

Важливим показником при відборі дескрипторів у тематичні термінологічні словники є показник частоти вживання даного слова чи словосполучення. Цілком очевидно, що ймовірність приналежності деякого дескриптора до складу термінів певної предметної галузі прямо пропорційна його частоті й тим вища, чим більшою є кількість його слів. Отже показник

$$K = (F-1)*Q,$$

де  $K$  – імовірнісний коефіцієнт,  $F$  – частота дескриптора,  $Q$  – кількість слів у словосполученні, можна використовувати як оціночний для прийняття рішення щодо включення певного слова чи словосполучення до словника. Сортування складеного словника за частотою, або за отриманим показником, незважаючи на уявну простоту цього методу, дає непогану вершину списку на достатньо великих обсягах текстів. Звичайно, найбільш частотними будуть однослівні дескриптори, попри все – загальноживані слова і загальнонаукові слова типу *система, метод* тощо. Але ж ми вже зазначали, що дана методика не враховує ані семантичні, ані прагматичні властивості слів, а використовує лише формальні методи для відбору претендентів на включення до словника, а вирішальне слово завжди надається експертові.

#### Література

1. Антонов А.В. Информационно-поисковая система Galaktika-ZOOM с элементами анализа на гипермассивах информации // НТИ. Сер. 1. – 2001. – № 8. – С.12–21.
2. Лазарва О.Я. Про деякі продуктивні моделі науково-технічних термінів. // Українська термінологія і сучасність: Зб.наук.праць. Вип.4. – К: КНЕУ, 2001. – С.185–188.
3. Лукашевич Н.В. Автоматизированное формирование информационно-поискового тезауруса по общественно-политической жизни России // НТИ. Сер.2. – 1995. – №3. – С. 21–24.
4. Суперанская А.В., Подольская Н.В., Васильева Н.В. Общая терминология. Вопросы теории. – М.: УРСС, 2003.

**В. Шульгіна**, к. філол. н. \*  
УДК 81'322.373

### ПРОБЛЕМИ КОМП'ЮТЕРНОЇ ЛЕКСИКОГРАФІЇ У ГАЛУЗІ ІНФОРМАЦІЙНОЇ ДІЯЛЬНОСТІ

*Метою статті є аналіз можливостей створення електронного забезпечення лексикографічними виданнями такої сфери, як документознавство. Проблема забезпечення навчального процесу на факультетах документознавства необхідною літературою включає в себе і підготовку термінологічних словників цієї потрібної сфери. Термінологічний словник документознавчої галузі є першою спробою узагальнення термінологічних понять, необхідних в роботі документознавців.*

Прикладна лінгвістика як один з новітніх сучасних напрямків розвитку мовознавства виокремилася з традиційної лінгвістики не так давно. Серед проблем, якими займаються дослідники в галузі прикладної лінгвістики, проблема комп'ютерної лексикографії посідає одне з чільних місць [1]. Це пов'язано, насамперед, з активним розвитком комп'ютерної техніки, з її використанням у різних сферах науки та народного господарства. Комп'ютерна лексикографія є складовою частиною лексикографії як розділу традиційного мовознавства. Тому досвід лінгвістів-лексикографів, силами яких створювалося словникове багатство будь-якої мови, є надійним фундаментом для подальшого його розвитку у межах комп'ютерної лексикографії. Важливу роль у розвитку сучасної комп'ютерної лексикографії відіграють і новітні лексикографічні розробки, які є безпосередньою реакцією на вимоги сьогодення.

Останнім часом українська лексикографія поповнилася цілою низкою нових термінологічних словників. Поява цих видань зумовлена у певній мірі оновленням змісту багатьох наукових сфер українського суспільства, яке переживає часи набуття і зміцнення незалежності нашої держави, часи переходу від соціалістичної економіки до ринкової. Економіко-соціальні зміни, що відбуваються в Україні, не могли не позначитися і на розвиткові такої наукової галузі, якою є документознавство. Адже сама підготовка фахівців-документознавців для України стала справою зовсім новою – до розпаду СРСР таких спеціалістів готували лише в Московському історико-архівному інституті. Зараз, з урахуванням потреби в грамотних документознавцях, яку відчувають як державні установи, так і кожна приватна фірма, факультети документознавства та

\* © В. Шульгіна, 2006

інформаційної діяльності відкрито в багатьох вищих навчальних закладах України як державної, так і недержавної форми власності (Східноєвропейський університет економіки і менеджменту, Київський університет культури, Харківський університет культури, Гуманітарний інститут Національного авіаційного університету та ін.).

