

Lexical homonyms are interpreted as random phenomena, especially such as homofoms, homophones and homographs. Their rather wide representation in the composition of the Russian language dictionaries, the formation due to certain factors contributes to the identification of the typology of homonyms.

Attention is drawn to absolute synonyms as semantic oppositions of an identity, one-root synonymy, and variants of words; antonyms as oppositions of the equipollent type, as well as types of the opposition of the private type.

Private oppositions (inclusion oppositions) can manifest themselves both formally and in semantic links of words. In semantic terms, these are words related by generic relations. The private semantic oppositions are very important because they determine the structure of the vocabulary, its hierarchical nature.

Equipollent oppositions (crossing oppositions) bind words that have common features and at the same time their common specific components.

Key words: *absolute synonyms, antonymic opposition, borrowing in homonymy, complete homonymy, equivalence semantic opposition, invasive opposition, phonetic homonyms, semantic opposition, sociolinguistic character of homonyms, verbal opposition, word-formation homonyms, words variant.*

УДК 81'33+81'374.2:004

Грязнухіна Т. О.,
кандидат філологічних наук,
старший науковий співробітник відділу лінгвістики
Українського мовно-інформаційного фонду НАН України
E-mail: ukritaldb@gmail.com

Любченко Т. П.,
кандидат технічних наук,
старший науковий співробітник відділу інформатики
Українського мовно-інформаційного фонду НАН України
E-mail: t.lyubch@gmail.com

Олександрук І. В.,
аспірант Інституту мовознавства ім. О. О. Потебні НАН України,
E-mail: ira_oleksandruk@ukr.net

ФЕНОМЕНОЛОГІЧНИЙ ПІДХІД ДО ВИВЧЕННЯ СЕМАНТИКИ ВІДНОСНИХ ПРИКМЕТНИКІВ ТА СПОСОБУ ПРЕДСТАВЛЕННЯ ЇЇ В ЕЛЕКТРОННОМУ ТЛУМАЧНОМУ СЛОВНИКУ

У статті розглянуто феноменологічний підхід до вивчення і представлення семантичних властивостей відносних прикметників у електронному тлумачному словнику української мови. Досліджено семантичні особливості лексичного значення відносних прикметників з іменниковою похідною основою. Основну увагу приділено розробленню апарату експліцитного представлення в електронному тлумачному словнику лексичного значення відносних прикметників зі схемами «Прикметник до...» і «Стосовно до...».

Ключові слова: *комп'ютерна лексикографія, електронний тлумачний словник, теорія семантичних станів, експліцитне представлення інформації*

Зростання ступеня інформатизації та комп'ютеризації суспільного життя людей загострює проблему підвищення ефективності систем автоматичного оброблення інформації, зокрема тієї, що передається засобами природних мов. Успішного розв'язання цієї проблеми можна досягти шляхом підвищення ступеня інтелектуалізації лінгвістичних процесорів автоматичного розпізнавання змісту письмових текстів.

Зрозуміти зміст тексту означає, по-перше, ідентифікувати лексичні значення текстових одиниць-слів, по-друге, встановити зв'язки, в яких знаходяться референти, позначувані цими словами. Зробити це система автоматичного опрацювання тексту (АОТ) може за допомогою алгоритмів аналізу семантики текстових одиниць шляхом зіставлення із семантичною інформацією, автоматично екстрагованою з надійних лінгвістичних джерел, і представленою у формі, зрозумілій комп'ютерним програмам.

Для української мови таким надійним і репрезентативним джерелом семантичної інформації є Електронний тлумачний словник (ЕТЛС) української мови, створений як комп'ютерна версія 20-томного українського тлумачного Словника. ЕТЛС представлено Віртуальною лексикографічною лабораторією Словника української мови (ВЛЛ СУМ) [1], у межах якої здійснюється також й формування паперової версії словника. ВЛЛ виконує функцію потужного автоматизованого робочого місця (АРМ) для лексикографів-укладачів тлумачного словника і слугує інструментом приведення лексикографічної інформації, що міститься в паперовій версії, до форми, придатної до оброблювання її комп'ютерними видавничими системами.

