

Карнаух Ганна Володимирівна,
молодший науковий співробітник
Українського мовно-інформаційного
фонду НАН України

**НОРМАТИВНО-ПРАВОВІ АКТИ ЯК ОБ'ЄКТ
ДОСЛІДЖЕННЯ НЕОДНОЗНАЧНОСТІ СЛОВОФОРМ
(НА МАТЕРІАЛІ ТЕКСТІВ УКРАЇНСЬКОГО
НАЦІОНАЛЬНОГО ЛІНГВІСТИЧНОГО КОРПУСУ)**

Швидкий темп сучасного життя, спричинений бурхливим розвитком інноваційних технологій, супроводжується постійними змінами в мові, а це, у свою чергу, вимагає відображення цих змін (відповідно до вимог часу) у словниковій продукції. Тому оновлення фундаментальних лексикографічних праць - модифікація їхнього змісту й структури, а також приведення у відповідність з новими реаліями - виявилось не лише актуальним мовознавчим завданням, але й нагальною потребою українського суспільства.

Однак при цьому виникла ціла низка конкретних питань - починаючи від визначення принципів укладання нового лексикону «поверх» старого і закінчуючи технологією та організацією виконання такого лексикографічного проекту [4]. Зрозуміло, що ефективне вирішення зазначених проблем викликало потребу створення певної узагальнювальної теоретичної схеми, спроможної надати уніфіковані засоби до їх розв'язання, а також необхідність побудови комп'ютерного

інструментального комплексу для реалізації зазначеного проєкту¹.

Як відомо, найновіші лінгвістичні дослідження часто охоплюють предметне поле декількох наукових дисциплін або здійснюються в рамках різних парадигм. З огляду на це останнім часом чітко окреслився системно-функціональний підхід до вивчення мови, який дозволяє здійснювати всебічний аналіз усіх параметрів досліджуваного об'єкта. Саме у межах цього підходу варто розглядати комунікативну функцію мови. До того ж детальне вивчення дієвих механізмів, правил та особливостей комунікації дозволяє розробити уніфіковані норми, моделі спілкування в певних сферах соціального життя [2].

При цьому для систематичного опису лексики потрібне її якомога повніше зібрання. Цю проблему вирішено завдяки Українському національному лінгвістичному корпусу обсягом понад 64 млн слововживань, створеному в Українському мовно-інформаційному фонді НАН України [6]. Цей корпус замінив собою традиційну лексичну картотеку та став одним із основних джерел текстово-ілюстративного матеріалу до Словника української мови в 20 томах. Загалом, зазначена програмна бібліотека - це представлений в електронному вигляді, великий за обсягом, уніфікований, структурований, розмічений і філологічно компетентний масив мовних даних, доповнений системою керування - універсальними програмними засобами для пошуку різноманітної лінгвістичної інформації, перевагами якого є великий обсяг мовного матеріалу, що залучається до мовознавчого дослідження, комплексність, оперативність опрацювання зазначеного матеріалу та можливість прямого доступу до значної кількості лінгвістичних фактів [2].

Одним із важливих компонентів програмного забезпечення корпусу текстів є морфологічний аналіз словоформ, метою якого є визначення граматичних параметрів досліджуваної одиниці. Однак суттєвою перешкодою на шляху

¹ Тлумачний словник української мови в 20-ти томах є одним із центральних проєктів програми створення Національної словникової бази, ініційованої Указом Президента України від 7 серпня 1999 р. № 967, координатором якого є Український мовно-інформаційний фонд НАН України.

реалізації морфологічної розмітки корпусу є явище неоднозначності¹ слівформ. Тому постає потреба створення програмних засобів, спрямованих на вирішення цієї проблеми. Крім того, побудова універсального коду для адекватного сприйняття і надання потрібної інформації є важливою складовою сучасних технологій створення мов машинного програмування, пошукових систем, а також систем автоматичного опрацювання мови (переклад, автокорекція та ін.) [4].

Специфіка мови законодавчих документів² обумовлює певні особливості прояву неоднозначності слівформ у текстах нормативно-правових актів. Це дозволяє виробити певні правила фільтрації так званого шуму та проаналізувати типи можливих помилок при автоматичному маркуванні неоднозначної слівформи.

1. Неправильне визначення лем

Застосування мов в Україні гарантується Конституцією України та визначається законом³.

Промарковано: 'мов' - сполучник.

Має бути: 'мов' - іменник, жіночий рід, родовий відмінок, множина.

2. Неправильне визначення граматичного значення:

2.1. Відмінок

Народ здійснює владу безпосередньо і через органи державної влади та органи місцевого самоврядування.

