

ОРГКОМІТЕТ КОНФЕРЕНЦІЇ / ORGANIZING COMMITTEE

Андрущенко В. П. – доктор філософських наук, академік НАПН України, член-кореспондент НАН України, професор, ректор Національного педагогічного університету імені М. П. Драгоманова.

Торбін Г. М. – доктор фізико-математичних наук, професор, проректор з наукової роботи Національного педагогічного університету імені М. П. Драгоманова.

Кудін А. П. – доктор фізико-математичних наук, проректор з дистанційної освіти та інноваційних технологій навчання Національного педагогічного університету імені М. П. Драгоманова.

Зернецька А. А. – доктор філологічних наук, доцент, декан факультету іноземної філології Національного педагогічного університету імені М. П. Драгоманова.

Корольова А. В. – доктор філологічних наук, професор, завідувач кафедри загального мовознавства і германістики Національного педагогічного університету імені М. П. Драгоманова.

Леміш Н. Є. – доктор філологічних наук, доцент, завідувач кафедри філософії мови, порівняльного мовознавства та перекладу Національного педагогічного університету імені М. П. Драгоманова.

Карпіловська Є. А. – доктор філологічних наук, професор, завідувач відділу структурно-математичної лінгвістики Інституту української мови Національної Академії наук України.

Дарчук Н. П. – доктор філологічних наук, доцент, професор кафедри української мови та прикладної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка.

Біскуб І. П. – доктор філологічних наук, професор, завідувач кафедри прикладної лінгвістики Східноєвропейського національного університету імені Лесі Українки.

Зубань О. М. – кандидат філологічних наук, доцент, доцент кафедри української мови та прикладної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка.

Жуковська В. В. – кандидат філологічних наук, доцент, завідувач кафедри міжкультурної комунікації та прикладної лінгвістики Житомирського державного університету імені Івана Франка.

Dr. Judita Giparaitė – Ph.D., Associate Professor, Lithuanian University of Educational Sciences, Lithuania.

Dr. James Swartz – Associate Professor, University of Illinois, Chicago, the USA.

КОРПУСНА ЛІНГВІСТИКА ТА ЇЇ МАЙБУТНЄ



Шановна наукова громадо! Питання обговорення ролі й місця корпусних досліджень у сучасній лінгвістичній науці та інших сферах знання, зокрема у лінгводидактиці, у перекладознавстві тощо залишається актуальним, і сьогоднішня дискусія, в якій бере участь зацікавлена наукова спільнота, є тому підтвердженням.

Яким буде шлях майбутньої лінгвістики?, покаже час, але вже з упевненістю можна говорити, що саме корпусний підхід “очолить” її передові позиції, який уже, за словами В. О. Плуґяна, змінив “уявлення лінгвістів про мову”.

Проте заради справедливості маємо сказати, що не всі лінгвісти погоджуються з такою думкою і навіть Н. Хомський у 2004 р. в одному з інтерв'ю сказав, що “Corpus linguistics doesn't mean anything”, а корпусний підхід зводиться до тривіального спостереження за великим обсягом даних, що “не є методом наукового пізнання і не може забезпечити ані успішне вирішення пізнавальних і практичних проблем, ані накопичення нових знань”. Окрім того, радикально налаштовані лінгвісти вважають, що корпус є лише заміною картотеки. Але чи є коректними такі аналогії? Відповідь однозначна, що – Ні! Корпус (з розміткою, анутованням тощо) аж ніяк не можна порівнювати з механічною вибіркою матеріалу, якою є картотека, як за масштабами, так і за дослідницькими можливостями.

Наразі скептиків корпусної лінгвістики залишається ще чимало, серед яких і Ю. Д. Апресян, який, хоча і не заперечує значимість корпусних

технологій, але при цьому застерігає, що “надмірне захоплення корпусними даними призводить до фальсифікації результатів і зловживання кількісними даними”.

Тим не менш такі критичні погляди не зупинили розвиток корпусної лінгвістики, палким прихильником якої В. О. Плунгян та його однодумці, який разом зі своїми колегами переконаний у тому, що сучасне дослідження мови не може здійснюватися поза лінгвістикою корпусів. І майбутнє корпусної лінгвістики вони пов’язують із розробкою “корпусних словників”, “корпусних граматик” нового покоління тощо.

Корпусна лінгвістика має вже потужні школи за кордоном: у Ланкастері, Бірмінгемі, Ноттінгемі, Аризоні тощо, а також власні спеціальні академічні журнали (“International Journal of Corpus Linguistics”, “Corpus Linguistics and Linguistic Theory” та ін.).

У сучасному українському мовознавстві також намітилися основні “теоретичні й практичні передумови для становлення й розвитку корпусної лінгвістики”, за словами О. Демської-Кульчицької.

В Україні корпусний підхід активно розробляють як в академічних установах, так і у вищих навчальних закладах, зокрема в Інституті української мови Національної Академії наук України, Інституті мовознавства ім. О. О. Потебні Національної Академії наук України, у Київському національному університеті імені Тараса Шевченка, у Національному педагогічному університеті імені М. П. Драгоманова, у Київському національному лінгвістичному університеті, у Житомирському державному університеті імені Івана Франка, у Львівському національному університеті імені Івана Франка, у Національному університеті “Острозька академія”, у Національному технічному університеті України “Київський політехнічний інститут імені Ігоря Сікорського”, у Донецькому національному університеті імені Василя Стуса (м. Вінниця), у Національному університеті “Львівська політехніка”, у Східноєвропейському національному університеті імені Лесі Українки, в Одеському національному університеті імені І. І. Мечникова та ін.

Корпусна лінгвістика в Україні вже має кілька корпусів писемного й усного українського мовлення (сьогодні про це колеги скажуть окремо), які становлять науковий інтерес і можуть слугувати дослідницьким матеріалом. Водночас існує ще багато проблем, а головне, перспектив у цій сфері, основні з яких плануємо обговорити під час наукової дискусії і намітити плани для їхнього подальшого вирішення, сподіваючись на велике майбутнє цієї сфери наукового знання. Дякую за увагу!

НАУКОВА ДИСКУСІЯ

Академічна й навчально-методична інтеграція корпусних досліджень



Карпіловська Євгенія Анатоліївна – завідувач відділу структурно-математичної лінгвістики Інституту української мови Національної Академії наук України, доктор філологічних наук, професор

1. Корпус як збалансована, формалізована й структурована вибірка текстів з відповідними засобами роботи з нею. Параметри формування такої вибірки. Типи корпусів за призначенням.

2. Корпус як будь-яка комп'ютерна лінгвістична система має свій формат і сферу застосування. Чого можна і чого не слід чекати від корпусу. Досвід сучасних національних корпусів мов.

3. Корпус не тільки не заперечує інші інструменти опрацювання мовної інформації, а, навпаки, має взаємодіяти з ними: електронні картотеки, служби культури мови, служби відстежування мовних змін, системи орфографічного та стилістичного контролю текстів.

4. Місце корпусу в системі координат: мовна діяльність суспільства (система мови у мовній практиці) – текст – словник – граматики – картотека.

Технологія розроблення корпусів текстів для навчальних цілей



Дарчук Наталія Петрівна – професор кафедри української мови та прикладної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка, доктор філологічних наук, доцент

Корпус текстів – це інформаційний лінгвістичний ресурс, в якому користувач може автоматично створити вибірку мовних одиниць та контекстів їхнього вживання на матеріалі текстів певної мови, представлених в електронному варіанті. До створення корпусів ставляться такі вимоги: 1) відцифрованість текстів; 2) наявність певної процедури відбору текстів до Корпусу; 3) єдина методика представлення відомостей про тексти та їхні одиниці; 4) кінечний розмір Корпусу текстів.

В Україні створено кілька корпусів української мови, але в них не передбачено створення підкорусів навчальних текстів. Навіть до відомих європейських національних корпусів включено переважно тексти художньої літератури та публіцистики.

Необхідність створення корпусу навчальних українсько-, англійсько-, німецько-, італійськомовних тощо текстів покликана практичними потребами вдосконалення методики викладання цих мов у вищій, середній школі тощо. А при укладанні сучасних підручників, посібників допоможуть паралельні українсько-англійські, -німецькі, -італійські тощо корпуси навчальних текстів.

Для здійснення цієї роботи попередньо варто проаналізувати українські, англійські, німецькі тощо підручники для вищів (шкіл) з різних дисциплін (фізика, географія, математика, історія тощо); з'ясувати різні види інформації (теоретичні теми, аспекти “головне”, “нові поняття, терміни” тощо); деталізувати окремі теми для проведення дискусій; висновків, матеріалу для повторення; запитання і завдання з теми; вправи; історичні відомості; основні події, дати; словники термінів.

Паралельний українсько-англійський (-німецький, -італійський тощо) корпус навчальних текстів дозволить відбирати приклади вживання слів, словосполучень, речень у текстах досліджуваної мови; демонструвати на конкретних прикладах способи зняття двомовної омонімії; укладати автоматично навчальні словники з різних предметних галузей; створювати українсько-англійські термінологічні словники з різних предметних галузей.

Для проведення наукових досліджень у зіставному плані такий корпус дозволить: автоматично виділяти групи слів певної словозміни або словотвору; знаходити слова з певними граматичними характеристиками; проводити зіставний аналіз двох мов на синтаксичному рівні; автоматично виявляти і систематизувати помилки у письмових роботах студентів і школярів тощо.

Технологія роботи з корпусами у процесі вивчення іноземних мов



Жуковська Вікторія Вікторівна – завідувач кафедри міжкультурної комунікації та прикладної лінгвістики Житомирського державного університету імені Івана Франка, кандидат філологічних наук, доцент

Автор презентує досвід застосування технологій корпусної лінгвістики у професійній підготовці вчителів англійської мови в навчально-науковому інституті іноземної філології Житомирського державного університету імені Івана Франка. Буде продемонстровано можливості використання корпусів англійської мови 1) у формуванні мовно-комунікативної компетентності майбутніх фахівців, 2) у підготовці начально-методичної продукції та 3) у проведенні різноаспектних наукових досліджень. Особливу увагу буде приділено особливостям формування професійно-значимих умінь майбутніх учителів в аспекті використання корпусних технологій на заняттях із англійської мови. Відбудеться також презентація наукових і навчально-методичних праць із корпусної лінгвістики викладачів кафедри.

МАЙСТЕР-КЛАС

КОРПУС ТЕКСТІВ ЯК ОСОБЛИВИЙ ТИП ІНФОРМАЦІЙНО-ПОШУКОВОЇ, НАВЧАЛЬНОЇ ТА ДОСЛІДНИЦЬКОЇ СИСТЕМ



Зубань Оксана Миколаївна – доцент кафедри української мови та прикладної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка, кандидат філологічних наук, доцент

МЕТОДОЛОГІЯ УКЛАДАННЯ ПАРАЛЕЛЬНИХ КОРПУСІВ

Майстер-клас ставить за мету ознайомити учасників конференції із технологією укладання Корпусу паралельних текстів, розробленою в лабораторії комп'ютерної лінгвістики КНУ імені Тараса Шевченка.

Система укладання паралельного корпусу працює в режимі on-line. Учасникам майстер-класу буде запропоновано ввести тексти до корпусу за чіткою схемою процедури введення текстів і вирівнювання абзаців та речень. А тому учасники майстер-класу мають завчасно підготувати електронні тексти у форматі *.doc, *.docx. Для практичної роботи необхідно мати два паралельні (оригінальний і переклад) тексти: українською та англійською мовами.

Створення презентативного стилістично збалансованого паралельного корпусу текстів є надзвичайно складною роботою, яка вимагає цілеспрямованої праці великого колективу. І тому ми пропонуємо об'єднати зусилля всіх зацікавлених дослідників для оптимізації й ефективності такої роботи.

ЕЛЕКТРОННІ ЧАСТОТНІ СЛОВНИКИ – ІНСТРУМЕНТ СИСТЕМАТИЗАЦІЇ МАТЕРІАЛУ ЛІНГВІСТИЧНОГО ДОСЛІДЖЕННЯ

У цьому сегменті майстер-класу ставиться завдання ознайомити філологів із систематизацією лінгвістичної інформації у серії електронних частотних словників, автоматично укладених на матеріалі текстів Корпусу української мови. Попередньо із цими словниками можна ознайомитися за посиланням <http://www.mova.info/article.aspx?l1=210&DID=5215>.

Частотні словники такого типу – ефективні й раціональні інструменти лінгвістичних досліджень, оскільки вони передбачають різноманітні автоматичні інтерактивні класифікації мовних одиниць (лексем, словоформ, морфем, морфемних структур слів, словосполучень) на великих масивах українських текстів за запитами користувача в режимі on-line.

Учасникам майстер-класу пропонується використати частотні словники Корпусу для автоматичного формування матеріалу дослідження за пропонованими ними темами, що передбачають систематизацію лексики (за граматичними характеристиками), морфем, морфемних структур слів, словосполучень з урахуванням текстової реалізації цих одиниць.

