

Київський університет імені Бориса Грінченка
Інститут філології
Кафедра української мови

“ЗАТВЕРДЖУЮ”

Проректор з науково-методичної
та навчальної роботи

“ 08 ” “ 09 ” О.Б.Жильцов
2020 року



РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Лінгвістика комп'ютерного аналізу і обробки інформації:
вступ до комп'ютерної лінгвістики

для студентів

Галузі знань	03 Гуманітарні науки
Спеціальності	035 Філологія
Спеціалізації	035.01 Українська мова та література
Освітньої програми	035.01.01 Українська мова та література
Додаткової спеціалізація	Прикладна лінгвістика
Освітнього рівня	першого (бакалаврського)

Київ – 2020



Розробники:

Вінтонів Михайло Олександрович, доктор філологічних наук, професор кафедри української мови.

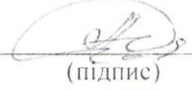
Викладачі:

Вінтонів Михайло Олександрович, доктор філологічних наук, професор кафедри української мови.

Робочу програму розглянуто і затверджено на засіданні
кафедри української мови _____

Протокол від 08.09.2021 № 2

Завідувач кафедри _____

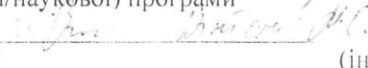


(С.О. Караман)

(підпис)

(ініціали, прізвище)

Гарант освітньої (професійної/наукової) програми
(керівник проектної групи) _____



(ініціали, прізвище)

Робочу програму перевірено

_____ 2020 р.

Заступник директора/декана _____



(О.І. Замарь)

Пролонговано:

на 2021/2022 н.р.



(Вінтонів Ф.О.)

«08»09 2021 р., протокол № 2

(підпис)

(ПІБ)

на 20___/20___ н.р.

(_____)

«___» 20___ р., протокол №

(підпис)

(ПІБ)

на 20___/20___ н.р.

(_____)

«___» 20___ р., протокол №

(підпис)

(ПІБ)

на 20___/20___ н.р.

(_____)

«___» 20___ р., протокол №

(підпис)

(ПІБ)

1. Опис навчальної дисципліни

Найменування показників	Характеристика дисципліни за формами навчання
	денна
Лінгвістика комп'ютерного аналізу і обробки інформації: вступ до комп'ютерної лінгвістики	
Вид дисципліни	<i>вибіркова</i>
Мова викладання, навчання та оцінювання	<i>українська</i>
Загальний обсяг кредитів / годин	20 /600
Курс	2-4
Семестр	3-7
Кількість змістових модулів з розподілом:	5
Обсяг кредитів	20
Обсяг годин, у тому числі:	600
Аудиторні	280
Модульний контроль	40
Семестровий контроль	30
Самостійна робота	280
Форма семестрового контролю	<i>залік (3,4, 6,7)</i>

Змістовий модуль 1. Вступ до комп'ютерної лінгвістики

Найменування показників	Характеристика дисципліни за формами навчання
	денна
Вид дисципліни	<i>вибіркова</i>
Мова викладання, навчання та оцінювання	<i>українська</i>
Загальний обсяг кредитів / годин	4/ 120
Курс	2
Семестр	3
Кількість змістових модулів з розподілом:	2
Обсяг кредитів	4
Обсяг годин, у тому числі:	120
Аудиторні	56
Модульний контроль	8
Семестровий контроль	
Самостійна робота	56
Форма семестрового контролю	<i>залік</i>

Київський університет імені Бориса Грінченка
Інститут філології
Кафедра української мови

“ЗАТВЕРДЖУЮ”
Проректор з науково-методичної
та навчальної роботи

О.Б.Жильцов
“ _____ ” _____ 2020 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

**Лінгвістика комп’ютерного аналізу і обробки інформації:
вступ до комп’ютерної лінгвістики**

для студентів

Галузі знань	03 Гуманітарні науки
Спеціальності	035 Філологія
Спеціалізації	035.01 Українська мова та література
Освітньої програми	035.01.01 Українська мова та література
Додаткової спеціалізація	Прикладна лінгвістика
Освітнього рівня	першого (бакалаврського)

Київ – 2020

Розробники:

Вінтонів Михайло Олексійович, доктор філологічних наук, професор кафедри української мови.

Викладач:

Вінтонів Михайло Олексійович, доктор філологічних наук, професор кафедри української мови.

Робочу програма розглянуто і затверджено на засіданні
кафедри української мови

Протокол від 02. 09. 2020 р. № 2

Завідувач кафедри _____ (С.О. Караман)
(підпис) (ініціали, прізвище)

Гарант освітньої (професійної/наукової) програми
(керівник проектної групи) _____ (М. О. Вінтонів)
(ініціали, прізвище)

Робочу програму перевірено
____.____. 2020 р.