¹Зрозуміло, що омонімія та полісемія - різні поняття. Однак не можна при цьому не визнавати того факту, що вони мають загальну ознаку - відмінність значень при тотожності форми, що дозволяє об'єднати їх у більш широке поняття неоднозначності [3, с. 2]

²Офіційність, конкретність змісту; чіткість, стислість, логічна послідовність, однозначність формулювань; несуперечність аргументації; усталеність та одноманітність форми документів; канцелярська лексика, відсутність емоційної образності, стандартні звороти (кліше); безособові та наказові форми; відсутність індивідуальних авторських рис; монологічна (рідше діалогічна) форма тексту П, с. 55].

Цей приклад (а також усі подальші) наведено з тексту Конституції України [Конституція України. Закон від 28.06.1996 № 254к/96-ВР / [Верховна Рада України]. - Режим доступу : <http://zakon2.rada.gov.ua/laws/>.]

Промарковано: ‘народ’ - іменник, чоловічий рід, знахідний відмінок, однаина.

Має бути: ‘народ’ - іменник, чоловічий рід, називний відмінок, однаина.

2.2. Відмінок, число

Норми Конституції України є нормами прямої дії.

Промарковано: ‘норми’ - іменник, жіночий рід, родовий відмінок, однаина.

Має бути: ‘норми’ - іменник, жіночий рід, називний відмінок, множина.

2.3. Число

Земля, її надра, атмосферне повітря, водні та інші природні ресурси, які знаходяться в межах території України, природні ресурси її континентального шельфу, виключної (морської) економічної зони є об’єктами права власності Українського народу.

Промарковано: ‘є’ - дієслово, недоконаний вид, теперішній час, третя особа, однаина.

Має бути: ‘є’ - дієслово, недоконаний вид, теперішній час, третя особа, множина.

2.4. Рід

Держава створює умови для повного здійснення громадянами права на працю, гарантує рівні можливості у виборі професії та роду трудової діяльності, реалізовує програми професійно-технічного навчання, підготовки і перепідготовки кадрів відповідно до суспільних потреб.

Промарковано: ‘професійно-технічного’ - прикметник, середній рід, родовий відмінок, однаина

Має бути: ‘професійно-технічного’ - прикметник, чоловічий рід, родовий відмінок, однаина.

2.5. Форма дієслова

Коаліція депутатських фракцій у Верховній Раді України відповідно до цієї Конституції вносить пропозиції Президенту України щодо кандидатури Прем’єр-міністра України, а також відповідно до цієї Конституції вносить пропозиції щодо кандидатур до складу Кабінету Міністрів України.

Промарковано: ‘вносить’ - дієслово, недоконаний вид, інфінітив.

Має бути: ‘вносить’ - дієслово, недоконаний вид, теперішній час, третя особа, однаина.

Отже, основне завдання законодавчої мови (владно приписувати суб'єктам права певну поведінку, формулювати вимоги, загальнообов'язкові норми) диктує певні правила та особливості оформлення нормативно-правових документів на всіх рівнях мови. Це у свою чергу дозволяє розробити певні програмні засоби автоматичного опрацювання текстів зазначеного виду. Очевидно, що обсяги мовного матеріалу в межах наукових досліджень, зокрема лінгвістичних, весь час зростатимуть, тому необхідність оперативного, комплексного, а головне - якісного його опрацювання, і, відповідно, потреба подальшого вдосконалення систем автоматичного оброблення природної мови залишаються актуальними й надалі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Колеснікова І. Є. Українське ділове мовлення : навч. посібник / І. Є. Колеснікова, Ю. Ф. Прадід / [за ред. Ю. Ф. Прадіда]. - Сімферополь : ВД «АРІАЛ», 2009. - 210 с.

2. Корпусна лінгвістика / [В. А. Широков, О. В. Бугаков, Т. О. Грязнухіна та ін.] ; під ред. В. А. Широкова. - К. : Довіра, 2005. - 407 с.

3. Кузьмина Р. В. К вопросу о разграничении явлений омонимии и полисемии и об особенностях их представления в орфоэпическом словаре [Электронный ресурс] / Р. В. Кузьмина // Вестник гуманитарного университета ИГХТУ. - Режим доступа : <https://www.isuct.ru/e-publ/vgf/2006/01/129>.

4. Лінгвістичні та технологічні основи тлумачної лексикографії / [В. А. Широков, В. М. Білоноженко, О. В. Бугаков та ін.] ; під ред. В. А. Широкова. - К. : Довіра, 2010. - 296 с.

5. Чулінда Л. І. Текст нормативно-правового акта: загальні підходи до аналізу / Л. І. Чулінда // Науковий вісник Національної академії внутрішніх справ України. - 2002. - № 3. - С. 120 - 126.

6. Шевченко И. В. Электронный грамматический словарь украинского языка. / Шевченко И. В., Рабулец А. Г., Широков В. А. // Труды Международной конференции «MegaLing-2005. Прикладная лингвистика в поиске новых путей», 27 июня - 2 июля 2005 г. - Меганом., 2005. - С. 124-129.