

INFOTECH IN HUMANITIES

Томчаковська Ю.О., Строченко Л.В.,
Швелідзе Л.Д., Ясніцька Г.В.



Міністерство освіти і науки України
Національний університет «Одеська юридична академія»
Кафедра іноземних мов

**Ю. О. Томчаковська, Л. В. Строченко,
Л. Д. Швелідзе, Г. В. Ясніцька**

INFOTECH in Humanities

Практикум для здобувачів другого (магістерського) рівня вищої освіти
спеціальності 035 «Філологія» / В 11 «Філологія (за спеціалізаціями)»



УДК 004:8(076)

I-74

Рекомендовано до видання Навчально-методичною радою Національного університету «Одеська юридична академія» (протокол № 6 від 29.04.2025 р.)

Укладачі:

Томчаковська Ю. О. – завідувачка кафедри іноземних мов, кандидат філологічних наук, доцент (<https://orcid.org/0000-0003-0117-2704>);

Строченко Л. В. – професор кафедри іноземних мов, доктор філологічних наук, доцент (<https://orcid.org/0000-0003-3619-3484>);

Швелідзе Л. Д. – доцент кафедри іноземних мов, кандидат філологічних наук, доцент (<https://orcid.org/0000-0002-8145-9483>);

Ясніцька Г. В. – викладач кафедри іноземних мов (<https://orcid.org/0009-0000-6373-7024>)

Рецензенти:

Мамич М. В. – доктор філологічних наук, доктор, завідувач кафедри прикладної лінгвістики Національного університету «Одеська юридична академія»;

Станко Д. В. – кандидат філологічних наук, доцент, доцент кафедри англійської філології Ужгородського національного університету

*Укладачі висловлюють глибоку подяку Президенту
Національного університету «Одеська юридична академія»,
академіку Сергію Васильовичу Ківалову за сприяння в публікації видання*

INFOTECH in Humanities : практикум для здобувачів другого
I-74 (магістерського) рівня вищої освіти спеціальності 035 «Філологія» /
В 11 «Філологія (за спеціалізаціями)» / уклад. Ю. О. Томчаковська,
Л. В. Строченко, Л. Д. Швелідзе, Г. В. Ясніцька. Суми : Університетська
книга, 2025. 68 с.

ISBN 978-617-521-112-0

Практикум для здобувачів другого (магістерського) рівня вищої освіти спеціальності 035 «Філологія» / В 11 «Філологія (за спеціалізаціями)». У практикумі пропонується план лекцій, питання для самостійного опанування, завдання для практичного застосування здобутих навичок з курсу «Інформаційно-комунікаційні технології у професійній діяльності» та вдосконалення навичок використання сучасних інформаційно-комунікативних технологій у науковій та викладацькій діяльності.

УДК 004:8(076)

© Ю. О. Томчаковська, Л. В. Строченко,
Л. Д. Швелідзе, Г. В. Ясніцька, 2025

ISBN 978-617-521-112-0

© ПФ «Видавництво “Університетська книга”», 2025

ПЕРЕДМОВА

Сучасна філологія дедалі більше інтегрується з інформаційними технологіями, що сприяє ефективному аналізу текстів, автоматизації процесів перекладу та дослідженню літературних і лінгвістичних особливостей. Використання цифрових інструментів дозволяє розширити можливості філологічних досліджень, забезпечуючи доступ до великих корпусів текстів, автоматизованих систем аналізу та обробки мовних даних.

Метою цього практикуму є надання здобувачам вищої освіти необхідних знань і навичок для застосування інформаційно-комунікаційних технологій у філологічних та наукових дослідженнях. Видання містить теоретичний матеріал та практичні завдання, що охоплюють ключові аспекти цифрової гуманітаристики, літературного аналізу, корпусної лінгвістики, машинного перекладу та автоматизованих засобів мовного аналізу. Особливу увагу приділено методології цифрових досліджень, аналізу літературних текстів та обробці великих масивів мовних даних.

Практикум призначений для студентів, викладачів і дослідників, які працюють у сфері філології, літературознавства, перекладу та цифрових гуманітарних наук. Окрім наукових досліджень, цей практикум стане у нагоді викладачам для вдосконалення навчальних методик, адаптації матеріалів до цифрового середовища та впровадження інноваційних підходів до викладання філологічних дисциплін. Використання сучасних інформаційних технологій сприятиме створенню інтерактивних навчальних матеріалів, персоналізованого навчального досвіду та ефективного оцінювання знань.

Матеріали структуровані за тематичними розділами, що охоплюють основні аспекти сучасних інформаційних технологій у філології, такі як штучний інтелект у літературознавстві, методи корпусного аналізу, спеціалізовані програмні засоби для обробки текстів, а також етичні аспекти цифрових досліджень.

Цей практикум є корисним інструментом для тих, хто прагне підвищити рівень своєї цифрової компетентності, розширити методологічний арсенал у галузі сучасної філології та застосувати передові технології у наукових дослідженнях і викладацькій діяльності.

ЗМІСТ

Новітні ІКТ у сучасній філології. Етика цифрових досліджень і використання ІКТ у науковій діяльності.....	5
Основи корпусної лінгвістики. Використання корпусів у лінгвістичних дослідженнях.....	9
Стратегії роботи з лінгвістичними корпусами (BNC, COCA). Аналіз даних та інтерпретація результатів дослідження.....	13
Особливості машинного та автоматизованого перекладів. Практичні підходи для перекладачів	18
Основи використання програм для аналізу текстів (AntConc, Voyant Tools), візуалізація та інтерпретація даних.....	20
Штучний інтелект у філології як інструмент нової епохи	23
Цифрові інструменти Google. Перспективи застосування ІКТ у філології та освіті	27
 ПРАКТИЧНА ЧАСТИНА	
Робота з British National Corpus (BNC)	33
Аналіз американського національного корпусу COCA.....	37
Принципи роботи з українським національним корпусом. Корпус ГРАК	43
Створення власного корпусу для досліджень	46
Аналітика текстів за допомогою AntConc.....	49
Візуалізація текстових даних за допомогою VOYANT TOOLS та інших інструментів	53
Інструменти ШІ для аналізу текстів і редагування.....	56
Основи роботи з програмами для перекладацької практики	58
Цифрові інструменти Google Workspace для освіти. Інтеграція з сервісами	62
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....	65

Новітні ІКТ у сучасній філології. Етика цифрових досліджень і використання ІКТ у науковій діяльності

<i>ІКТ</i>	<i>Штучний інтелект</i>	<i>Цифрова гуманітаристика</i>
<i>Великі мовні моделі (LLM)</i>		<i>Генеративний ІІІ (Generative AI)</i>
<i>Машинне навчання (ML)</i>		<i>Глибинне навчання (Deep Learning)</i>
<i>Корпусні менеджери</i>	<i>SAT-інструменти</i>	<i>Термінологічні менеджери</i>

1. **Значення ІКТ у філології.** Цифрова гуманітаристика як міждисциплінарна галузь

2. **Штучний інтелект у філологічних дослідженнях і перекладі.** Можливості ІІІ та переваги використання у науковій діяльності.
Типологія систем ІІІ в контексті філології:

- Великі мовні моделі (LLM)
- Генеративний штучний інтелект (Generative AI)
- Системи машинного навчання (ML)
- Нейромережі глибинного навчання (Deep Learning)

Спеціалізовані інструменти для філологічних досліджень:

- Корпусні менеджери (AntConc, SketchEngine, WordSmith Tools)
- Національні лінгвістичні корпуси та бази даних (BNC, COCA, Google Books NgramViewer)
- Програми лінгвістичного аналізу (NLTK, Stanford NLP, SpaCy)
- Інструменти візуалізації (Voyant Tools, GraphColl, R)

3. **Сучасні технології перекладу**

- Системи машинного перекладу (DeepL, GNMT, Microsoft Translator)

- CAT-інструменти (SDL Trados Studio, MemoQ, Memsource, Wordfast)
- Термінологічні менеджери та інструменти забезпечення якості (MultiTerm, Xbench).

4. Методологія застосування ІКТ у наукових дослідженнях

- Основні етапи дослідження (підготовчий, основний, завершальний).
- Верифікація результатів (тріангуляція, експертна оцінка, статистична валідація).

5. Етичні аспекти використання ІКТ у науковій діяльності

- Нормативно-правові засади (захист персональних даних, академічна доброчесність).
- Методологічна прозорість (документування, валідація методів, відтворюваність).
- Практичні рекомендації (ліцензійне ПЗ, конфіденційність).

6. Перспективи розвитку

- Мультимодальні підходи, нейролінгвістичні моделі, аналіз Big Data
- Методологічні виклики (репрезентативність, мовна упередженість, стандартизація).

Перелік визначень

ІКТ (Інформаційно-комунікаційні технології) – сукупність технологій та засобів, що забезпечують пошук, збирання, обробку, зберігання та розповсюдження інформації в цифровій формі.

Цифрова гуманітаристика (Digital Humanities) – міждисциплінарна галузь, що поєднує традиційні методи дослідження в гуманітарних науках з інноваційними комп'ютерними технологіями.

Штучний інтелект (ШІ) – галузь інформатики, що займається розробкою систем, здатних виконувати завдання, які потребують інтелектуальної діяльності людини (розпізнавання мовлення, прийняття рішень, обробка природної мови).

Великі мовні моделі (LLM) – алгоритми з обробки природної мови, збудовані на трансформерній архітектурі (GPT-4, Claude, BERT), здатні аналізувати та генерувати текст із високим ступенем зрозумілості й контекстуалізації.

Генеративний штучний інтелект (Generative AI) – клас ШІ-систем, що можуть створювати новий контент (текст, зображення, аудіо) на основі навчальних даних (ChatGPT, Bard, DALL-E).

Системи машинного навчання (Machine Learning) – алгоритми, які «навчаються» на великих масивах даних, виявляючи закономірності та роблячи прогнози без детального програмування кожного кроку.

Нейромережі глибокого навчання (Deep Learning) – багаторівневі штучні нейронні мережі, здатні до високоточного розпізнавання патернів у великих обсягах даних завдяки значній кількості шарів і вагових коефіцієнтів.

CAT-інструменти (Computer-Assisted Translation) – програмні засоби для комп'ютерного супроводу перекладу, які зберігають напрацьовані переклади в базах пам'яті (Translation Memory), що підвищує ефективність і якість роботи перекладача.

Тріангуляція методів – підхід у науковому дослідженні, що передбачає використання кількох методів аналізу одних і тих же даних з метою підвищення достовірності результатів.

Методологічна прозорість – практика чіткого опису методів дослідження, використаних інструментів і процедур, що забезпечує відтворюваність та перевірку результатів.

Питання для самостійного дослідження та самоперевірки

1. Визначити, як ІКТ сприяють підвищенню ефективності навчального процесу у філології.
2. Охарактеризувати роль цифрової гуманітаристики (Digital Humanities) в сучасних лінгвістичних дослідженнях.
3. Перелічити основні типи систем штучного інтелекту, що використовуються у філології, та визначити їхні особливості.
4. Навести приклади використання великих мовних моделей (LLM) для компаративного аналізу текстів.
5. Визначити різницю між генеративним ШІ та системами машинного навчання.
6. Перелічити можливості, які надають філологу корпусні менеджери (AntConc, SketchEngine).
7. Перелічити переваги та недоліки машинного перекладу порівняно з професійним людським перекладом.
8. Охарактеризувати призначення CAT-інструментів та визначити, як саме вони допомагають перекладачеві.
9. Перелічити етичні виклики, що виникають під час використання ІКТ у науковій діяльності.
10. Визначити поняття методологічної прозорості та пояснити її важливість для відтворюваності досліджень.

