УДК 81'374.8-022.218(0.034.2)
DOI https://doi.org/10.32841/2409-1154.2019.40.2.32

Черниш О. А.,
кандидат філологічних наук,
доцент кафедри теоретичної та прикладної лінгвістики
Державного університету «Житомирська політехніка»

ОСОБЛИВОСТІ СТВОРЕННЯ ЕЛЕКТРОННОГО БАГАТОМОВНОГО АКАДЕМІЧНОГО ГАЛУЗЕВОГО СЛОВНИКА

Анотація. Статтю присвячено аналізу особливостей створення електронного багатомовного академічного галузевого словника. Необхідність електронного багатомовного академічного галузевого словника зумовлена не лише потребою об'єктивізації досліджень, оптимізації й раціоналізації професійної роботи фахівця, а й зростаючим попитом пересічного користувача на адаптовану об'єктивну достовірну інформацію мовознавчого характеру. Укладання такого словника передбачає здійснення теоретичних узагальнень та низки лексикографічних робіт під час його моделювання на комп'ютері, окреслення алгоритму його укладання, проведення багатокритеріального аналізу лексичної бази для формування реєстру словника, а також уніфікації терміносистеми з дотриманням вимог цілісності, повноти та несуперечливості.

У статті розглянуто різновиди електронних словників та наголошено на їх особливостях та перевагах. Встановлено, що укладання електронного словника здійснюється у декілька етапів, які передбачають відбір лексичного мінімуму, відбір слів мови оригіналу (реєстру) та слів-відповідників, а також використання відповідного програмного забезпечення для ведення бази даних словника. У науковій розвідці проаналізовано методи створення електронного багатомовного академічного галузевого словника, схарактеризовано особливості системи для укладання багатомовних мультимедійних комп'ютерних словників. Доведено, що електронний багатомовний академічний галузевий словник сприятиме самостійній науково-дослідній роботі фахівців різних спеціальностей. Він уможливить змістовний, точний, логічний та правильний виклад думок, забезпечить доказовість, коректність і доречність уживання лінгвістичного матеріалу. Створення такого словника допоможе систематизувати та стандартизувати термінологію, сприятиме оптимальному засвоєнню знань, формуванню умінь і навичок за мінімальної затрати часових ресурсів. Використання такого словника допоможе у професійній діяльності, розширить кругозір фахівця, сприятиме встановленню міжнародних зв'язків, відкриє нові перспективи кар'єрного зростання. Таким чином, багатомовний академічний галузевий словник є запорукою успіху наукової та професійної діяльності фахівця.

Ключові слова: електронний словник, тезаурус, лексикографія, реєстр, лексикографічна діяльність, комп'ютерна база даних.

Постановка проблеми. Стрімкий розвиток науки та інформаційних технологій вимагає об'єднання та уніфікації наявної наукової інформації на єдиній концептуальній та методико-процедурній основі, що дозволить швидко та якісно здійснювати інформаційно-довідкову, дослідну, навчальну та редакційно-видавничу діяльність. Автоматизація лексикографічної діяльності є вимогою сьогодення. Вона дозволяє швидко реагувати на зміни вживання лексики у текстах, генерувати тексти з їх подальшим багатократним та різноплановим вживанням, здійснювати теоретичні узагальнення та проводити різноманітну лексикографічну роботу в автоматичному режимі, зокрема й створювати обґрунтовану базу лексики певної мови, що у свою чергу допоможе фасилітувати якісну наукову співпрацю.

Аналіз останніх досліджень і публікацій. Над темою розробки та втілення електронних словників працювала значна когорта визначних мовознавців як українських, до яких належать О. Балалаєва, Т. Борщевська, Н. Дарчук, І. Кульчицький, В. Морковкін, В. Перебийніс, М. Пещак, В. Широков, так і зарубіжних, до яких належать Л. Зіміна, В. Вуль, Л. Колодяжна та інші. Наразі наголошено на особливостях навчального словника [5, с. 44], на перевагах та недоліках його використання [3]. систематизовано підходи до визначення поняття «електронний словник» [9, с. 242], доведено важливість автоматизації лексикографічної діяльності [6, с. 28], окреслено критерії його укладання [8, с. 128] тощо.

