

Аннотация

Предлагаемая статья является частью исследования эффективных путей повышения речевой компетенции будущих юристов. Задача примененного констатирующего эксперимента заключалась в определении результатов изучения языкового курса в ВНЗ такого профиля.

Ключевые слова: языковой курс, речевая компетенция, речевое действие, речевые умения.

Summary

The given article is a part of the effective methods investigation of the communicative competence improvement of the law school students. The aim of the stating experiment was to find out the results of the adapted language course at Higher Educational Institution.

Keywords: language course, communicative competence, communicative activity, communicative ability.

УДК 81'373:004:371.315

Сидоренко О.М.,
кандидат філологічних наук,
Київський національний
університет ім. Тараса Шевченка

УКРАЇНЬСЬКА КОМП'ЮТЕРНА ЛЕКСИКОГРАФІЯ ЯК ВАЖЛИВИЙ ІННОВАЦІЙНИЙ ЧИННИК НАВЧАЛЬНОГО ПРОЦЕСУ

Постановка проблеми. Безсумнівним є той факт, що у вік бурхливого розвитку комп'ютерно-інформаційних технологій перспективним напрямком подальшої словникарської роботи можна вважати комп'ютерну лексикографію. Такий різновид видань значно спрощує і пришвидшує пошук необхідної інформації та її використання.

Аналіз досліджень і публікацій. Уже нині ми спостерігаємо велику кількість енциклопедичних словників, розміщених у мережі Інтернет. Одним із найбільших є *Вікіпедія*⁵⁹ – міжнародний проект, спрямований на створення універсальної енциклопедії. Словник є непрофесійним, адже наповнення та редагування може здійснювати кожен користувач, але ця певна недосконалість компенсується швидкістю розширення та відточування словникових статей. Також своєю актуальністю вирізняються словники, які є розділами пошукових систем (наприклад, *Янедкс*⁶⁰), в них можна отримати розгорнуту довідку майже про будь-яке потрібне слово, а також – його переклад на кілька мов за електронним словником *АВВУУ Lingvo*. Статті тут взято із найвідоміших енциклопедій та лінгвістичних словників, щоправда, російських або радянських, оскільки сам ресурс є російським (наприклад, *“Толковый словарь русского языка”* в 4 томах Д.Н. Ушакова (1935–1940), *“Большая советская энциклопедия”* в 30 томах (1969–1978), *“Новейший философский словарь”* під редакцією А.А. Грицанова (2003) тощо).

⁵⁹ <http://uk.wikipedia.org>

⁶⁰ <http://www.yandex.ru>

Загалом, в РуНеті (російському Інтернеті) такого типу ресурси становлять досить розгалужену систему. Наприклад, вражає високою якістю довідково-інформаційний портал *Русский язык*⁶¹, створений у 2000 році, що містить величезну кількість різнотипних лінгвістичних словників.

В Україні подібна практика лише набирає обертів. На сьогодні співробітниками Лабораторії комп'ютерної лінгвістики Інституту філології КНУ ім. Т. Шевченка розробляється некомерційний мовний портал *mova.info*⁶², який уже зараз містить такі словники, як “*Відкритий словник*” (новітніх термінів), “*Відкритий словник виправлень суржику*” (з можливістю для користувачів вносити правки), “*Електронний словник порівнянь*”, частотні словники (сучасної української прози та поезії), декілька перекладних словників тощо.

Загалом же, в Україні існують галузеві словники енциклопедичного типу при відповідних тематичних сайтах. Наприклад, на сайті Інтернет-видання *Проза* створено “*Енциклопедію української літератури*”⁶³, з актуальним, але поки що, на жаль, необ'ємним наповненням. Частина енциклопедії “*Українська мова*” опублікована на сайті *Ізборник*⁶⁴, який подає та систематизує тексти давньої української літератури та супровідні матеріали. “*Великий тлумачний словник сучасної української мови*”⁶⁵, що базується на основі одноіменного словника за редакцією В. Бусела (170 000 слів), має реєстр в понад 207 000 словникових статей та інтерактивно поповнюється кожного дня. Кілька українсько-іншомовних словників також розміщено на сайті *Novamova*⁶⁶ та на сайті Українського мовно-інформаційного фонду “*Лінгвістичний портал*”⁶⁷.

Мета статті – проаналізувати стан і перспективи комп'ютерної лексикографії в Україні, окреслити шляхи оптимізації діяльності науковців-лексикографів.

