

Megjegyzések a nyelvészeti adatok és evidencia problémájáról folyó vita jelenlegi állásához*

„A nyelvészet jövőjét valószínűleg metodológiai kérdések fogják meghatározni [...]”
(GEERAERTS 2006: 21).

1. Bevezetés. – A) Célkitűzés. – MARTINA PENKE és ANETTE ROSENBACH 2007-ben megjelent, „What counts as evidence in linguistics?” c. kötetük előszavát a következő beszámolóval kezdik: „Jelen kötet legtöbb tanulmánya azon a workshopon hangzott el, amelyre ’Mi számít evidenciának a nyelvészetben?’

* A tanulmány az MTA-DE-PTE-SZTE Elméleti Nyelvészeti Kutatócsoport és az OTKA NI 68436 sz. projektje támogatásával készült. Köszönetet mondunk CSATÁR PÉTERnek, KELEMEN JÁNOSnak, NÉMETH T. ENIKŐnek, VECSEY ZOLTÁNNak, valamint egy anonim lektornak a kézírathoz fűzött konstruktív megjegyzéseikért. Ugyancsak köszönettel tartozunk CHRISTIAN LEHMANNnak, MARGA REISnak és ANETTE ROSENBACHnak írásunk angol nyelvű változatának kommentálásáért. Az angol nyelvű idézetek magyar fordításának az eredetivel való egybevetéséért GEORGE SEELnek mondunk köszönetet. Írásunk hibáiért a felelősség kizárólag bennünket terhel.

címmel 2003. február 26–28. között Münchenben a 'Deutsche Gesellschaft für Sprachwissenschaft' 25. éves kongresszusán került sor. Meglepetéssel tapasztaltuk azt a reakciót, amelyet a szétküldött meghívók kiváltottak. Nyilvánvaló, hogy kényes témára tapintottunk rá, és a szakma megérett a nyelvészeti evidenciának szentelt vitára" (2007: vii).

A nyelvészeti adatok és evidencia problémája tehát benne van a levegőben, ezért egyetlen szikra is elegendő ahhoz, hogy az érdeklődés fellángoljon. És valóban: jelenleg szinte hónapról hónapra jelennek meg a nyelvészeti adatok és evidencia mibenlétét firtató publikációk, kerül sor újabb és újabb konferenciák szervezésére, és csapnak össze – olykor igen élesen – az eltérő álláspontokat képviselő nyelvészeti irányzatok szószólói. Az, hogy „mi számít evidenciának a nyelvészetben”, a nyelvtudomány egyik leglényegesebb témája, mely közvetlenül és érzékenyen érinti a tudományág alapjait – és ez egyáltalán nem véletlen. Könnyű belátni, hogy az egyes – egymással versengő, de olykor együttműködésre is hajlandó, bonyolult hálót alkotó – nyelvészeti elméletek működőképességének megítélése jelentős mértékben azon múlik, hogy milyen adatok fogadhatók el a hipotéziseiket alátámasztó vagy cáfoló evidenciaként.

Írásunk célja, hogy

a) a nyelvészeti evidenciáról folyó jelenlegi vita egyik markáns, megítélésünk szerint a nyelvtudomány fejlődése szempontjából releváns és progresszív vonulatához tartozó nézeteket bemutassa és rendszerezze,

b) a rendszerezés alapján a nézetek kritikai elemzését nyújtsa, és

c) a kritikai elemzésből a nyelvészeti adatok/evidencia problémájának jövőbe-ni kezelésére vonatkozó következtetéseket vonjon le.

Nem térhetünk ki sem a jelenlegi vita több száz évre visszanyúló tudománytörténeti gyökereinek feltárására, sem a jelenlegi nyelvtudomány különböző részterületeinek empiria-problémájára. Kizárólag azt a vonulatot emeljük ki, amely közvetlenül motiválja a legújabb szakirodalom egy jól körülhatárolható részében (l. alább a D) szakaszban az a)–g) pontokat) megfogalmazott problémafelvetést.