Заступник директора/декана _____ ()
Пролонговано:

на 20__/20__ н.р. _____ (______). «__»__20__р., протокол №

(підпис) (ПІБ)

на 20__/20__ н.р. _____ (______). «__»__20__р., протокол №

(підпис) (ПІБ)

на 20__/20__ н.р. _____ (______). «__»__20__р., протокол №

(підпис) (ПІБ)

на 20__/20__ н.р. _____ (______). «__»__20__р., протокол №

(підпис) (ПІБ)

1. Опис навчальної дисципліни

Найменування показників	Характеристика дисципліни за формами навчання
	денна
Лінгвістика комп'ютерного аналізу і обробки інформації: вступ до комп'ютерної лінгвістики	
Вид дисципліни	<i>вибіркова</i>
Мова викладання, навчання та оцінювання	<i>українська</i>
Загальний обсяг кредитів / годин	<i>20 /600</i>
Курс	<i>2-4</i>
Семестр	<i>3-7</i>
Кількість змістових модулів з розподілом:	<i>5</i>
Обсяг кредитів	<i>20</i>
Обсяг годин, у тому числі:	<i>600</i>
Аудиторні	<i>280</i>
Модульний контроль	<i>40</i>
Семестровий контроль	<i>30</i>
Самостійна робота	<i>280</i>
Форма семестрового контролю	<i>залік (3,4, 6,7)</i>

Змістовий модуль 1. Вступ до комп'ютерної лінгвістики

Найменування показників	Характеристика дисципліни за формами навчання
	денна
Вид дисципліни	<i>вибіркова</i>
Мова викладання, навчання та оцінювання	<i>українська</i>
Загальний обсяг кредитів / годин	<i>4/ 120</i>
Курс	<i>2</i>
Семестр	<i>3</i>
Кількість змістових модулів з розподілом:	<i>2</i>
Обсяг кредитів	<i>4</i>
Обсяг годин, у тому числі:	<i>120</i>
Аудиторні	<i>56</i>
Модульний контроль	<i>8</i>
Семестровий контроль	
Самостійна робота	<i>56</i>
Форма семестрового контролю	<i>залік</i>

ПОЯСНЮВАЛЬНА ЗАПИСКА

II. Мета та завдання навчальної дисципліни

Мета навчальної дисципліни «Вступ до комп'ютерної лінгвістики»: навчити ознайомити студентів із сучасним станом і досягненнями в галузі комп'ютерної лінгвістики, навчити розмежовувати типи інформації, знати принципи і вимоги до моделювання мовних одиниць та конструювання нових лінгвістичних об'єктів, способи графічного зображення моделей, типи лінгвістичних баз даних різної будови та призначення, аналізувати особливості презентації мовних одиниць у лінгвістичних базах даних, володіти навиками роботи з вітчизняними і зарубіжними лінгвістичними програмними продуктами (електронними словниками та бібліотеками, системами машинного перекладу та аналізу тексту, лінгвістичними базами даних і базами знань – новими лінгвістичними продуктами тощо).

Основними завданнями вивчення дисципліни «Вступ до комп'ютерної лінгвістики» є:

1) здатність розв'язувати складні спеціалізовані завдання та практичні проблеми в галузі філології, що передбачає застосування теорій та методів відповідних комп'ютерних технологій.

2) здатність до самостійного пошуку та оброблення інформації з різних джерел для розгляду конкретних питань; до ефективного використання інформаційних технологій в соціальній і професійній діяльності. Сервіс-орієнтування.

3) здатність виконувати навчально-дослідні завдання на основі аналітико-синтетичної мисленнєвої діяльності. Володіння науково-дослідними методами відповідного фахового спрямування. Уміння виявляти, ставити та вирішувати проблеми, добирати аргументи, ухвалювати рішення. Обізнаність в особливостях проектної діяльності, готовність до її реалізації. Критична обізнаність та інтелектуальна чесність. Здатність нестандартно виконувати типові й неординарні завдання. Здатність застосовувати знання у практичних ситуаціях.

4) здатність до самостійної пізнавальної діяльності, самоорганізації та саморозвитку. Спрямованість на розкриття особистісного потенціалу та самореалізацію. Прагнення до особистісно-професійного лідерства та успіху.

5) здатність аналізувати лінгвістичні явища з погляду фундаментальних філологічних принципів і знань, а також на основі відповідних загальнонаукових методів. Знання природи і сутності мови, походження й закономірностей розвитку мов, особливостей структури мови, класифікації мов світу, шляхів і методів вивчення мовного матеріалу. Знання системи сучасної української літературної мови, еволюції її граматичної будови.

6) здатність застосовувати нелінгвістичні знання в практичній філологічній діяльності. Уміти застосовувати знання ІКТ-технологій для виконання лінгвістичних завдань.

Під час занять, індивідуальної навчально-дослідницької та самостійної роботи студенти, майбутні бакалаври з комп'ютерної лінгвістики, *набувають таких програмних компетентностей: ЗК3, ЗК7, ФК1, ФК 2, ФК 5, ФК 6.*

• **ЗК3:** Інформаційна компетенція Здатність до самостійного пошуку та оброблення інформації з різних джерел для розгляду конкретних питань; до ефективного використання інформаційних технологій в соціальній і професійній діяльності. Сервіс-орієнтування.

ЗК7: Самоосвітня компетенція Здатність до самостійної пізнавальної діяльності, самоорганізації та саморозвитку. Спрямованість на розкриття особистісного потенціалу та самореалізацію. Прагнення до особистісно-професійного лідерства та успіху.

ФК2 Лінгвістична (глибокі знання та розуміння в галузі лінгвістики) Здатність аналізувати лінгвістичні явища з погляду фундаментальних філологічних принципів і знань, а також на основі відповідних загальнонаукових методів. Знання природи і сутності мови, походження й закономірностей розвитку мов, особливостей структури мови, класифікації мов світу, шляхів і методів вивчення мовного матеріалу. Знання системи сучасної української літературної мови, еволюції її граматичної будови

ФК5 Прикладна лінгвістично-орієнтована Здатність застосовувати нелінгвістичні знання в практичній філологічній діяльності. Уміти застосовувати знання ІКТ-технологій для виконання лінгвістичних завдань. Володіти базовими знаннями з юриспруденції, вміти застосовувати їх під час аналізу мовної політики та мовної ситуації в державі, проведення експертиз юридично значимих текстів.

ФК6 Комунікативна професійно-орієнтована Здатність застосовувати в професійній діяльності нормативні засоби мови в усному та писемному мовленні з урахуванням змістового наповнення, соціально-демографічних особливостей співрозмовника, специфіки ситуації спілкування та контексту. Здатність вільно й ефективно використовувати мову(и), що вивчається(ються), для розв'язання комунікативних завдань.

Організація навчальної діяльності в центрі культури української мови

Реформування системи вищої освіти на сучасному етапі зумовляє концептуальні зміни в напрямках, цілях, завданнях, змісті освіти, філологічної зокрема, які зорієнтовують передовсім на розвиток особистості, формування творчої ініціативи, мобільності, підвищення конкурентоздатності майбутнього фахівця.

Задля ефективної реалізації концепції нової освітньої стратегії університету програмою передбачено застосування можливостей Центру культури української мови.

Для цього впроваджено систематичне застосування інноваційних тренінгів, дослідницьких завдань, що передбачають карткування укладання мультимедійних словників, опрацювання спеціальної літератури тощо.

Крім того, студентів залучено до роботи в проблемних групах, до участі в наукових конференціях; мовно-інформаційних конкурсах тощо.

У Центрі заплановано презентації наукових видань з проблем культури української мови, зустрічі з науковцями.

3. Результати навчання за дисципліною

Опанувавши дисципліну «Вступ до комп'ютерної лінгвістики», студенти повинні мати такі *програмні результати навчання*:

ПРН-3-6: Володіння методами і прийомами здійснення усної та писемної комунікації; методами і прийомами створення різних типів і видів текстів (мовних творів); методами і прийомами аналізу й інтерпретації різних типів і видів текстів (мовних творів); методами і прийомами обробки та трансформації різних типів і видів текстів (творів).

ПРН-3-7: Знання основних бібліографічних покажчиків і пошукових систем, володіння прийомами бібліографічного опису різних джерел; сучасними методами і прийомами обробки інформації в галузі професійної діяльності.

ПРН-3-9: Здатність застосовувати нелінгвістичні знання у філологічній діяльності. Знання мовного законодавства, уміння застосовувати його у філологічній та педагогічній діяльності, аналізувати конфліктні та юридично значимі тексти. Оволодіння навичками застосовувати ІКТ-технології, лінгвістичні програмні продукти для створення баз даних, автоматизованих лексикографічних систем, систем машинного перекладу і користування ними.

ПРН-У-1: Вільне володіння базовим професійним категоріальнопоняттєвим та дослідницьким апаратом сучасної філологічної науки. Уміння пояснювати професійно з історичних позицій явища і факти української мови та літератури

ПРН-У-7: Ефективно застосовувати ІКТ-технології у науково-дослідній і науково-педагогічній діяльності.

ПРН-У-8: Уміння працювати самостійно, працювати з метою підвищення фахової компетентності, самоосвіти; уміння працювати в умовах обмеженого часу з акцентом на професійну сумлінність та унеможливлення плагіату.

ПРН-С-1: Фахово формувати й аргументувати власні теоретичні і практичні судження, професійні позиції

ПРН-С-3: Здатність формувати судження про значення і наслідки своєї професійної діяльності з урахуванням соціальних, професійних і етичних позицій.

III. Структура навчальної дисципліни

Тематичне планування навчальної дисципліни

Назви змістових модулів і тем	Кількість годин						
	денна форма						
	Усього	у тому числі					МК
		л	пр	сем.	семестр. контроль	с.р.	
1	2	3	4		6	7	8
Змістовий модуль 1. Комп'ютерні лінгвістичні системи							
Тема 1. Вступне	6		2	2		2	
Тема 2. Комп'ютерна лінгвістика як галузь прикладної лінгвістики	7	2		2		3	
Тема 3. Інформація та дані.	5			2		3	
Тема 4. Структурна та математична лінгвістики	7	2		2		3	
Тема 5. Моделювання як основний метод прикладної лінгвістики	9		2	2		3	2
Тема 6. Дані і знання.	7	2		2		3	
Тема 7. Інтегрована лексикографічні системи	5			2		3	
Тема 8. ПроЛінг Офіс для WINDOWS	5			2		3	
Тема 9. Морфемно-словотвірний фонд української мови в Інституті мовознавства ім.О.О.Потебні НАН України	7	2		2		3	
Тема 10. Комп'ютерні лінгвістичні системи, їх типи	9	2		2		3	2

Всього за 1 модуль	67	10	4	20		29	4
--------------------	----	----	---	----	--	----	---

