

ти і зацікавити читача [4, с. 28]. У рекламі часті, насамперед, випадки делексікалізації, тобто оновлення колишніх значень, наприклад *Ból gardła odbiera ci głos? Weź "Strepzil"*. Переносне значення фразеологізму: *zabrać komuś głos* (тобто не дозволити продовження виступу) залишається у тексті співвідносним з буквальним і актуальним завдяки контексту (рекламованим продуктом, тому що таблетки від болю у горлі).

Варто зупинитись на прикладі тексту реклами шампуню проти лупи "Head&Shoulders", в якому використаний цілий ряд фразеологічних зворотів, що мають у своєму складі компонент *głowa*, наприклад, *suszyć głowę, rwać włosy z głowy, mieć (coś) nagłowie*. Тут ми можемо побачити взаємодію буквального і переносного значень зворотів.

Іншим прикладом вживання модифікованого фразеологічного звороту є рекламний текст *Masz świat w zasięgu karty*. Цей текст виник внаслідок заміни компоненту фразеологізму *w zasięgu ręki*. Подібним чином відбулася модифікація фразеологізму *wziąć coś na rozum* рекламному тексті заспокійливого препарату *Weź to na spokoj!*

Дослідження показало, що фразеологізми широко використовуються в усіх частинах сучасного тексту ЗМІ: у заголовках, у рекламних та інформаційних блоках.

Особливості фразеологізмів:

- вони відомі більшості людей, тому інформаційні та рекламні тексти, що містять ФО, легко сприймаються мовцями;

- образність фразеологізмів, їх метафоричність має достатній вплив на адресата повідомлення;

- трансформація або модифікація відомих фразеологізмів обов'язково привертає увагу читача чи слухача повідомлення.

Аналіз показав, що прийоми трансформацій по-різному впливають на формування актуального значення фразеологізму. Структурно-семантичний прийом заміни компонентного складу є дієвим стилістичним засобом. Прийом скорочення компонентного складу фразеологізмів рідко зустрічається в газетних текстах. Прийоми трансформацій можуть виступати в газетному тексті комплексно. Це дозволяє "навантажити" текст додатковими смислами, що сприяє успішному вирішенню творчих завдань, які стоять перед журналістом [3, с. 222]. Аналіз фразеологічних одиниць у засобах масової інформації показав, що вони є необхідним будівельним матеріалом для створення експресивності та образності тексту.

Список використаних джерел

1. Ганопольская Е.В. Свободное слово или эзопов язык? // Актуальные проблемы теории коммуникации: Сборник научных трудов. – Спб.: Изд-во СПб ГПУ, 2004. – С. 68-74.
2. Казанцев А.И. Фразеологизмы в рекламных текстах Челябинских СМИ // Вестник челябинского государственного университета, 2009. – С. 108-114.
3. Lewiński P. Retoryka reklamy, Wrocław 1999. , s. 222.
4. Lusińska A. Reklama a frazeologia. Teksty reklamowe jako źródło nowych frazeologizmów, Toruń 2007.

Надійшла до редколегії 20.06.14

Н. Демьяненко, канд. філол. наук, доц.
КНУ імені Тараса Шевченка, Київ

ФРАЗЕОЛОГИЧЕСКИЕ ЕДИНИЦЫ В СРЕДСТВАХ МАССОВОЙ ИНФОРМАЦИИ (НА МАТЕРИАЛЕ ПОЛЬСКОГО И РУССКОГО ЯЗЫКОВ)

В статье рассматривается специфика функционирования фразеологизмов в газетных и рекламных текстах, а также изменения, которые происходят во фразеологических единицах при использовании их в СМИ.

Ключевые слова: фразеологическая единица, средства массовой информации, фразеологические трансформации.

