

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Львівський національний університет імені Івана Франка

Філологічний факультет

Катедра загального мовознавства

Затверджено

на засіданні катедри
загального мовознавства
філологічного факультету
Львівського національного університету імені
Івана Франка
(протокол № 1 від 29 серпня 2023 р.)

Завідувач катедри  Ф.С. Бацевич

Силабус з навчальної дисципліни

«Корпусна лінгвістика та аналіз тексту»,

що викладається в межах ОПП «Українська мова та література»,
«Літературна творчість», «Фольклористика», «Прикладна лінгвістика»

другого (магістерського) рівня вищої освіти
для здобувачів зі спеціальності **035 Філологія,**

спеціалізацій: 035.01 Українська мова та література,
035.09 Фольклористика,
035.10 Прикладна лінгвістика

Львів 2023 р.

Назва дисципліни	Корпусна лінгвістика та аналіз тексту
Адреса викладання дисципліни	Львівський національний університет ім. Івана Франка м. Львів, вул. Університетська, 1
Факультет та катедра, за якою закріплена дисципліна	Філологічний факультет, катедра загального мовознавства
Галузь знань, шифр та назва спеціальності	Галузь знань – 03 Гуманітарні науки. Спеціальність – 035 Філологія. Спеціалізація – 035.01 Українська мова та література; 035.09 Фольклористика, 035.10 Прикладна лінгвістика
Викладач дисципліни	професор катедри загального мовознавства Бук Соломія Несторівна, докт. філол. наук, доцент
Контактна інформація викладача	solomiia.buk@lnu.edu.ua https://philology.lnu.edu.ua/employee/buk-solomiya
Консультації із цієї дисципліни відбуваються	Консультації відбуваються перед і після занять, а також за попередньою домовленістю за адресою: вул.Університетська, 1, ауд.343 або онлайн на платформі Zoom
Сторінка дисципліни	https://philology.lnu.edu.ua/course/korpusna-linhvistyka-ta-analiz-tekstu
Інформація про дисципліну	Дисципліна «Корпусна лінгвістика та аналізу тексту» належить до вибіркового навчального предмету для магістрантів першого курсу філологічного факультету зі спеціальності 035 Філологія, спеціалізацій 035.01 Українська мова та література; 035.09 Фольклористика; 035.10 Прикладна лінгвістика. Ця дисципліна викладається для освітніх програм «Українська мова та література», «Літературна творчість», «Фольклористика», «Прикладна лінгвістика» в 2-му семестрі в обсязі 4 кредити ЄКТС.
Коротка анотація дисципліни	Впродовж курсу розглянемо корпусну лінгвістику як інноваційний інформаційно-семіотичний напрям сучасного мовознавства, навчимося користуватися корпусами української та інших мов, застосовувати їхні результати в різних напрямках мовознавства і для аналізу тексту, а також створювати власні корпуси текстів. Базуючись на використанні величезних масивів емпіричного матеріалу та залученні комп'ютерних технологій для отримання об'єктивних висновків про функціонування мовних одиниць, електронні корпуси текстів швидко перетворилися на неоціненний лінгвістичний ресурс. Саме корпуси текстів стали незамінним інструментом в руках сучасних фахівців текстотворення й лінгвістичної експертизи, фольклористів, дослідників української мови та її викладачів.
Мета та цілі дисципліни	Мета вивчення дисципліни – сформувані у студентів-магістрантів знання про теоретичні і практичні можливості застосування корпусів для сучасної гуманітаристики, зокрема для текстотворення й лінгвістичної експертизи, фольклористики, дослідження української мови загалом. Основні завдання курсу: Розкрити можливості використання електронних корпусів текстів у гуманітаристиці загалом і прикладній

	<p>лінгвістиці зокрема. Ознайомити студентів з електронними корпусами текстів української та інших мов, їх типологією, складом, будовою, роззначенням, інтерфейсом, можливостями пошуку фонетичної, морфологічної, лексичної, синтаксичної, метамовної інформації, а також даних про час і місце написання тексту, гендер та регіон його автора тощо. Показати потенціал застосування корпусів текстів для розв'язання проблем аналізу тексту, гендерної, соціо-, комунікативної когнітивної, історичної, судової лінгвістики, літературознавства, фольклористики тощо.</p>
<p>Література для вивчення дисципліни</p>	<p>Основна:</p> <ul style="list-style-type: none"> • Дарчук Н. Корпусна лінгвістика: проблеми, методи, перспективи. Навчальна програма. Київ: Київський національний університет ім. Тараса Шевченка, 2013. • Демська-Кульчицька О. Текстовий корпус: ідея іншої форми. Київ: НаУКМА, 2011. • Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013. • Прикладна лінгвістика: навчальний посібник/ МОН України, Уманський держ. пед. ун.-т. імені Павла Тичини/ упорядник: Н. А. Цимбал. Умань: Візаві, 2019. • Селіванова О.О. Корпусна лінгвістика // Сучасна лінгвістика: напрями та проблеми: підручник. Полтава: Довкілля-К, 2008. • М. Шведова, Р. фон Вальденфельс. Регіональна розмітка в ГРАКу, великому референтному корпусі української мови: питання і проблеми [Regional Annotation within GRAC, a Large Reference Corpus of Ukrainian: Issues and Challenges]. CEUR Workshop Proceedings. Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021). Volume I: Main Conference. Kharkiv, Ukraine, April 22-23, 2021. • Широков та ін. 2018: Лінгвістично-інформаційні студії : праці Українського мовно-інформаційного фонду НАН України : у 5 т. / Широков В. А., Бугаков О. В., Заїка Н. М., Грязнухіна Т. О., Кригін М. О., Надутенко М. В., Рабулець О. Г., Сидорчук Н. М., Терещенко Л. О., Шевченко І. В., Шипнівська О. О., Ющенко С. С., Якименко К. М. Т. 4 : Корпусна та когнітивна лінгвістика. Київ. Український мовно-інформаційний фонд НАН України, 2018. <p>Додаткова:</p> <ul style="list-style-type: none"> • Демська-Кульчицька О. Основи національного корпусу української мови. Київ : Інститут української мови НАНУ, 2005. • Українська корпусна лінгвістика: сьогодення і перспективи : [інтерв'ю із канд. філолог. наук, доц. каф. укр. мови Нац. ун-ту "Києво-Могилянська академія" Орисьою Мар'янівною Демською] // Дивослово. 2011. № 10. С. 32-34. <https://dyvoslovo.com.ua/wp-content/uploads/2016/03/6-1011.pdf> • Людмила Дика, Марія Шведова. Історія і нормативний статус прийменника / префікса од- в українській мові нового часу. Slavia Orientalis, 2022, vol. LXXI, No 4. • Бук С. Архітектура польсько-українського та українсько-польського паралельного корпусу автоперекладів Івана Франка // Slavia Orientalis. 2012. Т. LXI. N 2.

