, є зростання рівня їх інтелектуальності [13, 15, 16]. Їх створення пов'язане з використанням методів,

які традиційно відносять до застосування теорії та практики штучного інтелекту в комп'ютерній лінгвістиці. Концептуальною основою цих методів є так звана гіпотеза про інтегральну обробку, за якою комп'ютерне опрацювання мови забезпечує її надійне розуміння лише за спільного застосування і взаємодії синтаксичних та семантичних структур разом з репрезентантами знань про предметні області, які становлять зміст мовних об'єктів, що підлягають моделюванню. Тобто прикладні мовознавчі студії, що орієнтовані на моделювання процесів розуміння, повинні застосовувати для подання загальнозживаних знань комплексні нелінійні структури когнітивних наук, які є формами існування інформації [6].

Такий підхід вимагає відповідного інструментарію у вигляді розвинених формальних моделей природної мови, відповідних алгоритмів і програм, системотехнічних засад створення цілісних інформаційних технологій лінгвістики тощо. Саме у словниках теоретичні проблеми мовознавчої науки, зокрема проблема опису значення слова, знаходять своє розв'язання, реалізуються і перевіряються результати теорії. Лексикографічні описи самі по собі містять результати лінгвістичного аналізу, що зумовлює їх наукове значення [16].

Внутрішню організацію семантики слів подають у словнику відповідно до типу останнього, залежно від його призначення, обсягу, спеціалізації. На противагу словникам енциклопедичним, які доволі широко залучають до своїх реєстрів фахові й спеціальні знання та термінологію, що стосується багатьох галузей техніки, науки і культури, і пропонують стисле, але якомога вичерпне наукове визначення кожного терміна, тлумачні загальномовні словники характеризують слово засобами філологічного тлумачення, розкриваючи суть його семантичної будови. Адже основним завданням тлумачного словника є подання якомога точнішої семантичної характеристики лексики певної мови в повному обсязі, який відповідає його завданням і призначенню. Тлумачні загальномовні словники характеризують слово засобами філологічного тлумачення, розкриваючи суть його семантичної будови. У тлумачному словнику лексику системно описують, що передбачає виявлення і репрезентацію зв'язків і відношень між словами у лексико-семантичних групах. Тлумачення відображає лексичне значення у тлумачному словнику і тому є найважливішим компонентом словникової статті, а розроблення метамови словника, створення несуперечливих тлумачень слів є найважливішими лексикографічними проблемами [16]. Лексикографи різних країн обговорюють теоретично і вирішують практично питання створення універсальної метамови, яка давала б змогу, користуючись певним набором семантичних компонентів, створювати тлумачення слів [21].

Традиційно семантику слів вивчають у межах певних об'єднань лексем. Відповідно до теорії поля в мовознавстві, крім функціональних та синтагматичних полів, виділяють також парадигматичні поля, до яких належать різноманітні семантичні класи слів, зокрема і лексико-семантичні групи. У межах цих полів формуються семантичні моделі тлумачення.

Важливою проблемою у сучасній лексикографічній науці є розроблення представлення семантики окремих семантичних об'єднань слів. Доцільність розроблення однотипних тлумачень слів вбачала Н. Ю. Шведова для окремо взятої частини мови: „Для того, щоб уникнути різнобою у словнику, лексикограф повинен працювати... з певними лексичними групами в межах окремих частин мови. Тільки на цьому шляху можна досягти єдності у розмежуванні значень, однотипності тлумачень” [14].

З погляду лексичної системи найважливішим типом класів слів є лексико-семантичні групи. Вони об'єднують слова, що належать до однієї частини мови і в яких є як мінімум одна загальна сема – архісема. Архісема передбачає певні аспекти свого уточнення, так звані диференційні семи. У кожній лексико-семантичній групі є специфічний набір диференційних сем.