Основи корпусної лінгвістики. Використання корпусів у лінгвістичних дослідженнях

Корпус	Загальний (репрезентативний) корпус.	Розмітка корпусу
Спеціалізований корпус	Паралельний корпус	Корпус усної мови
Частотний аналіз	Лексикографія	
Колокації	Конкорданс	Розмітка корпусу



1. Вступ до корпусної лінгвістики: визначення, завдання, історичний розвиток
2. Особливості корпусного методу у мовознавстві.
3. Одиниці корпусного аналізу. Частотний аналіз, колокації, конкорданс.
4. Лексикографія та її зв'язок із корпусним аналізом.
5. Типологія корпусів:
 - Загальні (репрезентативні) корпуси

- Спеціалізовані корпуси
 - Паралельні корпуси
 - Діалектні корпуси
 - Корпуси усної мови та корпуси соціальних мереж
- 6. Методологічні засади корпусних досліджень та інструменти корпусної лінгвістики**
- Збір даних та принципи формування корпусів
 - Морфологічна та синтаксична розмітка
- 7. Програмні засоби корпусного аналізу**
- Автоматизовані засоби аналізу AntConc, Sketch Engine, WordSmith Tools, Treetagger тощо)
- 8. Використання корпусів у лінгвістичних дослідженнях**
- Аналіз частотності та вивчення колокацій
 - Перекладацькі дослідження: паралельні корпуси
 - Лексикографія та створення електронних словників
 - Дослідження мовної варіативності
- 9. Практичні приклади корпусних досліджень**
- Аналіз політичної риторики, нових запозичень
 - Освітні проекти
- 10. Сучасні перспективи та виклики корпусної лінгвістики**
- міждисциплінарні підходи; розвиток технологій (машинне навчання, штучний інтелект)
 - інтеграція в освітній процес та цифрову гуманітаристику

Перелік визначень

Корпус – зібрання письмових або усних текстів, спеціально відібраних та упорядкованих для лінгвістичного аналізу.

Текстовий корпус – добірка текстів одного жанру, стилю чи тематики. Дає змогу зосередитися на певному типі мовних матеріалів для детальнішого дослідження.

Загальний (репрезентативний) корпус – великий набір текстів, що охоплює різноманітні жанри, теми та стилі. Використовується для виявлення загальних мовних закономірностей у мові.

Спеціалізований корпус – корпус, укладений із текстів певної галузі (наукової, юридичної, медичної тощо). Дозволяє глибоко дослідити професійну лексику й термінологію.

Паралельний корпус – набір текстів на двох чи більше мовах із взаємними перекладами. Використовується для порівняльного аналізу оригінальних та перекладних текстів.

Діалектний корпус – зібрання текстів, що відображають особливості різних діалектів чи регіональних варіантів однієї мови. Дозволяє вивчати діалектну варіативність.

Корпус усної мови – містить транскрипти усних виступів, розмов, інтерв'ю. Сприяє аналізу природного (спонтанного) мовлення та його специфічних ознак (інтонація, паузи, повтори).

Корпус соціальних мереж – містить пости, коментарі, повідомлення з платформ соціальних медіа (Twitter, Facebook тощо). Корисний для дослідження сучасних тенденцій у мові, інтернет-сленгу та новітніх комунікативних практик.

Частотний аналіз – метод підрахунку та порівняння частоти вживання конкретних слів, фраз чи граматичних конструкцій у корпусі.

Колокація – стійке сполучення слів, що часто зустрічається в певному контексті (наприклад, “приймати рішення”, “важливе значення”).

Конкорданс – перелік усіх вживань певного слова чи фрази з невеликими контекстами з корпусу; дозволяє швидко побачити оточення шуканого елемента.

Лексикографія – наука про укладання словників, яка, опираючись на корпусні дані, аналізує реальне вживання слів, їхні семантичні та граматичні характеристики.

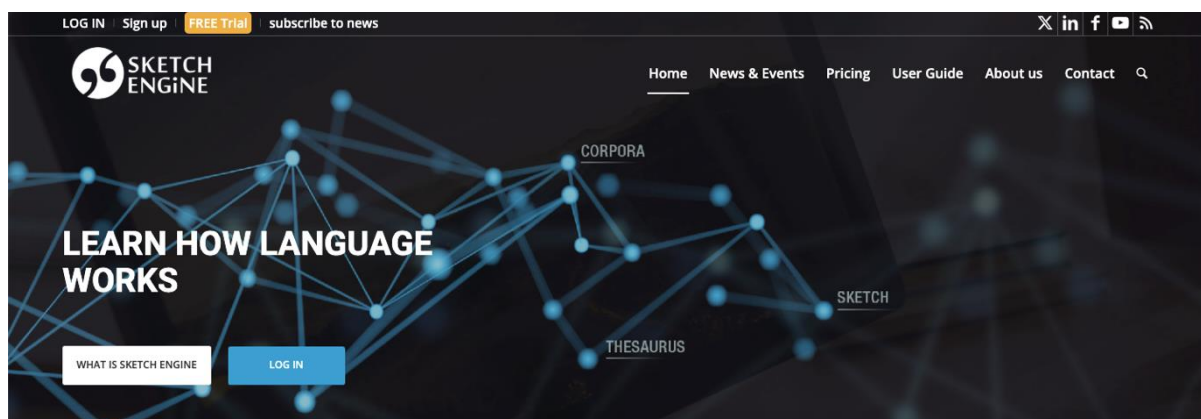
Розмітка корпусу – процес додавання спеціальних міток (частин мови, синтаксичних структур) до текстів у корпусі, що полегшує детальний лінгвістичний аналіз.

Питання для самостійного дослідження та самоперевірки

1. Визначити поняття корпусної лінгвістики та вказати її головні відмінності від традиційного мовознавства.
2. Визначити різницю між загальним (репрезентативним) та спеціалізованим корпусом; навести приклади спеціалізованих корпусів.
3. Пояснити, для чого лінгвісти застосовують аналіз частотності, та перелічити практичні результати цього методу.
4. Визначити поняття «колокація» та охарактеризувати її значення для лексикографії й викладання іноземних мов.
5. Перелічити основні переваги та недоліки паралельного корпусу й визначити, у яких випадках його використання є найбільш доцільним.
6. Охарактеризувати необхідність розмітки корпусу та перелічити основні її види.

7. Перелічити відомі вам інструменти для корпусного аналізу та визначити їхні основні функції й особливості.
8. Перелічити етапи та принципи, які потрібно враховувати під час створення власного корпусу.
9. Визначити, як корпусна лінгвістика може сприяти розвитку автоматизованих систем перекладу та розпізнавання мовлення.
10. Охарактеризувати перспективи розвитку корпусної лінгвістики у контексті міждисциплінарних досліджень та освіти.

Стратегії роботи з лінгвістичними корпусами (BNC, COCA). Аналіз даних та інтерпретація результатів дослідження



<i>Corpus</i>	<i>BNC</i>	<i>COCA</i>	<i>Exact Match</i>	<i>Lemma</i>
<i>POS search</i>	<i>MI-score</i>	<i>T-score</i>	<i>NGRAMs</i>	<i>Dispersion</i>

1. Вступ до корпусної лінгвістики

- Поняття лінгвістичного корпусу
- Значення корпусів у сучасних лінгвістичних дослідженнях.

2. Національні корпуси

British National Corpus (BNC) та його особливості:

історія створення та обсяги даних; структура корпусу (жанри, письмові та усні тексти); типи розмітки (морфологічна, синтаксична, соціолінгвістична).

Corpus of Contemporary American English (COCA):

опис та масштаби корпусу; баланс жанрів та систематичні оновлення; можливості аналізу сучасної американської англійської.

3. Методи пошуку та інструменти аналізу.

- Базові методи пошуку (точний, за лемою, за частиною мови)
- Розширені методи пошуку (регулярні вирази, контекстуальний пошук).
- Колокації, частотні списки, n-грами, конкорданс, дисперсійний аналіз.

4. Аналіз та інтерпретація даних.

- Кількісний аналіз (частота, статистичні методи)
- Якісний аналіз (контекстуальний та семантичний)
- Принципи інтерпретації результатів і типові помилки.

5. Практичні рекомендації для корпусних досліджень.

- Планування дослідження (цілі, методологія, часові рамки)
- Документування процесу (протокол, збереження даних).
- Валідація результатів (перехресна перевірка, порівняння з іншими джерелами).

6. Висновки

- Значущість корпусних ресурсів BNC та COCA
- Необхідність комплексного підходу до аналізу та інтерпретації
- Подальші перспективи досліджень.

Перелік визначень

Лінгвістичний корпус (Corpus) – це структурована електронна база текстів, зібраних з метою вивчення мовних закономірностей і тенденцій. Корпуси зазвичай мають розмітку (морфологічну, синтаксичну тощо) та відображають реальні мовні дані з різних джерел.

British National Corpus (BNC) – великий лінгвістичний корпус (понад 100 млн слів) британської англійської, що охоплює письмові та усні тексти різних жанрів, створений у 1990-х роках для відображення сучасної британської англійської.

Corpus of Contemporary American English (COCA) – найбільший корпус американської англійської мови (понад 1 млрд слів), що регулярно оновлюється і збалансований за жанрами (розмовне мовлення, художня література, журнали, газети, академічні тексти).

Точний пошук (Exact Match) – пошук конкретної словоформи з урахуванням пунктуації та регістру. Наприклад, пошук “run” знайде лише “run” без “runs” чи “running”.

Пошук за лемою (Lemma Search) – пошук усіх словоформ певної лексики. Наприклад, [run] охопить “run”, “runs”, “ran”, “running”.

Пошук за частиною мови (POS Search) – пошук слів певних частин мови з використанням тегів, наприклад [vv] для всіх дієслівних форм.*

Колокації (Collocations) – статистично значуще співвживання двох або більше слів, яке можна вимірювати коефіцієнтами *MI-score*, *t-score* та іншими показниками.

MI-score (Mutual Information) – показник сили асоціації між словами. *MI* > 3 зазвичай свідчить про значущу колокацію.

t-score – показник надійності колокації, де $t > 2$ вважається значущим. Враховує розмір корпусу.

Log-likelihood – альтернативна міра асоціації колокацій, менш чутлива до розміру корпусу, добре працює з рідкісними вживаннями.

Частотні списки (Frequency Lists) – списки слів з указанням частоти їх вживання у корпусі, можуть бути абсолютними або нормалізованими (на мільйон слів), а також порівняльними.

N-грами (N-grams) – послідовності з *N* слів, наприклад біграми (2 слова), триграми (3 слова), що дозволяють вивчати фразеологію та типові мовні патерни.

Конкорданс (Concordance) – відображення слова у всіх контекстах його вживання (*KWIC* – Key Word In Context), що допомагає швидко аналізувати значення та вживання слова.

Дисперсійний аналіз (Dispersion) – дослідження рівномірності розподілу слова чи фрази по корпусу, аналіз жанрових відмінностей та можливих кластерів вживання.

