

**СЕМАНТИЗАЦІЯ ТЕРМІНА
В ЕЛЕКТРОННИХ ГАЛУЗЕВИХ ТЛУМАЧНИХ СЛОВНИКАХ
(на матеріалі онлайн-ових англо-мовних ресурсів)**

*Купріянов Євген Валерійович,
здобувач*

Харківський національний університет імені В.Н. Каразіна

При підготовці електронних словників загально-мовної або термінологічної лексики використовуються як досягнення традиційної лексикографії, так і нові ідеї, які не можна реалізувати у друкованих виданнях. Це стосується передусім впорядкування системи інформації про значення лексичної одиниці, тобто її семантизації. У статті зроблено спробу системно дослідити особливості семантизації термінологічної лексики в різних онлайн-ових тлумачних словниках.

Ключові слова: онлайн-овий словник, комп'ютерна лексикографія, гіпертекст, семантизація

Лексикографія, зокрема термінологічна, набула нової якості із появою комп'ютерних словників та нових технологій їх укладання. Електронні словники та енциклопедії розробляються сьогодні як автономні або мережні програмні продукти. Не викликає сумніву той факт, що електронні словники надають користувачеві додаткові можливості порівняно з друкованими аналогами: 1) зберігання великого обсягу інформації за рахунок гіперпосилань: словник передбачає включення до своєї структури кілька словників різних типів та жанрів (словника сполучуваності, термінології, граматичних норм та ін.); 2) доступність словника за рахунок ефективної системи пошуку (повнотекстовий пошук, одночасний пошук у кількох словниках, швидкість пошуку); 3) використання засобів мультимедіа для семантизації лексики: озвучення заголовкових слів, введення ілюстративного матеріалу з фотографіями, анімацією, відеофрагментами; 4) використання словників у локальній та глобальній мережах, причому робота зі словниками може проводитися багатьма користувачами одночасно; 5) економія часу та матеріальних витрат під час створення комп'ютерних словників [Бессонова 2006, 91].

Проблему семантизації слова в електронних словниках загально-мовної лексики досліджують Л.Є. Бессонова, І.І. Заварусьва, О.Ю. Чепік, В.А. Широков, К.М. Якименко та інші. Але поза увагою науковців залишилися недослідженими принципи упорядкування семантичної інформації про термін в електронних галузевих словниках, що зумовлює актуальність нашого дослідження. Мета дослідження – виявити засоби та особливості семантизації терміна в електронних словниках, що розміщуються в мережі Інтернет. За об'єкт дослідження взято онлайн-ові англо-мовні словники, які можна вважати підвидом електронних словників, розміщених на різноманітних локальних носіях (лазерних, жорстких дисках тощо). Предмет аналізу – принципи упорядкування семантичної інформації про термін в тлумачних електронних словниках.

Наукова новизна дослідження полягає у тому, що в роботі вперше зроблено спробу системно дослідити особливості семантизації термінологічної лексики в різних електронних тлумачних словниках.

Під семантизацією розуміють виявлення значення мовної одиниці, процес або результат повідомлення необхідної інформації про змістовну сторону мовної одиниці.

В електронних словниках загальноживаної або термінологічної лексики використовуються як досягнення традиційної, тобто "паперової" лексикографії, так і нові ідеї, які не можна реалізувати у друкованих виданнях. Це стосується передусім впорядкування системи інформації про значення лексичної одиниці, тобто її семантизації. Електронний словник використовує ті ж самі інтерпретаційні параметри, що і паперовий, але форми їх подання в електронному словнику є більш різноманітними і мають деякі особливості:

1) різні семантичні параметри у паперових словниках подаються у лінійній послідовності, проте в електронних словниках вони можуть розташовуватися у будь-якому порядку завдяки використанню технології гіпертексту, який, на думку сучасних дослідників, є новим інструментом та новою технологією розуміння тексту. Практична цінність гіпертексту в комп'ютерній лексикографії полягає у тому, що він описує тип інтерактивного середовища з можливостями а) переходів за допомогою посилань, якими можуть бути слова, фрази або малюнки, б) вибору користувачем того чи іншого посилання і в) виводу на екран пов'язаних з ним текстових та мультимедійних матеріалів. Так, у кожній словниковій статті Вікіпедії існують "внутрішні" і "зовнішні" гіперпосилання. Внутрішні посилання відображають структуру словникової статті та вказують на інформацію енциклопедичного характеру, а зовнішні – вказують на зв'язки терміна з іншими словами, які подаються в окремих словникових статтях. Як приклад, розглянемо словникову статтю **Xylene** у "Wikipedia". До внутрішніх посилань відносяться такі: *Chemical properties* (хімічні властивості), *Application* (застосування), *Health effects* (вплив на здоров'я) та інші. За допомогою зовнішніх гіперпосилань виділяються різні види відношень, наявних у семантичній мережі:

*The term **xylene** or **xylol** refers to a mixture of three aromatic hydrocarbon isomers which is used as a solvent* [Wikipedia].

