

УДК 81 ' 42:004.2 (075.8)

ТЕРМІН У ЛІНГВІСТИЧНІЙ ІНФОРМАТИЦІ**Дарчук Н.П., Алексієнко Л.А., Сорокін В.М.****Київський національний університет імені Тараса Шевченка**

Стаття присвячена опису лексикографічної й енциклопедичної електронної бази українських лінгвістичних термінів, яка складається з чотирьох словників: алфавітного, перекладного, тлумачного та тезаурусного обсягом 3400 термінів (з морфології, синтаксису, лексики, семантики, комп'ютерної лінгвістики) та технології інформаційно-пошукової системи-тезауруса, що може функціонувати як довідково-енциклопедична система або у складі іншої інтелектуальної системи.

Ключові слова: лінгвістичне моделювання, термін, інформаційно-пошукова система, тезаурус, тезаурусні зв'язки.

Використання сучасних інформаційних технологій в лінгвістиці було поштовхом до формування в останні десятиріччя такої науки, як лінгвістична інформатика, що вивчає закони та методи організації і перероблення за допомогою комп'ютера лінгвістичної інформації, конкретніше – це сукупність законів, методів і засобів одержання, збереження, передавання, розповсюдження, перетворення інформації про мову і закони її функціонування за допомогою комп'ютера, які втілюються у такі комплексні завдання, як створення систем штучного інтелекту, автоматичного перекладу, автоматичного реферування й анотування текстів, інформаційно-пошукових систем, розроблення різних баз даних, в тому числі термінологічних тощо. При цьому основним методом вирішення різних завдань інформаційного характеру був і залишається метод моделювання.

Логіко-понятійне моделювання терміносистем окремих наук і галузей знань – одне з актуальних міжгалузевих завдань нашого часу, тому що моделі терміносистем необхідні при укладанні термінологічних словників, інформаційних тезаурусів, класифікаторів, рубрикаторів, створенні автоматизованих інформаційних систем, баз даних, систем штучного інтелекту. Частковим випадком моделювання знань можна вважати побудову інформаційно-пошукового тезауруса (ІПТ), який, з одного боку, є способом формалізованого представлення термінології, оскільки досить строго представляє семантичні відношення між термінами, а, з іншого – вважається важливим інструментом постійного вдосконалення систем знань конкретних наук.

Наш досвід науково-викладацької роботи свідчить, що в Україні цій проблемі не приділяється належної уваги. Сьогодні мовознавці мають справу з концептуально й технологічно застарілими книжковими словниками, які не відображають здобутків лінгвістики, принаймні за останні 50 років. Можна констатувати відсутність лінгвістичних та інформаційних технологій, які передбачають створення тезауруса як спосіб систематизації лінгвістичної термінології і як інструмент інформаційного пошуку. З праць у галузі лінгвістичної термінології, виконаних у Росії, Німеччині, Америці, Канаді та ін., можна відзначити тезаурус з мовознавства [5], фрагмент тезауруса на матеріалі словника О.С.Ахманової [6], роботи А.А.Війтухновської [1], в яких розглядаються деякі аспекти побудови блочних тезаурусів з мовознавства та загальної психології, тезаурус з теоретичної і прикладної лінгвістики С.Є.Нікітіної [4], тезаурус з мовознавства Гадумода Буссмана [7], тезаурус фасетного типу Харріса, Сайна, і Ласковського [8] та ін. Отже, усвідомлення такого стану української лінгвістичної термінографії, з одного боку, а також напрацювання авторів у сфері тезаурусотворення – з другого, стали внутрішніми і зовнішніми стимулами Проекту, який виконується в лабораторії комп'ютерної лінгвістики Інституту філології Київського національного університету ім.Т.Шевченка.

Мета проекту: 1) укладання електронного **Словника лінгвістичних термінів** з використанням **нової формалізованої методики конструювання тезауруса**, що відповідає сучасним стандартам термінографії, та представлення його в мережі Інтернет; 2) верифікація теоретичної тезаурусної моделі згідно з розробленими комп'ютерними технологіями шляхом застосування її для аналізу корпусу текстів з різних розділів лінгвістики.

Робота над проектом здійснювалася у два етапи. На **першому етапі** створювалася інформаційно-пошукова система (ІПС) у вигляді лексикографічної й енциклопедичної електронної бази українських лінгвістичних термінів, яка складається з чотирьох словників: **алфавітного, перекладного, тлумачного і тезаурусного**. В **алфавітному словнику** для кожного слова-терміна або термінологічного словосполучення (3400 термінів) надається **англійський еквівалент** (у перспективі планується російський, німецький, французький, італійський еквіваленти) і **тлумачення** (рис. 1). **Тезаурусний словник** представляє собою перелік логіко-семантичних відношень між лінгвістичними термінами (список запозичено з роботи [4], але доповнено і модифіковано нами). Тобто, розроблена ІПС включає не тільки множину окремих термінів, представлених у вигляді алфавітного списку з їхніми тлумаченнями, а й самі моделі представлення зв'язків між термінами (рис. 1).

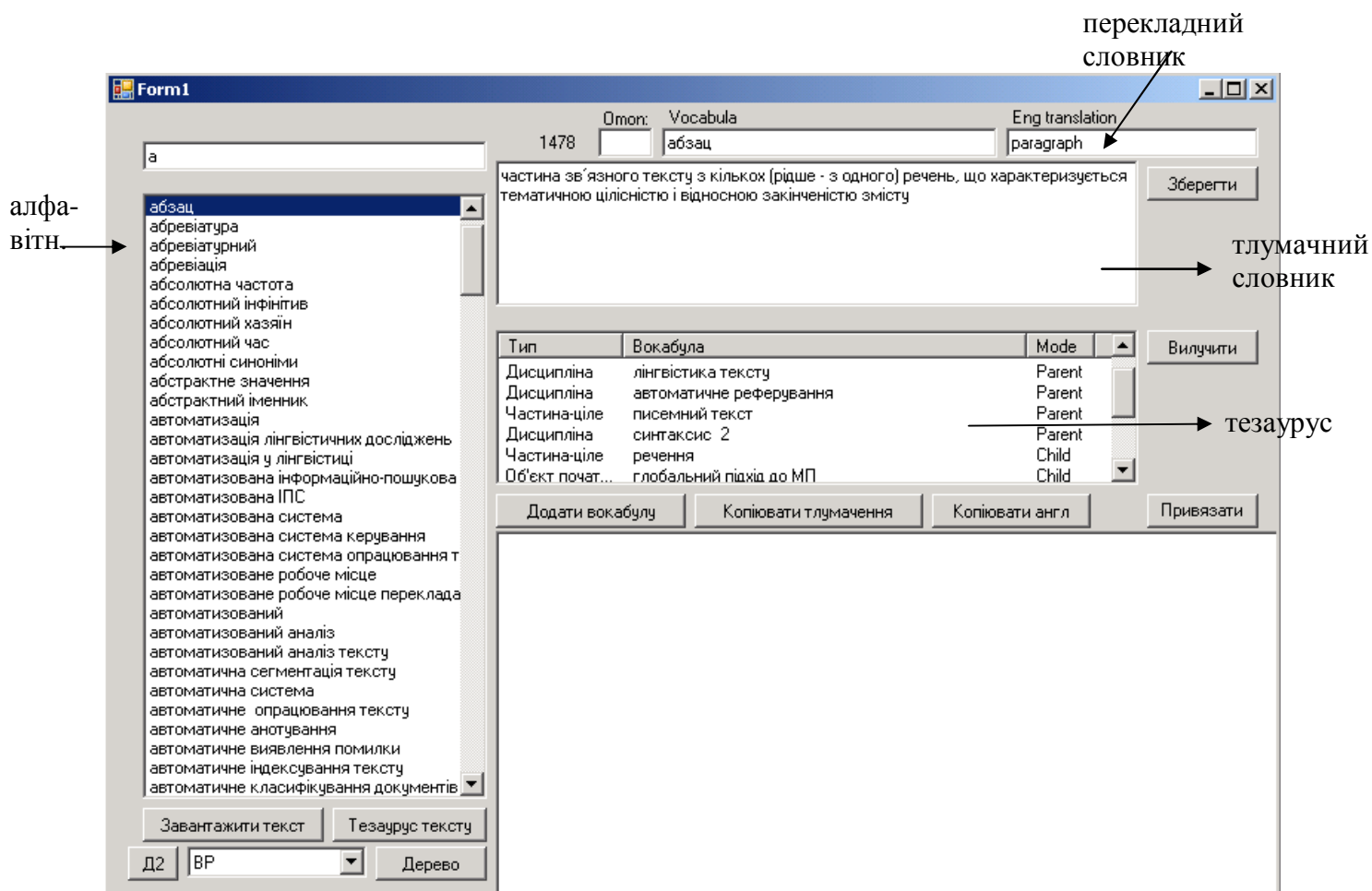


Рис. 1. Фрагмент електронного словника лінгвістичної термінології

На основі досягнень сучасного мовознавства в компактній та доступній формі подано тлумачення термінологічних одиниць з авторитетних джерел (близько тридцяти): термінологічних словників, граматик, монографій. До словника включаються загальнолінгвістичні терміни, переважно іменники або іменникові словосполучення, з усіх розділів граматики, лексикології, прикладної та комп'ютерної лінгвістики. Словниковими одиницями є: вживані загальнолінгвістичні терміни з морфеміки, словотвору, парадигматики, синтаксису, лексикології, семантики; терміни окремих прикладних напрямів,

відомих в українській, російській, зарубіжній лінгвістиці; терміни з комп'ютерної лінгвістики, пов'язані з автоматизацією лінгвістичних процесів. Відбір словникових одиниць до бази даних здійснювався на евристичних засадах (знання укладачів тезауруса, експертів-лінгвістів).