Кількісний аналіз – аналіз мовних явищ на основі підрахунку частот, використання статистичних тестів (χ^2 , *t*-тест, логарифмічна функція правдоподібності).

Якісний аналіз – зосереджується на семантичних та контекстуальних характеристиках мовних явищ, включно з аналізом жанрової специфіки та відтінків значення.

Інтерпретація результатів – процес пояснення та оцінки отриманих даних. Вимагає об'єктивності, врахування контексту та статистичної значущості.

Питання для самостійного дослідження та самоперевірки



1. Дати визначення корпусам BNC і COCA за структурою та складом текстів.
2. Встановити переваги використання корпусних методів дослідження порівняно з традиційними (наприклад, ручний аналіз тексту).
3. Описати варіанти розмітки корпусів (морфологічна, синтаксична, метатекстова) та її вплив на результати пошуку.
4. Встановити які різновиди пошуку (точний, за лемою, за частиною мови) є найдоцільнішими для вирішення певних лінгвістичних завдань? Навести приклади.
5. Як інтерпретувати MI-score, t-score та log-likelihood при вивченні колокацій.

6. Розкрити різницю між простою (raw) та нормалізованою частотою (frequency per million words) та запити використання кожної з них.
7. Дати визначення KWIC (Key Word In Context) і як конкорданс може допомогти у вивченні полісемії.
8. Встановити чому важливо проводити як кількісний, так і якісний аналіз даних корпусу.
9. Виокремити основні помилки які можуть виникати при інтерпретації результатів корпусного дослідження.
10. Описати планування та документування корпусного дослідження, щоб результати були відтворюваними та верифікованими.

Особливості машинного та автоматизованого перекладів. Практичні підходи для перекладачів

<i>Машинний переклад (МП)</i>	<i>Автоматизований переклад</i>	<i>CAT-tools</i>
<i>Нейронні мережі</i>	<i>Постредагування</i>	<i>Контроль якості</i>
<i>Пам'ять перекладів</i>	<i>SDL Trados Studio</i>	<i>DeepL</i>

1. **Роль технологій у сучасній перекладацькій діяльності.** Можливості та виклики машинного й автоматизованого перекладу
2. **Визначення основних понять.** Машинний переклад (МП). Автоматизований переклад (АП)
3. **Еволюція машинного перекладу.** Історичний розвиток та сучасний стан.
4. **Особливості нейронного машинного перекладу.** Принципи роботи. Переваги та обмеження.

5. **Системи автоматизованого перекладу (CAT-tools).** Основні компоненти. Функціональність пам'яті перекладів.
6. **Практичні підходи для перекладачів.** Інтеграція МП у робочий процес. Використання CAT-інструментів. Постредагування машинного перекладу
7. **Майбутні тенденції:** Розвиток технологій. Зміни у професії перекладача.
8. **Практичні рекомендації.** Вибір інструментів та організація робочого процесу. Шляхи вдосконалення автоматизованого перекладу.

Перелік визначень

Машинний переклад (МП) – процес автоматичного перекладу тексту з однієї мови на іншу за допомогою комп'ютерних програм.

Автоматизований переклад (АП) – переклад, який виконується з використанням програмного забезпечення, що допомагає перекладачеві.

CAT-tools – системи автоматизованого перекладу, що включають пам'ять перекладів, термінологічні бази, редактори перекладу та інструменти контролю якості.

Пам'ять перекладів – база даних, що зберігає попередні переклади для їх повторного використання.

Нейронний машинний переклад (НМП) – підхід до перекладу, який базується на використанні штучних нейронних мереж для аналізу тексту.

Питання для самостійного дослідження та самоперевірки

1. Визначити основну різницю між машинним і автоматизованим перекладом.
2. Охарактеризувати ключові етапи розвитку машинного перекладу.

3. Дати визначення пам'яті перекладів та її використання в САТ-інструментах.
4. Перелічити переваги нейронного машинного перекладу порівняно зі статистичним.
5. Назвати основні обмеження нейронного машинного перекладу.
6. Перелічити основні функції забезпечують сучасні САТ-інструменти.
7. Як інтеграція машинного перекладу змінює робочий процес перекладача?
8. Тенденції розвитку у сфері машинного перекладу.
9. Проаналізувати процес постредагування перекладу.
10. Охарактеризувати найпопулярніші інструментів машинного перекладу.

Основи використання програм для аналізу текстів (AntConc, Voyant Tools), візуалізація та інтерпретація даних

<i>Лінгвістичний аналіз</i>	<i>Стилістичний аналіз</i>	<i>Дискурс-аналіз</i>
<i>Літературознавчий аналіз</i>	<i>Соціолінгвістика</i>	
<i>Перекладознавство</i>	<i>AntConc</i>	<i>Voyant Tools</i>
	<i>Візуалізація даних</i>	<i>Інтерпретація даних</i>

1. Вступ

- Сучасна філологія та інтеграція цифрових технологій.
- Основні цілі та завдання аналізу текстів.

2. Напрямки використання аналізу текстів

- Лінгвістичний аналіз: фонетичний, морфологічний, синтаксичний, лексичний, семантичний.

- Стилiстичний аналіз: тропи, фігури мовлення, стилістичні прийоми.
- Літературознавчий аналіз: композиція, сюжет, персонажі, хронотоп.
- Соціолінгвістика: мовні зміни, діалекти, мовні норми.
- Дискурс-аналіз: мовленнєві акти, комунікативні стратегії.
- Історична лінгвістика: еволюція мови, зміни в лексиці та граматиці.
- Перекладознавство: порівняння оригіналів та перекладів.
- Фольклористика: усна народна творчість, мотиви, сюжети.

3. Комп'ютерні інструменти аналізу тексту

- AntConc: конкордансний аналіз, частотний аналіз, колокації, кластерний аналіз.
- SketchEngine: пошук по корпусу, аналіз колокацій, word sketches.
- Voyant Tools: хмари слів, графіки, мережеві графи, теплові карти.
- Інші програми: WordSmith Tools, LancsBox, Gephi.

4. Візуалізація та інтерпретація даних

- Методи візуалізації: хмари слів, графіки, мережеві графи, теплові карти.
- Інтерпретація результатів: контекст, жанрові особливості.
- Обмеження методів.

5. Перспективи розвитку комп'ютерного аналізу тексту

- Інтеграція машинного навчання та штучного інтелекту.
- Глибокий аналіз тексту (deep text analysis).
- Освітні й дослідницькі застосування.

6. Висновки

- Синтез традиційних методів та новітніх технологій.
- Важливість коректної інтерпретації даних.
- Цифрова грамотність філолога XXI століття.

Перелік визначень

Лінгвістичний аналіз – вивчення мовних одиниць та їх функціонування в тексті.

Стилістичний аналіз – дослідження особливостей використання мовних засобів для досягнення естетичного або прагматичного ефекту.

Літературознавчий аналіз – вивчення художніх особливостей тексту, включаючи композицію, сюжет, персонажів та хронотоп.

Соціолінгвістика – вивчення мовних змін, діалектів, мовних норм та їхнього відображення в текстах.

Дискурс-аналіз – дослідження тексту в його соціальному та культурному контексті.

Історична лінгвістика – вивчення еволюції мови, змін у лексиці та граматиці.

Фольклористика – вивчення усної народної творчості, мотивів та сюжетів.

Конкордансний аналіз – дослідження контекстів вживання слів та словосполучень.

Частотний аналіз – визначення частотності вживання слів у тексті.

Колокації – стійкі словосполучення та особливості сполучуваності слів.

Кластерний аналіз – знаходження повторюваних послідовностей слів.

Хмара слів – візуальне представлення частотності слів у тексті.

Мережевий граф – візуалізація зв'язків між елементами тексту.

Питання для самостійного дослідження та самоперевірки

1. Назвати основні зміни в методах філологічного аналізу, що відбулися з появою цифрових інструментів.
2. Перелічити переваги та обмеження автоматизованого аналізу тексту порівняно з традиційним.

3. Охарактеризувати основні функції програми AntConc та їх застосування для стилістичного аналізу художніх текстів.
4. Назвати основні методи візуалізації даних та пояснити їх роль в інтерпретації результатів текстового аналізу.
5. Перелічити інструменти Voyant Tools та опишіть їх застосування для аналізу тематичної структури тексту.
6. Охарактеризувати можливості застосування програм аналізу текстів у перекладознавчих дослідженнях.
7. Назвати фактори, які необхідно враховувати при інтерпретації результатів комп'ютерного аналізу тексту.
8. Перелічити основні напрямки використання штучного інтелекту в філологічному аналізі.
9. Описати методiku використання мережевих графів для аналізу системи персонажів у художньому творі.
10. Охарактеризувати найбільш ефективні методи візуалізації результатів лінгвістичного аналізу.

Штучний інтелект у філології як інструмент нової епохи

<i>Токенізація</i>	<i>Стилометричний аналіз</i>	<i>Нейромережеві системи</i>
<i>Мовні моделі</i>	<i>Deep Learning</i>	<i>Нейронний машинний переклад</i>

1. Вступ

2. Теоретичні засади та історичний контекст

- Еволюція комп'ютерної лінгвістики
- Фундаментальні принципи роботи нейромережевих систем
- Технологічна база сучасних мовних моделей

3. Основні напрямки застосування ШІ у філології

- Лінгвістичний аналіз
- Машинний переклад

4. Практичні інструменти та системи

- Системи аналізу тексту
- Системи машинного перекладу
- Інструменти стилOMETричного аналізу
- Практичні кейси застосування

5. Порівняльний аналіз традиційних та ШІ-інструментів

- Переваги використання ШІ
- Обмеження та недоліки

6. Трансформація філологічної практики

- Нові методології дослідження
- Зміна ролі філолога

7. Перспективи розвитку

- Найближчі перспективи
- Довгострокові перспективи
- Інноваційні напрямки досліджень
- Технологічні тренди

8. Етичні та професійні виклики

- Етичні аспекти
- Професійна адаптація
- Інституційні зміни
- Соціокультурний вплив

9. Практичні рекомендації

- Методологічні аспекти
- Професійний розвиток

10. Висновки

Перелік визначень

Комп'ютерна лінгвістика – галузь науки, що виникла у середині XX століття і займається машинною обробкою природної мови.

Корпусна лінгвістика – напрям мовознавства, що займається вивченням текстових корпусів.

Технології – сукупність методів, процесів, інструментів та матеріалів, що використовуються для досягнення конкретного результату у виробництві, науці чи іншій сфері діяльності. Це систематичне застосування наукових знань для практичних цілей, включаючи перетворення вхідних ресурсів у бажані вихідні продукти чи результати (штучний інтелект, нейронні мережі, deep learning, трансформери).

Deep learning – технологія глибинного навчання, що змінила підходи до обробки природної мови.

Трансформери – архітектура нейронних мереж, що дозволяє ефективно обробляти послідовності слів та враховувати контекст.

Нейронний машинний переклад (NMT) – система перекладу на основі нейронних мереж.

Напрямки застосування – морфологічний аналіз, синтаксичний розбір, семантичний аналіз, стилOMETричний аналіз.

Стилометричний аналіз – метод дослідження стилістичних особливостей тексту.

Токенізація – процес розбиття тексту на окремі елементи (токени).

Професійні аспекти – філологічна практика, компетенції, етичні виклики, академічна доброчесність.

Перспективи – мультимодальний аналіз, квантові обчислення, біологічно-інспіровані моделі.

Питання для перевірки та самоперевірки

1. Визначити основні етапи еволюції комп'ютерної лінгвістики від її зародження до сучасності.
2. Назвати ключові компоненти технологічної інфраструктури сучасних мовних моделей.
3. Виокремити основні переваги та недоліки використання ІІІ у філологічних дослідженнях.
4. Сформулювати основні етичні виклики, пов'язані з використанням ІІІ у філології.
5. Охарактеризувати зміни у ролі філолога в контексті впровадження ІІІ-технологій.
6. Проаналізувати основні напрямки застосування ІІІ у лінгвістичному аналізі.
7. Порівняти традиційні методи філологічного дослідження з методами, що використовують ІІІ.
8. Визначити ключові компетенції, необхідні сучасному філологу для роботи з ІІІ-інструментами.
9. Описати перспективи розвитку мультимодального аналізу в філологічних дослідженнях.
10. Систематизувати основні напрямки трансформації філологічної освіти під впливом ІІІ.

Цифрові інструменти Google. Перспективи застосування ІКТ у філології та освіті

<i>Google Classroom</i>	<i>Moodle</i>	<i>Canvas</i>
<i>Google Meet</i>	<i>Gmail</i>	<i>Google Scholar</i>
<i>Google Books</i>	<i>ResearchGate</i>	<i>Google Translate</i>
<i>Grammarly</i>	<i>Ngram Viewer.</i>	<i>Google Docs</i>
<i>Google Sheets</i>	<i>Google Slides</i>	

1. Вступ

Актуальність використання ІКТ у філології та вищій школі

2. Екосистема Google для освіти та науки (Google Workspace for Education)

Google Classroom – віртуальне навчальне середовище, що передбачає: створення онлайн-курсів; організація дистанційного навчання; керівництво навчальними матеріалами; оцінювання; забезпечення зворотнього зв'язку.

Google Meet проведення відеолекцій; організація семінарів; індивідуальні консультації; запис занять; інтеграція з іншими сервісами Google.

Google Scholar можливості пошуку наукових публікацій; створення бібліографічних посилань; відстеження цитувань; формування наукового профілю; моніторингу публікаційної активності.

Google Books доступ до оцифрованих видань; здійснення повнотекстового пошуку; створення власної електронної бібліотеки; експорту бібліографічних даних.

Спеціалізовані інструменти для філологів (лінгвістичні інструменти):

Google Translate – машинний переклад текстів; розпізнавання мови; переклад документів; інтеграція з іншими сервісами.

Google Ngram Viewer (Створення глосаріїв) – аналіз частотності вживання слів; діахронічні дослідження; порівняльний аналіз лексики; візуалізація мовних тенденцій.

Інструменти для створення та обробки контенту (текстові редактори та документи):

Google Docs – спільне редагування документів; відстеження змін; додавання коментарів; вбудована перевірка правопису; інтеграції з іншими сервісами.

Google Sheets – створення корпусів текстів; аналіз лінгвістичних даних; ведення статистики; створення діаграм.

Презентаційні інструменти:

Google Slides – створення інтерактивних презентацій; спільна робота над матеріалами; інтеграція мультимедіа; онлайн-демонстрація.

3. Перспективи розвитку ІКТ у філології

- **Штучний інтелект та машинне навчання**
 - Розвиток систем обробки природної мови
 - Вдосконалення машинного перекладу
 - Створення інтелектуальних асистентів
 - Автоматичний аналіз текстів
- **Корпусна лінгвістика та великі дані**
 - Створення масштабних мовних корпусів
 - Автоматична обробка текстів
 - Лінгвістичний аналіз соціальних медіа
 - Дослідження мовних варіацій

4. Методика впровадження ІКТ в освітній процес

- Інтеграція цифрових інструментів у навчальний процес
- Змішане та дистанційне навчання
- інтерактивні навчальні матеріали
- організація онлайн-взаємодії

- Оцінювання та моніторинг: (тести, аналітика, моніторинг результатів).

5. Grammarly як інструмент філологічного аналізу

• Характеристики Grammarly

- перевірка граматики та пунктуації
- стилістичний аналіз тексту
- перевірка чіткості викладу
- виявлення плагіату
- адаптація тону повідомлення

• Професійне застосування

- лінгвістичний аналіз текстів
- редагування наукових робіт
- створення навчальних матеріалів
- перевірка робіт;
- дослідження стилістичних особливостей.

• Перспективи розвитку

- розширення підтримки мов
- вдосконалення алгоритмів аналізу
- інтеграція з освітніми платформами
- розвиток функцій для наукових досліджень
- впровадження нових інструментів аналізу тексту

6. Висновки

Переваги застосування ІКТ та сервісів у філології й освіті. Сучасні виклики та необхідність постійного вдосконалення технологій. Значення цифрових компетентностей для фахівців-гуманітаріїв.

Термін	Визначення
ІКТ	Інформаційно-комунікаційні технології – сукупність методів, виробничих процесів і програмно-технічних засобів, інтегрованих з метою збирання, обробки, зберігання, розповсюдження, відображення і використання інформації
Google Workspace	Набір хмарних програм і інструментів для спільної роботи від компанії Google
Корпусна лінгвістика	Розділ мовознавства, що займається створенням, обробкою та дослідженням корпусів текстів
Машинне навчання	Підрозділ штучного інтелекту, що вивчає методи побудови алгоритмів, здатних навчатися на даних
LMS	Learning Management System (Система управління навчанням)
Цифрова компетентність	Здатність використовувати цифрові технології для вирішення професійних завдань

Питання для самостійного дослідження

1. Вплив використання ІКТ на розвиток сучасної філологічної науки.
2. Перспективи застосування штучного інтелекту в лінгвістичних дослідженнях.
3. Зміни методик викладання мови та літератури за допомогою цифрових інструментів.
4. Створення ефективного цифрового освітнього середовища для вивчення філологічних дисциплін.
5. Можливості надані корпусними дослідженнями для сучасної лінгвістики.

6. Забезпечення академічної доброчесності при використанні цифрових інструментів.
7. Переваги та недоліки використання машинного перекладу в перекладацькій практиці.
8. Ефективна організація дистанційного навчання.
9. Можливості застосування методів цифрової гуманітаристики у філологічних дослідженнях.
10. Оцінка ефективності використання ІКТ в освітньому процесі.

Питання для самоперевірки

1. Визначити основні компоненти Google Workspace for Education та їх призначення.
2. Назвати ключові інструменти Google для наукової роботи філолога.
3. Виокремити основні напрямки застосування Grammarly у філологічній практиці.
4. Сформулювати принципи організації ефективного дистанційного навчання з використанням ІКТ.
5. Охарактеризувати перспективи розвитку штучного інтелекту в філологічній галузі.

ПРАКТИЧНА ЧАСТИНА

Робота з British National Corpus (BNC)

Аналіз текстів, побудова словосполучень, частотні словники.



Мета заняття:

Сформувати у здобувачів практичні навички роботи з корпусом British National Corpus (BNC); навчити здійснювати пошук та аналіз лінгвістичних даних, будувати словосполучення, створювати частотні словники; розвинути здатність до критичного аналізу мовних явищ на основі корпусних даних; формувати вміння інтерпретувати результати корпусного аналізу та застосовувати їх у філологічних дослідженнях.

British National Corpus (BNC) – це збалансований корпус британського варіанту англійської мови об'ємом 100 мільйонів слів, який містить зразки письмової та розмовної мови. Корпус включає різноманітні жанри та стилі: художню літературу, газети, академічні тексти, повсякденні розмови тощо.

Ключові характеристики BNC:

- Репрезентативність: охоплює різні типи текстів та жанри
- Розмітка: морфологічна, синтаксична, метатекстова

- Можливість пошуку за різними параметрами: лексичними, граматичними, стилістичними
- Статистичні інструменти для аналізу частотності та колокацій

Функціональні можливості BNC

- **Базовий пошук**
 - пошук окремих слів та словоформ
 - використання регулярних виразів
 - фільтрація за частинами мови
- **Розширений пошук**
 - пошук колокацій та словосполучень
 - встановлення відстані між словами
 - пошук за граматичними категоріями
- **Статистичні інструменти**
 - побудова частотних словників
 - аналіз сполучуваності
 - визначення статистичної значущості колокацій
- **Робота з підкорпусами**
 - Створення власних підкорпусів
 - Фільтрація за жанрами та періодами
 - Порівняльний аналіз підкорпусів

Практичні завдання для самостійного виконання

1. Дослідження семантичного поля

- Обрати абстрактне поняття (наприклад, "happiness", "freedom", "love")
- Знайти всі можливі колокації з цим словом
- Побудувати семантичне поле на основі найчастотніших сполучень

- Проаналізувати отримані результати та зробіть висновки про особливості вживання слова

2. Порівняльний аналіз синонімів

- Вибрати пару синонімів (наприклад, "beautiful/pretty", "smart/clever")
- Дослідити частотність їх вживання в різних жанрах
- Проаналізувати сполучуваність
- Визначити контексти, де ці синоніми не є взаємозамінними

3. Діахронічне дослідження

- Обрати сучасне лексичну одиницю або вираз
- Дослідити частотність його вживання в текстах різних періодів
- Проаналізувати зміни в контекстах вживання
- Зробити висновки про еволюцію значення

4. Стилістичний аналіз

- Створити підкорпуси текстів різних стилів
- Порівняти частотні словники цих підкорпусів
- Визначити характерні колокації для кожного стилю
- Описати стилістичні особливості вживання певних слів

5. Граматичне дослідження

- Обрати граматичну конструкцію
- Знайти всі варіанти її реалізації в корпусі
- Проаналізувати частотність різних варіантів
- Визначити фактори впливу на вибір варіанту

Приклад формулювання результатів дослідження:

1. Вступ:

- Представити обрану пару слів, пояснити причину вибору.

- Сформулювати мету дослідження та гіпотези щодо можливих відмінностей у вживанні слів.

2. Методологія:

- Описати процес роботи з BNC та використані функції (наприклад, ‘Collocates’, ‘Compare’).
- Вказати налаштування пошуку (позиція колокацій, частина мови тощо).

3. Результати:

- Представити отримані дані щодо частотності кожного лексичної одиниці.
- Навести графіки або таблиці, що відображають використання слів у різних жанрах.
- Перелічити основні колокації для кожного слова та порівняти їх.
- Навести приклади контекстів з корпусу, що ілюструють знахідки.
- Проаналізувати отримані результати та співвіднести їх з початковими гіпотезами.
- Визначити можливі причини виявлених відмінностей у вживанні слів.
- Розглянути вплив контексту та жанру на вибір лексичної одиниці.

5. Висновки:

- Зробити висновки дослідження.
- Підтвердити або спростувати гіпотези.
- Запропонувати ідеї для подальших досліджень або практичного застосування знахідок.

Онлайн-ресурси та бази даних

1. British National Corpus (BNC) Офіційний веб-сайт

- <https://www.english-corpora.org/bnc/>

2. Sketch Engine

- <https://www.sketchengine.eu/>

Аналіз американського національного корпусу СОСА

Пошук лексичних та граматичних патернів, інтерпретація даних для досліджень

The great general balance (and large size) of COCA also allows you to see the frequency of related phrases across genres. For example, the following table shows "soft NOUN".