ТЕЗИ ДЛЯ ОБГОВОРЕННЯ ПРОБЛЕМНИХ ПИТАНЬ НАУКОВОЇ ДИСКУСІЇ

ЕПІСТЕМІЧНІ АДВЕРБИ В РАННЬО НОВОАНГЛІЙСЬКІЙ МОВІ (дослідження на базі *Corpus of Historical American English*)

Андрушенко Олена Юріївна,

кандидат філологічних наук,

доцент кафедри міжкультурної комунікації та прикладної лінгвістики

Житомирського державного університету імені Івана Франка

Ярощук Надія Андріївна, магістрант

Житомирського державного університету імені Івана Франка

Дослідження на базі *Corpus of Historical American English* (далі – СОНА) спрямоване на визначення тенденцій появи та розвитку епістемічних адвербів в англійській мові в діяхронії.

Англійська мова має широкий спектр епістемічних адвербів порівняно з іншими мовами. Їх значення та функції детерміновані як мовними, так і культурологічними факторами, що зумовлює труднощі для не носіїв мови розмовляти, обираючи епістемічні адверби з потрібним ступенем впевненості. [2, с. 249]. Пропонуємо зосередити увагу на прислівниках *certainly*, *undoubtedly* та *unquestionably*.

Використовуючи корпус СОНА [1], була виявлена така тенденція розвитку окреслених адвербів: найбільша частота вживання *certainly* зареєстрована у 1900 р. – 3847 слововживань, тоді як у 1810 р. їх налічувалося 160, а у 2000 р. – 2938 одиниць. Подібна тенденція властива *undoubtedly*: 717 слововживань зареєстровано у 1900 р., 22 – у 1810р. та 251 – у 2000 році. Пік вживання *unquestionably* припадає на 1990 р. – 183 випадки, 25 слововживань зафіксовано у 1810 р. та 40 – зареєстровано у 2000 р.

Наведені кількісні данні дають можливість стверджувати про нисхідну тенденцію вживання епістемічних адвербів у сучасній англійській мові.

ЛІТЕРАТУРА

1. *Corpus of Historical American English*. URL: <http://corpus.byu.edu/coha/>
2. Wierzbicka, A. *English: Meaning and Culture*. Oxford: Oxford University Press, 2006. 363 p.

КОРПУСНЕ ДОСЛІДЖЕННЯ ЛОКАТИВУ *THERE* В АМЕРИКАНСЬКОМУ ВАРІАНТІ АНГЛІЙСЬКОЇ МОВИ XIX СТ.

Андрушенко Олена Юріївна,

кандидат філологічних наук,

доцент кафедри міжкультурної комунікації та прикладної лінгвістики

Житомирського державного університету імені Івана Франка

Тищенко Романія Василівна, магістрант

Житомирського державного університету імені Івана Франка

Дослідження Corpus of Historical American English [1] спрямоване на вивчення інформаційного навантаження локативу *there* з використанням моделі Е. Принс [2] при визначенні колишньої та нової інформації для дискурсу.

У 1810 р. корпус надає 738 випадків із локативом *there*. У нашому дослідженні увага сфокусована на таких моделях (410 прикладів): SV + *there* + (O) – 140, *There* + SVO – 78, S + *there* + V (O) – 64, *There* + VOS – 60, SVO + *there* – 28, *There* + VS – 12, VS + *there* + O – 12, *There* + OSV – 6, VO + *there* – 6, V + *there* + VS – 2, SO + *there* – 2. Варто зауважити, що протягом 1810-1850 рр. спостерігається кількісне зменшення таких структур, як VO + *there* (1820 – 3,7%, 1850 – 0%) та V + *there* + VS (1820 – 2,9%, 1840 – 1,6%, 1850 – 0%), а також редукція *There* + OSV та SO+ *there*. Це можна пояснити тим, що такий порядок слів був непритаманний окресленому періоду.

Корпусні дослідження показують, як місцезростащування *there* впливає на вираження топіку та фокусу.

Під час історичного розвитку американського варіанту англійської мови протягом 1810-1850-х рр. локатив частіше маркував колишню інформацію, тобто відому дискурсу раніше.

ЛІТЕРАТУРА

1. Corpus of Historical American English. URL: <http://corpus.byu.edu/coha/>
2. Prince E. F. The ZPG Letter: Subjects, Definiteness, and Information Status // Discourse Description. Diverse Linguistic Analyses of a Fund Raising Text. Amsterdam; Philadelphia: John Benjamins Publishing Company, 1992. P. 295–325.

ДОСЛІДЖЕННЯ КОГНІТИВНО-СЕМАНТИЧНИХ УНІВЕРСАЛІЙ ТА ПАРАЛЕЛЬНІ КОРПУСИ ТЕКСТІВ

*Балабан Олена Олександрівна,
кандидат філологічних наук, доцент,
докторант кафедри загального мовознавства і германістики
Національного педагогічного університету імені М. П. Драгоманова*

До дослідження когнітивно-семантичних універсалій, на нашу думку, можна долучити і паралельні корпуси текстів. Це корпуси, де тексти, написані різними мовами на одну й ту ж тему, зіставляються між собою. Окрім вирівнювання в паралельних корпусах існує і розмітка: метатекстова, граматична або семантична. Це окремі теги, що припускаються окремим словам або всьому тексту в цілому [1]. В нашому дослідженні для виокремлення когнітивно-семантичних універсалій більш доцільною буде семантична розмітка, тобто можна буде запросити видати всі контексти, в яких наявні, наприклад, всі позначення предметів одягу, знайти всі ці слова і подивитись, як різні перекладачі в різних контекстах з ними працюють.

Ми підтримуємо думку О. Голубкової, що корпусні дослідження вдало komponуються з когнітивними, оскільки й ті, й інші:

- спираються на лінгвістичні докази та знаходять їх у текстах/дискурсі різного характеру та об'єму але, головне, в автентичному дискурсі;
- вивчають мову в дії та на базі цих даних утворюють картину компетенцій носіїв мови (так званий bottom-up approach – підхід “знизу вгору”);
- займаються пошуками еталона (узусу) і девіацій від нього (за допомогою корпусних даних когнітивна лінгвістика з високою вірогідністю встановлює прототипні члени категорій як найбільш частотних та типових);
- вивчають дистрибутивні особливості розбіжностей значень слова з різною метою (наприклад, для проведення концептуального аналізу для складання сітки значень полісеманта) [2, с. 52–53].

ЛІТЕРАТУРА

1. Добровольский Д. О. Корпус параллельных текстов: архитектура и возможности использования / Национальный корпус русского языка. Результаты и перспективы. М.: Индрик, 2004-2005. С. 263-296.
2. Методы когнитивного анализа семантики слова: компьютерно-корпусный подход / под общ ред. В. И. Заботкиной. М.: Языки славянской культуры, 2015. 334 с.

ОСНОВНІ ПРИНЦИПИ УКЛАДАННЯ БРИТАНСЬКОГО НАЦІОНАЛЬНОГО КОРПУСУ ТА ЇХНІЙ ТРАНСФЕР У МЕТОДОЛОГІЮ РОЗРОБКИ ІНШИХ КОРПУСІВ

Бобер Наталія Миколаївна,

старший викладач

Київського національного економічного університету імені Вадима Гетьмана

Сучасні комп'ютерні мережі сприяють і водночас зумовлюють необхідність пошуку якісно нових способів і засобів опрацювання інформації, які будуть перспективно розвиватися, особливо у напрямі інтелектуалізації [2]. Наразі серед способів швидкого опрацювання інформації дослідники розглядають різні корпуси і, особливо, Британський національний корпус, який вважають еталоном у корпусній лінгвістиці (Дж. Синклер).

Говорячи про концепцію укладання Британського національного корпусу, слід зазначити, що вона спиралася на загальну ідею розробки корпусів, що базується на певних принципових критеріях, яким має відповідати зібраний масив текстів. О. М. Демська-Кульчицька у своїй статті "Базові поняття корпусної лінгвістики", покликаючись на праці У. Френсіса та Г. Кучери, перерахувала визначені цими дослідниками основні критерії відбору текстів, яких мають дотримуватися укладачі корпусу:

1. Походження і склад тексту.
2. Синхронізація.
3. Продумане співвідношення різноманітних жанрів та відбір окремих текстів.
4. Доступність для комп'ютерної обробки (спеціальні посилання для передавання графічних особливостей тексту) [1].

Концепцію розробки Британського національного корпусу було розпочато з ретельного етапу планування, на якому було визначено принципи його проектування й укладання та систематизовано перелік критеріїв відбору текстового матеріалу. Основоположним при цьому був принцип укладання корпусу текстів так, щоб інформацію зручно було опрацьовувати на комп'ютері та щоб було забезпечено автоматичний пошук потрібних для користувача даних. І ще одна принципова позиція укладачів Британського національного корпусу полягала в тому, що вони обмежилися лише британським стандартом англійської мови, оскільки значну частину витрат на створення корпусу покривав Британський уряд.

Варто розглядати ці принципові вимоги до створення корпусу як загальні й засадничі, які повинні бути, окрім інших конкретних принципів, покладені в основу розробки будь-якого корпусу. Особливо звертаємо увагу на те, що важливим є перший етап – етап планування і проектування, тобто концептуальна ідея створення корпусу, від якої в подальшому залежатиме якість готового інтелектуального продукту, його затребуваність та соціальна, наукова, навчальна тощо значимість.

ЛІТЕРАТУРА

1. Демська-Кульчицька О. М. Базові поняття корпусної лінгвістики // Українська мова. 2003. № 1. С. 42–47.
2. Широков В. А., Бугаков О. В., Грязнухіна Т. О., Костишин О. М., Кригін М. Ю., Любченко Т. П., Рабулець О. Г., Сидоренко О. О., Сидорчук Н. М., Шевченко І. В., Шипнівська О. О., Якименко К. М. Корпусна лінгвістика: монографія. К.: Довіра, 2005. 471 с.

ЗАЛУЧЕННЯ ДАНИХ КОРПУСІВ ПРИ ДОСЛІДЖЕННІ ХУДОЖНЬОЇ КАРТИНИ СВІТУ

*Брик Марина Мурманівна,
аспірантка кафедри загального мовознавства і германістики
Національного педагогічного університету імені М. П. Драгоманова*

Художня картина світу найяскравіше виявляється у своєрідності образної системи, створеної автором. Образна аналогія – один із найважливіших прийомів пізнання зовнішнього світу, в результаті якого виникає нове знання про нього: троп “створює не просто семантичне зміщення, а принципово нову і парадоксальну семантичну ситуацію” [1, с. 15]. Вибір мовних засобів зумовлений світосприйняттям письменника та структурою його картини світу. Надаючи велику репрезентативну вибірку текстів, корпус дозволяє виділити в авторському дискурсі типові для національної літературної традиції метафоричні сполучення. Припускаємо, що “корпусний метод є надійним інструментом для аналізу реальних слововживань у природному мовному середовищі” [2, с. 11]. Оснащення комп’ютеризованих корпусів морфологічною розміткою уможливує вивчення різних граматичних структур, якими виражені метафори, та надає підстави для визначення їх особливостей у конкретному тексті. Відсутність фіксації досліджуваних художніх образів у корпусі є свідченням їх унікальності та основою для вивчення індивідуально-авторської картини світу.

ЛІТЕРАТУРА

1. Лакофф Дж., Джонсон М. Метафоры, которыми мы живем // Теория метафоры; вступ. ст. и сост. Н. Д. Арутюнова. М.: Прогресс, 1990. 387 с.
2. Молодкина Ю. Н. Синестетическая метафора запаха: корпусное исследование: автореф. дис. канд. филол. наук: 10.02.19. Курск, 2010. 20 с.

КЛАСИФІКАЦІЯ ПРАВОВИХ ТЕКСТІВ ТА ТЕРМІНІВ ООН, НАТО І СОТ ВІДПОВІДНО ДО НОРМ МОВИ СПЕЦІАЛЬНОГО ПРИЗНАЧЕННЯ

*Дем'янчук Юлія Ігорівна,
кандидат економічних наук,
викладач кафедри технічного перекладу
Львівського державного університету безпеки життєдіяльності*

Включення в тематичний список загальноправових, галузевих, міжгалузевих юридичних термінів, а також інших одиниць спеціальної мови, які наявні в правових текстах ООН, СОТ та НАТО, здійснюється як на основі методичних (класифікаційних), так і лінгвістичних принципів, серед яких: принцип семантичної цінності, принцип лексичної цінності, принцип однозначності чи багатозначності слова. Застосовується також принцип термінотворення, зокрема аббревіація, яка активно використовується в юридичній німецькомовній термінології ООН (GG: Grundgesetz; BGB: Bürgerliches Gesetzbuch). Одним із критеріїв розміщення спеціалізованих лексичних одиниць у тематичному словнику – це критерій обліку частин мови. До найбільш поширених у юридичній термінології належать субстантивні ТЕ [1]: 1) субстантивні прості (Recht, Zoll), афіксальні (Verjährung, Gläubiger), складні терміни (Eigentums garantie, Rechtsfähigkeit); 2) субстантивні термінологічні словосполучення (Verjährung einer Klage, böswillige Verletzung). Також у юридичній термінології виділяються терміни-дієслова (beurkunden, vererben), іменникові терміни (belastend, antragsberechtigt), терміни-прикметники (vererblich, vorsätzlich). Інший важливий лінгвістичний принцип, яким необхідно керуватися при відборі активної спеціалізованої лексики, – це принцип врахування багатозначності слова.