N. Demyanenko, PhD, Associate Professor
Taras Shevchenko National University of Kyiv, Kyiv

PHRASEOLOGICAL UNITS IN THE MEDIA (THE MATERIAL POLISH AND RUSSIAN)

The article reviews the specificity of functioning of phraseology in newspapers and advertising texts, changes occurring in phraseological units when used in the media.

Key words: phraseological units, media, phraseological transformations.

УДК 811.161.2'01:801.82

О. Ніка, д-р філол. наук, доц.
КНУ ім. Т. Шевченка, Київ

ІСТОРИЧНИЙ КОРПУС УКРАЇНСЬКОЇ МОВИ: ОСНОВНІ ПРИНЦИПИ СТВОРЕННЯ

У статті визначаються основні принципи створення історичного корпусу українських текстів для вирішення різноманітних дослідницьких завдань.

Ключові слова: корпусна лінгвістика, історичний корпус українських текстів, принципи побудови історичного корпусу.

Як відомо, створення національного корпусу є обов'язковим щодо рідної мови. У сучасній лінгвістиці частіше виникає питання не про створення окремого корпусу (чи підкорпусу) текстів певної мови, а про його максимальне наповнення і використання для різноманітних дослідницьких завдань (колекції корпусів на www.corpus.leeds.ac.uk, www.uow.edu.au). Історичні корпуси (підкорпуси) текстів уже створені на матеріалі різних мов, слов'янських і неслов'янських: російської, нідерландської, ісландської, американської англійської, валлійської мов. У паралельному корпусі перекладів давніх текстів представлено "Слово о полку Ігоревім". Словник-конкорданс укладений на основі зібрання творів Г. Сковороди.

На часі обговорення концепції побудови, наповнення і використання історичного корпусу української мови, що визначили актуальність цієї наукової проблеми і завдання статті.

У вітчизняній лінгвістиці основні засади створення дослідницького корпусу української мови викладено в монографічних дослідженнях Н. Дарчук [1], Є. Карпіловської [4].

Корпус сучасних українських текстів, розміщений на www.mova.info, підготовлений лабораторією комп'ютерної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка. У корпус включено: законодавчі, наукові, науково-технічні, художні (поетичні і прозові), публіцистичні тексти, а також фольклорні.

Пропозицію "спеціалізувати підкорпуси за хронологією, чи за історичними періодами розвитку української

мови" висловлювала О. Демська-Кульчицька, яка, покликаючись на періодизацію української мови В. Німчука, запропонувала охопити "текстовий матеріал давньоукраїнського (XII–XIV), ранньоукраїнського (XV–XVI), і середньоукраїнського (XVI–XVIII) періодів та сучасну українську мову (XIX–XXI) з виокремленням періоду кінця XX – початку XXI ст." [3].

Без сумніву, основний принцип історичного корпусу текстів – синхронно-діахронний, тобто описати динаміку мови можна через зіставлення двох (чи більше) синхронних зрізів, проте в цьому разі актуалізується визначення назв і точних меж синхронних зрізів української мови, відповідно, їх представлення в текстовому вимірі.

Історичні корпуси будуються за хронологічним критерієм (він, безсумнівно, є основним), а також доповнюватися жанрово-стильовим критерієм.

У цьому плані показовий Регенсбурзький діахронічний корпус (на його основі виконується наукова тема "Корпусна лінгвістика і діахронічний синтаксис" (науковий керівник проекту – проф., д-р Р. Майер).

У складі національного корпусу російської мови є декілька історичних підкорпусів (www.ruscorpora), як-от: "Рукописні пам'ятки Давньої Русі" (берестянські грамоти, літописи), "СКАТ" (агіографічні тексти), "Манустрипт" (літописи, життя, мінеї та ін.). До цього переліку також додано корпус публіцистичних текстів першої половини XIX ст., що вже виходить за межі давнього періоду, який традиційно окреслюється до XVIII ст.

Одним із найбільш наповнених історичних корпусів російської мови є "Манускрипт", який повно репрезентує історичні ресурси кирилицею і глаголицею, що охоплюють X–XIV ст.