	CONTEXT	ALL	BLOG	WEB-GENL	TV/MOVIES	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	SOFT TISSUE	1120	62	67	74	39	36	100	35	707
2	SOFT DRINKS	1109	123	123	42	90	83	304	296	48
3	SOFT MONEY	790	21	34	12	446	8	78	153	38
4	SOFT SPOT	867	133	110	166	63	159	135	86	15
5	SOFT DRINK	721	43	60	48	77	68	191	199	35
6	SOFT VOICE	546	12	39	10	11	351	53	52	18
7	SOFT POWER	421	49	53	1	64	1	51	32	170
8	SOFT TISSUES	328	13	23	5	6	10	41	6	224
9	SOFT LANDING	274	42	19	18	38	17	56	63	21
10	SOFT LIGHT	247	22	25	9	1	121	38	21	10
11	SOFT PEAKS	216	2	4		1	2	144	62	1
12	SOFT SKIN	207	9	24	28	6	114	19	4	3
13	SOFT TOUCH	213	22	26	28	15	43	51	25	3
14	SOFT SKILLS	219	41	41		7		13	13	104

You can also compare two genres (or sets of genres). For example, the following table shows collocates (nearby words) of *chair* in ACADEMIC (left) and FICTION (right).

EC 1 (ACADEMIC): 119,790,456 WORDS							SEC 2 (FICTION): 118,322,084 WORDS						
	WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2	RATIO		WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1	DIVISION	56	1	0.5	0.0	55.3	1	KITCHEN	293	3	2.5	0.0	98.9
2	ENGINEERING	53	1	0.4	0.0	52.4	2	WING	111	0	0.9	0.0	93.8
3	ABA	55	0	0.5	0.0	45.9	3	WAY	87	1	0.7	0.0	88.1
4	PROGRAM	46	1	0.4	0.0	45.4	4	FEET	220	3	1.9	0.0	74.2
5	EDUCATION	42	0	0.4	0.0	35.1	5	FOOT	72	1	0.6	0.0	72.9
6	RESEARCH	34	1	0.3	0.0	33.6	6	MOTHER	134	2	1.1	0.0	67.8
7	FACULTY	38	0	0.3	0.0	31.7	7	COMMAND	65	1	0.5	0.0	65.8
8	ASME	36	0	0.3	0.0	30.1	8	LEATHER	409	7	3.5	0.1	59.2
9	COMMISSION	30	1	0.3	0.0	29.6	9	CANVAS	58	1	0.5	0.0	58.7
10	ROLE	33	0	0.3	0.0	27.5	10	BEACH	67	0	0.6	0.0	56.6
11	SUBCOMMITTEE	24	1	0.2	0.0	23.7	11	FINGERS	53	1	0.4	0.0	53.7

Мета заняття:

Сформувати у здобувачів практичні навички роботи з американським національним корпусом СОСА; навчити здійснювати пошук лексичних та граматичних патернів, інтерпретувати отримані дані з урахуванням контексту; розвинути здатність до критичного аналізу мовних явищ за допомогою цифрових корпусних інструментів; сформувати вміння застосовувати результати корпусного аналізу у філологічних дослідженнях.

1. Концептуальні основи корпусної лінгвістики

Корпусна лінгвістика – це міждисциплінарний напрямок досліджень, що передбачає системне вивчення мови через Large Language Data Sets (великі масиви мовних даних).

COCA (Corpus of Contemporary American English) виступає унікальним інструментом, який надає дослідникам широкий спектр можливостей для аналізу мовних явищ.

2. Напрямки філологічних досліджень з використанням СОСА

• Лексикологічні дослідження

Діахронічна еволюція лексики – простеження змін значень слів; аналіз появи нових лексичних одиниць; дослідження семантичних трансформацій.

Приклад: Вивчення еволюції слова "cool" від температурної характеристики до молодіжного сленгу

Неологізми та запозичення – ідентифікація нових слів; аналіз джерел запозичень; швидкість входження нової лексики.

Приклад: Дослідження інтернет-термінів, запозичень з інших мов

• Соціолінгвістичні студії

Гендерні лінгвістичні особливості – відмінності у вживанні лексики; комунікативні стратегії.

Приклад: Аналіз відмінностей у мові чоловіків і жінок у різних жанрах

Вікові мовні характеристики – мовлення різних поколінь; молодіжний сленг.

Приклад: Дослідження динаміки молодіжної лексики

• Прагмалінгвістичні дослідження

Комунікативні стратегії – аналіз мовленнєвих актів; дослідження непрямої комунікації.

Приклад: Вивчення форм ввічливості, непрямих прохань

Міжкультурна комунікація – порівняння комунікативних моделей;
дослідження культурно маркованої лексики

Приклад: Аналіз концептуальних метафор

- **Граматичні дослідження**

Еволюція граматичних конструкцій – зміни в дієслівних формах;
трансформація синтаксичних моделей

Приклад: Вивчення варіативності пасивних конструкцій

Синтаксичні особливості різних жанрів – порівняння синтаксису в
академічних, розмовних, художніх текстах

Приклад: Аналіз складності синтаксичних структур

- **Стилістичні та дискурсивні дослідження**

Жанрова варіативність мови – лексичні та граматичні особливості
жанрів; стильові маркери.

Приклад: Порівняння мови наукових статей та блогів

Дискурс-аналіз – вивчення текстових стратегій; Риторичні прийоми

Приклад: Аналіз аргументативних стратегій у медійних текстах

3. Методологічні переваги використання СОСА

- Репрезентативність даних
- Можливість кількісного та якісного аналізу
- Діахронічна перспектива
- Мультижанровий підхід
- Миттєвий доступ до величезного масиву даних

4. Інструменти та функціонал СОСА

- **Основний інтерфейс пошуку**
 - Пошук за словом/фразою
 - Використання регулярних виразів
 - Налаштування параметрів виведення результатів

- Фільтрація за жанрами та періодами
- **Розширені можливості пошуку**
 - Пошук за частинами мови (POS-теги)
 - Пошук за леммами
 - Пошук за семантичними категоріями
 - Створення користувацьких списків
- **Інструменти аналізу**
 - Compare (порівняння частотності)
 - Collocates (аналіз колокацій)
 - Charts (візуалізація даних)
 - Lists (створення частотних списків)

Практичні завдання для самостійного виконання

Завдання 1:

Частотний аналіз: Дослідити частотність вживання синонімічних прикметників "beautiful", "gorgeous", "stunning" у різних жанрах корпусу. Створити порівняльну таблицю та проаналізувати отримані дані. Звернути увагу на:

- Загальну частотність
- Розподіл за жанрами
- Динаміку вживання за роками

Завдання 2:

Колокаційний аналіз: Провести аналіз колокацій для слова "climate" за останні 5 років. Визначити:

- Найчастотніші прикметники-модифікатори
- Дієслова, що вживаються з цим іменником
- Типові іменникові словосполучення Порівняти результати з більш раннім періодом (1990-1995)

Завдання 3:

Граматичні патерни: Дослідити конструкцію "keep + V-ing".

Використовуючи POS-теги

- Визначити найчастотніші дієслова в цій конструкції
- Проаналізувати типові підмети
- Описати семантичні групи дієслів
- Порівняти вживання в різних жанрах

Завдання 4:

Семантичний аналіз: Дослідити метафоричне вживання слова "heart"

в академічних текстах

- Знайти типові дієслова
- Визначити прикметники-модифікатори
- Проаналізувати стійкі словосполучення
- Порівняти з вживанням у художній літературі

Завдання 5:

Комплексне дослідження: Оберати будь-який модальний іменник (possibility, necessity, likelihood) та провести комплексний аналіз:

- Частотність у різних жанрах
- Типові граматичні конструкції
- Колокації та стійкі словосполучення
- Семантичні особливості вживання
- Діахронічні зміни

Методичні рекомендації щодо виконання завдань

1. Перед початком роботи уважно ознайомтеся з інтерфейсом та можливостями пошуку
2. Використовуйте розширені налаштування для уточнення результатів

3. Зберігайте проміжні результати пошуку
4. Звертайте увагу на статистичну значущість даних
5. Документуйте всі етапи дослідження
6. Формулюйте чіткі висновки на основі отриманих даних

Критерії оцінювання

- Правильність використання інструментів COCA – 20%
- Повнота виконання завдання – 30%
- Глибина аналізу отриманих даних – 30%
- Обґрунтованість висновків – 20%

Електронні ресурси та бази даних

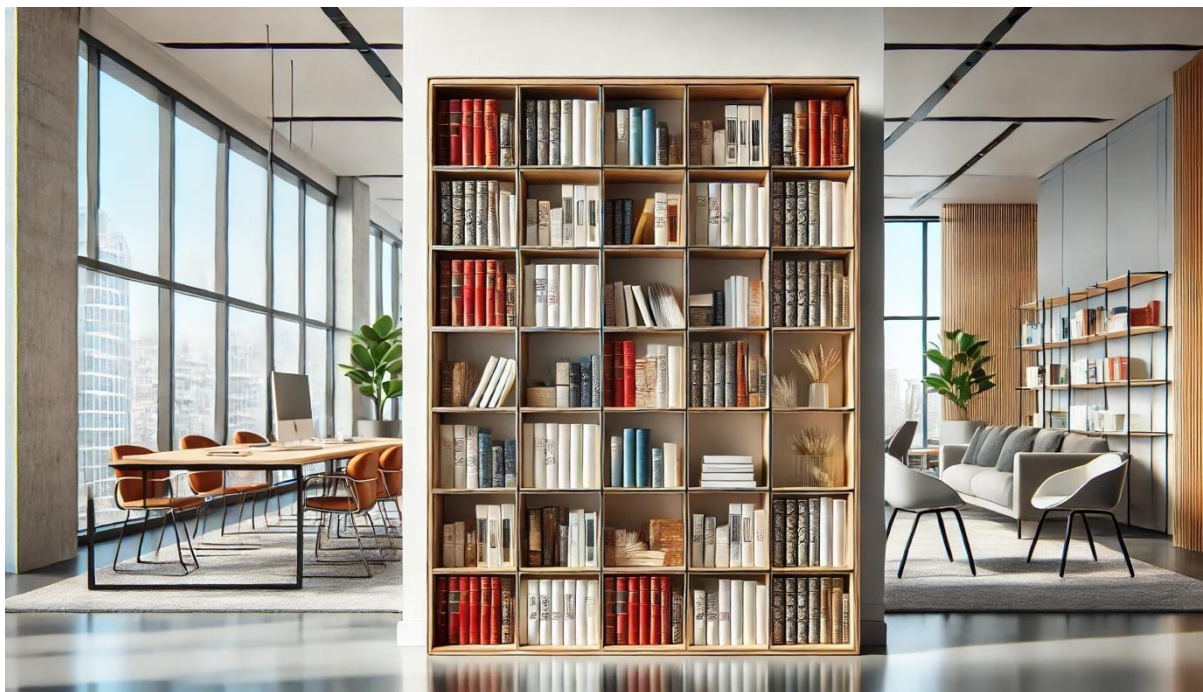
1. Brigham Young University Corpora

- Офіційний сайт: <https://www.englishcorpora.org/>
- Унікальний набір лінгвістичних корпусів
- Постійне оновлення баз даних

2. The Corpus of Contemporary American English (COCA)

- Офіційний сайт: <https://www.englishcorpora.org/coca/>
- Найбільший розмічений корпус американської англійської мови

Принципи роботи з українським національним корпусом. Корпус ГРАК



Мета заняття:

Сформувати у здобувачів уміння ефективно працювати з Українським національним корпусом, зокрема корпусом ГРАК; розвинути навички формулювання запитів для лінгвістичного пошуку, здійснення морфологічного та контекстуального аналізу, зіставлення мовних варіантів і фільтрації результатів за заданими параметрами; навчити критично осмислювати корпусні дані, інтерпретувати отримані результати в контексті філологічних досліджень, а також застосовувати корпусні інструменти для вирішення прикладних завдань у лексикографії, перекладі, стилістиці й мовній нормі.