Поширене явище полісемії в мові права ООН призводить дуже часто до проблем нерозуміння в міжнародному правовому середовищі. Існують причини виникнення багатозначних лексичних одиниць: 1) процес термінологізації загальноновживаної лексики; 2) наявність однакових термінів

у двох або більше галузях права, що відображають різні правові поняття. Зокрема загальноживана лексика може переходити в розряд юридичних термінів і використовуватися в термінологічному значенні, наприклад, Niederlassung [1]: 1) поселення, селище (загальноживане значення); 2) відділення, філія (спеціальна лексика); 3) дозвіл приватної (адвокатської) практики (юридичний термін).

ЛІТЕРАТУРА

1. Official Document System Search // United Nations. URL: <http://www.un.org/ru/documents/index.html>.

ПАРАЛЕЛЬНИЙ КОРПУС ТВОРІВ АЛЬБЕРА КАМЮ У ПЕРЕКЛАДІ АНАТОЛЯ ПЕРЕПАДІ

Дюндик Олена Андріївна,

*кандидат філологічних наук, молодший науковий співробітник
відділу лексикології, лексикографії та структурно-математичної лінгвістики
Інституту української мови Національної Академії наук України*

Формуючи загальномовний корпус українських текстів, не можна оминати увагою тексти художніх перекладів, адже “перекладна література виступає невід’ємною ланкою національного літературного процесу” [1, с. 9]. Формування паралельних корпусів нині активізується в українських корпусних студіях, адже відкриває шлях до міждисциплінарних досліджень: “Паралельні корпуси дають можливість заповнити прогалину між лінгвістикою і теорією перекладу, з потенційними перевагами для обох галузей, оскільки вони роблять дані перекладу доступними для наукового дослідження” [2, с. 96].

Анатоль Перепадя – знаний український перекладач, учень М. Лукаша. Перекладав із французької, іспанської, каталонської та інших романських мов. Сформований, протегований, анотований та вирівняний паралельний корпус творів Альбера Камю у перекладі Анатолія Перепаді дозволяє проводити спостереження за морфологією, лексикою та синтаксисом як української та французької частини окремо, так і досліджувати ці тексти у взаємозв’язку і взаємозалежностях. Корпусні технології дозволяють генерувати індекси слів за частотою, морфологічною будовою, порівнювати сполучуваність окремих слів за допомогою конкордансів; двомовний корпус дозволяє вивчати всі ці та інші особливості мови перекладів у зіставленні з мовою-оригіналом.

ЛІТЕРАТУРА

1. Коломієць Л. В. Український художній переклад та перекладачі 1920-30-х років: матеріали до курсу “Історія перекладу”: навчальний посібник. Вінниця: Нова книга, 2015. 360 с.
2. Salkie R. How can linguistics profit from parallel corpora? // Parallel corpora, parallel worlds / ed. Lars Borin. Amsterdam – New-York, 1999. P. 93–111.

ПЕРСПЕКТИВИ РОЗРОБЛЕННЯ СПЕЦІАЛІЗОВАНОГО КОРПУСУ НА ОСНОВІ БАЗИ ДАНИХ “ЛІНГВОАНТРОПОГЕНЕЗ”

*Железко Алла Миколаївна,
кандидат історичних наук,
старший науковий співробітник відділу організації наукових досліджень
Національного педагогічного університету імені М. П. Драгоманова*

Розвиток сучасної лінгвістики характеризується синтезом мовознавства та кібернетики, що сприяв появі комп'ютерної лінгвістики, а згодом – і корпусної лінгвістики [3, с. 9]. Можливість опрацювання величезної кількості емпіричних мовних даних відкриває величезні перспективи: створення спеціалізованих корпусів та корпусів текстів [2, с. 224].

Однією з перших спроб створити спеціалізований корпус є науковий фундаментальний проект “Германо-романо-слов'янський корпус аутентичних наукових текстів з лінгвоантропогенезу: розробка технологій нового покоління”, який виконується на факультеті іноземної філології в НПУ імені М. П. Драгоманова під керівництвом доктора філологічних наук, професора Алли Валер'янівни Корольової [4, с. 119].

Корпусна лінгвістика виокремилася зі слов'янознавства (славістики), германістики та романістики як прикладна наукова дисципліна, що описує реальні мовні явища, їхні елементи та структуру на основі аналізу автентичних письмових або усних текстів, зібраних у лінгвістичні текстові корпуси [3, с. 10].

Корпус – це зібрання текстів певною мовою (мовами) в електронній формі, відібраних за визначеним критерієм, яке певним чином впорядковане, але основною особливістю корпусу є його розмітка (анотування, маркування) [1, с. 209]. На сьогодні існують різні типи розмітки: 1) екстралінгвістична: метатекстова, структурна, типографська; 2) лінгвістична: морфологічна, синтаксична, семантична, анафорична, просодична розмітки [2, с. 226].

У нашому проекті створена база даних “Лінгвоантропогенез” містить інформацію, яка включає “широкий спектр питань, пов'язаних зі становленням людського роду в історико-еволюційному аспекті”, а також головне питання – виникнення людської мови [4, с. 120]. Тому, на нашу думку, важливо створити таку пошукову систему (корпусний менеджер), якою могли б користуватися не тільки лінгвісти, але й представники з інших галузей науки.

Одним із основних показників цінності корпусу є багатство пошукових можливостей, які надаються користувачеві. Для надання доступу користувачам, опрацювання, багаторазового використання корпусу необхідне відповідне стандартне оформлення метаданих текстів, які запропонувала до розроблюваного корпусу Н. Є. Леміш [5].

ЛІТЕРАТУРА

1. Войтко Г. Корпусна лінгвістика: історія формування і перспективи розвитку // Науковий вісник Східноєвропейського національного університету імені Лесі Українки. 2014. Вип. 5. С. 208–212.

2. Данилюк І. Корпус текстів для вивчення граматичної службовості // Лінгвістичні студії. 2013. Вип. 26. С. 224–229.

3. Ковбасюк Л.А. Корпусна лінгвістика та германістика: теоретичні засади і перспективи // Наукові записки НДУ імені Миколи Гоголя. Філологічні науки. 2017. Кн. 1. С. 9–14.

4. Корольова А.В. Принципи розробки бази даних “лінгвоантропогенез”: концепція, структура і зміст // Науковий часопис НПУ імені М.П. Драгоманова. Серія 9. Сучасні тенденції розвитку мов. 2017. Вип. 15. С. 119–125.

5. Леміш Н. Є. Корпусна метарозмітка спеціальних текстів з лінгвоантропогенезу // Науковий часопис національного педагогічного університету імені М. П. Драгоманова. Серія 9. Сучасні тенденції розвитку мов. 2017. Вип. 15. С. 139–152.

КВАНТИТАТИВНО-КОРПУСНИЙ ПІДХІД ДО АНАЛІЗУ СЕМАНТИКИ ГРАМАТИЧНОЇ КОНСТРУКЦІЇ

*Жуковська Вікторія Вікторівна,
кандидат філологічних наук, доцент,
завідувач кафедри міжкультурної комунікації та прикладної лінгвістики
Житомирського державного університету імені Івана Франка*

Семантика граматичної конструкції досліджується на прикладі англійської абсолютної конструкції з дієприкметником I ([_{AC} [_{SubjNoun}] [_{PredParticiple I}]]) із застосуванням новітнього методу квантитативної корпусної лінгвістики – колострукційного аналізу [2; 3]. Цей корпусно-базований лінгвостатистичний метод пропонує об'єктивний підхід до встановлення значення граматичної конструкції шляхом обрахування ступеню взаємного притягнення / відштовхування між конструкцією та лексемами, що заповнюють її провідний слот. Зasadничою основою методу є принцип *семантичної сумісності* між значеннями лексеми і граматичної конструкції. Значення абсолютної конструкції визначається на основі семантичного аналізу колексем її суб'єктного слоту [_{AK Subj}].

Колострукційний (зокрема колексемний) аналіз був застосований до вибірки досліджуваної конструкції з корпусу BNC-BYU [4] (2203 прикладів). У результаті відповідних обрахунків із застосуванням статистичної системи аналізу даних R та скрипту C. Гріса [1] отримано список іменників, що виявили статистично значимий зв'язок із слотом [_{AK Subj}] конструкції. Подальший семантичний аналіз цих іменників показав, що серед них кількісно переважають номінації елементів будови тіла людини. Розгляд комунікативних функцій абсолютних конструкцій з іменниками зазначеної семантики дозволив попередньо виокремити у семантиці конструкції два значення, що представлені у підконструкціях [зовнішність X [_{MSubj}] має [_{AK Y} [_{Subj частина тіла}] у стані / процесі V [_{Pred Participle I}]] та [внутрішній стан X [_{MSubj}] має [_{AK Y} [_{Subj частина тіла}] у стані / процесі V [_{Pred Participle I}]].

ЛІТЕРАТУРА

1. Gries S. Th. Coll. Analysis 3.2a. A program for R for Windows 2.x. URL: <http://www.linguistics.ucsb.edu/faculty/stgries/teaching/groningen/>
2. Gries S. Th. More (old and new) misunderstandings of colostruational analysis: On Schmid and Küchenhoff. *Cognitive Linguistics*. 2015. № 26(3). P. 505–536.

3. Stefanowitsch A., Gries St. Th. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*. 2003. № 8(2). P. 209–243.

ДЖЕРЕЛА ІЛЮСТРАТИВНОГО МАТЕРІАЛУ

4. British National Corpus (BNC-BYU). URL: <http://corpus.byu.edu/bnc/>

СТРУКТУРНО-ПРАГМАТИЧНИЙ АСПЕКТ ТЕКСТІВ ДЕБАТІВ БРИТАНСЬКОГО ПАРЛАМЕНТУ

Зернецький Павло Васильович,

кандидат філологічних наук,

професор кафедри англійської мови

Національного університету “Київо-Могилянська Академія”

У процесі дослідження використовується авторська класифікація мовленнєвих актів (топікальних та функціональних), розроблена у монографії “Речевое общение на английском языке (Коммуникативно-функциональный анализ дискурса)” [1, с. 101–103], разом із моделлю аналізу мовленнєвого дискурсу Дж. Сінклера та М. Култхарда [2, с. 153], що застосовується для опису конструктивних елементів мовленнєвих кроків (Pre-Head, Head, Post-Head) для структурно-прагматичного аналізу мовленнєвої діяльності учасників Британських парламентських дебатів.

В результаті дослідження встановлюються структурно-прагматичні типи функціональних зв'язків між різними конструктивними елементами мовленнєвих кроків і надається опис типології характерних для Британського парламентського дискурсу типів структурно-прагматичних конфігурацій смислової побудови дискурсу цього типу.

ЛІТЕРАТУРА

1. Зернецький П. В. Речевое общение на английском языке (Коммуникативно-функциональный анализ дискурса). К.: Лыбидь, 1992. 144 с.

2. Sinclair J. M., Coulthard R. M. *Toward an analysis of discourse: the English used by teachers and pupils*. London: Oxford University Press, 1975. 163 p.

**АВТОМАТИЧНЕ УКЛАДАННЯ
ЧАСТОТНИХ МОРФЕМНИХ СЛОВНИКІВ
(на матеріалі корпусу української мови)**

*Зубань Олена Миколаївна,
кандидат філологічних наук, доцент,
доцент кафедри української мови та прикладної лінгвістики Інституту філології
Київського національного університету імені Тараса Шевченка*

У межах наукового проекту “Корпус української мови” [1] колективом лабораторії комп’ютерної лінгвістики КНУ імені Тараса Шевченка було розроблено методику комп’ютерного моделювання морфемних структур слів, на основі якої створено автоматизовану систему морфемно-словотвірного аналізу (АСМСА). Система АСМСА працює на матеріалі морфемної бази даних (МБД) обсягом \approx 200 тис. слів. Функціонально система АСМСА може використовуватися як: 1) інформаційно-довідкова система з морфеміки та словотвору української мови; 2) автоматичний лінгвістичний конструктор частотних морфемних словників (ЧС) на базі лексики текстових вибірок у Корпусі української мови; 3) автоматичний морфемний сегментатор текстових слововживань Корпусу. На сьогодні за текстами Корпусу автоматично укладено 13 частотних морфемних словників на матеріалі текстових вибірок окремих авторів та стилів [2].

ЛЕКСИКОГРАФІЧНІ ДЖЕРЕЛА

1. Корпус української мови. URL: <http://www.mova.info/corpus.aspx>.
2. Частотні словники Корпусу. URL: <http://www.mova.info/article.aspx?l1=210&DID=5215>.

ПАРАЛЕЛЬНИЙ КОРПУС ТЕКСТІВ: СТРУКТУРА, АСПЕКТИ ПАРАМЕТРИЗАЦІЇ ТА ТЕХНОЛОГІЇ УКЛАДАННЯ

Зубань Олена Миколаївна,

кандидат філологічних наук, доцент,

*доцент кафедри української мови та прикладної лінгвістики Інституту філології
Київського національного університету імені Тараса Шевченка*

Корпус паралельних текстів – це новий проект лабораторії комп'ютерної лінгвістики КНУ імені Тараса Шевченка, розпочатий у 2017 р. викладачами та студентами у межах нормативного курсу "Основи автоматичного перекладу" та навчально-виробничої практики з машинного перекладу.