Представлення у ВЛЛ СУМ лексикографічної інформації у структурованій формі, доступній комп'ютерним програмам, інтегрування Словника з іншими лексикографічними джерелами в межах загальної інтегрованої лексикографічної системи (ІЛС), створюваної в Українському мовно-інформаційному фонді [5], є необхідною, але недостатньою умовою для можливості використання ЕТЛС як мовно-інформаційного інструментарію семантичних досліджень. Важливо, щоб автоматично екстрагована зі словника інформація (йдеться про тлумачення значень слів) насамперед відповідала загальній вимозі, висунутій до електронних словників, щодо експліцитності представлення в них інформації [3]. У нашому випадку ця вимога (відповідно до поняття «лексичне значення») інтерпретується як необхідність відображення в тексті тлумачення інформації про предмет чи явище навколишнього світу, позначуване конкретним словом, у явному вигляді. Виконання цієї вимоги забезпечує здійснення словником функції бути автоматизованою інформаційною системою, розрахованою на широке коло користувачів.

Для того, щоб ЕТЛС став ефективним знаряддям проведення за допомогою методів комп'ютерної лінгвістики наукових досліджень із семантики з подальшим застосуванням одержаних результатів у системах АОТ, словникова текстова форма опису значення слова повинна дозволяти виконувати над нею цілий ряд логічних операцій з поділу значення на семантичні складники, встановлення ієрархічних відношень між ними, операцій зіставлення значень різних слів з метою встановлення їхньої семантичної подібності чи розбіжності.

Попередньо проведений із зазначених позицій аналіз схем тлумачень, задіяних в ЕТЛС [5], дозволяє зробити наступні висновки:

1. Для переважної більшості тлумачень у словнику характерним є експліцитний спосіб опису лексичного значення. І саме в дефініціях цієї більшості слів реалізується форма описування семантичних характеристик слів через встановлення гіперонімічних зв'язків між позначуваними ними референтами. Такий спосіб тлумачення значень цілком відповідає прийнятому в усіх лексикографічних працях тлумачного типу, енциклопедіях, термінологічних словниках, а також у таких електронних продуктах як тезауруси, семантичні класифікатори. За цим самим принципом будуються предметні онтології. У базах знань конкретних підмов структура даних організована аналогічно: таксономічні відношення (гіперонім – гіпонім) відіграють центральну роль. До них можуть (але не обов'язково) додаватися відношення тотожності (синоніми), композиційні відношення (ціле – компоненти цілого), відношення належності до множини як її елемента тощо. Тому спосіб описування лексичного значення слів шляхом визначення гіперонімічних зв'язків між позначуваними референтами може бути рекомендованим для застосування й при розробці тлумачних статей для нових слів, які вводяться до словника.

2. Серед 34 схем, виділених як такі, що реалізують неекспліцитний спосіб опису лексичних значень, поширеними є відсильні схеми (довжина ланцюжка відсилань може сягати чотирьох): «Те саме, що ...» (*АБАК 2. архт. Те саме, що абáка.*), «Див.» (*КАПЛОУ́ХИЙ див. каплову́хий.*), «Жін. до ...» (*ЛІКАРКА, и, ж. Жін. до лі́кар.*), схема представлення тлумачення одним або двома синонімами (*ГОЛЯ́Р, á, ч., заст. Перукар.*), а також схеми, які встановлюють зв'язок між значеннями компонентів видових пар дієслова «Док. до ...» (*ВІГАНЬБИТИ, блю, биш; мн. віганьблять; кого. Док. до ганьбіти.*), «Однокр. до ...» (*АГА́КНУТИ, ну, неш, док., розм. Однокр. до ага́кати.*). Ці схеми можуть бути названими умовно неекспліцитними, тому що програмні засоби ІЛС дозволяють у лінгвістичній базі словника автоматично перенести інформацію із зони «Тлумачення» слова, записаного в схемі після «що...», «див...», «до...», у зону «Тлумачення» слова з відповідною схемою.