Побудова тезауруса (ТЗ) передбачає розкриття всіх типів відношень між поняттями, вираженими термінами, основними з яких є гіпонімія (рід - вид); супідрядність на одному рівні – парціація (частина - ціле); синонімія; кореляція; асоціація; локалізація об'єкта; його призначення; функція; способи вираження функції; відношення тощо. Зміст відношень розширено настільки, щоб можна було охопити максимально широкий пласт термінів, з якими зв'язаний аналізований термін як реєстровий. Оскільки зміст тлумачення нерідко був недостатнім для здобуття всіх істотних для термінів відношень, ми орієнтувалися на енциклопедичні словники, наукові праці з конкретної проблематики, знання власні та фахівців-лінгвістів. Словникова стаття побудована у вигляді анкети, яка заповнювалася для кожного терміна. В анкеті вміщено стандартний перелік відношень, які щодо реєстрового слова є поняттєвими. Назва відношення є **двомісним предикатом R (A,B)**, який зв'язує заголовне слово статті (A) і введений цим предикатом термін (B) [3, 22].

Тезаурус містить **3394** терміни, які охоплені семантичною мережею із **9265** семантичних відношень (табл. 1).

Табл.1.

Кількісні характеристики семантичних відношень

Тип семантичних відношень	Термін А (приклад)	Термін В (приклад)	Частота реалізації семантичних відношень у ТЗ
1	2	3	4
Рід – Вид (<i>A</i> є родовим до <i>B</i>)	частини мови	несамостійні частини мови	1843
Дисципліна (<i>A</i> розглядається в дисципліні <i>B</i>)	абзац	синтаксис	1157
Синоніми (<i>A</i> синонімічний <i>B</i>)	автоматичний переклад	машинний переклад	1016
Частина – Ціле (<i>B</i> складається з <i>A</i>)	абзац	писемний текст	934
Корелят (<i>A</i> протилежний <i>B</i>)	агент	пацієнс	822
Дивись...(<i>про A</i> дивись <i>B</i>)	акузатив	знахідний відмінок	529
Асоціація (<i>A</i> асоціюється з <i>B</i>)	субморф	морф	312
Операція/процедура (для <i>A</i> процедурою / операцією буде <i>B</i>)	абревіатура	абревіація	263
Об'єкт початковий (<i>B</i> здійснюється над <i>A</i>)	речення	автоматична сегментація тексту	232
Інструмент/Метод (<i>A</i> із застосуванням <i>B</i>)	автоматичне анотування	мережеве моделювання	221
Параметр (<i>A</i> характеризується <i>B</i>)	аломорф	додаткова дистрибуція	205
Носій параметра (носієм параметра <i>A</i> є <i>B</i>)	абревіація	іменник	210
Об'єкт кінцевий (<i>A</i> здійснюється над <i>B</i>)	автоматичний синтез тексту	текст	202

