

In: Kugler N., Laczkó K., Tátrai Sz.: *A megismerés és az értelmezés konstrukciói. Tanulmányok Tolcsvai Nagy Gábor tiszteletére*. Budapest: Tinta Kiadó, 2013, 63-74

Kertész András
Rákosi Csilla

Az adattípusok integrációjának tudomány módszertani problémái az elméleti nyelvészetben¹

Absztrakt

Az elméleti nyelvészetben jelenleg éles vita folyik arról, hogy milyen adattípusok és milyen adatforrások fogadhatók el. A vita kiváltója az elméleti nyelvészet fő áramában évtizedek óta domináns szerepet játszó, az introspekcióban mint adatforrásban gyökerező grammatikalitási ítéletek megbízhatóságába vetett bizalom megrendülése. A jelen tanulmány amellel érvel, hogy bármily problematikusak is az introspektív adatok, nem iktathatók ki az elméleti nyelvészetből. Ezért az alapkérdés nem az, hogy milyen adattípussal válthatók fel, hanem az, hogy milyen módon integrálhatók egymással különböző adatforrásokból származó adattípusok az elméleti nyelvészetben. Ugyanakkor a jelenlegi elméleti nyelvészet tudomány módszertana alkalmatlan arra, hogy e kérdésre elfogadható, a nyelvészeti kutatómunka gyakorlatában alkalmazható választ adjon. A tanulmány központi feltevése az, hogy Kertész – Rákosi (2012) p-modellje lehet egy olyan újfajta tudomány módszertan egyik lehetséges (de nem az egyedül lehetséges) kiindulópontja, amely a különböző adatforrásokból származó, különböző adattípusok integrációjának igényét a jelenleginél hatékonyabban ki tudja szolgálni.

Kulcsszavak: introspekció, adat, p-modell, plauzibilis érvelés

1. Az adatforrások és adattípusok problémája az elméleti nyelvészetben

Aligha vitatható, hogy a jelenlegi elméleti nyelvészet kialakulása jelentős változást hozott a nyelvtudományban, és egészen a legutóbbi évekig rendkívül sikeresnek látszott. Legfontosabb eredménye, hogy a felszínen megfigyelhető jelenségek mögött meghúzódó általános törvényszerűségeket feltárására törekedett, melyek magyarázatot adhatnak a megfigyelt jelenségek egyes tulajdonságaira. E törekvés megvalósításában fontos szerep jutott annak, hogy domináns adatforrásnak az introspekciót, domináns adattípusnak a grammatikalitási ítéletet tekintette.

Ugyanakkor az ezen adatforrás és ezen adattípus megbízhatóságába vetett hit az utóbbi évtizedben megrendült.² A bizalom megrendülésének oka – egyebek mellett – az, hogy az elméletek egyre nehezebben boldogultak azokkal a problémákkal, amelyek a grammatikalitási ítéletek inhomogenitásából és gradualitásából származtak. Annak a válsághangulatnak a jelentőségét, amely jelenleg az elméleti nyelvészetet az adatfogalom bizonytalansága miatt jellemzi, pontosan illusztrálja Hubert Haider diagnózisa (l. még Kertész 2011):

¹ Jelen írás az MTA-DE Elméleti Nyelvészeti Kutatócsoport és az OTKA K 77823 sz. projektje támogatásával készült.

² Schütze (1996) volt az első olyan mű, amely szisztematikusan tekintette át az introspekcióban gyökerező grammatikalitási ítéletek gyenge pontjait, és elindította azt a folyamatot, amely ezen adatforrás és adattípus megbízhatóságának megkérdőjelezéséhez vezetett. A folyamatot részletesen elemzi Kertész – Rákosi (2012).

“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 K.A./R.Cs.).

A diagnózis jelentőségét nem lehet eléggé hangsúlyozni. Ugyanis nyilvánvaló, hogy ha egy elmélet domináns adatforrásának megbízhatósága csökken, akkor csökken azon hipotézisek megbízhatósága is, amelyek ezen adatforrásra épülnek. Ha pedig a már elfogadott hipotézisek megbízhatósága csökken, akkor kicsúszik a talaj azon elméletek alól, amelyek felhasználják őket.

Az első lehetséges kiútnak nyilvánvalóan az kínálkozik, hogy az introspektív adatokat számúzzuk a nyelvészetből, és olyan új adattípusokkal váltjuk fel, amelyek az utóbbi években kerültek az elméleti nyelvészet látókörébe. Ilyenek a korpuszadatok, kísérleti adatok, statisztikai adatok, számítógépes modellezésből származó adatok, neurolingvisztikai adatok, diakrón adatok, a nyelvvelsajátítás tanulmányozásából származó adatok, beszédzavarok tanulmányozásából származó adatok, műszeres kísérletekből származó adatok stb. Haider például idézett tanulmányában azon az állásponton van, hogy az introspektív adatforrásból származó adatokat műszeres kísérletekből (mindenekelőtt EEG készülékkel végzett kísérletekből) származó adatokkal kell felváltani. Ez a megoldás azonban nem kielégítő, mert egyrészt minden adattípus problematikus, sajátos korlátai, gyenge pontjai vannak, egyik adattípus sem tökéletesen megbízható. Másrészt a nyelvi intuíció használata minden adattípus esetében elkerülhetetlenül jelen van (l. bővebben Kertész 2011, Kertész–Rákosi 2012, 13.2 fejezet). Lássunk az utóbbira két példát!

