

УДК 81-13: 81'322.5

**КОРПУС ТЕКСТІВ У ЛІНГВОДИДАКТИЦІ
(НА МАТЕРІАЛІ ОМОНІМІЇ У КОРПУСІ ВЕЛИКОЇ ПРОЗИ ІВАНА
ФРАНКА)**

Соломія Бук

*Львівський національний університет імені Івана Франка,
кафедра загального мовознавства,
вул. Університетська, 1/343, 79001, Львів, Україна,
тел.: (032) 239 47 56*

Узагальнено можливості використання корпусу текстів, зокрема корпусу текстів Івана Франка, для теоретичних та практичних мовознавчих досліджень. Застосування корпусу текстів на заняттях у вищій школі для студентів-філологів подано на прикладі завдань з омонімії.

Ключові слова: омонімія, корпус текстів, морфологічне маркування, лінгводидактика.

Комп'ютеризація гуманітарної сфери, використання нових комп'ютерних версій словників, а також створення зрівноважених електронних корпусів текстів як необхідного джерельного матеріалу для появи якісно нових наукових досліджень стали невід'ємним фактом життя кінця XX – початку XXI ст. Для здійснення різносторонніх мовознавчих досліджень в україністиці зараз усе більше використовують електронні ресурси, зокрема ті, що доступні в Інтернеті: Словник.net (портал української мови та культури) [30], Словники України on-line [31], лінгвістичний портал Mova.info [22], конкорданс Г. Сковороди [46] тощо, а також ті, які перебувають у внутрішньому користуванні наукових установ: Український національний лінгвістичний корпус, Віртуальну лексикографічну лабораторію «Словник української мови» (Український мовно-інформаційний фонд НАНУ) [32], Національний корпус української мови (Інститут української мови НАНУ) [15], корпус текстів Шевчука [23] тощо. Глибокі наукові описи як лексичної, так і граматичної будови мови сьогодні усе більше ґрунтуються на корпусах текстів (КТ).

Допомагаючи створювати паперові та електронні словники, автоматично перекладати, викладати мову тощо, корпусна лінгвістика стала важливим інструментом для промоції національних мов. КТ також має широке застосування і в лінгводидактиці. Правильно організований і маркований корпус також дає викладачеві, студентові та учню доступ до коректно представлених усіх функціональних стилів та підстилів певної мови; спрощує пошук необхідного мовного ілюстративного матеріалу для тем з фонетики та фонології, морфеміки та морфології, лексикології та фразеології, синтаксису, теорії тексту, комунікативної лінгвістики тощо. Причому можна швидко отримати кількісну характеристику мовленнєвих явищ для окремого відрізка тексту, автора, твору, стилю тощо.

Важливість пропонованої теми полягає також у тому, що сучасні методи дослідження мови тільки починають застосовувати для аналізу ідіолектів

українських письменників. Застосування КТ Івана Франка (КТФ) в навчальній аудиторії описано вперше.

За таких обставин важливим завданням університету є підготування висококваліфікованих філологів, що вільно можуть оперувати КТ. Тому введення в навчальні програми завдань із корпусної лінгвістики сприятиме переосмисленню як академічного статусу мовознавчих дисциплін, так і їх соціальної функції. Адже знання понять корпусної лінгвістики становить невід'ємну частину лінгвістичної компетенції студента.

КТ у лінгвістичних дослідженнях. Застосування комп'ютерних технологій у лінгвістиці відкриває значно ширші перспективи для лінгвістичних досліджень у порівнянні з "ручною працею", зокрема карткуванням. Тому корпусний підхід до дослідження мовних явищ активно розвивається фактично в усіх національних мовознавчих науках. Упровадження до лінгвістичного вжитку спеціально підготованого матеріалу – КТ – дає змогу: здійснювати автоматичний пошук будь-якої лексеми, фразеологізму, порівняння у будь-якій формі у всіх чи заданих текстах; фіксувати час першого та останнього використання певної лексеми в обстеженому матеріалі, здійснювати автоматичний пошук твору за будь-яким зареєстрованим у зовнішньому анотуванні параметром (наприклад, твори певного жанру, написані певного року (періоду), підписані певним псевдонімом, написані певною мовою, наявність чи відсутність епіграфа тощо); отримання конкордансів, що подають до кожного реєстрового слова (чи словоформи) усі або вибірккові контексти його вживання, достатні для ідентифікації його лексичного значення; автоматичне укладання частотних словників будь-якого конкретного чи групи творів тощо. КТ успішно застосовують у лексикографії, структурній і когнітивній, соціо- і психо-, етно- і судовій лінгвістиці, культурології, політології, адже завдяки йому можна отримати зразки певного граматичного явища, синтаксичної конструкції чи семантичної категорії разом із їх кількісною характеристикою.

