

# Застосування комп'ютерних технологій при дослідженні медичної термінології

Роксолана-Юстина Перхач

Кафедра прикладної лінгвістики, Національний університет «Львівська політехніка», вул. С. Бандери, 12, Львів, 79013, УКРАЇНА, E-mail: perkhach.roksolana@gmail.com

*This paper discusses the current and potential use of corpora in linguistic researches and the role of corpus-based method in analyzing the terminology of medicine.*

Ключові слова: термінологія, корпус текстів, корпусний менеджер, інструкція до лікарського засобу.

Фахова мова медицини достатньо своєрідна, оскільки медична лексика постійно поповнюється, вона вміщує слова, які не лише складають основу словесного фонду кожної мови, але й визначають фундаментальні поняття та цінності людства, соціальної діяльності, безпеки та життя людини.

Зі швидким розвитком медицини, збільшується і кількість медичних термінів, які опрацювати без застосування технічних засобів важко. Для швидшого та ефективнішого опрацювання мовної інформації лінгвістами впроваджуються комп'ютерні технології, які дають можливість перетворювати природню мову в форму електронної інформації.

Однією з таких форм електронної інформації є лінгвістичний корпус текстів. Під лінгвістичним корпусом текстів ми розуміємо представлені в автоматизованому вигляді, структуровані та філологічно компетентні мовні дані, призначені для вирішення конкретних лінгвістичних завдань, таким чином, дотримуємося визначення українського мовознавця О. Демської-Кульчицької, яка визначає *текстовий корпус* як, певним чином організоване електронне зібрання писемних і усних текстів довільної природньої мови, якому притаманні обов'язкові ознаки і яке призначене для здійснення наукового дослідження мови [3;89].

Чимало досліджень присвячено питанням корпусної лінгвістики та класифікації лінгвістичних корпусів, зокрема варто зазначити таких дослідників як, А. Баранов, М. Бейкер, О. Бугаков, О. Демська-Кульчицька, Б. Захаров, Є. Карпіловська, В.Риков, Л.Ричкова, Дж. Синклер, В.Широков [1,3-6,8-11].

Вищенаведені дослідники виділяють декілька типів лінгвістичних корпусів, які можна поділити: за типом мовних даних (письмові, усні, змішані); за «паралельністю» (одномовні, двомовні, багатомовні); за «літературністю» (літературні, діалектні, розмовні, термінологічні, змішані); за метою створення (багатоцільові та спеціалізовані); за жанром (літературні, фольклорні, драматургічні, публіцистичні); за доступністю (вільнодоступні, комерційні та закриті); за призначенням (дослідницькі й ілюстративні); за динамічністю (динамічні і статичні); за розміткою (розмічені і нерозмічені); за характером розмітки (морфологічні, семантичні,

синтаксичні, просодійні та інші); за об'ємом тексту (повнотекстові і «фрагментовані»)[11].

Невід'ємною частиною поняття *лінгвістичний корпус* є корпусний менеджер, завдяки якому маємо змогу вивчати лексичні одиниці в контексті, опрацьовувати списки колокацій, створювати словники частотності, групувати лексичні одиниці в лексико-семантичні поля.

Існує досить велика кількість корпусних менеджерів, зокрема: Ant Conc, EXMARaLDA Corpus-Manager (Co-Ma), KWIC Concordance for Windows, LEXA, Manatee (Bonito), XAIRA, Simple Concordance Program, Virtual Corpus Manager (VMC).

Результати пошуку в корпусному менеджері представлені у вигляді конкордансу, спеціалізованої лінгвістичної програми, за допомогою якої здійснюють автоматичну вибірку лексичних одиниць з текстових даних у електронному вигляді.

Завдяки інструментам корпус-менеджера ми маємо змогу розглянути, які слова і фрази зазвичай використовуються в різних контекстах (конкорданс); результати дослідження введеного користувачем ключового слова в контексті (колокації); кількість всіх слів в корпусі та представлення їх у впорядкованому списку за алфавітом, або за частотністю (список слів).

Таким чином завдяки вищенаведеним інструментам складання конкордансів ми маємо можливість отримати дані про частоту словоформ, частоту лем, граматичних категорій та список контекстів, в яких функціонують дані лексичні одиниці.

За допомогою лінгвістичного корпусу можна вирішувати проблеми з тлумаченням медичних термінів в інструкціях до лікарських засобів, які використовують як, і лікарі та фармацевти, так і нефахівці медичної галузі, які мають потребу у використанні лікарських засобів.

Інструкціям до лікарських засобів властиві змістовність і чіткість, що вигідно відрізняє їх від інших фармацевтичних текстів, які містять медичну термінологію. Окрім того, дослідження інструкцій до лікарських засобів дасть змогу простежити специфіку вживання як фармацевтичної, так і медичної термінології.