(a) A kísérleti adatok és az introspektív adatok elválaszthatatlanságát jól szemléltetik a metaforákkal kapcsolatos pszicholingvisztikai kísérletek. Például Keysar et al. (2000) egyik kísérletében a szerzők saját nyelvi intuíciójuk alapján feltételezik, hogy a *You’re infected with this disease; this relationship is on its last legs; a strong marriage; this relationship is about to flatline* kifejezések a LOVE IS A PATIENT fogalmi metafora megnyilvánulásaiaként interpretálhatók. A kísérleti adatok plauzibilitása azonban nyilvánvalóan attól függ, hogy elfogadjuk-e ezt az interpretációt.

(b) Korpusznyelvészeti példaként Anti Arppe egyik vizsgálatát említjük. Arppe (2009: 4ff.) a szemantikailag hasonló igék (near-synonyms) közötti lexikai választás modellálásakor abból a hipotézisből indult ki, hogy a nyelvészeti elemzésnek figyelembe kell vennie

- a kiválasztott igék és az őket tartalmazó igei láncok morfológiai jellemzőit;
- az igék teljes szintaktikai argumentumszerkezetét;
- az egyes argumentumtípusok szemantikai alosztályait;
- azoknak az igei láncoknak a szemantikai jellemzését, amelyekben az igék előfordulnak;
- extralingvisztikai tényezőket, mint pl. a médium.

A korpusz egy helsinki napilap kéthavi számai és két internetes vitafórum anyaga volt, melyeken Arppe a finn gondolkodást kifejező igéket vizsgálta. A korpusz elemzésének lépései a következők voltak:

- automatizált morfológiai és szintaktikai analízis a Functional Dependency Grammar egy számítógépes implementációja segítségével;
- a *gondolkodni*-lexémák és szintaktikai argumentumaik kézi validálása, korrekciója és szemantikai osztályozása pl. a WordNet lexikai adatbázis által használt 25 kezdő kategória [unique beginner] segítségével (body, natural object, quantity, shape, stb.);

- az alárendelt mellékmondatok és más, a *patiens* argumentumhelyet betöltő frázisok besorolása a hagyományos *participium*, főnévi igenév, indirekt kérdés, „hogy”-gyal bevezetett propozíciók, direkt idézetek stb. kategóriákba. Ily módon megkaphatók pl. az *agens*, *patiens*, forrás, cél, hely típusú szintaktikai argumentumok;
- további, a gondolkodással kapcsolatos lexémák környezetében gyakran előforduló szintaktikai argumentumtípusok (pl. mód, idő, időtartam, gyakoriság, mennyiség) szemantikai osztályozása egy *ad hoc*, a korábbi döntéseket folyamatosan felülvizsgáló, evidencia-vezérelt eljárás segítségével (vö. Hanks 1996):

“[...] as contextual examples accumulate, one generalizes semantic classes out of them, possibly reanalyzing the emergent classification if need be, without attempting to apply some prior theoretical model. Only in the case of manner arguments did several levels of granularity emerge at this stage in the semantic analysis.” (Arppe 2009: 6)

- az igei láncok modalitás és más releváns jellemzők szerinti osztályozása;
- az igei láncok további jellemzőinek azonosítása: polaritás (állító vs. tagadó alakok), mód, igenem, idő, szám, személy, stb.;
- frekvenciák megállapítása;
- azoknak a kontextuális változóknak a kiválasztása, amelyekre elvégezzük a statisztikai elemzést;
- statisztikai analízis, többváltozós statisztikai elemzés;
- de ez is heurisztikák alkalmazásával történik, meg kell találni a vizsgált nyelvészeti problémához legalkalmasabbakat.

A vázolt lépések többsége nem automatizálható, hanem a nyelvésznek a saját nyelvi intuíciójára támaszkodó döntését igényli. Különösen érdekes ebből a szempontból a szemantikai osztályozás, hiszen Arppe a korábbi döntéseket folyamatosan felülbíráló ciklikus folyamatként jellemzi, melynek során újabb adatok fényében újraértékeljük a korábbi kategorizálást.