- Бук С. Велика проза Івана Франка: електронний корпус, частотні словники та інші міждисциплінарні контексти : монографія / Львів : ЛНУ імені Івана Франка, 2021.
- Бук С. Учнівські корпуси в методиці викладання іноземної мови // Теорія і практика викладання української мови як іноземної. Львів : ЛНУ імені Івана Франка, 2007. Вип. 2.
- Čermák F. New generation corpus-based frequency dictionaries : The case of Czech // International Journal of Corpus Linguistics. 10. 2005.
- Divjak A., Sharoff S., and Erjavec T. Slavic Corpus and Computational Linguistics // Journal of the Slavic Linguistics Society. 2017. Vol. 25. N 2.
- Enery T. Mc., Xiao R., and Tono Y. Corpus-based language studies: an advanced resource book. Routledge, London, 2006.
- T. McEnery, A. Wilson Corpus Linguistics. An Introduction. 2-nd Edition. Edinburg: University Press, 2001.
- Lahjouji-Seppälä, M. – Rabus, A. – von Waldenfels, R. Ukrainian standard variants in the 20th century: stylometry to the rescue [on-line] <https://www.researchgate.net/publication/364332887_Ukrainian_standard_variants_in_the_20th_century_stylometry_to_the_rescue?fbclid=IwAR3bNDrrWLNUXtyPF0SZ5dID_NZaFYfVXCACULYtMOhjr5vqse2TY48u5c>
- Paquot M. & Gries S. T. A practical handbook of corpus linguistics. Berlin: Springer. 2020.
- Podstawy językoznawstwa korpusowego / Red. B. Lewandowska-Tomaszczyk. Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 2005. 306 s.
- Praktyczny przewodnik po korpusach języków słowiańskich. Red. Milena Hebal-Jeziarska. Warszawa: Zakład Graficzny Uniwersytetu Warszawskiego, 2014.<<http://slawistyka.uw.edu.pl/pl/2014/11/05/praktyczny-przewodnik-po-korpusach-jezykow-slowianskich-red-nauk-m-hebal-jeziarska/>>
- Sinclair J. EAGLES Preliminary recommendations on Corpus Typology. <www.ilc.cnr.it/EAGLES/pub/eagles/corpora/corpusyp.ps.gz>
- Viana V., & O'boyle A. Corpus linguistics for English for Academic Purposes. New York: Routledge. 2022.
- Zufferey S. Introduction to corpus linguistics. New Jersey: John Wiley & Sons. 2020.

Перелік деяких корпусів та інших електронних ресурсів:

- ГРАК: Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024 <www.uacorpora.org>.
- Український національний лінгвістичний корпус. Український мовно-інформаційний фонд <svc.ulif.org.ua>
- КУМ (Чтиво) = Корпус української мови <korpus.org.ua>
- Лабораторія української. Паралельні корпуси <https://mova.institute/%D0%BF%D0%B0%D1%80%D0%B0%D0%BB%D0%B5%D0%BB%D1%8C%D0%BD%D1%96_%D0%BA%D0%BE%D1%80%D0%BF%D1%83%D1%81%D0%B8>
- Корпус української мови Лепцизького університету / Corpus of Ukrainian by the University of Leipzig <https://corpora.uni-leipzig.de/en?corpusId=ukr_mixed_2014&fbclid=IwAR2n7pHaHOoOxkOQHaJsIkN9ynXCB6-dq6uQn_5eEKQ719h6YCxZQwh5ix8>
- Корпуси української мови на SketchEngine

	<p>https://www.sketchengine.eu/corpora-and-languages/ukrainian-text-corpora/</p> <ul style="list-style-type: none"> • Інструменти роботи зі системою SketchEngine (user guide) https://www.sketchengine.eu/guide/ • Національний корпус кримсько-татарської мови / National Corpus of the Crimean Tatar Language https://ctcorpus.org/index.php/en/ • Skovoroda Online Concordance / Онлайн конкорданс Сковороди http://www.arts.ualberta.ca/~ukr/skovoroda/NEW/index.php • Дарчук Н. П. Корпус текстів Тараса Шевченка як джерело для мовознавчих досліджень // Science and Education a New Dimension. Philology. V(32). Issue: 122. 2017. С. 33–37. • Дослідницький корпус Лесі Українки: лінгвістичні параметри // Науковий вісник Волинського національного університету імені Лесі Українки. Філологічні науки. Мовознавство. 2018. № 12. С. 30–36. • Загнітко А., Данилюк І., Ситар Г. Корпус текстів Юрія Шевельова: структура, функції, навігація // Мова: класичне – модерне – постмодерне. Київ: НУКМА. 2019. Вип. 5. С. 158–169. • Корпус текстів Богдана Лепкого / Кітаєва Віка (укладач). http://corpora.donnu.edu.ua/bonito/run.cgi/corp_info?corpname=kitaeva • Корпус текстів Михайла Коцюбинського / Бройко Н. (укладач). http://corpora.donnu.edu.ua/bonito/run.cgi/corp_info?corpname=brojk • Бук С., Ровенчак А. Онлайн-конкорданс роману Івана Франка «Перехресні стежки» http://ktf.franko.lviv.ua/users/andrij/science/Franko/concordance.html • British National Corpus http://www.natcorp.ox.ac.uk. • Český národní korpus http://ucnk.ff.cuni.cz • Korpus Języka Polskiego IPI PAN http://korpus.pl/index.php?page=welcome • Korpus Języka Polskiego Wydawnictwa Naukowego PWN https://sjp.pwn.pl/korpus • PAWUK = Polish Automatic Web corpus of Ukrainian language / Польський автоматичний вебкорпус української мови https://pawuk.ipipan.waw.pl/overview • PolUKR = polsko-ukraiński korpus równoległy / польсько-український паралельний корпус http://www.domeczek.pl/~polukr/index.php?option=welcome • Korpusomat https://korpusomat.eu/ • Laurence A. AntConc (Windows, MacOS, Linux). Build 4.1.4 https://www.laurenceanthony.net/software/antconc/releases/AntConc414/help.pdf The JRC-Acquis Multilingual Parallel Corpus / Корпус законодавства Європейського Союзу: Acquis Communautaire, що містить 231 мовну пару https://web.archive.org/web/20060619034515/http://langtech.jrc.it/JRC-Acquis.html • Von Waldenfels, Ruprecht and Meyer, Roland (2006-2023): ParaSol, a Corpus of Slavic and Other Languages https://www.parasolcorpus.org/Ursynow/
Тривалість курсу	Один семестр (2-й)

