

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б /ОК21-2023
	Екземпляр № 1	Арк 11 / 1

ЗАТВЕРДЖЕНО

Вченою радою факультету педагогічних технологій та освіти впродовж життя



31 серпня 2023 р.,

протокол № 11

Голова Вченої ради

Оксана ЧЕРНИШ

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «Комп'ютерна та корпусна лінгвістика»

для здобувачів вищої освіти освітнього ступеня «бакалавр»
спеціальності 035 «Філологія»
освітньо-професійна програма «Філологія (прикладна лінгвістика)»
факультет педагогічних технологій та освіти впродовж життя
кафедра теоретичної та прикладної лінгвістики

Схвалено на засіданні кафедри
теоретичної та прикладної лінгвістики
28 серпня 2023 р., протокол № 7

Завідувач кафедри

_____ Людмила МОГЕЛЬНИЦЬКА

Гарант освітньо-професійної програми

_____ Людмила МОГЕЛЬНИЦЬКА

Розробник: к.ф.н., доц., декан факультету педагогічних технологій та освіти впродовж життя ЧЕРНИШ Оксана, к.ф.н., доцент кафедри теоретичної та прикладної лінгвістики, АНДРУШЕНКО Олена

Житомир
2025 – 2026 н.р.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б/ОК21-2023
	Екземпляр № 1	Арк 11 / 2

1. Опис початкової дисципліни

Найменування показників	Галузь знань, спеціальність, освітній ступінь	Характеристика навчальної дисципліни
		денна форма навчання
Кількість кредитів – 3	Галузь знань 03 «Гуманітарні науки»	нормативна
Модулів – 1	Спеціальність 035 «Філологія»	Рік підготовки:
Змістових модулів – 3		3
Загальна кількість годин – 90		Семестр
		6
Тижневих годин для денної форми навчання: аудиторних – 3 самостійних – 2,6	Освітній ступінь «бакалавр»	Лекції
		16 год.
		Практичні
		32 год.
		Лабораторні
		–
		Самостійна робота
42 год.		
	Вид контролю: екзамен	

Співвідношення кількості годин аудиторних занять до самостійної та індивідуальної роботи становить:

для денної форми навчання – 53% аудиторних занять, 47% самостійної та індивідуальної роботи.

2. Мета та завдання навчальної дисципліни

Мета навчальної дисципліни полягає в ознайомленні студентів з основними поняттями комп'ютерної та корпусної лінгвістики та їхніми методами дослідження.

Завдання курсу: - закласти необхідні базові теоретичні знання про мову і мовлення з позицій комп'ютерної лінгвістики; ознайомити студентів з типовими методами та алгоритмами розв'язання лінгвістичних задач у комп'ютерних системах; - проінформувати студентів щодо типових лінгвістичних систем.

У результаті вивчення навчальної дисципліни студент повинен **знати:**

- особливості використання комп'ютерних та корпусних технологій у сучасному мовознавстві;
- пріоритетні напрямки комп'ютерних та корпусних досліджень у сучасному мовознавстві на теренах України та за її межами;

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б/ОК21-2023
	Екземпляр № 1	Арк 11 / 3

- основні поняття та категорійний апарат комп'ютерної та корпусної лінгвістики;
- теоретичні основи опису мови і мовлення з позицій комп'ютерної та корпусної лінгвістики.

вміти:

- застосовувати методи комп'ютерної та корпусної лінгвістики для розв'язання лінгвістичних завдань;
- працювати з комп'ютерними програмами автоматичного опрацювання текстів та сучасними корпусами;
- описати базові алгоритми опрацювання текстів.

Зміст навчальної дисципліни спрямований на формування наступних **компетентностей**, визначених стандартом вищої освіти зі спеціальності 035 «Філологія»:

ЗК 12. Знання основ архітектури апаратного забезпечення, основного функціоналу операційних систем, навички використання інформаційних і комунікаційних технологій.

ЗК 13. Здатність проведення досліджень на належному рівні.

ФК 2. Здатність використовувати в професійній діяльності знання про мову як особливу знакову систему, її природу, функції, рівні.

ФК 3. Здатність використовувати в професійній діяльності знання з теорії та історії мов(и), що вивчаються(ється).

