

	Назва: <i>035_Прикладна лінгвістика Спецкурс №4</i>
КОРПУСНА ЛІНГВІСТИКА	
Тип	дисципліна професійної та практичної підготовки за вибором студента
Код	К34
Семестр	7
Загальна кількість кредитів / годин	3 кредити / 90 годин
Форма контролю	Залік
Викладач	К.філол.н., доцент Данилюк І.Г.
Необхідні обов'язкові попередні та супутні навчальні дисципліни	«СУМ: Лексикологія», «СУМ: Морфологія»
Місце у структурно-логічній схемі	К34 викладається на 4 році навчання
Форми навчання	Лекції, практичні заняття, самостійна робота
Критерії оцінювання	Поточний контроль – 100 балів Підсумковий контроль (залік)
Мова викладання	українська

ЗМІСТ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Корпусна лінгвістика та комп'ютерна лінгвістика. Корпусні лексикографія. Застосування корпусів в галузі перекладу й перекладознавства. Історія корпусної лінгвістики. Завдання й основні напрями корпусної лінгвістики. Об'єкт, предмет, основні поняття корпусної лінгвістики. Поняття корпусу. Одиниця зберігання корпусу даних. Вимоги до корпусу текстів. Розвиток лінгвістичних корпусів у світі. Перше покоління корпусів. Браунівський корпус. Ланкастерського-Осло-Бергенський корпус. Лондонсько-Лундський корпус. Друге покоління корпусів. Бірмінгемський корпус The Bank of English (The COBUILD Project). Британський національний корпус. Корпус сучасної американської англійської мови (СОСА). Останнє покоління корпусів. Типи корпусів. Анотовані й неанотовані корпуси. Усні корпуси та письмові корпуси. Змішані корпуси. Одномовні корпуси. Двомовні корпуси. Багатомовні корпуси. Паралельні корпуси. Динамічні (моніторні) корпуси. Статичні корпуси. Розмічені корпуси. Нерозмічені корпуси. Багатоцільові корпуси. Спеціалізовані корпуси. Національні корпуси. Лінгвістична анотація й метадані. Частково розмічені корпуси. Повністю розмічені корпуси. Типи анотування (розмічування) мовного корпусу. Лінгвістичні дослідження на базі корпусу: вивчення лексики. Дослідження частотності лексеми в корпусі. Аналіз розподілу частотності вживання лексеми за жанрами. Дослідження лексичної сполучуваності. Дослідження лексичної синонімії за допомогою корпусу. Індекс винятковості корпусу. Дослідження частиномовного складу тексту

(корпусу текстів) за допомогою морфологічної розмітки. Методи добування інформації з корпусу. Типи інформації. Конкорданс. Програми для роботи з корпусом: корпусні менеджери, конкордансери. Функції корпусного менеджера. Запити за допомогою шаблонів. Мова регулярних виразів. Програми AntConc, WordSmith, Dialing Concordancer, Corsis. Створення дослідницького корпусу. Анотація корпусу. Планування корпусу. Добір джерел корпусу. Визначення жанрово-тематичної структури корпусу. Збирання та оцифрування даних. Розмічування тексту. Токенізація. Лематизація. Морфологічна розмітка. Синтаксична розмітка (парсинг). Семантична розмітка. Інструменти автоматичного розмічування корпусу.

Програмні результати навчання визначені в освітній програмі

Збирати, аналізувати, систематизувати й інтерпретувати мовні факти, перекладати тексти різних стилів і жанрів **(ПРН-К-4)**

Використовувати інформаційні технології для вирішення стандартних завдань професійної діяльності **(ПРН-У-2)**

Ефективно працювати з інформацією: добирати необхідну інформацію з різних джерел, критично аналізувати й інтерпретувати її, впорядковувати, класифікувати й систематизувати **(ПРН-У-6)**.

Розбивати інформацію на компоненти, розуміти їх взаємозв'язки та організаційну структуру, бачити помилки й огріхи в логіці міркувань, різницю між фактами і наслідками, оцінювати значущість даних **(ПРН-У-7)**.

Створювати лінгвістичний алгоритм для розв'язання поставлених завдань дослідження **(ПРН-У-8)**

Знати систему мов, що вивчаються **(ПРН-З-3)**.

Планувати і здійснювати дослідження в галузі лінгвістики та інформаційних технологій на належному рівні **(ПРН-В-1)**.