1	2	3	4
Відноситься до... (стосується термінів-а'єктивів: <i>A</i> відноситься до B)	<i>акузативний</i>	<i>акузатив</i>	189
Спосіб вираження (<i>A</i> виражається B)	<i>особове дієслово</i>	<i>особова форма дієслова</i>	184
Рівень мови (<i>A</i> розглядається на рівні, позначеному B)	<i>актантна структура</i>	<i>синтаксичний рівень</i>	176
Функція основна (<i>A</i> виражає B)	<i>вигук</i>	<i>експресивність</i>	157
Відношення (<i>B</i> зв'язує A)	<i>головний член</i>	<i>безпосередня синтаксична залежність</i>	130
Спосіб представлення об'єкта (<i>A</i> представляється через B)	<i>словотвірна структура слова</i>	<i>дериваційна історія слова</i>	109
Аспект (<i>A</i> розглядається в аспекті B)	<i>графіка</i>	<i>писемний текст</i>	99
Лінгвістичний об'єкт (<i>A</i> представляється у вигляді B)	<i>зв'язний текст</i>	<i>семантична мережа</i>	96
Імплікація (<i>якщо A, то B</i>)	<i>знахідний відмінок</i>	<i>перехідне дієслово</i>	81
Одиниця рівня (<i>одиноцею A є B</i>)	<i>морфологічний рівень</i>	<i>словоформа</i>	40
Клас (<i>A</i> входить до класу B)	<i>аломорф</i>	<i>морф</i>	36
Об'єкт науки (<i>A</i> є об'єктом B)	<i>природна мова</i>	<i>інженерна лінгвістика</i>	22

Найголовніші парадигматичні семантичні відношення (рід-вид, синонімія, частина-ціле, кореляти), як видно з таблиці, охоплюють значну частину термінів (майже 70 % всіх парадигматичних відношень). З точки зору теоретичної семантики, чим більше у словнику міститься семантичної інформації, тим краще, тому що багата система відношень в ТЗ надає користувачеві більше можливостей виражати в запиті свою інформаційну потребу. Оскільки словникова стаття являє собою синтез інформації лінгвістичної, перекладної, тлумачної та енциклопедичної, у зв'язку з інформаційним підходом, розрахованим на запит користувача в ПС, реалізовано діалогову систему на зразок:

Запит: «*який термін є родовим до терміна абстрактний іменник?*»

Відповідь: «*Родовим є термін іменник*».

Запит: «*З якими термінами і якими зв'язками пов'язаний термін абстрактний іменник?*»

Відповідь: «*Грамматичний рід*» - відношення: *носій параметра*; «*Абстрактне значення*» - відношення *носій параметра*; *відношення імплікації*.

Процедурно відповіді на запит здобуваються з тезаурусного графа, який має вигляд семантичної мережі, що представляє собою ієрархічно організовану структуру даних – термінів-вузлів і дуг, які виражають різні типи тезаурусних відношень, й автоматично подається з тезаурусу у текстовому вигляді (рис. 2).

Чорним кольором на екрані комп'ютера позначаються відношення «хазяїн (х.) – слуга (с.)», де реєстровий термін є «хазяїном», що підпорядковує, а синім – «слуга - хазяїн», тобто, навпаки, є слугою, отже, підпорядкованим.

У тезаурусі є терміни, охоплені розгалуженою мережею семантичних відношень: речення – 121 (х.= 24, с. = 97); слово – 62 (х.= 30, с.= 32) тощо.