Аналіз термінології, поданої в інструкціях до лікарських засобів, здійснено комплексно: зіставлено українську, польську та німецьку медичну термінологію.

Відповідно до держстандарту України створено українськомовні інструкції до лікарських засобів, польськомовні та німецькомовні інструкції згідно зі

правовими нормами структур Європейського союзу, оскільки Німеччина та Польща є членами Європейського Союзу.

Джерельну базу інструкцій до лікарських засобів сформовано за ознакою належності ліків до певних фармакологічних груп, а саме згідно з міжнародною системою класифікації лікарських засобів – анатомо-терапевтично-хімічною класифікацією. У фармакології та медицині використовують позначення АТС (від англ. Anatomical Therapeutic Chemical). Відповідно до класифікації АТС виокремлюють такі підгрупи:

Код А: Лікарські засоби, що впливають на травну систему та обмін речовин.

Код В: Лікарські засоби, що впливають на кровотворення та кров.

Код С: Лікарські засоби для лікування захворювань серцево-судинної системи.

Код D: Лікарські засоби для лікування захворювань шкіри.

Код G: Лікарські засоби для лікування захворювань уrogenітальних органів та статеві гормони.

Код H: Гормональні лікарські засоби для системного використання (виключаючи статеві гормони).

Код J: Протимікробні лікарські засоби для системного використання.

Код L: Протипухлинні лікарські засоби та імунomodulatory.

Код M: Лікарські засоби для лікування захворювань кістково-м'язевої системи.

Код N: Лікарські засоби для лікування захворювань нервової системи.

Код P: Протипаразитарні лікарські засоби, інсектициди та репелленти.

Код R: Лікарські засоби для лікування захворювань дихальної системи.

Код S: Лікарські засоби для лікування захворювань органів чуттів.

Код V: Інші лікарські засоби.

У нашому лінгвістичному дослідженні ми враховували всі фармакологічні групи, окрім лікарських засобів, які відносяться до підгрупи «Ветеринарні лікарські засоби (код Q)», оскільки досліджуємо інструкції до лікарських засобів, які використовують люди. Таким чином було відібрано сто дев'яносто п'ять інструкцій до лікарських засобів (шістдесят п'ять українськомовних, шістдесят п'ять німецькомовних та шістдесят п'ять польськомовних).

Результати проведеного лінгвістичного аналізу свідчать, що існування електронних корпусів текстів дає можливість значно розширити і автоматизувати аналіз мовного матеріалу, який є найважливішою базою лінгвістичного дослідження. Інструкції до лікарських засобів містять значну кількість термінологічних одиниць, які виконують роль своєрідної комунікації між лікарем, фармацевтом та пацієнтом-неспеціалістом, тому, актуальним є

зіставний аналіз текстів таких інструкцій до лікарських засобів.

## Література

1. Баранов А. Н. Введение в прикладную лингвистику. – М., 2003. – 360 с.
2. Браунівський корпус. [Електронний ресурс]. – Доступно з (<http://khnt.aksis.uib.no/icame/manuals/brown/>). 3
3. Демська-Кульчицька О. Текстовий корпус: ідея іншої форми / О. Демська. – К., 2011.
4. Демська-Кульчицька О. [Електронний ресурс]. – Доступно з ([http://www.ekmair.ukma.kiev.ua/bitstream/123456789/1704/1/Demska\\_Deshcho%20pro%20klasyfikatsiiu.pdf](http://www.ekmair.ukma.kiev.ua/bitstream/123456789/1704/1/Demska_Deshcho%20pro%20klasyfikatsiiu.pdf)). 3
5. Захаров В. П., Коваль С. А. Корпусная лингвистика и лингвистические базы данных // НТИ: Информационные процессы и системы. — 2002. — № 7. — С. 62-69
6. Карпіловська Є. Вступ до комп'ютерної лінгвістики / Є. Карпіловська. – Донецьк, 2003.
7. Корпусна лінгвістика / В. А. Широков, О. В. Бугаков, Т. О. Грязнухіна та ін. – К., 2005.
8. Рычкова Л.В. Проблема саставных аб'ектау у корпусах славянскому і лінгвістичных базах дадзеных // Мовознавства. Література. Культуралогія. Фалькларыстыка. XIII Міжнародный з'езд славыстау. Доклады беларускай дзлегацыі. —Мінськ, 2003. - С. 184-195.
9. Словники України on-line [Електронний ресурс]. – Доступно з: (<http://lcorep.ulif.org.ua/dictua>).
10. Baker M. 1995. Corpora in translation studies: An overview and some suggestions for future research. Target 7(2): 223–243
11. Sinclair J. Corpus, Concordance, Collocation / J.Sinclair – Oxford: Oxford University Press, 1991. – 170 p.