У вищенаведеному прикладі за допомогою гіперпосилань виділяються гіпероніми *aromatic hydrocarbon isomers* (ароматичні вуглеводні ізомери) та *solvent* (розчин), на які, в свою чергу зроблені гіперпосилання. Тобто відбувається віддзеркалення над множини до початкового.

В "Encyclopedia Britannica" значення терміна **Xylene** розкривається через зв'язані гіперпосиланнями тематичні групи *chemical formula* (хімічна формула), *raw material* (сировина для виготовлення фарб), *methyl group* (метилова група):

Xylene – any of three isomeric dimethylbenzenes [which have the same chemical formula, $C_6H_4(CH_3)_2$, but different molecular structure], used as solvents, as components of aviation fuel, and as raw materials for the manufacture of dyes, fibres, and films [Encyclopedia Britannica].

Крім того, у цій же енциклопедії за допомогою гіперпосилань виділяються гіпоніми до терміна **Xylene**: *ortho-xylene*, *meta-xylene*, *para-xylene*.

У "Chemistry Encyclopedia" семантизація терміна **Xylene** є визначенням його місця в таксономічній класифікації, яка включає такі елементи: *benzene*, *isomers*, *carbon*, *benzene ring*, *methyl*:

*The term **xlenes** refers to a group of 3 benzene derivatives which encompasses ortho-, meta-, and para- isomers of dimethyl benzene. The o-, m- and p- isomers specify to which carbon atoms (of the main benzene ring) are the 2 methyl groups attached* [Chemistry Daily 2008].

Гіперпосилання також використовуються для позначення партонімічних відношень, як у словниковій статті **Display screen**:

The display part of a monitor. Most display screens work under the same principle as a television, using a cathode ray tube (CRT) [Webopedia 2009].

У цьому прикладі терміни **monitor** й **cathode ray tube**, виділені за допомогою гіперпосилань, є меронімами до поняття **display screen**.

Окрім наведених вище відношень, можуть також виділятися асоціативні відношення:

Window – *An enclosed, rectangular area on a display screen. Most modern operating systems and applications have graphical user interfaces (GUIs) that let you divide your display into several windows* [Webopedia 2009].

Нелінійний характер гіпертексту дає можливість формування кожного елементу словника у вигляді розгалуженої структури, що дозволяє значною мірою розширити рамки словникової статті у порівнянні з паперовими словниками;

2) енциклопедична інформація як засіб семантизації в електронних словниках зустрічається частіше і подається у значно більшому обсязі, ніж у паперових словниках, де він із зрозумілих причин є обмеженим. Наприклад, у словнику термінів комп'ютерної техніки та Інтернету "Webopedia" об'ємно наводяться гіпотези щодо походження того чи іншого терміна, як, наприклад у словниковій статті **spam**:

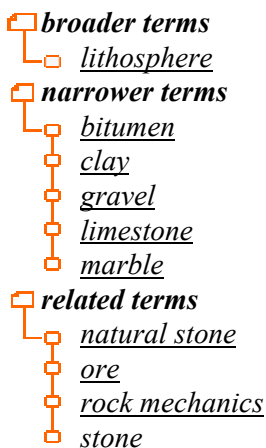
There is some debate about the source of the term, but the generally accepted version is that it comes from the Monty Python song, "Spam spam spam spam, spam spam spam spam, lovely spam, wonderful spam..." Like the song, spam is an endless repetition of worthless text. Another school of thought maintains that it comes from the computer group lab at the University of Southern California who gave it the name because it has many of the same characteristics as the lunchmeat Spam:

– *Nobody wants it or ever asks for it.*

– *No one ever eats it; it is the first item to be pushed to the side when eating the entree.*

Sometimes it is actually tasty, like 1% of junk mail that is really useful to some people [Webopedia 2009];

3) як ми зазначили вище, вказівка на видові (співвідносні й тематично пов'язані) терміни наявна, як у паперових, так і електронних словниках. Але в них застосовуються різні екстралінгвістичні засоби. У перших – видові терміни вказуються в межах статті курсивом, погрубленим шрифтом або виносяться в окрему словникову статтю. В електронному – спеціальні позначення реалізуються за допомогою кольору або анімації, що фіксує увагу користувача, або наводяться у вигляді ієрархічного дерева, як наприклад у Словнику-тезаурусі з екології "GEMET Thesaurus". Такий спосіб подання видових, співвідносних та тематичних термінів ілюструє термінологічну систему як цільну структуру. Розглянемо, наприклад, у "GEMET Thesaurus" словникову статтю **Rock**:

Rock

[GEMET 2008]

У наведеній вище словниковій статті у вигляді дерева показано родовий термін (**broader term**), видові (**narrower terms**) та співвідносні терміни (**related terms**). Встановлення відношень уточнюють і збагачують семантизацію терміна;

4) графічна ілюстрація використовується у паперових словниках переважно чорно-білих для показу конкретних та абстрактних понять, репрезентованих тим чи іншим терміном. Проте, в електронних словниках ілюстрація може бути більш складнішою, різнокольоровою, анімуватися, завдяки чому можуть бути показані динамічні процеси (наприклад, "Анімований словник з телекомунікацій" (<http://www.techtionary.com>) наочно показує фізичні процеси, що протікають під час роботи того чи іншого пристрою або зміна стану залежно від застосовуваних до нього настройок).