A fenti két példa jól szemlélteti azt, hogy nem lehetséges az introspektív adatforrás kiiktatása az elméleti nyelvészetből. Ha viszont nem lehetséges az introspekció felváltása egy másik adatforrással, akkor második lehetséges kiútként szükségképpen több, különböző adatforrás és adattípus integrációja kínálkozik. Ezért az alapkérdés az, hogy **milyen módon** integrálhatók egymással különböző adatforrásokból származó adattípusok az elméleti nyelvészetben.

2. Az elméleti nyelvészet jelenlegi tudományódszertanának korlátai

A jelenlegi elméleti nyelvészet tudományódszertana alkalmatlan arra, hogy e kérdésre elfogadható, a nyelvészeti kutatómunka gyakorlatában alkalmazható választ adjon. Ugyanis legalább a következő, jelenlegi apparátusával megoldhatatlan problémák lépnek fel.

1. probléma. A jelenlegi elméleti nyelvészet nem tudja reprezentálni az adatforrások megbízhatóságát befolyásoló tényezők és az adatok elfogadhatósága közötti kapcsolatot. Például egy történeti korpusz megbízhatóságát befolyásolja, hogy milyen régi nyelvallapotra vonatkozik, milyen terjedelmű, milyen nyelvi regisztereket ölel fel, milyen fizikai állapotban van a kézirat, stb. A következő nyitott kérdések merülnek fel:

- Hogyan fejezzük ki annak mértékét, ahogyan az egyes – egymástól jelentősen eltérő jellegű – tényezők befolyásolják az adott forrásból származó adatok elfogadhatóságát?
- Hogyan húzzuk meg az elfogadható és a nem elfogadható adatok (adatjelöltek) közötti határvonalat?

2. probléma. Nem tudja reprezentálni a különböző forrásokból származó, nem tökéletesen megbízható adatok és az elmélet egy adott hipotézise közötti viszonyt. Például lehetséges, hogy a nyelvi intuíciónk helytelennek ítél egy mondatot, amelyet ugyanakkor megtalálunk egy korpuszban. Gyakran előfordul, hogy az egyik kísérlet olyan adatokkal szolgál, amelyek alátámasztanak egy introspektív adatok alapján felállított hipotézist, egy másik kísérlet viszont cáfolja. Nyitott kérdések:

- Hogyan írható le az adatok elfogadhatósága és a hipotézisek közötti viszony, a) ha a különböző adattípusok egyaránt alátámasztják a hipotézist (mennyire támasztják alá együttesen?); b) ha egységesen cáfolják (milyen mértékben cáfolják együtt?); c) ha nem egységesen ítélik meg?

3. probléma. Nem tudja eldönteni, hogy mely esetekben tolerálhatók az eltérő adatforrások alkalmazásából eredő ellentmondások, és mely esetekben nem. Sokféle és egymással összeegyeztethetetlen nézet van jelen a nyelvészetben azzal kapcsolatban, hogy hogyan bánjunk az ellentmondásokkal, ellenpéldákkal, kivételekkel:

- Erős falszifikacionizmus.* Az ellentmondások nem tolerálhatók, azonnal el kell vetni az elméletet/hipotézist, mert az elmélet logikai káoszba torkollik (Popper 1959).
- Gyenge falszifikacionizmus Chomsky-féle változata.* Az ellentmondások hibák, azonban nem végzetesek, és nem igényelnek azonnali megoldást. Chomsky szerint gyakran akkor járunk el helyesen, ha feljegyezzük és félretesszük őket mindaddig, amíg az elmélet meg nem tudja magyarázni őket (Chomsky 1980; l. még Kertész 2011). Az inkonzisztencia időleges tolerálása lehetővé teszi, hogy egy-egy ellenpélda ne akassza meg az elméletalkotási folyamatot. Bizonyos esetekben racionális lehet így eljárni, amennyiben remény van rá, hogy az elmélet fejlődésének egy későbbi fázisában elérhetővé válnak majd olyan eszközök, amelyek lehetővé teszik az ellentmondás feloldását.
- Gyenge falszifikacionizmus egy másik változata* (pl. Penke–Rosenbach 2004: 484, Kepser–Reis 2005: 3). Az ellentmondások nem hibák. Nem félretenni kell őket, hanem az elméletalkotási folyamat motorjaiként, inspirálóiként, motiváló erőként kell rájuk tekinteni. Azaz azonnali reakciót kívánnak, de szerepük egyértelműen pozitív: megmutatják, hogy merre kell haladnia az elméletalkotási folyamatnak.
- Gyenge falszifikacionizmus statisztikus változata.* Ha a nyelv szabályait nem szigorú előírásokként, hanem statisztikai tendenciákként értelmezzük, akkor egyetlen ellenpélda vagy statisztikailag ritka előfordulás nem elegendő ahhoz, hogy falszifikáljanak egy hipotézist, csupán a statisztikailag szignifikáns mértékű előfordulásuk esetén kell elvetnünk az érintett hipotézist (Penke–Rosenbach 2004: 486).
- A Moravcsik (2011) által leírt stratégiák* három nagy csoportra oszthatók:
 - *Szeparálás.* Az ellentmondó állításokat elkülönítjük pl. kettős ágrajz, a mély- és felszíni struktúra megkülönböztetése, különböző reprezentációs szintek vagy eltérő nyelvi modulokhoz való hozzárendelésük segítségével.
 - *Regularizálás.* Az ellentmondó állításokat újraelemezzük oly módon, hogy a kivételes esetek nem kivételesek többé, vagy legalábbis jól körülírható, egységes alosztályt alkotnak. Ez történhet például úgy, hogy mind a szabályos, mind a szabálytalan

