

## Adattípusok integrációja az elméleti nyelvészetben

### 1. Problémafelvetés

Jelenleg a nyelvtudomány nemzetközi élvonalában éles hangú vita zajlik, melynek központi kérdése az, hogy milyen adattípusok alkalmazhatók és milyen módon használhatók fel evidenciaként a nyelvészeti elméletekben. Ez a *nyelvtudományi alap kutatások egyik leglényegesebb problémája*. Könnyű belátni ugyanis, hogy a nyelvészeti elméletek működőképességének megítélése jelentős mértékben azon múlik, hogy mi fogadható el a hipotéziseiket alátámasztó vagy cáfoló adattípusként, és mi nem. A vitában a *válsághangulat* jelei tapasztalhatók, mert egyes mértékadó vélemények szerint az elméleti nyelvészet fejlődése megrekedni látszik a nyelvészeti adatok tisztázatlan státusa miatt (Featherston 2007, 2009, Schlesewsky 2009, Schütze 1996 stb.). A jelen kutatás e vitához kíván egy lényeges, ugyanakkor eddig szisztematikusan nem vizsgált aspektusból a *nemzetközi élvonal szintjén hozzászólni, és hozzájárulni a válságból kivezető út megtalálásához*.

A vita erőteljesen megkérdőjelezte *minden* – mindenekelőtt az introspekcióból, korpuszokból, kísérletekből származó – adattípus megbízhatóságát és ennek következtében azoknak a nyelvészeti hipotéziseknek az elfogadhatóságát is, amelyek ezen adattípusok valamelyikét használják evidenciaként. Az egyik nagyon fontos tanulság az a felismerés, hogy egy adott kutatás egyetlen adattípus felhasználására való korlátozása sokkal inkább gátolja, semmint serkenti az elméleti nyelvészet fejlődését. E felismerés jelentőségét jól illusztrálja, hogy az egyik vezető elméleti nyelvész, akinek egész eddigi életműve arra az előfeltevésre épült, hogy az introspektív adatok különösebb módszertani szabályok nélkül alkalmazhatók és a nyelvészeti elméletalkotás szilárd alapját jelentik, az alábbi, ezzel ellentétes következtetésre jut:

“The theoreticians must acknowledge that the practice that proved successful in the pioneering phase of the past decades, namely *introspection* and *eclectic feed-back from informants*, has reached its *limits*.” (Haider 2009: 97; kiemelés tőlünk)

Ugyanakkor hiba lenne ebből a megállapításból azt a következtetést levonni, hogy az anyanyelvi beszélők intuícióját mint adatforrást a jövőben teljesen ki kell zárni a nyelvészeti elméletalkotásból. Egyrészt ugyanis ez az utóbbi évtizedek kutatási eredményeinek elvetésével járna. Másrészt a nyelvészeti adatok szerkezetéről és funkciójáról folyó jelenlegi viták egyik nagyon fontos tanulsága az a felismerés, hogy noha az introspekció bizonytalan és megbízhatatlan adatforrás, egyéb adatforrások – mindenekelőtt kísérletek és korpuszok – nem működőképesek az anyanyelvi intuíció legalább részleges felhasználása nélkül. Például a vizsgált szintaktikai szerkezetet, szemantikai, pragmatikai vagy nyelvtörténeti jelenséget stb. tartalmazó nyelvi információk megkeresése egy korpuszban, közülük a megbízhatónak minősíthető előfordulások meghatározása, vagy egy adott kísérletben stimulusként használt megnyilatkozások kiválasztása és elemzése egyaránt szükségessé teszi a nyelvész nyelvi intuíciójának felhasználását. Az adatforrások mindegyikének megvannak a maga „gyenge pontjai”, és az általuk szolgáltatott adatok megbízhatóságát befolyásoló tényezők nem kontrollálhatók kimerítő módon. Abból, hogy egyetlen adattípus kizárólagos alkalmazása „elérkezett határaihoz”, az adatforrások és adattípusok *sokféleségének, egyenrangúságának* elismerése és a kutatómunkába való *minél szélesebb körű bevonása* következik:

„Der wohl wichtigste Punkt ist, dass wir wesentlich stärker Phänomene mit unterschiedlichen **Methoden systematisch vergleichen** und dabei *alle Datenmuster gleichberechtigt behandeln und in die Interpretation mit einbeziehen*. Die immer noch zu findende Auffassung, dass es Methoden gibt, die mehr über bestimmte sprachliche Domänen auszusagen gestatten als andere, oder die Meinung, dass es bessere und schlechtere Methoden zur Untersuchung von sprachlichen Phänomenen gibt, sind in höchstem Maße hinderlich und dem wissenschaftli-

chen Fortschritt abträglich. Nur in der *Interaktion der Methoden, ihrer Koppelung* liegt der Schlüssel für zukünftigen Erkenntnisgewinn und die Hoffnung adäquatere Modelle über Sprache zu finden [...]” (Schlesewsky 2009: 177)

Az idézet felhívja a figyelmet arra, hogy az elméleti nyelvészet jelenlegi teljesítménye nem kielégítő; ugyanakkor megfogalmazza azt a felismerést is, hogy hatékonysága javításának eszköze a különböző módszerekkel nyert, különféle adattípusok *integrációja*. Azonban hiába vetődik fel kívánalomként többféle adatforrás evidenciaként való felhasználásának a szükségessége, a szakirodalomban jelenleg nincs utalás arra, hogy *milyen módon* valósítható meg ez a kívánalom. Az elméleti nyelvészet (egyre szembetűnőbben erodálódó) tudománymetodológiája eleve nem terjed ki az adattípusok integrációjának lehetőségeit és mikéntjét meghatározó releváns összefüggésekre. Így például:

- nem tudja reprezentálni az adatforrások megbízhatóságát befolyásoló tényezők és az adatok elfogadhatósága közötti kapcsolatot;
- nem tudja reprezentálni a különböző forrásokból származó, nem tökéletesen megbízható adatok és az elmélet egy adott hipotézise közötti viszonyt;
- nem tudja eldönteni, hogy mely esetekben tolerálhatók az eltérő adatforrások alkalmazásából eredő ellentmondások és mely esetekben nem;
- nem tisztázza, melyek azok az érvelési stratégiák, amelyek alkalmazásával meghozhatók a (legalábbis részben) eltérő adatforrásokból származó adatokkal alátámasztott, alternatív hipotézisek közötti döntések;
- nem ismeri azokat az érvelési stratégiákat, amelyek alapján később bevont adatforrások fényében újraértelmezhetők a korábbi döntések;
- nem tudja explicitálni az adatforrások és -típusok differenciálásának az elméleti apparátusra gyakorolt hatását.

Következésképpen a vita válsághangulatából való kilábaláshoz olyan *új tudomány módszertanra* van szükség, amely feltárja a fenti és a hozzájuk hasonló összefüggések megragadására alkalmas *heurisztikákat*. A tervezett kutatás központi problémája ezért a következő lesz:

(P) Melyek az adattípusok integrációjának heurisztikái az elméleti nyelvészetben?

## 2. Kutatási előzmények

A tervezett projekt a kutatócsoport 2007-2011-ben végzett kutatásainak közvetlen folytatása. Kiinduló tézise a következő:

(T) *Azt az új tudomány módszertant, amely alkalmas (P) megoldására, a kutatócsoport 2007-2011-ben végzett kutatásainak eredményeként kidolgozott p-modellnek az adattípusok integrációjára való kiterjesztésével kapjuk meg.*

A p-modell a nyelvészeti adatok és evidencia problémakörének nemzetközi szinten is *első átfogó megoldási kísérlete*, mely az elméleti nyelvészet új tudomány módszertanát körvonalazza. Összefoglalása a Cambridge University Pressnél jelenik meg *Data and Evidence in Linguistics: A Plausible Argumentation Model* címmel (Kertész, Rákosi 2012; l. (D10)b)-t is).<sup>1</sup> A p-modell az alábbi *kiindulópontokat* kínálja (P) megoldásához:

---

<sup>1</sup> A p-modell jelentőségét az anonim bírálók többek között így emelik ki:

“[...] the text is in my opinion a very valuable contribution to a real problem in linguistic theorising. [...] the book at issue [...] can become an important source of reference for linguistic researchers who are in search of advice to answer questions

- A p-modell alapgondolata az, hogy a nyelvészeti elméletalkotás *ciklikus és prizmatikus, retrospektív újraértékelésre* épülő plauzibilis érvelési folyamat. A nyelvészeti adatok és evidencia tulajdonságai, szerkezete és funkciói ezen érvelési folyamat részeiként értelmezhetők.
- A p-modell az adatok elfogadhatóságának mértékét *plauzibilitási érték*ként interpretálja. A plauzibilitási értéket kezdetben az *adatforrás megbízhatósága* határozza meg: minél megbízhatóbb a forrás, annál plauzibilisabb az adat. Előfordulhat ugyanakkor, hogy egy másik forrás nem támasztja alá ezt az adatot, hanem ellene szól, vagyis implauzibilissé teszi.
- Amennyiben egy (nem tökéletesen megbízható) adat egy olyan *plauzibilis következtetés* egyik premisszája, amely pozitív plauzibilitási értéket rendel egy nyelvészeti hipotézishez, akkor ez az adat e *hipotézis melletti evidenciaként* funkcionál. Ha különböző adatfajták szolgálnak evidenciaként egy adott hipotézis mellett, akkor együttesen nagyobb plauzibilitási értéket kölcsönöznek e hipotézisnek, mint bármelyikük önmagában. Ezért arra kell törekednünk, hogy *minél többféle típusú adat integrációjával tudjuk alátámasztani az elméletünk hipotéziseit*.
- Gyakran előfordul azonban, hogy más adatokból kiindulva az adott hipotézis *elleni* evidenciához jutunk. A p-modell az eltérő adatforrások alkalmazásából eredő *ellentmondások feloldását* komplex érvelési folyamatként írja le, amely a rendelkezésünkre álló információk lehető legbővebb körét figyelembe véve kísérli meg a konfliktus minél többféle lehetséges megoldásának kidolgozását és összevetését.

### 3. A kutatás végrehajtása

(P) megoldásához két párhuzamos síkon kívánunk eljutni.

A metaelméleti keret síkján a p-modellt *továbbfejlesztjük* és *kiterjesztjük* a nyelvészeti adattípusok integrációját lehetővé tevő heurisztikák feltárására.

Az alkalmazás síkján nagy számú *esettanulmányt* dolgozunk ki elsősorban a szemantika-pragmatika metszéspontja, a formális és kognitív szemantika, a generatív szintaxis és a történeti nyelvészet területéről származó elméletekre vonatkozóan.

A tervezett kutatás (P)-t az alábbi *részproblémákra bontja*, melyeket az esettanulmányok síkján, a felsorolt témákba sorolható konkrét nyelvészeti vizsgálatok alapján kívánunk megoldani:

- (P)(a) Milyen heurisztikák segítségével azonosíthatók az adott kutatás szempontjából releváns adatforrások és határozható meg a belőlük származó adatok kezdő plauzibilitása?
- (b) Milyen adattípusok milyen feltételek mellett és milyen módon integrálhatók egymással? Melyek az egyes adattípusok integrációjának korlátai?
- (c) Hogyan befolyásolja az elmélet hipotéziseinek plauzibilitását az olyan adattípusok együttes használata, amelyek a vizsgált nyelvi jelenség más-más aspektusát ragadják meg, illetve az olyan adattípusoké, amelyek ugyanarra az aspektusra vonatkoznak?
- (d) Milyen tényezők befolyásolják az egyes inkonzisztencia-kezelési stratégiák alkalmazhatóságát?
- (e) Milyen új adattípológia válthatja fel a jelenleg eklektikusan és reflektálatlanul posztulált adattípusokat?

---

like: ‘what are my data worth’ or ‘what is the (best) evidence for my claims (if there is any)’?”

“The solution to the problems of linguistic evidence posed by the authors provides a *new interpretation of linguistic theory construction and testing of linguistic theories*. The book is *highly innovative*, and there can be *no doubt that it is original*.”