КТ як організація мовного матеріалу спеціальним чином в електронному вигляді, що дає змогу автоматизувати лінгвістичні дослідження, має фундаментальне значення для лінгвістичної теорії та практики.

КТ у лінгводидактиці. У світовій лінгвістиці є плідна практика застосування КТ у навчальному процесі. Існує навіть окремий тип КТ – учнівський, що складається з текстів, написаних учнями чи студентами, які вивчають певну мову [7; 37; 39–41]. Лінгводидактичний потенціал корпусу текстів Івана Франка у викладанні української мови як іноземної вже було описано [4]. Проте КТ знайшов своє місце у викладанні не лише іноземної, а й рідної мови [16; 26; 36]. Часто КТ використовують для написання завдань і вправ, організації самостійної дослідницької роботи студентів та школярів, верифікації даних підручників і словників тощо [16, с. 336; 19; 38].

Існують спеціальні розвідки та статті про КТ для педагогів, що полегшують їм роботу. Класик корпусної лінгвістики Дж. Сінклар став одним з редакторів збірника праць "Як використовувати корпуси у викладанні мови" [42], М. Маккарті зосереджує увагу на тому, як навчити викладачів користуватися корпусом, серед іншого, отримувати й інтерпретувати корпусну інформацію [45], С. Крилов у статті про специфіку пошуку в КТ звертає увагу на те, що важливо знати про це викладачеві-словеснику, зокрема, він виділяє "пошук точних форм" і "лексико-граматичний пошук" [21]. На сайті російського національного корпусу є навіть освітній портал Studiorum, мета якого допомогти викладачам і вчителям засвоїти використання КТ у їхній повсякденній роботі, стати інструментом для створення

нових методик навчання, нових словників, довідників і посібників. Ця інформаційно-довідкова система містить розділи “Допомога початковому користувачеві”, “Методичні розробки”, “Корпусні методи в русистиці”, “Семінар у комп’ютерному класі”, “Питання до корпусу” і т. п. [26].

Часто студентам не лише демонструють можливості та переваги КТ як потужного лінгвістичного ресурсу, а й пропонують самим брати участь у їх створенні та маркуванні, як, наприклад, у польському, російському, словацькому, чеському [18; 24; 35; 44].

Омонімія в корпусі текстів. Омонімія – цікаве лінгвістичне явище, що відображає двобічність знака. Це збіг у формі різних за значенням слів, тому у словниках слова-омоніми “подаються в окремих статтях з позначкою біля слова арабськими цифрами. Напр.: мати 1 – іменник, мати 2 – дієслово ...” [29, Т. I, с. VIII].

Розрізнення омонімії в КТ є важливою і самодостатньою проблемою, оскільки коректне її вирішення впливає на якість комп’ютерного перекладу, синтаксичного аналізу та на інші типи автоматичного опрацювання усного та письмового мовлення. Особливо це стосується слов’янських мов, у яких багата морфологія, тому вони мають велику кількість слів з різною граматичною інтерпретацією: наприклад, пол. *miał* (мав і м’яв), рос. *разогнать* (у різні боки і швидкий рух на дорозі), чеське *vinu* (який стосується вина і який має вину) тощо. Людина безпомилково визначає потрібне значення з контексту, тоді як комп’ютер треба “навчити” робити це, оскільки він не володіє знаннями про базові складові комунікації (пресупозиції, інтенції мовців і т. д.).

