

## Korpuszok a nyelvészeti kutatásban szeminárium

Oktató: Dr. Vincze Veronika

Időpont: hétfő 14-15:30

Elérhetőség: <http://www.inf.u-szeged.hu/~vinczev/>

A kurzus célja, hogy a hallgatókkal megismertesse, miként használatosak a korpuszok és egyéb nyelvi adatbázisok a nyelvészeti kutatásban. A félév során áttekintjük a nyelvi adatgyűjtés módszertani követelményeit, sorra vesszük a korpuszépítés különböző lépéseit, illetve a korpuszból való adatgyűjtés módjait, továbbá az adatok statisztikai elemzésére is figyelmet fordítunk. A kurzus során minden hallgató elkészíti saját minikorpuszát, amely az elsajátított elméleti ismeretek gyakorlati alkalmazására is kiváló lehetőséget nyújt.

### Tematika

Szeptember 3.: *Bevezetés, követelmények ismertetése*

Szeptember 15.: *Nyelvi adatok, az adatgyűjtés módszertana*

Szeptember 2.: *Korpuszok és adatbázisok*

Szeptember 29.: *Korpuszok létrehozása*

Október 6.: *Annotáció, nyelvi elemzés*

Október 13.: *Annotációs eszközök*

Október 20.: *Saját korpuszépítő projektek megbeszélése*

Október 27.: *Java I. (alapok, Eclipse használata)*

November 3.: *Java II. (adatgyűjtés korpuszból)*

November 10.: *Statisztikai elemzések*

November 17.: *LaTeX*

November 24.: *Projektbemutatók*

December 1.: *A félév zárása*

### Követelmények

Az órákon való részvétel kötelező (max. 2 hiányzás megengedett).

Teljesítés: a félév végére saját minikorpusz létrehozása és az abból kigyűjtött adatok rendszerezése, statisztikai elemzése. Rövid projektbemutató tartása a félév végén. Beadandó: a kész korpusz + az elemzés leírása.

**Értékelés:** ötfokozatú minősítés (1-5).