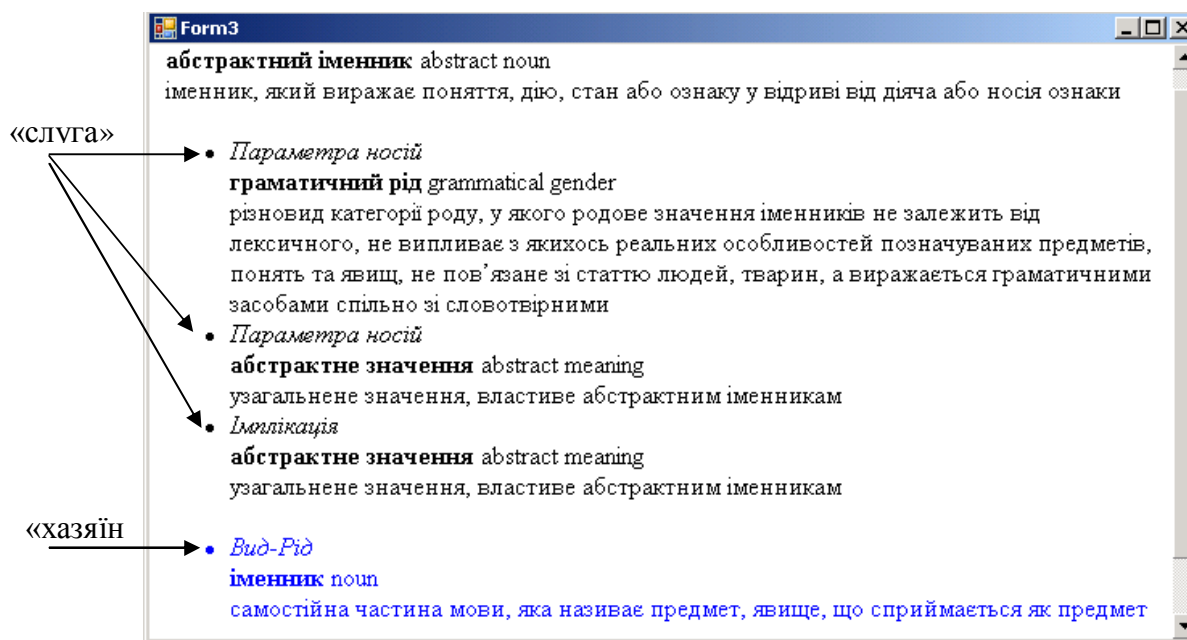


Рис.2. Тезаурусний граф

Оскільки кожне з відношень має свою мітку і номер в електронному форматі, це дозволяє одержувати дані, згруповані так, як це потрібно користувачеві. Представлену модель тезауруса можна розглядати як семантичний та інформаційний опис лінгвістичної термінології, яка модифікується як: а) алфавітний словник з дефініціями; б) блочний термінологічний словник за дисциплінами (морфологія, когнітивістика, комп'ютерна лінгвістика...); в) словник синонімів; г) словник родо-видових відношень; д) автоматичний інформаційний довідник тощо.

Мережеве представлення даних має не тільки чисто прикладне значення, а дозволяє глибше проникнути в логічні відношення даної науки, точніше змоделювати аналізовану терміносистему.

Створений тезаурус є **статичною моделлю** логіко-поняттєвих відношень між термінами лінгвістичної метамови. Можна розглянути цю модель у плані відображення **динамічних аспектів** структури наукового знання, а саме як верифікацію теоретичної тезаурусної моделі шляхом застосування її для аналізу корпусу текстів з різних розділів лінгвістики. Важливість такого дослідження пояснюється тим, що будь-яке знання має текстове вираження і пізнається через тексти. Текст – якісно інша форма існування наукового знання. Тут знання функціонує вже як семантика тексту, а саме знання поступово набуває статусу наукової інформації [2, 50]. Безперечно, найбільш повно знання про об'єкт, його ознаки і властивості відображається у жанрі наукової монографії (підручника, дисертації), тому що в науковій монографії всебічно розглядаються міждисциплінарні зв'язки. Однак не можна нехтувати й іншими типами текстів – доповідями, статтями, дискусіями тощо, які по-різному виражають наукове знання.

Врешті будь-яка енциклопедична модель наукового знання є похідною від множини реальних текстів і є репрезентацією цих текстів на рівні семантичної моделі. Семантична модель мови науки – це інваріант мікромоделей окремих часткових текстів, по відношенню до яких тексти виступають як його варіанти. По відношенню до множини термінів певної науки логіко-понятійна система галузі знання – це модель плану змісту галузі знання. Логіко-понятійна структура тексту відображає основні елементи семантичної парадигматики текстів. Моделювання семантики наукових текстів передбачає виявлення всіх типів семантичних відношень між усіма словами – назвами понять, їхніх ознак і властивостей тощо.

При дослідженні терміна у тексті звичайно застосовуються два підходи: від тексту до терміна (тобто термінологічний аналіз тексту), і від терміна до тексту (текстовий аналіз

терміна), які дають різні теоретичні й практичні результати. Терміни формуються у тексті, а фіксуються у терміносистемі (словнику). І якщо у сфері функціонування термін існує у вигляді парадигматичних і синтагматичних варіантів, то у терміносистемі – тільки у вигляді парадигматичних варіантів [3, 152].

Видається перспективним поєднати ці підходи, розробивши методику, за якою створена тезаурусна модель лінгвістичних термінів у вигляді ієрархічної класифікаційної схеми – графа накладається на словниковий реєстр наукового тексту. В результаті ми одержуємо також ієрархічний класифікаційний граф конкретного аналізованого тексту з абсолютною частотою вживання терміна у конкретному тексті. Такий аналіз забезпечує другий підхід – від термінологічного словника до тексту. Наступний методичний крок – створення додаткового реєстру слів за текстом з відповідними частотами без урахування тих термінів, які увійшли до тезаурусного графа. Шляхом перегляду слів одержаного реєстру можна сформулювати список слів, які вважатимуться претендентами на терміни. Цим забезпечується підхід – від тексту до словника, причому частота вживання і контексти дають можливість концентрувати розпорошену термінологічну інформацію для розв'язання різноманітних термінологічних завдань, зокрема аналізу доцільності включення претендента до створеного словника термінів.

