

ПЕДАГОГІКА

УДК 81'33

DOI: 10.31548/hspedagog2020.01.006

З ІСТОРІЇ РОЗВИТКУ ЕЛЕКТРОННИХ СЛОВНИКІВ: ЗАРУБІЖНИЙ І ВІТЧИЗНЯНИЙ ДОСВІД

БАЛАЛАЄВА О. Ю., кандидат педагогічних наук,

доцент кафедри журналістики та мовної комунікації

*Національний університет біоресурсів і природокористування України*E-mail: olena.balalaeva@gmail.com

ORCID: 0000-0002-2675-5554

Анотація. Метою статті є простежити динаміку розвитку електронних словників у зарубіжній і вітчизняній практиці. З'ясовано, що електронні словники є відносно новим явищем на ринку лексикографічної продукції, які за кілька десятиріч років еволюціонували від машиночитних словників, точних копій паперових видань до складних цифрових лексикографічних систем з потужним арсеналом функцій. Наведено етапи розвитку автономних та онлайн словників, які завдяки розширеним можливостям пошуку, швидкості, простоті та зручності у використанні, доступності та компактності набули популярності серед широкого загалу користувачів. Дослідження у галузі вітчизняної комп'ютерної лексикографії становить перспективний напрям подальших наукових розвідок.

Ключові слова: електронний словник, комп'ютерна лексикографія, цифрова лексикографічна система, бази даних.

Актуальність. На думку вітчизняних науковців, інтенсивний розвиток комп'ютерної лексикографії і активне створення електронних словників в Україні зумовлені потребами як власне науковими (об'єктивація досліджень, впорядкування локалізованої потужної бази фактичного матеріалу, зручний та швидкий пошук інформації), так і екстранауковими (задоволення попиту широкого кола користувачів на адаптувану, об'єктивну й достовірну інформацію) [4].

Електронні словники є відносно новим явищем, яке стрімко розвивається і набуває популярності серед широкого загалу користувачів. Проте ана-

ліз сучасного рівня розвитку вітчизняних електронних ресурсів свідчить про наявний дефіцит електронних словників як загальної, так і спеціальної лексики. Брак електронних словників зумовлений низкою об'єктивних проблем як практичного, так і теоретичного характеру.

Аналіз останніх досліджень та публікацій. Теоретичні та практичні аспекти створення електронних словників вже були предметом дослідження багатьох сучасних науковців. Зокрема історичний аспект висвітлювали у своїх працях Альцина А., Балог В., Баранов А., Беляєва Л., Дубічинський В., Ка-

рпіловська Є., Купріянов Є., Несі Х., Марчук Ю., Палкова А., Пастор В., Перебийніс В. та ін.

Мета статті – простежити динаміку розвитку електронних словників у зарубіжній і вітчизняній практиці. У дослідженні застосовано **методи** аналізу наукових джерел, порівняння, узагальнення та систематизації.

Результати. Комп'ютери почали використовувати в лексикографії в 1960-х роках. Один із редакторів *Random House Dictionary of the English Language* Л. Урданг розробив систему баз даних для категоризації та сортування одиниць інформації словника, назвавши її «набір даних» [8, с. 458]. Одночасно в межах лексикографічного проєкту корпорації *System Development Corporation* було створено версії словників *Webster's 7th New Collegiate Dictionary* та *New Merriam-Webster Pocket Dictionary* на магнітній стрічці. Спочатку машиночитні словники були точними копіями паперових, дослідники витрачали чимало часу на інтерпретацію кодів набору тексту. З розвитком технологій видавці розділили компоненти друку та бази даних словників.

Прогрес технологій у 1970-х роках сприяв ширшому застосуванню комп'ютерів у лексикографічних проєктах. Комп'ютерні системи компіляції використовували для сортування та перевірки записів у словниках *Longman Dictionary of Contemporary English* (1978), який вважають першим дійсно комп'ютеризованим словником, та *Collins English Dictionary* (1979).

Щойно комп'ютери почали застосовувати у створенні словників, електронні компанії почали працювати над розробкою кишенькових гаджетів для перекладу. Перші електронні «словники» з інтерфейсами, призначеними для користувачів (*LK-3000* та *Craig M100*), технічно були нащадками портативних електронних калькуляторів і персональних цифрових помічників. *LK-3000*, або *Lexicon*, розроблений у

1976 р. і запатентований як «електронний словник та перекладач мови» демонстрував на екрані відповідності між словами англійської та низки інших мов (французької, німецької, грецької, італійської, португальської, іспанської). Незважаючи на певну лексичну обмеженість та технічну недосконалість (мовні засоби перекладу продавали окремо на знімних картриджах) пристрій було обрано офіційним засобом перекладу для Олімпіади-80 [8, с. 463]. Випущений незабаром *Craig M100* мав схожий дизайн, але більший словниковий запас і надавав можливість перекладу трьома мовами одночасно.

Виробники почали досліджувати додаткові засоби зберігання словникової інформації на знімних ІС-картах з інтегральними схемами. Із розвитком технологій видавництва випускали більше електронних продуктів, збільшувався розмір пам'яті портативних пристроїв. Як зауважує Х. Несі, у цей період почав зникати традиційний поділ між словником у мобільному пристрої (призначеним для швидкої довідки в практичному контексті) та академічним словником (який цінують за якість лексикографічної інформації). Електротехнічні компанії поступово почали купувати ліцензії на лексикографічні товари, як-от словники *Collins COBUILD*, *Longman* та *Oxford*, додаючи до них переваги останніх технічних винаходів [8, с. 467].

Розроблення технологій компресії та декомпресії дали можливість зберігати та інтегрувати все більшу кількість повнотекстових словників. Поява дисків спричинила новий виток у зберіганні інформації. Компакт-диск першого випуску 12-томного видання Оксфордського словника англійської мови з'явився в 1988 році, а в 1990-х зростає кількість словників із удосконаленими інструментами. Так, у десятому виданні словника *Concise Oxford English Dictionary* запроваджено інструмент «швидкого пошуку», інший лексикографічний інструмент *BOOKcase* надавав доступ до зовнішніх програм, пошукових

систем, а також інтегрував окремі комплементарні електронні довідкові праці, дозволяючи відкрити декілька одночасно. Одними з перших словників на компакт-диску для тих, хто вивчає англійську були інтерактивні англійський та американський словники *Longman* (1993) і *Collins COBUILD* (1995), які містили не лише словник, а й посібники з англійської фонетики та граматики. Найбільш використовуваним словником на компакт-диску у світі вважають двомовний англо-китайський словник *Kingsoft Powerword* (у період з 1997 до 2002 рр. було розповсюджено вісім мільйонів примірників) [8, с. 472].

Згодом ще одним осередком розміщення величезної кількості інформаційних ресурсів став інтернет. Спочатку видавництва, які інвестували значні кошти в розробку словників, не бажали безкоштовно розповсюджувати свою продукцію в інтернеті. До того ж дослідники вказували на ненадійність онлайн-словників загального доступу порівняно з паперовими, адже ніхто не відповідав за точність інформації, яку надавали онлайн-словники, а веб-адреси, і вміст сторінок постійно змінювались. Збільшення кількості та покращення якості онлайн-словників частково пояснювалося зміною політики видавців, які почали стягувати плату за власні онлайн-продукти або застосовували їх як інструменти маркетингу, щоб заохотити придбати книгу, компакт-диск чи електронний доступ до тексту (*Oxford English Dictionary Online*, *Cambridge Dictionaries Online* тощо). Загальною практикою для багатьох видавництв стало пропонувати власні словники у різних форматах: паперовому, на дисках, онлайн.

Використання високошвидкісних широкосмугових технологій на початку 2000-х рр. розширило доступ до інформації для багатьох пересічних користувачів. Іншою тенденцією розвитку словників в інтернеті стало створення багатифункціональних багатомовних словників довільного наповнення, як-от Ві-

кісловник. Перший *Wiktionary* (спрямований на опис усіх слів усіх мов) створений англійською 2002 р., за його схемою створювалися словники іншими мовами. Станом на 2020 р. український розділ Вікісловника містить 49 352 статті, й кожен може стати його співавтором, написати власну статтю, або доповнити чи відредагувати наявні.

Отже, як зауважують дослідники, спочатку комп'ютери лише асистували процесу укладання словників, потім сприяли перенесенню паперової інформації на комп'ютерний носій і, нарешті, стали основою складання електронних словників «з нуля» [6].