Аспекти параметризації та структура паралельного корпусу визначалися спрямуванням на виконання таких функцій: 1) ілюстративно-перекладацької: використання корпусу як текстоорієнтованого перекладного словника; 2) дослідницької: використання корпусу для вирішення дослідницьких завдань у різних галузях порівняльного мовознавства та літературознавства; 3) навчальної: використання корпусу у вивченні іноземної мови, а також автоматичне укладання лексичних мінімумів різних навчальних дисциплін; 4) автоматичного перекладу: використання корпусу у системах машинного перекладу із перекладацькою пам'яттю.

ЕФЕКТИВНИЙ ПОШУК ПО КОРПУСУ ЗА ДОПОМОГОЮ РЕГУЛЯРНИХ ВИРАЗІВ

Івашкевич Леся Сергіївна,

кандидат філологічних наук, доцент,

доцент кафедри теорії, практики та перекладу німецької мови

Національного технічного університету України

"Київський політехнічний інститут імені Ігоря Сікорського"

Для ефективного пошуку у власному чи вже готовому та завантаженому корпусі зручно користуватися спеціалізованою мовою програмування – так званими "регулярними виразами" (англ. *regular expressions*). У регулярних виразах для певних типів пошуку використовуються різні технічні символи, наприклад, `.,|+{}[]()?./*`: та ін. Комбінації цих знаків дозволяють знаходити, приміром, всі

слова з певною кількістю букв, всі слова, що містять кілька певних морфем, наприклад, бажаний префікс та бажаний суфікс одночасно, всі цифри, всі слова, що відрізняються лише певними літерами та ін. У протегованому за частинами мови корпусі регулярні вирази допомагають знаходити потрібні сполучення певних частин мови. Регулярні вирази інтегровані також у деякі корпусні менеджери (наприклад, AntConc) або платформи для роботи з корпусами. Також за допомогою регулярних виразів можна опрацьовувати текст, застосовуючи до нього певні алгоритми, що може бути зручним при створенні власного корпусу (наприклад, усунення зайвої анотації після тегування тексту у парсері) [1, с. 733]. Покроковий інтерактивний онлайн-тьюторіал з регулярних виразів можна знайти на сторінці <https://regexone.com>, тренуватися шукати за допомогою регулярних виразів можна в онлайн-редакторі <https://regexr.com>.

ЛІТЕРАТУРА

1. Івашкевич Л. С. Інструменти для укладення та обробки власного корпусу // Наукові записки. Серія: Філологічні науки. Кропивницький, 2017. Вип. 153. С. 729–733.

МІЖНАРОДНА БАЗА ДАНИХ “ВАВИЛОНСКАЯ БАШНЯ” ЯК ПОШУКОВА СИСТЕМА ДЛЯ ВИВЧЕННЯ НОСТРАТИЧНИХ ЕТИМОЛОГІЙ

Капранов Ян Васильович,

кандидат філологічних наук, доцент

кафедри англійської філології і перекладу

Київського національного лінгвістичного університету

На сьогодні в мережі Інтернет представлена міжнародна база даних з відкритим доступом – “Вавилонская Башня” – глобальна лексикостатистична база даних (далі – ГЛБД), яка була запущена в 1998 р. як персональна сторінка Сергія Анатолійовича Старостіна – одного із засновників макрокомпаративістики. “Вавилонская Башня” перетворилася на міжнародний інформаційний ресурс. Однією зі складових ГЛБД є “Ностратична етимологія”, яка була укладена С. А. Старостіним за даними як наявних на той час ностратичних словників (В. М. Ілліч-Світлич та ін.), так і з залученням нових етимологій, запропонованих самим С. А. Старостіним й іншими учасниками Московської школи компаративістики [1].

Для пошуку ностратичних етимологій використовується ієрархічна система, що включає списки слів різних рівнів, від вищого до нижчих, що, з одного боку, не лише полегшує роботу з величезним обсягом інформації, але і відповідає концепції генеалогічного дерева, згідно з якою з мов-предків виростають численні мови-нащадки, на основі яких методами історичної лінгвістики можна реконструювати їх спільну мову-предок [2].

Пошук ностратичних етимологій здійснюється за трьома рівнями [2]:

Перший рівень – мовні групи (балтійська, тюркська та ін.) – складають порівняно невеликі бази даних, кожна з яких містить списки слів мов, що розійшлися не більше 3000 років тому і близька спорідненість яких не викликає сумнівів, а також список слів прамови, що є їхнім спільним предком.

Другий рівень – мовні сім'ї (індоевропейська, уральська та ін.) – бази, що містять списки лише реконструйованих слів прамови, які достовірно чи хоча б імовірно споріднені між собою. На цьому рівні також представлено список слів прамови, що є спільним предком представлених у даній базі прамови.

Третій рівень – мовні макросім'ї (ностратична та ін.) – база, де зіставляються форми кількох прамов різних сімей – у тому разі, якщо існує припущення про те, що між цими сім'ями є дуже глибоке генетичне споріднення.

ЛІТЕРАТУРА

1. Вавилонская башня. Проект “Эволюция языка”. URL: <http://starling.rinet.ru/intrab.php?lan=ru>
2. Глобальная лексикостатистическая база данных: общее описание. URL: <http://starling.rinet.ru/new100/GLDr.htm>

МОДЕЛЮВАННЯ ЯК КОГНІТИВНИЙ МЕТОД КОРПУСНОЇ ЛІНГВІСТИКИ

Карлова Вікторія Олександрівна,

аспірантка кафедри загального мовознавства і германістики

Національного педагогічного університету імені М. П. Драгоманова

На початку ХХІ ст. у лінгвістиці почали активно оперувати такими поняттями, як “модель / моделювання”. У результаті лінгвісти дійшли методологічної єдності у визначенні цих понять як наукових універсалій.

Тимчасом у корпусній лінгвістиці поняття модель розглядають як інструмент знання, деталізації і типологізації методів моделювання [2, с. 322].

Сучасна корпусна лінгвістика тісно пов'язана із когнітивною лінгвістикою, які разом зорієнтовані на вивчення мовленнєвої діяльності [1, с. 12]. У цьому ракурсі застосовується процедура моделювання для “створення гіпотетичної моделі явища із подальшою її верифікацією на мовному матеріалі”.

Зокрема, когнітивне моделювання на матеріалі даних корпусів може застосовуватися 1) як унаочнення кореляції між мовними структурами тексту і структурами знань автора [5, с. 201]; 2) як представлення смислових структур тексту: чуттєво-емоційних образів – метафорично згорнутих концептів тощо [2, с. 327] тощо.

ЛІТЕРАТУРА

1. Копотев М. В., Мустайоки А. Современная корпусная русистика / под ред. А. Мустайоки, М. В. Копотева, Л. А. Бирюлина, Е. Ю. Протасовой. Хельсинки: Helsinki University Press, 2008. 24 с.

2. Романова Т. В. Моделирование как метод верификации в когнитивной лингвистике. Когнитивные исследования языка. 2014. № 16. С. 322–327.

КОРПУС СЕРЕД ІНШИХ ІНСТРУМЕНТІВ ОПРАЦЮВАННЯ МОВИ

Карпіловська Євгенія Анатоліївна,

доктор філологічних наук, професор,

завідувач відділу структурно-математичної лінгвістики

Інституту української мови Національної Академії наук України

Кожен інструмент опрацювання мовної інформації має своє призначення та спектр використання. Практика створення й використання корпусів мови навіть у найвичерпнішому їхньому форматі (фундаментальний, або національний корпус на протигагу пошуковому, дослідницькому) переконливо підтвердила цю аксіому. Не можна не погодитися із С. Лендау, який нагадує дослідникам мови про вимоги до формування корпусу: показовість, збалансованість, структурованість, наявність анотування та спеціального інтерфейсу [1, с. 316].

Корпус для свого розвитку й об'єктивності потребує інструменту відстежування змін у мові, її динаміки. Таким інструментом може стати спеціальна картотека “портретів слова” чи будь-якої мовної або текстової

одиниці. У цій картотеці дослідник накопичує нову інформацію про мову, об'єктивність якої має довести корпус. Така електронна картотека й корпус взаємодіють як "сполучені посудини" інформації про мову.

ЛІТЕРАТУРА

1. Лендау С. І. Словники: мистецтво та ремесло лексикографії. К.: К.І.С., 2012. 479 с.

КОРПУСНА МЕТОДОЛОГІЯ У ВИВЧЕННІ ІНОЗЕМНИХ МОВ

Качанов Валерій Іванович,

кандидат військових наук,

старший викладач кафедри загального мовознавства і германістики

Національного педагогічного університету імені М. П. Драгоманова

У сучасному процесі навчання іноземних мов активно використовуються інноваційні методи і підходи. Одним із таких підходів є корпусний методологічний підхід, що має чимало переваг [1, с. 9]. По-перше, – це емпіричний процес, який надає змогу вивчати зразки мовленнєвих конструкцій в аутентичних текстах. А тому можемо говорити про те, що ключовий елемент корпусного підходу – використання аутентичних матеріалів [2, с. 14]. Важливим є не те, які тексти використовуються (фантастика, документальна проза, журнали, наукові статті, світова література, газети) або інші акти мовленнєвої комунікації (розмови, телефонні дзвінки), а їх аутентичність. По-друге, – корпусний підхід у вивченні іноземних мов надає змогу значно збільшити обсяг навчальної інформації, якою є аутентичні тексти, а відповідно, і відпрацювання на їхньому матеріалі особливостей функціонування тих чи інших мовних явищ [1, с. 9]. По-третє, – корпусний підхід передбачає інтенсивне використання нових технологій для вивчення іноземних мов [1, с. 9].

ЛІТЕРАТУРА

1. Gena B. Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers. University of Michigan Press ELT, 2010. P. 9–11.

2. Reppen R. Using corpora in the language classroom. Cambridge: Cambridge University Press, 2010. P. 14–15.

**СОЦІОЛІНГВІСТИЧНИЙ АСПЕКТ РОЗГЛЯДУ ПОНЯТТЯ
“ПРЕДМЕТНА ГАЛУЗЬ”
(на основі корпусу MICASE)**

*Комаренко Вероніка Василівна,
аспірант кафедри англійської філології та перекладу
Житомирського державного університету імені Івана Франка*

Соціолінгвістичні дослідження дискурсу, увагу яких зосереджено на аналізі учасників й обставин спілкування, спрямовані на вивчення соціального середовища мовців та умов, в яких відбувається фахова комунікація. Ідентифікація у соціальному просторі з'являється та виявляється у процесі мовленнєвої комунікації. Комуникант окреслює “своє коло”, відмежовуючи себе за яким-небудь “фактором ідентифікації – віковим, етнічним, соціальної чи гендерної приналежності, професії тощо” [1].

Однією з найважливіших якісних ознак будь-якого корпусу є його здатність коректно відобразити предметну галузь чи той мовний рівень, що його задекларовано корпусом. Під *предметною галуззю* розуміється рівень реалізації мовної системи, яка містить феномени, що підлягають лінгвістичному описові [2, с. 101–102].

Одним із ресурсів для соціолінгвістичних досліджень є корпус академічного мовлення (MICASE – Michigan Corpus of Academic Spoken English), що містить 139 затранскрибованих текстів загальним обсягом 1 695 540 слововживань. Предметна галузь, в якій відбувається академічна комунікація, поділяється на точні науки, до яких входять Biological and Health Sciences, Physical Sciences and Engineering, та гуманітарні науки, до яких відносять Humanities and Arts, Social Sciences and Education. Пошук в зазначеному корпусі також може здійснюватися відповідно до наступних параметрів: гендерна приналежність мовця, вік, академічна роль комуникантів, статус носія мови, рідна мова комуникантів, тип мовної ситуації, предметні дисципліни, академічна роль учасників та рівень інтерактивності.

Отже, корпус академічного мовлення MICASE – Michigan Corpus of Academic Spoken English дає можливість здійснювати соціолінгвістичний пошук відповідно до різних параметрів.

ЛІТЕРАТУРА

1. Ачкасов В. А. Этническая идентичность в ситуациях общественного выбора // Журнал социологии и социальной антропологии. 1999. Том II. URL: <http://www.soc.pu.ru/publications/jssa/1999/1/8achkas.html>

2. Демська-Кульчицька О. Деяко про класифікацію текстових корпусів // Наукові записки. Серія: Мовознавство, 2004. 1 (11). С. 153–157.

БАЗА ДАНИХ "ЛІНГВОАНТРОПОГЕНЕЗ" ЯК КОНЦЕПТУАЛЬНА ПЛАТФОРМА ДЛЯ СТВОРЕННЯ ГЕРМАНО-РОМАНО-СЛОВ'ЯНСЬКОГО КОРПУСУ АУТЕНТИЧНИХ НАУКОВИХ ТЕКСТІВ

*Корольова Алла Валер'янівна,
доктор філологічних наук, професор,
завідувач кафедри загального мовознавства і германістики
Національного педагогічного університету імені М. П. Драгоманова*

І Міжнародна науково-прикладна відео-Інтернет-конференція "Корпусна лінгвістика: розроблення технологій нового покоління" у своїй програмі передбачає проведення майстер-класу з проблематики обговорення практичних проблем, пов'язаних із розглядом корпусів текстів як особливих інформаційно-пошукових, навчальних та дослідницьких систем.