B) Az evidencia fogalma a tudományelméletben. – A probléma megértéséhez abból a közismert tényből kell kiindulnunk, hogy a huszadik század domináns tudományszemléletének kialakulásában meghatározó szerepet játszott az analitikus tudományelmélet standard nézete. E nézet tömören így jellemezhető (l. még ehhez LAKI 1998. és 2006: 24–7 is): „A 60-as évekig az analitikus tudományfilozófia lényegében két irányzatot foglalt magában: a Bécsi Kör kezdeményezte logikai empirista pozitivista és a Popper-féle irányzatot. E két felfogás a jelentős különbségek ellenére (például, hogy az előbbi az induktív verifikációt, az utóbbi a deduktív falszifikációt tekintette a legalapvetőbb módszertani elvnek) számos fontos kérdésben azonos nézeten volt. Mindenekelőtt abban, hogy a tudomány pszichológiai, szociológiai és történeti vonatkozásával nem foglalkoztak, a tudományos megismerés elemzését a logikai analízisre korlátozták. Elvileg is megalapozott szándékuk ezen belül az volt, hogy az ismereteket csupán igazolásuk, és nem felfedezésük szempontjából vegyék szemügyre. Azokat az empirikus és logikai eljárásokat, módszertani előírásokat igyekeztek megadni, amelyekkel a tudományos és a nem tudományos ismeretek egymástól elhatárolhatók,

illetve amelyek révén az előbbieik igazságértéke (igaz vagy hamis volta) véglegesen, egyszer s mindenkorra megállapítható” (FEHÉR 2000: 236).

Az idézet többek között utal a „felfedezés kontextusa” és az „igazolás kontextusa” REICHENBACH-tól (1938.) származó megkülönböztetésére. A „felfedezés kontextusának” (azaz a tudományos elméletalkotás folyamatának, beleértve annak társadalmi, pszichológiai és történeti aspektusait is) irrelevánsá bélyegzésével és az „igazolás kontextusának” (vagyis a felfedezési folyamat végeredményeként kapott elméletek rekonstrukciójának és értékelésének) kizárólagos figyelembevételével összhangban az analitikus tudományelmélet standard nézete mindenekelőtt azt tekintette feladatának, hogy az empirikus tudományos elméleteket logikai szerkezetük rekonstrukciója révén elemezze és értékelje.

Az említett két irányzat második, jelen írás problémafelvetése szempontjából fontos (és az idézetben említett) közös kiindulópontja a tudományos és a nem-tudományos ismeretek elválasztásának szándéka volt. Ennek egyik eszközeként fogalmazták meg azt a követelményt, hogy az empirikus elméleteknek olyan kijelentésekből kell állniuk, amelyek az adatok egy – evidenciának nevezett – kitüntetett osztálya alapján tesztelhetők. Ezt az alap gondolatot például HEMPEL nagyon egyszerűen és világosan így fogalmazza meg: „[...] az empirikus tudomány minden kijelentésének tesztelhetőnek kell lennie olyan evidencia alapján, amely nyilvános, azaz amely különböző megfigyelők által rögzített, és nem függ lényegesen a megfigyelőtől” (HEMPEL 1952: 22; kiemelés: K. A. – R. Cs.). Az evidencia eszerint – első megközelítésben és nagyon leegyszerűsítve – a következő sajátosságokkal rendelkezik, melyek az analitikus tudományelmélet különböző irányzataihoz tartozó nézetek közös intuitív magját alkotják, és melyek kitüntetett szerepét motiválják:

(E) Az evidencia:

- a) objektív,
- b) alkalmas arra, hogy az alternatív hipotézisek/elméletek közötti döntés semleges alapjául szolgáljon,
- c) feladata a hipotézisek igazolása (verifikációja, konfirmációja, falszifikációja),
- d) közvetlenül adott,
- e) elsődleges az elmélethez képest, és
- f) megbízható.

Az evidencia ezen ismérveiről és az ismérvek különböző explikációiról tömör áttekintést nyújt pl. KELLY (2006.) a logikai pozitivizmustól egészen napjainkig. Az analitikus tudományelmélet standard nézetén belüli álláspontokat részletesen elemző klasszikus áttekintés STEGMÜLLER (1970.).

E sajátosságok mindegyikének értelmezése azonban erősen problematikus. Így az evidencia fogalmával kapcsolatos huszadik századi tudományelméleti viták célja a felsorolt ismérvek explikációja, valamint annak eldöntése volt, hogy az ismérvek közül melyek azok, amelyek valóban relevánsak. E tudományelméleti viták során az (E)-ben felsorolt ismérvek mindegyikét kétségbe vonták valamilyen szempontból, és mindegyiket sokféle, olykor egymástól jelentősen eltérő módon explikálták. Például RUSSELL az evidenciát érzéki adatokra vezette vissza, az analitikus tuda-

mányelmélet standard nézetének logikai pozitivista irányzata „protokollkijelentésekre”, POPPER pedig „bázisállításokról” beszél, miközben e fogalmak értelmezése is heves viták tárgya volt. (Áttekintésükhöz l. pl. KELLY 2006. és STEGMÜLLER 1970.). Bizonyos esetekben olyan javaslatok is kikristályosodtak, amelyek általánosan elfogadottá váltak; például az objektivitást interszjektív ellenőrizhetőségként értelmezték (l. POPPER 1959: 44). A viták során folyamatosan felmerültek megoldhatatlannak látszó, paradoxon-típusú problémák is (l. GOODMAN 1983. [1955.]).