Обсяг курсу	<i>Денна форма здобуття освіти:</i> 120 годин, зокрема 48 год. аудиторних, з них – 16 лекційних, 32 практичних; 72 год. – самостійна робота. 4 кредити ЄКТС
Очікувані результати навчання	<p>У результаті вивчення дисципліни магістранти будуть знати:</p> <ul style="list-style-type: none"> – об'єкт і предмет дослідження корпусної лінгвістики. – передумови виникнення та етапи становлення корпусної лінгвістики як науки. – теоретичне підґрунтя корпусної лінгвістики. – основні напрямки корпусної лінгвістики. – історію становлення корпусної лінгвістики – корпусні дослідження в Україні та за кордоном – принципи створення корпусу текстів – типологію та основні сфери застосування корпусу текстів <p>вміти:</p> <ul style="list-style-type: none"> – пояснити завдання, методи й основний понятійним апарат корпусної лінгвістики – представити історію корпусних досліджень, її теоретичне підґрунтя і шляхи розвитку – користуватися комп'ютерним інструментарієм корпусної лінгвістики – застосовувати корпуси української та інших мов залежно від поставленого завдання – аналізувати будову, структуру, наповнення корпусів української та інших мов – правильно інтерпретувати якісні та кількісні дані, отримані в результаті корпусного пошуку – продемонструвати можливості роботи з корпусними базами даних в лінгвістичних дослідженнях на прикладі власної наукової роботи.
Ключові слова	Корпус текстів, роззначення (анотація), тег, запит, репрезентативність корпусу, конкорданс, частотність, ключові слова, лема, метадані, токен, слововживання
Формат курсу	Очний
Теми	* СХЕМА КУРСУ
Підсумковий контроль, форма	Залік в кінці семестру
Пререквізити	Для вивчення курсу студенти потребують базових знань з філологічних дисциплін
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	<u>Методи навчання:</u> пояснювально-ілюстративний, метод проблемного викладання, евристичний метод, дослідницький метод. <u>Методи викладу навчального матеріалу</u> – лекція, консультація, евристична бесіда, дискусія. <u>Наочні методи навчання</u> – ілюстрування (художні тексти), демонстрування з використанням технічних засобів (слайди). <u>Техніки:</u> навчальної дискусії, створення проблемної ситуації, інтелектуального експерименту, командна гра.
Необхідне обладнання	Комп'ютер/ноутбук, проектор, екран, доступ до інтернету
Критерії оцінювання	Максимальна оцінка, яку може отримати магістрант в кінці курсу, – 100 балів.

<p>(окремо для кожного виду навчальної діяльності)</p>	<p>Робота магістранта оцінюється за такими показниками:</p> <p>а) активність на заняттях – 20 балів (середньостатистичний бал за відповіді на практичних);</p> <p>б) застосування корпусних технологій для написання індивідуально-дослідницької роботи – 30 балів;</p> <p>в) відповідь на заліку – 50 балів.</p> <p>Політика виставлення балів. Враховуються бали за відповіді на практичних заняттях, за усне та письмове опитування, за виконання письмових завдань, за самостійну роботу з корпусами текстів для свого наукового дослідження, за відповідь на заліку. При цьому обов'язково враховуються присутність на лекційних і практичних заняттях, активність студента під час практичного заняття; недопустимість запізень на заняття; користування мобільним телефоном, планшетом або іншими мобільними пристроями під час заняття в цілях, не пов'язаних з навчанням і т. ін.</p> <p>Академічна доброчесність. Очікується, що письмові творчі роботи студентів будуть їх оригінальними дослідженнями, у яких вони висловлюють власну думку.</p> <p>Прикладами можливої академічної недоброчесності можуть вважатися відсутність посилань на використані джерела, фабрикування джерел, списування, підглядання під час усної відповіді, втручання в роботу інших студентів тощо. Виявлення ознак академічної недоброчесності в письмовій чи усній роботі студента є підставою для її незарахування незалежно від масштабів плагіату чи обману.</p> <p>Студенти зобов'язані дотримуватися строків, визначених для виконання усіх видів робіт, передбачених курсом.</p> <p>Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідуватимуть лекції і практичні заняття курсу. Пропущені практичні заняття студент повинен відпрацювати у письмовому вигляді (формат письмового відпрацювання визначає викладач індивідуально залежно від теми і форми пропущеного заняття).</p> <p>Література. Усю літературу, яку студенти не зможуть знайти самостійно, буде надано викладачем виключно в освітніх цілях без права її передання третім особам. Заохочуємо студентів використовувати також іншої літератури та джерел, яких немає серед рекомендованих.</p>
<p>Питання до підсумкової контрольної роботи (заліку)</p>	<ul style="list-style-type: none"> – Корпусна лінгвістика як інноваційний інформаційно-семіотичний напрям сучасного мовознавства: сутність, базові поняття, дослідницькі принципи й методи – Поняття корпусу текстів (КТ). Його основні ознаки – Об'єкт і предмет дослідження корпусної лінгвістики. – Передумови виникнення та етапи становлення корпусної лінгвістики як науки. – Теоретичне підґрунтя корпусної лінгвістики. – Основні напрямки корпусної лінгвістики. Створення КТ, використання КТ. Національні КТ, мультилінгвальні ресурси, навчальні корпуси (learner corpora), корпуси мов, що зникають (indangered corpora). – Дослідницький потенціал КТ. Інформаційно-пошукові, дослідницько-мовознавчі та інші можливості текстового корпусу в різних галузях мовознавства. – Розвиток корпусної лінгвістики у світі. – Стан корпусної лінгвістики в Україні. Проекти створення корпусів текстів української мови. Спільне та відмінне між ними.

- Робота з референтним корпусом української мови ГРАК.
- Збір та підготовка текстів для корпусу ГРАК: метаінформація.
- Корпусні студії: історична перспектива та сучасний стан. За і проти.
- Історія становлення корпусної лінгвістики: від паперових конкордансів і картотек до перших електронних корпусів
- Корпусна лінгвістика з 60-х років ХХ ст. до пост 2000х
- Корпусні дослідження в Україні та за кордоном.
- Критика корпусного аналізу мови і мовлення.
- Корпусні студії: історична перспектива та сучасний стан
- Критика корпусного аналізу мови і мовлення.
- Принципи створення корпусу текстів. Основні вимоги до створення КТ. Загальний огляд.
- Репрезентативність КТ
- Анотованість (роззначення, індексованість розмітка). Параметри зовнішньої (екстралінгвістичної), структурної та внутрішньої анотації. Традиційні види внутрішньої анотації: морфологічна (тегінг, тагування), синтаксична (парсінг), лематизація (до початкової форми), просодична (для усного мовлення інтонаційні характеристики). Permission and using.
- Стандартність.
- Одномовні та багатомовні корпуси текстів.
- Паралельні (колатеральні, collateral, simultaneous) та порівнювані (з подібною чи ідентичною будовою у відношенні заздалегідь відомих параметрів).
- Ілюстративні та дослідницькі корпуси.
- Корпуси текстів національної мови та конкретної галузі. Повнотекстові, фрагментовані.
- Динамічні / статичні (авторські, тут машинні фонди рос. мови) корпуси текстів. Динамічні (моніторні) (автор теорії Джон Синклер) для лінгвістичного відслідковування матеріалу із врахуванням діахронії. Бірмінгемський корпус.
- Корпуси текстів у системно-структурному та семантичному дослідженні мови.
- Корпуси текстів та конкорданси у визначенні значення слова та в укладанні тлумачних словників.
- Учнівські корпуси текстів. Корпуси текстів у навчанні іноземних мов.
- Корпуси текстів і переклад.
- Основні сфери застосування корпусів текстів
- Огляд та використання КТ української мови. Лінгвістичний корпус текстів української мови (Мовно-інформаційний фонд НАН України)
- Генеральний регіонально анотований корпус української мови (Грак). Загальна характеристика та функційні можливості.
- Польський автоматичний веб-корпус української мови (ПАВУК). Загальна характеристика та функційні можливості.
- Дослідницький корпус української мови (nova.info). Загальна характеристика та функційні можливості.
- Корпус української мови Лейпцизького університету. Загальна характеристика та функційні можливості.
- Паралельні корпуси української мови. Загальна характеристика та функційні можливості.
- Як шукати і знаходити потрібну інформацію у КТ? CQL (Common Query Language) універсальна формальна мова запитів – її

	застосування на прикладі корпусів української та інших мов. – Функційні можливості AntConc 4.1.4 для обробки та аналізу даних корпусу.
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано після завершення курсу.

Схема курсу

Тиж. / дата / год.-	Тема, план, короткі тези	Форма діяльності (заняття)	Література	Завдання, год	Термін виконання
Модуль I Змістовий модуль 1. Корпусна лінгвістика: сутність, становлення, основні принципи, критика					
1-й т. 2 год	<p>Тема 1. Корпусна лінгвістика як інноваційний інформаційно-семіотичний напрям сучасного мовознавства: сутність, базові поняття, дослідницькі принципи й методи</p> <p>1. Поняття текстового корпусу. Корпус текстів як складна глибока розгалужена система.</p> <p>2. Об'єкт і предмет дослідження корпусної лінгвістики.</p> <p>3. Передумови виникнення та етапи становлення корпусної лінгвістики як науки.</p> <p>4. Теоретичне підґрунтя корпусної лінгвістики.</p> <p>5. Основні напрямки корпусної лінгвістики. Створення КТ, використання КТ. Національні КТ, мультлінгвальні ресурси, навчальні корпуси (learner corpora), корпуси мов, що зникають (indangered corpora).</p> <p>6. Дослідницький потенціал КТ. Інформаційно-пошукові, дослідницько-мовознавчі та інші можливості текстового корпусу в різних галузях мовознавства.</p>	Лекція	<p>Дарчук Н. Корпусна лінгвістика: проблеми, методи, перспективи. Навчальна програма. Київ: Київський національний університет ім. Тараса Шевченка, 2013.</p> <p>Демська-Кульчицька О. Текстовий корпус: ідея іншої форми. Київ : НаУКМА, 2011.</p> <p>Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013.</p> <p>Прикладна лінгвістика: навчальний посібник/ МОН України, Уманський держ. пед. ун.-т. імені Павла Тичини/ упорядник: Н. А. Цимбал. Умань: Візаві, 2019.</p> <p>Селіванова О.О. Корпусна лінгвістика // Сучасна лінгвістика: напрями та проблеми: підручник. Полтава: Довкілля-К, 2008. С. 667–669.</p> <p>Українська корпусна лінгвістика: сьогодення і перспективи : [інтерв'ю із канд. філолог. наук, доц. каф. укр. мови Нац. ун-ту "Києво-Могилянська академія" Орисьою Мар'янівною Демською] // Дивослово. 2011. № 10. С. 32-34.</p>	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3 год.	1-й т.

