

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 2024 р.)

Завідувач катедри  Ф.С. Бацевич

Силабус з навчальної дисципліни

«Квантитативні методи аналізу тексту»,

що викладається в межах ОПП **«Прикладна лінгвістика»**

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

спеціалізації **035.10 Прикладна лінгвістика**

Львів 2024 р.

Назва курсу	Квантитативні методи аналізу тексту
Адреса викладання курсу	Львівський національний університет ім. Івана Франка м. Львів, вул. Університетська, 1
Факультет та кафедра, за якою закріплена дисципліна	Філологічний факультет Катедра загального мовознавства
Галузь знань, шифр та назва спеціальності	Галузь знань - 03 Гуманітарні науки. Спеціальність - 035 Філологія. Спеціалізація - 035.10 Прикладна лінгвістика
Викладач курсу	Доктор філологічних наук, професор катедри загального мовознавства Бук Соломія Несторівна. Кандидат фізико-математичних наук, доцент катедри прикладної математики Чирун Любомир Вікторович.
Контактна інформація викладача	solomiia.buk@lnu.edu.ua +38 050 1734XX77
Консультації по курсу відбуваються	Консультації відбуваються за попередньою домовленістю за адресою: вул.Університетська, 1, ауд.343 або онлайн
Сторінка курсу	https://philology.lnu.edu.ua/course/kvanytyatyvni-metody-analizu-tekstu-2023
Інформація про курс	Дисципліна «Квантитативні методи аналізу тексту» належить до нормативних навчальних предметів для магістрантів другого року навчання. Ця дисципліна викладається в обсязі 3 кредити (ECTS). Курс покликаний поглибити теоретичні і практичні знання магістрантів зі спеціальності прикладна лінгвістика, сформувані чіткі уявлення про особливості застосування квантитативних методів у гуманітарних науках.
Коротка анотація курсу	У лінгвістиці зламу ХХ–ХХІ століть особливо актуальні міждисциплінарні дослідження мови, і застосування квантитативних методів для аналізу тексту є саме таким напрямком. Впродовж цього курсу навчимося створювати, користуватися і правильно інтерпретувати частотні словники; знаходити необхідні кількісні характеристики різних стилів і жанрів, письменників і їхніх творів; визначати авторство, прочитність і складність тексту; з'ясуємо, скільки слів потрібно, щоби оволодіти мовою; ознайомимося з основними статистичними інструментами текстових корпусів, програмам Sketch Engine, Stylo, графіками CQL, Word, Excel тощо, сформуємо навички їх використання та візуалізації кількісних текстових даних тощо. Практичні заняття передбачають використання корпусних та інших цифрових технологій, формат академічної дискусії, командну роботу, соціолінгвістичний експеримент, укладання частотного словника, кількісний аналіз мови і тексту. Курс є важливим в системі освіти сучасного мовознавця, без нього складно уявити висококваліфікованого фахівця-лінгвіста. Він сприяє переосмисленню академічного статусу лінгвістичних дисциплін і їх соціальної функції.

<p>Мета і цілі курсу</p>	<p>Мета курсу – сформувати у магістрантів знання щодо застосування основних квантитативних (кількісних) методів у філології: мовознавстві та літературознавстві, лінгвістичній експертизі, стилометрії, криміналістичній, соціо-, когнітивній та історичній лінгвістиці, лінгводидактиці та оволодінні іноземною мовою, етнолінгвістиці, фольклористиці, діалектології, а також у культурології, гендерних дослідженнях тощо.</p> <p>Завдання курсу: (1) осмислити карколомний кількісний поворот у лінгвістиці: протягом останніх 20 років у світовій лінгвістиці частка кількісно- статистичних методів неухильно зростає, а частка якісного підходу різко зменшується (Kortmann 2021);</p> <p>(2) дати студентам уявлення про те, як покращити якість своїх досліджень, застосовуючи основні кількісні інструменти, і таким чином зробити молоде покоління більш конкурентоспроможним в міжнародній галузі досліджень;</p> <p>(3) проаналізувати основні статистичні інструменти текстових корпусів, програм Sketch Engine, Stylo, графіків CQL, Word, Excel тощо та сформувати навички їх використання; навчитися візуалізувати текстові дані;</p> <p>(4) з'ясувати роль квантитативного опису мови і тексту, а також визначити місце цієї галузі серед цифрових гуманітарних наук;</p> <p>(5) показати значення основоположних кількісних понять (частота, відносна частота та її порівняння; ключові слова, вимірювання колокацій, багатовимірний аналіз тощо) та їхнє значення в гуманітарних галузях, згаданих вище.</p>
<p>Література для вивчення дисципліни</p>	<p>Основна</p> <ul style="list-style-type: none"> • Альтман Г. Мода та істина в лінгвістиці // Проблеми квантитативної лінгвістики. Чернівці: Рута, 2005. С. 3–11. • Бук С. Основи статистичної лінгвістики: Навчально-методичний посібник. Видавничий центр ЛНУ імені Івана Франка, 2008. 124 с. • Васильєв О., Чалий О., Васильєва І. Математичні методи й моделі в лінгвістиці // Україна модерна. Справа Verba et numeri (кількісні підходи до аналізу мови й тексту). 2019. Ч. 27. С. 9–28. • Дарчук Н. П. Статистичні характеристики лексики як відображення структури тексту // Мовознавчі студії. К.: Наукова думка, 1976. С. 97–102. • Перебийніс В. С. Статистичні методи для лінгвістів: Навчальний посібник. Вінниця: «Нова книга», 2002. С. 3–8. • Перебийніс В. С., Муравицька М. П., Дарчук Н. П. Частотні словники та їх використання. К: Наукова думка, 1985. 204 с. • Статистичні параметри стилів / за ред. В. С. Перебийніс. К.: Наукова думка, 1967. 260 с. • Структура мови і статистика мовлення / Відп. ред. канд. філол. наук М. П. Муравицька і Л. А. Олексієнко. К.: Наукова думка, 1974. 176 с. • Burrows J. 'Delta' a Measure of Stylistic Difference and a Guide to Likely Authorship. Literary and Linguistic Computing. 2002. 17/3, 1 September: 267–287. • Eder M., Rybicki J., Kestemont M. Stylometry with R: a package for computational text analysis // R Journal. 2016. 8(1). 107–121. <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html> • Kortmann B. Reflecting on the Quantitative Turn in Linguistics // Linguistics. 59 (5): P. 1212. • Lahjouji-Seppälä M., Rabus A., von Waldenfels R. Ukrainian standard variants in the 20th century: stylometry to the rescue <https://www.researchgate.net/publication/364332887_Ukrainian_standard_vari-

ants_in_the_20th_century_stylometry_to_the_rescue?fbclid=IwAR3bNDrrWLN
UXtyPF0SZ5dID_NZaFYfVXCACULYtMOhjt5vqse2TY48u5c>