ФК 13. Здатність забезпечувати лінгвістичний супровід інформаційних систем, здійснювати автоматизовану обробку текстової інформації.

Отримані знання з навчальної дисципліни стануть складовими наступних **програмних результатів** навчання за спеціальністю 035 «Філологія»:

ПРН 6. Використовувати інформаційні й комунікаційні технології для вирішення складних спеціалізованих задач і проблем професійної діяльності.

ПРН 7. Розуміти основні проблеми філології та підходи до їх розв'язання із застосуванням доцільних методів та інноваційних підходів.

ПРН 12. Аналізувати мовні одиниці, визначати їхню взаємодію та характеризувати мовні явища і процеси, що їх зумовлюють.

ПРН 17. Збирати, аналізувати, систематизувати й інтерпретувати факти мови й мовлення й використовувати їх для розв'язання складних задач і проблем у спеціалізованих сферах професійної діяльності та/або навчання.

ПРН 20. Уміння створювати електронні мовні ресурси: корпуси, словники, перекладачі.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б /ОК21-2023
	Екземпляр № 1	Арк 11 / 4

3. Програма навчальної дисципліни

Змістовий модуль 1. Комп'ютерна та корпусна лінгвістика як провідні галузі сучасного мовознавства.

Тема 1. Основні поняття та категорійний апарат комп'ютерної та корпусної лінгвістики. Сутність понять «комп'ютерна лінгвістика» та «корпусна лінгвістика». Об'єкт, предмет і завдання комп'ютерної та корпусної лінгвістики. Зв'язок із математичними та лінгвістичними дисциплінами.

Тема 2. Історія виникнення комп'ютерної та корпусної лінгвістики. Місце комп'ютерної та корпусної лінгвістики у сучасному мовознавстві. Нагальні питання комп'ютерної та корпусної лінгвістики у східній та західній лінгвістичних традиціях.

Тема 3. Комп'ютерна та корпусна лінгвістика в Україні. Сучасний стан комп'ютерної та корпусної лінгвістики в Україні.

Тема 4. Методологія досліджень комп'ютерної та корпусної лінгвістики. Методи і методики комп'ютерної та корпусної лінгвістики.

Змістовий модуль 2. Теоретичні основи та основні напрями практичного втілення комп'ютерної лінгвістики.

Тема 5. Комп'ютерне моделювання мислення. Лінгвістична модель. Поняття лінгвістичної моделі. Вимоги до моделі. Особливості лінгвістичної моделі. Поняття мислення. Структури знань (фрейм, сценарій, скрипт, семантична сітка).

Тема 6. Практичне застосування комп'ютерної лінгвістики. Міжнародні проекти Wordnet та Framenet.

Змістовий модуль 3. Теоретичні основи та основні напрями практичного втілення корпусної лінгвістики.

Тема 7. Корпус текстів природної мови. Корпусна розмітка. Параметри корпусу текстів. Типологія корпусів. Типи корпусної розмітки. Вимоги до корпусної розмітки. Етапи створення корпусу. Стандартизація та актуалізація даних корпусу.

Тема 8. Практичне застосування корпусної лінгвістики. Особливості та сфери використання лінгвістичних корпусів. Переваги та недоліки застосування конкордансів. Лінгвістичні дослідження на матеріалі корпусів.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05-05.01/035.00.1/ Б/ОК21-2023
	Екземпляр № 1	Арк 11 / 5