В Україні становлення комп'ютерної лексикографії бере відлік у 80-х роках ХХ ст.: хоча дослідження у цьому напрямку велися і раніше, відправним пунктом традиційно вважають [4, с. 28]. Всесоюзну конференцію з проблем створення машинного фонду (1983), в якій взяли участь відомі українські дослідники – співробітники Інституту мовознавства ім. О. О. Потебні та Інституту кібернетики ім. В. М. Глушкова В. Перебийніс, М. Пещак, І. Білецька та затвердження програми Державного комітету з науки і техніки СРСР зі створення машинних фондів мов народів СРСР. У межах цієї програми науковці Львівського політехнічного інституту розробили систему «Слово» для підтримки багатомовних термінологічних словників [5].

Після отримання Україною незалежності розвиток вітчизняної комп'ютерної лексикографії був детермінований процесами інтеграції країни до світової інформаційної спільноти – з одного боку, та реалізації державної мовної політики – з іншого. Важливим завданням стало створення національної словникової бази.

На думку Є. Купріянова, піонером у цій галузі став заснований у 1991 р. Український мовно-інформаційний фонд НАН України, де «було створено базові моделі комп'ютерної лексикографії, які склали підґрунтя для розробки

відповідних технологій» [5, с. 14], започатковано серію лексикографічних праць нового покоління – «Словники України» відповідно до Указу Президента України від 7 серпня 1999 року «Про розвиток національної словникової бази».

Станом на 2020 рік система «Словники України» в оновленій версії 4.1 складається з п'яти словникових модулів, виправлених і розширених: «Словозміна», «Транскрипція», «Синонімія», «Антонімія», «Фразеологія», а її генеральний реєстр містить близько 258 тис. слів.

Український мовно-інформаційний фонд НАН України також пропонує ресурси: «Тлумачний словник української мови в 20 томах», «Система лінгвістичної взаємодії "ВЛЛ"», «Український національний лінгвістичний корпус», «Словник української мови за редакцією Б. Д. Грінченка», «Українсько-російсько-англійський словник зі зварювання», «Український біографічний архів», віртуальні лексикографічні системи «MONDILEX» і «Turkic», віртуальні лексикографічні лабораторії «Російсько-українсько-англійський словник з механіки», «Російсько-українсько-англійський словник з радіоелектроніки», «Українсько-кримськотатарський словник» та віртуальні термінографічні лабораторії з фізики та біології.

Потужною ланкою вітчизняної комп'ютерної лексикографії є відділ структурно-математичної лінгвістики Інституту мовознавства ім. О. О. Потебні, фахівці якого розробили комп'ютерний морфемно-словотвірний фонд української мови та систему автоматизованого редагування наукового тексту. Як заявлено на офіційному сайті, нині відділ працює над дослідженням інноваційних процесів у сучасній українській мові, їх комп'ютерним моделюванням, вивченням впливу соціодинаміки на систему та структуру мови, тенденцій мовних змін, найважливішими проєктами є: розроблення лінгвістичного та алгоритмічного забезпечення для систем

автоматичного опрацювання інформації, створення бази даних про будову сучасного українського слова та розроблення засобів автоматизованого укладання на її основі нових типів українських словників. За допомогою комп'ютерних програм укладено «Обернений частотний словник сучасної української художньої прози», «Словник афіксальних морфем української мови», «Кореневий гніздовий словник української мови», «Активні ресурси сучасної української номінації: Ідеографічний словник нової лексики», «Електронний граматичний словник української літературної мови. Словозміна». Розробки співробітників покладено в основу лексикографічних продуктів: редактора для перевірки правопису українських текстів «Рута» та системи перекладу текстів з російської на українську та навпаки «Плай».

Останній є спільним проєктом з Лабораторією комп'ютерної лінгвістики Інституту філології Київського університету імені Тараса Шевченка. З-поміж інших досліджень і продуктів цієї лабораторії: частотні словники української художньої прози, публіцистики, сучасної поетичної мови, морфемно-словотвірна база української мови, граматичний словник українських дієслів, програма автоматичного синтезу парадигми англійського іменника та дієслова. Наразі в лабораторії тривають дослідження принципів англо-українського та українсько-англійського машинного перекладу, розробляють перекладні та термінологічні словники, програми аналізу та синтезу людського мовлення.