3. Значення відносних прикметників із похідною іменниковою основою представлені схемами «Прикметник до..» (*АБАЖУРНИЙ, а, е. Прикм. до абазур*), «Стосовно до ...» (*КАБАЛІСТИЧНИЙ, а, е. Стос. до кабалістики*), зумовлені семантикою іменників, від яких вони утворені, а позначувана ними ознака стосується характеру зв'язків, які можуть встановлюватися між предметами-референтами цих іменників з іншими предметами. Число реалізацій схем в ЕТЛС 14 тисяч і 8 тисяч відповідно. Для цих прикметників експліцитне представлення значення можливе після аналізу їхніх семантичних властивостей із врахуванням вказаних факторів.

Саме ця група відносних прикметників і є об'єктом нашого дослідження.

Предмет дослідження – семантичні особливості лексичного значення відносних прикметників.

Завдання дослідження – створення лінгвістично-інформаційного інструментарію опису (аналізу) семантики відносних прикметників з іменниковою похідною основою.

Мета дослідження – розробка апарату експліцитного представлення лексичного значення відносних прикметників зі схемами «Прикметник до...» і «Стосовно до...» в ЕТЛС.

Матеріал дослідження – автоматично екстраговані з Електронного українського тлумачного словника обсягом 199 тис. слів тлумачні статті, в яких містяться значення з формулами «Прикм. до..», «Стос. до ...».

Методи дослідження – дистрибутивно-статистичний і теорія семантичних станів.

Вивчення специфіки семантики відносних прикметників – зумовленість її семантикою іменника, від якого походить прикметник, і реляційний характер позначуваної ним ознаки – в дослідженні здійснюється за допомогою формалізму, розробленого в межах теорії семантичних станів [8, 9], яка репрезентує феноменологічний підхід до вивчення мовних об'єктів різних рівнів [2]. У випадку аналізу семантичних властивостей слів згідно з феноменологічним підходом окреме значення слова (лексико-семантичний варіант) розглядається як його конкретний лінгвістичний стан, детермінований певним семантичним контекстом, представленим сумою ознак граматичної та лексичної семантики. За допомогою формалізму контекстуальна обумовленість семантичного стану описується оператором F , який інтерпретується як оператор сукупності значень певних семантичних категорій. Дія оператора F становить сумарну дію операторів: оператор M – задає граматичне значення лексеми L , тобто ідентифікує частину мови, Z – задає семантичну тему, яка вказує на референта ЛСВ, та множина X_i – оператори, які репрезентують диференційні семантичні ознаки відповідної семантичної теми. Тобто, дію оператора F як повного інтелектуального механізму представлення семантичних станів мовних одиниць описуємо за формулою:

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

де S – ЛСВ, $g(S)$ – семантичний стан ЛСВ.

Такий підхід до опису семантики слова цілком сумісний з принципами загальноновизнаного в лінгвістиці методу компонентного семантичного аналізу, суть якого полягає в розкладанні значення слова на семи (елементарні смислові одиниці), між якими встановлюються ієрархічні відношення. Але якщо семне представлення структури значення слова передбачає перетворення тлумачного словника в словник семантичний на кшталт словника Н. Ю. Шведової [6], то застосування формалізму теорії семантичних станів до вивчення семантичних властивостей слів вимагатиме лише певного коригування тексту тлумачення в ЕТЛС з позицій виконання умови обов'язкової присутності в тлумаченні лексем-представників певних семантичних класів, заданих операторами $\{X_i\}$. Наприклад, при аналізі тлумачень слів, що належать до тематичного поля рослин, семантична тема Z «Рослини» може бути ідентифікована в результаті знаходження в тексті тлумачення однієї з лексем {дереву, кущ, чагарник, трав'яниста рослина, овоч, суцвіття, бур'ян ...}.