Попри прогресивний розвиток наукових досліджень у галузі таких лінгвістичних дисциплін, як комп'ютерна лексикографія та корпусна лінгвістика, досі все ще не створено єдиного об'єктивного та достовірного доступу до масиву фактичного мовного матеріалу, що нагально необхідний для подальшої наукової робити фахівців різних галузей. Зазначені вище факти переконливо свідчать про необхідність проведення додаткових досліджень, спрямованих на подолання лакун у розробці комп'ютерних алгоритмів, програм, систем та технологій для укладання, наповнення та використання електронних багатомовних академічних галузевих словників

Метою нашої наукової розвідки є аналіз особливостей створення електронних багатомовних академічних галузевих словників.

Виклад основного матеріалу. Електронний словник розуміється як особливий лексико-граматичний об'єкт [9], чітка послідовна структура з передмовою, правилами використання, переліком скорочень, корпусом, граматичним описом [2]; комп'ютерна база даних, що містить закодовані словникові статті, які здійснюють швидкий пошук необхідних слів з урахуванням морфологічних форм пошуку словосполучень [9]. Відповідно до трактувань науковців розрізняють автоматичний, машинний, комп'ютерний та електронний словники, однак усі ці терміни в науковій літературі вживаються синонімічно.

Електронні словники розрізняються відповідно до цілі та кінцевого споживача [11] та поділяються на такі словники:

– для широкого загалу (Google Translate, Systran 7 Home Translator);

– для корпоративного користування (Prompt NET Professional);

– для професійних перекладачів (PT 2008 Professional, Translation Memory: Transit, Star, Trados, Translation Manager, Eurolang Optimizer);

– для перекладачів вебсайтів (Inter Tran Web Site Translation Server, Systran Web Translator);

– для користувачів портативних пристроїв (Prompt Mobile, Speereo Voice Translator);

– для синхронних перекладачів (Google Translate for Android).

– Електронний багатомовний академічний галузевий словник (тезаурус) має низку переваг, зокрема:

– актуальність та динамічність, що забезпечує можливість постійного оновлення інформації, а також вилучення застарілих даних;

– значний обсяг словникової бази та широкий доступ до необхідної інформації завдяки використанню гіперпосилань;

– універсальність, оскільки програми уможливлюють роботу з декількома мовами водночас та пропонують різні варіанти перекладу;

– багатофункціональність, що забезпечується низкою додаткових функцій, які спрощують роботу зі словником;

– мультимедійність, що полягає у використанні графічного й ілюстративного матеріалу, аудіо- та відеофрагментів;

– варіативність у використанні як у локальній, так і глобальній мережах, а також у використанні оффлайнової та онлайнової версій;

– зручний пошук завдяки ефективній системі повнотекстового пошуку, високій швидкості пошуку та одночасного пошуку у декількох словниках. Електронний словник передбачає застосування різноманітних лінгвістичних технологій, зокрема морфологічного та синтаксичного аналізу, повнотекстового пошуку, розпізнавання і синтезу мов тощо.

З огляду на зазначені переваги електронний багатомовний академічний галузевий словник є незамінним у науковій роботі фахівців різних спеціальностей.

