



ПРИЛОЖЕНИЕ № 18

Договор: BG051PO001-3.3.06-0045

Име на проект: „Развитие на капацитета на специализанти, постдокторанти и млади учени за преподаване на академични курсове по и на чужд език със съвременни методи и информационни и компютърни технологии”

Бенефициент: СУ „Св. Кл. Охридски“ – Факултет по класически и нови филологии

Индивидуална учебна програма

за представителите на целева група А¹

Име: „Езиковите корпуси в работата на филолога ”

Ръководител на дейност: Доц. д-р Цветомира Венкова

Хорариум: 8 учебни часа

1. Цели на учебната програма/план

Трудно можем да си представим съвременната филологическа практика без използването на големи текстови масиви в електронен вид, наречени корпуси. Този курс има за цел да покаже най-новите постижения в тази област и да подобри уменията за работа с тях. Корпусите се разглеждат от гледна точка на тяхното конкретно приложение в различни филологически разработки. За тази цел се обсъждат редица практически ориентирани въпроси, като например: кои основни параметри са показателни при подбора на нужния корпус сред многообразието от ресурси; как да се направи първоначална статистическа обработка на материала и какво може да покаже тя; как бързо да се извлекат необходимите за конкретната цел данни; как да се работи с корпуси, които съдържат допълнителна информация, а в по-задълбочен аспект – и какво е филологическото познание, закодирано в някои съвременни корпуси.

2. Предварителни изисквания

Основна компютърна грамотност (Microsoft Word, Internet Explorer)

3. Теоретична подготовка

3.1. Тема 1 „Електронните езикови корпуси като ресурс ”

Съдържание брой часове/занятия 1

¹ Група А: преподаватели по чужд език; група Б: преподаватели на чужд език; излишните да се изтрият



Схема BG051PO001-3.3.06 „Подкрепа на развитието на докторанти, постдокторанти, специализанти и млади учени”

3.2. Тема 2 „Основни параметри на езиковия корпус с оглед на неговите приложения“

Съдържание брой часове/занятия 1

3.3. Тема 3 „Представяне и дискусия на конкретни езикови корпуси“

Съдържание брой часове/занятия

3.4. Тема 4 „Основни резултати от статистическата обработка на корпуси и какво показват те“

Съдържание брой часове/занятия 1

3.5. „Тема 5 „Автоматично извличане и подреждане на нужните данни от корпуса

Съдържание брой часове/занятия 1

3.6. „Тема 6 „Проблеми на граматичното аотиране на корпуси – многозначност и подбор на маркери“

Съдържание брой часове/занятия 1

3.7. Тема 7 „Корпусите на изреченско ниво – модели и подходи с оглед на изследователските цели“

Съдържание брой часове/занятия 1

3.7.1. Тема 8 „Преглед на бази, в които текстът е снабден със специализирана информация, и възможностите, които те предлагат“

Съдържание брой часове/занятия 1

4. Очаквани резултати

В резултат на обучението в този курс се очаква филолозите да развият умения за работа с електронни езикови корпуси и системи за корпусна обработка с оглед на индивидуалните си научноизследователски и академични задачи.

5. Критерии за оценяване (според създадената матрица)

6. Библиография

6.1. Основна

- McEnery and Hardie (2012): McEnery, T. and Hardie, A. *Corpus Linguistics: Method, Theory and Practice*, Cambridge University Press.



Схема BG051PO001-3.3.06 „Подкрепа на развитието на докторанти, постдокторанти, специализанти и млади учени”

6.2. Допълнителна

- Jurafsky and Martin (2008), Jurafsky, D. and Martin, J. *Speech and Language Processing*. (2nd edition). Prentice-Hall.
- Коева и Стоянова. (2009). Коева, Св., Стоянова, Ив. Български национален корпус. *Български език* 3. 137-145
- Паскалева (2007), Паскалева, Е., Компютърна морфология: ресурси и инструменти, Институт за паралелна обработка на информацията, БАН София,
- Венкова, Цв. (под печат), Един електронен корпус на българската разговорна реч в ретроспекция и развитие, *Езикът във времето и пространството*, МСД, София.
- Венкова, Цв. (2008): Уличният език в рап музиката – митове и реалност. – *Агресията на уличния език*. МСД, София, 176-182.
- Венкова, Цв. (2004): Съпоставима корпусна извадка на подчинените обстоятелствени изречения за време в българската и английската разговорна реч. - *Проблеми на българската разговорна реч VI*, Изд. “Св. Св. Кирил и Методий”, В. Търново, 67-73.
- Венкова, Цв. (2002): Bilingual corpora as a platform for cross-linguistic treebank development. - *Proceedings of TLT-1*, 264-274.
- Венкова, Цв. (1996): Corpus-Based Study of the Words Governing Da-Constructions in Bulgarian Fiction and Speech. - *Working Papers in Linguistics* 28, University of Trondheim, 319-328.

Дата: 30.01.2014

Изготвил: Доц. д-р Цветомира Венкова

Одобрил: