

Квантитативна истраживања језика

Наставник: Душица М. Ђурђевић-Филиповић, Петар Д. Милин

Статус предмета: изборни

Број ЕСПБ: 6

Услов:

Циљ предмета

а) Упознавање студената са основама савременог интердисциплинарног истраживачког програма који се бави језиком и који обухвата лингвистику, психологију, рачунарство, математику, неуронауке.

б) Оспособљавање студената за самостално извођење квантитативних истраживања на различитом језичком материјалу.

Исход предмета

Од студента се очекује да на крају курса буде способан да репродукује знања о основама савременог интердисциплинарног истраживачког програма који се бави језиком. Такође, према сопственим интересовањима и потребама студент ће бити оспособљен да спроводи базична и/или примењена квантитативна истраживања језика, у једној или више могућих подобласти: рачунарској лингвистици, квантитативној лингвистици, психолингвистици, истраживањима индивидуалних разлика, психодијагностици, психопатологији итд., као и да примењује напредну методологију карактеристичну за квантитативна истраживања језика.

Садржај предмета

Теоријска настава

I Увод у истраживања језика: Кратак поглед у лингвистику Чомског. Нове емпиријске чињенице и пробабилистичке граматике. Језик као категоријални или континуирани (пробабилистички) феномен? II Претпоставке за квантитативна истраживања језика: језички корпуси; елементи теорије вероватноће у истраживањима језика; статистика у истраживањима језика. III Одабране теме из квантитативне и рачунарске лингвистике: Зипфов закон и његове модификације; проблем карактеристичне константе текста; конкордансе; колокације; проблем утврђивања ауторства и форензичка психолингвистика. IV Квантитативно описивање различитих нивоа језика и евалуација изведених мера: опис фонетских/фонолошких, ортографских, морфолошких, семантичких и синтаксичких карактеристика језика; теорија информације и њена примена у психолингвистици. V Векторски заснована семантичка анализа. VI Напредне теме: непараметарски и параметарски модели дистрибуције учесталости речи; n-грам модели; Марковљеви модели; пробабилистичке граматике и њихова употреба у аутоматској обради језика.

Практична настава (вежбе, ДОН, студијски истраживачки рад)

Спровођење истраживања и опсежног прегледа литературе као основе за израду семинарског рада. Активности обухватају и прикупљање и обраду података и тумачење налаза у оквиру конкретног теоријског модела. Усмена и писмена презентација резултата истраживања.

Литература

Baayen, R. H. (2001). *Word Frequency Distributions*. Dordrecht: Kluwer Academic Publishers.

Biber, D., Conrad, S. and Reppen, R. (2000). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Bod, R., Hay, J. and Jannedy, S. (2003). *Probabilistic Linguistics*. Cambridge, MA: The MIT Press.

Manning, C. D. and Schuetze, H. (2000). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press.

Jurafsky, D. and Martin, J. H. (2000). *Speech and Language Processing*. New Jersey: Prentice Hall.