	7. Розвиток корпусної лінгвістики у світі. 8. Стан корпусної лінгвістики в Україні. Проекти створення корпусів текстів української мови. Спільне та відмінне між ними.		https://dyvoslovo.com.ua/wp-content/uploads/2016/03/6-1011.pdf Широков та ін. Лінгвістично-інформаційні студії : праці Українського мовно-інформаційного фонду НАН України : у 5 т. Т. 4 : Корпусна та когнітивна лінгвістика. Київ. Український мовно-інформаційний фонд НАН України, 2018.		
1-й т. 2 год	Робота з референтним корпусом української мови ГРАК. Ознайомлення. Стилі, тематика і жанри; мова оригіналу; Датування; Правопис; Відомості про автора тексту; Відомості про медіа; Регіональна розмітка; Морфологічна розмітка; Граматична омонімія; Джерело тексту; Семантична розмітка	Практичне заняття	М. Шведова, Р. фон Вальденфельс. Регіональна розмітка в ГРАКУ, великому референтному корпусі української мови: питання і проблеми [Regional Annotation within GRAC, a Large Reference Corpus of Ukrainian: Issues and Challenges]. CEUR Workshop Proceedings. Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021). Volume I: Main Conference. Kharkiv, Ukraine, April 22-23, 2021. ГРАК: Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024 < www.uacorpus.org >.	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел. 3 год.	1-й т.
2-й т. 2 год	Збір та підготовка текстів для корпусу ГРАК: метаінформація. Загальні принципи і правила. Інформація про корпус: розробники, використані ресурси, подяки, версії. Догляд за корпусом.	Практичне заняття	ГРАК: Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024 < www.uacorpus.org >.	Навчитися збирати і готувати тексти для корпусу ГРАК. Знати загальні принципи і правила. 3 год.	2-й т.
3-й т. 2 год	Тема 2. Корпусні студії: історична перспектива та сучасний стан. За і проти. 1. Історія становлення корпусної	Лекція	Демська-Кульчицька О. Основи національного корпусу української мови. Київ : Інститут української мови НАНУ,	Опрацювання відповідних розділів	3-й т.

	<p>лінгвістики: від паперових конкордансів і картотек до перших електронних корпусів</p> <p>2. Корпусна лінгвістика з 60-х років ХХ ст. до пост 2000х</p> <p>3. Корпусні дослідження в Україні та за кордоном.</p> <p>4. Критика корпусного аналізу мови і мовлення.</p>		<p>2005.</p> <p>Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013.</p> <p>Прикладна лінгвістика: навчальний посібник/ МОН України, Уманський держ. пед. ун.-т. імені Павла Тичини/ упорядник: Н. А. Цимбал. Умань: Візаві, 2019. 106 с.</p> <p>Селіванова О.О. Корпусна лінгвістика // Сучасна лінгвістика: напрями та проблеми: підручник. Полтава: Довкілля-К, 2008.</p> <p>Широков та ін. Лінгвістично-інформаційні студії : праці Українського мовно-інформаційного фонду НАН України : у 5 т. Т. 4 : Корпусна та когнітивна лінгвістика. Київ. Український мовно-інформаційний фонд НАН України, 2018.</p>	<p>рекомендованої літератури, інтернет-джерел, конспекту лекції.</p> <p>3 год.</p>	
3-й т. 2 год.	<p>Корпусні студії: історична перспектива та сучасний стан.</p> <p>1. Історія становлення корпусної лінгвістики: від паперових конкордансів і картотек до перших електронних корпусів</p> <p>2. Корпусна лінгвістика з 60-х років ХХ ст. до пост 2000х</p> <p>3. Корпусні дослідження в Україні та за кордоном.</p> <p>4. Критика корпусного аналізу мови і мовлення.</p>	Практичне заняття	Л-ра до лекції 2.	<p>Підготувати відповіді на питання, вміти їх аргументувати, навести конкретні приклади.</p> <p>3 год.</p>	3-й т.
4-й т. 2 год.	<p>Критика корпусного аналізу мови і мовлення.</p>	Практичне заняття	Л-ра до лекції 2.	<p>Порівняти традиційний та корпусний підходи до</p>	4-й т.

				аналізу текстів. Відповідь обґрунтувати. 3 год.	
5-й т. 2 год	<p>Тема 3. Принципи створення корпусу текстів</p> <p>1. Основні вимоги до створення КТ. Загальний огляд. Адресованість. Машиночитаність. Репрезентативність.</p> <p>2. Анотованість (роззначення, індексованість розмітка).</p> <p>3. Параметри зовнішньої (екстралінгвістичної), структурної та внутрішньої анотації.</p> <p>4. Традиційні види внутрішньої анотації: морфологічна (тегінг, тагування), синтаксична (парсінг), лематизація (до початкової форми), просодична (для усного мовлення інтонаційні характеристики). Permission and using.</p> <p>5. Стандартність. Міжнародні системи індексування текстів.</p>	Лекція	<p>Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013.</p> <p>Eney T. Ms., Xiao R., and Tono Y. Corpus-based language studies: an advanced resource book. Routledge, London, 2006. 286 p.</p> <p>T. McEnery, A. Wilson Corpus Linguistics. An Introduction. 2-nd Edition. Edinburg: University Press, 2001. 235 p.</p> <p>Paquot M. & Gries S. T. A practical handbook of corpus linguistics. Berlin: Springer. 2020.</p> <p>Podstawy językoznawstwa korpusowego / Red. B. Lewandowska-Tomaszczyk. Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 2005. 306 s.</p> <p>Sinclair J. EAGLES Preliminary recommendations on Corpus Typology. <www.ilc.cnr.it/EAGLES/pub/eagles/corpora/corpusstyp.ps.gz></p> <p>Viana V., & O'boyle A. Corpus linguistics for English for Academic Purposes. New York: Routledge. 2022.</p> <p>Zufferey S. Introduction to corpus linguistics. New Jersey: John Wiley & Sons. 2020.</p>	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3 год.	5-й т.
5-й т. 2 год	<p>Основні вимоги до створення КТ.</p> <p>Загальний аналіз на прикладі обраних корпусів української і/або інших мов</p>	Практичне заняття	Л-ра до лекції 3.	Проаналізувати основні вимоги до створення КТ на прикладі українських КТ.	5-й т.