Тому метою даної статті є аналіз можливостей створення електронного забезпечення лексикографічними виданнями такої сфери, як документознавство. Проблема забезпечення навчального процесу на факультетах документознавства необхідною літературою включає в себе і підготовку термінологічних словників цієї потрібної сфери. Термінологічний словник документознавчої галузі є першою спробою узагальнення термінологічних понять, необхідних в роботі документознавців. Враховуючи необхідність сучасного документознавця мати міцні знання не лише у межах власне документознавства, нами було залучено до словника ще і терміни суміжних з документознавством дисциплін: діловодства, архівознавства, менеджменту, інформатики, комп'ютерного забезпечення діловодства, інформаційна діяльність та ін.

Терміни, представлені в словнику, подаються в алфавітному порядку з обов'язковим посиланням на джерело, де тлумачення цього терміна може бути подано в більш широкому викладі. Певні терміни у різних джерелах мають різне тлумачення. Такі поняття і в словнику подаються у кількох варіантах, що передбачає можливість студентів порівняти існуючі уявлення про документознавче явище, обрати найбільш оптимальне для конкретної виробничої ситуації. Список використаних при укладанні словника джерел подається в кінці словника.

Вважаємо, що поява термінологічного словника документознавчої галузі не лише допоможе в формуванні високопрофесійних фахівців з документознавства у межах вищих навчальних закладів, а й знадобиться в практичній роботі багатьом працівникам управлінської сфери. У цілому, вміння спеціаліста-документознавця користуватися словниками, в тому числі і термінологічним, вчасно знайти потрібну інформацію повинні розглядатися як одна з його найважливіших професійних якостей.

Термінологічний словник документознавчої сфери доцільно розглядати як перший крок для створення комп'ютерного тезауруса галузі документознавства. У такому тезаурусі матеріал, зібраний у словнику, має бути систематизованим за певними напрямками, його компоненти мають підпорядковуватися один одному в залежності від тих ієрархічних відношень, у яких вони перебувають саме у терміносистемі документознавчої галузі.

Створення такого тезауруса відповідає тенденціям, які простежуються на сучасному етапі розвитку нашого суспільства, яке поступово набуває статусу суспільства інформаційного. Адже для сучасного розвитку проблем інформації та інформованості характерним є те, що на одне з перших місць виходить проблема удосконалення інформаційних технологій, комп'ютеризація, створення новітніх баз даних та нових електронних видань, які мають значні переваги перед виданнями паперовими.

До таких електронних видань можна віднести і такі одиниці комп'ютерної лексикографії, як електронні тезауруси, створення яких є нагальною потребою сьогодення. Під терміном "тезаурус" можна розуміти таку одиницю системології, складові елементи якої входять у певні відношення між собою. При такому розумінні тезауруса особливої актуальності набуває проблема дослідження як самих компонентів тезауруса, так і тих відношень, зв'язків, які між ними виникають.

Статтю тезауруса можна гіпотетично ототожнити з моделлю хімічної речовини з шкільного курсу органічної хімії, де речовина, що має свої фізичні і хімічні властивості, складається з окремих молекул різнорідних хімічних речовин, що в свою чергу мають власні фізичні і хімічні ознаки за рахунок своєї внутрішньої структури – атомів та їх складових частин. Усі компоненти складної речовини при цьому перебувають у складних відношеннях не лише з сусідніми компонентами, а й з найбільш віддаленими. Окрім цієї аналогії напрошується аналогія побудови структури Всесвіту, де його компоненти – планети та інші космічні тіла, перебувають у постійних специфічних зв'язках між собою. Тож основні принципи побудови тезауруса є природними і універсальними.

