

## **Sass Bálint**

tudományos munkatárs  
Magyar Tudományos Akadémia  
Nyelvtudományi Intézet  
Nyelvtechnológiai és Alkalmazott Nyelvészeti Osztály  
Nyelvtechnológiai Kutatócsoport

Fő kutatási területei a korpuszlekérdező felületek és a korpuszépítés. Egyik megalkotója a Magyar Nemzeti Szövegtárnak, a Magyar Történeti Szövegtárnak, az Ómagyar Korpusznak és a Budapesti Szociolingvisztikai Interjúból készített számítógépes korpusznak, illetve ezek lekérdezőfelületének. A magyar igei bővítményszerkezetek vizsgálatára kifejlesztett egy lekérdezőeszközt Mazsola néven, mellyel számos sikeres kutatást folytattak. Foglalkozik még számítógéppel segített szótáralkotással és a magyar Braille-rövidírással is. Egyetemi hallgatóknak több nyelvtechnológiai kurzust is tartott.

## **Keresés korpuszban**

Az előadás során ismertetjük a NoSkE korpuszkezelő rendszer funkcióit, a szűrési lehetőségeket, és azt, hogy miként kaphatunk a céljainknak megfelelő gyakorisági listákat. Betekintést adunk a reguláris kifejezések és a CQL lekérdezőnyelv, valamint a Magyar Történeti Szövegtár és a Magyar Nemzeti Szövegtár használatába. Bemutatjuk, hogyan hajtható végre keresés elemzetlen és elemzett korpuszban, és hogy a korpuszkeresés során milyen elveket érdemes szem előtt tartani. Végül mindezekre példákat mutatunk a cigány eredetű szavaktól az „automatikus” versírásig.

**kulcsszavak:** korpusz, szövegtár, keresés, lekérdezés, CQL