Саме тому було передбачено **другий етап** цього Проекту – побудова динамічної логіко-понятійної моделі кожного конкретного тексту з корпусу текстів лінгвістичної тематики.

Методика складається з таких основних етапів: 1) лематизація та впорядкування лем за частиномовною приналежністю; 2) визначення для кожної лєми (іменникової або прикметникової) абсолютної частоти вживання; 3) побудова тезаурусного графа термінів конкретного тексту з абсолютними частотами вживання у тексті шляхом накладання тезаурусного графа терміносистеми; 4) зняття омонімії значень термінів (напр., **граматика 1** (*будова мови, тобто система морфологічних одиниць, категорій і форм, синтаксичних одиниць і категорій, словотвірних одиниць і способів словотворення*); **граматика 2** (*розділ мовознавства, що вивчає граматичну будову мови*) і **граматика 3** (*формальна граматики, що є частиною системи автоматичного опрацювання тексту*)) через звертання до контексту; 5) побудова допоміжного реєстру слів з абсолютними частотами, які не увійшли до ТЗ; 6) пошук слів-претендентів на терміни і фіксація їх з ілюстративним контекстом у додатковому термінологічному реєстрі. Робота цих етапів забезпечується пакетом алгоритмів і програм 1 – 3 і 5-го етапів в автоматичному, а 4-го і 6-го – в автоматизованому режимі. Перевірка роботи методики здійснювалася на матеріалі корпусу наукових статей з мовознавства, тому що стаття як одна з форм існування наукового знання подає обговорення одного з аспектів, однієї зі сторін, пов'язаної з тим чи іншим науковим поняттям, з тією чи іншою науковою проблемою, забезпечуючи цим широкий спектр наукових знань, їх семантику, репрезентовану термінологічно (тексти 72 статей збірника «Українське мовознавство»; довжина тексту 82418 слововживань).

У результаті автоматичного накладання тезаурусного словника у вигляді орієнтованого графа на загальний реєстр корпусу наукових статей з мовознавства було одержано реєстр також у вигляді орієнтованого графа. Якщо вважати попередньо створений ТЗ інваріантною моделлю лінгвістичної галузі знань, то по відношенню до аналізованих текстів ТЗ виступає як модель плану змісту: в ньому семантична мережа представляє ієрархічно організовану структуру даних, у якій виділяються вузли – терміни – найменування понять, а дугами відображаються відношення між цими поняттями (рис.3).

Реєстр обсягом у 821 термін покриває 16,12 % тексту, (кумулятивно 13287 термінів від 82418 слововживань тексту). У той же час у корпусі текстів відтворено 24% термінів з лінгвістики, зафіксованих у тезаурусі (821 термін проти 3340 термінів у ТЗ).