Змістовий модуль 2.							
Бази даних. Системи машинного перекладу							
Тема 1. Бази даних і бази знань	7		2	2		3	
Тема 2. Лінгвістичні проблеми створення баз знань	5			2		3	
Тема 3. База даних українських прийменників	5			2		3	
Тема 4. Комп'ютерна лексикографія	5			2		3	
Тема 5. Можливості і принципи роботи з програмним продуктом Lingvo	7			2		3	2
Тема 6. Системи машинного перекладу	7		2	2		3	
Тема 7. Машинний переклад за допомогою програми Pragma	5			2		3	
Тема 8. Корпуси текстів: основні типи і структура	5			2		3	
Тема 9. Використання можливостей мережі Internet у лінгвістичних дослідженнях	7			2		3	2
Всього за 2 модуль	53		4	18		27	4
Всього	120	10	8	38		56	8

IV. Програма навчальної дисципліни

Змістовий модуль 1. Комп'ютерні лінгвістичні системи

Тема 1. Вступне

Основні питання теми:

Комп'ютерна лінгвістика в системі лінгвістичних дисциплін
Комп'ютерна лінгвістика: діахронний та синхронний аспекти

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 2 Комп'ютерна лінгвістика як галузь прикладної лінгвістики

Основні питання теми:

Комп'ютерна лінгвістика як різновид прикладної лінгвістики.
Специфіка предмета її дослідження.

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 3 Інформація та дані.

Основні питання теми:

Типи та функції інформації.

Мовна інформація.

Структурування та формалізація інформації.

Мовні дані

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 4 Структурна та математична лінгвістики

Основні питання теми:

Структурна та математичної лінгвістика: основні поняття

Поняттєвий та методико-процедурний апарат сучасних прикладної та комп'ютерної лінгвістики

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 5 Моделювання як основний метод прикладної лінгвістики.

Основні питання теми:

Лінгвістична модель.

Вимоги до побудови лінгвістичних моделей, типи моделей

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 6. Дані і знання.

Основні питання теми:

Дані і знання.

Лінгвістичний алгоритм та лінгвістичний процесор

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 7 Інтегровані лексикографічні системи

Основні питання теми:

Інтегровані лексикографічні системи
Система "Словники України"

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 8 ПроЛінг Офіс для WINDOWS

Основні питання теми:

ПроЛінг Офіс для WINDOWS

Система орфографічного контролю тексту РУТА

Система російсько-українського та українсько-російського машинного перекладу ПЛАЙ

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 9 Морфемно-словотвірний фонд української мови в Інституті мовознавства ім.О.О.Потебні НАН України

Основні питання теми:

Морфемно-словотвірний фонд української мови в Інституті мовознавства ім.О.О.Потебні НАН України: джерела формування, принципи організації та виконувани функції

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 10 Комп'ютерні лінгвістичні системи, їх типи

Основні питання теми:

Комп'ютерна лінгвістика як різновид прикладної лінгвістики.

Специфіка предмета її дослідження.

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Змістовий модуль 2.

Бази даних. Системи машинного перекладу

Тема 1. Бази даних і бази знань

Основні питання теми:

Загальна типологія баз даних

Методологія створення баз даних

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 2 Лінгвістичні проблеми створення баз знань

Основні питання теми:

Лінгвістичні проблеми створення баз знань

База даних українських прийменників

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 3 База даних українських прийменників

Основні питання теми:

Можливості бази даних українських прийменників, принципи і джерела укладання.

Структура словникової статті.

Можливості одержання лінгвістичної інформації, створення запитів у базі.

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 4 Комп'ютерна лексикографія, предмет, завдання, напрями

Основні питання теми:

Комп'ютерна лексикографія

Предмет, завдання, напрями комп'ютерної лексикографії

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 5 Програмний продукт Lingvo

Основні питання теми:

Загальна характеристика програмного продукту Lingvo

Можливості і принципи роботи з програмним продуктом Lingvo

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 6 Системи машинного перекладу

Основні питання теми:

Редагування автоматичного перекладу тексту, встановлення типу помилки та її причини.

Укладання словникових статей і внесення їх до словника користувача програми автоматичного перекладу.

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 7 Машинний переклад за допомогою програми Pragma

Основні питання теми:

Загальна характеристика програмного продукту Pragma: словники, мінімальні системні вимоги.

Можливості Pragma: переклад слів і словосполучень.

Типові помилки автоматичного перекладача Pragma та їхні причини.

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 8 Корпуси текстів: основні типи і структура

Основні питання теми:

Корпусна лінгвістика як галузь комп'ютерної лінгвістики. Предмет і завдання корпусної лінгвістики

Типи корпусів текстів, їх можливості та використання в лінгвістичних дослідженнях.

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

Тема 9 Використання можливостей мережі Internet у лінгвістичних дослідженнях

Основні питання теми:

Лінгвістичні ресурси в мережі Internet.

Робота в мережі Internet як спосіб роботи лінгвіста з мовним матеріалом: принципи пошуку мовних одиниць, встановлення контексту, стильової належності, частотності слів або їхніх складників чи комбінацій і под.

Рекомендована література:

Основна: 1–5.

Додаткова: 1, 14, 17, 33, 53, 57, 58, 75, 76, 86.