Однією з таких систем буде створюваний на факультеті іноземної філології НПУ імені М.П. Драгоманова "Германо-романо-слов'янський корпус наукових аутентичних текстів з лінгвоантропогенезу: розробка технологій нового покоління" під керівництвом проф. А.В. Корольової.

Наразі уже створена База даних "Лінгвоантропогенез", яка оснащена автоматизованою пошуковою системою і яка стане концептуальною платформою для подальшої її трансформації у корпус текстів із цієї проблематики.

База даних "Лінгвоантропогенез" містить інформацію наукового характеру, що включає розгляд проблем, які стосуються 1) різних гіпотез походження людської мови, теорій моногенезу і полігенезу, 2) типів спорідненості мов світу, мовних сімей, мовних союзів і макросімей, 3) теорій мовних контактів, 4) ностратичної теорії, 5) зв'язків мови і мислення, мови і свідомості, мови і культури, мови і етносу, мови і релігії, 6) процесів концептуалізації і категоризації явищ навколишньої дійсності, відображених у людській

свідомості, 7) механізмів формування мовних і концептуальних картин світу, різних концептосфер, 8) методик розкриття етимології слів у відповідних лексикографічних та інших джерелах, 9) процедур реконструкції системних зв'язків між мовними явищами в діахронії і т. ін.

Окреслений спектр різномірної інформації з проблематики лінгвоантропогенезу часто ускладнює пошук в мережі Інтернет конкретного і потрібного інформаційного ресурсу, актуального на даний момент для користувача-дослідника, викладача. Тому створюваний корпус задуманий як мобільна інформаційно-пошукова, навчальна та дослідницька система, яка буде працювати на основі інформаційних запитань користувача, який бажає отримати відповідь на сформульоване запитання у вигляді ключових слів. Автоматично система починає обробку цього запиту і видачу потрібної інформації, а результати пошуку заносяться в бланк запиту інформації. Попередня інформація щодо концепції розробки корпусу представлена в одній із моїх публікацій на тему: "Принципи розробки Бази даних "Лінгвоантропогенез": концепція, структура і зміст" [1].

ЛІТЕРАТУРА

1. Принципи розробки Бази даних "Лінгвоантропогенез": концепція, структура і зміст // Науковий часопис Національного педагогічного університету імені М. П. Драгоманова. Серія 9 Сучасні тенденції розвитку мов. К., 2017. Вип. 15. С. 119–125.

ПРОБЛЕМЫ СОЗДАНИЯ МЕТАФОРИЧЕСКОГО КОРПУСА

*Кравцова Юлия Валентиновна,
доктор филологических наук, профессор,
заведующая кафедрой русского языка*

Национального педагогического университета имени М. П. Драгоманова

В связи с интенсивным развитием корпусных технологий и появлением национальных корпусов разных языков (английского, чешского, русского и др.) стало актуальным создание метафорических корпусов (Amsterdam Metaphor Corpus, 2010; MetaNet, 2012; Русскоязычный корпус концептуальной метафоры, 2013), так как значительная роль метафоры в познании, концептуализации и репрезентации мира является общепризнанной. Однако при формировании электронных ресурсов метафорики и инструментария их автоматической

обработки неизбежно возникает ряд проблем (А. Н. Баранов, А. Д. Плисецкая, А. С. Самигуллина, С. Б. Кураш), к основным из которых, на наш взгляд, можно отнести следующие: 1) обоснование принципов и методов создания метафорического корпуса; 2) обеспечение его репрезентативности и сбалансированности; 3) рациональное использование национальных языковых корпусов; 4) подбор базовых толковых, семантических, метафорических словарей и их качественных онлайн-версий; 5) адекватный выбор способов лингвистической (семантической) разметки узуальных метафор; 6) разработка алгоритмов идентификации в текстах разной жанрово-стилевой квалификации окказиональных (инновативных) метафор и их интерпретации; 7) поиск эквивалентных метафорических употреблений в параллельных корпусах. Успешное решение перечисленных проблем позволит создать метафорический корпус, который крайне необходим для эффективного осуществления исследовательской, учебной и переводческой деятельности.

A KEYWORD APPROACH TO COMPARABLE SUB-CORPORA

Anna Kryvenko,

Doctor of Philosophy in Linguistics, Associate Professor,

*Professor Oleksandr Morokhovsky's Chair of English Philology and Philosophy of Language,
Kyiv National Linguistic University*

In corpus linguistics, keywords are statistically significant words which are used in the texts of a target corpus more frequently (or less frequently, if negative keyness is investigated) than in the texts of a reference corpus [2]. Keywords analysis has increasingly been used for the exploration of specialized discourses with the intention to unearth the propositional content of a corpus and its overall semantic cohesion as well as sets of meanings and values, which are of importance to some particular communities or institutions. One of the fundamental challenges in applying a keyword approach is narrowing down a list of automatically generated candidate items, which might contain hundreds or even thousands of keywords, for further analysis, even when tight statistical constraints are imposed [1, p. 60–61]. This contribution aims to lay out a comprehensible methodology for singling out relevant keywords in comparable specialized sub-corpora on European integration, which allows reducing subjectivity during the selection process.

LITERATURE

1. Groom N. Closed-class keywords and corpus-driven discourse analysis // *Keyness in Texts*. Amsterdam / Philadelphia: John Benjamins, 2010. P. 59–78.
2. Scott M., Tribble C. *Textual Patterns: Keyword and Corpus Analysis in Language Education*. Amsterdam / Philadelphia: John Benjamins, 2006. x, 203 p.

РОЗМІТКА ПАРАЛЕЛЬНОГО КОРПУСУ

Леміш Наталія Євгенівна,

доктор філологічних наук, доцент,

завідувач кафедри філософії мови, порівняльного мовознавства та перекладу

Національного педагогічного університету імені М. П. Драгоманова.

Паралельні корпуси разом з багатомовними суміжними і зіставними є найбільш важливими типами корпусів для перекладу, що пояснюється їх практичним призначенням. Загальновідомо, що для побудови корпусу паралельних текстів, а саме: для встановлення відповідності фрагментів оригінальних текстів їх перекладним варіантам іншими мовами необхідно розробити спеціальне програмно-технічне забезпечення. Розмітка текстів паралельного корпусу, як правило, включає три традиційні види: метатекстову, структурну і лінгвістичну, але здійснюється вона як для тексту оригіналу, так і для перекладного варіанту з одночасним виведенням після запиту результатів на екран.

Метатекстова розмітка [1, с. 76; 2, с. 146] включає назву твору, жанр, тематику, рік і місце видання, відомості про автора / перекладача (ім'я та прізвище, рік народження, вік, стать). *Структурна* розмітка приписує спеціальні мітки розділам, абзацам, реченням, словоформам. *Лінгвістична* розмітка для паралельного корпусу складається з морфемно-морфологічної (маркування префіксів, суфіксів, закінчень, частин мови) і синтаксичної (визначення типу сполуки + формалізованої схеми речення).

Отже, коректна розмітка паралельних текстів уможливорює здійснення пошуку в паралельному корпусі не лише за мовою оригіналу чи перекладу, за хронологією, за ім'ям автора тощо, а й за окремими абзацами, частинами мови, словоформами, морфемами, типами сполук, схемами речень.

ЛІТЕРАТУРА

1. Жуковська В. В. Вступ до корпусної лінгвістики: навч. посіб. Житомир: Вид-во ЖДУ імені І. Франка, 2013. 142 с.
2. Леміш Н. Є. Корпусна метарозмітка спеціальних текстів з лінгвоантропологенезу // Науковий часопис національного педагогічного університету імені М. П. Драгоманова. Серія 9. Сучасні тенденції розвитку мов. 2017. Вип. 15. С. 139–152.

СПЕЦИФІКА МУЛЬТИМОДАЛЬНИХ КОРПУСІВ

Макарук Лариса Леонідівна,

кандидат філологічних наук,

старший викладач кафедри прикладної лінгвістики

Східноєвропейського національного університету імені Лесі Українки

Мультимодальні корпуси текстів – одні з найнеординарніших та найменш досліджених у сучасних лінгвістичних студіях. Незважаючи на це, у найближчі роки вони можуть стати найбільш затребуваними та вагомими, оскільки складаються з різнорідних елементів, які дають змогу розв'язати одночасно низку проблем, врахувавши особливості кількох модальностей як єдиного комплексу семіотично навантажених модусів. З-поміж інших, вони відрізняються тим, що окрім текстового (вербального) складника також вміщують й аудіо- та відеозаписи. Враховуючи це, мультимодальні корпуси текстів є одними з найскладніших не лише за будовою, а й з огляду на їх наповнення, структуру та покликання.

Анотування таких корпусів кардинально відрізняється від традиційних письмових корпусів, в яких усе залежить від словесного модусу. Складність мультимодальних корпусів полягає у тому, що вони ґрунтуються на різних модусах, наприклад, жестах, міміці, темпі й тембрі, які є невід'ємними характеристиками усного мовлення, темпоральних і просторових характеристиках. До мультимодальних корпусів текстів належать: АМІ, SmartKom and HuComTech тощо. Анотування мультимодальних корпусів текстів залежить від кількох пов'язаних між собою рівнів анотації, в основі яких лежать різні модуси. Мультимодальну розмітку вважають непростим і довготривалим процесом, який спрямований на ґрунтовне дослідження й урахування усіх модусів, залучених до процесу інтеракції. Кожен мультимодальний корпус має свою специфіку й покликання, тому в одних із них є більш вираженою, наприклад, аудіальна модальність, у той час як в інших – візуальна.

ВИМОГИ ДО УКЛАДАННЯ СПЕЦІАЛІЗОВАНИХ МОВНИХ КОРПУСІВ

Маслова Тетяна Борисівна,

старший викладач кафедри англійської мови технічного спрямування № 1

Національного технічного університету України

“Київський Політехнічний Інститут імені Ігоря Сікорського”

Завдяки сучасним комп'ютерним технологіям з'явилися нові технічні можливості для обробки, збереження і відбору лінгвістичних даних у вигляді “корпусу”, тобто репрезентативного зібрання упорядкованих текстів, що може розглядатися як зменшена модель мови чи підмови [1, с. 367]. Текстові корпуси знаходять застосування у різних галузях: їх використовують в комп'ютерній лінгвістиці, в корпусній стилістиці, в лексикографії, а також у методиці викладання іноземних мов [2]. Проте наявність великої кількості текстів, які є доступними для лінгвістичного аналізу, напевно, є найбільш корисною для проведення лінгвістичних досліджень, оскільки автоматичне опрацювання реальних мовних явищ допомагає оптимізувати і об'єктивізувати пошук та статистичну обробку мовних даних.

Існує чимало національних корпусів, які репрезентують мову в цілому, проте зараз спостерігається зацікавленість у створенні спеціалізованих корпусів із метою відобразити певну підмову і вирішити конкретні лінгвістичні завдання. Ключовими вимогами є репрезентативність і збалансованість корпусів. У випадку відсутності репрезентативності будь-які результати дослідження, отримані в межах корпусу, не можна буде екстраполювати на відповідну ділянку мови, бо буде незрозуміло, як корпус з нею співвідноситься. Збалансованість передбачає пропорційність різних частин корпусу з огляду на такі характеристики, як автори, реципієнти і жанри текстів. Забезпечення репрезентативності й збалансованості можливе лише за умови чіткого визначення параметрів сукупності текстів, з якої потрібно зробити вибірку [1; 3]. Параметризація корпусу, тобто оптимальне поєднання різних підходів, методів і критеріїв із урахуванням специфіки конкретного корпусу включає: широту (охоплення максимальної кількості текстів), глибину (виокремлення підкатегорій серед обраних текстів), пропорційність (добирання відсутніх текстів у певній пропорції) і часовий вимір.

ЛІТЕРАТУРА

1. Коцюк Л., Очеретович О. Параметризаційні характеристики корпусу текстів біатлону сучасної англійської мови // Наукові записки Національного університету “Острозька академія”. Серія: Філологічна. 2016. Вип. 62. С. 366–368.

2. Маслова Т. Інструменти корпусної лінгвістики у навчанні і дослідженні іноземних мов // Smart-освіта: ресурси та перспективи: матеріали II Міжнар. наук.-метод. конф., Київ, 23 лист. 2016 р.: тези доповідей. К.: Вид-во КНТЕУ, 2016. С. 277–280.

3. Старко В., Чейлитко Н. Параметризація корпусу як спосіб підвищення його репрезентативності та збалансованості // Українське мовознавство. 2013. Вип. 43. С. 87–94.