Щодо розробки електронних лінгвістичних словників в Україні, то на сьогодні це надзвичайно актуальна проблема. Диктує умови інтенсивного розвитку не лише власне наукова потреба об'єктивації досліджень (оптимізація та раціоналізація професійної роботи мовознавця), а й позанаукові – необхідно задовольнити зростаючий попит пересічного користувача на адаптовану для нього об'єктивну, достовірну інформацію мовознавчого характеру у вигляді електронних словників різного типу.

Початковим етапом подібних комп'ютерних досліджень в Україні можна вважати Всесоюзну конференцію з проблем створення машинного фонду для автоматизованої системи лексикографічних досліджень (1983 рік) за участі українських науковців – працівників Інституту української мови ім. О.О. Потебні та Інституту кібернетики ім. В.М. Глушкова (В.С. Перебийніс, М.М. Пещак та інших).

⁶¹ <http://www.gramota.ru>

⁶² <http://www.mova.info>

⁶³ <http://proza.com.ua/enc>

⁶⁴ <http://litopys.org.ua/ukrmova/um.htm>

⁶⁵ <http://www.slovyk.net>

⁶⁶ <http://novamova.com.ua>

⁶⁷ <http://www.ulif.org.ua>

Нині існує декілька наукових установ, де проводиться наукова робота з впровадження комп'ютерних технологій у царину лінгвістичних досліджень.

Відділ структурно-математичної лінгвістики Інституту мовознавства ім. О.О. Потебні НАН України можна назвати фундатором комп'ютерної лінгвістики в українському мовознавстві. Тут створено *Морфемно-словотвірний фонд української мови*, який має розгалужену архітектоніку і складається з трьох основних підфондів:

1) текстової бази, що містить на сьогодні близько 700 000 слововживань, оснащена процедурами орфографічного контролю текстів, аналізу їхньої морфологічної, синтаксичної та семантичної структури;

2) генерального реєстру українських слів, що становить 166385 лексем, зведеного за матеріалами 5-ти найґрунтовніших українських словників-джерел;

3) текстових процесорів, що виконують морфологічний, синтаксичний та логіко-семантичний аналіз фактичного матеріалу.

За матеріалами фонду укладені комп'ютерні: *Словник символних моделей морфемної будови слова*, *Словник афіксальних морфем української мови* (виданий також у паперовому вигляді), *Кореневий гніздовий словник української мови* Є.А. Карпіловської, *Ідеографічний словник іменників української мови* Н.В. Сніжко, *Ідеографічний словник дієслів переміщення української мови* А.Я. Середняцького. Здійснюються спроби створити комп'ютерні версії *Словника староукраїнської мови XIV–XV століття* та декількох словників лінгвістичний термінів.

Національний мовно-інформаційний фонд НАН України. Тут під керівництвом В.А Широкова створюється Національна словникова база України. У 2001 році випущено компакт-диск інтегрованої лексикографічної системи *Словники України*, яка має абетковий принцип побудови та пошукову систему обсягом близько 152000 лексем (виданий також друком). Користувачам запропоновано здійснювати роботу у 5-ти режимах: “парадигма”, “транскрипція”, “фразеологія”, “синонімія” та “антонімія”. В основу кожного з цих розділів покладено по декілька найґрунтовніших словників відповідного різновиду. У цьому фонді також започатковано універсальну систему підтримки дослідницьких лексикографічних проєктів – *Український лінгвістичний корпус*, що на сьогодні нараховує 36 млн. слововживань з українських текстів різної стильової спрямованості. Подальша робота фонду зорієнтована на видання нової версії тлумачного *Словника української мови* обсягом на рівні 20 томів.

Лабораторія комп'ютерної лінгвістики кафедри сучасної української мови Київського Національного Університету ім. Т. Шевченка має такі основні напрямки діяльності, як комп'ютерна лексикографія, створення навчальних програм та машинний переклад. Основні здобутки лабораторії у створенні *Морфемно-словотвірної бази української мови* (170000 слів); виданні електронного підручника української мови з інтерактивним тестуванням для широкого загалу; створенні граматичного словника українських дієслів з перекладом на італійську мову (близько 3000 одиниць); створенні частотних словників художньої прози та

публіцистики та частотного словника сучасної поетичної української мови кінця ХХ століття обсягом у 300000 слововживань. Тут також досліджуються принципи англо-українського та українсько-англійського машинного перекладу, розробляються прикладні та термінологічні словники, програми аналізу та синтезу людського мовлення.