У національних загальномовних КТ розрізнення омонімії здійснюють на стохастичних засадах із аналізом правого та лівого контексту одиниці. Для цього застосовують автоматичні алгоритми та програми [1; 2; 11; 12, с. 63–68, 72–79; 17, с. 133–134; 20, с. 309–329, 338–343; 27, с. 447; 28, с. 66–72; 34] із подальшою “ручною” верифікацією принаймні частини результатів [24; 35; 43; 44; 47]. Також КТ використовують як дослідницьку базу для омонімії [1; 2; 10, 34]. На основі КТ укладено словники омонімів [13; 25; 34].

Омонімія в корпусі текстів Івана Франка. КТФ спроектовано відповідно до усіх світових вимог до такого типу конструкту: машиночитаності, репрезентативності, маркованості, автентичності [3; 14]. На першому етапі створено КТ великої прози письменника, що охоплює 9 українськомовних творів: “Петрії й Довбушки”, “Воа constrictor”, “Борислав сміється”, “Захар Беркут”, “Не спитавши броду”, “Для домашнього огнища”, “Основи суспільності”, “Перехресні стежки” (а також онлайн конкорданс), “Великий шум”. До кожного з них укладено частотний словник (ЧС).

У КТФ здійснено зовнішню та внутрішню анотацію текстів [6]. Важливим елементом у морфологічному маркуванні КТФ було розрізнення омонімії перед тим, як лематизувати (зводити до початкової, словникової форми) словоформи. Це було необхідно не лише для вивчення особливостей ідіолекту письменника, а й для загальномовних теоретичних, практичних, дидактичних досліджень. Омонімію розрізнено шляхом контекстного аналізу оточення мовної одиниці, оскільки застосування автоматичних програм, розроблених для української мови, видавалося не зовсім коректним, бо вони зорієнтовані на літературний її варіант, а мова І. Франка представляє її західний варіант зламу XIX–XX ст.

У КТФ виявлено майже всі типи омонімів, зокрема такі, що не зафіксовані в сучасному словнику омонімів української мови [13]: *адресу* (ім. ж. р. 3. в. і ч. р. Р. в.), *академічний* (загальна назва) і *Академічний* (власна назва: пляц Академічний (“Для

домашнього огнища”)), *антихристові* (ім. ч. р. Д. в. та присв. прикм.), *апостола* (особа) і *Апостола* (книга), *барвінкові* (ім. ч. р. Д. в. та прикм.) та ін. Частина з них є потенційними, проте для правильної лематизації їх треба було обов’язково виявити.

Підручники сучасної української мови представляють здебільшого онтологію, походження та місце омонімії в системі мови загалом або у системі конкретної мови зокрема. За допомогою КТ дослідник може побачити це явище у світлі функціонування мови, тобто у мовленні. Тут стануть очевидними такі чинники, як співвідношення частотності омонімів між собою, їх дистрибуція, поява в площинах прямої чи авторської мови творів, особливості взаємодії зі знаками пунктуації (!, ? –) синтаксичними явищами (зворотами, вставними елементами, звертаннями) та ін. Враховуючи кількісні характеристики функціонування цієї категорії слів у текстах, услід за О. Невзоровою [25], в КТФ зафіксовано регулярні омоніми (які трапляються в ролі різних частин мови приблизно порівну), нерегулярні омоніми (які трапляються переважно в ролі однієї частини мови) та потенційні омоніми (які трапляються в тексті лише в ролі однієї частини мови).

Квантитативні дані можна знайти вже в “Частотному словнику сучасної української художньої прози”, де кожний омонім має окрему словникову статтю з прецизійною інформацією про кількість його вживань у різних формах у прямій чи авторській мові обстежених текстів [33]. Тому можна запропонувати студентам вправи на його основі.

Завдання 1.

Подано перелік омонімів: *а* (спол., частка, вигук), *мати* (дієсл. та ім.), *коса* (волосся, інструмент, вузька смуга суходолу в морі), ...

1) З’ясуйте, яке слово з кожного омонімічного ряду виявляє більшу активність у художній прозі.