Додаткова

- Бук С. Велика проза Івана Франка: електронний корпус, частотні словники та інші міждисциплінарні контексти : монографія . Львів : ЛНУ імені Івана Франка, 2021. 424 с.
- Бук С. Н. Лексична основа української мови: виділення та системно-структурна організація : дис. ... канд. філол. н. Львівський національний університет імені Івана Франка, Львів, 2004. 281 с.
- Бук С. Н. Корпусно-лексикографічні та лінгвостатистичні виміри великої прози Івана Франка: словник і текст : дис. ... д-ра філол. наук : 10.02.01 ; Київ. нац. ун-т ім. Тараса Шевченка. Київ, 2021. 500 с.
- Бойко Ю. Диференційні параметри речення як детермінанта авторського стилю // Проблеми квантитативної лінгвістики. Чернівці: Рута, 2005. С. 292–305.
- Гіков Л. В. Особливості вживання лексико-граматичних і лексичних одиниць в авторському стилі : автореф. дис. ... канд. філол. наук. Львів : Чернівецький НУ ім. Ю. Федьковича, 2003. 20 с.
- Муравицька М. П. Статистичні лінгвістичні дослідження та їх розвиток в українському мовознавстві // Мовознавство. 1967. № 5. С. 47–58.
- Основні проблеми структурних і статистичних досліджень мови і мовлення // Розвиток мовознавства в УРСР 1967–1977. К.: Наукова думка, 1980. С. 177–185.
- Пасічник В., Щербина Ю., Висоцька В., Шестакевич Т. Математична лінгвістика. [Книга 1. Квантитативна лінгвістика] : [Навчальний посібник]. Львів : «Новий світ-2000», 2012. 359 с.
- Пасічник В., Щербина Ю., Висоцька В., Шестакевич Т. Математична лінгвістика. [Книга 2. Комбінаторна лінгвістика] : [Навчальний посібник]. Львів : «Новий світ-2000», 2019. 240 с.
- Перебийніс В. С. Математична лінгвістика: Навчальний посібник. К.: Вид. центр КНЛУ, 2014. 125 с.
- Перербийніс В. С. Кількісні та якісні характеристики системи фонем сучасної української літературної мови. К.: Наукова думка, 1970. С. 168–174.
- Перебийніс В. С. Теоретичні та прикладні проблеми структурно-математичної лінгвістики // Мовознавство. 1981. № 4.
- Розвиток структурно-математичної лінгвістики // Мовознавство на Україні за п'ятдесят років. К.: Наукова думка, 1967. С. 403–410.
- Тищенко В. Частота частин мови в різних функціональних стилях сучасної української мови // Питання структурної лексикології. К.: Наукова думка, 1970. С. 215–224.
- Caldeira S. M. G., Petit Lobão T. C., Andrade R. F. S., Neme A., Miranda J. G. V. The network of concepts in written texts // European Physical Journal. 2006. B 49. P. 523–529.
- Hammerl R., Sambor J. O statystycznych prawach językowych. Warszawa: Polskie towarzystwo semiotyczne, 1993. 161 s.
- Hammerl R., Sambor J. Statystyka dla językoznawców. Warszawa: PWN, 1990. 504 s.
- Köhler R. Bibliography of Quantitative Linguistics / with the assistance of Hoffmann C. Amsterdam ; Philadelphia, PA : J. Benjamin, 1995. 780 p.
- Lord R. D. Studies in the History of Probability and Statistics. VIII De Morgan and the statistical study of literary style // Biometrika. 1958. 45. P. 282.

- McMahon A., McMahon R. Language Classification by Numbers. Oxford : Oxford University Press, 2005. xvii, 265 pp.
- Oakes M. M. Statistics for Linguistics. Edinburg University Press, 2003. S. 251–257.
- Quantitative Linguistics: An International Handbook. Ed. by by Köhler R., Altmann G., Piotrowski R. Berlin ; New York : Walter de Gruyter, 2005. 1041 p.
- Rasinger S. M. Quantitative Research in Linguistics: An Introduction Research Methods in Linguistics. A&C Black, 2013. 304 p.
- Ruszkowski M. Statystyka w badaniach stylistyczno-składniowych. Kielce: Wydawnictwo Świętokrzyskie, 2004. 144 s.
- Těsitelová M. Quantitative Linguistics, John Benjamins Publishing Company, 2000.
- Wimmer G., Altmann G., Hřebíček L., Ondrejovič S., Wimmerová S. Úvod do analýzy textov. Bratislava: Veda, 2003. S. 18.

Перелік деяких електронних ресурсів:

- ГРАК: Генеральний регіональний анотований корпус української мови (ГРАК), 2017–2024 <www.uacorp.us.org>.
- КУМ (Чтиво) = Корпус української мови <korpus.org.ua>
- 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_5eEKQ7I9h6YCxZQwh5ix8>
- Корпуси української мови на SketchEngine <<https://www.sketchengine.eu/corpora-and-languages/ukrainian-text-corpora/>>
- Лабораторія української. Паралельні корпуси <https://mova.institute/%D0%BF%D0%B0%D1%80%D0%B0%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>
- Інструменти роботи зі системою SketchEngine (user guide) <<https://www.sketchengine.eu/guide/>>
- Skovoroda Online Concordance / Онлайн конкорданс Сковороди <<http://www.arts.ualberta.ca/~ukr/skovoroda/NEW/index.php>>
- Частотні словники різних стилів. Mova.info. Лінгвістичний портал (Інституту філології Київського національного університету імені Тараса Шевченка) <<http://www.mova.info/Page.aspx?l1=57>>
- Електронний словник мови Т. Шевченка (Інституту філології Київського національного університету імені Тараса Шевченка) <<http://www.mova.info/page.aspx?l1=220>>
- Електронні частотні словники текстів Б. Лепкого, М. Коцюбинського, І. Огієнка, Ю. Андруховича, Т. Прохаська, В. Стуса, Ю. Шевельова, І. Роздобудько <http://corpora.donnu.edu.ua/bonito/run.cgi/first_form> (Донецький національний університет ім. В. Стуса)
- Електронні частотні словники різних стилів (Донецький національний університет ім. В. Стуса) <http://corpora.donnu.edu.ua/bonito/run.cgi/first_form>
- Бук С., Ровенчак А. Онлайн-конкорданс роману Івана Франка «Перехресні стежки» <<http://ktf.franko.lviv.ua/users/andrij/science/Franko/concordance.html>>
- 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

	<p>паралельний корпус</p> <p><http://www.domeczek.pl/~polukr/index.php?option=welcome></p> <ul style="list-style-type: none"> • 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/> • Stylo: R package for stylometric analyses https://github.com/computationalstylistics/stylo
Тривалість курсу	Один семестр (3-й)
Обсяг курсу	Загальний обсяг - 90 годин, зокрема 32 години аудиторних, з них - 16 лекційних, 16 - практичних; 58 годин самостійної роботи; 3 кредити ЄКТС.
Очікувані результати навчання	<p>У результаті вивчення курсу студент/ка повинен/нна знати:</p> <ul style="list-style-type: none"> • найважливіші поняття і категорійний апарат квантитативної лінгвістики; • лексикографічні та електронні джерела знаходження кількісної інформації про мовні одиниці; • основні статистичні інструменти текстових корпусів, програм Sketch Engine, Stylo, графіків CQL, Word, Excel тощо; • ключові інструменти візуалізації статистичної інформації про текстові дані; • частотні словники української мови та їх визначальні характеристики; • основні підстави застосування квантитативних методів у мовознавстві; • розвиток та становлення квантитативних методів у мовознавстві; <p>поняття статистичної структури тексту, типологію та проблеми укладання частотних словників, методику їх порівняння;</p> <ul style="list-style-type: none"> • лінгвостатистичні характеристики, закони, залежності; • кількісні характеристики одиниць різних рівнів мови та мовлення. <p>уміти:</p> <ul style="list-style-type: none"> • знаходити необхідну кількісну інформацію про мовні одиниці в лексикографічних та електронних джерелах; • застосовувати квантитативні методи для аналізу тексту у своїй дослідницькій роботі та для розв'язання різних типів лінгвістичних завдань; • укладати і зіставляти частотні словники, правильно інтерпретувати їх інформацію; • виявляти і правильно інтерпретувати квантитативні дані про мовну одиницю, отримані з лексикографічних та електронних джерел, основні їх залежності і закони; • користуватися основними статистичними інструментами текстових корпусів, програм Sketch Engine, Stylo, графіків CQL, Word, Excel тощо; • візуалізувати статистичну інформацію про текстові дані за допомогою різних ресурсів; • оперувати найважливішими поняттями квантитативної лінгвістики; • аналізувати мовлення (текст) на предмет кількісних характеристик; • послуговуватися корпусами текстів для виявлення зазначених