ТОКЕНІЗАЦІЯ ЯК СПОСІБ ОБРОБКИ КОРПУСНОГО ТЕКСТУ

Матвеева Світлана Анатоліївна,

кандидат філологічних наук, доцент

професор кафедри філософії мови, порівняльного мовознавства та перекладу

Національного педагогічного університету імені М. П. Драгоманова

Первинне автоматичне розмічування тексту для подальшої обробки в межах корпусу передбачає виконання таких основних операцій: сегментація тексту (визначення абзаців та речень), токенізація, стемінг, лематизація, морфологічний аналіз. Під токенізацією розуміють процес “сегментації тексту, що є послідовністю символів, таких як букви, пробіли, знаки пунктуації та цифри, на слова і фрази” [1, с. 5]. Завданням токенізації є відокремлення слів від синтаксичних знаків, цифр, комплексів літер та цифр, інтернет-адрес, нікнеймів, знаків %, +, -, // тощо. При цьому створення досконалого алгоритму машинної ідентифікації та відмежування токенів є невирішеним питанням через наявність у текстах великої кількості одиниць та комбінацій, однозначна автоматична класифікація яких на цьому етапі є неможливою, наприклад: багатослівні токени (*Переяслав-Хмельницький*); імена, що містять знаки (*Григор'єв*); числа, що містять пробіли; наявність у тексті крапки, яка не є ознакою кінця речення (пряма мова, скорочення *м. – місто, с. – село*), розділових знаків, що містять більше одного символу (напр., три крапки) тощо. На сучасному етапі ці та багато інших випадків не можна класифікувати автоматично, такі процедури виконуються механічним способом. Розв'язання зазначених питань уможливить належний рівень первинного автоматичного опрацювання текстового матеріалу і подальше якісне розпізнавання смислу окремих одиниць та всього тексту у складі певного корпусу.

ЛІТЕРАТУРА

1. Teodorescu M. H. Machine Learning Methods for Strategy Research. HBS Working Paper 18-011. Harvard Business School, 2017. 59 p.

ПЕРСПЕКТИВА І ПОДАЛЬШИЙ РОЗВИТОК КОРПУСНИХ ДОСЛІДЖЕНЬ

Назаров Богдан Олександрович,

аспірант кафедри загального мовознавства і германістики

Національного педагогічного університету імені М. П. Драгоманова

Корпусна лінгвістика є одним із напрямів сучасних прикладних лінгвістичних досліджень. Пройшовши складний шлях від майже повного неприйняття в 60-70-ті роки минулого століття до провідного напрямку науки про мову в сьогоденні, корпусна лінгвістика визнана самостійною галуззю прикладного мовознавства [1, с. 122; 2, с. 89].

На сьогодні спостерігається стрімкий розвиток корпусних студій у всьому світі, загальна проблематика яких розробляється у двох напрямках: 1) вивчення загальної теорії корпусної лінгвістики (Д. Байбер, Дж. Синклер, В. Тойберт), 2) розроблення загальних засад створення корпусів природних мов тощо (праці Б. Алтенберга, М. Баньки, У. Френсіса, Г. Кеннеді, Г. Ліча, М. Михайлова, С. Шарова та ін.).

Сьогодні лінгвістичні корпуси письмових чи усних текстів успішно використовуються для навчання іноземних мов та в лінгвістичній педагогіці. На базі корпусів формуються списки автентичної активної лексики студентів, частотні списки термінів для використання у професійних курсах [3, с. 231]. Корпусна лінгвістика має широкі перспективи не лише в рамках лінгвістики, а й в інших галузях гуманітарної науки, зокрема соціології, психології, педагогіці та ін. Перспективними є міждисциплінарні дослідження з вивчення когнітивних механізмів, соціальної визначеності та історичної зумовленості функціонування мовних одиниць та явищ на базі корпусів. Інноваціями в корпусній лінгвістиці є створення мультилінгвальних, паралельних та мультимодальних корпусів.

Корпусна лінгвістика є частиною методологічного інструментарію сучасної лінгвістики. Тому подальший розвиток корпусних досліджень потребує детального дослідження та є на часі в сучасній лінгвістиці.

ЛІТЕРАТУРА

1. Демська-Кульчицька О. Основи національного корпусу української мови: монографія. К.: Інститут української мови НАН України, 2005. 218 с.
2. Рыков В. В. Прагматически ориентированный корпус текстов // Тверской лингвистический меридиан. 1999. Вып. 3. С. 89–96.
3. Baker M. Corpora in translation studies. An overview and suggestions for future research // Target. 1995. No. 7 (2). P. 223–243.

СТРУКТУРА І ФУНКЦІОНАЛЬНИЙ СТАТУС АБСОЛЮТНОЇ КОНСТРУКЦІЇ В СУЧАСНИХ АНГЛІЙСЬКІЙ ТА ФРАНЦУЗЬКІЙ МОВАХ: КОРПУСНИЙ АНАЛІЗ

Папіжук Валентина Олександрівна,

кандидат педагогічних наук, доцент,

доцент кафедри міжкультурної комунікації та прикладної лінгвістики

Житомирського державного університету імені Івана Франка

Мельник Каріна Сергіївна, магістрант

Житомирського державного університету імені Івана Франка

Абсолютна конструкція (далі – АК) є особливою структурою вторинної предикації, що наявна в багатьох сучасних індоєвропейських мовах, зокрема в англійській і французькій. Це складна за своєю будовою та семантикою граматична конструкція, що має бінарну структуру та складається з іменної групи в ролі вторинного суб'єкта та з неособової форми дієслова чи недієслівної частини мови в ролі вторинного предикату. Вважають, що АК в сучасних англійській та французькій мовах мають спільне походження від латинського звороту із власним підметом *ablativus absolutus* [1, с. 115; 2, с. 51], але які в подальшому по-різному розвивалися у кожній з мов.

Зіставний аналіз структурно-морфологічної та функціональної варіативності конструкції в досліджуваних мовах здійснювався із залученням даних Британського національного корпусу (BNC-BYU) [3] та Французького Національного Центру текстових та лексичних ресурсів (Корпусів) (CNRTL) [4]. Результати проведеного аналізу свідчать, що досліджувані конструкції вирізняються низкою подібних структурних і функціональних характеристик, а саме: переважання дієслівних АК над недієслівними, засоби вираження недієслівної АК (прикметникова або іменникова фраза), засоби вираження суб'єкта АК та їх кількісні характеристики (загальний іменник або особовий

займенник, з суттєвим переважанням іменника). Відмінними в англійській та французькій мовах виявилися варіанти вираження предикатів дієслівної АК, зокрема, в англійській мові – це Participle I, Participle II та to-Infinitive, а у французькій – це Participe та, рідше, Gérondif.

ЛІТЕРАТУРА

1. Алексеева В. А. Функционирование беспредложных абсолютных конструкций в английском дискурсе // Вестник МГЛУ. 2011. № 58 С. 113–121.
2. Жуковська В. В. Абсолютна конструкція в англійській мові: сутність, походження, функціонування // Studia Philologica (Філологічні студії): зб. наук. пр. К.: КУ імені Б. Грінченка, 2013. С. 48–53.
3. British National Corpus (BNC). URL: <https://corpus.byu.edu/bnc/>
4. Corpus de Référence du Français parlé. URL: <http://sites.univprovence.fr/delic/corpus/index.html>

ТУРЕЦЬКОМОВНІ САКРАЛЬНІ ТЕКСТИ ЯК ОБ'ЄКТ КОРПУСНОЇ ЛІНГВІСТИКИ

Підвойний Володимир Миколайович,

кандидат філологічних наук, доцент,

доцент кафедри східних мов та методики викладання східних мов

Національного педагогічного університету імені М. П. Драгоманова

Утилітарність корпусної лінгвістики (далі – КЛ) профілює уведення в навчальний контент текстових блоків різних типів, ступенів складності, дидактичних сценаріїв і стилістичної маркованості. Гетерогенність текстової бази (турецька – Т, українська – У) детермінує “паралельність” як базову категорію КЛ у функціонально-теоретичному вимірі (ТПП, функціональна стилістика, порівняльна лексикологія). Сегментація, морфологічне моделювання та допрезентаційне електронне опрацювання сакральних текстових масивів в умовах культурно неспорідненої ситуації (модель Т→У) реалізується профільними завдання навчального процесу. Сакральні тексти (турецькомовний варіант Корану, гіпертекстовий характер сур, хадисний текстовий корпус, збірки дидактико-моралізаторських наративів) виконують важливу культурно-детермінаційну роль і функцію соціокультурного феномена, мають смислоформує значення [1, с. 56]. Термінологічне опрацювання турецькомовних сакральних текстів (рівень метафорики, ономазіологічні аспекти, експлікація категорії прецедентності) у зіставному форматі профілює

доцільність використання паралельних текстів для підвищення дидактичного ефекту і об'єктивує дослідницьку актуальність компаративних студій (модель $U \leftrightarrow T$) [2, с. 67–71].

Використання паралельних текстових корпусів (модель $T \leftrightarrow U$) у навчальному процесі уможливить еластичність архітекτονіки програмного контенту профільних дисциплін та полегшить процес сприйняття когнітивних моделей, релевантних для текстових презентацій турецької лінгвокультури.

ЛІТЕРАТУРА

1. Görgün T. Anlam ve yorum: Dini metinlerin anlaşması ve yorumlanması. İstanbul: Külliyyat yayınları, 2016. 224 s.
2. Yeşilyurt T. Söz'ün anlamı: Teolojik dilin paradoksal görünümü üzerine bir çözümleme. İstanbul: İnsan yayınları, 2016. 238 s.

ПЕРСПЕКТИВИ ПОЄДНАННЯ КВАНТИТАТИВНОГО І КОРПУСНОГО МЕТОДІВ (на матеріалі кількісно-квантитативного обчислення ергонімів як знаків соціокультурного простору України)

*Позніхіренко Юлія Іванівна,
старший викладач кафедри загального мовознавства і германістики
Національного педагогічного університету імені М. П. Драгоманова*

Соціокультурний простір України характеризується особливим історично сформованим синтезом різнорідних етнокультурних традицій, а також соціокультурними процесами, які дозволяють говорити про його цілісність і створюють можливість для лінгвістичного вивчення, об'єктом якого є назви закладів харчування, що стали знаками цього простору.

Відповідно до номінативного розмаїття назв закладів харчування України за допомогою квантитативного методу було укладено ергонімійну карту.

Ергонімійна карта назв закладів харчування України демонструє процентне співвідношення назв закладів харчування українською мовою та іншими мовами. Процентне співвідношення було виявлено шляхом обчислення в таблиці excel за формулою $x = 100 \cdot n / y$ (де x – процент (%) назв закладів харчування будь-якої групи обраного регіону (групи: назви укр. М., назв інш. М. (англ., рос., фр., ісп., італ., груз. Та ін.) Чи назви гібр.); y – загальна кількість (од.) Назв мереж харчування обраного регіону; n – кількість (од.) Назв закладів харчування будь-

якої групи обраного регіону (групи: назви укр. М., назв інш. М. (англ., рос., фр., ісп., італ., груз. Та ін.) Чи назви гібр.)

Наприклад: по Київській області у – загальна кількість (од.) Назв мереж харчування становить 643, з них п – кількість (од.) Назв закладів харчування українською мовою – 109, п – кількість (од.) Назв закладів харчування іншими мовами (російською, англійською, французькою, іспанською, італійською та грузинською мовами) – 509, п – кількість (од.) Назв-гібридів – 25. Щоб знайти х – процент (%) назв закладів харчування українською мовою підставимо значення у формулу $x = 100 \cdot 109 / 643 = 17\%$, х – процент (%) назв закладів харчування іншими мовами (російською, англійською, французькою, іспанською, італійською та грузинською мовами) підставимо значення у формулу $x = 100 \cdot 509 / 643 = 79\%$, х – процент (%) назв-гібридів закладів харчування підставимо значення у формулі $x = 100 \cdot 25 / 643 = 4\%$. Ці дані було нанесено на карту України.

На західній Україні виявлено найбільшу кількість назв закладів харчування українською мовою, а на сході України популярними є назви закладів харчування переважно російською та англійською мовами.

Назви-гібриди більшою мірою зафіксовано в містах, популярних серед туристів із різних країн: м. Київ, м. Одеса, м. Львів та м. Івано-Франківськ. Для подальшого обчислення специфіки функціонування цих мовних знаків пропонуємо залучити матеріал Національного корпусу української мови.

ІНСТРУМЕНТАРІЙ ГУРТУ R2U ДЛЯ СТВОРЕННЯ УКРАЇНСЬКИХ КОРПУСІВ

Старко Василь Феодосійович,

кандидат філологічних наук, доцент,

докторант кафедри прикладної лінгвістики

Східноєвропейського національного університету імені Лесі Українки

Донедавна задум створити новий український корпус неunikно впирався в проблему браку комп'ютерних засобів для автоматичного опрацювання тексту. Цю прогалину заповнив гурт r2u, що розробляє низку проектів для української мови [4]. Створено й удоступнено онлайн набір засобів [5] для, зокрема, таких операцій: розбиття тексту на речення, токенізація, лематизація, морфологічний аналіз словоформ. Найважливіший компонент, без якого ці операції були б неможливі, – Великий електронний словник української мови (ВЕСУМ) [1],

поточна версія 4.1 якого налічує 332 тис. лем. ВЕСУМ — найбільший словник такого типу, й гурт r2u постійно поповнює його новою лексикою. З використанням вказаного інструментарію вже створено корпус ГРАК [6] та корпуси групи lang-uk [2]. Триває робота над корпусом БРУК [3] і модулем автоматичного знімання омонімії. Гурт r2u відкритий до співпраці над новими українськими корпусами й спільного розвитку NLP-технологій для української мови.

