

О. В. Іванова,

Національний університет біоресурсів і природокористування України, м. Київ

ТЕЗАУРУС МОВИ І ЛІНГВІСТИЧНЕ КОНСТРУЮВАННЯ

Висвітлено питання важливості та необхідності укладання тезаурусів, що задовольняють потреби сучасного інформатизованого суспільства; подано загальну класифікацію тезаурусів; проаналізовано структурні параметри словника-тезауруса й розкрито сутність кожного з них; описано поняття «лінгвістичне конструювання» та його роль у створенні нових «об'єктів».

Ключові слова: тезаурус, концепт, контекст, знак, семантичні зв'язки, дескриптор, лінгвістичне конструювання.

Раскрыто вопросы важности и необходимости составления тезаурусов, которые удовлетворяют потребности современного информатизированного общества; представлена общая классификация тезаурусов; проанализировано структурные параметры словаря-тезауруса и раскрыто сущность каждого из них; описано понятие «лингвистическое конструирование» и его роль в создании новых «объектов».

Ключевые слова: тезаурус, концепт, контекст, знак, семантические связи, дескриптор, лингвистическое конструирование.

Questions of necessity and importance of composing thesauruses that satisfy the demands of modern computerized society are opened; a general thesauruses classification is provided; structural parameters of thesaurus are presented and the essence of each of them is disclosed; the concept of «linguistic construction» and its role in the creation of new «objects» is described.

Key words: thesaurus, concept, context, sign, semantic connections, descriptor, linguistic construction.

Постановка проблеми. В умовах стрімких соціокультурних змін, динамічного зростання обсягу знань, значних досягнень техніки, новітніх інформаційних технологій виникають нові завдання для лінгвістів, а саме: необхідність експериментувати з мовою. На сьогодні одним із важливих завдань лексикографії є проектування таких словників, що б на рівні світових стандартів задовольнили потреби інформатизованої спільноти в систематизованій лінгвістичній інформації. У слові, як у краплі води океан, відображена вся мовна система, і пізнання її змістовних властивостей виявляється найбільш ефективним, якщо зосередитись на певному конкретному об'єкті, що відбиває ці системні відносини. Таким об'єктом у лексиці й є зокрема тезаурус, який не лише інвентаризує, а й систематизує лексичні одиниці у межах певної мовної підсистеми та є об'єктом підвищеної уваги фахівців.

Питання необхідності тезауруса, зокрема тезауруса літературної мови, може здатися недоречним і навіть абсурдним з точки зору наукової діяльності. До ідеї створення такого словника вчених спонукає сама логіка розвитку загального мовознавства та лексикографії. Тезаурус розширює число можливих «точок доступу» до словникового складу мови, збільшуючи тим самим розмір семантичного простору лексики. Він дає наближення до цілісного уявлення про лексичну систему в її ідеальному розумінні. Саме це завжди виступає головним аргументом при обґрунтуванні необхідності укладання такого роду словників.

Аналіз останніх досліджень із даної проблеми. Серед наукових праць, актуальних і на сьогодні, найбільш відомі такі тезауруси: словник П. Роже для англійської мови, П. Буассьєра – для французької, Ф. Дорнзайфа – для німецької, Х. Касареса – для іспанської. Проблемі створення тезаурусів присвячено низку наукових досліджень, що висвітлені у працях Ш. Баллі, Л. Щерби, Н. Шведової, В. Морковкіна, Ю. Караулова, Ю. Апресяна, О. Баранова, І. Мельничук, М. Кронгауз. Серед вітчизняних дослідників тезауруса доцільно відзначити роботи Н. Драчук, В. Дубічинського, І. Гетьмана, М. Сніжко, Н. Сніжко.

Мета статті – дослідити й узагальнити відомості щодо створення тезаурусів мови, а також способів і прийомів компіляції та комбінування «зразків вирішення лінгвістичних завдань».

Реалізації поставленої мети сприяє розв'язання таких **завдань**:

1) розкриття сутності основних дефініцій зазначеної проблеми; 2) аналіз питання структурних параметрів словника-тезауруса; 3) з'ясування поняття «лінгвістичного конструювання» та його ролі у створенні нових «об'єктів»; 4) дослідження тезауруса як семантичної матриці з чотирма входами, що задає лексичну систему мови в чотирьох вимірах.