Архісеми також називають інтегральними (З. Д. Попова, Й. А. Стернін), або ідентифікувальними семами (Л. М. Васильєв). Конкретніші семи, які відрізняють індивідуальне значення слова від інших протиставлених йому значень, визначають як диференціовальні семи (Л. М. Васильєв), диференційні семи (В. Г. Гак), індивідуалізувальні, діагностувальні (І. В. Арнольд) [16, с. 28].

Наявність спільної архісеми передбачає те, що всі елементи лексико-семантичної групи семантично співвідносяться. Це безпосередньо стосується поняття лексикографічного типу: до однієї групи лексем, об'єднаних хоча б однією якоюсь спільною властивістю, застосовуються однакові правила опису в словнику. Тобто прослідковується типове представлення елементів значення слова у тлумаченнях слів однієї лексико-семантичної групи, а сама формула тлумачення є певним чином стандартизованою, уніфікованою структурою, у межах якої реалізується однотипне функціонування семантичних компонентів. «Отже, актуальними є поділ словникового складу на лексико-семантичні об'єднання і створення у їх складі уніфікованих формул тлумачення. Такий напрямок досліджень є актуальним і конструктивним для формування наукової парадигми лексико-семантичних відношень мови, їхніх класифікації та комбінації, подання у вигляді цілісної картини» [16].

Формальна модель відображення семантичних характеристик дієслів у словниках може знайти своє застосування у семантичному маркуванні дієслів у корпусах текстів.

Для сучасних лінгвістичних досліджень проведення семантичної розмітки корпусу та створення ефективних комп'ютерних засобів її обробки є важливим, оскільки дає змогу досліджувати семантику на репрезентативних мовних масивах, а не лише на окремих прикладах. Зокрема, семантично маркований корпус надає засоби перевірки гіпотез та верифікації алгоритмів зняття семантичних неоднозначностей, що не може бути переоцінене з огляду на необхідність створення ефективних технологій опрацювання знань [7].

Дж. Ліч визначає маркування корпусу як: “практику додавання лінгвістичної інформації до корпусу” [18]. Спочатку він [19] розглядав такі типи розмітки: орфографічна, фонетична, просодична, граматична, синтаксична, семантична, дискурсивна і прагматична / стилістична. Проте в пізнішій публікації лінгвіст виділяє такі типи лінгвістичної розмітки в корпусі [18]:

1. Граматична / частиномовна розмітка – додавання граматичної інформації до слова.
2. Фонетична розмітка – додавання інформації про те, як слово було вимовлено в корпусі усного мовлення. Просодична розмітка – знову ж в корпусі усного мовлення – додавання інформації про просодичні елементи, такі як наголос, інтонація та пауза.
3. Семантична розмітка – додавання інформації про семантичні категорії слів – наприклад, один і той самий іменник може належати до різних семантичних категорій, хоча може нічим не відрізнятися за написанням або вимовою.
4. Прагматична розмітка – додавання інформації про види мовленнєвих актів, які відбуваються в розмовному діалозі – тобто одне і те саме висловлювання може фігурувати як офіційна заява, запит на зворотний зв'язок, схвалення або прагматичний маркер ініціювання нового етапу обговорення.
5. Дискурсивна розмітка – додавання інформації про анафоричні посилання у тексті, наприклад, займенник у цьому реченні з'єднаний з іменником у попередньому реченні.
6. Стилійстична розмітка – додавання інформації про вираження мовлення та думок (пряма мова, непряма мови, вільна непряма думка і т.д.).
7. Лексична розмітка – додавання леми, тобто основної форми слова до кожної словоформи в тексті.