Багатопараметричність формули семантичного стану, необмежувана кількість її складників, зміст яких зумовлений семантичним контекстом, який є відображенням денотативного рівня значення, наділяє формалізм значними потенціями для узагальнень, робить його досить гнучким інструментом у дослідженнях лексичної семантики.

Так, для відображення специфіки семантики відносних прикметників із іменниковою похідною основою в загальний оператор F , який інтерпретується як оператор сукупності значень певних семантичних категорій, вводимо в ролі репрезентантів реляційної ознаки, позначуваної відносним прикметником, новий оператор V (семантичне відношення), якому підпорядковується множина

операторів $\{X_i(V)\}$, якими представлені 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 утворений прикметник) і X_2 (семантичний клас означуваного цим прикметником іменника).

Віднесення семантичних відношень до числа диференційних ознак семантичного контексту, за якими можуть визначатися лексико-семантичні варіанти відносного прикметника є цілком правомірним. Саме за допомогою їх експлікується (розкривається) реляційна сутність позначуваної прикметником ознаки – характеризувати означуваний предмет опосередковано через зв'язок його з предметом, позначуваним іменником, від якого утворено цей прикметник.

Число значень (відтінків у межах значення) відносного прикметника перебуває у прямій залежності від кількості семантичних відношень, які реалізуються ним у сполученнях з іменником.

У мові визначальною для встановлення конкретних типів семантичних відношень є інформація про лексичну сполучуваність прикметника з іменниками, які належать до певних семантичних класів. Цим самим в описі семантичних ознак відносних прикметників експлікується ще одна складова їхнього лексичного значення – контекстна обумовленість. Наприклад, для прикметника *АБРАЗІВНИЙ* (схема тлумачення в ТЛС «Прикм. до абразиви») у результаті аналізу його лексичної сполучуваності з іменниками, встановленої на рівні семантичних класів (СК) «Устрій/ інструмент» (СК іменника, від якого утворено прикметник) і семантичних класів «Підприємство» (завод, фабрика) і «Дія» (шліфування, відрізання, личкування), до яких належать іменники із словосполучень *А* (абразивний)+*І*, зафіксованих у тексті, неекспліцитне тлумачення «Прикм. до абразиви» було замінено двома експліцитними ЛСВ відповідно – «який має об'єктом своєї діяльності абразиви» і «виконуваний за допомогою абразивів».

Єдиним репрезентативним джерелом для встановлення лексичної сполучуваності прикметників є текст. У нашому дослідженні, яке проводиться у межах Інтегрованої лексикографічної системи, компонентом якої є Український національний лінгвістичний корпус [4] обсягом понад 190 мільйонів слововживань, вибір контекстів здійснюється автоматично із залученням програми автоматичного морфологічного аналізу (АМА) тексту і програми лематизації. Інструментом АМА і лематизації в ІЛС слугує Електронний граматичний словник української мови [7].

Як показує досвід здійснюваної в Українському мовно-інформаційному фонді НАНУ роботи зі створення електронних семантичних словників, проводити аналіз на великих масивах лінгвістичного матеріалу за багатьма параметрами із залученням комп'ютерних методів доцільно шляхом створення спеціальних лінгвістичних баз даних (ЛБД), орієнтованих на опис конкретного семантичного явища (в нашому випадку йдеться про створення ЛБД «Семантичні властивості прикметників»).

Побудована як програмно-інформаційний інструментарій проведення семантичних досліджень, ЛБД поєднує в собі функцію автоматизованого робочого місця (АРМ) дослідника, масив мовного матеріалу, на якому проводиться дослідження, диференційні ознаки, за допомогою яких описується семантичне явище, а також проміжні результати певних етапів аналізу і кінцеві результати.

Інтегрування ЛБД з електронним тлумачним словником, сервісні програми АРМ забезпечують дослідникові під час роботи з ЛБД візуалізацію на екрані комп'ютера інформації певної її зони, можливість відкривати кілька вікон, можливість вносити потрібні виправлення, доповнення, проміжні результати, вилучати інформацію.