2) Яких слів немає у словнику? Як називаються омоніми, яких немає в аналізованому тексті?

3) Що таке регулярні та нерегулярні омоніми?

У результаті студент має дійти висновку, що омоніми *а* нерегулярні: лексема *а* (спол.) із частотністю 6251 в десятки разів активніша у тексті, ніж *а* (частка) із частотністю 24, і *а* (вигук) із частотністю 21. Окрім того, студент зверне увагу, що *а* (частка) жодного разу не трапилась в авторському мовленні, і *а* (вигук) також властивий винятково прямій мові, бо в словах автора її частота 1 [33, с. 25].

Цю вправу із залученням ЧС художньої прози в університетській аудиторії практикувати складно, оскільки потрібно забезпечити всю групу зазначеним словником, а електронної його версії, на жаль, немає. Значно ефективніше сформулювати ці ж завдання для КТФ, який кожен студент може переписати собі на комп’ютер, або для роману І. Франка “Перехресні стежки”, електронний конкорданс якого є у вільному доступі в Інтернеті [9]. Ще однією перевагою роботи з електронним КТ є те, що результати пошуку лексеми він отримує у всіх контекстах її вживання, а значить, має ширший матеріал для аналізу (див. табл. 1, 2).

Табл. 1

Конкорданс лексеми МАТИ (ім.)

Форма контексту – +/- 7 слів

Адже його батько не мав тої слабості, **мати** не мала, він сам парубком був

долівку і довго молилася, плачучи гарячими слізьми.	“Мати	здоровісінський, Божа, -- молилася я, -- дай, щоб
і скажи виразно: “Во грісіх роди мя	мати	моя”, жию в свинстві і роблю щохвиля
себе самих, коли вірите, що та інша	мати	полюбить і пригорне вас як своїх рідних?

Табл. 2

Конкорданс лексеми МАТИ (дієсл.)
Форма контексту – +/- 7 слів

обороні, не тільки увільнені трибуналом, але надто	мали	надію в дорозі цивільного процесу виграти те
та сей збув їх коротко, навчив, що	мають	робити далі, і вийшов із темнуватого судового
отсе тільки що перенісся. Від нинішньої оборони	мало	залежати його дальше поводження на новім ґрунті,
але, держачися старого правила “ <i>aequam servare mentem</i> ”,	мав	вид не то байдужно-спокійний, не то занятий
одягнений у чорний витертий сурдут. Д-р Рафалович	мав	бистре око і добру пам'ять, але не
	– Мате	тут яку посаду?
	...	
я вжити інших, правних способів. Ось тут	маю	честь вручити пану старості донесення про віче,

На матеріалі наведеного конкордансу студент підсумує, що омоніми *мати* нерегулярні: лексема *мати* (дієсл.) із частотністю 396 суттєво активніша в тексті, ніж *мати* (ім.) із частотністю 4.

Після спостереження над поведінкою омонімів наступного омонімічного ланцюжка студент виявить, що у романі лексему *коса* два рази вжито у значенні “волосся”, один раз – у значенні “знаряддя” і жодного разу у значенні “вузька смуга суходолу в морі”. Тобто останній приклад – це зразок потенційного омоніма.

Як показує практика, розрізнення омонімії дієслово/іменник, іменник/прикметник, іменник/прийменник у студентів не викликає труднощів. Складніше їм із омонімами типу займенник/сполучник/частка, прислівник/сполучник, ... У цьому ключі можна запропонувати деякі завдання.

Завдання 2.

1) Опрацюйте тлумачення слова *що* у Словнику української мови (11 тт.). Якими частинами мови воно може бути? Який основний критерій розрізнення цих частин мови?

2) На заданому відрізьку КТФ диференціюйте лексеми *що* займенник/сполучник/частка шляхом додавання до слова позначки “#” таким чином:

що (займенник) – залишайте без змін,

що (сполучник) – замініть на що#,

що (частка) – замініть на що##

3) Звірте свій результат із відповідним відрізком КТФ, де омонімія розрізнена. Які труднощі виникали під час роботи?