При побудові формальної моделі відображення семантики дієслів припустимо, що семантичні ознаки кожного дієслова подають у словниках тлумачного типу, позаяк структура статей тлумачних словників імпліцитно містить структуру більшості одномовних філологічних словників [5]. Таке припущення дасть нам змогу, по-перше розглядати семантичну характеристику дієслів як один з лексикографічних параметрів [16], по-друге, застосувати при моделюванні поняття елементарної лексикографічної моделі даних та її конкретної реалізації – елементарної лексикографічної системи (ЕЛС) з інформаційної теорії лексикографічних систем [6, 15, 16]. За цією теорією основу словника BS_d [6] подаємо як ЕЛС [16]:

$$BS_d = \{I_d, \delta_d, V_d, \sigma[\beta]_d, \mathcal{F}_d, \mathcal{C}_d, \mathcal{H}_d\},$$

де

- $I_d = \cup_{i,j=1-N} x_{i,j}$ – множина сутностей реального світу, що підпадають лексикографуванню відповідно до мети словника (елементарні лексикографічні одиниці (ЕЛЮ) BS_d); N – обсяг словника;
- δ_d – відображення, що вербалізує ЕЛЮ;
- $V_d = \cup_{x \in I_d} V_d(x) \equiv (\Lambda_d; P_d)$ – множина словникових статей, яку, своєю чергою, можна подати як агрегат множин лівих ($\Lambda_d = \cup_{x \in I_d} \Lambda_d(x)$) та правих ($P_d = \cup_{x \in I_d} V_d(x)$) частин статті словника;
- β_d – структура основи словника;
- $\sigma[\beta]_d$ – макроструктура статей словника;
- \mathcal{F}_d – відображення, що дає змогу з V_d отримати Λ_d ;
- \mathcal{E}_d – відображення, що дає змогу з V_d отримати P_d ;
- \mathcal{H}_d – ізоморфне відображення між Λ_d і P_d ;

Скористаємось рекурсивною редукцією лексикографічної системи і подамо праву частину словникової статті як ЕЛС відносно набору атрибутів, що характеризують ліву частину словникової статті:

$$LS_p = \{I_p, \delta_p, V_p, \sigma[\beta]_p, \mathcal{F}_p, \mathcal{E}_p, \mathcal{H}_p\},$$

де

- $I_p = \cup_{i,j=1-N} p_{i,j}$ – множина атрибутів-характеристик лівої частини словникової статті (елементарні лексикографічні одиниці (ЕЛЮ) LS_p); N – їхня кількість;
- δ_p – відображення, що вербалізує ЕЛЮ;
- $V_p = \cup_{x \in I_p} V_p(x) \equiv (\Lambda_p; P_p)$ – множина лексикографічних параметрів, що творять праву частину словника і яку, своєю чергою, можна подати як агрегат множин лівих ($\Lambda_p = \cup_{x \in I_p} \Lambda_p(x)$) та правих ($P_p = \cup_{x \in I_p} V_p(x)$) частин статті лексикографічної системи (у нашому випадку $\Lambda_p(x)$ має вироджений вигляд, її неявно подають у статті поліграфічним чи композиційним оформленням);
- β_p – загальна структура правої частини словника;
- $\sigma[\beta]_p$ – макроструктура кожного лексикографічного параметра;
- \mathcal{F}_p – відображення, що дає змогу з V_p отримати Λ_p ;
- \mathcal{E}_p – відображення, що дає змогу з V_p отримати P_p ;
- \mathcal{H}_p – ізоморфне відображення між Λ_p і P_p ;

Розглянемо, наприклад, статтю зі Словника української мови [11]:

БАНУВАТИ, ую, уеш, без додатка і за ким, чим, *недок., діал.* Сумувати, журитися. – Почав я банувати. А найгірше за Марією (*Михайло Коцюбинський, I, 1955, 141*); Наче за мерцем, за Петриком банувала [баба] (*Марко Черемшина, Тв., 1960, 38*); – Діду, ви не бануйте. От побачите, що все буде наше: і земля, і заводи, всі ліси і полонини, – говорить Василько дідові (*Агата Турчинська, Зорі..., 1950, 133*).