У межах АРМ реалізовано автоматичне формування текстових контекстів конкретної реєстрової одиниці з Лінгвістичного корпусу, інтегрованого в ІЛС. У результаті аналізу контекстів встановлюються характеристики лексичної сполучуваності.

В автоматизованому режимі заповнюється також зона ілюстративного матеріалу. З комп'ютера дослідник може записувати свої коментарі в спеціально відведену для цього зону в ЛБД, а також кінцеві результати аналізу.

Далі дається опис зон сформованої ЛБД.

Зони 1-9 заповнюються автоматично з ЕТЛС:

зона 1 – номер лексеми в ЛБД;

зона 2 – лексема;

зона 3 – номер значення в тлумачній статті ЕТЛС;

зона 5 – тлумачення значення (відтінка значення) зі словника;

зона 6 – іменник, від якого утворено прикметник;

зона 7 – значення іменника, від якого утворено прикметник.

Передбачається, що якщо значення в тлумачному словнику містить у собі відсилання до тлумачення інших реєстрових одиниць, у зоні повинен відображатися весь ланцюжок цих відсилань, доведений до тексту значення без них.

Заповнення зони покажемо на прикладі прикметника *БАГАТІРСЬКИЙ*:

багатирський (у 1-му значенні) – прикм. до *багатир*:

багатир – те саме, що *багач* →

багач – той, хто має багатство (у 1-му значенні) →

багатство 1 – велике майно, цінності, гроші →

багатство 1.1 – достаток усього, розкіш.

Зона 9 – стилістична мітка.

Зони 10, 11 орієнтовані на використання ЛБД для дослідження ролі словотвірної семантики у формуванні лексичного значення слів з похідною основою. У зоні 10 фіксується словотвірний формант (СФ) – суфікс прикметника, 11 – значення СФ.

Зона 12 – модифіковане значення із зон 5 і 7. Для *БАГАТІРСЬКИЙ* це «прикм. до *багатир* у значенні людина, що має великі достатки, майно, цінності».

Зона 13 – семантичний клас мотивувальної основи прикметника.

Зона 14 – семантичний клас іменника, з яким сполучається прикметник (визначено в результаті аналізу лексичної сполучуваності).

Зона 15 – семантичне відношення як репрезентант реляційної ознаки.

Зона 16 – Ілюстрації (словосполучення А+І з тексту).

Зона 17 – експліковане в результаті аналізу лексичне значення прикметника.

Зона 18 – входження прикметника в стійкі словосполучення (фразеологізми).

Результати дослідження. У дослідженні розроблено принципи формування експліцитних тлумачень для слів ад'єктивного класу, представлених у тлумачному словнику неекспліцитними схемами опису лексичного значення, за допомогою формалізму теорії семантичних станів. Здійснено певну адаптацію формалізму теорії до особливостей досліджуваного класу відносних прикметників. Завдання подолання неекспліцитного представлення семантичної інформації в ЕТЛС зводиться до завдання розробки дефініцій для нових слів, якими постійно поповнюється лексичний склад мови.

Побудовано спеціалізовану ЛБД «Семантика відносних прикметників», яка виконує функцію мовно-інформаційного інструмента для проведення досліджень семантичних властивостей ад'єктивів.

Дослідження пов'язане з розв'язанням такої важливої проблеми комп'ютерної лексикографії як створення електронних тлумачних словників і застосування їх в системах АОТ.

Перелік умовних скорочень

АМА – автоматичний морфологічний аналіз

АОТ – автоматичне оброблювання (опрацювання) тексту

АРМ – автоматизоване робоче місце

ВЛЛ – віртуальна лексикографічна лабораторія

ІЛС – інтегрована лексикографічна система

ЛБД – лінгвістична база даних

ЛСВ – лексико-семантичний варіант

СК – семантичний клас

СУМ – словник української мови

Список використаної літератури

1. Віртуальна лексикографічна лабораторія «Словник української мови». Український мовно-інформаційний фонд НАН України, 2001–2018. – [Електронний ресурс]. – Режим доступу: <http://lcorp.ulif.org.ua/LSlist/>

2. Граматичні системи. Феноменологічний підхід : [монографія] / В. А. Широков, Т. П. Любченко, І. В. Шевченко, К. В. Широков. – К. : Наукова думка, 2018. – 312 с.
3. Грязнухина Т. А. Эксплицитное представление семантической информации в Электронном толковом словаре украинского языка (синонимическая и паронимическая параметризация) / Т. А. Грязнухина, Т. П. Любченко, К. М. Якименко // Прикладная лингвистика и лингвистические технологии: MegaLing-2012: Сборник научных трудов. – К., 2012. – С. 79–91.
4. Корпусна лінгвістика: [монографія] / В. А. Широков, О. В. Бугаков, Т. О. Грязнухіна та ін. – К. : Довіра, 2005. – 470 с.
5. Лінгвістичні та технологічні основи тлумачної лексикографії: [монографія] / В. А. Широков, В. М. Білоноженко, О. В. Бугаков та ін. – К. : Довіра, 2010. – 293 с.
6. Русский семантический словарь : толковый словарь, систематизированный по классам слов и значений : [в 6 т.] / Российской академия наук, Отделение литературы и языка, Институт русского языка им. В.В. Виноградова ; под общей редакцией Н. Ю. Шведовой.
7. Шевченко И. В. Электронный грамматический словарь украинского языка / И. В. Шевченко, А. Г. Рабулец, В. А. Широков // Труды Международ. конфер. «Прикладная лингвистика в поиске новых путей». – Спб. : Изд-во Осипов, 2005. – С. 124–129.
8. Широков В. А. Семантичні стани мовних одиниць та їх застосування в когнітивній лексикографії / В. А. Широков // Мовознавство. – № 3 – 4. – 2005. – С. 47–62.
9. Широков В. А. Системна семантика тлумачних словників / В. А. Широков // Акцентологія. Етимологія. Семантика. До 75-річчя В. Г. Скляренка. – К. : Наукова думка, 2012. – С. 487–510.

Грязнухина Т. А., Любченко Т. П., Олександрук І. В. Феноменологічний підхід к изучению семантики относительных прилагательных и способа представления ее в электронном толковом словаре.

Анотація

В статье рассматривается феноменологический подход к изучению и представлению семантических свойств относительных прилагательных в электронном толковом словаре украинского языка. Исследованы семантические особенности лексического значения относительных прилагательных с именной производной основой. Разработан аппарат эксплицитного представления в электронном толковом словаре лексического значения относительных прилагательных со схемами «Прилагательное к ...» и «Применительно к ...».

Ключевые слова: компьютерная лексикография, электронный толковый словарь, теория семантических состояний, эксплицитное представление информации.

Hryaznukhina T. O., Lyubchenko T. P., Oleksandruk I. V. Phenomenological approach to the study of the semantics of relative adjectives and the way of its representation in the electronic explanatory dictionary.

Summary

The Object of research are relative adjectives.

The Subject of research are semantic features of the lexical meaning of relative adjectives.

The task of the study is to create a linguistic-information tool for describing (analyzing) the semantics of relative adjectives with a noun derivative basis.

The purpose of the study is to develop the apparatus of an explicit representation of the lexical meaning of relative adjectives with the schemes "Adjective to ..." and "In relation to ..." in the electronic explanatory dictionary.

The Material of the research - automatically extracted from the Electronic Ukrainian Explanatory Dictionary (circa 199 thousand words) the explanatory articles, which contain the meaning with the formulas "Adj. to ...", "In rel. to ...".

Methods of research - distributive-statistical method and theory of semantic states.

The semantic features of the lexical meaning of relative adjectives with noun derivative basis are investigated. The device of the explicit representation of the lexical meaning of relative adjectives with the schemes "Adjective to ..." and "In relation to ..." is developed.

Key words: computer lexicography, electronic explanatory dictionary, theory of semantic states, explicit representation of information.