alosztályokat újraelemezzük, hogy egységes osztályt alkossanak. De eljárhatunk úgy is, hogy vagy kvantitatív, vagy kvalitatív szempontból megerősítjük a kivételeket tartalmazó alosztályt. Előbbi azt jelenti, hogy további hasonló eseteket találunk, és megmutatjuk, hogy nem elszigetelt, sporadikus kivételekről van szó, utóbbi pedig azt, hogy a kivételesen viselkedő elemek újabb közös tulajdonságait azonosítjuk. Ezzel lehetővé válhat, hogy egységes magyarázatot találjunk a viselkedésükre.

- *Magyarázat.* A regularizálást kiegészítjük a magyarázattal. Nem csupán közös tulajdonságokat azonosítunk a kivételeknek tekintett eseteknél, hanem olyan okokat találunk, amelyek segítségével magyarázatot tudunk kínálni a viselkedésükre.
- f) *Diakrón magyarázat.* A nem szabályosan viselkedő nyelvi elemről kimutatjuk, hogy korábbi nyelvállapot maradványa. Például Nübling (2011) a német igék egyes inflexiós szabálytalanságait nyelvtörténeti okokra vezeti vissza, és négyféle morfológiai változás eredményének tekinti őket.
- g) *Performancia alapú magyarázatok.* Newmeyer (2011) szerint az univerzális grammatika nem alkalmas sem a tipológiai szabályszerűségekre, sem a kivételeik magyarázatára. Ezért performancia alapú magyarázatokat kell keresnünk, mivel az univerzális nyelvtan nem engedi meg a kivételek jelenlétét, míg a performancia-elvek igen: a szabályok megszegése – legalábbis az egyének szintjén – utóbbiak esetében nem cáfolja a szabályokat.
- h) *Optimalitáselmélet.* A szabályok helyett megsérthető korlátokat tételezünk. Ekkor a szabálysértések nem kivételeknek minősülnek, hanem a nyelv működésének természetes velejáróiként interpretálhatók. Az optimalitáselmélet szerint minden kimeneti forma megsért legalább egy korlátot. A korlátokkal való konfliktusok száma és súlyossága azonban döntő tényező: az egyazon bemeneti formához tartozó kimeneti forma-jelöltek közül az a legjobb, amelyik a legkisebb rangú korlátot sérti meg.
- i) *Grammatikalitási skálák.* Többfokozatú skálákat alkalmazunk a grammatikus/nem-grammatikus vagy helyes/helytelen dichotómiák helyett. Featherston (2011) azzal indokolja a használatukat, hogy a szintaktikai szabályok mellett egy sor egyéb korlát hatása is összeadódik, melyek elfedik a generatív grammatika szabályainak működését.

Látva e stratégiák sokféleségét a szakirodalomban, a következő nyitott kérdések merülnek fel:

- Melyek az egyes stratégiák alkalmazásának kritériumai?
- Ha több stratégia is alkalmazható bizonyos körülmények között, hogyan járjunk el?
- Hogyan alakítható ki ezekből a stratégiákból egy egységes módszertani eszköz?

4. probléma. A jelenlegi elméleti nyelvészet nem tisztázza, melyek azok a stratégiák, amelyek alkalmazásával meghozhatók a (legalábbis részben) eltérő adatforrásokból származó adatokkal alátámasztott, alternatív hipotézisek közötti döntések. Például gyakran úgy döntenek két rivális hipotézis között, hogy valamilyen új adat(típus) bevonásával azt választják, amelyiket ezek az adatok alátámasztanak, a másik hipotézist pedig elvetik. A nyitott kérdések:

- Ha egyszerűen feladjuk az egyik hipotézist, noha adatok egy körét jól írta le, akkor információt veszítünk. Hogyan akadályozható meg vagy legalábbis minimalizálható az információvesztés?
- Hogyan alkossunk olyan hipotézisrendszert, amely minden rendelkezésünkre álló adatot képes leírni?