A CARNAP nevével fémjelzett induktivista irányzat *verifikáció*ként értelmezte az evidencia funkcióját, amely szerint az evidencia bizonyítja az elmélet hipotéziseinek igazságát. A POPPER-féle – kritikai racionalizmusnak nevezett – hipotetikus-deduktív vonulat *falszifikáció*ként kezelte abban az értelemben, hogy az evidencia nem bizonyíthatja ugyan a hipotéziseket, azonban lehetővé teszi cáfolatukat (egy empirikus hipotézist térben és időben lejátszódó, megfigyelhető eseményeket leíró báziskijelentések falszifikálhatnak, l. POPPER 1959: 86, 103). Végül pedig a hipotézist megerősítő, valószínűbbé tevő, de igazságát nem bizonyító *konfirmáció*ként explikálták. (A konfirmációelmélet rendkívül bonyolult és vitatott, ezért nem vállalkozhatunk tárgyalására e cikkben. Tömör ismertetéséhez l. pl. HOWSON 2000: 108). E három explikáció – melyek mindegyikének számos, egymással is vitában álló változatát ismerjük – történetileg különböző időszakokban és heves tudományelméleti viták közepette alakult ki ugyan, azonban logikailag összefügg egymással. Ismét nagyon leegyszerűsítve azt mondhatjuk, hogy az analitikus tudományelmélet standard nézete szerint ha E a H hipotézis melletti evidencia, akkor E valószínűbbé teszi, hogy H igaz. Ekkor E konfirmálja H -t. Ha viszont E a H hipotézis elleni evidencia, akkor E kevésbé valószínűvé teszi H igazságát, azaz E diszkonfirmálja H -t. A verifikáció abban az értelemben a konfirmáció egyik szélső esete, hogy E bizonyítja H igazságát. A diszkonfirmáció szélső esete a falszifikáció; ekkor E bizonyítja H hamisságát (vö. ehhez KELLY 2006.).

A későbbi hivatkozások megkönnyítése érdekében az analitikus tudományelmélet standard nézetének a jelen tanulmány érvelése szempontjából releváns elemeit a következőképpen foglaljuk össze:

(ATSN)

- a) Meg kell különböztetni a „felfedezés kontextusát” az „igazolás kontextusától”. Az elméletek igazolása szempontjából csupán az utóbbinak van jelentősége.
- b) A tudományelméleti rekonstrukció célja a tudományos elméletek igazolása.
- c) A két domináns irányzat az induktivizmus és a hipotetikus-deduktivizmus.
- d) Az evidencia (E)-ben felsorolt tulajdonságai révén kitüntetett szerepet játszik a hipotézisek és az elméletek igazolásában.

C) Az elméleti nyelvészet empiriavitái. – A huszadik század közepének meghatározó nyelvtudományi irányzatai – noha gyakran különböző módon és eltérő mértékben – az analitikus tudományelmélet standard nézetéből merítették metodológiai normáikat (l. pl. RINGEN 1975., ALLAN 2003.). A generatív nyelvészet fellépése egymásba ágyazott, egymással bonyolultan összefüggő tudomány módszertani viták sorozatát váltotta ki. E viták a nyelvészeti elméletek empirikus megalapozottságával kapcsolatos problémákra irányultak. A nyelv-

vészeti elméletek empirikus megalapozottságának kérdését ugyanakkor vitáról vitára másként explikálták, más részkérdésekre redukálták, és eltérő szempontok alapján közelítették meg. Annak érdekében, hogy világossá tehessük a jelen írás probléma-felvetésének motivációját, e viták közül *n é g y r e* kell röviden utalnunk.