	<p>характеристик на різних рівнях;</p> <ul style="list-style-type: none"> • критично оцінювати кількісні дані текстів, отримані з частотних словників, корпусів текстів та комп'ютерних програм. <p>Під час вивчення дисципліни студенти можуть сформувані такі загальні компетентності:</p> <p>ЗК 1. Здатність спілкуватися державною мовою як усно, так і письмово.</p> <p>ЗК 2. Здатність бути критичним і самокритичним.</p> <p>ЗК 3. Здатність до пошуку, опрацювання та аналізу інформації з різних джерел.</p> <p>ЗК 4. Уміння виявляти, ставити та вирішувати проблеми.</p> <p>ЗК 5. Здатність працювати в команді та автономно.</p> <p>ЗК 7. Здатність до абстрактного мислення, аналізу та синтезу.</p> <p>ЗК 8. Навички використання інформаційних і комунікаційних технологій.</p> <p>ЗК 12. Здатність генерувати нові ідеї (креативність).</p> <p>Фахові компетентності:</p> <p>ФК 4. Здатність здійснювати науковий аналіз і структурування мовного / мовленнєвого й літературного матеріалу з урахуванням класичних і новітніх методологічних принципів.</p> <p>ФК 6. Здатність застосовувати поглиблені знання з обраної філологічної спеціалізації «Прикладна лінгвістика» для вирішення професійних завдань.</p> <p>ФК 7. Здатність вільно користуватися спеціальною термінологією в обраній галузі філологічних досліджень.</p> <p>ФК 9. Здатність виявляти семантичні, синтаксичні, прагматичні й власне комунікативні чинники й категорії, притаманні різним типам мовленнєвих жанрів.</p> <p>ФК 10. Здатність виявляти імпліцитні (приховані) наміри мовця (адресанта, автора тексту, наративу) і відповідні комунікативні смисли їх утілення.</p> <p>Після вивчення дисципліни буде сформовано такі програмні результати навчання:</p> <p>ПРН 3. Застосовувати сучасні методиками і технології, зокрема інформаційні, для успішного й ефективного здійснення професійної діяльності та забезпечення якості дослідження в конкретній філологічній галузі.</p> <p>ПРН 13. Доступно й аргументовано пояснювати сутність конкретних філологічних питань, власну точку зору на них та її обґрунтування як фахівцям, так і широкому загалу, зокрема особам, які навчаються.</p> <p>ПРН 15. Обирати оптимальні дослідницькі підходи й методи для аналізу конкретного лінгвістичного чи літературного матеріалу.</p> <p>ПРН 16. Використовувати спеціалізовані концептуальні знання з обраної філологічної галузі для розв'язання складних задач і проблем, що потребує оновлення та інтеграції знань, часто в умовах неповної/недостатньої інформації та суперечливих вимог.</p> <p>ПРН 17. Планувати, організовувати, здійснювати і презентувати дослідження та/або інноваційні розробки в конкретній філологічній галузі.</p> <p>ПРН 19. Володіти методами, методиками і прийомами аналізу текстів (дискурсів, наративів) з метою виявлення імпліцитних (прихованих) намірів їх авторів.</p>
Ключові слова	Міждисциплінарні дослідження, частотний словник, вибірка, репрезентативність, квантитативна лінгвістика, статистична лінгвістика,

	частотність одиниць мови та мовлення, залежність, кореляція, візуалізація кількісних даних мови та мовлення, інтерпретація квантитативних характеристик мови та мовлення, стилеметрія,
Формат курсу	Очний/онлайн залежно від воєнної ситуації в Україні, лекційний курс з практичними заняттями
Теми	<p>* СХЕМА КУРСУ</p> <p>Модуль 1.</p> <p>Тема 1. Кількісний поворот у лінгвістиці. Квантитативна лінгвістика як наука та її основні завдання.</p> <p>Тема 2. Що і для чого рахувати у мові? – Основні поняття та категорійний апарат квантитативної лінгвістики: вибірка, частота, розподіл.</p> <p>Тема 3. Що показують частотні словники, як ними користуватися і укладати самим? – Частотні словники, їх типологія та проблеми укладання. Статистична структура тексту. Електронні інструменти укладання частотних списків.</p> <p>Тема 4. Який текст багатший? – Методика порівняння частотних словників. Лінгвостатистичні характеристики, закони, залежності.</p> <p>Модуль 2.</p> <p>Тема 5. Як визначити автора тексту? – Стилеметрія. Атрибуція тексту, її види (авторська, часова тощо). Визначення авторства тексту. Методи й інструменти атрибуції тексту.</p> <p>Тема 6. «Відбиток пальця» І. Франка у текстах – Лінгвостатистична модель ідіолекту: приклад І. Франка.</p> <p>Тема 7. Скільки слів потрібно, щоби вивчити мову? – Частотність як основний критерій компресії словника. Словники-мінімуми. Лексична основа української мови.</p> <p>Тема 8. Як поміряти мову? – Лінгвометрія. Кількісні параметри мови та мовлення. Лінгвостатистичні закони (переваги, Ципфа, Ципфа-Мандельброта). Лінгвометричні методи у виявленні варіативності української мови ХХ ст. на прикладі корпусу текстів.</p>
Підсумковий контроль, форма	Іспит
Пререквізити	«Теоретичне і прикладне мовознавство», «Іноземна мова за професійним спрямуванням», «Комунікативно-прагматичні аспекти текстотворення і текстосприйняття», «Семантико-текстуальна експертиза», «Міжкультурна комунікація та проблеми перекладу тексту».
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	<p><u>Методи навчання:</u> пояснювально-ілюстративний метод, метод проблемного викладання, евристичний метод, дослідницький метод, дискусія.</p> <p><u>Методи викладу навчального матеріалу</u> – лекція, консультація; дослідні роботи. <u>Наочні методи навчання</u> – ілюстрування (художні тексти), демонстрування з використанням технічних засобів (слайди), використання онлайн ресурсів (програм, текстових корпусів, електронних словників тощо), візуалізація кількісних текстових даних тощо.</p> <p><u>Техніки:</u> навчальної дискусії, створення проблемної ситуації, інтелектуального експерименту, командна робота/гра, соціолінгвістичний експеримент.</p>
Необхідне	Комп'ютер/ноутбук, проєктор, екран, доступ до інтернету

<p>обладнання</p> <p>Критерії оцінювання (окремо для кожного виду навчальної діяльності)</p>	<p>Максимальна оцінка, яку може отримати магістрант/ка, – 100 балів. Робота магістранта/ки оцінюється за такими показниками: а) поточна успішність за участь у практичних заняттях – 30 балів; б) письмові роботи за темою практичних занять – 20 балів; в) відповідь на іспиті – 50 балів.</p> <p><u>Бали поточної успішності за участь у практичних заняттях: 30 балів</u> – середньостатистичний бал за відповіді, доповнення і коментарі на практичних заняттях.</p> <p>«5» (27-30 балів) – студент повною мірою володіє навчальним матеріалом, самостійно та аргументовано його викладає під час усних і письмових відповідей, глибоко та аргументовано розкриває зміст теоретичних питань, послідовно виконує усі практичні завдання.</p> <p>«4» (22-26 балів) – студент добре володіє навчальним матеріалом, обґрунтовує свою думку під час усних і письмових відповідей, розкриває зміст теоретичних питань (однак недостатньо впевнено аргументує, недостатньо впевнено відповідає, допускає несуттєві неточності та незначні помилки), виконав більшість практичних завдань.</p> <p>«3» (21-15 балів) – студент володіє навчальним матеріалом на достатньому рівні, розуміє і викладає його основний зміст під час усних та письмових відповідей, однак невичерпно, без глибокого аналізу, обґрунтування та аргументації, допускає окремі суттєві неточності та помилки.</p> <p>«2» (9-14 балів) – студент засвоїв навчальний матеріал фрагментарно, поверхово, не може аргументувати та навести власний ілюстративний матеріал під час усних та письмових відповідей, слабо орієнтується в теоретичних питаннях та виконав лише кілька практичних завдань, допускає суттєві неточності.</p> <p>«1» (4-8 балів) – студент частково володіє навчальним матеріалом, не здатен викласти зміст більшості питань під час усних і письмових відповідей, допускає суттєві помилки.</p> <p>«0» (0-3 бали) – студент не володіє навчальним матеріалом, не може викласти зміст більшості питань теми під час усних і письмових відповідей. Не вирішив тестові завдання.</p> <p><u>Письмові роботи за темою практичних занять – 20 балів:</u> Студент вибирає за бажанням 4 теми практичних і пише письмові роботи із акцентом, як ці теми він/вона застосовує у своєму науковому дослідженні. Кожна письмова робота оцінюється у 5 балів.</p> <p><u>Іспит – 50 балів</u></p> <table border="1" data-bbox="422 1657 1506 2074"> <thead> <tr> <th data-bbox="422 1657 603 1697">Бали</th> <th data-bbox="603 1657 1506 1697">Критерії оцінювання</th> </tr> </thead> <tbody> <tr> <td data-bbox="422 1697 603 1809">40–50</td> <td data-bbox="603 1697 1506 1809">Відповіді, в яких навчальний матеріал відтворюється в повністю правильно, обґрунтовано, логічно, які містять аналіз і си аргументовані висновки. Засвідчено глибоке володіння матеріалом</td> </tr> <tr> <td data-bbox="422 1809 603 1888">25–39</td> <td data-bbox="603 1809 1506 1888">Відповіді, в яких відтворюється значна частина навчального матеріалу. Виявлено знання і розуміння основних положень навчальної дисципліни</td> </tr> <tr> <td data-bbox="422 1888 603 1966">14–24</td> <td data-bbox="603 1888 1506 1966">Відповіді, в яких основні положення навчального матеріалу відтворюються на рівні заучування без достатнього його розуміння.</td> </tr> <tr> <td data-bbox="422 1966 603 2045">1–13</td> <td data-bbox="603 1966 1506 2045">Відповіді, які засвідчують, що навчальний матеріал відсутній чіткого і логічного формулювання</td> </tr> <tr> <td data-bbox="422 2045 603 2074">0</td> <td data-bbox="603 2045 1506 2074">Відсутність відповіді.</td> </tr> </tbody> </table>	Бали	Критерії оцінювання	40–50	Відповіді, в яких навчальний матеріал відтворюється в повністю правильно, обґрунтовано, логічно, які містять аналіз і си аргументовані висновки. Засвідчено глибоке володіння матеріалом	25–39	Відповіді, в яких відтворюється значна частина навчального матеріалу. Виявлено знання і розуміння основних положень навчальної дисципліни	14–24	Відповіді, в яких основні положення навчального матеріалу відтворюються на рівні заучування без достатнього його розуміння.	1–13	Відповіді, які засвідчують, що навчальний матеріал відсутній чіткого і логічного формулювання	0	Відсутність відповіді.
Бали	Критерії оцінювання												
40–50	Відповіді, в яких навчальний матеріал відтворюється в повністю правильно, обґрунтовано, логічно, які містять аналіз і си аргументовані висновки. Засвідчено глибоке володіння матеріалом												
25–39	Відповіді, в яких відтворюється значна частина навчального матеріалу. Виявлено знання і розуміння основних положень навчальної дисципліни												
14–24	Відповіді, в яких основні положення навчального матеріалу відтворюються на рівні заучування без достатнього його розуміння.												
1–13	Відповіді, які засвідчують, що навчальний матеріал відсутній чіткого і логічного формулювання												
0	Відсутність відповіді.												

	<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 іспиті. При цьому обов'язково враховуються присутність на лекційних і практичних заняттях, активність студента/ки 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/ки 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 студенти повинні 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 іспиту</p>	<ul style="list-style-type: none"> • Підстави застосування квантитативних методів у гуманітаристиці й мовознавстві. • Об'єкт, предмет та завдання квантитативної лінгвістики. • Сфери застосування квантитативних методів у мовознавстві. • Зв'язок квантитативної лінгвістики з математичними науками. • Зв'язок квантитативної лінгвістики з філологічними науками. • Співвідношення понять «статистична лінгвістика», «обчислювальна лінгвістика», «квантитативна лінгвістика», «стилеметрія», «прикладна лінгвістика», «комп'ютерна лінгвістика». • Кількісний поворот у лінгвістиці (Kortmann 2021) • Основні тези статті Альмана І. „Мода та істина в лінгвістиці”. • Сучасний стан квантитативних досліджень мови у світовому мовознавстві. • Сучасний стан квантитативних досліджень мови в українському мовознавстві. • Основні поняття та категорійний апарат квантитативної лінгвістики. • Поняття частотності, її різновиди. Фактори впливу на коливання частотності. • Поняття вибірки та підвибірки. Правила організації вибірки. Різновиди вибірок. • Лінгвістична однорідність вибірки, її типи. Величина вибірки. • Обчислення коефіцієнта варіації абсолютної частоти. Коефіцієнт

стабільності.

- Поняття точності статистичного дослідження. Відносна похибка. Стандартна похибка.
- Залежності між явищами у вибірці: детермінація, кореляція (пряма, обернена), спряжені ознаки.
- Коефіцієнт рангової кореляції Спірмена, коефіцієнт кореляції.
- Поняття непараметричних методів дослідження. Метод експертних оцінок.
- Закон Ципфа та його інтерпретації в інших галузях знань.
- Необхідні умови для порівняння частотних словників.
- Залежність «словник–текст». Закон переваги.
- Закон Менцера(–Альтманна).
- Закон Крилова.
- Визначення відносної неточності дослідження.
- Кількісні характеристики одиниць різних рівнів мови та мовлення.
- Співвідношення якісних і кількісних характеристик одиниць мови та мовлення.
- Кількісні характеристики звуків та літер.
- Кількісні характеристики морфем та граматичних категорій.
- Кількісні характеристики слів. Закон Ципфа.
- Кількісні характеристики синтаксичних конструкцій.
- Темп мовлення-думки.
- Індекс синтетичності мови.
- Стилеметрія. Атрибуція тексту, її види (авторська, часова тощо). Визначення авторства тексту.
- Проблеми укладання частотних словників мови письменника.
- Аналіз частотних словників мови письменників та їх окремих творів.
- Індекс графоманства (теорія Юнгварта–Десса).
- Мовне профілювання. Питання авторства Гомера, В. Шекспіра, Марка Вовчка.
- Лінгвостатистична модель ідіолекту: приклад І. Франка. Її основні параметри.
- Квантитативні морфологічні характеристики великої прози І. Франка (індекси номіналізації, дієслівних означень, епітетизації);
- Індекс діалогічності великої прози І. Франка (співвідношення прямого та авторського мовлення).
- Частотні словники та проблеми їх укладання.
- Проблема співвідношення текстів різних стилів (жанрів, форм) для укладання частотного словника усієї мови. Приклади її вирішення в зарубіжному та українсько-му мовознавстві.
- Ручне versus формалізоване укладання частотного словника.
- Поняття лематизації слів. Проблеми розрізнення лексичних та граматичних омонімів; окремих значень багатозначних слів.
- Характеристика Частотного словника сучасної української художньої прози / За ред. В. С. Перебийніс.
- Характеристика частотних словників публіцистичного, розмовно-побутового, наукового, офіційно-ділового стилів української мови.
- Обчислення індексів різноманітності та концентрації, середньої повторюваності слова у тексті.
- Вимоги до порівняння статистичних характеристик частотних словників.
- Частотні словники мови письменника.

	<ul style="list-style-type: none"> • Частотність як об'єктивний критерій для компресії словника. Мінімізовані словники. • Співвідношення понять «словник-мінімум», «компресований словник», «мінімізований словник», «лексична основа мови». • Лексична основа української мови. • Поняття покриття тексту. Його обчислення. • Принципи укладання мінімізованих словників. • Словник-мінімум і розуміння тексту. • Кількісний склад корпусних ресурсів української мови, їх орієнтовна статистика. • Квантитативний аналіз вибірок (підкорпусів) електронних текстових ресурсів української мови. • Частотні словники української мови, їх аналіз та основні параметри. • Методи й інструменти атрибуції тексту. • Ключові інструменти візуалізації статистичної інформації про текстові дані. • Електронні ресурсів візуалізації кількісних даних сполучуваності мовних одиниць.
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

СХЕМА КУРСУ

Тиж. / дата / год.-	Тема, план, короткі тези	Форма діяльності (заняття)	Література	Завдання, год	Термін виконання
Модуль I Змістовий модуль 1. Кількісний поворот у лінгвістиці. Основні поняття та категорійний апарат квантитативної лінгвістики					
1-й т. 2 год.	<p>Тема 1. Кількісний поворот у лінгвістиці. Квантитативна лінгвістика як наука та її основні завдання.</p> <p>1. Динамічні та статистичні закони. Поняття про статистичний метод та статистику. Загальні риси статистичного методу в різних галузях знань.</p> <p>2. Кількісний поворот у лінгвістиці. Підстави застосування квантитативних методів у мовознавстві.</p> <p>3. Квантитативна лінгвістика як дисципліна, що вивчає кількісні закономірності природної мови. Об'єкт, предмет та основні поняття статистичної лінгвістики.</p> <p>4. Співвідношення понять «статистична лінгвістика», «лінгвістична статистика», «математична лінгвістика», «квантитативна лінгвістика», «комбінаторна лінгвістика», «комп'ютерна лінгвістика», «обчислювальна лінгвістика», «стилеметрия», «лінгвومتрія», «прикладна лінгвістика».</p> <p>5. Зв'язок квантитативної лінгвістики</p>	Лекція	<p><i>Основна література:</i></p> <p>1. Альтман Г. Мода та істина в лінгвістиці // Проблеми квантитативної лінгвістики. Чернівці: Рута, 2005. С. 3–11.</p> <p>2. Бук С. Основи статистичної лінгвістики: Навчально-методичний посібник. ЛНУ імені Івана Франка, 2008. 124 с.</p> <p>3. Перебийніс В. С. Статистичні методи для лінгвістів: Навчальний посібник. Вінниця: «Нова книга», 2002. С. 3–8.</p> <p>4. Перебийніс В. С. Математична лінгвістика // Українська мова: Енциклопедія / Редкол.: В. М. Русанівський та інші. К.: Українська енциклопедія, 2000. С. 302–287.</p> <p>5. Перебийніс В. С. Теоретичні та прикладні проблеми структурно-математичної лінгвістики // Мовознавство. 1981. № 4.</p> <p>6. Щербина Ю. М. Предмет</p>	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3,6 год.	1-й т.