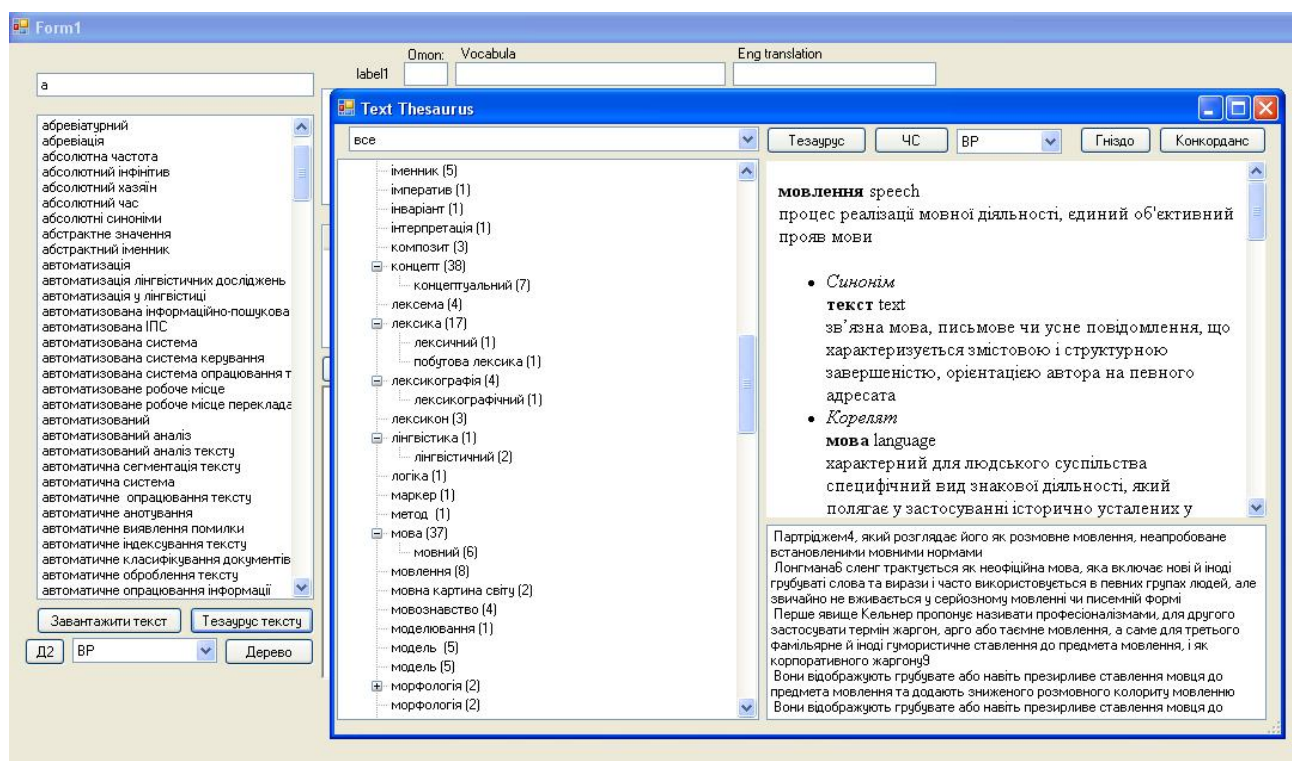


Рис. 3. Тезаурус термінів збірника статей «Українське мовознавство»

На наступному етапі перевірявся автоматичний допоміжний реєстр на предмет наявності в ньому термінів/терміносполук, не зафіксованих у загальнолінгвістичному ТЗ з кількох причин: по-перше, можливі авторські терміни, які вводилися у текст у зв'язку з необхідністю викласти позицію автора з певного теоретичного питання; по-друге, ніколи словник не може бути повним, тому що процес становлення, розвитку науково-технічної термінології не припиняється. Текст є тим джерелом, де формується термін формально і семантично, тільки звертаючись до тексту, можна прослідкувати «життя» терміна, остаточно вирішивши питання про доцільність включення його до термінологічного словника.

Пропонована методика дає можливість в автоматизованому режимі виявити, а потім вивчити термін у сфері функціонування. У текстах терміни народжуються разом з ідеями, поняттями, іноді супроводжуються формулюваннями на зразок «під цим терміном будемо розуміти...» тощо, що може бути підставою для створення реєстру претендентів на термін разом з контекстами, достатніми для остаточного вирішення питання доцільності, а в разі наведення тлумачення, відмінного від загальноприйнятого, підставою для збереження його разом з тлумаченням у спеціально створеному термінологічному банку даних (ТБД).

Тезаурусне представлення терміносистеми у вигляді семантичної мережі – це також частина опису мови науки. ІПТ представляє собою модель лексичної структури, оскільки тезаурус у явному вигляді відображає парадигматичну організацію лексики і дає можливість систематизувати під парадигматичним кутом усю основну лексику досліджуваного матеріалу, в даному випадку корпусу мовознавчих статей.

Тезаурус ілюструє семантичну безперервність: у словнику немає і не може бути термінів, ізольованих в семантичному відношенні. Кожний термін десятками і сотнями ниток зв'язаний зі значеннями багатьох інших. Це зв'язки родо-видові, парцелятивні, синонімічні, корелятивні тощо. Від семантичної мережі можна перейти і до тематичної на такій підставі: чим більше термінів (з їхніми кількісним показниками) у тексті пов'язані у дереві ТЗ семантичними зв'язками, тим вони більше тематично значущі у тексті. Темати є, по суті, назви тематичних груп тезаурусного дерева, з іншого боку, тематична мережа становить ієрархію провідних і додаткових тем, а вихідними будуть такі методичні положення: 1) тематично значимі термінологічні групи будуються на основі тезаурусного дерева;

2) ієрархія тем тексту визначається залежно від кількості термінів, які входять до тематичних груп.