4) У якій частині мови *що* має найбільшу частотність? Чи можна було би це передбачити, зважаючи на тематику твору? Який кількісний показник, на вашу думку, не залежить від теми твору.

5) Навіщо розрізняти омонімію в КТ? Які переваги має КТ із знятою омонімією над КТ із незнятою омонімією?

Для цього завдання в навчальних умовах можна запропонувати будь-який уривок з КТФ. Наприклад, у романі “Основи суспільності” *що* (займ.) має частоту 768, *що* (спол.) – 690, *що* (частка) – 83. Порівняння отриманих кількісних характеристик з ЧС інших творів І. Франка дає підстави стверджувати, що вони (характеристики) від теми твору не залежать [5].

Завдання 3.

1) У КТФ проаналізуйте певну кількість контекстів (наприклад, 200) лексеми *як*.

2) Підберіть до заданої лексеми *як* синоніми. Яку синтаксичну роль вона виконує? Про що це свідчить? Зробіть висновок про її частиномовну належність.

3) Порівняйте свої висновки із Словником української мови в 11 томах.

4) Зіставте свої висновки із відповідним відрізком КТФ, де омонімія розрізнена. Які труднощі виникали під час роботи?

Тут студент повинен здійснити класифікаційну роботу, виявити закономірності, які би дали статусну характеристику омонімічній частині мови [10] і прийти до висновку, що лексема *як* може бути прислівником, сполучником, часткою. У першому випадку вона має синтаксичну роль обставини, у двох останніх – не виступає жодним членом речення. Слід зазначити, що висновок, зроблений у проблемній ситуації самостійно, вимагає у студента значно більшої концентрації думки, набагато краще запам’ятовується у порівнянні з тим випадком, коли відповідь попередньо взято зі словника чи підручника.

Завдання 4.

1) Подано омонім *так* прислівник/сполучник/частка. У якій частині мови, на вашу думку, він має найбільшу частотність? Чи суттєво відрізняється його частотність у цих частинах мови? Від чого це може залежати?

2) Результат звірте із КТФ або частотними словниками відповідних романів.

3) Чи правильними були Ваші міркування? Що викликало найбільші труднощі?

Виконання цього завдання потребує інтроспекції, застосування мовної інтуїції майбутнього лінгвіста. ЧС повісті “Воа constrictor”, наприклад, показує такі дані: так (присл.) 61, так (спол.) 7, так (частка) 20 [8]. Ця інформація корелює з іншими ЧС творів І. Франка. Для вирішення вправи студент повинен мати вже неабияке мовне чуття, вміння логічно мислити і моделювати текстові ситуації. У правильності чи хибності своїх умовиводів можна переконатися лише на КТ, оскільки інформацію про частотність функціонування слів у тлумачних словниках української мови поки

що не подають, хоча вона дуже суттєва для розуміння місця мовної одиниці у системі. До речі, у словниках інших мов (англійської, польської) її вже залучають. Завдання розвиває мовну інтуїцію майбутнього лінгвіста.

Завдання 5.

- 1) Які типи омонімів та суміжні з омонімією явища ви знаєте?
- 2) Поміркуйте, чи всі типи потрібно розрізняти в КТ, чому?
- 3) Для яких типів зробити це швидше? Чому?

Завдання створює евристичну ситуацію на занятті. Для правильних відповідей студент мусить уже мати достатні знання як з теорії мови, так і з корпусної лінгвістики, добре орієнтуватися в особливостях морфологічного маркування КТ.

Значна кількість сучасних студентів люблять працювати в Інтернеті, тому запропоновані вправи з дослідницьким елементом мають шанс на успіх.

Оскільки КТФ має морфологічне маркування, його можна залучати як додаткове джерело матеріалу до складання граматичних завдань з кожної конкретної теми, починаючи від граматичного значення і закінчуючи частинами мови і граматичними категоріями. У статті наведено деяке його практичне застосування у лінгводидактиці на прикладі омонімії.