				3 год.	
6-й т. 2 год	Анотованість (роззначення, індексованість розмітка). Параметри зовнішньої (екстралінгвістичної), структурної та внутрішньої анотації. Традиційні види внутрішньої анотації: морфологічна (тегінг, тагування), синтаксична (парсінг), лематизація (до початкової форми), просодична (для усного мовлення інтонаційні характеристики).	Практичне заняття	Л-ра до лекції 3.	Проаналізувати параметри зовнішньої, структурної та внутрішньої анотації на прикладі українських КТ. 3 год.	6-й т.
7-й т. 2 год	Тема 4. Типологія та основні сфери застосування корпусу текстів 1. Одномовні та багатомовні корпуси текстів. Паралельні (колатеральні, collateral, simultaneous) та порівнювані (з подібною чи ідентичною будовою у відношенні задалегідь відомих параметрів). 2. Ілюстративні та дослідницькі корпуси. 3. Корпуси текстів національної мови та конкретної галузі. Повнотекстові, фрагментовані. 4. Динамічні / статичні (авторські, тут машинні фонди рос. мови) корпуси текстів. Динамічні (моніторні) (автор теорії Джон Синклер) для лінгвістичного відслідковування матеріалу із врахуванням діахронії. Бірмінгемський корпус. 5. Корпуси текстів у системно-структурному та семантичному дослідженні мови. 6. Корпуси текстів та конкорданси у визначенні значення слова та в укладанні тлумачних словників. 7. Учнівські корпуси текстів.	Лекція	Демська-Кульчицька О. Основи національного корпусу української мови. Київ : Інститут української мови НАНУ, 2005. Демська-Кульчицька 2011: Демська-Кульчицька О. Текстовий корпус: ідея іншої форми. Київ : НаУКМА, 2011. Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013. 142 с. Бук С. Учнівські корпуси в методиці викладання іноземної мови // Теорія і практика викладання української мови як іноземної. Львів : ЛНУ імені Івана Франка, 2007. Вип. 2. С. 19–23. Praktyczny przewodnik po korpusach języków słowiańskich. Red. Milena Hebal-Jeziarska. Warszawa: Zakład Graficzny Uniwersytetu Warszawskiego, 2014. 231 s. < http://slawistyka.uw.edu.pl/pl/2014/11/05/praktyczny-przewodnik-po-korpusach-jezykow-slowianskich-red-nauk-m-hebal-jeziarska/ > Sinclair J. EAGLES Preliminary	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3 год.	7-й т.

	8. Корпуси текстів у навчанні іноземних мов. 9. Корпуси текстів і переклад.		recommendations on Corpus Typology. < www.ilc.cnr.it/EAGLES/pub/eagles/corpora/corpusstyp.ps.gz >		
7-й т. 2 год	Класифікація корпусів текстів	Практичне заняття	<p>Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024 <www.uacorporus.org>.</p> <p>Український національний лінгвістичний корпус. Український мовоно-інформаційний фонд <svc.ulif.org.ua></p> <p>КУМ (Чтиво) = Корпус української мови <korpus.org.ua></p> <p>Лабораторія української. Паралельні корпуси <https://mova.institute/%D0%BF%D0%B0%D1%80%D0%B0%D0%BB%D0%B5%D0%BB%D1%8C%D0%BD%D1%96_%D0%BA%D0%BE%D1%80%D0%BF%D1%83%D1%81%D0%B8></p> <p>Корпус української мови Лейпцизького університету / Corpus of Ukrainian by the University of Leipzig <https://corpora.uni-leipzig.de/en?corpusId=ukr_mixed_2014&fbclid=IwAR2n7pHaHOoOxkOQHaJsIkN9ynXCB6-dq6uQn_5eEKQ7l9h6YCxZQwh5ix8></p> <p>Національний корпус кримсько-татарської мови / National Corpus of the Crimean Tatar Language <https://ctcorpus.org/index.php/en/></p> <p>Skovoroda Online Concordance / Онлайн конкорданс Сковороди <http://www.arts.ualberta.ca/~ukr/skovoroda/NEW/index.php></p> <p>Дарчук Н. П. Корпус текстів Тараса Шевченка як джерело для мовознавчих досліджень // Science and Education a New</p>	Визначити тип запропонованих корпусів. 3 год.	7-й т.

		<p>Dimension. Philology. V(32). Issue: 122. 2017. С. 33–37.</p> <p>Дослідницький корпус Лесі Українки: лінгвістичні параметри // Науковий вісник Волинського національного університету імені Лесі Українки. Філологічні науки. Мовознавство. 2018. № 12. С. 30–36.</p> <p>Загнітко А., Данилюк І., Ситар Г. Корпус текстів Юрія Шевельова: структура, функції, навігація // Мова: класичне – модерне – постмодерне. Київ: НУКМА. 2019. Вип. 5. С. 158–169.</p> <p>Корпус текстів Богдана Лепкого / Кітаєва Віка (укладач). <http://corpora.donnu.edu.ua/bonito/run.cgi/corp_info?corpname=kitaeva></p> <p>Корпус текстів Михайла Коцюбинського / Бройко Н. (укладач). <http://corpora.donnu.edu.ua/bonito/run.cgi/corp_info?corpname=brojk></p> <p>British National Corpus <http://www.natcorp.ox.ac.uk>.</p> <p>Český národní korpus <http://ucnk.ff.cuni.cz></p> <p>Korpus Języka Polskiego IPI PAN <http://korpus.pl/index.php?page=welcome></p> <p>Korpus Języka Polskiego Wydawnictwa Naukowego PWN <https://sjp.pwn.pl/korpus></p> <p>PAWUK = Polish Automatic Web corpus of Ukrainian language / Польський автоматичний вебкорпус української мови <https://pawuk.ipipan.waw.pl/overview></p> <p>PolUKR = polsko-ukraiński korpus</p>		
--	--	---	--	--