Українська корпусна лінгвістика є важливою складовою сучасних філологічних досліджень. Національний корпус мови — це репрезентативний електронний масив текстів різних жанрів, стилів та

періодів, що дозволяє проводити комплексний лінгвістичний аналіз. Корпус Грак (HRYAK) – унікальний проєкт, який являє собою повнотекстову базу українських текстів, призначену для глибокого наукового вивчення мови.

1. Структурні особливості:

- Обсяг корпусу: понад 100 млн слововживань
- Хронологічний діапазон: від кінця XVIII до початку XXI століття
- Жанрова диверсифікація: художня література, наукові праці, публіцистика, офіційно-ділові документи
- Типи текстів: друковані та рукописні джерела
- Формати представлення: повнотекстові документи, розмічені за лінгвістичними параметрами

2. Основні характеристики Корпусу Грак

- Електронний архів україномовних текстів
- Методологічно вивірена система класифікації та анотації текстів
- Можливість багаторівневого пошуку та аналізу мовних одиниць
- Репрезентативність текстових масивів різних історичних періодів

3. Інструменти та функціонал

• Основні інструменти корпусу:

- Повнотекстовий пошук
- Контекстний аналіз
- Статистичні розрахунки
- Лексико-граматичний аналіз
- Діахронічне дослідження мовних явищ

• Функціональні можливості:

- Повнотекстовий контекстуальний пошук
- Частотний аналіз лексичних одиниць

- Вивчення дериваційних моделей
- Реконструкція мовної еволюції
- Аналіз синтаксичних трансформацій
- Дослідження термінологічних систем
- Порівняльний аналіз текстів різних жанрів та стилів

Практичні завдання для самостійного виконання

Завдання 1: Лексико-статистичне дослідження

Обрати художній твір українського письменника ХХ століття. За допомогою Корпусу Грак провести аналіз частотності використаних лексем, укласти словник найуживаніших слів та визначити лексичне навантаження тексту.

Завдання 2: Діахронічне вивчення мовних змін

Простежте еволюцію вживання певного граматичного явища (наприклад, форми давального відмінка) впродовж трьох історичних періодів за матеріалами Корпусу Грак. Підготуйте аналітичну доповідь про виявлені трансформації.

Завдання 3: Жанрово-стильове порівняння

Здійснити контрастивний аналіз використання синтаксичних конструкцій у текстах наукового та художнього стилів. Визначити специфічні граматичні особливості кожного стилю.

Завдання 4: Авторська лексикографія

Змодельовати мовну палітру обраного українського письменника, використовуючи інструменти корпусного аналізу. Укласти тематичний словник авторських лексем.

Завдання 5: Дослідження мовної динаміки

Проаналізувати зміни в українській науковій термінології впродовж останніх п'ятдесяти років. Простежити входження нових термінів, трансформацію значень існуючих лексичних одиниць.

- Офіційний сайт Українського мовно-інформаційного фонду НАН України
- Електронний архів наукових публікацій «Наукова електронна бібліотека»
- Міжнародний репозитарій лінгвістичних корпусів

Створення власного корпусу для досліджень Налаштування та аналіз спеціалізованих корпусів для філологічних досліджень

1. Платформи для створення лінгвістичних корпусів

- **Sketch Engine** – професійна платформа для створення та аналізу лінгвістичних корпусів з потужним інструментарієм
- **AntConc** – безкоштовний інструмент для корпусного аналізу з простим інтерфейсом
- **UIMA (Unstructured Information Management Architecture)** – відкрита платформа для розвиненого текстового аналізу

2. Основні характеристики корпусних платформ:

- Підтримка різних мовних форматів
- Інструменти морфологічного та синтаксичного аналізу
- Функції статистичної обробки текстів
- Можливості текстової розмітки
- Експорт результатів у різні формати

3. Інструменти та функціонал

- **Ключові інструменти корпусного аналізу:**
 - Токенізація – розділення тексту на окремі одиниці
 - Лематизація – приведення слів до базової форми
 - Частотний аналіз – підрахунок вживаності мовних одиниць
 - Конкорданс – контекстний пошук слів/виразів
 - Колокаційний аналіз – вивчення сполучуваності лексем
- **Функціональні можливості:**
 - Імпорт текстових документів
 - Автоматична розмітка частин мови
 - Побудова статистичних графіків
 - Порівняння підкорпусів
 - Генерація конкордансів та частотних словників

4. Інструкція зі створення корпусу

- **Алгоритм підготовки корпусу:**
 - Формування цільової колекції текстів
 - Уніфікація текстового формату
 - Попередня обробка документів
 - Додавання метаданих
 - Здійснення лінгвістичної розмітки
 - Налаштування параметрів пошуку та аналізу
- **Методичні рекомендації**
 - Приділяти увагу якості та репрезентативності текстової колекції
 - Дотримуватися авторських прав під час формування корпусу
 - Використовувати актуальне програмне забезпечення
 - Документувати всі етапи роботи
 - Критично оцінювати проміжні результати

Практичні завдання для самостійного виконання

1. Проектування корпусу

- Обрати тематичну спрямованість корпусу
- Розробити критерії добору текстових матеріалів
- Підготувати колекцію з 50-100 документів

2. Технічна підготовка

- Здійснити уніфікацію текстових форматів
- Виконати первинне очищення текстів
- Розробити систему метаданих для корпусу

3. Лінгвістична розмітка

- Провести морфологічну розмітку текстів
- Створити частотний словник корпусу
- Побудувати графік розподілу частиномовних одиниць

4. Аналітичне дослідження

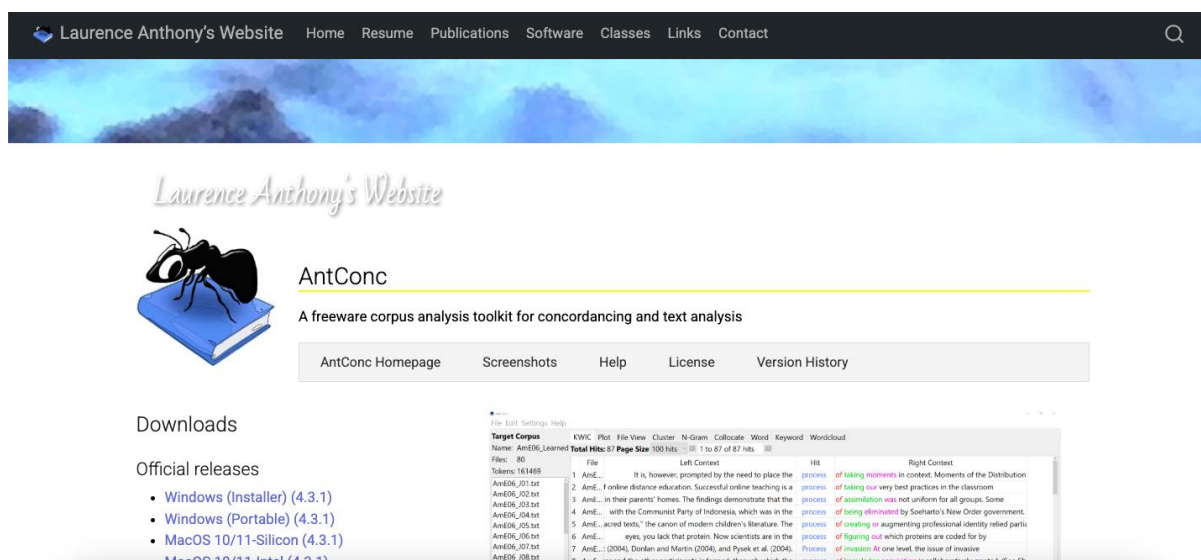
- Виконати конкордансний аналіз обраних лексем
- Підготувати статистичний звіт про особливості корпусу
- Сформулювати лінгвістичні гіпотези на базі отриманих даних

5. Презентація результатів

- Підготувати презентацію про методологію створення корпусу
- Представити основні лінгвістичні характеристики
- Обґрунтувати наукову цінність розробленого корпусу

Аналітика текстів за допомогою AntConc

Налаштування аналізу, робота з кластерами слів, створення частотних словників



Мета заняття:

Сформувати у здобувачів практичні навички роботи з програмою AntConc для аналізу текстів; навчити налаштовувати параметри лінгвістичного аналізу, працювати з кластерами слів та створювати частотні словники; розвинути здатність до критичного осмислення текстових структур на основі кількісних даних; формувати вміння інтерпретувати результати комп'ютерного аналізу текстів та застосовувати їх у філологічних дослідженнях.

1. Значення комп'ютерної лінгвістики для філологічних досліджень

- Розкриття структурних особливостей текстів.
- Виявлення стилістичних та жанрових особливостей.
- Підтримка емпіричних досліджень та об'єктивне підтвердження гіпотез.

2. Основні поняття

- **Корпус текстів:** збірка текстових даних, яка використовується для лінгвістичного аналізу.
- **Частотний аналіз:** метод вивчення повторюваності лексичних елементів у текстах.
- **Кластери слів (word clusters):** групування слів або виразів, які з'являються разом у тексті, що дозволяє виявляти типові схеми використання мови.

3. Опис функціоналу AntConc

AntConc –безкоштовна програма для аналізу корпусів текстів, розроблена з метою допомогти дослідникам проводити комплексний аналіз текстових даних.

- **Пошук за ключовими словами (Concordance)**
Забезпечує пошук контексту вживання певного слова чи виразу в текстовому корпусі.
- **Створення частотних словників (Word List/Keyword List)**
Дозволяє отримати список слів із зазначенням їхньої частоти. Це сприяє виявленню домінуючих лексичних одиниць та побудові статистичних моделей.
- **Аналіз кластера слів (Clusters/N-Grams)**
Модуль для пошуку й аналізу найбільш типових комбінацій слів. Використовується для виявлення стійких словосполучень та аналізу контексту їх вживання.
- **Налаштування параметрів аналізу**
Можливість задавати параметри, такі як мінімальна та максимальна довжина кластерів, обмеження по частоті, використання регулярних виразів для пошуку тощо.

- **Графічний інтерфейс** Інтуїтивно зрозумілий інтерфейс з можливістю візуалізації результатів аналізу, що сприяє ефективній роботі з даними.

4. Практичні аспекти роботи

- **Завантаження та підготовка текстового корпусу:**
Формати файлів, попередня обробка даних (очищення тексту, токенизація).
- **Вибір та налаштування інструментів аналізу:**
Налаштування параметрів пошуку, визначення діапазону частот, адаптація інтерфейсу під потреби дослідження.
- **Інтерпретація отриманих даних:**
Аналіз отриманих частотних словників та кластерів, виявлення закономірностей та лінгвістичних аномалій.