Ілюстрація як засіб семантизації слова в словнику майже не обговорюється лексикографією як теоретична проблема. В той же час зарубіжна "паперова" лексикографія використовує образотворчі засоби давно, з початку XVIII століття.

Питання щодо функції ілюстрації та обґрунтованості її використання вирішується неоднозначно в американській і англійській лексикографії. Так, наприклад, діаметрально протилежними є думки у С. Ландау і Л. Згуста на те, чи замінює ілюстрація дефініцію простих, конкретних понять. На думку С. Ландау, "*прості слова не потребують картинок, а потребують визначення*" (*simple words need no pictures, they certainly do need definitions*). Ладислав Згуста наголошує на тому, що основне призначення ілюстрації – це зображення незвичайних і незнайомих речей. Пізніше в статті "Ornamental Pictures in Dictionaries Л. Згуста" пише про те, що образотворчі засоби корисні для наочності і звичайно використовуються на додаток до визначення [Табанакова 2001, 131–132]. Прикладом використання ілюстрацій є *Illustrated Architecture dictionary*, у якому наводяться зображення архітектурних споруд, елементів будівельних конструкцій.

Поява Всесвітньої павутини сприяла виникненню виняткових можливостей отримання вичерпної інформації за будь-яким питанням, отже семантизація будь-якого терміна відбувається точніше та всеохоплююче. Завдяки існуванню онлайн-словників, енциклопедій (Wikipedia, Webopedia, GEMET Thesaurus, MedicineNet, Illustrated Architecture dictionary та інші) скорочуються часові витрати на пошук, за-

безпечується оперативний доступ до потрібної інформації у будь-який момент часу, об'єм отримуваних даних не обмежується одним книжковим томом і навіть бібліотекою – він охоплює інформаційні ресурси світового виміру. Цілком зрозуміло, що нові технології породжують і нові проблеми, пов'язані з умінням користуватися комп'ютерною технікою, орієнтуватися у світовій мережі, правильно оформлювати запит для отримання релевантної відповіді і т. ін. Але це не є проблемою саме семантизації, тому ми не і не зупиняємося на розгляданні цих моментів.

Перспективним напрямком дослідження, на нашу думку, є проектування та створення українських галузевих онлайн-словників та енциклопедій, використовуючи та удосконалюючи досвід зарубіжних лексикографів у семантизації термінологічних одиниць.

При подготовке электронных словарей общеупотребительной или терминологической лексики используются как достижения традиционной лексикографии, так и новые идеи, которые нельзя реализовать в печатных изданиях. Это касается, в первую очередь, упорядочения системы информации о значении лексической единицы, то есть ее семантизации. В статье сделана попытка систематически исследовать особенности семантизации терминологической лексики в разных онлайн-словарях.

Ключевые слова: онлайн-словарь, компьютерная лексикография, гипертекст, семантизация.

While compiling general or special-purpose dictionaries both the achievements in traditional lexicography and new ideas, which can't be implemented in printed dictionaries, are applied. The above mentioned statement concerns the systematization of the information about lexical unit meanings. The article is devoted to the systematic analysis of term semantization in various online dictionaries.

Key words: electronic dictionary, computer lexicography, hypertext, semantization.

Література:

1. Бессонова Л.Е. Систематизация виртуального лексикографического пространства: лексикографический аспект / Л.Е. Бессонова // Культура народов Причерноморья. – 2006. – № 94, – С. 90–94.
2. Заваруєва І.І. Тлумачні електронні словники російської мови: автореф. дис. на здоб. наук. ступ. канд. філол. наук: 10.02.02 "Російська мова" / І.І. Заваруєва. – Дніпропетровськ, 2008. – 19 с.
3. Табанакова В.Д. Идеографическое описание научной терминологии в специальных словарях: дис. на соиск. учен. степ. доктора филол. наук: 10.02.21 / Вера Дмитриевна Табанакова. – Тюмень, 2001. – 288 с.
4. Чепик Е.Ю. Политическое слово в структуре электронного словаря / Е.Ю. Чепик. – Режим доступу: <http://workinggroup.org.ua/publchepik1.shtml>.
5. Широков В.А. Феноменология лексикографических систем / Володимир Анатолійович Широков. – К.: Наукова думка, 2004. – 328 с.
6. Britannica Online Encyclopedia. – Режим доступу: <http://www.britannica.com>.
7. General Multilingual Environmental Thesaurus (GEMET). – Режим доступу: <http://www.eionet.europa.eu/gemet>.
8. On-line Computer Dictionary for Computer and Internet Terms (Webopedia). – Режим доступу: <http://www.webopedia.com>.
9. The Chemistry Encyclopedia. – Режим доступу: <http://www.chemistrydaily.com>.