З кожним новим науковим текстом у корпусі мовознавчої тематики модель ППТ оновлюється. Ця модель передбачає: 1) автоматичний відбір термінологічних лексем на основі збігу із загальним ТЗ; 2) автоматизований контроль з боку спеціаліста: виявлення термінів, не передбачених у ТЗ; надання їм тлумачення, перекладу, встановлення тезаурусних зв'язків.

Процес укладання лексико-тематичної моделі включав такі етапи:

1. лематизація текстових словоформ, створення частотного реєстру тексту та впорядкування його за частотомовною приналежністю;
2. підрахування абсолютної частоти, яка вказує, скільки термінів зустрічається у гілці та яка сукупна частота, з якою термін зустрівся у текстах;
3. зняття омонімії термінологічних лексем і перерахування частотних характеристик.

Важливість корпусного дослідження функціонування термінів не підлягає сумніву, оскільки в текстах наукового спрямування викладаються теорії, концепції, описуються методи тощо, а згодом терміни фіксуються у тексті словника (або стандарту, класифікатора). Тобто головним у термінознавстві є принцип первинності сфери функціонування терміна у так званих терміностворюючих і терміновикористовуючих текстах.

У процесі термінологічного аналізу тексту і текстового аналізу терміна застосовувалися:

- логічний метод класифікування і конструювання терміносистем зі встановленням зв'язків між виявленими термінами;
- логічний, лінгвістичний, термінологічний методи виявлення фактів сумісного використання термінів з нетермінами і загальнонауковими термінами, вивчення сполучень термінів з термінами у тексті.
- кількісні методи визначення вживаності термінів.

Розроблений текстозорієнтований ППТ з лінгвістичним апаратом і програмним забезпеченням може бути використаний при створенні термінологічних банків даних (ТБД) і термінологічних банків знань (ТБЗ), оскільки розроблено значні обсяги термінологічної інформації (3400 од.). Засоби автоматизації відбору, збереження і пошуку цієї інформації: зберігаючись у пам'ять комп'ютера, термінологічний словник лінгвістичних термінів може збагачуватися й розвиватися. Переваги такого словника перед паперовим у тому, що електронний ППТ уможливує вхід через будь-яку характеристику терміна. З іншого боку, термін не обмежений рамками своєї «словникової статті», як у паперовому словнику – в ТБД встановлюються логічні й асоціативні зв'язки між термінами: відношення роду і виду, парціації, синонімії, кореляції тощо. Програмно будуються семантичні мережі, які наочно відображають структуру багаторівневих класифікацій термінів у межах галузевої, лінгвістичної, терміносистеми.

Важливість Проекту полягає в тому, що, по-перше, ППС у мультимедійному просторі забезпечить широке коло мовознавців сучасним стандартизованим словником лінгвістичних термінів; по-друге, здобутком Проекту є методика конструювання тезаурусів, а також комп'ютерний інструментарій для реалізації цієї методики; по-третє, ППС сумісний з інтелектуальними системами опрацювання текстової інформації, в яких він може бути використаний як база знань та інструмент розпізнавання смислу текстів (автоматичне реферування й анування наукового тексту лінгвістичної тематики).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Вийтухновская А.А. О терминологической связи тезаурусов по языкознанию и общей психологии // Проблемы библиографии, библиотековедения и детского чтения, вып. 8. – Л., 1976.
2. Герд А.С. Прикладная лингвистика / А.С.Герд. – С.-Пб.: Издательство С.-Петербургского университета, 2005. – 266 с.

3. Лейчик В.М. Терминоведение: Предмет, методы, структура / В.М. Лейчик. – М.: КомКнига, 2006. – 256 с.
4. Никитина С.Е. Тезаурус по теоретической и прикладной лингвистике / С.Е. Никитина. – М., 1978.
5. Тезаурус информационно-поисковый по языкознанию. – М., 1977.
6. Шелов С.Д. Опыт семантического анализа лингвистической терминологии при построении информационно-поискового тезауруса: Автореф. канд. дис. – М., 1976.
7. Bußmann Hadumod. Lexikon der Sprachwissenschaft. – Stuttgart. – 2002.
8. Harris B., Sines. FIRL (Facetted Information Retrieval for Linguistics). Summary of paper submitted for the 1973 International Conference on Computational Linguistics. Univ. of Ottawa, May, 1973.