Електронний тезаурус документознавчої галузі, на нашу думку, повинен відповідати таким основним вимогам:

1. тезаурус має бути багатоаспектним електронним виданням, кожний аспект якого повинен нести інформацію про концепти певної галузі;
2. послідовність словникових статей має бути на алфавітна, а тематична;

3. концепт – це узагальнене найменування для цілого класу явищ реальної дійсності;
4. вершиною словникової статті має бути артефакт;
5. кожна стаття концептного тезауруса має будуватися на принципі ієрархічності, в основу якого покладено гіпо-гіперонімічні, родо-видові відношення;
6. найдрібнішим елементом статті словника концептів може бути максимально конкретне найменування;
7. між артефактом і максимально конкретним найменуванням назви вибудовуються у межах послідовної шкали наростання деталізації за принципом: від найбільш загального до найбільш конкретного;
8. у межах комп'ютерного тезауруса концептів слід забезпечити можливість динамічних процесів між елементами однієї словникової статті як в бік узагальнення, так і в бік конкретизації.

Сфери застосування тезауруса можуть бути досить різноманітними.

Однією з основних сфер їх використання вважається інформаційно-пошукова робота, при якій відбувається “ заміна лексичних одиниць тексту стандартизованими словами і висловлюваннями (дескрипторами) при індексуванні документів і використанні родо-видових і асоціативних зв'язків між дескрипторами при автоматизованому інформаційному пошуку документів” [7: 506 – 507].

Поряд з цим електронний тезаурус може активно використовуватися в навчальному процесі, в тому числі при застосуванні дистанційної форми навчання, адже кожний концепт у такому тезаурусі має супроводжуватися цілим „деревом” конкретизації його певних сторін та аспектів, що може сприйматися як свого роду схема, за якою можуть бути структуровані знання з певної проблеми.

Електронні тезауруси, що ґрунтуються на потужних можливостях сучасної комп'ютерної техніки для зберігання інформації, здатні містити значно більше інформації, ніж окремий паперовий словник специфічного спрямування. Йдеться про можливість реалізації критеріїв комплексності і вичерпності у подачі та зберіганні інформації, яка стосується кожного окремого поняття реальної дійсності.

Системи концептів, які функціонують у природному людському мовленні і являють собою систему понять у їх понятійному, концептуальному аспекті, є тим спільним, що об'єднує мову людини і мову машини. Тому можна вважати, що порушена проблема може бути цікавою для розробників діалогових систем “людина – ПЕОМ”, а значить створення електронних тезаурусів концептів певних галузей можна розглядати лише як перший крок на шляху до створення штучного інтелекту.

#### Література

1. Богданов В.В., Бондарко Л.В., Буторов В.Д., Герд А.С. Прикладная лингвистика и теория языка // Структурная и прикладная лингвистика. – Вып 3 Л: ЛГУ, 1987. .
2. Клименко Н.Ф. Нови підходи до укладання комп'ютерних словників// Мовознавство, 1996, №4-5.
3. Морфологический анализ научного текста на ЭВМ. – Киев, 1989.
4. Грязнухина Т.О. Лінгвістичні засади автоматичного синтаксичного аналізу// Мовознавство. – 1994. – №2.
5. Апресян Ю.Д. О языке толкований и семантических примитивах// Избран. Труды. В 2-х т. – М.: Школа «Языки русской культуры», 1995. – Т.2: Интегральное описание языка и системная лексикография.
6. Шайкевич А.Я. Тезаурус // Лингвистический энциклопедический словарь. – М.: Сов. энциклопедия, 1990.
7. Чурсин Н.Н. Популярная информатика. – К.: «Техника», 1982.

**В. Перебийніс**, д.філол. н., **С. Назаров**, к.філол. н., **Т. Бобкова** к.філол. н.\*  
Київський національний лінгвістичний університет (Київ)  
УДК 81'374.33

#### КОМПЛЕКС ЕЛЕКТРОННИХ НАВЧАЛЬНИХ СЛОВНИКІВ „ГЛОСА”

*Learner's dictionaries are intended for students and teachers of a foreign language and serve the purpose of selecting and activizing vocabulary to be studied. Lexical minimums are aimed at selecting most frequent words to be taught at different stages of studying a foreign*