Виклад основного матеріалу дослідження. Л.В. Щерба вживав термін «тезаурус» в його першому споконвічному значенні – «скарбниця», збірка усіх слів мови [9, с.4]. По відношенню до розглянутого типу словників він користувався терміном «ідеологічний», характеризуючи словники цього типу за їх властивостями, що, насамперед, відрізняють їх від звичайних, тобто тлумачних словників. У теоретичному відношенні тезаурус має невичерпну значущість оскільки структурує уявлення людей про світ.

Під словниками тезаурусного типу за Є. Архиповою, Л. Шишкіною, ми розуміємо такий словник, у якому представлені ключові лексичні одиниці певної предметної галузі разом із їх взаємозв'язками: синонімія, родові та видові відношення, асоціативні зв'язки, які класифікуються в узагальнених схемах, за допомогою яких користувач може скласти структуру предметного змісту в різних видах мовленнєвої діяльності, виражати різноманітні семантичні відносини, розвивати думку за певною логічною схемою [1, с. 11]. Лексикон тезаурусного типу слід розглядати як спосіб організації, введення й закріплення змістової структури знань окремої галузі у вигляді системи лексичних засобів з їх взаємозв'язками і відношеннями [3, с. 20].

Вітчизняний лексикограф І. Гетьман тлумачить поняття тезаурус, що втілює у своїй семантиці систему загальнозживаних слів, термінів, крилатих висловів та інших одиниць, у яких реалізовано наукову й повсякденну картину світу [2, с.109].

Характерною й разом з тим відмінною рисою тезаурусів є те, що слова в них розміщені не за алфавітним порядком, а на основі логіко-філософської класифікації предметів і явищ дійсності.

К. Марелло виділяє три види тезаурусів:

- 1) кумулятивні (групуєть слова без визначення їх значень);
- 2) дефінітивні (тлумачать кожен лексичну одиницю групи слів);
- 3) дво- та багатомовні тезауруси для подорожуючих [4, с. 82].

Словники, що створені за алфавітним принципом, руйнують семантичні зв'язки між словами й їх контекстами, у тезаурусах збережено семантичні зв'язки між одиницями мови. Головна мета тезауруса – подати правильне слово для концепту в необхідному контексті. Концепт відображає категоріальні й ціннісні характеристики знань про певні фрагменти світу. Саме концепти й є основними носіями змісту, що актуалізуються за допомогою мовних одиниць, і виникають в процесі когнітивної діяльності людини.

Створення докладного галузевого тезауруса – це значна за обсягом робота, що вимагає багато часу і залучення широкого кола фахівців. Тезаурус має вигляд словника, в якому в певному порядку розташовані терміни і зафіксовані зв'язки між ними. Як правило, тезаурус складається з алфавітного і систематичного покажчика. Алфавітний покажчик містить алфавітний перелік дескрипторних статей. Систематичний покажчик служить для розкриття, обліку та контролю парадигматичних відносин між дескрипторами. Для кожного інформаційно-пошукового дескриптора повинні бути знайдені відповідні родові та видові дескриптори.

Тезаурус фактично визначає семантику мови (національної мови, мови конкретної науки чи формалізованої мови для автоматизованої системи управління). Тезаурус представляє лексичний інструмент інформаційно-пошукових систем. Він складається з контрольованого, але змінюваного словника термінів, між якими існують змістові зв'язки. Такий словник, вичерпно покриває певну специфічну галузь знань і поєднує перелік дескрипторів і недескрипторів (допоміжних термінів). Центральне місце у будь-якому тезаурусі займає його систематична частина, що є алфавітно-впорядкованим переліком дескрипторів (тобто основних понять з відповідними до них визначеннями) та ключових слів (також є різновидом основних понять, позбавлених самостійної «дескрипторної» статті, оскільки пов'язані з ними слова входять до включаючого їх дескриптору). Ключові слова є так званіми нитками, що об'єднують внутрішню структуру тезауруса. В повному обсязі тезаурус складається з чотирьох частин – систематичної, або лексико-семантичного покажчика термінів; класифікаційної, або покажчика ієрархічних відносин дескрипторів; алфавітного покажчика термінів; пермутаційного покажчика термінів (дескрипторів) [10].