6. Контроль навчальних досягнень

5.1. Система оцінювання навчальних досягнень студентів

Вид діяльності студента	Максимальна кількість	Модуль 1		Модуль 2	
		кількість одиниць	максимальна	кількість одиниць	максимальна
Відвідування лекцій	1	5	5		
Відвідування семінарських занять	1	10	10	9	9

Відвідування практичних занять	1	2	2	2	2
Робота на семінарському занятті	10	10	100	9	90
Робота на практичному занятті	10	2	20	2	20
Лабораторна робота (зокрема допуск, виконання, захист)	-	-	-	-	-
Виконання завдань для самостійної роботи	5	4	20	4	20
Виконання модульної роботи	25	2	50	2	50
Виконання ІНДЗ	-	-	-	-	-
	Разом	-	207	-	191
Максимальна кількість балів:	398				
Розрахунок коефіцієнта:	4				

VI. САМОСТІЙНА РОБОТА

VI.1. Завдання для самостійної роботи

Опрацювання наукової літератури (5 балів за кожну опрацьовану статтю) Модуль 1.

Тема 1. Комп'ютерна лінгвістика як галузь прикладної лінгвістики

Городецкий Б.Ю. Компьютерная лингвистика: моделирование языкового общения // Новое в зарубежной лингвистике. Вып. 24: Компьютерная лингвистика. М., 1989. – С. 5-29.

Марчук Ю.Н. Информатика и компьютерная лингвистика // Марчук Ю.Н. Основы компьютерной лингвистики. – М.: Народный учитель, 2000.– С. 7-23.

Тема 2. Інформація та дані. Типи та функції інформації. Мовна інформація. Структурування та формалізація інформації. Мовні дані

Васильев В.И., Шевченко А.И. Искусственный интеллект: Формирование и опознавание образов: Изд. 2-ое, испр., дополн. – Донецк: ДонГИИИ, 2000. – С. 57-79. Або: Искусственный интеллект. – В 3-х кн. – Кн. 2: Модели и методы: Справочник / Под ред. Д.А.Поспелова. – М.: Радио и связь, 1990. – С. 7-64.

Тема 3. Внесок структурної та математичної лінгвістики у розроблення поняттєвого та методико-процедурного апарату сучасних прикладної та комп'ютерної лінгвістики

Кочерган М.П. Загальне мовознавство. – К.: Академія, 1999. – С. 224-240, 246-255 (Структурний метод. Застосування математичних методів у мовознавстві)

Хомский Н. Несколько методологических замечаний о порождающей грамматике // Вопросы языкознания. – 1962. – № 4. Або: Шаумян С.К. Структурная лингвистика. – М.: Наука, 1965. – С. 151-182.

Тема 4. Моделювання як основний метод прикладної та комп'ютерної лінгвістики. Лінгвістична модель, вимоги до її побудови, типи моделелей

Карпов В.А. Язык как система. – М.: Едиториал УРСС, 2003. – С. 256-263.

Марчук Ю.Н. Моделирование в лингвистической информатике // Марчук Ю.Н. Основы компьютерной лингвистики. – М.: Народный учитель, 2000.– С. 134-170.

Модуль 2.

Опрацювання наукової літератури

Тема 1. Комп'ютерні лінгвістичні системи, їх типи

Сніжко Н.В., Сніжко М.Д. "Ідеографічний тезаурус" як інформаційно-довідкова система при вивченні закономірностей структурно-функціональної організації лексики // Мовознавство. – 1996. – № 4-5. – С.23-28.

Марчук Ю.Н. Основы компьютерной лингвистики. – М.: Народный учитель, 2000. – С. 42-80.

Тема 2. Бази даних і бази знань

Берко А., Верес О. Організація баз даних: практичний курс. – Львів: Національний університет "Львівська політехніка", 2004. – 152 с.

Русанівський В.М., Тараненко О.О., Широков В.А. Теоретико-лінгвістичні засади та інформаційно-комп'ютерне забезпечення україномовних лінгвістичних інтелектуальних систем // Мовознавство. – 1996. – № 4-5.

Тема 3. Лінгвістичні проблеми створення баз знань

Коссак О., Коруд О., Хвищун Л. Система комп'ютерної підтримки словникових баз даних СЛОВО // Проблеми українізації комп'ютерів. – К., 1993. – С. 73-76.

Клименко Н.Ф., Русанівський В.М. Від універсальної бази лінгвістичних знань до комп'ютерного укладання словників // Мовознавство. – 1995. – № 4-5. – С. 3-10.

Тема 4. Комп'ютерна лексикографія, предмет, завдання, напрями

Автоматичний словник // Українська мова: Енциклопедія. – К.: Українська енциклопедія, 2000, 1-е вид або 2004, 2-е вид. – С. 10.

Ардан Р.В., Бацевич Ф.С., Партико З.В. Комп'ютерний словник-мінімум української мови // Мовознавство. – 1996. – № 4-5.

Критерії оцінювання виконання студентами завдань для самостійної роботи

Самостійна робота студентів з навчальної дисципліни «Вступ до комп'ютерної лінгвістики» оцінюється за наступними параметрами:

5 балів – завдання для самостійної роботи виконані в повному обсязі, лаконічно, змістовно, обґрунтовано, із наведенням прикладів і застосуванням методу аналізу та творчого підходу. Студент самостійно працює з науковою літературою; вміє формулювати й обґрунтовувати власну позицію; володіє правилами написання та підготовки виступів, доповідей, бесід; креативно підходить до їх підготовки; має комунікативні вміння і навички; прагне до самовдосконалення і саморозвитку.

