Усачова К. С., кандидат філологічних наук, доцент кафедри юридичного документознавства Національної академії внутрішніх справ

ЗАСТОСУВАННЯ ПРИЙОМІВ КОРПУСНОЇ ЛІНГВІСТИКИ В ДОСЛІДЖЕННІ ЮРИДИЧНОЇ ЛЕКСИКИ

Термінологія сучасної української мови належить водночас до найбільш активно досліджуваних у науково-теоретичних роботах та найбільш проблемних у сфері практичного функціонування. Складність питань у дослідженнях та функціонуванні термінології зумовлена її унікальним становищем на межі лінгвістики та певної конкретної галузі науки та техніки. Таким чином, окрім суто мовних проблем (надмірні запозичення, наявність синонімів та багатозначних термінів тощо), розвиток певної терміносистеми може гальмуватися проблемами тієї науки, поняття якої вона називає.

Розбіжності між теоретичним концептом та практичним слововживанням можуть стати проблемою як для лінгвіста-дослідника термінології, так і для науковця, який досліджує теоретичні та практичні проблеми певної наукової галузі. Адже точно та коректно сформульований з погляду науки, термін може виявитися незручним у використанні та породжувати численні професіоналізми, чи заміщатися загальновживаними відповідниками. З іншого боку, більш зручні та загальнозрозумілі терміни можуть із часом набувати конотацій, що впливають на правильне розуміння терміна мовцями (ситуація, важлива під час опанування певного фаху).

Комп'ютеризація сучасного суспільства та науки утворює нові можливості дослідження практичного функціонування фахової лексики у мовленні різних верств населення та текстах різних стилів та жанрів. До таких, зокрема, належать прийоми корпусної лінгвістики.

Метою даної розвідки є розгляд перспектив використання прийомів корпусної лінгвістики у дослідженнях практичного вживання юридичної лексики у текстах різних сфер.

Корпусна лінгвістика належить до відносно нових галузей мовознавства. Відповідно, існують проблеми як зі створенням загальноприйнятих визначень базових понять галузі

(«лінгвістичний корпус текстів»), так і з фундаментальним завданням створення національного корпусу української мови.

У найбільш загальному значенні, корпусна лінгвістика займається вивченням принципів побудови та обробки даних корпусів текстів із використанням комп'ютерних технологій. До характерних ознак, що дозволяють виокремити корпус текстів (на відміну від звичайних збірок текстів у електронній формі) належать: відібраність, репрезентативність, збалансованість, машиночитаність, стандартність [1, с. 56–58]. Подібні корпуси (або національний корпус певної мови) надають безпосередні та об'єктивні дані про реалізацію мовної системи (у синхронії та діахронії) і, таким чином, виникає можливість використовувати їх для оновлення чи створення словників та граматик.

Традиційні напрямки використання прийомів корпусної лінгвістики: лексикографія, навчання іноземній мові, переклад (автоматизований). Проте корпуси можуть використовуватися також у соціології та соціолінгвістиці, літературознавстві, історичних дослідженнях тощо. У правовій сфері прийоми корпусної лінгвістики активно використовують у сфері криміналістичної лінгвістики.

Одним із найбільш доступних для загального використання нині є корпус українських художніх, публіцистичних, наукових та законодавчих текстів, сформований у лабораторії комп'ютерної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка (http://www.mova.info). Корпус дає можливість отримати інформацію двох типів: контексти вживання шуканих одиниць (конкорданси), кількісні характеристики вживання мовних одиниць у певному тексті. Проте використання даного корпусу має певні обмеження (як з боку гнучкості використання, так і з боку доступності різних типів даних).