5. probléma. nem ismeri azokat a stratégiákat, amelyek alapján később bevont adatforrások fényében újraértékelhetők a korábbi döntések. Például tipikus problémahelyzet a nyelvészeti elméletalkotás során, hogy bizonyos adatok nem támasztják alá egy sejtésünket, ezért egy rivális hipotézissel próbálkozunk helyette; ezt azonban egy másik (esetleg megbízhatóbbnak ítélt) adattípusból származók cáfolják, az elvetett hipotézist viszont nem. Az analitikus tudományelmélet standard nézete nem nyújt támpontokat az ilyen helyzetek kezeléséhez, mivel vagy eleve a „felfedezés kontextusához” sorolja őket, vagy csupán annyit mond, hogy egy hipotézist falszifikálni kell tekinteni, ha bármilyen adat ellentmond neki.³ Pedig lehetséges, hogy a felhasznált adatok forrásának megbízhatóságát érdemes felülvizsgálni, vagy az is előfordulhat, hogy mindkét hipotézist érdemes megőrizni és kombinálni egymással. A megoldandó kérdések a következők:

- Mikor érdemes visszatérni egy korábban már elvetett hipotézishez?
- Általánosítva: Elméletileg korábbi döntéseink bármelyike felülvizsgálható; hogyan döntsük el, hogy melyikük revidálásával érdemes próbálkozunk?

6. probléma. Nem tudja explikálni az adatforrások és -típusok differenciálásának az elméleti apparátusra gyakorolt hatását. A szakirodalom azt mutatja, hogy egyre több olyan adattípust alkalmaznak a nyelvészek, amelyeket az adott elmélet módszertana korábban nem ismert (el). Schütze (1996) például a következő adattípusokat javasolja az introspektív adatok felváltására:

- a kísérleti alanyoktól magyarázat kérése az ítéleteikről;
- megkérdezni a kísérleti alanyokat, hogy milyen feltételek mellett lehet az adott mondat grammatikus (ha az lehet egyáltalán);
- a mondatok sorrendbe rendeztetése az alanyokkal aszerint, hogy mennyire tartják őket grammatikusnak;
- a hibásnak ítélt mondatok összehasonlításabból a szempontból, hogy milyen típusú hiba van bennük;
- kétértelmű mondatokhoz parafrázisok kérése;
- elektrofiziológiai adatok (pl. az eseményfüggő potenciálok mérése, EKP).

Világos, hogy egyik adattípus sem illeszthető be az elméletalkotási folyamatba anélkül, hogy a generatív grammatika fő téziseihez és háttérfeltevéseihez való hozzáilleszhetőségét ellenőriztük vagy megteremtettük. Problémát jelent, hogy nem vagy csupán részlegesen történt meg azoknak az elméleti feltevéseknek az azonosítása és vizsgálata, amelyek alapján az új adattípusok alkalmazhatók.

3. Az elméleti nyelvészet egy új tudomány módszertana felé

Mivel az elméleti nyelvészet tudomány módszertana számára mindezek a nyelvészeti elméletalkotás hatékonyságát aláásó problémák megoldhatatlanok, olyan új tudomány módszertanra van szükség, amely feltárja a megoldásukra alkalmas heurisztikákat. Noha az említett problémák a nyelvészeti elméletalkotás legkülönbözőbb aspektusait érintik, mégis szorosan összefüggnek, és átfogó megoldást kívánnak. A nyelvészeti problémamegoldás technikáinak megújítására van szükség – a meglévő gyakorlat mélyebb megértéséből kiindulva, szisztematikusan és koncepciózusan. Ennek során a cél nem az

³ Az analitikus tudományelmélet standard nézetének rövid ismertetéséhez a nyelvészeti adat-evidencia problémakör szempontjából l. Kertész–Rákosi (2012).

uniformizálás, egyetlen, a nyelvészet minden területén érvényes szabályrendszer kidolgozása, hanem egy stratégiagyűjtemény kialakítása lehet, melynek egyes elemei nem feltétlenül alkalmazhatók minden területen és minden esetben. A kérdés nem az, hogy „hogyan kell”, hanem „hogyan lehet, hogyan érdemes”.

Feltevésünk az, hogy Kertész – Rákosi (2012) p-modellje lehet egy olyan újfajta tudomány módszertan egyik lehetséges (de nem az egyedül lehetséges) kiindulópontja, amely a különböző adatforrásokból származó, különböző adattípusok integrációjának igényét a jelenleginél hatékonyabban ki tudja szolgálni. A p-modell alap gondolata az, hogy a nyelvészeti elméletalkotás ciklikus és prizmatikus, retrospektív újraértékelésre épülő plauzibilis érvelési folyamat, melynek célja a plauzibilis kijelentéseként explikált adatok, illetve az adatok és az elmélet többi hipotézise közötti ellentmondások feloldási lehetőségeinek feltárása és összevetése. A p-modell jelen formájában újfajta kiindulópontokat kínál az adatok integrációjával kapcsolatos problémákra, azonban nem nyújt kész megoldásokat. A feladat tehát: a p-modell továbbfejlesztése több irányból, sokféle perspektívából, az eredmények szisztematizálására törekedve.