			<p>równoległy / польсько-український паралельний корпус http://www.domeczek.pl/~polukr/index.php?option=welcome</p> <p>The JRC-Acquis Multilingual Parallel Corpus / Корпус законодавства Європейського Союзу: Acquis Communautaire, що містить 231 мовну пару https://web.archive.org/web/20060619034515/http://langtech.jrc.it/JRC-Acquis.html</p>		
8-й т. 2 год	Основні сфери застосування корпусів текстів	Практичне заняття	Л-ра до попереднього практичного заняття	Визначити сферу застосування запропонованих корпусів. 3 год.	8-й т.
Модуль II					
Змістовий модуль 2. Корпус текстів як інноваційний інформаційно-семіотичний інструмент аналізу мови і мовлення					
9-й т. 2 год	<p>Тема 5. Аналіз та використання КТ української мови</p> <ol style="list-style-type: none"> 1. Лінгвістичний корпус текстів української мови (Мовно-інформаційний фонд НАН України) 2. Генеральний регіонально анатований корпус української мови (Грак) 3. Національний корпус української мови (Інститут української мови НАН України) 4. Польський автоматичний веб-корпус української мови (ПАВУК) 5. Дослідницький корпус української мови (mova.info) 6. Корпус української мови Лепцизького університету 	Лекція	<p>Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024 www.uacorporus.org.</p> <p>Український національний лінгвістичний корпус. Український мовно-інформаційний фонд svc.ulif.org.ua</p> <p>Корпус української мови. Лабораторія комп'ютерної лінгвістики Київського національного університету імені Тараса Шевченка. http://www.mova.info/corpus.aspx?11=209</p> <p>PAWUK = Polish Automatic Web corpus of Ukrainian language / Польський автоматичний вебкорпус української мови https://pawuk.ipipan.waw.pl/overview</p> <p>Корпус української мови Лепцизького</p>	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3 год.	9-й т.

			університету / Corpus of Ukrainian by the University of Leipzig < https://corpora.uni-leipzig.de/en?corpusId=ukr_mixed_2014&fbclid=IwAR2n7pHaHOoOxkOQHaJsIkN9ynXCB6-dq6uQn_5eEKQ7I9h6YCxZQwh5ix8 >		
9-й т. 2 год	Аналіз та зіставлення корпусів текстів української мови: 1.Лінгвістичний корпус текстів української мови (Мовно-інформаційний фонд НАН України) 2.Генеральний регіонально анатований корпус української мови (Грак) 3.Націонільний корпус української мови (Інститут української мови НАН України) 4.Дослідницький корпус української мови (mova.info)	Практичне заняття	Л-ра до теми 5.	Проналізувати та зіставити вказані КТ української мови. 3 год.	9-й т.
10-й т. 2 год	Практична робота з корпусами текстів: 1. Польський автоматичний веб-корпус української мови (ПАВУК) 2. Корпус української мови Лепцизького університету	Практичне заняття	PAWUK = Polish Automatic Web corpus of Ukrainian language / Польський автоматичний вебкорпус української мови < https://pawuk.ipipan.waw.pl/overview > Корпус української мови Лепцизького університету / Corpus of Ukrainian by the University of Leipzig < https://corpora.uni-leipzig.de/en?corpusId=ukr_mixed_2014&fbclid=IwAR2n7pHaHOoOxkOQHaJsIkN9ynXCB6-dq6uQn_5eEKQ7I9h6YCxZQwh5ix8 >	Проаналізувати будову та функціонал вказаних КТ. 3 год.	10-й т.
11-й т. 2 год	Тема 6. Паралельні корпуси української мови 1. Паралельний корпус української мови (ПарКУМ, Н. Дарчук, М. Лангенбах, В. Сорокін, Я. Ходаківська) 2. Паралельний корпус автоперекладів І.	Лекція	Дарчук Н. П., Лангенбах М. О., Сорокін В. М., Ходаківська Я. В. Паралельний корпус текстів ПАРКУМ // Науковий часопис НПУ імені М. П. Драгоманова. 2017. Вип. 15. С. 28–35. PolUKR = polsko-ukraiński korpus równoległy	Опрацювання відповідних розділів рекомендованої літератури, інтернет-	11-й т.