Відділ лексикології та комп'ютерної лексикографії Інституту української мови НАН України здебільшого спеціалізується на теоретичній роботі з проблем комп'ютерної лінгвістики. Одним із основних здобутків його є електронна картотека, яка перебуває у перманентній стадії наповнення текстами. Також започатковано велику роботу над створенням *Національного корпусу української мови* з мінімальною кількістю слововживань в 1 млн. одиниць, призначеного для суто наукових завдань: збереження текстового матеріалу, забезпечення наукових досліджень лексичної та граматичної структури мови, а також – простеження динаміки і якості змін у мовній системі протягом певного хронологічного періоду.

Кафедра української мови Донецького національного університету. У 2000 році тут упорядковано та видано на компакт-диску фонотеку *Українські говірки Донеччини* – базу діалектологічних даних у звуковій та графічній формах. До диска увійшли аудіозаписи зв'язаних текстів, здійснених у 65 населених пунктах Донеччини протягом 1997–2000 років.

У стінах **Університету “Львівська політехніка”** створено лексикографічний процесор *Слово*, який дозволяє на основі бази даних термінологічних одиниць створювати одно- та багатомовні термінологічні словники. В основу *Слова* покладений *Англо-українсько-російський словник з інформатики та обчислювальної техніки*.

Висновки. Попри безумовно прогресивний розвиток лексикографічних досліджень через призму комп'ютерних технологій, поки що не створено єдиного комплексу (так званого “робочого місця” лінгвіста) з доступом до масиву фактичного матеріалу, що оптимізував би аналітичну роботу вчених. Як слушно зазначає Є.А. Карпіловська, “накопичений в україністиці досвід створення лінгвістичних баз даних, формування машинних копій та версій різнотипних традиційних (“паперових”) словників, розроблення лінгвістичних словникових та текстових процесорів ставить на часі завдання об'єднання наявної інформації на єдиній концептуальній та методико-процедурній основі в загальнодержавний комп'ютерний фонд української мови, який виконував би всі властиві такій інституції функції: інформаційно-довідкову, дослідницьку, навчальну та редакційно-видавничу” [3, 101].

Подальша перспектива дослідження полягає у створенні єдиного лексикографічного комплексу, і розв'язання цього завдання передбачає співробітництво усіх фахівців – лінгвістів, математиків-програмістів, фахівців із різних галузей знань, а також вироблення правової основи для такої співпраці.

Література

1. Балог В. Сучасний стан української комп'ютерної лінгвістики / В. Балог // Лексикографічний бюлетень (ІУМ). – Вип. 11. – К., 2005.

2. Дарчук Н. П. Комп'ютерна лінгвістика (автоматичне опрацювання тексту) : підручник [для студ. вищих навч. Закладів] / Н. П. Дарчук ; КНУТШ. – Київ : Київський університет, 2008. – 352 с.
3. Дубічинський В. В. Українська лексикографія : історія, сучасність та комп'ютерні технології : навч. посібник / В. В. Дубічинський. – Харків, 2004.
4. Корпусна лінгвістика / відпов. ред. Широков В. А. – К., 2005.
5. Карпіловська Є. Вступ до комп'ютерної лексикографії / Є. Карпіловська. – К., 2004.
6. Широков В. А. Елементи лексикографії / В. А. Широков. – К., 2005.

Анотація

У статті розглядаються проблеми перспективної галузі лексикографічних досліджень – комп'ютерної лінгвістики, аналізуються стан і перспективи комп'ютерної лексикографії в Україні, зокрема наукові установи, де проводиться робота із впровадження комп'ютерних технологій у лексикографію, окреслюються шляхи оптимізації діяльності науковців-лексикографів.

Ключові слова: комп'ютерна лексикографія, корпусна лінгвістика, лінгвістичні бази даних.

Аннотация

В статье рассматриваются проблемы перспективной отрасли лексикографических исследований – компьютерной лингвистики, анализируются состояние и перспективы компьютерной лексикографии в Украине, в частности научные учреждения, где проводится работа из внедрения компьютерных технологий в лексикографию, очерчиваются пути оптимизации деятельности ученых-лексикографов.

Ключевые слова: компьютерная лексикография, корпусная лингвистика, лингвистические базы данных.

Summary

The problems of perspective industry of lexicographic researches are examined in the article – computer linguistics, consisting and prospects of computer lexicography is analyzed of Ukraine, in particular scientific establishments, where work from introduction of computer technologies is conducted in lexicography, the ways of optimization of activity of scientists-lexicographers are outlined.

Keywords: computer lexicography, cabinet-type linguistics, linguistic databases.