Створюють програмні продукти для редагування та перекладу й інші комерційні організації. Компанія Трайдент Софтвр розробила лінійки продуктів *Language Master i Pragma*, що базується на власній багатомовній технології перекладу. Останньою версією є Pragma 6.x для української, англійської, німецької та інших мов, яка загалом має 56 напрямів перекладу.

ABBYY Україна пропонує пакет *ABBYY Lingvo x6 Багатомовний* (професійна версія), який містить 224 загальнолексичні, тлумачні та тематичні словники для 19 мов, серед яких українська, англійська, німецька, французька, іспанська, китайська тощо.

Відомими є також розроблені харківським фахівцями перекладна англо-російська / російсько-англійська система ПАРС, оснащена комплексом термінологічних комп'ютерних словників більшості галузей людської діяльності, її модифікація ПАРС / D для російсько-німецького та німецько-російського автоматичного перекладу, система машинного українсько-російського / російсько-українського перекладу РУМП, оснащена не лише загальноживаною лексикою, а термінологічною (на базі електронних словників авіаційних, екологічних, комп'ютерних і фінансових термінів).

Варто погодитися з В. Пастор та А. Альцина, що сьогодні термінологія є ключовим фактором роботи перекладачів. Розвиток спеціалізованих галузей зростає разом із прогресом науки та техніки, і перекладачам доводиться вирішувати виклики, пов'язані з високим рівнем спеціалізації тексту. Ці вимоги ринку пояснюють, чому перекладачі шукають ресурси, щоб швидко та ефективно задовольнити свої термінологічні потреби [9].

Як зауважує С. Амеліна, переклад текстів аграрного спрямування належить до підвиду технічного перекладу. Він має основні риси науково-технічного стилю, але водночас відрізняється своїми особливостями, які властиві тільки цьому підвиду перекладу [1, с. 121].

Низку продуктів для перекладачів аграрної галузі (від електронних словників до електронних термінологічних баз) створено фахівцями Національного університету біоресурсів і природокористування України [2; 3; 7].

Висновки і перспективи. Електронні словники є відносно новим явищем на ринку лексикографічної продукції, які за кілька десятиріч еволюціонували від машиночитних словників, точних копій паперових видань до складних цифрових лексикографічних систем з потужним арсеналом функцій. Завдяки розширеним можливостям пошуку, швидкості, простоти та зручності у використанні, доступності та компактності, електронні словники набули популярності серед широкого загалу користувачів. Сьогодні їх застосовують у багатьох сферах людської діяльності – науковій, навчальній, професійній, побутовому спілкуванні, тому дослідження в галузі вітчизняної комп'ютерної лексикографії становить перспективний напрям подальших наукових розвідок.

Список використаних джерел

1. Амеліна С. М., Верченко Н. П. Особливості редагування перекладів текстів аграрної тематики. *Науковий вісник Національного університету біоресурсів і природокористування України. Серія: Філологічні науки*. 2014. № 206. С. 121—126.
2. Амеліна С. М., Тарасенко Р. О. Вивчення технологій створення електронних термінологічних баз у процесі професійної підготовки перекладачів. *Інформаційні технології і засоби навчання*. 2017. Т. 60, № 4. С. 105—115.
3. Балалаева Е. Ю. Электронный словарь: сущность, структура, классификации. *Современная педагогика*. 2014. № 4. URL: <http://pedagogika.snauka.ru/2014/04/2238> (дата звернення 10.08.2020).
4. Балог В. Сучасний стан української комп'ютерної лінгвістики. *Лексикографічний бюлетень*. Київ: Ін-т української мови НАН України, 2005. № 11. С. 28—35.
5. Купріянов Є. В. Комп'ютерна лексикографія як проблема сучасного мовознавства (історичний аспект). *Віс-*

ник Харківського національного університету ім. В. Н. Каразіна. 2008. № 53. С. 12–16.

6. Палкова А. Основные понятия электронной лексикографии. *Вестник ТвГУ. Серия «Филология»*. 2015. № 4. С. 88–93.