	<p>з математичними науками: математичною статистикою, теорією інформації, теорією ймовірності та ін.</p> <p>6. Зв'язок квантитативної лінгвістики з філологічними науками: фонетикою, морфе-мікою, лексикою, лексикографією, синтаксисом, теорією тексту. Застосування її результатів в інших філологічних науках.</p>		<p>математичної лінгвістики // Вісник НУ «Львівська політехніка». Інформаційні системи та мережі. 2002. № 464. С. 340–349.</p> <p>7. Kortmann В. Reflecting on the Quantitative Turn in Linguistics // Linguistics. 59 (5): P. 1212.</p> <p><i>Додаткова література:</i></p> <p>1. Пасічник В., Щербина Ю., Висоцька В., Шестакевич Т. Математична лінгвістика. [Книга 1. Квантитативна лінгвістика] : [Навчальний посібник]. Львів : «Новий світ-2000», 2012. 359 с.</p> <p>2. Пасічник В., Щербина Ю., Висоцька В., Шестакевич Т. Математична лінгвістика. [Книга 2. Комбінаторна лінгвістика] : [Навчальний посібник]. Львів : «Новий світ-2000», 2019. 240 с.</p> <p>3. Köhler R. Bibliography of Quantitative Linguistics / with the assistance of Hoffmann С. Amsterdam ; Philadelphia, PA : J. Benjamin, 1995. 780 p.</p> <p>4. Wimmer G., Altmann G., Hřebíček L., Ondrejovič S., Wimmerová S. Úvod do analýzy textov. Bratislava: Veda, 2003. S. 18.</p>		
2-й т. 2 год.	Квантитативна лінгвістика як наука та її основні завдання. Кількісний поворот у лінгвістиці.	Практичне заняття	1. Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024	Навчитися аналізувати кількісний склад	2-й т.

	<p>Статистика корпусних ресурсів української мови. Визначення орієнтовної теми індивідуального проєкту.</p>		<p><www.uacorporus.org>. 2. КУМ (Чтиво) = Корпус української мови <korpus.org.ua> 3. 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_5eEKQ7l9h6YCxZQwh5ix8> 4. Частотні словники різних стилів. Mova.info. Лінгвістичний портал (Інституту філології Київського національного університету імені Тараса Шевченка) <http://www.mova.info/Page.aspx?l1=57> 5. Електронні частотні словники текстів Б. Лепкого, М. Коцюбинського, І. Огієнка, Ю. Андруховича, Т. Прохаська, В. Стуса, Ю. Шевельова, І. Роздобудько <http://corpora.donnu.edu.ua/bonito/run.cgi/first_form> (Донецький національний університет ім. В. Стуса) 6. Електронні частотні словники різних стилів (Донецький національний університет ім. В. Стуса) <http://corpora.donnu.edu.ua/bonito/run.cgi/first_form> +Л-ра до лекції 1</p>	<p>корпусних ресурсів української мови, визначати їх статистику. 3,6 год.</p>	
3-й т. 2 год.	<p>Тема 2. Що і для чого рахувати у мові? – Основні поняття та категорійний апарат квантитативної лінгвістики: вибірка, частота, розподіл. Звідки брати статистику мовних</p>	Лекція	<p><i>Основна література:</i> 1. Бук С. Основи статистичної лінгвістики: Навчально-методичний посібник. Видавничий центр ЛНУ імені Івана Франка, 2008. 124 с.</p>	<p>Опрацювання відповідних розділів рекомендованої літератури,</p>	3-й т.

	<p>одиниць і про що вона свідчить. Аналіз вибірок (підкорпусів) електронних текстових ресурсів української мови.</p> <ol style="list-style-type: none"> 1. Поняття генеральної сукупності. Поняття вибірки та підвибірки. Правила організації вибірки, її різновиди. Правила організації механічної, випадкової, зональної вибірок. 2. Лінгвістична однорідність вибірки (хронологічна, жанрова, тематична). 3. Поняття частотності. Абсолютна й відносна частота. 4. Середня частота. Середнє квадратичне відхилення. Міра коливання середньої частоти. Стандартна похибка відхилення середньої. 5. Розподіл одиниць у вибірці. Рівномірність розподілу. Коефіцієнт розповсюдженості. 6. Поняття відносної неточності статистичного дослідження. Відносна похибка. 7. Варіація, однорідність, залежності між явищами у вибірці, непараметричні методи дослідження 		<ol style="list-style-type: none"> 2. Пасічник В., Щербина Ю., Висоцька В., Шестакевич Т. Математична лінгвістика. [Книга 1. Квантитативна лінгвістика] : [Навчальний посібник]. Львів : «Новий світ-2000», 2012. 359 с. 3. Пасічник В., Щербина Ю., Висоцька В., Шестакевич Т. Математична лінгвістика. [Книга 2. Комбінаторна лінгвістика] : [Навчальний посібник]. Львів : «Новий світ-2000», 2019. 240 с. 4. Перебийніс В. С. Математична лінгвістика: Навчальний посібник. К.: Вид. центр КНЛУ, 2014. 125 с. 5. Перебийніс В. С. Статистичні методи для лінгвістів: Навчальний посібник. Вінниця: «Нова книга», 2002. С. 9–60. <p><i>Додаткова література:</i></p> <ol style="list-style-type: none"> 1. Hammerl R., Sambor J. Statystyka dla językoznawców. Warszawa: PWN, 1990. 504 s. 2. Hammerl R., Sambor J. O statystycznych prawach językowych. Warszawa: Polskie towarzystwo semiotyczne, 1993. 161 s. 3. Guiraud P. Zagadnienia i metody statystyki językoznawczej. Warszawa: PWN, 1966. 159 s. 4. Glossary // Oakes M. M. Statistics for Linguistics. Edinburg University Press, 2003. S. 251–257. 5. Wimmer G., Altmann G., Hřebíček L., Ondrejovič S., Wimmerová S. Úvod do analýzy textov. Bratislava: Veda, 2003. S. 18. 	<p>інтернет-джерел, конспекту лекції. Квантитативний аналіз вибірок (підкорпусів) електронних текстових ресурсів української мови. 3,6 год.</p>	
4-й т. 2 год.	Основні поняття та категорійний апарат статистичної лінгвістики: вибірка,	Практичне заняття	1. Генеральний регіонально анотований корпус української мови (ГРАК), 2017–2024	Виявити основні	4-й т.