4. A p-modell által kínált kiindulópontok

Noha a p-modell kiindulópontokat kínál a 2. fejezetben felsorolt nyitott problémák megoldásához, melyek kijelölik a további kutatások irányát, újabb problémákat is felvet. Az alábbiakban röviden körvonalazzuk egyrészt azt, hogy a p-modell milyen módon teszi kezelhetővé az elméleti nyelvészet jelenlegi tudomány módszertanával nem vagy nem kielégítően megoldható problémákat, másrészt azokat az új kérdésfeltevéseket, melyek tételeiből levezethetők.

Az **1. probléma** az adatforrások megbízhatóságát befolyásoló tényezők és az adatok elfogadhatósága közötti kapcsolatra vonatkozik. A p-modell kiindulópontja a megoldáshoz az, hogy az adatok elfogadhatóságának mértékét plauzibilitási értéként interpretálja. A plauzibilitási értéket kezdetben az adatforrás megbízhatósága határozza meg. Ez a kiinduló érték az elméletalkotási folyamat során dinamikusan változhat: új információk hatására növekedhet, csökkenhet, vagy akár az adatforrás bizonytalanná válásával el is tűnhet. Kertész–Rákosi (2012) részletesen leírja a plauzibilitási értékek változásának mechanizmusait. További kutatások feladata lehet ugyanakkor az alábbi kérdések megválaszolása:

- Milyen heurisztikák segítségével azonosíthatók az adott kutatás szempontjából releváns adatforrások, és hogyan határozható meg a belőlük származó adatok kezdő plauzibilitása?
- Hol és hogyan érdemes felülvizsgálni a kialakult értékeléseket az adatok megbízhatóságát befolyásoló tényezőkkel kapcsolatos újabb kutatások fényében?

A **2. probléma** a különböző forrásokból származó, nem tökéletesen megbízható adatok és az elmélet egy adott hipotézise közötti viszony reprezentálása. A lehetséges megoldás kiindulópontjaként a p-modell bevezeti az adott hipotézis melletti, illetve elleni gyenge/relatív/erős evidencia, a forrásosztály és a p-inkonzisztencia fogalmát. Ezen fogalmi háló lehetővé teszi az alábbi kérdések vizsgálatát:

- Milyen adattípusok milyen feltételek mellett és milyen módon integrálhatók egymással? Melyek az egyes adattípusok integrációjának korlátai?

- 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?

A **3. probléma** az eltérő adatforrások alkalmazásából eredő ellentmondások tolerálhatóságának kritériumaira vonatkozik. A p-modell szerint nem tolerálhatók az erős ellentmondások (egyazon p-kontextusban, szeparálatlanul hagyott ellentmondások), ideiglenesen tolerálhatók azok a gyenge ellentmondások, amelyekben az ellentmondás két tagjának szétválasztása nem jól motivált, és tartósan tolerálhatók a gyenge, jól motivált módon szétválasztott ellentmondások. Felvetődik a kérdés:

- Hogyan integrálható ez a felosztás például a Moravcsik Edit által kidolgozott rendszerrel, és más, a nyelvészeti szakirodalomban felvetett eljárásokkal?

A **4. probléma** az eltérő adatforrásokból származó adatokkal alátámasztott, alternatív hipotézisek közötti döntések meghozatalát lehetővé tevő stratégiák kidolgozása. A megoldás kiindulópontjaként 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. A p-modell szerint nem egyes hipotézisek között hozunk döntést, hanem az őket tartalmazó elméletváltozatok (p-kontextus változatok) között. A p-modellben a kontrasztív stratégia döntésorientált, azaz a megalapozott döntés lehetőségeit keresi, a kombinatív stratégia kombinációorientált, azaz az együttes fenntartás lehetőségeit keresi. Az inkonzisztencia-kezelési stratégiák alapmechanizmusainak leírásán túllépve például az alábbi kérdések megválaszolása lehet az új tudományelméleti modell feladata:

- Mikor mondhatjuk azt, hogy a döntési folyamatot lezárhatjuk, és megkaptuk az adott p-probléma feloldását?
- Azaz: mikor jobb egy megoldás, mint egy másik/a többi? Melyik szempontot tekintjük fontosabbnak, és melyeket kevésbé lényegesnek a döntéskor?
- Milyen tényezők befolyásolják az egyes inkonzisztencia-kezelési stratégiák alkalmazhatóságát?

Az **5. probléma** a korábbi döntések felülvizsgálatának stratégiáira vonatkozott. A p-modellben a 'p-kontextus' fogalma felöleli mindazon tényezőket, amelyek relevánsak lehetnek az adott nyelvészeti probléma megoldásában: a forrásokat, az adatokat, a módszertani elveket, stb. A p-modell az elméletalkotási folyamatot ciklikus és prizmatikus érvelési folyamatként interpretálja. Feltárásra várnak azonban az újraértékelési folyamat során meghozandó döntések finomabb lépései:

- A p-modell egyelőre nem tartalmaz olyan heurisztikákat, amelyek segítenének abban, hogy egy adott helyzetben a ciklikus visszatérés során a p-kontextus melyik elemének módosításával érdemes először próbálkoznunk.