Права частина цієї статті містить чотири лексикографічні параметри ($I_p = \{\text{частиномовна характеристика; стилістична ремарка; тлумачення; приклади вживання}\}$). Λ_p (частиномовна характеристика) задана своїм місцем у статті (перше після коми, яка відділяє реєстрове слово; Λ_p (стилістична ремарка) задана місцем у статті (після частиномовної характеристики) та способом відображення; Λ_p (тлумачення) задана місцем (після ремарок); Λ_p (приклад вживання) задана місцем (після тлумачення). P_p (частиномовна характеристика); P_p (стилістична ремарка) = (діал.);

P_p (тлумачення) = (Сумувати, журитися); P_p (прикладні вживання) (Почав я банувати. А найгірше за Марією (Михайло Коцюбинський, 1, 1955, 141); Наче за мерцем, за Петриком банувала [баба] (Марко Черемшина, Тв., 1960, 38); – Діду, ви не бануйте. От побачите, що все буде наше: і земля, і заводи, всі ліси і полонини, – говорить Василько дідові (Агата Турчинська, Зорі..., 1950, 133)).

Продовжуючи рекурсивну редукцію лексикографічної системи кожний лексикографічний параметр, з якого складається права частина словникової статті можна подати як елементарну лексикографічну систему, де роль ЕЛЮ матимуть його значення. За необхідності цей процес можна продовжити, деталізуючи структуру кожного значення лексикографічного параметра.

Для побудови ЕЛС, що відповідає лексикографічному параметру «відображення семантики дієслова» скористаємося поділом дієслів на лексико-семантичні групи, який запропонувала С. М. Дишлева [4] та побудуємо на його ґрунті ряд елементарних лексикографічних систем:

$$LS_s = \{I_s, \delta_s, V_s, \sigma[\beta]_s, \mathcal{F}_s, \mathcal{E}_s, \mathcal{H}_s\}.$$

Для цієї ЕЛС множина I_s міститиме 12 таких елементів: {Дієслова дії; Дієслова руху та переміщення; Дієслова релятивної семантики; Дієслова стану і процесу; Дієслова ментальних і соціальних дій суб'єкта; Дієслова звучання та звуконаслідування; Дієслова буття; Дієслова на позначення виникнення, появи, настання; Дієслова характеристики; Дієслова просторової локалізації; Модальні дієслова; Фазові дієслова}. Для окремих елементів I_s необхідно побудувати свої ЕЛС з власними елементарними лексикографічними одиницями:

$I_{\text{Дієслова дії}} = \{\text{Дієслова загальної фізичної дії; Дієслова на позначення дії, яка спричиняє зміни в об'єкті; Дієслова на позначення дії, яка створює об'єкт}\};$

$I_{\text{Дієслова релятивної семантики}} = \{\text{Дієслова ставлення; Дієслова впливу; Посесивні дієслова; Дієслова залежності}\};$

$I_{\text{Дієслова стану і процесу}} = \{\text{Дієслова психічного стану; Дієслова фізіологічного стану; Дієслова на позначення переходу із одного стану в інший; Дієслова на позначення фізіологічних процесів}\};$

$I_{\text{Дієслова ментальних і соціальних дій суб'єкта}} = \{\text{Дієслова фізичної чи інтелектуальної діяльності або відпочинку; Дієслова мовлення; Дієслова мислення; Дієслова пізнання; Дієслова сприйняття}\};$

$I_{\text{Дієслова сприйняття}} = \{\text{активного; пасивного}\};$

$I_{\text{Дієслова характеристики}} = \{\text{Позначення кваліфікативних властивостей особи; Позначення зовнішнього вияву ознаки дії або стану; Позначення зменшення або збільшення вияву ознаки}\};$

$I_{\text{Модальні дієслова}} = \{\text{волевияву; бажання; наміру; можливості}\}$

Висновки і перспективи подальших наукових розвідок

Окреслено поняття моделі у лінгвістиці, а також важливі напрямки застосування комп'ютерів у лексикографії. Обґрунтовано доцільність подання семантичних характеристик дієслів у словниках тлумачного типу. На основі інформаційної теорії лексикографічних систем побудовано елементарну лексикографічну модель, яка є репрезентативною формальною моделлю відображення семантичних характеристик дієслів.