	Франка 3. Паралельний польсько-український та українсько-польський корпус (Н. Коциба, М. Турська)		/ польсько-український паралельний корпус < http://www.domeczek.pl/~polukr/index.php?option=welcome > Бук С. Архітектура польсько-українського та українсько-польського паралельного корпусу автоперекладів Івана Франка // Slavia Orientalis. 2012. Т. LXI. N 2. С. 213–230.	джерел, конспекту лекції. 3 год.	
11-й т. 2 год	Характеристика, аналіз і практична робота з корпусами текстів: 1. Паралельний корпус української мови (ПарКУМ, Н. Дарчук, М. Лангенбах, В. Сорокін, Я. Ходаківська) 2. Паралельний корпус автоперекладів І. Франка	Практичне заняття	Дарчук Н. П., Лангенбах М. О., Сорокін В. М., Ходаківська Я. В. Паралельний корпус текстів ПАРКУМ // Науковий часопис НПУ імені М. П. Драгоманова. 2017. Вип. 15. С. 28–35. Бук С. Архітектура польсько-українського та українсько-польського паралельного корпусу автоперекладів Івана Франка // Slavia Orientalis. 2012. Т. LXI. N 2. С. 213–230.	Проаналізувати структуру і функціонал вказаних КТ. 3 год.	11-й т.
12-й т. 2 год	Характеристика, аналіз і практична робота з корпусами текстів (продовження): Паралельний польсько-український та українсько-польський корпус (Н. Коциба, М. Турська)	Практичне заняття	PolUKR = polsko-ukraiński korpus równoległy / польсько-український паралельний корпус < http://www.domeczek.pl/~polukr/index.php?option=welcome >	Проаналізувати структуру і функціонал вказаного КТ. 3 год.	12-й т. 2 год
13-й т. 2 год	Тема 7. Міжнародні корпусні проекти із задіянням української мови 1. Parasol – паралельний корпус слов'янських мов 2. Корпус слов'янських мов Лепцизького університету 3. Інші міжнародні проекти	Лекція	Корпус української мови Лепцизького університету / Corpus of Ukrainian by the University of Leipzig < https://corpora.uni-leipzig.de/en?corpusId=ukr_mixed_2014&fbclid=IwAR2n7pHaHOoOxkOOHaJsIkN9ynXCB6-dq6uQn_5eEKQ7I9h6YCxZQwh5ix8 > Von Waldenfels, Ruprecht and Meyer, Roland (2006-2023): ParaSol, a Corpus of Slavic and Other Languages < https://www.parasolcorpus.org/Ursynow/ >	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3 год.	13-й т.
13-й т.	Характеристика, аналіз і практична	Практичне	Von Waldenfels, Ruprecht and Meyer, Roland	Проаналізувати	13-й

2 год	робота з корпусами текстів: Parasol – паралельний корпус слов'янських мов	заняття	(2006-2023): ParaSol, a Corpus of Slavic and Other Languages < https://www.parasolcorpus.org/Ursynow/ >	структуру і функціонал вказаного КТ. 3 год.	т.
14-й т. 2 год	Характеристика, аналіз і практична робота з корпусами текстів: Корпус слов'янських мов Лепцизького університету й інші міжнародні проекти	Практичне заняття	Корпус української мови Лепцизького університету / Corpus of Ukrainian by the University of Leipzig < https://corpora.uni-leipzig.de/en?corpusId=ukr_mixed_2014&fbclid=IwAR2n7pHaHOoOxkOQHajSIkN9ynXC B6-dq6uQn_5eEKQ719h6YCxZQwh5ix8 >	Проаналізувати структуру і функціонал вказаних КТ. 3 год.	14-й т.
15-й т. 2 год	Тема 8. Корпуси текстів українських та зарубіжних письменників 1. Онлайн конкорданс Г. Сковороди 2. Конкорданс Т. Шевченка 3. Корпус текстів великої прози І. Франка 4. Корпус текстів Лесі Українки 5. Корпус текстів Біблії, Шекспіра, Дж. Джойса, можливості English-corpora.org тощо.	Лекція	Skovoroda Online Concordance / Онлайн конкорданс Сковороди < http://www.arts.ualberta.ca/~ukr/skovoroda/NEW/index.php > Дарчук Н. П. Корпус текстів Тараса Шевченка як джерело для мовознавчих досліджень // Science and Education a New Dimension. Philology. V(32). Issue: 122. 2017. С. 33–37. Дослідницький корпус Лесі Українки: лінгвістичні параметри // Науковий вісник Волинського національного університету імені Лесі Українки. Філологічні науки. Мовознавство. 2018. № 12. С. 30–36. Бук С. Архітектура польсько-українського та українсько-польського паралельного корпусу автоперекладів Івана Франка // Slavia Orientalis. 2012. Т. LXI. N 2. С. 213–230. Бук С. Велика проза Івана Франка: електронний корпус, частотні словники та інші міждисциплінарні контексти : монографія / Львів : ЛНУ імені Івана Франка, 2021. 424 с.	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3 год.	15-й т.

			<p>Бук С. Учнівські корпуси в методиці викладання іноземної мови // Теорія і практика викладання української мови як іноземної. Львів : ЛНУ імені Івана Франка, 2007. Вип. 2. С. 19–23.</p> <p>Загнітко А., Данилюк І., Ситар Г. Корпус текстів Юрія Шевельова: структура, функції, навігація // Мова: класичне – модерне – постмодерне. Київ: НУКМА. 2019. Вип. 5. С. 158–169.</p> <p>Корпус текстів Богдана Лепкого / Кітаєва Віка (укладач). <http://corpora.donnu.edu.ua/bonito/run.cgi/corp_info?corpname=kitaeva></p> <p>Корпус текстів Михайла Коцюбинського / Бройко Н. (укладач). <http://corpora.donnu.edu.ua/bonito/run.cgi/corp_info?corpname=brojk></p> <p>Бук С., Ровенчак А. Онлайн-конкорданс роману Івана Франка «Перехресні стежки» <http://ktf.franko.lviv.ua/users/andrij/science/Franko/concordance.html></p>		
15-й т. 2 год	<p>Як шукати і знаходити потрібну інформацію у КТ</p> <p>SQL (Common Query Language) універсальна формальна мова запитів – її застосування на прикладі корпусів української та інших мов</p>	Практичне заняття	<p>ГРАК: Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024 <www.uacorpora.org>.</p> <p>Інструменти роботи зі системою SketchEngine (user guide) <https://www.sketchengine.eu/guide></p>	Проаналізувати варіативність запитів мовою SQL на прикладі корпусу ГРАК. 3 год.	15-й т.
16-й т. 2 год	<p>Практичні інструменти для створення власного КТ</p> <ol style="list-style-type: none"> 1. Застосування програми Korpusomat 2. Потенціал КТ ГРАК. 	Практичне заняття	<p>Korpusomat <https://korpusomat.eu/></p> <p>Laurence A. AntConc (Windows, MacOS, Linux). Build 4.1.4 <https://www.laurenceanthony.net/software/an</p>	Створити свій КТ за певними параметрами за допомогою	16-й т.

	3. Функційні можливості AntConc 4.1.4 для обробки та аналізу даних корпусу.		tconc/releases/AntConc414/help.pdf >	ресурсу Corpusomat, проаналізувати AntConc. 3 год.	
--	---	--	--	--	--