ЛІТЕРАТУРА

1. ВЕСУМ. URL: <https://r2u.org.ua/vesum/>
2. Корпуси групи lang-uk. URL: <http://lang.org.ua/uk/corpora/>
3. Корпусна група БрУК. URL: <https://r2u.org.ua/corpus>
4. Старко В. Комп'ютерні лінгвістичні проекти гурту r2u: стан та застосування // Українська мова. 2017. № 3. С. 86–100.
5. Утиліти NLP гурту r2u. URL: https://github.com/brown-uk/nlp_uk
6. Шведова М. та ін. ГРАК: Генеральний регіонально анотований корпус української мови. URL: <http://uacorporus.org/>

СЕМІОМЕТРИЧНИЙ АНАЛІЗ АКСІОКОНЦЕПТОСФЕР ІЗ ЗАЛУЧЕННЯМ МОЖЛИВОСТЕЙ ПОШУКОВИХ СИСТЕМ КОРПУСІВ

*Стефанова Наталія Олександрівна,
кандидат філологічних наук, доцент,
доцент кафедри англійської філології*

Національного педагогічного університету імені М. П. Драгоманова

Вивчення співвідношення ментальних і мовних структур відкрило нові перспективи для методик опрацювання мовних явищ, однією з яких є семіометрія ціннісних смислів (Е. Ф. Серебренникова та ін. [1]) концептосфер представників різних етнокультур як дослідницька методологічна процедура.

Попередньо розроблюваний семіометричний аналіз лінгвокультурних концептів як специфічно організованих компонентів аксіоконцептосфери кожного етносу спрямований на вимірювання етнічних констант і культурно-ціннісних домінант цього ментального утворення за допомогою зіставно-параметричного методу (Й.А. Стернін, В.Т. Тітов та ін.), а також на основі можливостей пошукових систем англо- та україномовних корпусів.

Суть цієї комплексної методологічної процедури полягає у розробці й застосуванні формалізованих параметрів аналізу етнічних констант і культурно-ціннісних домінант концептосфер носіїв англійської і української мов, які сприятимуть виявленню національної специфіки кожної з мов та ціннісних пріоритетів представників обох етнокультур на певних етапах розвитку їх соціумів.

Для того, щоб проаналізувати значення імен та вербалізаторів лінгвокультурних концептів, необхідно реконструювати їхню семантичну структуру за п'ятьма параметрами: 1) етимологічним (аналіз етимонів імен лінгвокультурних концептів, що утворюють етнічні константи); 2) функціонально-семантичним (ідентифікація й аналіз ціннісного компонента семантики у семантичній структурі вербалізаторів лінгвокультурних концептів з виявленням домінантних смислів), 2) синтагматичним (аналіз фразеологізмів – вербалізаторів лінгвокультурних концептів, оскільки за допомогою фразеології маркуються концепти культури конкретного етносу, і відбір аксіологічно-маркованих слів, що є найбільш частотними у складі фразеосполук, дасть змогу виявити ядро синтагматично активних і культурно значимих слів), 3) епідигматичним (аналіз дериваційної активності ціннісно-маркованих слів, що сприятиме виявленню в одного слова кількох ціннісних значень), 4) парадигматичним (аналіз наявності синонімічних рядів).

Обчислення показників за кожним параметром буде здійснюватися переважно на основі можливостей пошукових систем англо- та україномовних корпусів. “Характеристика того чи іншого мовного або когнітивного явища в рамках відповідного параметра буде зафіксована у вигляді певного індексу, представленого в абсолютних числах або відсотках. Зіставлення однойменних індексів в англійській та українській мовах дасть підстави зробити припущення про наявність чи відсутність національної специфіки за даним параметром, а також про ступінь її вияву” [2, с. 3].

ЛІТЕРАТУРА

1. Лингвистика и аксиология: этносемиометрия ценностных смыслов: коллективная монография / отв. ред. Л. Г. Викулова. М.: ТЕЗАУРУС, 2011. 352 с.
2. Стернин И. А., Стернина М. А. Сопоставительно-параметрический метод: новый этап развития // Сопоставительные исследования. 2010. Воронеж: Истоки, 2010. С. 3–9.
3. Титов В. Т. Выделение параметрического ядра лексики: методические указания. Воронеж: Изд-во Воронежского университета, 2006. 55 с.

ПОШУКОВА СИСТЕМА І КОРПУС ЯК ІНСТРУМЕНТ ДОСЛІДЖЕННЯ СЕМАНТИКИ СЛІВ

*Таран Алла Анатоліївна,
кандидат філологічних наук,
доцент кафедри українського мовознавства і прикладної лінгвістики
Черкаського національного університету імені Богдана Хмельницького*

Поява Інтернету й створення корпусів текстів різних національних мов значно спростили підготовчий, але надзвичайно значущий для подальшого розгортання дослідження етап. Досліднику достатньо ввести слово в пошукову систему відповідного корпусу, щоб отримати повний індекс його словоформ у найближчих і розширених контекстах із відповідною адресацією. Створення емпіричної бази дослідження в такий спосіб забезпечує мінімум витрат зусиль і часу на зібрання даних, обсяг яких водночас суттєво збільшиться. З одного боку, це спростило й полегшило, але, з іншого боку, ускладнило процес аналізу й осмислення такої кількості прикладів.

Поява нового значення слова зумовлена контекстом функціонування слова й внутрішніми семантичними можливостями, адже стрижень значення – комплекс ознак чогось – залишається в усіх словосполуках. Ми протестували пошукові системи формування нового значення слова *синдром*. ГУГЛ за 52 секунди запропонував 19 000 000 результатів пошуку. Як зазначав Сидні Лендау, загальна проблема, що постає під час роботи з цитатною картотекою, – це брак прикладів конкретного слововжитку; загальна проблема під час роботи з корпусом – надмір прикладів (“*влучень*” *hits*) [2, с. 27]. На перших 10–15 сторінках нових контекстів функціонування слова немає. Можливості розширеного пошуку ГУГЛ передбачають уведення словосполуки: тому *донецький синдром* засвідчив 35 результатів пошуку за 23 с. Як бачимо, особливого полегшення пошуку немає, бо, виходить, треба знати, що шукати. Корпус текстів української мови (mova.info) у підкорпусі “публіцистичні тексти” засвідчує окремі словосполуки на зразок *трамвайний синдром*.

Для порівняння протестували базу даних Neomat (Інститут чеської мови АН Чеської Республіки), де не подано тлумачення нового слова *синдром*, однак контекст уживання лексеми відображає її нову лексичну або синтаксичну сполучуваність, що дає змогу сформуванню дефініцію слова. У Чеському національному корпусі на основі формул [tag="A.*"] [lemma="syndrom"] та [lemma="syndrom"] [tag="N.*"] визначаємо контексти вживання лексеми.

Як бачимо, важливою особливістю даних Інтернету як лінгвістичного джерела є можливість безпосереднього звернення до контекстів, у яких зафіксовані вживання потрібного слова. Це важливо для переходу від кількісних характеристик до якісних, від абстрактного інваріанта, репрезентованого в квантитативному коефіцієнті слова, до його семантичних, стилістичних і інших варіантів. Проте на сьогодні, на жаль, пошукові системи потребують удосконалення, адже не відомо, як упорядкувати результати пошуку, було би добре, якби функціонували фільтри того, що ми шукаємо. Або ж мовознавцям треба створювати свої, спеціалізовані пошукові системи, зорієнтовані на виконання наукових завдань.

ЛІТЕРАТУРА

1. Карпіловська Є. А. Вступ до прикладної лінгвістики: комп'ютерна лінгвістика: підручник. Донецьк: ТОВ "Юго-Восток, ЛТД", 2006. С. 153–165.
2. Сидні І. Лендау Словники: мистецтво та ремесло лексикографії. К.: КІС, 2012. 480 с.

СИСТЕМА "АРХІВНА КАРТОТЕКА УКРАЇНСЬКОЇ МОВИ" ЯК ДОСЛІДНИЦЬКИЙ КОРПУС

*Тищенко Оксана Миколаївна,
кандидат філологічних наук,
старший науковий співробітник*

Інституту української мови Національної Академії наук України

Архівна картотека (далі – АК) – це лексико-ілюстративний матеріал Російсько-українського словника, за ред. А. Кримського та С. Єфремова (1924–1933), упорядкований через призму філологічної думки відбиток мовної ситуації ХІХ–поч. ХХ ст. [1]. Зацифрування АК й створення системи "Архівна картотека української мови", окрім збереження й удоступнення цієї пам'ятки української мови й мовознавчої думки, передбачає формування дослідницького ілюстративного корпусу з формалізованою мега-, макро- і мікроструктурою, а також із можливостями взаємодії з іншими системами опрацювання української мови.

Корпус ілюстрацій може не лише слугувати фактичним підґрунтям для укладання різного типу словників, а й становити самостійний об'єкт

лінгвістичних студій різного плану. Система АК загалом становить корпус інформації про українську мову з довідковими та дослідницькими функціями.

ЛІТЕРАТУРА

1. Тищенко О. М. Архівна картотека як лексико-ілюстративна база “Російсько-українського словника” / за ред. А. Ю. Кримського, С. О. Єфремова // Мікро- та макроструктура архівної картотеки / Українська мова. 2016. № 3. С. 57–78.

КОРПУСНІ МЕТОДИ У ДОСЛІДЖЕННЯХ ГЕНДЕРНОЇ ЛІНГВІСТИКИ

*Ткачик Олена Володимирівна,
кандидат філологічних наук,
доцент кафедри теорії, практики та перекладу англійської мови
Національного технічного університету України
“Київський політехнічний інститут імені Ігоря Сікорського”*

Комп’ютерна епоха надає нові можливості для сучасних гендерних досліджень, додаючи до культурологічного, соціолінгвістичного, психолінгвістичного, лінгвокогнітивного та ін. наукових підходів якісно та кількісно нові результати. У західних дослідженнях гендеру та мови корпусні методи є досить популярними, але для вітчизняної лінгвістики вони ще доволі нові та мало застосовані [2].

Сучасні українські гендерні дослідження з використанням методів корпусної лінгвістики нечисленні і проводяться, зокрема, проф. О. П. Левченко, яка відмічає, що “використання значних масивів даних є важливою умовою отримання релевантних результатів” [1]. Базуючись на дослідницькому корпусі з уривків художньої прози українських письменниць та письменників, авторка виявляє типові порівняльні стратегії різних статей.

Звертаючись до досвіду використання корпусних методів для дослідження гендеру в мові західними лінгвістами, зазначимо роботу Пола Бейкера “Introduction: Virtual Special Issue of Gender and Language on corpus approaches” [3], в якій він робить огляд таких джерел: Пол Бейкер, Джонсон і Енслін (2007), Хартіс-Чорний та Сіле (2009), Холмгрен (2009), Бейкер (2010) та Кінг (2011), де автори використовували методіку корпусної лінгвістики в своїх дослідженнях, поєднуючи якісні та кількісні підходи.

ЛІТЕРАТУРА

1. Левченко О. П. Порівняльні стратегії в сучасній українській художній прозі (корпуснобазований підхід) // Граматичні студії: зб. наук. праць / ред. А. П. Загнітко. Вінниця: ТОВ “Нілан-ЛТД”, 2015. С. 194–201.
2. Ткачик О. В. Використання корпусних методів у гендерних дослідженнях // Наукові записки НУ “Острозька академія”. Серія Філологічна: зб. наук. пр. Острог: В-во нац. ун-ту “Острозька академія”, 2018. Вип. 69.
3. Baker P. Introduction: Virtual Special Issue of Gender and Language on corpus approaches // Gender and Language. 2013. URL: <https://journals.equinoxpub.com/index.php/GL/13506>

ОБЧИСЛЕННЯ ЧАСТОТИ ВЖИВАННЯ ТЕРМІНІВ TRADEMARK І BRAND NAME В АНГЛОМОВНИХ КОРПУСАХ

*Толчеева Тетяна Станіславівна,
доктор філологічних наук, професор,
професор кафедри загального мовознавства і германістики
Національного педагогічного університету імені М. П. Драгоманова*

Терміни, що належать до сфери комерційної номінації, серед яких такі англomовні гіпероніми, як “trademark” і “brand name”, продовжують викликати дискусії у плані не лише їх термінотворчого потенціалу, а й юридичного статусу, і, головне, – пріоритетності вживання на тлі суміжних утворень.

У праці І.С. Березіна “Маркетинговий аналіз. Рынок. Фирма. Товар. Продвижение” термін “торгова марка” витлумачують як словесну кальку з англomовного словосполучення “trademark” і розуміють назву, під якою фірма, підприємство рекламує і продає свою продукцію. Але при цьому, на думку дослідника, цей термін активно вживається споживачами і власниками фірм і комерційних установ, тоді як юридична система послуговується поняттями “товарний знак” і “знак обслуговування” [1], що підлягають правовій охороні.

Водночас думки фахівців у цій сфері не є однозначними. Так, М. Портер термін “торгова марка” ототожнює з терміном “словесний товарний знак” [3], Г. Чармессон – із терміном “комерційна назва” [4], а Ю. С. Бернадська – з терміном “ім'я брэнда” (brand name) [2].

