

# A Nyelvtechnológiai és Élőnyelvi Osztály kutatási célkitűzései a 2011. évre

## Élőnyelvi kutatócsoport:

- Feltáró vizsgálat a magyarországi kisebbségi közösségek nyelvcserejéről longitudinális és összehasonlító módszerrel. Ennek eredményei egy könyvben lesznek bemutatva. Forrás: OTKA K 81574.
- A BUSZI-3 és BUSZI-4 adatbázisok rossz hangminőségű interjúinak újradigitalizálása, javítása, majd ezt követően a tesztadatok kódolása. A tesztadatok első kódolása befejeztével az első ellenőrzés megkezdése.
- A BUSZI-2 anonimizálási, illetve publikálási tervének kidolgozása, az anonimizálás megkezdése.

## Nyelvtechnológiai kutatócsoport:

- A hivatkozási adatbázis feltöltése. Forrás: OTKA pályázat.
- A nyelvi portál próbaüzembe helyezése, a szótárak beépítése és a webfelület elkészítése. Saját forrás.
- Megkezdeni a koordinátori és nyelvtechnológiai munkálatokat a CESAR (Közép- és kelet-európai erőforrások) pályázat keretén belül. A Magyar Nemzeti Szövegtár új, bővített változatának megtervezése, a szükséges nyelvtechnológiai eszközök számba vétele illetve kifejlesztése. Forrás: ICT-PSP pályázat.
- Az AMICUS projekt koordinálása, a projekt második munkaszakaszának megvalósítása, mely magában foglalja a soron következő workshop szervezését is.
- Sikeresen együttműködni a CLARIN projektumban, majd lezárni annak előkészítő szakaszát.
- Sikeresen együttműködni az ECOKNOBS projektben, sikeresen beadni a pályázatot.
- Korpusznyelvészeti, számítógépes nyelvészeti kutatások: szemantikai elemző technológiák kifejlesztése kulturális örökségi szövegekre, az ehhez tartozó erőforrások fejlesztése. Multimodális számítógépes dialógusrendszer alapjainak kidolgozása.
- Sikeresen együttműködni az iTranslate4.eu projektben
- Folytatni az ómagyar és elkezdni a középmagyar nyelvemlékek digitalizálását és morfológiai elemzését. A gépi normalizáló elkészült prototípusának továbbfejlesztése, felhasználása a gyakorlatban. Forrás: OTKA.
- Folytatni a szervező munkát a Mindentudás Egyeteme 2.0 nyelvtechnológia témakörében; az előadás, a kerekasztal-beszélgetés és a magazinműsor előkészítése. Forrás: külső támogatás.
- A Nyelv- és Beszédtechnológiai Platform által megkezdett szervező és kapcsolatépítő munka folytatása; a szakterület eredményeinek terjesztése és érdekeinek képviselése a döntéshozók, az ipari szféra és a nagyközönség felé. Saját forrás.
- 2011-ben folytatódik az EFNILEX projekt, melynek célja a módszer kiterjesztése többszavas kifejezésekre valamint a szinonimák automatikus detekciója által javítani a szótár lefedettségét.
- Eredményesen közreműködni "Az Ember-gép kommunikáció technológiájának elméleti alapjai" című TÁMOP projektum lezárásában. Forrás: TÁMOP.

- Az Uralonet - Uráli Etimológiai Adatbázis hálózati alkalmazás elkészítése.

### **Nyelvművelő és Nyelvi Tanácsadó kutatócsoport:**

1. A nyelvművelés, nyelvi normativitás, helyesírás, stilsztika, utónévas egyes kérdéseinek további kutatása. Az eredmények közzététele konferenciákon, folyóiratokban, kötetekben.
2. A nyelvi tanácsadói tevékenységi körök folytatólagos technológiai megújítása
  - részvétel a helyesírási közönségszolgálat új webfelületének kialakításában
  - a nyelvi szakvéleményezéseknek mint szolgáltatásoknak a rendezése
3. Az utónév-szakvéleményezés, a névkutatás, a nevek gondozása területén
  - A honlapon a névadással kapcsolatos információk folyamatos bővítése, javítása
  - Az adható nevek szócikkszerű leírása, adatainak számítógépes rögzítése
  - Az internetes névkereső program elindítása
  - Az adható nevek kiadványszerű internetes megjelentetése
  - A névas trendjeinek, jellemző motívumainak további kutatása
  - A névasi szabályozás kutatása
4. Hazai és nemzetközi konferencián való részvétel
5. Nyelvi, nyelvtudományi ismeretterjesztés
6. Kapcsolattartás a médiával közérdekű kérdésekben

Forrás: központi támogatás.