7. Хоменко Л. О., Балалаєва О. Ю. Досвід розробки методів дистанційного навчання іноземних мов у НАУ. *Наукові доповіді Національного аграрного університету*. 2005. Вип. 1 (1). URL: <http://nd.nubip.edu.ua/2005-1/05klomin.pdf>.

8. Nesi H. Dictionaries in electronic form. *The Oxford History of English Lexicography Oxford / Cowie, A. P. (ed.)*. Oxford UPress, 2008. P. 458–478.

9. Pastor V., Alcina A. Search techniques in electronic dictionaries: a classification for translators. *International Journal of Lexicography*. 2010. Vol. 23, No 3. P. 307–354.

References

1. Amelina, S. M., Verchenko, N. P. (2014). Osoblyvosti redahuvannia perekladiv tekstiv ahrarnoi tematyky [Specificities of editing translations of agricultural texts]. *Naukovyi visnyk NUBiP. Seriya: Filolohichni nauky*, Issue 206, 121–126.

2. Amelina, S. M., Tarasenko, R. O. (2017). Vyvchennia tekhnolohii stvorennia elektronnykh terminolohichnykh baz u protsesi profesiinoi pidhotovky perekladachiv. [Studying technologies for creating electronic terminological bases in the process of professional training of translators]. *Information Technologies and Learning Tools*. Vol. 60, no 4, 105–115.

3. Balalaieva E. Yu. (2014). Electronic dictionary: essence, structure, classifications. *Sovremennaja pedagogika*, No 4. Retrieved from: <http://pedagogika.snauka.ru/2014/04/2238>.

4. Baloh V. (2005). Suchasnyi stan ukrainskoi kompiuternoi linhvistyky [Current state of Ukrainian computer linguistics]. *Leksykohrafichnyi buleten*, No 11, 28–35.

5. Kupriianov Ye. V. (2008). Kompiuterna leksykohrafiia yak problema suchasnoho movoznavstva (istorychnyi aspekt) [Computer lexicography as a problem of modern linguistics]. *Visnyk KNU*. Issue 53, 12–16.

6. Palkova A. (2015). Osnovnye ponjatija elektronnoj leksikografii [Basic definitions of electronic lexicography]. *Vestnik TvGU. Series «Filology»*. No 4, 88–93.

7. Khomenko L. O., Balalaieva O. Yu. (2005). Dosvid rozrobky metodiv dystantsiinoho navchannia inozemnykh mov u NAU [Experience in the methods development of distance learning of foreign languages at NAU], *Naukovi dopovidi NAU*. No 1. Retrieved from: <http://nd.nubip.edu.ua/2005-1/05klomin.pdf>

8. Nesi H. (2008). Dictionaries in electronic form in: Cowie, A. P. (ed.). *The Oxford History of English Lexicography Oxford*, Oxford University Press, 458–478.

9. Pastor, V., Alcina, A. (2010). Search techniques in electronic dictionaries: a classification for translators. *International Journal of Lexicography*, 23(3), 307–354.

**FROM THE HISTORY OF THE
DEVELOPMENT OF ELECTRONIC DICTIONARIES: FOREIGN AND DOMESTIC
EXPERIENCE**

Balalaieva O. Y.

Abstract. *The purpose of the article is to study the dynamics of electronic dictionaries development abroad and in Ukraine using methods of analysis of scientific sources, comparison, generalization and systematization. Electronic dictionaries have been found to be a relatively new phenomenon in the lexicographic market, evolving from machine-readable dictionaries, exact copies of paper editions to complex digital lexicographic systems with a powerful arsenal of functions over the decades. The stages of development of autonomous and online dictionaries are described. Electronic dictionaries due to the advanced search capabilities, speed, simplicity, ease of use, accessibility and compactness have gained popularity among a wide range of users. Today they are used in many spheres of human activity – scientific, educational, professional, everyday communication. However, the analysis of the current level of development of Ukrainian electronic resources indicates a shortage of electronic dictionaries both common and terminological vocabulary. The lack of electronic dictionaries is due to a number of objective problems, both practical and theoretical, that is why research in the field of domestic computer lexicography is a promising area of further research.*

Keywords: *electronic dictionary, computer lexicography, digital lexicographic systems, databases*