Понятійна класифікація як спосіб систематизації термінів при тезаурусному описі зумовлює структурні параметри словника-тезауруса. Ю.Н. Караулов виділяє в структурі тезауруса декілька взаємопов'язаних частин. Класифікаційна схема розділяє усю сукупність понять на великі понятійні зони чи області й системно їх впорядковує. В ідеографічній частині під кожним поняттям згрупована уся лексика, що до нього відноситься. У пермутаційний покажчик понять-дескрипторів включені всі значущі слова, що входять до складу термінів тезауруса. Після кожного значущого слова в алфавітному порядку наводяться усі терміни тезауруса, що містять значуще слово у якості терміноелемента, при чому кожен термін зустрічається в пермутаційному покажчику стільки разів, скільки значущих слів він містить. Пермутаційний покажчик є відображенням синтагматичних відношень термінів певної галузі, їх сумісності.

Термінологічний словник-тезаурус відрізняється від інших словників і тим, що він представлений у вигляді «об'ємної» когнітивної моделі певної галузі знань чи людської діяльності за допомогою декількох входів у словник:

- 1) від «концепту до концепту», тобто від одного поняття до іншого з поданням ієрархічних (вертикальних) і кореляційних (горизонтальних) зв'язків;
- 2) від «концепту до знака» через ідеографічну частину тезауруса з поданням понятійно-лексичних зв'язків;
- 3) від «знака до концепту» за допомогою алфавітного покажчика адрес слів у понятійних полях дескрипторів з представленням лексико-понятійних зв'язків;
- 4) від «знака до знака» через пермутаційний покажчик, що використовується для пошуку словосполучень за одним із складових компонентів [7].

Таким чином, структура тезауруса, представляючи собою своєрідну семантичну матрицю з чотирма входами, задає лексичну систему мови в чотирьох вимірах: «концепт-знак», «концепт-концепт», «знак-концепт», «знак-знак»[6].

Створити, тобто побудувати якусь «конструкцію», означає не тільки вміти пояснити ті властивості мови, що в ній використані і на яких вона базується, не лише показати певні закономірності мовної структури, а й з'ясувати нові властивості побудованого об'єкта, які так чи інакше характеризують досліджувану мову, а значить – розширити знання про людську мову загалом. Таким чином, новим лінгвістичним об'єктом можна називати предствалення фактів чи даних, що генерують нову інформацію про мову. Як правило, такі об'єкти виникають не в результаті опису певного матеріалу, а з'являються в результаті експерименту.

Лінгвістичне конструювання – це сукупність узагальнених способів і прийомів компіляції та комбінування «зразків вирішення лінгвістичних завдань», екстраполяції вже наявних, готових теоретичних і практичних знань, отриманих у різних областях лінгвістики, і їх прямого або евристичного використання для подолання труднощів та вирішення проблем, що виникають в процесі побудови нових лінгвістичних об'єктів [6, с.84]. Особливістю лінгвістичного конструювання як комп'ютерної технології створення нових лінг-

вістичних об'єктів, є його межування з традиційною лінгвістичною практикою в тих областях, де ступінь комп'ютеризації незначний. При цьому можливими є наступні варіанти: а) електронне лінгвістичне видання створюється на основі традиційного, тобто, наприклад, поліграфічно виданий словник трансформується в електронну (комп'ютерну) форму; б) традиційне поліграфічне видання створюється в результаті спеціальної обробки електронного лінгвістичного матеріалу, наприклад, база даних словника – основа для формування оригінал-макету його поліграфічного видання, система баз даних текстів джерел, картотек та словників – основа поліграфічного видання, в) складне переплетення першого і другого варіантів, результатом якого є як електронні, так і поліграфічні версії лінгвістичних видань.

Отже, лінгвістичне конструювання – це не лише спосіб створення нових «об'єктів», а й метод вивчення, виявлення основних положень мовного матеріалу, що передбачає як аналіз так і його синтез [6, с.87]. Тобто, при лінгвістичному конструюванні, в ході побудови нових лінгвістичних об'єктів, сама мова виступає як предмет експерименту і як суб'єкт наукового дослідження. Фактично мова йде про те, що ми маємо два типи нових лінгвістичних об'єктів: по-перше, новим лінгвістичним об'єктом є результат, отриманий за допомогою інформаційних лексикографічних технологій в електронному або традиційному вигляді, умовно його можна назвати об'єктом першого типу; по-друге, новим лінгвістичним об'єктом є автоматизована (комп'ютеризована) діяльність лексикографа, вміщена в наборі описових інструкцій, операцій та інших документах, а також у виробничих навичках конкретних осіб і колективів. Таку діяльність слід відрізнити від діяльності, спрямованої на використання об'єктів першого типу, вона інша та призначена для внутрішнього, «цехового» використання, тобто є інструментальною. Часто для позначення такої діяльності використовується поняття «know-how». Такий лінгвістичний об'єкт називають об'єктом другого типу.