Висновки та перспективи дослідження. Отже, омонімію у навчальних закладах варто викладати не лише з погляду теорії (дефініція, приклади, вправи), а й на практиці: звернути увагу студента власне на прикладне значення розрізнення омонімії в тексті. КТ, зокрема КТФ, дає змогу продемонструвати поведінку мовних одиниць, зокрема й омонімів, у живому тексті, їх поширеність, частотність, залежність функціонування від типу тексту (пряма, непряма мова) тощо.

Запропоновані завдання допоможуть студентам глибше усвідомити та зрозуміти сутність омонімії як явища, значення її розрізнення для автоматичного опрацювання тексту, присвоєння морфологічних категорій різним частинам мови тощо. Отже, КТФ допомагає організувати тренувальну, дослідницьку, а також самостійну роботу студентам.

Використання КТ у майбутньому дасть уточнення до лексикографічного опису омонімів, у тому числі доповнення до СУМу. Застосування комп'ютерних технологій у лінгвістиці вищої школи видається природним і необхідним етапом розвитку сучасної філологічної освіти. КТ – могутній інструмент, але без детального вивчення та глибокого розуміння його будови можна втратити значну частину його потенціалу. КТ – потужний ресурс у лінгводидактиці, проте, звичайно, не самодостатній. Спілкування з комп'ютером і КТ у жодному випадку не може замінити рекомендацій викладачів та дискусій між студентами.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Бугаков О.В.* Аналіз граматичної омонімії применників у мові й у тексті / О. В. Бугаков // *Мовознавство.* – 2004. – № 5–6. – С. 87–98.
2. *Бугаков О.В.* Усунення граматичної омонімії применників в українському тексті / О. В. Бугаков // *Мовні і концептуальні картини світу: Збірник наукових праць.* – Вип. 11. Кн. 1. – К., 2004. – С. 50–53.

3. Бук С. Корпус текстів Івана Франка: спроба визначення основних параметрів / С. Бук // Прикладна лінгвістика та лінгвістичні технології: MegaLing-2006 : 36. наук. пр. – К., 2007. – С. 72–82.
4. Бук С. Лінгводидактичний потенціал корпусу текстів Івана Франка у викладанні української мови як іноземної / С. Бук // Теорія і практика викладання української мови як іноземної. – Львів, 2010. – Вип. 5. – С. 70–74.
5. Бук С. Статистичні характеристики роману Івана Франка “Основи суспільності” (на основі частотного словника твору) / С. Бук // Вісник : Проблеми української термінології. – Львів, 2010. – № 676. – С. 90–93.
6. Бук С. Структурне анотування у корпусі текстів (на прикладі прози Івана Франка) / С. Бук // Українська мова. – 2009. – № 3. – С. 59–71.
7. Бук С. Учнівські корпуси в методиці викладання іноземної мови / С. Бук // Теорія і практика викладання української мови як іноземної. – Львів, 2007. – Вип. 2. – С. 19–23.
8. Бук С. Частотний словник повісті Івана Франка “Воа constrictor” (редакція 1884 р.) / С. Бук // Стежками Франкового тексту: комунікативні, лінгвосоціотичні, когнітивні, лінгвостатистичні та лексичні виміри прози. Львів, 2012 (*у друці*).
9. Бук С. Он-лайн конкорданс роману Івана Франка “Перехресні стежки” [Електронний ресурс] / С. Бук, А. Ровенчак. – Доступно з : <<http://www.ktf.franko.lviv.ua/~andrij/science/Franko/concordance.html>>. – [Цит. 16.12.2011].
10. Вепрева И. Т. Неснятая омонимия в текстах Национального корпуса русского языка как исследовательская база для студентов-филологов на практических занятиях по морфологии / И. Т. Вепрева // Национальный корпус русского языка и проблемы гуманитарного образования: материалы международной научной конференции. Москва 19–20 апреля 2007. – Москва, 2007. – С. 17–19.
11. Грязнухина Т. О. Шляхи усунення граматичної омонімії в системі автоматичного морфологічного аналізу наукового тексту / Т. О. Грязнухина, Л. Г. Братищенко, Н. П. Дарчук та ін. // Мовознавство. – 1989. – № 5. – С. 3–12.
12. Дарчук Н. П. Комп’ютерна лінгвістика (автоматичне опрацювання тексту) : підручник / Н. П. Дарчук. – К., 2008.
13. Демська О. М. Словник омонімів української мови / О. М. Демська, І. М. Кульчицький. – Львів, 1996.
14. Демська О. Текстовий корпус: ідея іншої форми / О. Демська. – К., 2011.
15. Демська-Кульчицька О. Основи національного корпусу української мови / О. Демська. – К., 2005.
16. Добрушина Н. Р. Корпусные методики обучения русскому языку / Н. Добрушина // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. – Санкт-Петербург, 2009. – С. 335–352.
17. Карпіловська Є. А. Вступ до комп’ютерної лінгвістики / Є. А. Карпіловська. – Донецьк, 2003.
18. Карчова А. Морфологічна анотація текстів Словацького національного корпусу [Електронний ресурс] / А. Карчова, М. Шимкова. – Доступно з : <korpus.juls.savba.sk/.../block1/.../2006-karcova-simkova-kyjev.pdf>. – [Цит. 16.12.2011].
19. Кольцова Л. М. Ресурси національного корпусу русского языка как основа подготовки научных и учебно-методических работ по современному русскому языку / Л. М. Кольцова, Ж. В. Грачев // Национальный корпус русского языка и