Завдання для самостійного виконання (з інструкцією)

Завдання 1. Завантаження та підготовка корпусу

Інструкції:

1. Обрати декілька текстових документів англійською мовою (можна використати художню літературу або наукові статті).
2. Зберегти документи у форматі TXT.
3. Використовуючи AntConc, завантажити створений корпус.
4. Оцінити якість даних, перевірити наявність небажаних символів та здійснити базову очистку текстів.

Завдання 2. Створення частотного словника

Інструкції:

1. Використовуючи модуль «Word List», згенерувати частотний словник для завантаженого корпусу.

2. Визначити найбільш частотні слова та вирішити, чи потребують вони додаткової попередньої обробки (наприклад, виключення стоп-слів).
3. Сформулювати висновки щодо загальної лексичної структури текстів.

Завдання 3. Аналіз ключових слів за допомогою Concordance

Інструкції:

1. Обрати 3–5 ключових слів, які, на вашу думку, є характерними для досліджуваного корпусу.
2. За допомогою функції Concordance знайти приклади вживання обраних слів у тексті.
3. Проаналізувати контексти, звернувши увагу на сусідні слова та фрази, що можуть свідчити про типові мовні конструкції.

Завдання 4. Побудова та аналіз кластерів слів (N-Grams)

Інструкції:

1. Використати модуль «Clusters/N-Grams» для генерації кластерів з 2–3 слів.
2. Задати різні параметри (наприклад, мінімальну частоту появи) та порівняти отримані результати.
3. Вибрати найбільш характерні кластери та обґрунтувати лінгвістичне значення.

Завдання 5. Порівняльний аналіз текстів за допомогою Keyword List

Інструкції:

1. Підготувати два окремих текстових корпуси (наприклад, твори двох різних авторів або тексти різних жанрів).

2. За допомогою функції Keyword List побудувати частотні словники для обох корпусів.
3. Провести порівняльний аналіз, визначивши ключові відмінності у вживанні лексичних елементів.
4. Сформулювати висновки щодо стилістичних та жанрових особливостей текстів.

Візуалізація текстових даних за допомогою VOYANT TOOLS та інших інструментів



Мета заняття: Сформувати у здобувачів практичні навички роботи з інструментами візуалізації текстових даних, розвинути здатність до критичного аналізу текстів за допомогою цифрових інструментів, навчити інтерпретувати візуальні представлення текстових даних та застосовувати отримані результати у філологічних дослідженнях.

Ключові слова

<i>текстова аналітика,</i>	<i>візуалізація даних,</i>	<i>частотний аналіз,</i>
<i>конкордансний аналіз,</i>	<i>колокації,</i>	<i>n-грами, хмара слів</i>

Voyant Tools – це веб-платформа з відкритим кодом для аналізу та візуалізації текстових даних, яка надає широкий спектр інструментів для дослідження текстів. Платформа дозволяє завантажувати тексти різними способами (пряме введення, завантаження файлів, URL-посилання) та підтримує різні формати файлів (txt, doc, docx, pdf, xml). Програма автоматично генерує набір візуалізацій та аналітичних інструментів, включаючи частотні списки, графіки розподілу слів, візуалізації зв'язків між словами, хмари слів та інші види представлення даних.

Визначення

Текстова аналітика – комплекс методів та підходів до аналізу неструктурованих текстових даних з метою видобування значущої інформації та виявлення прихованих закономірностей.

Дистантне читання – метод аналізу текстів, запропонований Франко Моретті, який передбачає вивчення великих текстових корпусів за допомогою комп'ютерних інструментів замість традиційного уважного читання окремих текстів.

Хмара слів (Cirrus) – візуальне представлення частотності слів у тексті, де розмір слова пропорційний частоті його вживання в корпусі. Один з основних інструментів візуалізації у Voyant Tools.

Частотний розподіл (Trends) – графічне відображення частоти появи слів у різних частинах тексту, що дозволяє простежити зміни у використанні термінів протягом тексту.

Ключові слова в контексті (KWIC) – формат відображення результатів пошуку, який показує кожне входження шуканого слова разом із певною кількістю слів до і після нього, що дозволяє вивчати безпосередній контекст вживання.

Контекстне вікно – кількість слів до і після ключового слова, які враховуються при аналізі колокацій та побудові конкордансу в Voyant Tools.

Стоп-слова – список слів, які виключаються з аналізу (зазвичай службові частини мови та інші високочастотні слова), що дозволяє зосередитись на змістовно важливій лексиці.

Практичні завдання

1. Порівняльний аналіз текстів

- Завантажити два тексти різних авторів одного періоду до Voyant Tools
- Створити порівняльну характеристику частотності вживання ключових слів
- Проаналізувати відмінності у словнику авторів
- Підготувати візуалізацію результатів аналізу

2. Дослідження динаміки тексту

- Обрати художній текст великого обсягу
- Створити графік розподілу ключових слів по тексту
- Визначити основні тематичні кластери
- Проаналізувати зміни в використанні лексики протягом розвитку сюжету

3. Аналіз колокацій

- Обрати 3-5 ключових слів з тексту
- Побудувати мережу колокацій для кожного слова
- Визначити найбільш значущі словосполучення
- Проаналізувати семантичні зв'язки між словами

4. СтилOMETричний аналіз

- Завантажити корпус текстів одного автора
- Створити частотний список службових слів

- Побудувати візуалізацію стилOMETричних характеристик
- Визначити індивідуальні особливості авторського стилю

5. Тематичне моделювання

- Підготувати корпус текстів однієї тематики
- Визначити основні тематичні групи
- Створити візуалізацію тематичної структури
- Проаналізувати взаємозв'язки між темами

Інструменти ШІ для аналізу текстів і редагування

Можливості використання Grammarly

<i>Автоматизоване редагування</i>		<i>Граматична перевірка</i>
<i>Стилістичний аналіз</i>	<i>Семантична корекція</i>	<i>Плагіат-контроль</i>
<i>Контекстуальні пропозиції</i>		<i>Тональність тексту</i>

Мета заняття:

- Сформувати практичні навички роботи з програмою Grammarly для аналізу та редагування текстів
- Навчити студентів використовувати розширені функції програми для покращення якості академічних текстів
- Розвинути компетенції з цифрової грамотності та автоматизованого редагування

Grammarly – це інтелектуальний програмний засіб для перевірки граматики, пунктуації, стилю та чіткості тексту. Програма використовує алгоритми штучного інтелекту для аналізу тексту та надання рекомендацій щодо його покращення. Застосунок доступний як веб-сервіс, розширення для браузера та десктопна програма. Grammarly підтримує різні стилі

письма та може адаптувати свої рекомендації відповідно до мети тексту (академічний, діловий, неформальний стиль).

Визначення

Лінгвістичний алгоритм – система правил та інструкцій для автоматизованого аналізу мовних конструкцій та їх корекції.

Семантична когерентність – змістова цілісність тексту, логічний зв'язок між його частинами.

Стилістична консистентність – послідовність у використанні мовних засобів відповідно до обраного стилю.

Контекстуальна релевантність – відповідність мовних одиниць контексту їх використання.

Лексична варіативність – різноманітність використання словникового запасу в тексті.

Практичні завдання

1. Завантажити академічний текст обсягом 1000 слів до Grammarly та проведіть повний аналіз. Скласти звіт про виявлені помилки за категоріями (граматичні, пунктуаційні, стилістичні). Обґрунтувати доцільність запропонованих виправлень.
2. Порівняти два варіанти академічного тексту (до і після корекції Grammarly). Проаналізувати різницю в показниках читабельності, залучення та чіткості. Скласти таблицю змін та їх впливу на якість тексту.
3. Налаштувати особисті цілі та вподобання в Grammarly відповідно до вимог академічного стилю. Створити власний словник термінів та виключень. Обґрунтувати свій вибір налаштувань.

4. Використовуючи функцію синонімів та перефразування Grammarly, покращити лексичну варіативність наукового тексту. Скласти перелік оригінальних та заміненних термінів з поясненням доцільності заміни.
5. Провести аналіз тональності та емоційного забарвлення академічного тексту за допомогою Grammarly. Визначити відповідність тону академічному стилю та запропонувати необхідні корективи.

Основи роботи з програмами для перекладацької практики Cat Tools

Мета заняття:

Сформувати у здобувачів базові практичні навички роботи з комп'ютеризованими системами підтримки перекладу (CAT tools), зокрема з програмами OmegaT та MemoQ; навчити створювати та налаштовувати перекладацькі проєкти, працювати з пам'яттю перекладів і глосаріями; розвинути вміння використовувати цифрові інструменти для підвищення якості, послідовності та ефективності перекладу; сформувати здатність інтегрувати отримані навички у професійну перекладацьку діяльність.

CAT Tools (Computer-Assisted Translation Tools) – це спеціалізовані програмні інструменти для професійних перекладачів, які автоматизують та оптимізують процес перекладу. Вони працюють за принципом сегментації тексту та створення бази перекладацької пам'яті, що дозволяє повторно використовувати раніше перекладені фрагменти, забезпечувати послідовність термінології та значно прискорювати роботу перекладача.

Що таке OmegaT:

- Відкрите програмне забезпечення для перекладачів.

- Основні принципи роботи: використання пам'яті перекладів, підтримка численних форматів файлів, можливості інтеграції з глосаріями та машинним перекладом.

Історія та спільнота:

- Проект з відкритим кодом, активна спільнота користувачів та розробників.
- Постійне вдосконалення завдяки зворотному зв'язку та підтримці користувачів.

Переваги використання OmegaT:

- Безкоштовний доступ.
- Гнучкість налаштувань, підтримка численних мов та форматів файлів.
- Простий інтерфейс для початківців, проте з можливістю розширення функціоналу за допомогою плагінів.

Опис функціоналу OmegaT

1. Основні можливості та модулі:

- **Пам'ять перекладів (Translation Memory, TM):**
 - Автоматичне збереження раніше перекладених сегментів.
 - Функція пошуку точних та частково збіжних фраз (fuzzy matching).
- **Глосарії:**
 - Зберігання термінів та їх перекладів.
 - Автоматична підказка термінології під час роботи.
- **Сегментація тексту:**
 - Розбиття документа на логічні сегменти для полегшення перекладу.
 - Налаштування правил сегментації відповідно до специфіки мови чи документа.

- **Підтримка різних форматів файлів:**
 - Робота з форматами, такими як .docx, .odt, .html, .xml, та інші.
 - Збереження форматування та структури документа.
- **Інтеграція з іншими інструментами:**
 - Підключення до машинного перекладу (наприклад, Google Translate API) для попередніх варіантів перекладу.
 - Плагіни для додаткових функцій (перевірка правопису, стилістична корекція).
- **Управління проектами:**
 - Створення, збереження та організація перекладацьких проектів.
 - Відстеження прогресу роботи над документом.

2. Практичні аспекти використання:

- Демонстрація інтерфейсу OmegaT.
- Огляд типового робочого процесу: створення проекту, імпорт файлів, переклад сегментів, збереження та експорт готового документу.
- Налаштування глосаріїв та пам'яті перекладів.

Завдання для самостійного виконання (з інструкцією)

Завдання 1: Створення нового проекту, імпорт документу та налаштування базових параметрів мови.

Інструкція:

1. Запустити OmegaT.
2. Створити новий перекладацький проект, вказавши назву, джерельну та цільову мови.
3. Додати тестовий документ (наприклад, короткий текст у форматі .txt або .docx) до проекту.