4 балів – завдання для самостійної роботи виконані в повному обсязі, лаконічно, змістовно, обґрунтовано, але з певними помилками щодо викладу матеріалу та без наведення прикладів до окремих фактів і явищ. Студент працює із запропонованою науковою літературою; робить висновки; володіє правилами написання та підготовки виступів, доповідей, бесід і втілює їх у практику. Однак у процесі відтворення самостійно дібраного матеріалу простежується брак власних суджень, прояву креативності й ініціативності.

3 бали – завдання для самостійної роботи виконані в повному обсязі, але з порушенням логіки й послідовності викладу матеріалу та без ілюстрування прикладами. У роботі допускаються стилістичні помилки, неточне вживання термінів, довільне витлумачення фактів. Студент працює із запропонованою науковою літературою, але аналізує її примітивно, без висвітлення власної позиції; володіє правилами написання та підготовки виступів, доповідей, бесід, але втілює їх у практику допускаючи грубі помилки та уникаючи використання фахової термінології.

2 бали – завдання для самостійної роботи виконані не в повному обсязі, частково самостійно і за певним зразком. Студент володіє матеріалом на початковому рівні, викладає його хаотично і необґрунтовано, без дотримання мовних норм; має фрагментарні навички роботи з науковими джерелами і не вміє робити висновки.

1 бал – виконано менше половини завдань для самостійної роботи. Матеріал фрагментарний, не систематизований, не ілюстрований прикладами, порушена структура викладу. Його основу складає необґрунтоване копіювання чужих думок, фраз і позицій. Студент не вміє працювати з науковими джерелами і робити власні висновки; не володіє правилами написання та підготовки виступів, доповідей, бесід; відсутній творчий компонент даного виду діяльності; комунікативні вміння і навички мають дуже низький рівень розвитку.

0 балів – більшу частину самостійних завдань не виконано. Студент лише частково володіє навчальним матеріалом, не вміє чітко, лаконічно й послідовно висвітлювати його; не працює з науковими і навчальними джерелами і не володіє правилами написання та підготовки виступів, доповідей, бесід; майже відсутні творчі та комунікативні уміння і навички.

VI.2. Форми проведення модульного контролю та критерії оцінювання

Кожний модуль включає бали за *модульну контрольну роботу* (модульний контроль).

Виконання модульних контрольних робіт здійснюємо в системі "Мудл".

Модульний контроль здійснюємо після завершення вивчення навчального матеріалу модуля.

Критерії оцінювання виконання студентами модульних контрольних робіт

Модульні контрольні роботи студенти виконують на платформі "Мудл" у формі тестів. Один варіант містить 13 запитань. 12 запитань оцінено по 2 бали за кожну правильну відповідь (максимальна оцінка 24 б) і одне запитання оцінено на 1 бал за правильну відповідь (максимальна оцінка 1 б). Всього 25 балів.

VI.3. Форми проведення семестрового контролю та критерії оцінювання

Семестровий контроль із навчальної дисципліни «Вступ до комп'ютерної лінгвістики» здійснюється у формі *заліку*.

Орієнтовний перелік питань для семестрового контролю

1. Комп'ютерна лінгвістика як різновид прикладної лінгвістики. Специфіка предмета її дослідження.
2. Інформація і дані. Визначення і відмінність.
3. Поняття структурування (поділу, диференціації) та формалізації інформації.
4. Типи інформації за класифікацією А.О.Білецького.
5. Типи інформації за джерелом її створення і за її формою. Приклади.
6. Типи інформації за її зв'язком із ситуацією і за джерелом її створення. Приклади.