проблеми гуманитарного образования. Материалы международной научной конференции. Москва 19–20 апреля 2007. – Москва, 2007. – С. 37–38.

20. Корпусна лінгвістика / В. А. Широков, О. В. Бугаков, Т. О. Грязнухіна та ін. – К., 2005.

21. Крылов С. А. Специфика поиска в Национальном корпусе русского языка: что важно знать о ней преподавателю-словеснику [Электронный ресурс] / С. А. Крылов. – Доступно з: <<http://sergej-krylov.livejournal.com/16268.html>>. – [Цит. 16.12.2011].

22. Лінгвістичний портал Mova.info [Электронный ресурс]. – Доступно з: <<http://www.mova.info>>. – [Цит. 16.12.2011].

23. Монахова Т. В. Мова Валерія Шевчука: ключові концепти, корпус, тезаурус / Т. В. Монахова : Автореф. дис. ... канд. філол. наук. – К., 2008.

24. Национальный корпус русского языка [Электронный ресурс]. – Доступно з: <<http://www.ruscorpora.ru>>. – [Цит. 16.12.2011].

25. Невзорова О. А. К разработке нового словаря функциональных омонимов на основе национального корпуса русского языка / О. А. Невзорова, Ю. В. Зинькина // Национальный корпус русского языка и проблемы гуманитарного образования. Материалы международной научной конференции. Москва 19–20 апреля 2007. – Москва, 2007. – С. 61–63.

26. Образовательный портал Национального корпуса русского языка [Электронный ресурс]. – Доступно з: <<http://studio.ruscorpora.ru>>. – [Цит. 16.12.2011].

27. Резникова Т. И. Славянская корпусная лингвистика: современное состояние ресурсов / Т. И. Резникова // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. – Санкт-Петербург, 2009. – С. 402–461.

28. Романюк Ю. В. Прикладна морфологія: навчально-методичний посібник / Ю. В. Романюк. – Черкаси, 2010.

29. Словник української мови : в 11-ти томах. – К., 1970–1980.

30. Словник.net: портал української мови та культури [Электронный ресурс]. – Доступно з: <<http://slovyk.net>>. – [Цит. 16.12.2011].

31. Словники України on-line [Электронный ресурс]. – Доступно з: <<http://lcorer.ulif.org.ua/dictua>>. – [Цит. 16.12.2011].

32. Український мовно-інформаційний фонд НАНУ [Электронный ресурс]. – Доступно з: <<http://www.ulif.org.ua/UMIF/projects>>. – [Цит. 16.12.2011].

33. Частотний словник сучасної української художньої прози : в 2-х томах / За ред. В. С. Перебийніс – К., 1981.