4. Структура (тематичний план) навчальної дисципліни

Назви змістовних модулів і тем	Кількість годин			
	Усього	У тому числі		
		Лекції	Практичні	Самост. та інд. робота
Змістовий модуль 1. Комп'ютерна та корпусна лінгвістика як провідні галузі сучасного мовознавства.				
<u>Тема 1.</u> Основні поняття та категорійний апарат комп'ютерної та корпусної лінгвістики.	10	2	4	4
<u>Тема 2.</u> Історія виникнення комп'ютерної та корпусної лінгвістики.	12	2	4	6
<u>Тема 3.</u> Комп'ютерна та корпусна лінгвістика в Україні.	10	2	4	4
<u>Тема 4.</u> Методологія досліджень комп'ютерної та корпусної лінгвістики.	12	2	4	6
Разом за змістовим модулем	44	8	16	20
Змістовий модуль 2. Теоретичні основи та основні напрями практичного втілення комп'ютерної лінгвістики.				
<u>Тема 1.</u> Комп'ютерне моделювання мислення. Лінгвістична модель.	12	2	4	6
<u>Тема 2.</u> Практичне застосування комп'ютерної лінгвістики.	10	2	4	4
Разом за змістовим модулем	22	4	8	10
Змістовий модуль 3. Теоретичні основи та основні напрями практичного втілення корпусної лінгвістики.				
<u>Тема 1.</u> Корпус текстів природної мови. Корпусна розмітка.	12	2	4	6
<u>Тема 2.</u> Практичне застосування корпусної лінгвістики.	12	2	4	6
Разом за змістовим модулем	24	4	8	12
Усього годин	90	16	32	42

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б /ОК21-2023
	Екземпляр № 1	Арк 11 / 6

5. Теми лекційних занять

№ з/п	Назва теми	Кількість годин
1	Основні поняття та категорійний апарат комп'ютерної та корпусної лінгвістики.	2
2	Історія виникнення комп'ютерної та корпусної лінгвістики.	2
3	Комп'ютерна та корпусна лінгвістика в Україні.	2
4	Методологія досліджень комп'ютерної та корпусної лінгвістики.	2
5.	Комп'ютерне моделювання мислення. Лінгвістична модель.	2
6.	Практичне застосування комп'ютерної лінгвістики.	2
7.	Корпус текстів природної мови. Корпусна розмітка.	2
8.	Практичне застосування корпусної лінгвістики.	2
Разом		16

6. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Основні поняття та категорійний апарат комп'ютерної та корпусної лінгвістики.	4
2	Історія виникнення комп'ютерної та корпусної лінгвістики.	4
3	Комп'ютерна та корпусна лінгвістика в Україні.	4
4	Методологія досліджень комп'ютерної та корпусної лінгвістики.	4
5.	Комп'ютерне моделювання мислення. Лінгвістична модель.	4
6.	Практичне застосування комп'ютерної лінгвістики.	4
7.	Корпус текстів природної мови. Корпусна розмітка.	4
8.	Практичне застосування корпусної лінгвістики.	4
Разом		32

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б/ОК21-2023
	Екземпляр № 1	Арк 11 / 7

8. Завдання для самостійної роботи

Самостійна робота 1

Перегляд та коментування відео «SYN_1957 – Syntactic Structures (N. Chomsky)».

Режим доступу:

https://www.youtube.com/watch?v=BsMbQvVQAqQ&ab_channel=TheVirtualLinguisticsCampus

Самостійна робота 2

Перегляд та коментування відео «Syntax 2 - Trees: Crash Course Linguistics № 4».

Режим доступу: <https://www.youtube.com/watch?v=n1zpnN-6pZQ>

Самостійна робота 3

Перегляд та коментування відео «Lecture 6. Dependency Parsing».

Режим доступу: <https://www.youtube.com/watch?v=PVShkZgXznc>

Самостійна робота 4

Перегляд та коментування відео «Corpus linguistics for language lovers. - Ondřej Matuška | PG 2019».

Режим доступу: <https://www.youtube.com/watch?v=PtKMjj1yaCY>

9. Індивідуальні завдання

Підготувати усне повідомлення на одну із запропонованих тем:

1. Комп'ютерна лексикографія.
2. Комп'ютерні лінгвістичні системи (на вибір — системи комп'ютерного перекладу, комп'ютерні словники, пошукові системи, системи редагування, навчальні системи).
3. Історія розвитку комп'ютерної лінгвістики.
4. Історія розвитку корпусної лінгвістики.
5. Проблеми створення систем штучного інтелекту.
6. Границі й можливості лінгвістичного програмування.
7. Міжнародні проєкти Wordnet та Framenet.
8. Особливості автоматичного перекладу.
9. Створення електронних лінгвістичних словників і корпусів.
10. Лінгвістичне програмне забезпечення.
11. Статус комп'ютерної лінгвістики в системі лінгвістичних наук.
12. Статус корпусної лінгвістики в системі лінгвістичних наук.
13. Конкорданс у системі вивчення значення слова.
14. Використання конкордансів та корпусних технологій у сучасних лінгвістичних дослідженнях.
15. Особливості укладання конкордансу розміченого тексту.
16. Автоматизований конкорданс та його особливості.
17. Сучасні лінгвістичні технології і системи.
18. Лінгвістичний корпус як новітній інформаційно-дослідницькій інструментарій сучасної лінгвістики.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б /ОК21-2023
	Екземпляр № 1	Арк 11 / 8

10. Методи навчання

I. Методи організації та здійснення навчально-пізнавальної діяльності

1) за джерелом інформації:

- *словесні*: лекція із застосуванням комп'ютерних інформаційних технологій, семінари, пояснення, розповідь, бесіда, дискусія;
- *наочні*: спостереження, ілюстрація, демонстрація;
- *практичні*: обговорення проблемних питань, виконання тестових завдань.

2) за логікою передачі і сприймання навчальної інформації: індуктивні, дедуктивні, аналітичні, синтетичні.

3) за ступенем самостійності мислення: репродуктивні, продуктивні, пошукові, дослідницькі.

4) за ступенем керування навчальною діяльністю: під керівництвом викладача; самостійна робота студентів: з книгою; виконання індивідуальних навчальних проектів.

II. Методи стимулювання інтересу до навчання і мотивації навчально-пізнавальної діяльності:

1) методи стимулювання інтересу до навчання: навчальні дискусії; створення ситуації пізнавальної новизни; створення ситуацій зацікавленості (метод цікавих аналогій тощо).

11. Методи контролю

Навчальні досягнення студентів з дисципліни «Комп'ютерна та корпусна лінгвістика» оцінюються за модульно-рейтинговою системою, в основу якої покладено принцип поопераційної звітності, обов'язковості модульного контролю, накопичувальної системи оцінювання рівня знань, умінь та навичок.

Підсумкове оцінювання включає бали за поточну роботу студента під час лекційних та семінарських занять, виконання самостійної та індивідуальної роботи, модульну контрольну роботу.

Виконання модульної контрольної роботи здійснюється у режимі комп'ютерної діагностики або з використанням роздрукованих завдань.

Модульний контроль знань студентів здійснюється після завершення вивчення навчального матеріалу модуля.

Кількість балів за роботу з теоретичним матеріалом, на семінарських заняттях, під час виконання самостійної та індивідуальної навчально-дослідної роботи залежить від дотримання таких вимог:

- своєчасність виконання навчальних завдань;
- повний обсяг їх виконання;
- якість виконання навчальних завдань;
- самостійність виконання;
- творчий підхід у виконанні завдань;
- ініціативність у навчальній діяльності.

Методи контролю засвоєння знань з дисципліни: індивідуальне та

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б /ОК21-2023
	Екземпляр № 1	Арк 11 / 9

фронтальне усне опитування; виступ на практичних заняттях з презентацією, участь у дискусії; співбесіда; виконання модульної контрольної роботи; самоконтроль: уміння самостійно оцінювати свої знання, самоаналіз; екзамен.

У дисципліні існує можливість врахування результатів неформальної освіти. При наявності сертифікату, посвідчення чи іншого документу, що підтверджує здобуття знань, компетенцій та результатів навчання з питань, пов'язаних з даною дисципліною, можливе перезарахування окремих тем або призначення додаткових балів (не більше 10).

12. Розрахунок рейтингових балів за видами поточного (модульного) та підсумкового контролю

Навчальні досягнення студентів з дисципліни «Комп'ютерна та корпусна лінгвістика» оцінюються за модульно-рейтинговою системою, в основу якої покладено принцип поопераційної звітності, обов'язковості модульного контролю, накопичувальної системи оцінювання рівня знань, умінь та навичок.

Підсумкове оцінювання включає бали за поточну роботу студента під час лекційних та семінарських занять, виконання самостійної та індивідуальної роботи, модульну контрольну роботу.

Виконання модульної контрольної роботи здійснюється у режимі комп'ютерної діагностики або з використанням роздрукованих завдань.

Модульний контроль знань студентів здійснюється після завершення вивчення навчального матеріалу модуля.

Кількість балів за роботу з теоретичним матеріалом, на практичних заняттях, під час виконання самостійної та індивідуальної навчально-дослідної роботи залежить від дотримання таких вимог:

- своєчасність виконання навчальних завдань;
- повний обсяг їх виконання;
- якість виконання навчальних завдань;
- самостійність виконання;
- творчий підхід у виконанні завдань;
- ініціативність у навчальній діяльності.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б/ОК21-2023
	Екземпляр № 1	Арк 11 / 10

Розподіл балів

Оцінювання досягнень здобувачів за дисципліною за кількісним критерієм здійснюється за 100-бальною шкалою та шкалою ЄКТС (А, В, С, D, E, FX, F). Бали розподілені за темами дисципліни наступним чином:

Поточне тестування та самостійна робота								Сума
Змістовий модуль №1				Змістовий модуль №2		Змістовий модуль №3		
T1	T2	T3	T4	T5	T6	T7	T8	
12	12	12	12	13	13	13	13	100

Шкала оцінювання

За шкалою	Екзамен	Бали
A	Відмінно	90-100
B	Добре	82-89
C		74-81
D	Задовільно	64-73
E		60-63
FX	Незадовільно	35-59
F		0-34

13. Рекомендована література

Базова література:

1. Biber D., Reppen R. English corpus linguistics. Cambridge University Press. 2015. 642 p.
2. Lange C., Leuckert S. Corpus Linguistics for World Englishes: A Guide for Research, New-York: Routledge, 2020. 236 p.
3. Rühlemann Ch. Corpus Linguistics for Pragmatics: A Guide for Research – New York, London : Routledge, Taylor & Francis Group, 2019. 206 p.
4. Stefanowitsch, A. Corpus linguistics: a guide to the methodology. Berlin: Language Science, 2020. 490 p.

Додаткова:

1. Волошин В.Г. Комп'ютерна лінгвістика: Навч. посібник. Суми: Університетська книга, 2004.
2. Дарчук Н. П. Комп'ютерна лінгвістика (автоматичне опрацювання тексту):

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-31.05- 05.01/035.00.1/ Б /ОК21-2023
	Екземпляр № 1	Арк 11 / 11

- підручник. К.: Видавничо-поліграфічний центр “Київський університет”, 2008.
3. Карпіловська Є. А. Вступ до комп’ютерної лінгвістики. Донецьк: Юго-Восток, 2003.
 4. Лук’янчук С. Комп’ютерна модель парадигматичних класів дієслів // Українське мовознавство. 2000. Вип. 22. С. 82-85.
 5. Пещак М. М. Нариси з комп’ютерної лінгвістики. Ужгород: Видавництво закарпаття, 1999. 200 с.
 6. Andrushenko, O. Iu. Lancsbox software options for the prospective investigation of the multilingual corpus for European studies. Вісник КНЛУ. Серія Філологія, 2023. Вип. 26 (1). С. 16–32.
 7. Garcia Fernandes, L. Sources and Steps of Corpus Lemmatization. Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics. (2020). Vol. 33:2. P. 416–442
 8. Geyken, A. & Gloning, T. (2015). A living text archive of 15th-19th-century German: Corpus strategies, technology, organization. In J. Gippert & R. Gehrke (Eds.), Historical corpora. Challenges and perspectives (pp. 165–179). Tübingen: Narr.
 9. Lüdeling, A., J. Ritz, M. Stede, A. Zeldes. Corpus linguistics and information structure research, in: C. Fery, Sh. Ishihara (Eds.), The Oxford Handbook of Information Structure, OUP, Oxford, 2016, 3. 599–620.
 10. McEnery, T. & Hardie, A. Corpus Linguistics. New-York: Routledge. 2015.
 11. Toupin F., Gatelais S., Sasu I. (Eds.). Studies in Linguistic Variation and Change 3. Corpus Based Research in English Syntax and Lexis. Cambridge: Cambridge Scholar Publishing, 2020. 214 p.
 12. Weisser M. Practical Corpus Linguistics: An Introduction to Corpus-Based Language Analysis. Wiley, Blackwell, 2016. 287 p.

Інформаційні ресурси:

1. Big data and Language 1 <https://www.coursera.org/learn/big-data-language-1>
2. <https://www.english-corpora.org>
3. <https://www.ice-corpora.uzh.ch/en.html>
4. <https://childes.talkbank.org>