1. Андрющенко В. М. Концепція і архітектура машинного фонду російського мови. – М.: Наука, 1989. – 200 с. 2. Ахманова О. С. Словарь лингвистических терминов. – 2-е изд., стер. – М.: Сов. энциклопедия, 1969. – 608 с. 3. Вихованець І. Р. Нариси з функціонального синтаксису української мови / Іван Романович Вихованець. – К.: Наук. думка, 1992. – 220 с. 4. Дишлева С. М. Адвербіальна дистрибуція лексико-семантичних груп українських дієслів: автореф. дис... канд. філол. наук: 10.02.01 / С. М. Дишлева; НПУ ім. М.П. Драгоманова. – К., 2008. – 18 с. 5. Кульчицький І. Деякі аспекти типології словників / І. Кульчицький, А. Костенко, Н. Осідач // Прикладна лінгвіс-

тика та лінгвістичні технології: *MegaLing-2008*: зб. наук. пр. / НАН України, Укр. мовно-інформ. фонд [та ін.]; редкол.: Ю. Д. Апресян [та ін.]. – К.: Довіра, 2009. – С. 135–145. 6. Кульчицький І. М. Комп'ютерно-технологічні аспекти створення сучасних лексикографічних систем. – К.: НБУ ім. В. І. Вернадського, 2002. – 59 с. 7. Корпусна лінгвістика: Моногр. / В. А. Широков, О. В. Бугаков, Т. О. Грязнухіна, О. М. Костишин, М. Ю. Кригін; НАН України, Укр. мов.-інформ. фонд. – К.: Довіра, 2005. – 472 с. 8. Леута О. І. Дієслівні речення в українській літературній мові: структура, семантика, моделі. : автореф. дис. ... д-ра філол. наук / О. І. Леута. – К.: [б. и.], 2009. – 35 с. – Б. ц. 9. Лінгвістический энциклопедический словарь / Гл. ред. В. Н. Ярцева. – М.: Сов. энциклопедия, 1990. – 683 с. 10. Лосев А. Ф. Введение в общую теорию языковых моделей / Под ред. И. А. Василенко. – 3-е. изд. – М.: Едиториал УРСС, 2010. – 296 с. 11. Словник української мови: В 11 т. – К.: Наук. думка, 1970–1980. 12. Сучасна українська літературна мова : морфологія / за заг. ред. акад. АН УРСР І. К. Білодіда. – К.: Наук. думка, 1969. – 583 с. 13. Феноменологія лексикографічних систем: моногр. / В. А. Широков; НАН України. Укр. мов.-інформ. фонд. – К.: Наук. думка, 2004. – 327 с. 14. Шведова Н. Ю. Лексическая система и ее отражение в толковом словаре // Русистика сегодня. Язык: система и ее функционирование. – М.: Наука, 1988. – 224 с. 15. Широков В. А. Інформаційна теорія лексикографічних систем. – К.: Довіра, 1998. – 331 с. 16. Широков В. А. Комп'ютерна лексикографія / В. А. Широков. – К.: Наук. думка, 2011. — 347 с. 17. Archer D., Wilson A., Rayson P. Introduction to the USAS category system. Benedict project report. – October 2002. – Available from: <http://ucrel.lancs.ac.uk/usas/usas%20guide.pdf>. 18. Leech G. Adding linguistic annotation. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice*. – Oxford: Oxbow Books, 2004. – P. 17–29 – Available from: <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/chapter2.htm>. 19. Leech G. Introducing corpus annotation. In R. Garside, G. Leech & T. McEnery (Eds.), *Corpus annotation: Linguistic information from computer text corpora*. – London: Longman, 1997. – P. 1–18. 20. UCREL Semantic Analysis System (USAS). – Available from: <http://ucrel.lancs.ac.uk/usas/>. 21. Wierzbicka A. *Semantics: Primes and Universals*. – Oxford University Press, 1996. – 512 p.