Az amerikai strukturalizmus fő tudomány módszertani eszköze köztudottan az indukción alapuló taxonómia volt. CHOMSKY az ötvenes-hatvanas években éles támadást intézett e módszer ellen, szembeállítva vele az (ATSN) hipotetikus-deduktív irányzatának metodológiáját (vö. pl. CHOMSKY 1969. [1957.], 1965.). Alkalmazásával azt a mintát kívánta követni, amelyet az utóbbi a természettudományos elméletek számára előírt: „A nyelvészet metaelméletében két makroparadigma uralkodott: a fenomenológiai induktívizmus és a hipotetikus-deduktív módszer. [...] A huszadik század közepén az amerikai nyelvészetben áttevődött a hangsúly a fenomenológiai induktívizmus paradigmájáról a hipotetikus-deduktív paradigmára. Ez a hangsúlyeltolódás 'chomskyánus forradalomként' ismert, mivel CHOMSKY művének hatására következett be” (ALLAN 2003: 538). Ily módon az ötvenes-hatvanas években elkezdődött egy olyan vita, amelynek középpontjában – nagyon leegyszerűsítve – az a tudomány módszertani kérdés állt, hogy az empiricitás igényével fellépő nyelvtudománynak elsősorban *i n d u k t í v* általánosításokra vagy a hipotézisek *d e d u k t í v* ellenőrzésére kell-e épülnie.

A hetvenes években új szempontok bevonásával vált éles hangvételű összecsapások tárgyává az elméleti nyelvészet empirikus megalapozottságának problémája. Ekkor a vita súlypontja az (ATSN)-n belüli irányzatok közötti rivalizálásról a hermeneutika (l. mindenekelőtt ITKONEN 1978.) és az (ATSN) szembenállására helyeződött át. Az analitikus tudományelmélet standard nézetének hipotetikus-deduktív vonulatát képviselő nyelvészek a generatív grammatikát a természettudományokhoz hasonló, magyarázó és empirikus elméletnek tekintették, míg a hermeneutika hívei nem-empirikus, normatív, értelmező tudománynak (vö. pl. ALLAN 2003., WUNDERLICH ed. 1976., PERRY ed. 1980. stb.). A vita a hetvenes évek végére terméketlenségbe torkollott, mivel a résztvevők mereven és mechanikusan alkalmazott absztrakt tudományfilozófiai elvekre hivatkozva védtek álláspontjukat, miközben figyelmen kívül hagyták mind a gyakorló nyelvész számára releváns szempontokat, mind az ellentábor érveit (vö. KERTÉSZ 1991., 2004., REIS et al. 1998., PENKE–ROSENBAACH 2004. stb.).

A hetvenes évek közepétől – e vita fokozatos elhalásával párhuzamosan – több, egymástól független, de egy irányba ható nyelvtudomány-történeti fejlemény következtében az empiriavita újabb fázisa bontakozott ki. A kérdés immár nem az volt, hogy a hermeneutika vagy az analitikus tudományelmélet standard nézetének hipotetikus-deduktív irányzata alkalmasabb-e a nyelvészeti elméletek szerkezetének megragadására. Az érdeklődés homlokterébe ismét az induktív és a deduktív metodológiának az ötvenes-hatvanas években már exponált szembeállítása került, de más nézőpontból (vö. pl. MCENERY–WILSON 2001: 1–20, LEMNITZER–ZINSMEISTER 2006: 14–39). Az egyik – de korántsem az egyetlen –, a generatív nyelvészet introspektív adatfogalmával szembeforduló, jelentős tendenciát a számítógépes módszerek megjelenése és robbanásszerű elterjedése váltotta ki, mely meghatározó módon és mértékben járult hozzá a jelenlegi korpusznyelvészet kialakulásához. A „Brown