Завдання 2: Використання пам'яті перекладів (TM) та fuzzy matching

Інструкція:

1. Виконати переклад кількох сегментів тексту.
2. Зберегти переклади, щоб вони з'явилися в пам'яті.
3. Створити новий документ із схожим текстом, де частина сегментів збігається з попереднім перекладом.
4. Проаналізувати, як OmegaT пропонує варіанти перекладу через fuzzy matching.

Завдання 3: Налаштування та використання глосарію

Інструкція:

1. Створити файл глосарію у форматі, підтримуваному OmegaT (наприклад, .csv або .txt згідно з документацією).
2. Додати до глосарію кілька термінів з їхніми перекладами.
3. Підключити глосарій до проекту.
4. Переконайтесь, що під час перекладу термінологічні підказки з'являються автоматично.

Завдання 4: Налаштування сегментації та редагування правил

Інструкція:

1. Дослідити правила сегментації, які застосовуються у OmegaT (знайти відповідний файл конфігурації).
2. Внести зміни до правил сегментації для певного типу тексту (наприклад, для текстів з великою кількістю діалогів).
3. Застосувати нові налаштування до тестового документу та проаналізувати результат.

Завдання 5: Експорт готового перекладу та інтеграція з машинним перекладом

Інструкція:

1. Завершити переклад тестового документу.
2. Зберегти та експортувати готовий переклад у відповідному форматі (наприклад, .docx або .html).
3. Налаштувати інтеграцію з машинним перекладом (за наявності доступу до API, наприклад, Google Translate) для попереднього варіанту перекладу.
4. Провести тестове використання машинного перекладу для декількох сегментів та порівняти результати з ручним перекладом.

Цифрові інструменти Google Workspace для освіти. Інтеграція з сервісами

Мета заняття : Сформувати у здобувачів практичні навички роботи з інструментами Google Workspace для освіти, розвинути вміння інтегрувати різні сервіси для ефективної організації освітнього процесу та наукової діяльності.

Ключові слова:

<i>Google Workspace</i>	<i>хмарні технології</i>	<i>освітні інструменти</i>
<i>цифрова колаборація</i>		<i>синхронізація даних</i>
<i>спільний доступ</i>	<i>навчальне середовище</i>	
<i>автоматизація освітнього процесу</i>		<i>академічна комунікація</i>

Google Workspace для освіти – це комплексне рішення для організації освітнього процесу, що включає набір інтегрованих онлайн-інструментів для продуктивної роботи та навчання. Платформа забезпечує можливість

створення єдиного цифрового простору для викладачів та студентів, підтримує різні формати навчання та сприяє ефективній комунікації між учасниками освітнього процесу.

Визначення

Google Workspace для освіти – інтегрована система хмарних сервісів, спрямована на забезпечення освітньої діяльності через цифрові інструменти співпраці та комунікації.

Інтеграція сервісів – процес об'єднання різних цифрових інструментів та додатків у єдину функціональну систему для досягнення освітніх цілей.

Цифрове освітнє середовище – сукупність інформаційних систем та технологічних засобів, що забезпечують організацію освітнього процесу в електронному форматі.

Академічна колаборація – спільна діяльність учасників освітнього процесу з використанням цифрових інструментів для досягнення навчальних та наукових цілей.

Синхронізація освітніх ресурсів – процес забезпечення актуальності та узгодженості навчальних матеріалів across різних цифрових платформах.

Практичні завдання

Завдання 1. Організація навчального простору Створити структуровану систему папок у Google Drive для зберігання навчальних матеріалів з філологічних дисциплін. Налаштувати різні рівні доступу для різних категорій користувачів (викладачі, студенти, адміністратори).

Продемонструвати можливості автоматичної синхронізації даних між різними сервісами.

Завдання 2. Розробка інтерактивного навчального модуля За допомогою Google Sites створити освітній веб-ресурс, інтегрувавши в нього різні типи контенту: презентації з Google Slides, опитування через Google Forms, навчальні відео з YouTube та інтерактивні завдання. Забезпечити можливість відстеження прогресу студентів.

Завдання 3. Організація групової роботи Розробити систему спільної роботи над науковим проєктом з використанням Google Docs та Google Sheets. Налаштувати автоматичне сповіщення про зміни, створити систему коментування та відстеження версій документів. Інтегрувати календар для планування етапів роботи.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Ваховська О. В. Основи комп'ютерної лінгвістики: навчально-методичний посібник. Київ: Видавничий центр КНЛУ, 2023. – 112 с. [Електронний ресурс]. Режим доступу: <https://rep.knlu.edu.ua/xmlui/handle/787878787/3115>
2. Гладченко О., Ратушняк Т., Беспарточна О. Сучасні інформаційно-комунікаційні технології для забезпечення якісного дистанційного навчання [Електронний ресурс] // *Теорія і практика управління соціальними системами*. – 2022. – № 2. – С. 56–63. – Режим доступу: <http://tipus.khpi.edu.ua/article/view/260402>
3. Жуковська В. В. Лінгвістичний корпус як новітній інформаційно-дослідницький інструментарій сучасного мовознавства [Електронний ресурс] // *Вчені записки ТНУ імені В. І. Вернадського. Серія: Філологія. Соціальні комунікації*. – 2020. – Т. 31 (70), № 3 Ч. 1. – С. 113–119. – Режим доступу: <https://www.academia.edu/44519093/>
4. Іваненко Н. В., Денека М. Т., Сидоренко О. М. Цифровізація освітнього процесу в галузі філологічної науки (українські реалії) [Електронний ресурс] // *Академічні візії*. – 2023. – Вип. 18. – С. 45–58. – Режим доступу: <https://academy-vision.org/index.php/av/article/download/296/283/286>
5. Методологія використання інформаційних систем у дослідженнях: систематичний огляд [Електронний ресурс] // *Frontiers in Psychology*. – 2022. – Т. 13. – Режим доступу: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.927636/full>
6. Перекладацькі інновації: матеріали XIII Всеукраїнської студентської науково-практичної конференції [Електронний ресурс] / редкол.: С. В. Баранова, І. К. Кобякова, О. В. Бровкіна та ін. – Суми: Сумський державний університет, 2023. – 259 с. – Режим доступу: https://gf.sumdu.edu.ua/images/naukova-diyalnist/Zbirnyk_2023.pdf
7. Проблема виокремлення перекладацьких трансформацій в українському перекладознавстві [Електронний ресурс] // *Вісник філології*. – 2022. – № 62, ч. 2. – Режим доступу: http://vestnik-philology.mgu.od.ua/archive/v62/part_2/28.pdf
8. Свідер І. А. Інформаційні технології у філології та перекладі: навчально-методичний посібник. Кам'янець-Подільський: ТОВ «Друкарня Рута», 2021. – 184 с. [Електронний ресурс]. Режим

- доступу: <https://kum.pnu.edu.ua/wp-content/uploads/sites/45/2024/10/informatsijni-tekhnologii-u-filolohichnij-nautsi.pdf>
9. Селіванова О. О. Корпусна лінгвістика // *Сучасна лінгвістика: напрями та проблеми: підручник*. Полтава: Довкілля-К, 2008. – С. 667–669. [Електронний ресурс]. Режим доступу: https://moodle.znu.edu.ua/pluginfile.php/587137/mod_resource/content/1/Селіванова.pdf
 10. Barth D., Schnell S. *Understanding Corpus Linguistics*. London: Routledge, 2021. – 248 p. [Електронний ресурс]. Режим доступу: <https://www.taylorfrancis.com/books/mono/10.4324/9780429269035/understanding-corpus-linguistics-danielle-barth-stefan-schnell>
 11. Graziosi B., Haubold J., Cowen-Breen C., Brooks C. *Machine Learning and the Future of Philology: A Case Study* [Електронний ресурс] // *Transactions of the American Philological Association*. – 2023. – Режим доступу: <https://muse.jhu.edu/article/901022>
 12. Makarenko V., Aleksieieva O., Fysiuk A., Filimonova T., Tsypliak N. *Google's Digital Tools for Education: A Selection of Tools* [Електронний ресурс] // *Journal of Curriculum and Teaching*. – 2023. – Режим доступу: <https://www.sciencedirect.com/journal/index.php/jct/article/view/26305>
 13. Mittelstadt B. D., Allo P., Taddeo M., Wachter S., Floridi L. *Ethical reasoning methods for ICT: What they are and when to use them* [Електронний ресурс] // *Information and Organization*. – 2024. – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S0169023X24000971>
 14. Stefanowitsch A. *Corpus Linguistics: A Guide to the Methodology*. Berlin: Language Science Press, 2020. – 510 p. [Електронний ресурс]. Режим доступу: <https://langsci-press.org/catalog/book/148>
 15. SYSTRAN. *Machine Translation vs. CAT Tools—Collaboration or Competition?* [Електронний ресурс] // *SYSTRAN Blog*. – 2023. – Режим доступу: <https://blog.systransoft.com/machine-translation-vs-cat-tools/>
 16. Viana V., Zyngier S., Barnbrook G. (Eds.). *Perspectives on Corpus Linguistics. Studies in Corpus Linguistics*. Vol. 48. Amsterdam: John Benjamins Publishing, 2021. – 256 p. [Електронний ресурс]. Режим доступу: <https://benjamins.com/catalog/scl.48>
 17. Wang S. *Research on Corpus Linguistics and English Language Teaching* [Електронний ресурс] // *Education Reform and Development*. – 2023. –

- Режим доступу: https://www.researchgate.net/publication/380063738_Research_on_Corpus_Linguistics_and_English_Language_Teaching
18. Zhou J., Xia Y. Register-based distribution of expressions of modality in COCA [Електронний ресурс] // *Humanities and Social Sciences Communications*. – 2024. – Режим доступу: <https://www.nature.com/articles/s41599-024-03134-x>
 19. Етика через призму інформаційних і комунікаційних наук. Етичні питання, викликані використанням цифрових технологій, охоплюють усе суспільство [Електронний ресурс] // *RFSIC*. – 2022. – № 25. – Режим доступу: <https://journals.openedition.org/rfsic/12114>

Навчальне видання

**ТОМЧАКОВСЬКА Юлія Олегівна
СТРОЧЕНКО Леся Василівна
ШВЕЛІДЗЕ Лідія Дмитрівна
ЯСНІЦЬКА Ганна В'ячеславівна**

INFOTECH in Humanities

Практикум для здобувачів другого (магістерського) рівня вищої освіти
спеціальності 035 «Філологія» / В 11 «Філологія (за спеціалізаціями)»

Підписано до друку 02.05.2025 р.
Формат 60х84 1/16. Папір офсетний.
Друк цифровий. Ум. друк. арк. 3,95. Обл.-вид. арк. 1,95.
Тираж 300 прим. Замовлення № 19-05/26.

Відділ реалізації. Тел.: (067) 542-08-01. E-mail: info@book.sumy.ua
ПФ «Видавництво “Університетська книга”»
40000, м. Суми, площа Покровська, 6
Тел.: (0542) 65-75-85. E-mail: publish@book.sumy.ua
www.book.sumy.ua, newlearning.com.ua

Свідоцтво суб'єкта видавничої справи ДК № 7461 від 05.10.2021
Віддруковано на обладнанні ПФ «Видавництво “Університетська книга”»