Наприклад, аналіз функціонування синонімічної пари *самогубство/суїцид* у підкорпусі законодавчих текстів демонструє, що лексема *суїцид* відсутня у слововжитку взагалі, тоді як лексема *самогубство* використовується обмежено, як частина терміносполуки *доведення до самогубства*. Аналіз підкорпусу наукових текстів загалом засвідчує подібну тенденцію – лексема *самогубство* вживається суттєво частіше та ширше за контекстом, ніж лексема *суїцид*. Проте значно відрізняється специфіка слововживання синонімічної пари у

підкорпусі публіцистичних текстів. Обидві лексеми вживаються з однаковою частотністю, проте конкорданси лексеми *суїцид* належать до більш офіційного стилю викладу із виразними психологічними (психіатричними) асоціативними зв'язками, тоді як конкорданси лексеми *самогубство* засвідчують більш розмовно-розповідний характер. Недостатня репрезентованість фахових текстів та обмеженість можливостей розробленої програми не дають змогу поглибити аналіз у даному випадку.

Варто зазначити, що окрім випадків, коли корпус є ключовим фокусом дослідження, використання прийомів корпусної лінгвістики може бути корисним допоміжним прийомом. Проте таке використання потребує двох ключових факторів: навичок зі створення корпусу текстів (корпус, навіть у мікроваріанті, має відповідати всім вищезгаданим ознакам) та обробки отриманих даних, а також комп'ютерної програми, що використовується для аналізу даних корпусу (конкордансер чи комп'ютерний менеджер, напр.: Wordsmith, AntConc).

Прикладом може слугувати аналіз викликаної екстралінгвістичними чинниками спроби адаптації лексики кримінально-виконавчої галузі права під новий вектор розвитку сфери виконання покарань (зміна властивого радянській системі акценту на покаранні на властивий європейській моделі акцент на ресоціалізації).

Задля такого аналізу було створено корпус на основі текстів трьох тематичних категорій: офіційні документи зазначеної галузі; наукові статті з пенітенціарної проблематики у фахових збірниках; тексти новин на сайті ДПтСУ. Даний підбір давав змогу проаналізувати офіційний дискурс відповідної галузі у межах певного часового зрізу.

Створення простого списку слів, вживаних в усіх текстах продемонструвало умовний баланс між частотою вживання лексики, що демонструє акцент на системі покарань (напр., «покарання», «виконання», «влада») та лексики, що демонструє акцент на особистості та виправленні засудженого (напр., «засуджений», «виправлення»).

Ситуація помітно змінилася, коли було проаналізовано дані слова як лексеми, використовуючи конкорданси. Використання лексем «виконувати», «карати», «влада», «кримінальний» майже рівномірне за частотою вживання в усіх типах текстів. Тоді як лексема «виправити» використовується

виключно у наукових текстах, лексема «засуджений» має найбільшу частоту вживання у наукових текстах та значно меншу – у текстах новин.

Слова, найбільш тісно асоційовані з новим вектором розвитку сфери виконання покарань – «ресоціалізація» та «пенітенціарний» – засвідчили мінімальну частоту вживання в усіх текстах. Лексема «ресоціалізація» вживалася виключно у наукових текстах. Лексеми «соціальний» та «суспільний» також мали мінімальну частоту вживання, переважно у наукових текстах, а також у текстах новин.

Наведений аналіз специфіки функціонування лексики у текстах пенітенціарної (кримінально-виконавчої) тематики виявив тенденції до переважання (певної консервації) слововжитку, пов'язаного з орієнтацію на систему та покарання, тоді як слова, пов'язані з орієнтацією на людину та виправлення функціонують майже виключно у наукових текстах, де вживаються поряд із більш традиційною для галузі лексикою.

Метою аналізу було охарактеризувати специфіку проблемного явища, тоді як глибокий науковий аналіз специфіки слововживання у зазначеній сфері потребував би створення динамічного корпусу із залученням ширшого кола текстів, а також наявності референтного корпусу сучасної української мови права. Проте навіть обмежене застосування прийомів дозволяє висвітлити окремі специфічні риси слововжитку та терміновжитку у фаховій сфері.

Для якісного дослідження особливостей функціонування юридичної лексики у текстах різних стилів, тематики, структури та призначення, корпус має слугувати моделлю мовної системи, а це потребує значної підготовчої роботи. З іншого боку, один раз створений корпус мови права може надалі використовуватися для різноманітних досліджень.

Список використаних джерел

1. Демська-Кульчицька О. Основи національного корпусу української мови : монографія / О. Демська-Кульчицька. – Київ : Ін-т укр. мови НАН України, 2005. – 218 с.