A **6. probléma** az adatforrások és -típusok differenciálásának az elméleti apparátusra gyakorolt hatása. A p-modell a megoldáshoz azt a kiindulópontot nyújtja, hogy az adatokat nem fizikai vagy mentális entitásokkal azonosítja, hanem a nyelvi viselkedés/nyelvtudás/nyelv valamely aspektusát leíró plauzibilis kijelentéseként értelmezi – azaz olyan kijelentéseként, amelyeket valamilyen forrás (nyelvi intuíció, korpusz, kísérlet, stb.) alátámaszt, azonban nem tesz tökéletesen bizonyossá. Ezek a plauzibilis kijelentések

inkonzisztensek lehetnek az elmélet hipotéziseivel, háttérfeltevéseivel. Mindez felveti az alábbi kérdést, melynek megválaszolása további kutatások feladata lehet:

- Milyen új adattípológia válthatja fel a jelenleg eklektikusan és reflektálatlanul posztulált adattípusokat?

5. Egy példa az adattípusok integrációjára: a gondolat kísérletek és a valódi kísérletek viszonya

A kísérleti pragmatika megjelenésével egyre több nyelvész véli úgy, hogy a nyelvész pragmatikai intuíción alapuló ítéleteket és a gondolat kísérleteket fel kell váltani kísérleti adatokkal. Például Noveck és Sperber (2007: 185f.) szerint a kísérleti adatok egyértelműen jobbak, mint a pragmatikai intuíción alapuló adatok (a gondolat kísérletek eredményei vagy az akceptálhatósági ítéletek), mivel – az utóbbiakkal ellentétben – megbízhatóak, és lehetővé teszik, hogy rivális elméletek között döntést hozzunk. Mások – mint például Meibauer (2011) – ezzel szemben úgy vélik, hogy a kétféle adatforrás együttes alkalmazására van szükség.

A p-modell eszköztárával elemezve a gondolat kísérleteket és a valódi kísérleteket azt láthatjuk, hogy eltéréseik ellenére szoros kapcsolat van közöttük:

- A gondolat kísérletek és a valódi kísérletek szerkezete hasonló.
- Mind a valódi, mind a gondolat kísérletek plauzibilis, de nem bizonyosan igaz adatokat szolgáltatnak.
- A gondolat kísérletek eredménye gyakran szolgál valódi kísérletek kiindulópontjaként, így indirekt folytonosság van köztük. Ezért a gondolat kísérletekből származó adatok plauzibilitása közvetve befolyásolja a valódi kísérlet eredményének megbízhatóságát.
- Ugyanakkor a valódi kísérlet során módosulhat (növekedhet vagy csökkenhet) a kiindulópontul szolgáló gondolat kísérlet eredményének plauzibilitása.
- A gondolat kísérletek konstitutív részei minden valódi kísérletnek, hiszen a kísérleti design nem más, mint a kísérlet minden apró részletének és lefolyásának lehető legalaposabb elképzelése. Ez a gondolat kísérlet nem más, mint a kutató előzetes elgondolásai arról, hogy működőképes lesz-e a kísérlet. Feladata annak felmérése, hogy a kísérleti eljárás alkalmasnak ítélik-e arra, hogy megbízható eredményeket produkáljon.
- Ugyanakkor lényeges eltérések is vannak a valódi és a gondolat kísérletek között: más-más erősségeik és gyenge pontjaik vannak, részben másfajta tényezők határozzák meg a belőlük származó adatok plauzibilitását.

A felsorolt tényezők mind egy irányba mutatnak: a két adattípus együttes felhasználása, integrációja, tulajdonságaik feltárása és komplementaritásuk mértékének megállapítása nagyban növelheti a pragmatikai elméletek hipotéziseinek plauzibilitását.

6. Összegzés

Az elméleti nyelvészet zsákutcába kerülésének egyik lényeges kiváltó oka az volt, hogy nem alakult ki párbeszéd a tudományelméleti megfontolások és a gyakorlati kutatómunka között. A jelenlegi adat-evidencia vita legfontosabb hozadéka e párbeszéd szükségességének felismerése, mivel élesen vetette fel az elméleti nyelvészet több, évtizedeken át szőnyeg alá

söpört tudományelméleti alapkérdését (l. pl. Kepser–Reis (eds.)(2005); Penke–Rosenbach (eds.)(2004); Kristiansen, Achard, Dirven, de Mendoza Ibáñez (eds.)(2006); Schütze 1996, Sternefeld (ed.)(2007); Featherston–Winkler (eds.)(2007), Winkler–Featherston (eds.)(2007), Horst–Wiese (2011)). E felismerést tudatosítja és aknázza ki a p-modell. Lehetővé teszi, hogy a metaelméleti reflexió és a nyelvészeti kutatómunka gyakorlatának folyamatos egymásra vonatkoztatásával próbáljunk meg választ keresni a 4. fejezetben felvetett kérdésekre.