Corpus of Standard American English”, melyet W. N. FRANCIS és H. KUČERA hozott létre a 60-as évek legelején, fontos szerepet játszott e fordulatban. Hasonlóan nagy jelentőséggel bírt a QUIRK nevéhez fűződő SEU-projekt („Survey of English Usage”), valamint a J. SINCLAIR és társai irányításával összeállított, neo-firthiánus COBUILD-projekt a Birminghami Egyetemen. A generatív nyelvészet és a korpusnyelvészet képviselői között antagonisztikus ellentét alakult ki a nyolcvanas évektől, melynek forrása a relevánsnak tekintett adatok gyökeresen eltérő meghatározása volt (vö. LEMNITZER–ZINSMEISTER 2006: 32). A generatív nyelvészek triviálisnak és érdektelennek tekintették a nyelvhasználatból származó, nagy mennyiségű adat gyűjtését és elemzését. CHOMSKY élesen elutasította az indukció jelentőségét az elméletalkotás folyamatának minden pontján: sem az elméletek „felfedezésében”, azaz adatgyűjtési technikaként nem tulajdonított neki kiemelt szerepet, sem a kész elméletek „értékelésében” nem tartotta relevánsnak. CHOMSKY műveiben számtalan erre vonatkozó megfogalmazás található, melyek közül illusztrációképpen a következőket említjük: „A nyelvtudás, mint a legtöbb fontos és érdekes tény, nem figyelhető meg közvetlenül, és nem nyerhető ki adatokból semmilyen ismert induktív eljárás segítségével” (1965: 18); „Az adatgyűjtés informális módon történik; (a fonetikán kívüli területeken) nem sok hasznát vesszük a kísérleti módszereknek, vagy a komplex adatgyűjtési és -elemzési technikáknak, melyek könnyen megtervezhetők, és melyeket széles körben használnak a magatartástudományok. [...] a jelenleg legkritikusabbnak tűnő elméleti problémák szempontjából semmilyen nehézséget sem okoz, hogy releváns adatok tömkelegéhez jussunk ilyen technikák alkalmazása nélkül is” (1969: 56). A korpuszadatokat preferálók ezzel szemben az indukció nélkülözhetetlensége és az induktív adatok kezelésére vonatkozó szigorú módszertani szabályok követésének szükségessége mellett érveltek, és nem tekintették empirikusnak a nyelvész introspekcióján alapuló adatokat. Például SAMPSON már a hetvenes években így írt: „Ha a nyelvészet valóban az intuíción alapulna, akkor nem lenne tudomány” (1975: 60).

Ily módon az adat/evidencia-probléma megítélése szempontjából két tábor jött létre: egyfelől a korpusnyelvészeké, másfelől az introspektív adatok mellett érvelő generativistáké. CHOMSKY ma is differenciálatlanul jelenti ki, hogy „a korpusnyelvészet semmit sem jelent” (ANDOR 2004: 95). A „Corpus Linguistics and Linguistic Theory” c. folyóirat 2007-es kötetében (STEFANOWITSCH–GRIES eds. 2007.) közreadott vitában pedig SAMPSON hasonlóan differenciálatlanul nyilatkoztatja ki, hogy „a nyelvészet mindaddig nem fejlődik egészségesen, amíg a generatív nyelvészetet áltudományként félre nem söprik, maradék követői vissza nem vonulnak a tudományos életből, és a tudományág új képviselői el nem vetik éthoszát, feltevéseit és állítólagos eredményeit” (2007: 122).

Ugyanakkor e két tábor nézetei, bármily eltérők is, a következő közös elemekkel rendelkeznek, melyek a nyelvészeti adatok standard felfogását körvonalazzák:

(SF)

- a) Nem minden nyelvészeti elmélet tekinthető empirikusnak, hanem csupán azok, amelyek a megfelelő adattípus t használják. Csupán bizonyos adattípusok tekinthetők relevánsnak és legitimnek:

a generativisták szerint az introspektív adatok, a korpusznyelvészek szerint pedig elsősorban a korpuszadatok.

- b) Kizárólag az adatok eredete dönti el, mi számít megfelelő adattípusnak.
- c) Saját adatkezelési eljárását mindkét tábor problémamentesnek ítéli. Elegendőnek tartják, hogy az adatok megbízhatóságát néhány általános tudomány-metodológiai szabály alapján ellenőrizzék.
- d) Az adatok és az elmélet hipotézisei közötti kapcsolatot mindkét tábor szigorú, általános módszertani szabályok által rögzítettnek és egyirányúnak tételezte: vagy az adatoktól a hipotézisekhez vezető indukcióként, vagy a hipotézisektől az adatokhoz vezető dedukcióként írta le.
- e) Az evidenciát olyan empirikus adatnak tekintik, amely kitüntetett szerepet játszik. Kitüntetett szerepe az (E)-ben felsorolt ismérvekben rejlik. Mindenekelőtt abban, hogy közvetlenül (azaz bármilyen elméletre való hivatkozás nélkül) ellenőrizhető, valamilyen, adottnak feltételezett értelemben megbízhatónak minősített, és ezért felülvizsgálhatatlannak tartott tényként alkalmas az elméletek hipotéziseinek igazolására (alátámasztására vagy elvetésére).