#### 4. Résztermék és ütemezés

Az eddigi kutatásainkban követett munkamódszernek megfelelően a metaelméleti keret síkja és az alkalmazás síkja között *folyamatos oda-visszacsatolás* lesz. Ennek megfelelően a kutatás a résztermék közötti állandó információcserével, párhuzamosan zajlik. Ezért a részterméket nem éves bontásban, hanem három blokkba sorolva adjuk meg:

##### I. blokk: A p-modell kiterjesztése

Ütemezés	Feladat	Személyek	Ellenőrzési pont
2012. júl. – 2016. jún.	(P) megoldása a p-modell folyamatos továbbfejlesztésével az esettanulmányokból kapott visszacsatolások alapján	Kertész, Rákosi, Bihari	Tanulmányok folyamatos publikációja nemzetközi kiadványokban. Az éves workshopokon az eredmények bemutatása

##### II. blokk: Esettanulmányok

Ütemezés	Feladat	Személyek	Ellenőrzési pont
2012. júl. – 2016. dec.	Adattípusok integrációja a pragmatikában	Nagy, Kertész, Bihari, Komlósi (l. (D10)a)-t is)	A részterméknek megfelelően esettanulmányok folyamatos publikációja nemzetközi kiadványokban. Az éves workshopokon az eredmények bemutatása. Tanulmánykötetek publikálása nemzetközi kiadóknál
	Adattípusok integrációja a kognitív szemantikában	Rákosi, Tóth, Majoros (l. (D10)a)-t is)	
	Adattípusok integrációja a formális szemantikában	Vecsey	
	Adattípusok integrációja a történeti nyelvészetben	Nagy, Nemes	
	Adattípusok integrációja a generatív nyelvészetben	Kertész, Rákosi (l. (D10)a)-t is)	

##### III. blokk: Az eredmények kiértékelése

Ütemezés	Feladat	Személyek	Ellenőrzési pont
2017. jan. – 2017. jún.	Az eredmények részletes bemutatása	Az összes résztvevő, valamint meghívott külföldi szakértők	Nemzetközi konferencia megszervezése és megrendezése
2017. jan. – 2017. jún.	Az eredmények összegzése	Kertész, Rákosi	A kutatás eredményeit összefoglaló kötet publikálása egy nemzetközi kiadónál

## 5. Várható eredmény

- Általánosítható következtetések levonása az eltérő adatforrásokból származó adattípusok nyelvészeti feldolgozásának tudomány módszertani alapjaira vonatkozóan a *p-modell* segítségével.
- Az integráció különböző lehetőségeinek és korlátainak számbavétele esettanulmányok alapján és annak elemzése, hogy e lehetőségek milyen mértékben és milyen módon befolyásolják a nyelvészeti hipotézisek plauzibilitását.
- A heurisztikák feltárása, ami közvetlen hozzájárulást jelent a nyelvészeti elméletalkotás hatékonyságának növeléséhez.

## Hivatkozott irodalom

- Featherston, S. (2007): Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33, 269-318.
- Featherston, S. (2009): A scale for measuring well-formedness: Why syntax needs boiling and freezing points. In: Featherston, Winkler (eds.), 47-73.
- Featherston, S., Winkler, S. (eds.) (2009): *The fruits of empirical linguistics. Vol. 1: Process*. Berlin & New York: de Gruyter.
- Haider, H. (2009): The Thin Line Between Facts and Fiction. In: Featherston, Winkler (eds.), 75-102.
- Kertész, A., Rákosi, Cs. (2012): *Data and Evidence in Linguistics: A Plausible Argumentation Model*. Cambridge: Cambridge University Press.
- Schlesewsky, M. (2009): Linguistische Daten aus experimentellen Umgebungen: Eine multi-experimentelle und multimodale Perspektive. *Zeitschrift für Sprachwissenschaft* 28: 169-178.
- Schütze, C.T. (1996): *The Empirical Base of Linguistics. Grammaticality Judgments and Linguistic Methodology*. Chicago/London: The University of Chicago Press.