7. Типи інформації за способом її оформлення, за сферою її призначення, характером її функціонального навантаження. Приклади.
8. Мовна інформація. Основні та додаткові її ознаки.
9. Поняття лінгвістичної моделі. Співвідношення моделі та її натурального об'єкта (оригіналу, прототипа, прообразу). Вимоги до побудови лінгвістичних моделей.
10. Поняття вирізняльних (диференційних) ознак мовних одиниць. Поняття опозиції мовних одиниць. Підстава опозиції (спільна ознака для зіставлення). Приклади.
11. Принципи побудови моделей фонем Р.Г.Піотровського.
12. Характеристика різних способів графічного зображення моделей Р.Г.Піотровського (деревна та кубічна (стереометрична). Табличний спосіб представлення даних. Пояснити на прикладі класифікацій голосних та приголосних фонем української мови.
13. Поняття інтерфейса, меню, режиму роботи.
14. Символьна модель морфемної будови слова. Принципи її побудови.
15. Типи моделей будови українського слова. Активні та унікальні моделі.
16. Знання. Їхнє співвідношення з поняттями "інформація" та "дані". Знання декларативні і знання процедурні. Способи представлення. Приклади.
17. Поняття алгоритму. Блок-схема, або граф-схема алгоритму та її складники. Крок алгоритму, правила алгоритму. Вимоги до побудови алгоритму.
18. Лінгвістичний алгоритм та лінгвістичний процесор.
19. Поняття комп'ютерної лінгвістичної системи, її типи.
20. Комп'ютерні лексикографічні системи (на прикладі "Інтегрованої лексикографічної системи" Мовно-інформаційного фонду НАН України).
21. Комп'ютерні системи опрацювання тексту (на прикладі систем РУТА, ПЛАЙ).
22. Загальна характеристика лексикографічної системи "Словники України": авторський колектив, загальна концепція, мінімальні системні вимоги, джерельна база, режими роботи з системою, характеристика типів одержуваної лінгвістичної інформації. Поняття *інтегральна автоматизована, або комп'ютерна лексикографічна система*.
23. Структура і можливості лексикографічної системи "Словники України": основний реєстр слів, транскрипція, парадигма, частиномовний статус, синоніми, антоніми, фразеологізми, можливості поповнення бази даних системи – додавання нових слів з їхніми характеристиками.
24. Характеристика словників, які забезпечують функціонування системи "Словники України" в режимах "Парадигма", "Транскрипція", "Синонімія", "Фразеологія", "Антонімія".
25. Загальна характеристика ПроЛінг Офіс для WINDOWS: авторський колектив, загальна концепція, мінімальні системні вимоги. Характеристика систем опрацювання тексту в складі ПроЛінг Офісу: системи орфографічного контролю РУТА та системи машинного перекладу ПЛАЙ.
26. Можливості програми РУТА: основний словник, пошук синонімів, парадигма слова, перевірка орфографії та граматики, розставлення переносів.
27. Можливості програми РУТА: додавання нових слів до словника, пошук та заміна слів, система слів-підказок. Поняття помилконебезпечного місця.
28. Переклад слів і текстів за допомогою програми ПЛАЙ. Характеристика словникової бази для забезпечення роботи цієї системи з текстом.
29. Типові помилки при автоматичному перекладі за допомогою програми ПЛАЙ та їх причини.
30. База даних морфемно-словотвірного фонду Інституту мовознавства імені О.О.Потебні НАН України: принципи укладання, структура, типи мовної інформації.

31. База даних морфемно-словотвірного фонду Інституту мовознавства імені О.О.Потебні НАН України: функції роботи фонду та можливості одержання лінгвістичної інформації у базі. Нові типи словників української мови – продуктів використання фонду: "Словник афіксальних морфем української мови" та "Кореневий гніздовий словник української мови".

Шкала оцінювання

<i>Кількість балів</i>		<i>Значення оцінки</i>
A	100-90	Відмінно – відмінний рівень знань (умінь) в межах обов'язкового матеріалу з можливими незначними недоліками
B	82-89	Дуже добре – достатньо високий рівень знань (умінь) в межах обов'язкового матеріалу без суттєвих (грубих) помилок
C	75-81	Добре – в цілому добрий рівень знань (умінь) з незначною кількістю помилок
D	69-74	Задовільно – посередній рівень знань (умінь) із значною кількістю недоліків, достатній для подальшого навчання або професійної діяльності
E	60-68	Достатньо – мінімально можливий допустимий рівень знань (умінь).
F _x	35-59	Незадовільно – з можливістю повторного складання – незадовільний рівень знань, з можливістю повторного перескладання за умови належного самостійного доопрацювання
F	1–34	Незадовільно з обов'язковим повторним вивченням курсу – досить низький рівень знань (умінь), що вимагає повторного вивчення дисципліни. Якщо дисципліна вивчається протягом двох і більше семестрів, то семестрові оцінки визначаються в установленому

VII. Навчально-методична картка дисципліни
«Лінгвістика комп'ютерного аналізу і обробки інформації»:

вступ до комп'ютерної лінгвістики

Разом 120 годин: лекції – 10 годин, практичні заняття – 8 годин, семінарські заняття - 38 годин, самостійна робота – 56 годин, модульний контроль – 8 годин.

Теми лекцій	Комп'ютерна лінгвістика як галузь прикладної лінгвістики (1 бал)	Структурна та математична лінгвістики (16)	Дані і знання. (16) Морфемно-словотвірний фонд української мови в Інституті мовознавства ім.О.О.Потебні НАН України (16)	Комп'ютерні лінгвістичні системи, їх типи (16)	
Теми практичних	Вступне (16+106)	Моделювання як основний метод прикладної лінгвістики(16+106) (16+106)	Системи машинного перекладу (16+106)	Бази даних і бази знань (16+106)	

Теми семінарських	<p>Вступ 16+106)</p> <p>Комп'ютерна лінгвістика як галузь прикладної лінгвістики (16+106)</p> <p>Інформація та дані. (16+106)</p> <p>Структурна та математична лінгвістики (16+106)</p> <p>Моделювання як основний метод прикладної лінгвістики (16+106)</p> <p>Дані і знання (16+106)</p> <p>Інтегрована лексикографічні системи (16+106)</p> <p>ПроЛінг Офіс для WINDOWS (16+106)</p> <p>Морфемно-словотвірний фонд української мови в Інституті мовознавства ім.О.О.Потебні НАН України (16+106)</p> <p>Комп'ютерні лінгвістичні системи, їх типи (16+106)</p> <p>Бази даних і бази знань (16+106)</p> <p>Лінгвістичні проблеми створення баз знань (16+106)</p> <p>База даних українських прийменників (16+106)</p> <p>Комп'ютерна лексикографія (16+106)</p> <p>Можливості і принципи роботи з програмним продуктом Linguo (16+106)</p> <p>Системи машинного перекладу (16+106)</p> <p>Машинний переклад за допомогою програми Pragma (16+106)</p> <p>Корпуси текстів: основні типи і структура (16+106)</p> <p>Використання можливостей мережі Internet у лінгвістичних дослідженнях (16+106)</p>	
Самостійна робота	4*5=20 балів	4*5=20 балів
Види поточного контролю	Модульна контрольна робота 1 (25 балів)	Модульна контрольна робота 2 (25 балів)
Підсумковий контроль	залік	