Яскравими прикладами електронних словників є такі:

– ABBYY Lingvo, який містить 13 тематичних словників та здійснює переклад російською з англійської, німецької, французької, італійської й іспанської мов;

– ГЛОСА, що становить сукупність електронних навчальних словників з використанням подвійного перекладу з української на англійську мову та з англійської мови на українську. Реєстр цього словника складає близько 3 000 слів, понад 11 600 словосполучень, 1 500 прикладів слововживань. Словник розрахований на учнів старших класів середніх загальноосвітніх шкіл, а також усіх охочих вивчати англійську мову;

– MultiLex, який передбачає понад 40 загальних, тлумачних та тематичних словників зі спеціалізованої лексики у галузі будівництва, економіки, банківської справи, права, інформаційних технологій, медицини тощо;

– чотиримовний науковий словник, який містить широковживані наукові терміни й термінологічні сполучення з хімії, теоретичної хімії та фізики. Програмне забезпечення словника дозволяє здійснювати прямий та зворотний переклади англійською, німецькою, українською та російською мовами;

– навчальний трьохмовний тлумачний словник з понад 700 англійських, 800 українських та близько 900 російських термінів у галузі комп'ютерної лінгвістики;

– словник-довідник з труднощів англійського слововживання, що надає понад 2 000 найуживаніших англійських слів з їхніми 3 000 українськими відповідниками та інші.

Особливістю нашого електронного багатомовного академічного галузевого словника (тезауруса) вважаємо винятковий добір робочих мов (українська, російська, англійська, німецька, французька, китайська мови), а також його наукову термінологічну спрямованість. Такий словник буде незамінним для фахівця у будь-якій галузі.

Процес створення електронного багатомовного академічного галузевого словника є доволі складним, потребує залучення висококваліфікованих спеціалістів у галузі мовознавства й програмування та передбачає вирішення низки таких *завдань*:

– проаналізувати принципи укладання електронного багатомовного академічного галузевого словника;

– дослідити різновиди електронних багатомовних академічних видань, окреслити їхні переваги та недоліки;

– здійснити теоретичні узагальнення та низку лексикографічних робіт в автоматичному режимі під час втілення та моделювання електронного багатомовного академічного галузевого словника на комп'ютері;

– здійснити багатокритеріальний аналіз лексичної бази для формування реєстру електронного багатомовного академічного галузевого словника;

– уніфікувати терміносистеми робочих мов, необхідних для створення електронного багатомовного академічного галузевого словника, з дотриманням вимог цілісності, повноти та несуперечливості;

– створити автоматизовані банки термінів робочих мов, необхідних для створення електронного багатомовного академічного галузевого словника;

– уніфікувати та впорядкувати величезні об'єми фактологічних даних робочих мов для підготовки електронних багатомовних академічних галузевих словників;

– оцифрувати величезну кількість фактичного лінгвістичного матеріалу, що передбачає процедури сканування, розпізнавання тексту, редагування тексту та виправлення помилок;

– створити відповідну модель даних, структура якої дозволяє відобразити структурно-семантичні зв'язки між реєстровими одиницями;

– здійснити процедуру парсингу, що полягає у створенні лексикографічної бази даних;

– трансформувати лексикографічну систему у віртуальний об'єкт, що функціонує з використанням мережі та будується за багаторівневою архітектурою з основним принципом для побудови технологією вебсервісів.

Виконання цих завдань уможливить створення електронного багатомовного академічного галузевого словника, що сприятиме здійсненню інформаційно-довідкової, дослідницької та редакційно-видавничої діяльності, а отже, стане у нагоді фахівцям різних спеціальностей.

Укладання електронного багатомовного академічного галузевого словника передбачає створення єдиного лексикографічного комплексу. *Першочерговий етап* укладання електронного

багатомовного академічного словника полягає у створенні реєстру. Реєстровим вважається слово, яке пояснюється чи перекладається. Отже, реєстр розуміють як сукупність заголовних слів [10]. Для зручності пошуку відповідної фактичної мовної інформації реєстрове слово, як правило, подається іншим шрифтом, містить позначки з роз'ясненням та перекладом, морфологічний та синтаксичний аналіз, а також додаткову інформацію граматичного та стилістичного характеру. Для пошуку реєстру застосовуються методи аналогії, класифікації та системного пошуку.