Що ж являють собою лінгвістичні об'єкти першого типу? У роботі Ю.Н.Караулова наведений «тезаурус літературної російської мови, побудований за допомогою електронно-обчислювальної машини, діюча модель створення словозмінних форм (тобто автоматичний морфологічний словник, призначений як для аналізу, так і для синтезу)» [6, с.219]. На сьогоднішній день це широкий набір лінгвістичних інформаційно-програмних засобів різного призначення, а також лінгвістичних виробів у поліграфічній формі, отриманих з використанням комп'ютерних технологій.

Лексикографічний і тезаурусний опис відносять до методів формалізації знань, тобто методів подачі знань суб'єкта про предметну галузь з використанням знакових систем. При лексикографічному (словниковому) описі об'єкти і процеси описані непослідовно й незв'язно у вигляді системи слів і їх визначень, пояснень, а результатом лексикографічної формалізації знань виступають словники, конкорданси. При тезаурусному описі мови знання фрагментуються й структуруються таким чином, що вони виявляються розділеними на окремі групи понять, пов'язаних між собою певними відношеннями [8]. Тезаурус призначений відновити семантичні зв'язки між одиницями мови, що порушені в словниках, створених за алфавітним принципом [4, с. 80].

Висновки. Підводячи підсумки нашого дослідження, ми дійшли висновку, що в словники виконують важливі функції, дозволяють найкоротшим способом долучитися до накопичених знань, даючи користувачам необхідні знання з рідної або іноземної мови. Фіксує значення та особливості вживання слів, словники сприяють вдосконаленню та уніфікації мови як засобу спілкування. Словники виникли в глибоку давнину і з часом набували все більшого й більшого значення, відіграючи неабияку роль у накопиченні й передачі інформації. Також особливе місце займають тезауруси, що роблять лексичне конструювання невід'ємною частиною розвитку мовознавства, лексикографії та лінгвістики в цілому.

Перспективи подальших розвідок ми вбачаємо у дослідженні етапів

створення електронних словників-тезаурусів, які, відображаючи всі нові процеси суспільного життя, задовольнятимуть потреби різноманітних користувачів.

Література:

1. Архипова Е. И. Формирование иноязычного лексикона специалиста в интегративном обучении иностранному языку и общепрофессиональным дисциплинам: дисс. ... канд. пед. наук. Екатеринбург, 2007. – 242 с.
2. Гетьман И. М. Общая идеография: Конспекты лекций по спецкурсу / И. М. Гетьман. – Нежин, 1993. – 154 с.
3. Дубичинский В. В. Теоретическая и практическая лексикография. / В. В. Дубичинский. – Вена – Харьков, 1998. – 156 с.
4. Дубичинский В. В. Лексикография русского языка: учебное пособие / Дубичинский В. В. – М. : Наука: Флинта, 2008. – 432 с.
5. Зарецкая Е. Н. Риторика: Теория и практика речевой коммуникации. – 4-е изд. – М. : Дело, 2002. – 480 с.
6. Караулов Ю. Н. Лингвистическое конструирование и тезаурус литературного языка / Ю. Н. Караулов. – М. : Наука, 1981. – 363 с.
7. Кобрин Р. Ю., Антонова М. В. Терминологические системы и их когнитивные модели // Очерки научно-технической лексикографии. – СПб. : Изд-во С.-Петерб. ун-та, 2002. – С. 47-66.
8. Филиппович Ю.Н., Прохоров А. В. Семантика информационных технологий: Опыт словарно-тезаурусного описания / Ю. Н. Филиппович, А. В. Прохоров. – М. : МГУП, 2002. – 368 с.
9. Щерба Л. В. Предисловие ко второму изданию «Русско-французского словаря» / Л. В. Щерба // Русско-французский словарь. Сост. Л. В. Щерба, М. И. Матусевич / Под общ. ред. Л. В. Щербы. – М. : Высшая школа, 1959. – С. 4-7.
10. Тезаурус [Електронний ресурс]. – Режим доступу : <http://ru.wikipedia.org/Тезаурус>