Для того, щоб переконатися в тому, наскільки тотожними є англomовні терміни “trademark” і “brand name” та чи можна вважати повним еквівалентом

їхній україномовний термін-відповідник “товарний знак”, пропонуємо скористатися даними англomовних корпусів і визначити як частотну пріоритетність уживання кожного з них, так і обґрунтувати коректність вибору правничою системою україномовного відповідника.

ЛІТЕРАТУРА

1. Березин И. С. Маркетинговый анализ. Рынок. Фирма. Товар. Продвижение. 3-е изд., испр. и доп. М.: Вершина, 2008. 369 с.
2. Бернадская Ю. С. Текст в рекламе: учебная помощь. М.: ЮНИТИ-ДАНА, 2008. 288 с.
3. Портер М. Е. Конкурентное преимущество: как достичь высокого результата и обеспечить его устойчивость / пер. с англ. М.: Альпина Бизнес Букс, 2005. 714 с.
4. Чармэссон Г. Торговая марка: как создать имя, которое принесет миллионы. СПб.: Изд-во “Питер”, 1999. 224 с.

ГРАК: ПРИНЦИПИ І ПРОБЛЕМИ ПОБУДОВИ КОРПУСУ

*Шведова Марія Олексіївна,
кандидат філологічних наук,
докторант кафедри російської мови і літератури
Київського національного лінгвістичного університету*

Генеральний регіонально анотований корпус української мови (ГРАК) – це велика, репрезентативна, структурована колекція текстів українською мовою у супроводі програми, яка дозволяє будувати на базі корпусу власні підкорпуси, шукати слова, граматичні форми та їх сполучення, а також обробляти результати пошуку, сортувати, робити збалансовані вибірки і одержувати різну статистичну інформацію.

Корпус охоплює період з 1818 по 2018 р. і містить понад 25 тисяч текстів близько 3850 авторів.

Напрями та жанри. Приблизно 60% корпусу становлять художні тексти. Деякі жанри художніх текстів виділені окремо: дитячі, фольклорні, драматичні твори та сценарії. З нехудожніх текстів великий підкорпус становлять публіцистичні тексти, зокрема колекції газет 1888–1918 рр., 1919–1943 рр., сучасні газети різних регіонів, а також тексти з інформаційних сайтів в Інтернеті.

В корпусі також розміщено мемуари, листи та щоденники, зокрема чималий корпус текстів Facebook, в якому представлені блоги людей з усіх областей України та діаспори. Інший великий підкорпус складають наукові та навчальні тексти: монографії, дисертації, наукові статті, підручники, окремо виділено значні за обсягом підкорпуси наукової літератури з історії, етнографії, філософії, права. Є окремий підкорпус релігійних текстів, до якого зокрема увійшли два українські переклади Біблії. Невеликі підкорпуси творів усних жанрів: промови та інтерв'ю. Також до корпусу залучено деякі словники, де є фразові приклади і фразеологія, зокрема “Словарь української мови” Б. Грінченка і “Російсько-український словник сталих виразів” І. Виргана, М. Пилинської. За допомогою інструментів корпусу в словниках можна шукати не тільки слова, а й лексико-граматичні моделі у складі прикладів та фразеологічних виразів.

Перекладні тексти. Приблизно третина текстів в корпусі – переклади. Корпус містить переклади з 38 мов, найбільше – з англійської та російської. Треба зауважити, що інколи українські переклади зроблено не з мови оригіналу, а з російського перекладу, не в усіх виданнях це зазначено. У випадках, коли точно відомо, що переклад зроблено з російського перекладу, в корпусі вказується, що текст перекладено з російської мови. У випадках, коли є серйозні підстави вважати, що переклад зроблено зі значним використанням російського перекладу, в корпусі вказано дві мови оригіналу (EN_RU). Але не всі перекладні тексти в корпусі було проаналізовано з точки зору можливого впливу іншого перекладу.

У складі корпусу є показова колекція текстів 20–30-х рр. ХХ ст., в якій представлені письменники і публіцисти того часу, у тому числі маловідомі, твори яких перевидають зараз.

Сучасні тексти. Більша частина корпусу – тексти післявоєнного радянського періоду і сучасні, які доступні в Інтернеті.

Розробники:

Концепція корпусу, підготовка текстів: Марія Шведова (corpus.textiv@gmail.com)

Наукові консультанти проекту: Рупрехт фон Вальденфельс, Дмитро Січінава, Василь Старко, Михайло Назаренко

Технічна реалізація: Рупрехт фон Вальденфельс, Міхал Возняк, Сергій Яригін, Михайло Крук, Андрій Рисін

На основі попередньої роботи Андрія Рисіна та гурту r2u

Просимо посилатися на ГРАК:

М. Шведова, Р. фон Вальденфельс, С. Яригін, М. Крук, А. Рисін, М. Возняк. Генеральний регіонально анотований корпус української мови (ГРАК). Київ, Осло, Йена, 2017–2018. URL: <http://uacorpus.org>.

СПЕЦІАЛІЗОВАНИЙ КОРПУС НАУКОВИХ ТЕКСТІВ ЯК ПОШУКОВА СИСТЕМА ДЛЯ ЛІНГВІСТИЧНИХ ДОСЛІДЖЕНЬ

*Щербина Світлана Миколаївна,
кандидат філологічних наук, доцент,
докторант кафедри загального мовознавства і германістики
Національного педагогічного університету імені М. П. Драгоманова*

Створення та використання спеціалізованих корпусів наукових текстів в освітній та професійній сферах є актуальним завданням сучасних студій із прикладної лінгвістики. Засновниками цієї сфери вважають англосаксонську лінгвістичну традицію, результатом якої стало створення спеціалізованих корпусів англійської мови як зразків усного й писемного академічного мовлення. Зокрема, це такі корпуси: CANCODE (Cambridge and Nottingham Corpus of Discourse in English – Кембриджський і Нотингемський письмовий корпус англійської мови) MiCASE (Michigan Corpus of Academic Spoken English – Мічиганський корпус академічної англійської мови), LIBELCASE (Corpus of Spoken Academic English – Лімерік-Белфас корпус усної академічної англійської мови).

В Україні також намітилися серйозні перспективи щодо створення спеціалізованих корпусів наукових текстів. Принаймні, можемо говорити про три проекти. Одним із таких проектів є науковий проект факультету іноземної філології Національного педагогічного університету імені М.П. Драгоманова “Германо-романо-слов’янський корпус аутентичних наукових текстів з лінгвоантропогенезу: розробка технологій нового покоління”, започаткований у січні 2017 р. під керівництвом доктора філологічних наук, професора Алли Валер’янівни Корольової. Основним завданням проекту є розробка різних типів розмітки для систематизації наукових текстів з проблематики лінгвоантропогенезу.

Заслуговує на увагу і проект корпусу анотацій наукових статей із комп’ютерної лінгвістики, створений на базі Київського національного лінгвістичного університету. Корпус складається з анотацій наукових статей,

що мають метатекстову розмітку, опублікованих у журналі “Computational linguistics”, який є офіційним виданням міжнародної асоціації комп’ютерної лінгвістики ACL (Association for Computational Linguistics). Загальний обсяг створеного корпусу – 1475 текстів анотацій, які містять близько 150 000 слововживань [1, с. 32]. Ще один проект корпусу текстів спеціалізованого наукового мовлення напряму “Філологія” розроблюваний дослідниками з Національного університету “Острозька академія” (керівник проекту – Коцюк Леся Миколаївна). Метою проекту є пошук та впорядкування у вигляді корпусу наукових статей, надрукованих у фахових збірниках напряму “Філологія” [2].

Таким чином, на сьогодні в Україні простежується позитивна динаміка щодо застосування корпусних підходів до вивчення наукового мовлення, яка спирається на традиції англосаксонської корпусної лінгвістики. Тому є всі підстави говорити про новий етап розвитку корпусної лінгвоукраїністики, яка зорієнтована на створення корпусів фахових текстів.

ЛІТЕРАТУРА

1. Коломієць В., Орел В. Корпус анотацій наукових статей із комп’ютерної лінгвістики // Комп’ютерна лінгвістика: сучасне і майбутнє. Матеріали Міжнародної науково-практичної конференції. К.: Вид. центр КНЛУ, 2012. С. 32–34.
2. Проект корпусу текстів спеціалізованого наукового мовлення напрямку “Філологія”. URL: <https://www.oa.edu.ua>

ЗМІСТ

Корольова А. В.	Корпусна лінгвістика та її майбутнє.	4
-----------------	--	---

Наукова дискусія

Карпіловська Є. А.	Академічна й навчально-методична інтеграція корпусних досліджень в Україні	6
Дарчук Н. П.	Технологія розроблення корпусів текстів для навчальних цілей	7
Жуковська В. В.	Технологія роботи з корпусами у процесі вивчення іноземних мов	8

Майстер-клас

КОРПУС ТЕКСТІВ ЯК ОСОБЛИВИЙ ТИП

ІНФОРМАЦІЙНО-ПОШУКОВОЇ, НАВЧАЛЬНОЇ ТА ДОСЛІДНИЦЬКОЇ СИСТЕМ

Зубань О. М.	Електронні частотні словники – інструмент систематизації матеріалу лінгвістичного дослідження	9
	Методологія укладання паралельного корпусу текстів	9

Тези для обговорення

проблемних питань наукової дискусії

Андрушенко О. Ю, Ярощук Н. А.	Епістемічні адверби в ранньоновоанглійській мові (дослідження на базі Corpus of Historical American English)	11
Андрушенко О. Ю., Тищенко Р. В. Балабан О. О.	Корпусне дослідження локативу <i>there</i> в американському варіанті англійської мови XIX ст. Дослідження когнітивно-семантичних універсалій та паралельні корпуси текстів	12 13
Бобер Н. М.	Основні принципи укладання Британського національного корпусу та їхній трансфер у методологію розробки інших корпусів	14
Брик М. М.	Залучення даних корпусів при дослідженні художньої картини світу.	15
Дем'янчук Ю. І.	Класифікація правових текстів та термінів ООН, НАТО і СОР відповідно до норм мови спеціального призначення	16
Дюндик О. А.	Паралельний корпус творів Альбера Камю у перекладі Анатолія Перепаді	17

Железко А. М.	Перспективи розроблення спеціалізованого корпусу на основі бази даних "Лінгвоантропогенез"	18
Жуковська В. В.	Квантитативно-корпусний підхід до аналізу семантики граматичної конструкції	20
Зернецький П. В.	Структурно-прагматичний аспект текстів дебатів британського парламенту	21
Зубань О. М.	Автоматичне укладання частотних морфемних словників (на матеріалі корпусу української мови) .	22
Зубань О. М.	Паралельний корпус текстів: структура, аспекти параметризації та технології укладання	23
Івашкевич Л. С.	Ефективний пошук по корпусу за допомогою регулярних виразів	23
Капранов Я. В.	Міжнародна база даних "Вавилонская башня" як пошукова система для вивчення ностратичних етимологій	24
Карлова В. О.	Моделювання як когнітивний метод корпусної лінгвістики	25
Карпіловська Є. А.	Корпус серед інших інструментів опрацювання мови	26
Качанов В. І.	Корпусна методологія у вивченні іноземних мов ..	27
Комаренко В. В.	Соціолінгвістичний аспект розгляду поняття "Предметна галузь" (на основі корпусу MICASE) ..	28
Корольова А. В.	База даних "Лінгвоантропогенез" як концептуальна платформа для створення германо-романо-слов'янського корпусу аутентичних наукових текстів	29
Кравцова Ю. В.	Проблемы создания метафорического корпуса ...	30
Крувенко А.	A Keyword Approach to Comparable Sub-Corpora ..	31
Леміш Н. Є.	Розмітка паралельного корпусу	32
Макарук Л. Л.	Специфіка мультимодальних корпусів	33
Маслова Т. Б.	Вимоги до укладання спеціалізованих мовних корпусів	34
Матвеева С. А.	Токенізація як спосіб обробки корпусного тексту ..	35
Назаров Б. О.	Перспектива і подальший розвиток корпусних досліджень	36
Папіжук В. О., Мельник К. С.	Структура і функціональний статус абсолютної конструкції в сучасних англійській та французькій мовах: корпусний аналіз	37
Підвойний В. М.	Турецькомовні сакральні тексти як об'єкт корпусної лінгвістики	38

Позніхіренко Ю. І.	Перспективи поєднання квантитативного і корпусного методів (на матеріалі кількісно-квантитативного обчислення ергонімів як знаків соціокультурного простору України)	39
Старко В. Ф.	Інструментарій гурту r2u для створення українських корпусів	40
Стефанова Н. О.	Семіотричний аналіз аксіоконцептосфер із залученням можливостей пошукових систем корпусів	41
Таран А. А.	Пошукова система і корпус як інструмент дослідження семантики слів	43
Тищенко О. М.	Система “Архівна картотека української мови” як дослідницький корпус	44
Ткачик О. В.	Корпусні методи у дослідженнях гендерної лінгвістики	45
Толчєєва Т. С.	Обчислення частоти вживання термінів <i>trademark</i> і <i>brand name</i> в англомовних корпусах	46
Шведова М. О.	ГРАК: принципи і проблеми побудови корпусу . . .	47
Щербина С. М.	Спеціалізований корпус наукових текстів як пошукова система для лінгвістичних досліджень	49