	<p>частота, розподіл. Звідки брати статистику мовних одиниць і про що вона свідчить. Аналіз вибірок (підкорпусів) електронних ресурсів української мови.</p>		<p><www.uacorpus.org>. 2. КУМ (Чтиво) = Корпус української мови <korpus.org.ua> 3. 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_5eEKQ7l9h6YCxZQwh5ix8> 4. Частотні словники різних стилів. Mova.info. Лінгвістичний портал (Інституту філології Київського національного університету імені Тараса Шевченка) <http://www.mova.info/Page.aspx?l1=57> 5. Електронні частотні словники текстів Б. Лепкого, М. Коцюбинського, І. Огієнка, Ю. Андруховича, Т. Прохаська, В. Стуса, Ю. Шевельова, І. Роздобудько <http://corpora.donnu.edu.ua/bonito/run.cgi/first_form> (Донецький національний університет ім. В. Стуса) 6. Електронні частотні словники різних стилів (Донецький національний університет ім. В. Стуса) <http://corpora.donnu.edu.ua/bonito/run.cgi/first_form> + Л-ра до лекції 2</p>	<p>квантитативні характеристики текстів на прикладі запропонованих електронних корпусів та словників. Відповідь обґрунтувати. 3,6 год.</p>	
<p>5-й т. 2 год.</p>	<p>Тема 3. Що показують частотні словники, як ними користуватися і укладати самим? – Частотні словники, їх типологія та проблеми укладання. Статистична структура тексту. Електронні інструменти укладання частотних списків.</p>	<p>Лекція</p>	<p>1. Частотний словник сучасної української художньої прози / За ред. Перейбийніс В. С. К.: Наукова думка, 1981. Т. 1, 2. 2. Частотний словник. Розділ "художня проза" корпусу <http://www.mova.info/cfq1.aspx?fdid=hproz20></p>	<p>Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел,</p>	<p>5-й т.</p>

<p>1. Поняття статистичної структури тексту. Статистичні параметри стилів на рівні фонем, складів, дієслівних словоформ, речень.</p> <p>2. Частотний словник (ЧС) і його роль у виявленні статистичної структури тексту на лексичному рівні. Типологія частотних словників (за одиницею підрахунку, характером вибірки, кількістю мов, за способом фіксації мовлення (письмове чи усне) і т. д.).</p> <p>3. Методика укладання частотних словників. Паперове versus електронне укладання ЧС.</p> <p>4. Проблема джерел ЧС ідіотнічної мови та їх обсягу.</p> <p>5. Проблема співвідношення текстів різних стилів (жанрів, форм) для укладання ЧС усієї мови. Приклади її вирішення в зарубіжному та українському мовознавстві.</p> <p>6. Поняття лематизації. Обсяг словника словоформ, обсяг словника лексем. Проблеми розрізнення лексичних та граматичних омонімів; окремих значень багатозначних слів.</p> <p>7. Поняття рангу слова, проблема присвоєння рангу словам з однаковими часто-тами.</p> <p>8. Аналіз ЧС в українській (ЧС сучасної української художньої прози, української публіцистики, розмовно-побутового, наукового та офіційно-ділового стилів української мови.</p>		<p>18></p> <p>3. Частотний словник сучасної української публіцистики / Дарчук Н. П. та ін. // Лінгвістичний портал MOVA.info. 2003. <http://www.mova.info/freqcard.aspx?l1=98&sl=publicist>.</p> <p>4. Дарчук Н. П. Частотний словник сучасної поетичної української мови. <http://www.mova.info/Page2.aspx?l1=89>.</p> <p>5. Частотний словник. Розділ "наукові тексти" корпусу <http://www.mova.info/cfq1.aspx?fdid=nauk2018>.</p> <p>6. Частотний словник. Розділ "законодавчі тексти" корпусу <http://www.mova.info/cfq1.aspx?fdid=zakon2018>.</p> <p>7. Бук С. 3 000 найчастотніших слів наукового стилю сучасної української мови. Львів: ЛНУ імені Івана Франка, 2006.</p> <p>8. Бук С. 3 000 найчастотніших слів розмовно-побутового стилю сучасної української мови. Львів: ЛНУ імені Івана Франка, 2006.</p> <p>9. Бук С. Частотний словник офіційно-ділового стилю: принципи укладання та статистичні характеристики // Лінгвістичні студії: Зб. наук. праць. 2006.</p>	<p>конспекту лекції. 3,6 год.</p>	
--	--	--	---------------------------------------	--

	Обернений словник) та в зарубіжній лексикографії.		Випуск 14. С. 184–188. 10. Обернений частотний словник сучасної української художньої прози / Т. О. Грязнухіна (уклад.). К.: Спалах, 1998. 11. Korpusomat < https://korpusomat.eu/ > 12. Laurence A. AntConc (Windows, MacOS, Linux). Build 4.1.4 < https://www.laurenceanthony.net/software/antconc/releases/AntConc414/help.pdf >		
6-й т. 2 год.	Частотні словники, їх типологія та проблеми укладання. Статистична структура тексту. Електронні інструменти укладання частотних списків. Практичне укладання частотного словника обраного тексту. Командна робота – співвідношення стилів у репрезентативному частотному словнику мови.	Практичне заняття	Л-ра до Лекції 3.	Проаналізувати запропоновані частотні словники за усіма можливими параметрами. Обговорити співвідношення стилів у репрезентативному частотному словнику мови. 3,6 год.	6-й т.
7-й т. 2 год.	Тема 4. Який текст багатший? – Методика порівняння частотних словників. Лінгвостатистичні характеристики, закони, залежності. 1. Співвідношення якісних і кількісних характеристик одиниць мови та мовлення. Закон переваги. 2. Кількісні характеристики звуків та літер. 3. Кількісні характеристики морфем	Лекція	<i>Основна література:</i> 1. Бук С. Основи статистичної лінгвістики: Навчально-методичний посібник. Видавничий центр ЛНУ імені Івана Франка, 2008. 124 с. 2. Дарчук Н. П. Статистичні характеристики лексики як відображення структури тексту // Мовознавчі студії. К.: Наукова думка, 1976. С. 97–102. 3. Васильєв О., Чалий О., Васильєва І.	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3,6 год.	7-й т.

	<p>та граматичних категорій.</p> <p>4. Кількісні характеристики слів. Функціональна залежність частотність–ранг (закон Ципфа–Мандельброта). Залежність частотність–кількість значень у слові. Співвідношення частотності абстрактної й конкретної лексики. Різниця між частотністю слова та його наявністю в пам'яті людини (за Р. Мішеа). Індекс синонімічності.</p> <p>5. Кількісні характеристики синтаксичних конструкцій.</p> <p>6. Роль статистичних параметрів різних рівнів у розмежуванні стилів.</p> <p>7. «Темп мовлення-думки».</p> <p>8. Індекс синтетичності мови.</p>		<p>Математичні методи й моделі в лінгвістиці // Україна модерна. Справа Verba et numeri (кількісні підходи до аналізу мови й тексту). 2019. Ч. 27. С. 9–28.</p> <p><i>Додаткова література:</i></p> <p>1. Широков В. Еволюція як універсальний природний закон (пролегомени до майбутньої загальної теорії еволюції). Лекція 03.04.2018, 23.04.2018, 26.04.2018. <http://ulif.org.ua/page/konfieriensiia-ukrayinska-nauka-v-informatsiinih-vimirakh-aktual-ni-vikliki-s-oghodiennia></p> <p>2. Buk S. N., Rovenchak A. A. Rank–Frequency Analysis for Functional Style Corpora of Ukrainian // Journal of Quantitative Linguistics. 2004. V. 11, No. 3. P. 161–171.</p> <p>3. Ferrer-i-Cancho R., Elvevåg B. Random texts do not exhibit the real Zipf's law-like rank distribution. PLoS ONE. 2010. 5 (3). e941. P. 1–10.</p> <p>4. Pawłowski A. Metody kwantytatywne w sekwencyjnej analizie tekstu. Warszawa: UW, 2001. 168 s.</p>		
<p>8-й т. 2 год.</p>	<p>Методика порівняння частотних словників. Лінгвостатистичні характеристики і закони. Визначення найважливіших характеристик свого укладеного власноруч</p>	<p>Практичне заняття</p>	<p>Л-ра до Лекції 4.</p>	<p>Опрацювання л-ри. Визначення найважливіших характеристик свого</p>	<p>8-й т.</p>

	частотного словника.			укладеного власноруч частотного словника. 3,6 год.	
Модуль II					
Змістовий модуль 2. Лінгвометрія, стилеметрія, словник-мінімум.					
9-й т. 2 год.	<p>Тема 5. Як визначити автора тексту? – Стилеметрія. Атрибуція тексту, її види (авторська, часова тощо). Визначення авторства тексту. Методи й інструменти атрибуції тексту.</p> <ol style="list-style-type: none"> 1. Об'єкт і предмет дослідження стилеметрії. 2. Сфери застосування стилеметрії. 3. Атрибуція тексту, її види (авторська, часова тощо). Визначення авторства тексту. 4. Пакет комп'ютерних програм Stylo і кластеризація текстів 5. Формула визначення графоманії 6. Огляд ЧС мови письменників та їх окремих творів. 7. Мовно-правописні проблеми джерел укладання ЧС. 8. Мовне профілювання. Питання авторства Гомера, В. Шекспіра, Марка Вовчка. 	Лекція	<p><i>Основна література</i></p> <ol style="list-style-type: none"> 1. Бук С. Основи статистичної лінгвістики: Навчально-методичний посібник. Видавничий центр ЛНУ імені Івана Франка, 2008. 124 с. 2. Бойко Ю. Диференційні параметри речення як детермінанта авторського стилю // Проблеми квантитативної лінгвістики. Чернівці: Рута, 2005. С. 292–305. 3. Гіков Л. В. Особливості вживання лексико-граматичних і лексичних одиниць в авторському стилі : автореф. дис. ... канд. філол. наук. Львів : Чернівецький НУ ім. Ю. Федьковича, 2003. 20 с. 4. Дарчук Н. П. Статистичні характеристики лексики як відображення структури тексту // Мовознавчі студії. К.: Наукова думка, 1976. С. 97–102. 5. Перебийніс В. С. Статистична стилістика // Українська мова: Енциклопедія, 2004. С. 644–645. 6. Тараненко О. О. Мови письменника словник // Українська мова: Енциклопедія, 2004. С. 326–327. 7. Burrows J. 'Delta' a Measure of Stylistic Difference and a Guide to Likely Authorship. Literary and Linguistic Computing. 2002. 17/3, 1 	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3,6 год.	9-й т.

		<p>September: 267–287.</p> <p>8. Eder M., Rybicki J., Kestemont M. Stylometry with R: a package for computational text analysis // R Journal. 2016. 8(1). 107–121. <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html></p> <p>9. Stylo: R package for stylometric analyses https://github.com/computationalstylistics/stylo</p> <p>10. Lahjouji-Seppälä M., Rabus A., von Waldenfels R. Ukrainian standard variants in the 20th century: stylometry to the rescue <https://www.researchgate.net/publication/364332887_Ukrainian_standard_variants_in_the_20th_century_stylometry_to_the_rescue?fbclid=IwAR3bNDRrWLNUXtyPF0SZ5dID_NZaFYfVXCACULYtMOhjr5vqse2TY48u5c></p> <p><i>Додаткова література:</i></p> <p>1. Білоус М. До питання мовно-правописного редагування Франкових творів // Іван Франко — письменник, мислитель, громадянин. Матеріали міжнародної наукової конференції (Львів 25–27 вересня 1996) / Ред. М. Ільницький, Б. Якимович. Львів: Світ, 1998. С. 633–636.</p> <p>2. Волошиновська І. А. Сильова, тематична й авторська атрибуція наукових і художніх текстів (на матеріалі англійської, німецької та української мов) : автореф. дис. ... канд. філол. наук. К, 2013. 18 с.</p> <p>3. Дарчук Н., Лангенбах М. Електронний</p>		
--	--	---	--	--

		<p>словник мови Тараса Шевченка: методика і технології укладання Українське мовознавство : міжвідом. наук. зб. / М-во освіти і науки України, Київ. нац. ун-т ім. Тараса Шевченка. Київ : Київський університет, 2014. Вип. 1 (44). С. 106–114. <http://www.mova.info/corpus_papers/darchuk-langenbakh-slovnyk2.pdf></p> <p>4. Електронний словник мови Тараса Шевченка <http://www.mova.info/cfqsh.aspx></p> <p>5. Невідомська Л. Лінгвістичні проблеми редагування мови поетичних творів Івана Франка // Іван Франко — письменник, мислитель, громадянин. Матеріали міжнародної наукової конференції (Львів 25–27 вересня 1996) / Ред. М. Ільницький, Б. Якимович. Львів: Світ, 1998. С. 714–721.</p> <p>6. Полюга Л. Статистичний аналіз лексики поетичних творів І. Франка // Іван Франко і національне відродження. Львів: ЛДУ ім. І Франка; Ін-т франкознавства, 1991. С. 164–166.</p> <p>7. Buk, S., Rovenchak, A. Statistical Parameters of Ivan Franko's Novel Perekhresni stezky (The Cross-Paths) // Quantitative Linguistics. V. 62: Exact Methods in the Study of Language and Text. Berlin; New York: Mouton de Gruyter, 2007. P. 39–48.</p>		
--	--	---	--	--

			<p>8. Hilpert M., Correia Saavedra D. Using token-based semantic vector spaces for corpus-linguistic analyses: From practical applications to tests of theoretical // <i>Corpus Linguistics and Ling. Theory</i> 2020. 16(2). 393–424.</p> <p>9. Literary detective work // Oakes M. M. <i>Statistics for Linguistics</i>. Edinburg University Press, 2003. 287 s. S. 199–246.</p> <p>10. Shapiro J. <i>Contested Will: Who Wrote Shakespeare?</i>— UK edition: Faber and Faber, 2010.</p>		
10-й т. 2 год.	<p>Стилеметрія. Атрибуція тексту, її види (авторська, часова тощо). Визначення авторства тексту.</p> <p>Практичні методи й інструменти атрибуції тексту.</p>	Практичне заняття	Л-ра до Лекції 5	Засвоїти теоретичні знання. Оволодіти методами й інструментами атрибуції тексту. 3,6 год.	10-й т.
11-й т. 2 год.	<p>Тема 6. «Відбиток пальця» І. Франка у текстах – Лінгвостатистична модель ідіолекту: приклад І. Франка. Кількісно-візуальний розподіл прозових творів письменника; морфологічні характеристики (індекси номіналізації, дієслівних означень, епітетизації); індекс діалогічності (співвідношення прямого та авторського мовлення великої прози І. Франка).</p> <p>Ключові інструменти візуалізації статистичної інформації про текстові дані.</p>	Лекція	<p><i>Основна література:</i></p> <p>1. Бук С. Велика проза Івана Франка: електронний корпус, частотні словники та інші міждисциплінарні контексти : монографія . Львів : ЛНУ імені Івана Франка, 2021. С.7–261.</p> <p>2. Бук С. Н. Корпусно-лексикографічні та лінгвостатистичні виміри великої прози Івана Франка: словник і текст : Автореф. дис. ... д-ра філол. наук : 10.02.01 ; Київ. нац. ун-т ім. Тараса Шевченка. Київ, 2021. 33 с.</p> <p>3. Бук С. Частотний словник повісті</p>	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3,6 год.	11-й т.

	<ol style="list-style-type: none"> 1. Застосування елементів кількісного аналізу до творчості І. Франка. 2. Методика дослідження. 3. Статистична структура тексту великої прози І. Франка. 4. Кількісно-візуальний розподіл прозових творів письменника. 5. Кількісні морфологічні характеристики великої прози І. Франка. Індекс номіналізації. Індекс дієслівних означень. Індекс епітетизації. 6. Співвідношення прямого та авторського мовлення великої прози І. Франка. 7. Зіставлення частотного словника великої прози І. Франка з Частотним словником художньої прози сучасної української літературної мови 		<p>Івана Франка «Voа constrictor» (редакція 1884 р.) // Стежками Франкового тексту (комунікативні, лінгвoseміотичні, когнітивні та лінгвостатистичні виміри прози). Львів : ЛНУ імені Івана Франка, 2013. С. 202–501.</p> <p>4. Бук С. Частотний словник роману Івана Франка «Основи суспільності»: Інтерпретація твору крізь призму статистичної лексикографії. Львів : ЛНУ імені Івана Франка, 2012. 264 с.</p> <p>5. Бук С., Ровенчак А. Частотний словник повісті І. Франка «Перехресні стежки» // Стежками Франкового тексту (комунікативні, стилістичні та лексикографічні виміри роману «Перехресні стежки»). Львів : ЛНУ імені Івана Франка, 2007. С. 138–369.</p> <p><i>Додаткова література:</i></p> <ol style="list-style-type: none"> 1. Ковалик І. І., Ощипко І. Й., Полюга Л. М. (уклад.) Лексика поетичних творів Івана Франка : метод. вказівки з розвитку лексики. Львів : ЛНУ, 1990. 264 с. 2. Терлак З. Словник мови поетичної збірки Івана Франка «Зів'яле листя» = Glossary of Ivan Franco's Collection of Poetry «Faded Leaves». Львів : ЛНУ імені Івана Франка, 2010. 394 с. 325. 3. Терлак З. Словник мови поетичної збірки Івана Франка «Мій Ізмарagd» : у 2 ч. Ч. 1 : А–М. Львів : Априорі, 2016. 424 с. Ч. 2 : Н–Я. Львів : ЛНУ імені Івана Франка, 2016. 428 с. 4. Частотний словник сучасної 		
--	---	--	---	--	--