Побудова історичних корпусів за хронологічним критерієм увиразнена сукупністю різних часових зрізів, пещрим з яких є синхронний зріз до XIV ст.

В історичному корпусі російської мови наступний синхронний зріз охоплює період XV – початку XVIII ст. Цей ресурс у відкритому доступі функціонує з 2013 року під назвою "Історичний корпус середньоросійських пам'яток XV – початку XVIII ст.". Як зазначають лінгвісти, він налічує 3 млн слововживань, що репрезентують літературні твори, літописи, життя, ділові грамоти, побутове листування цього часу.

Умовність назви "середньоросійська мова" на означення хронологічного періоду XV – початку XVIII ст. прокоментував філолог, учасник цього корпусного проекту Б. Орехов в інтерв'ю електронному ресурсу "ProScience" [5]. На його думку, вибір назви "середньоросійська мова" передусім мотивується тим, що фіксує певну послідовність у хронології синхронних зрізів, а тому в робочому порядку встановлює часову вертикаль в історичних корпусах.

Проаналізувавши основні принципи створення корпусів текстів, зокрема історичних, вважаємо, що оптимально історичний корпус українських текстів будується як дослідницький, загальнономовний, повнотекстовий, двомовний, анотований.

Як правило, корпуси текстів є дослідницькими, оскільки призначені для отримання адекватних результатів для різнопланових завдань науковців.

Загальнономовність корпусу текстів, на відміну від його спеціалізації, реалізується у повному представленні писемних пам'яток різних періодів, без особливого маркування виявленої в цих текстах діалектної, стилістичної та іншої інформації.

Якщо для корпусів сучасних текстів достатньою та показовою є фрагментарність, то історичні джерела в корпусі, на нашу думку, варто репрезентувати повнотекстово. Мотивація цього вибору зумовлена повнотою

представлення лінгвальної інформації в корпусі, фіксованим реєстром джерел. Відкритий характер історичного корпусу конкретизується відповідно до поставленої мети як поступове його розширення.

Попри полілінгвальність мовної ситуації в Україні корпус текстів зорієнтований на репрезентацію писемних джерел староукраїнською та церковнослов'янською мовами, що в такий спосіб характеризує співвідношення принципів мономовності/багатомовності.

Анотованість джерел у корпусі передбачає максимальне врахування загальної інформації про текст: автор, повна назва тексту (заголовок), рік написання (для рукописів), друкування, місце видання, назва друкарні, жанр, стиль, оригінал чи переклад (з якої мови), редакція, список. У паспорті джерела також зазначаються дослідницькі видання писемної пам'ятки, основні дослідження про неї.

Отже, у концепції історичного корпусу української мови передбачено:

- синхронно-діахронічний принцип, який оптимально відображає різні синхронічні зрізи в історії української мови, репрезентує її динаміку.

- принцип різностильової повноти, що репрезентує писемні тексти різних стилів і жанрів кожного синхронного зрізу.

Зокрема, писемні пам'ятки другої половини XVI – XVII ст. характеризують період розквіту староукраїнської літературної мови, відзначають повну картину жанрово-стильового розмаїття української мови старої доби, поступовий перехід до нового, структурно-функціонального принципу стильової диференціації.

Преваги побудови корпусу цього часу – у нових та оновлених жанрах, перекладах, барокових текстових інноваціях та ін., що відрізнятиме корпус української мови з-поміж інших мов. Така мотивація актуалізує створення окремого корпусу українських текстів, що у доступних раритети XVI–XVII ст. для вирішення дослідницьких завдань, дає можливість працювати з великим обсягом лінгвального матеріалу за допомогою спеціально заданого дослідником пошуку.