Другий етап створення електронного багатомовного академічного галузевого словника (тезауруса) полягає у пошуку відповідників [7; 4]. *Третій етап* передбачає структурування словника, яке може здійснюватися у декілька способів, а саме: абетково-гніздовий, ідеографічний, систематичний тощо. Найпоширенішим є абетково-гніздовий спосіб, коли у словниковій статті подається як ключове слово, так і похідні слова, а терміни розміщуються під реєстровим словом за алфавітом. Саме цей спосіб плануємо застосовувати під час укладання нашого електронного багатомовного академічного галузевого словника.

Під час створення електронного багатомовного академічного галузевого словника застосовуватимемо такі *методи:*

– *групування,* що передбачає об'єднання термінів у тематичні групи, широке використання термінів-словосполучень та термінів-зворотів, а також системи посилань;

– *систематизації*, що полягає у комплексному аналізі єдиного цілого, що функціонує завдяки узгодженості своїх складових частин;

– *концептуального аналізу*, що має на меті співвіднесення одиниць з певною категорією, концептом.

Вважаємо за доцільне скористатися *дедуктивним підходом*, який передбачає, що мовні одиниці розподіляються відповідно до готової схеми семантичної організації робочих мов [1, с. 66].

Для створення електронного багатомовного академічного галузевого словника скористаємося системою для укладання багатомовних мультимедійних комп'ютерних словників PolyDic ML 3.0, яка містить два окремих модулі для введення та редагування словникових баз та їхнього відтворення на дисках. Програма PolyDic ML 3.0 має низку додаткових переваг, зокрема:

– компактний обсяг програми та словникової бази;

– відсутність необхідності у додатковому програмному забезпеченні;

– можливість створення інформаційних вікон з додатковим лінгвістичним коментарем;

– доступність конвертації словникових баз у текстовий формат для корегування та редагування;

– наявність редактора словникових статей;

– підтримка кодовою таблицею Unicode символів усіх національних алфавітів;

– незалежне створення словникових баз на декількох комп'ютерах з їх подальшим злиттям в єдину базу;

– можливість відтворення синонімів, антонімів, транскрипції, транслітерації, граматичних, стилістичних коментарів, абревіатур, скорочень;

– наявність зображень, формул та графіків, завантаження аудіо- та відеофайлів;

– зручний пошук і простота у використанні тощо.

Використання системи для укладання багатомовних мультимедійних комп'ютерних словників PolyDic ML 3.0, а також застосування вищезазначених методів є запорукою успішного створення електронного багатомовного академічного галузевого словника.

Висновки. Таким чином, створення електронного багатомовного академічного галузевого словника (тезауруса) передбачає вирішення низки прикладних завдань, що полягають у формуванні лексикографічних баз, машинних фондів мов, лексиконів як додаткової інформації для лінгвістичних процесорів комп'ютерної обробки мови тощо. Знаходження алгоритмів вирішення цих завдань має значну наукову цінність. Створення такого словника (тезауруса) допоможе систематизувати та стандартизувати термінологію, сприятиме оптимальному засвоєнню знань, формуванню умінь і навичок за мінімальної затрати часових ресурсів. Використання такого словника допоможе у професійному зростанні, розширить кругозір фахівця, сприятиме налагодженню міжнародних зв'язків, відкриє нові перспективи кар'єрного зростання.

Література:

1. Андрощук О.С. Принципи укладання електронного словника українсько-англійських фразеологізмів на означення людини. *Мовні і концептуальні картини світу* : збірник наукових праць. Київ : Видавничо-поліграфічний центр «Київський університет». 2013. Вип. 46 (1). С.64–68.
2. Балалаєва О.Ю. Інформаційні технології як чинник формування нової парадигми освіти. *Оновлення змісту, форм та методів навчання і виховання в закладах освіти*. 2013. Вип. 7. С. 437–446.
3. Білятинська І. Комп'ютерна лексикографія як перспективний напрямок розвитку процесу укладання словників. *Прикладна лінгвістика* : каталог статей конференції. 2013. URL: https://prlingv.at.ua/publ/komp_juterna_leksikografija_jak_perspektivnij_naprjamok_rozvitku_procesu_ukladannja_slovnikiv/1-1-0-7.
4. Гордієнко Н. Українська тлумачна лексикографія: сучасний стан і перспективи розвитку. *Науковий вісник Херсонського державного університету. Серія «Лінгвістика»* : збірник наук. праць. Херсон: Вид-во ХДУ. 2010. Вип. 11. С.15–21.
5. Джонг Й.Д. Терміни НАВЧАЛЬНИЙ СЛОВНИК / УЧЕБНЫЙ СЛОВАРЬ та LEARNER'S DICTIONARY в сучасній науковій лексикографії. *Актуальні проблеми української лінгвістики: теорія і практика*. 2009. Вип. 18. С. 41–49. URL: http://nbuv.gov.ua/j-pdf/apyl_2009_18_9.pdf.
6. Кульчицький І.М. Комп'ютерно-технологічні аспекти створення сучасних лексикографічних систем ; Укр. мов.-інформ. фонд НАН України. Київ : Нац. б-ка України ім. В.І. Вернадського НАН України, 2002. 59 с.
7. Перебийніс В.І., Рукіна Е.П., Хідекель С.С. Англо-український навчальний словник з методичними коментарями та граматичними таблицями. Київ : Вежа, 2002. 424 с.
8. Сінкевич Н.М. Сучасний електронний український словник як база лінгвістичного аналізу слова. *Дослідження з лексикології і граматики української мови*. 2011. Вип. 10. С. 240–244.
9. Троссель М.В. Проблемы составления переводных терминологических словарей. *Journal of Siberian Federal University. Humanities & Social Sciences.* N. 8. 2015. URL: http://journal.sfu-kras.ru/article/19742.
10. Морковкин В.В. Типология филологических словарей. *Вестник Харьковского политехнического университета.* Вып. 1. 1994. С. 13–23.
11. Шиба А.В. Використання нових інформаційних технологій у процесі формування професійної компетентності майбутнього перекладача засобами інтерактивних технологій. *Педагогіка формування творчої особистості у вищій і загальноосвітній школах.* 2013. Вип. 28. С. 356–365.

Chernysh O. The peculiarities of electronic multilingual academic dictionary

Summary. The scientific article is devoted to the analysis of electronic multilingual academic dictionary peculiarities. The importance of an electronic multilingual academic dictionary is caused not only by the need of scientific research objectification, and professional work optimization and rationalization, but also by growing demand for objective reliable linguistic information.

The compilation of such electronic dictionaries implies theoretical generalizations and a number of lexicographic works, outlining its compilation algorithm, lexical base multicriteria analysis, as well as terminology system unification in compliance with the requirements of integrity.

The article considers the variety of electronic dictionaries and highlights their features and benefits. It is established that the compilation of an electronic multilingual academic dictionary is carried out in several stages. These stages involve lexical minimum selection, register and its correspondents' selection, as well as the use of appropriate software to deal with the dictionary database. The methods of an electronic multilingual academic dictionary compilation are analyzed. The features of the multilingual computer dictionaries compilation system are described.

It is proved that an electronic multilingual academic dictionary facilitates independent scientific research. It guarantees meaningful, accurate, logical and correct material presentation and ensures the provability, correctness and appropriateness of the use of linguistic material.

Such electronic dictionaries help to systematize and standardize terminology, optimize knowledge acquisition and skills formation in the shortest time possible. Electronic multilingual academic dictionaries are beneficial for students, research scholars, teachers, travellers and other people. They expand their outlook, promote international relations and open up new career prospects. Thus, an electronic multilingual academic dictionary is the key to scientific and professional success of a specialist as well as educational institutions.

Key words: electronic dictionary, thesaurus, lexicography, register, lexicographic activity, computer database.