			української художньої прози / За ред. Перебийніс В. С. К.: Наукова думка, 1981. Т. 1, 2.		
12-й т. 2 год.	<p>Лінгвостатистична модель ідіолекту. Приклад І. Франка. Кількісно-візуальний розподіл прозових творів письменника; морфологічні характеристики (індекси номіналізації, дієслівних означень, епітетизації); індекс діалогічності (співвідношення прямого та авторського мовлення великої прози І. Франка).</p> <p>Потенціал моделі на прикладі творчості інших письменників.</p> <p>Ключові інструменти візуалізації статистичної інформації про текстові дані.</p>	Практичне заняття	Л-ра до Лекції 6.	Засвоїти теоретичні знання. Опанувати використання ключових інструментів візуалізації статистичної інформації про текстові дані 3,6 год.	12-й т. 2 год
13-й т. 2 год.	<p>Тема 7. Скільки слів потрібно, щоби вивчити мову? – Частотність як основний критерій компресії словника. Словники-мінімуми. Лексична основа української мови.</p> <p>Електронні ресурси візуалізації кількісних даних сполучуваності мовних одиниць.</p> <p>1. Співвідношення понять «словник-мінімум», «компресований словник», «міні-мізований словник», «лексична основа мови», «редукований словник», «базовий словник», «словник для спеціальних цілей», «словник на перші потреби».</p> <p>2. Поняття покриття тексту та його</p>	Лекція	<p><i>Основна література:</i></p> <p>1. Ардан В. Р., Бацевич Ф. С., Партико З. В. Комп'ютерний словник-мінімум української мови // Мовознавство. 1996. № 4–5. С. 34–40.</p> <p>2. Бук С. Основи статистичної лінгвістики: Навчально-методичний посібник. Видавничий центр ЛНУ імені Івана Франка, 2008. 124 с.</p> <p>3. Бук С. Н. Лексична основа української мови: виділення та системно-структурна організація. Рукопис. Дисертація ... канд. філол. н. Львівський національний університет імені Івана Франка, Львів, 2004. 281 с.</p>	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3,6 год.	13-й т.

	<p>обчислення.</p> <p>3. Покриття та розуміння тексту.</p> <p>4. Принципи укладання мінімізованих словників.</p> <p>5. Проблеми застосування статистичних методів для викладання іноземної мови: парадокси «статистики і системи», «ймовірності та інформації», «конфлікт» статистичної структури ЧС та окремого тексту.</p> <p>6. Прочитність тексту і найчастотніша лексика.</p>		<p><i>Додаткова література:</i></p> <p>1. Бук С. До розрізнення деяких термінів на позначення мінімалізованого словника // Теорія і практика викладання української мови як іноземної. 2006. Вип. 1. С. 76–79. www.academia.edu.</p> <p>2. Бук С. Лексична основа української мови як компресована мовна модель // Вісник Львівського університету. Серія філологічна. Випуск 34. Львів: ЛНУ імені Івана Франка, 2004. С. 289–293. www.academia.edu.</p> <p>3. Бук С. Лексичний мінімум: Теоретичні засади і принципи виділення // Вісник: Проблеми української термінології. Львів: Національний університет «Львівська політехніка». 2002. № 453. С. 474–479. www.academia.edu.</p> <p>4. Бук С. Лексичні мінімуми в слов'янській лексикографічній традиції (на прикладі польської, російської та української практики) // Проблеми слов'янознавства. Випуск 54. Львів, 2004. С. 273–274. www.academia.edu.</p> <p>5. Manning C. D., Schütze H. Foundations of Statistical Natural Language Processing. Cambridge; Massachusetts; London: The MIT Press, 2000. 680 p.</p>		
--	--	--	---	--	--

14-й т. 2 год.	<p>Частотність – критерій компресії словника. Словники-мінімуми. Лексична основа української мови.</p> <p>Електронні ресурси візуалізації кількісних даних сполучуваності мовних одиниць. Тестування.</p>	Практичне заняття	<p>1. 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_5eEKQ719h6YCxZQwh5ix8></p> <p>2. Laurence A. AntConc (Windows, MacOS, Linux). Build 4.1.4 <https://www.laurenceanthony.net/software/antconc/releases/AntConc414/help.pdf> + Л-ра до лекції 7.</p>	Засвоїти теоретичний матеріал. Опанувати використання електронних ресурсів візуалізації кількісних даних сполучуваності мовних одиниць. 3,6 год.	14-й т.
15-й т. 2 год.	<p>Тема 8. Як поміряти мову? – ЛінгвOMETрія. Кількісні параметри мови та мовлення. Лінгвостатистичні закони (переваги, Ципфа, Ципфа-Мандельброта). ЛінгвOMETричні методи у виявленні варіативності української мови ХХ ст. на прикладі корпусу текстів.</p> <p>1. Співвідношення якісних і кількісних характеристик одиниць мови та мовлення. Закон переваги.</p> <p>2. Кількісні характеристики звуків та літер.</p> <p>3. Кількісні характеристики морфем та граматичних категорій.</p> <p>4. Кількісні характеристики слів. Функціональна залежність частотність–ранг (закон Ципфа–Мандельброта). Залежність частотність–кількість значень у слові. Співвідношення частотності</p>	Лекція	<p><i>Основна література:</i></p> <p>1. Бук С. Основи статистичної лінгвістики: Навчально-методичний посібник. Видавничий центр ЛНУ імені Івана Франка, 2008. 124 с.</p> <p>2. Перербийніс В. С. Кількісні та якісні характеристики системи фонем сучасної української літературної мови. К.: Наукова думка, 1970. С. 168–174.</p> <p>3. Статистичні параметри стилів / за ред. В. С. Перебийніс. К.: Наукова думка, 1967. 260 с.</p> <p>4. Структура мови і статистика мовлення / Відп. ред. канд. філол. наук М. П. Муравицька і Л. А. Олексієнко. К.: Наукова думка, 1974. 176 с.</p> <p>5. Тищенко В. Частота частин мови в різних функціональних стилях сучасної української мови // Питання структурної лексикології. К.: Наукова думка, 1970. С.</p>	Опрацювання відповідних розділів рекомендованої літератури, інтернет-джерел, конспекту лекції. 3,6 год.	15-й т.

	<p>абстрактної й конкретної лексики. Різниця між частотністю слова та його наявністю в пам'яті людини (за Р. Мішеа). Індекс синонімічності.</p> <p>5. Кількісні характеристики синтаксичних конструкцій.</p> <p>6. Роль статистичних параметрів різних рівнів у розмежуванні стилів.</p> <p>7. «Темп мовлення-думки».</p> <p>8. Індекс синтетичності мови.</p> <p>9. Лінгвометричні методи у виявленні варіативності української мови ХХ ст. на прикладі Генерального регіонально анотованого корпусу української мови (ГРАК).</p>		<p>215–224.</p> <p><i>Додаткова література:</i></p> <p>1. Вашак П., Перебийніс В. Розподіл довжини слова у слов'янських мовах // Питання структурної лексикології. К.: Наукова думка, 1970. С. 178–186.</p> <p>2. Дарчук Н. П., Грязнухіна Т. О. Частотний словник сучасної української публіцистики // Мовознавство. 1996. №4–5. С. 15–19.</p> <p>3. Перебийніс В.С., Кадомцева Л.О. Методи структурного дослідження мови. Київ: Наукова думка, 1968. 346 с.</p> <p>4. Широков В. А. Дослідження інформаційно-статистичних властивостей української мови. Розподіл ймовірностей символів українського тексту, отриманий на основі експериментальної вибірки // Широков В. А. Інформаційна теорія лексикографічних систем. К.: Довіра, 1998. С. 109–117; с. 268.</p> <p>5. Ruszkowski M. Statystyka w badaniach stylistyczno-składniowych. Kielce: Wydawnictwo Świętokrzyskie, 2004. 144 s.</p>		
16-й т. 2 год.	<p>Лінгвометрія. Кількісні параметри мови та мовлення.</p> <p>Представлення індивідуальних проєктів.</p>	Практичне заняття	Л-ра до Лекції 8.	Підготувати індивідуальний дослідницький проєкт застосування квантитативних методів у своїй науковій роботі 4 год.	16-й т.