

Fókuszterület: Nyelvtechnológia és bioetika

Kutatócsoport: Az integrált nyelvi adatforrásokból származó evidencia feldolgozásának tudomány-módszertani alapjai

Kutatócsoport-vezető: Kertész András

1. A kutatási előzmények

Az elméleti nyelvészet jelenlegi tudomány-módszertani vitáinak középpontjában az a kérdés áll, hogy a nyelvészet különböző hagyományait követő elméletek mit tekintenek *adatnak* és *evidenciának*, és milyen szerepet tulajdonítanak nekik a nyelvészeti elméletalkotásban. A probléma egy lehetséges megoldásaként az MTA-DE-PTE-SZTE Elméleti Nyelvészeti Kutatócsoportja kidolgozta a nyelvészeti adatfeldolgozás ú.n. *p-modelljét*.

A p-modell szerint nyelvészeti ciklikus és prizmatikus, retrospektív újraértékelésre épülő plauzibilis érvelési folyamat. A nyelvészeti adatok és evidencia tulajdonságai, szerkezete és funkciói ezen folyamat részeiként értelmezhetők, és annak sajátosságaitól függnnek:

- (a) A nyelvészeti adatok direkt forrásból származó, plauzibilitási értékkel rendelkező kijelentések. A nyelvészeti evidencia olyan adat, amely egy indirekt forrás részeként hozzájárul inkonzisztenciát okozó hipotézisek plauzibilitásának megítéléséhez és egymással való összevetéséhez.
- (b) Az adatoknak egyik funkciója az, hogy plauzibilitási értékeket táplálnak az elméletbe, amelyek plauzibilis következtetések segítségével további kijelentések plauzibilitásának megítélésére alkalmazhatók. Másik funkciójuk abban áll, hogy az információs túldetermináltság megszüntetésének eszközei lehetnek.

2. A kutatás célja

A jelen pályázatban tervezett kutatás a p-modellt kívánja alkalmazni azoknak az adatfeldolgozással összefüggő tudomány-módszertani problémáknak a megoldására, amelyeket a fókuszterület magyar nyelvészeti munkacsoportjainak kutatási témái felvetnek. Ily módon hozzá kíván járulni ezen kutatócsoportok munkájának hatékonyságához.

A nyelvészeti elméletalkotás egyik sajátossága az utóbbi évtizedekben az volt, hogy a nyelvészek a hipotézisek alátámasztásához felhasznált adataikat egy-egy adattípusra korlátozzák. Például a generatív nyelvészet különböző változatai elsősorban introspektív adatokat alkalmaznak, a konverzáció-elemzés hangfelvételeket tekint adatnak, a korpusznyelvészet digitalizált és annotált szövegekkel dolgozik, a kísérleti nyelvészet értelemszerűen kísérleti adatokkal. Ezzel a hagyománnyal szemben a legfontosabb új felismerések egyike *a nyelvi adatforrások integrációjának* szükségessége. Ezen azt értjük, hogy a nyelvészeti hipotézisek plauzibilitása növelhető, ha több, különböző forrásból származó adattípus kombinációjával támasztjuk alá őket. Ugyanakkor az, hogy milyen adatforrások milyen feltételek mellett milyen módon kombinálhatók egymással és a kombináció hogyan befolyásolja a hipotézisek plauzibilitását, olyan kérdés, amely esetről-esetre másképp válaszolandó meg. Ennek megfelelően a kutatás célja az, hogy a p-modell segítségével esettanulmányokon vizsgálja meg az integrált nyelvi adatforrásokból származó evidencia szerepét a nyelvészeti hipotézisek alátámasztásában. Ily módon a tervezett kutatás egységes nyelv- és tudományelméleti keretben megalapozza a fókuszterületen belüli nyelvészeti kutatásokat, majd értékeli az utóbbiak eredményeit, és visszacsatolja azokat az elméleti alapokhoz.

3. A kutatás kivitelezésének terve

A p-modell segítségével a következő vizsgálatokat tervezzük:

- (a) A p-modell alkalmazásainak bővítése az adatforrások integrálásának figyelembe vételével.
- (b) A generatív grammatikai vizsgálatban az adatok megbízhatóságának fontos kritériuma annak elemzése, hogy az introspektív, a kísérleti és a korpuszadatok milyen feltételek közepette és milyen mértékben kombinálhatók egymással annak érdekében, hogy az általuk alátámasztott hipotézisek plauzibilitása növelhető legyen.
- (c) A konverzáció-analitikai kutatás tudomány-módszertani alapkérdése az, hogy milyen szerepet játszik az anyanyelvi beszélő nyelvi intuícója a hangfelvételek kiértékelésében.
- (d) A p-modell tartalmaz olyan mechanizmust, amely képes plauzibilitási értékeket rendelni a nyelvtörténeti adatokhoz és ennek következtében hozzájárulni a nyelvtörténeti hipotézisek értékeléséhez.
- (e) A nyelvtudományban a kísérletek alkalmazása egészen új kezdeményezés, ezért tudományelméleti alapjuk kidolgozatlan. Vállaljuk a nyelvészeti kísérletek tudományelméleti alapjainak feltárását.
- (f) A kognitív nyelvészet adatkezelésének modellálása az egyik legaktuálisabb feladat.

5. *Várható eredmények*

Az eredményeknek két, egymást kiegészítő típusa várható:

- (a) Általánosítható következtetések levonása az integrált nyelvi adatforrásokból származó evidencia feldolgozásának tudomány-módszertani alapjaira vonatkozóan a p-modell segítségével.
- (b) Ezen általános felismerések közvetlenül kapcsolhatók a nyelvészeti kutatásokhoz, mivel hozzájárulhatnak a hatékonyabb adatkezeléshez.

Az eredményeket elsősorban angol és német nyelvű monográfiákban kívánjuk publikálni nemzetközi kiadóknál. Ezen túlmenően számos – magyar, angol és német nyelvű – tanulmányt készítünk.