VIII. Рекомендована література

Рекомендована література

1) основна:

- Баранов А.Н. Введение в прикладную лингвистику. – М.: Едиториал УРСС, 2003. – 360 с.
- Белецкий А.А. Семиотический аспект языковой системы // Структурная и математическая лингвистика. – К., 1979. – Вып.7. – С.11-18.
- Височанський В.С., Кардаш А.І., Костів О.В., Черняхівський В.В. Елементи інформатики. – Львів, 1990.
- Волошин В.Г. Комп'ютерна лінгвістика: Навчальний посібник. – Суми: ВТД "Університетська книга", 2004. – 382 с.
- Карпіловська Є.А. Вступ до комп'ютерної лінгвістики. – Донецьк: ТОВ "Юго-Восток, Лтд", 2003. – 184 с.
- Карпіловська Є.А. Машинні версії традиційних словників як основа для укладання комп'ютерних словників і тезаурусів // Мовознавство. – 1996. – № 4-5.
- Клименко Н.Ф., Карпіловська Є.А., Комарова Л.І. та ін. Морфемно-словотвірний фонд української мови як дослідницька та інформаційно-довідкова система // Мовознавство. – 1990. – № 6. – С. 41-50.
- Клименко Н.Ф., Карпіловська Є.А. Передмова // Клименко Н.Ф., Карпіловська Є.А., Карпіловський В.С., Недозим Т.І. Словник афіксальних морфем української мови. – К., 1998.
- Лавошникова Э.К. Компьютерная проверка орфографии: вчера, сегодня, завтра // Вестник МГУ. Серия 9. Филология. – 2003. – № 5.
- Лосев А.Ф. Введение в общую теорию языковых моделей. – Изд. 2-ое, М.: Едиториал УРСС, 2004.– 296 с.
- Пещак М.М. Стан і перспективи комп'ютерної лексикографії в Україні // Мовознавство. – 1996. – № 4-5.
- Пиотровский Р.Г. Лингвистический автомат и его речемыслительное обоснование. – Минск, 1999.
- Пиотровский Р.Г. Моделирование фонологических систем и методы их сравнения. – М.; Л., 1966.
- Русанівський В.М., Широков В.А. Інформаційно-лінгвістичні основи сучасної тлумачної лексикографії // Мовознавство. – 2002. – № 6. – С. 7-48.
- Русанівський В.М., Тараненко О.О., Широков В.А. Теоретико-лінгвістичні засади та інформаційно-комп'ютерне забезпечення україномовних лінгвістичних інтелектуальних систем // Мовознавство. – 1996. – № 4-5.
- Толковый словарь по вычислительным системам. – М., 1989.
- Українська мова: Енциклопедія. – К., 2000, 1-е вид. або 2004, 2-е вид. (Статті: Автоматизація лінгвістичних досліджень. Граматики формальні. Конкорданс. Математична лінгвістика. Машинна мова. Модель. Прикладна лінгвістика. Прогнозування лінгвістичне. Формалізація в лінгвістиці).

2) додаткова:

- Андреев Н.Д. Статистико-комбинаторные методы в теоретическом и прикладном языковедении. – Ленинград, 1967.
- Анисимов А.В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык. – К., 1991.
- Апресян Ю.Д. Идеи и методы современной структурной лингвистики. – М.: Просвещение, 1966.
- Зубов А.В., Зубова И.И. Информационные технологии в лингвистике. – М.: Академия, 2004. – 206 с.
- Караулов Ю.Н. Лингвистическое конструирование й тезаурус литературного языка. – М., 1981.
- Карпов В.А. Язык как система. – М.: Едиториал УРСС, 2003. – С. 57-68, 243-255.

- Клименко Н.Ф., Карпіловська Є.А. Морфемні структури слів у сучасній українській мові // Мовознавство. – 1991. – № 4. – С. 10-21.
- Пешак М.М. Нариси з комп'ютерної лінгвістики. – Ужгород, 1999.
- Ревзин И.И. Структура языка как моделирующей системы. – М., 1978.
- Слама-Казаку Т. Место прикладной лингвистики в системе наук: отношение ПЛ к "лингвистике" // Новое в зарубежной лингвистике. Вып. XII: Прикладная лингвистика. – М., 1983.
- Шайкевич А.Я. Дистрибутивно-статистический анализ текстов. – Ленинград, 1982.
- Шаумян С.К., Соболева П.А. Основания порождающей грамматики русского языка: Введение в генотипические структуры. – М., 1968.
- Широков В.А., Рабулець О.Г., Шевченко І.В., Костишин О.М., Якименко К.М. Технологічні основи сучасної тлумачної лексикографії // Мовознавство. – 2002. – № 6. – С. 49-86.

ІХ. Інформаційні ресурси

<http://uk.wikipedia.org>; <http://books.google.com/?hl=EN>; <http://www.donnu.edu.ua>