Послідовність створення такого історичного корпусу (підкорпусу) передбачає кілька етапів:

- 1) добір репрезентативних текстів за встановленими дослідницькими критеріями;

- 2) наукова підготовка текстів, що максимально наближує їх передавання до оригіналу, проте уніфікує деякі особливості накреслення літер, діакритику та ін.;

- 3) встановлення морфологічного анотування, що передає мовну інформацію про квазіфлексії та квазіоснови, приписує їм відповідні парадигматичні класи.

Добір текстів у межах окремого періоду проводиться за жанрово-стильовим критерієм, тобто сукупність текстів, що належать до одного жанру (стилю), становить окремих сегмент у текстовій репрезентації періоду.

Наукова підготовка писемних пам'яток для проведення морфологічного анотування в історичному корпусі ускладнюється через роботу зі значними за обсягом текстовими матеріалами, їх переведення в електронний вигляд.

Для цього потрібно уніфікувати спосіб передавання писемних пам'яток за встановленими однотипними правилами: набирати (розпізнавати за допомогою спеціальних програм) тексти, максимально наближаючи їх графічний та орфографічний запис до оригіналу, проте без збереження в електронному варіанті окремих графічних знаків (наприклад, діакритичних знаків), неістотних для морфологічного анотування староукраїнських оригіналів. Таке наукове опрацювання писемних джерел за однотипними правилами потребує тривалої підготовки та ускладнює роботу.

Електронні версії вже виданих на сьогодні писемних джерел, хоч й орієнтуються на прийняті правила публікації давніх текстів, проте не завжди вирізняються однотипністю.

Більш того, в українському історичному джерелознавстві усталився підхід до транслітерування текстів, що передбачав передавання окремих літер за допомогою символічного зображення, спеціальне вираження надрядкових літер у рядковому записі, збереження більшості діакритичних знаків (різні типи наголосів, титла), дослідницькі примітки про особливості прочитання складних (з погляду збереження оригіналу та його прочитання) місць. Факсимільні видання як найбільш надійні джерела передавання оригіналу не відповідають вимогам дослідницького пошуку в корпусі.

Для прикладу, в історичному корпусі агіографічних текстів "СКАТ" укладено словопоказки, тобто "списки слівформ з їхніми адресами (номерами аркушів і рядків) у рукописах" [1, с. 40], а життя двічі представлено – у текстовому форматі і форматі редактора Word: "у текстовому файлі передбачено кодування для всіх букв слов'янського алфавіту, текст рукопису поділений на слова і представлений у лінійному вигляді: виносні букви у круглих дужках вставлені на своє місце в слові за смислом, пропущені акцентні знаки. Особливими символами відзначаються кінці рядків, стовпчиків і аркушів рукопису; межі тексту, який вноситься з полів; помилкові написання. Текст у редакторі Word також створюється на основі текстового файлу, у нього вносяться діакритичні знаки, і всі виносні літери займають місце над рядком, за зовнішнім виглядом цей текст наближується до тексту рукопису, відрізняючись від нього тим, що розділений на слова" [1, с. 41].

Лінгвістичне опрацювання підготовленого тексту передбачає встановлення морфологічної інформації (анотування) кожному слову. Під морфологічною інформацією розуміється: частини мовна належність слова, його основні граматичні значення, розрізнення власних і загальних назв.

Принципи створення морфологічного аналізатора викладено в монографії Н. Дарчук [2, с. 58-66], зокрема другий, флексійний етап ґрунтується на інформації про квазіоснови і квазіфлексії, що дає підстави приписати номер парадигматичного класу кожній словозмінюваній лексемі.

Для історичного корпусу така інформація має бути створена: визначено квазіоснови і квазіфлексії (за наведеною термінологією), приписано нові номери парадигматичних класів.

На відміну від сучасного, в історичному корпусі враховуються диференційні граматичні значення староукраїнської мови порівняно із сучасною. Максимальне вираження диференційних значень у давній мові передається за допомогою кодів, які створюються окремо.

Наприклад, якщо взяти за основу морфологічні коди в системі АГАТ [2, с. 356], то для диференційних граматичних значень (двоїни, аориста, імперфекта) давньої мови додаються нові коди. Так, ІН – іменник чоловічого роду, називний відмінок двоїни. Граматичні коди, які використовуються для морфологічного аотування українського тексту, утворюються з двох символів (двозначні), у них перший символ вказує на частини мовну належність, другий – на граматичне значення слівформ.

Опрацювання і представлення отриманих результатів може проводитися в різних напрямках: від тексту (контексту) – до парадигми і, навпаки, від парадигми – до тексту (контексту). Вибір оптимального варіанта залежить від дослідницьких завдань і ступеня наповненості корпусу, опрацьованих текстів за допомогою морфологічного аотування.

Вирішення поставлених завдань передбачає розв'язання окремих проблемних питань щодо 1) варіативності і 2) динаміки мови.

Проблема варіативності вирішується в такий спосіб, що уодноманіюються ті графічні та орфографічні варіанти, які не впливають на фонетику. Наприклад, під час підготовки текстового варіанта уніфікуються різні варіанти написання тієї самої літери.

У парадигмі об'єднуються варіанти написання слівформ, а саме: рядкова літера зі зредукованим, рядкова літера без зредукованого, виносна літера без зредукованого.

Динаміка змін ураховується за допомогою порівняння різних синхронних зрізів, що вибудовують діахронію.

Отже, основні принципи створення історичного корпусу української мови ґрунтуються на вже відомих заходах, деталі – у специфіці роботи з текстовим українським матеріалом певного історичного зрізу, нове – у побудові парадигм за вказаними граматичними параметрами, у перспективі – не тільки морфологічна, але й синтаксична, лексична атрибуція цих текстів.

Список використаних джерел

1. Герд А.С. Корпус древнерусских агиографических текстов СКАТ: современное состояние и перспективы развития / А.С. Герд, И.В. Азарова, Е.Л. Алексеева, Е.С. Иванова // Современные информационные технологии и письменное наследие: от древних рукописей к современным текстам : материалы международной научной конференции. – Ижевск, 2006. – С. 38–42.
2. Дарчук Н.П. Комп'ютерне аотування українського тексту: результати і перспективи : монографія / Наталія Дарчук. – К. : Освіта України, 2013. – 544 с.
3. Демська-Кульчицька О.М. Український національний корпус / О. Демська-Кульчицька [Електронний ресурс]. – Режим доступу: www.rastko.rs.
4. Карпіловська Є.А. Вступ до прикладної лінгвістики: комп'ютерна лінгвістика: підручник. – Донецьк : "ТОВ Юго-Восток ЛТД", 2006. – 188 с.
5. Создан новый корпус среднерусских текстов [Электронный ресурс] // ProScience. – Режим доступа: <http://polit.ru/news>.

Надійшла до редколегії 18.06.13

О. Ника, д-р филол. наук, доц.
КНУ імені Тараса Шевченка, Київ

ИСТОРИЧЕСКИЙ КОРПУС УКРАИНСКОГО ЯЗЫКА: ОСНОВНЫЕ ПРИНЦИПЫ СОЗДАНИЯ

В статье определены основные принципы создания исторического корпуса украинских текстов для решения разноплановых исследовательских задач.

Ключевые слова: корпусная лингвистика, исторический корпус украинских текстов, принципы построения исторического корпуса.

O. Nika, Dr. Hab., Associate Professor
Taras Shevchenko National University of Kyiv, Kyiv

HISTORICAL CORPUS OF THE UKRAINIAN LANGUAGE: BASIC PRINCIPLES OF FORMING

The article defines the basic principles of forming the historical corpus of the Ukrainian language aimed at solving diverse research tasks.

Key words: corpus linguistics, historical corpus of the Ukrainian texts, basic principles of forming the historical corpus.