34. Шипнівська О. О. Формування лінгвістичної бази даних граматичних омонімів на матеріалі Українського національного лінгвістичного корпусу / О. О. Шипнівська, С. В. Стариков // Лексикограф. бюл. : зб. наук. праць / Відп. ред. д. філол. н., проф. В. Німчук. – К., 2006. – Вип. 14. – С. 188–194.

35. Český národní korpus [Электронный ресурс]. – Доступно з: <<http://ucnk.ff.cuni.cz>>. – [Цит. 16.12.2011].

36. Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods / Ed. by S. Braun, K. Kohn, J. Mukherjee. – Frankfurt am Main, 2006.

37. Corpus-Based Approaches to English Language Teaching / Ed. by M. C. Campoy, B. Bellés-Fortunato, M. Gea-Valor. – London ; New York, 2009.

38. From corpus to classroom: language use and language teaching / A. O’Keeffe, M. McCarthy, R. Carter. – Cambridge University Press, 2007.

39. *Granger S.* Computer learner corpus research: current status and future prospects / S. Granger // *Applied Corpus Linguistics: A Multidimensional Perspective* / Ed. by U. Connor and T. Upton. – Amsterdam ; Atlanta, 2004. – P. 123–145.

40. *Granger S.* The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation / S. Granger // *Corpora and Language Teaching* / Ed. by K. Aijmer. – Amsterdam ; Philadelphia, 2009. – P. 13–32.

41. *Granger S., Paquot M.* Language for specific purposes learner corpora / S. Granger, M. Paquot // *Language for specific purposes. The encyclopedia of applied linguistics* / Ed. by T. Upton & C. Chapelle. – Oxford, 2011.

42. *How to use corpora in language teaching* / ed. by J. McH. Sinclair (ed.) – Amsterdam [u.a.], 2004.

43. Hrvatski nacionalni korpus [Електронний ресурс]. – Доступно з : <<http://www.hnk.ffzg.hr>>. – [Цит. 16.12.2011].

44. Korpus języka polskiego IPI PAN [Електронний ресурс]. – Доступно з : <<http://korpus.pl>>. – [Цит. 16.12.2011].

45. *McCarthy M.* Accessing and interpreting corpus information in the teacher education context. – Cambridge, 2008.

46. Online Concordance to the Complete Works of Hryhorii Skovoroda [Електронний ресурс]. – Доступно з : <<http://www.arts.ualberta.ca/~ukr/skovoroda/NEW/index.php>>. – [Цит. 16.12.2011].

47. Slovenský národný korpus [Електронний ресурс]. – Доступно з : <<http://korpus.juls.savba.sk>>. – [Цит. 16.12.2011].

Стаття надійшла до редакції 15.10. 2011

доопрацьована 15. 11. 2011

прийнята до друку 10. 01. 2012

TEXT CORPUS IN LINGUODIDACTICS
(On the material of homonymy in the Corpus of the long prose fiction
by Ivan Franko)

Solomija Buk

*Ivan Franko National University of Lviv,
Department for General Linguistics,
Universytetska St., 1, room 343, 79000, Lviv, Ukraine,
tel.: (032) 239 47 56, e-mail: solomija@gmail.com*

In the article, the possibility of text corpus (in particular Ivan Franko text corpus) usage is summarized. It can be used in both theoretical and practical linguistic studies. The text corpus application in the university lectures is demonstrated on the example of homonymy.

Key words: homonymy, text corpus, morphological mark-up, linguodidactics.

КОРПУС ТЕКСТОВ В ЛИНГВОДИДАКТИКЕ
(На матеріалі омонимії в корпусі великої прози Івана Франка)

Соломія Бук

*Львівський національний університет імені Івана Франка,
кафедра загального мовознавства,
ул. Університетська, 1, к. 343, 79000, Львів, Україна,
тел.: (032) 239 47 56, e-mail: solomija@gmail.com*

В статті обобщені можливості використання корпусу текстів, в частині корпусу текстів Івана Франка, для теоретических і практичних досліджень в мовознавстві. Застосування корпусу текстів на заняттях в вищій школі для студентів-філологів продемонстровано на прикладі вправ по омонимії.

Ключевые слова: омонимия, корпус текстов, морфологическая разметка, лингводидактика.