Irodalom

- Arppe, Antti 2009. Linguistic choices vs. Probabilities – vs. how much and what can linguistic theory explain? In: Featherston, Sam – Winkler, Susanne (eds.), 1–24.
- Chomsky, Noam 1980. On binding. *Linguistic Inquiry* 11, 1–46.
- Featherston, Sam 2011. Three types of exceptions – and all of them all rule-based. In: Simon, Horst J. – Wiese, Heike (eds.), 291–323.
- Featherston, Sam – Winkler, Susanne (eds.) 2009. *The fruits of empirical linguistics. Vol. 1: Process*. Berlin – New York: de Gruyter.
- Haider, Hubert 2009. The thin line between facts and fiction. In: Featherston, Sam – Winkler, Susanne (eds.). 75–102.
- Hanks, Patrick 1996. Contextual dependency and lexical sets. *International Journal of Corpus Linguistics* 1(1): 75–98.
- Kepser, Stephan – Reis, Marga 2005. Evidence in linguistics. In: Kepser, Stephan – Reis, Marga (eds.), 1–6.
- Kepser, Stephan – Reis, Marga (eds.) *Linguistic evidence. Empirical, theoretical and computational perspectives*. de Gruyter. Berlin–New York.
- Kertész András 2011. Az episztemológiai tolerancia elve az elméleti nyelvészetben. *Magyar Nyelv* CVII, 129–147.
- Kertész, András – Rákosi, Csilla 2012. *Data, evidence and plausible argumentation in linguistics. A plausible argumentation model*. Cambridge. University Press Cambridge.
- Keysar, Boaz, Shen, Yeshayahu, Glucksberg, Sam, Horton, William S. 2000. Conventional language: How metaphorical is it? *Journal of Memory and Language* 43, 576–593.
- Kristiansen, Gitte, Achard, Michel, Dirven, René, de Mendoza Ibáñez, Francisco J. Ruiz (eds.) 2006. *Cognitive linguistics: Current applications and future perspectives*. Berlin – New York: de Gruyter.
- Meibauer, Jörg (2012): Pragmatic evidence, context, and story design. An essay on recent developments in experimental pragmatics. In: Kertész, András, Schwarz-Friesel, Monika & Consten, Manfred (eds.) *Converging data sources in cognitive linguistics* [=Special Issue of *Language Sciences*], 768-776.
- Moravcsik, Edith 2011. Coming to grips with exceptions. In: Horst J. Simon, Heike Wiese (eds.) *Expecting the unexpected: Exceptions in grammar*. (Trends in linguistics. Studies and monographs, 216.) de Gruyter Mouton. Berlin–New York. 31–55.
- Newmeyer, Frederick J. 2011. Three approaches to exceptionality in syntactic typology. In: Simon, Horst J. – Wiese, Heike (eds.): *Expecting the unexpected: Exceptions in grammar*. De Gruyter Mouton, 255–282.
- Noveck, Ira A., Dan Sperber. 2007. The why and how of experimental pragmatics. In: Burton-Roberts, Noël (ed.) *Pragmatics*. Basingstoke: Palgrave Macmillan. 184-212.
- Nübling, Damaris 2011. How do exceptions arise? On different paths to morphological irregularity. In: Simon, Horst J. – Wiese, Heike (eds.): *Expecting the unexpected: Exceptions in grammar*. De Gruyter Mouton, 139–162.

- Penke, Martina – Rosenbach, Annette 2004. What counts as evidence in linguistics? In: Penke, Martina – Rosenbach, Annette (eds.)(2004). 480–526.
- Penke, Martina – Rosenbach, Annette (eds.) 2004. *What counts as evidence in linguistics?* [= Studies in Language 28(3)].
- Popper, Karl 1959. *The logic of scientific discovery*. Hutchinson. London.
- Schütze, Carson T. 1996. *The empirical base of linguistics. Grammaticality judgments and linguistic methodology*. The University of Chicago Press. Chicago–London.
- Simon, Horst J. – Wiese, Heike (eds.) 2011. *Expecting the unexpected: Exceptions in grammar*. De Gruyter Mouton.
- Sternefeld, Wolfgang (ed.) 2007. *Data in generative grammar* (= *Theoretical Linguistics* 33(3)).
- Winkler, Susanne – Featherston, Sam (eds.) 2009. *The fruits of empirical linguistics. Vol. 2: Product*. Berlin & New York: de Gruyter.