A 80-as és 90-es években egyik tábor pozíciója sem ingott meg. Úgy látszott, a két ellentétes felfogás a másik kritikáját figyelmen kívül hagyva, párbeszédre és bármiféle együttműködésre képtelen módon él egymás mellett. Ez az empíriavita újbóli terméketlenné válásához vagy teljes elhalásához vezethetett volna, azonban ennek pontosan az ellenkezője történt: az adatokkal és az evidenciával kapcsolatos tudomány-metodológiai viták a 90-es évek második felétől kezdve fokozatosan kiszélesedtek, és új elemekkel gazdagodtak. Jelenleg már kiterjednek a nyelvészet számos különböző kutatási területére és irányzatára, és nem csupán az adatok eredetére, hanem struktúrájára, funkciójára, elfogadhatóságuk kritériumaira vonatkozó, mindeddig reflektálatlanul hagyott kérdések is előtérbe kerültek. Az ezredfordulón ugyanis olyan művek láttak napvilágot, amelyek ismét megváltoztatták a probléma fókuszát. Míg a hetvenes években a hermeneutika és az analitikus tudományelmélet hipotetikus-deduktív vonulatának ütközése állt a középpontban, mely a nyolcvanas-kilencvenes években az introspektív és a korpuszadatok dichotómiájára redukálódott, jelenleg az érdeklődés homlokterébe az a kérdés került, hogy a nyelvészet különböző részterületein milyen adattípusok használhatók, mely adatok tekinthetők evidenciának, és milyen funkciót tulajdoníthatunk az utóbbiaknak a nyelvészeti elméletekben. A probléma fókuszának áthelyezését jól jellemzi például a következő idézet: „Míg harminc évvel ezelőtt a nyelvészek még mindig arról vitatkoztak, hogy a nyelvészetnek ’empirikus tudománynak’ kell-e lennie [...], addig ma nyugodtan kijelenthetjük, hogy ez a kérdés lényegileg eldőlt, és hogy manapság a legtöbb nyelvész valószínűleg egyetért abban, hogy a nyelvészet valóban empirikus tudomány. Amiről a vita folyik, nem az, hogy vajon szabad-e vagy kell-e empirikus evidenciát alkalmazni, hanem sokkal inkább az, hogy milyen típusú evidenciát és hogyan alkalmazunk” (PENKE–ROSENBACH 2004: 480).

D) Problémafelvetés. – Ez az említett négy vita közül az utolsó, jelenleg folyó vita alapkérdése. A következőképpen rögzítjük:

(P1)

- a) Milyen adatok/adattípusok használhatók,
- b) mely adatok tekinthetők evidenciának, és
- c) milyen funkció tulajdonítható az utóbbiaknak a nyelvészeti elméletekben?

A szerzők egy része nem tekinti relevánsnak az empiricitás fogalmának explikációját és az azzal összefüggő terminológiai vitákat (PENKE–ROSENBACH 2004., KEPSEK–REIS 2005.). A fogalom széles és laza értelmezését előfeltételezik: „Amennyiben az ’empirikus’-t nagyon tág értelemben ’adatokat használó’-ként értelmezzük, akkor mindenfajta adatorientált kutatást beleérthetünk. Ebben az értelemben valószínűleg minden nyelvész empirikusan dolgozik – hiszen hogyan művelhetne bárki is nyelvészetet bármiféle nyelvi adat (azaz a nyelv valamely darabja) nélkül?” (PENKE–ROSENBACH 2004: 514–5). Mások (LEHMANN 2004., GEERAERTS 2006.) viszont relevánsnak tekintik, és az analitikus tudományelmélet standard nézetének induktivista vagy deduktivista irányzata szerinti explikációját fogadják el.

Az új problémafelvetést mindenekelőtt a következő publikációk fémjelzik:

a) SCHÜTZE (1996.) az az alapmű, amely az ezredfordulót közvetlenül megelőző években elindította az adatok szerkezetéről és funkciójáról való gondolkodást az elméleti nyelvészetben.

b) A „Studies in Language” c. folyóirat a „Deutsche Gesellschaft für Sprachwissenschaft” 2003. évi kongresszusának anyagából tematikus számot publikált, melynek tanulmányai igyekeznek szisztematikus, véleményeket ütköztető áttekintést nyújtani az adatok/evidencia problémájának sajátosságairól a nyelvtudomány több, eltérő részterületén (PENKE–ROSENBACH eds. 2004.). A tematikus szám anyaga tanulmánykötetként is megjelent (l. PENKE–ROSENBACH eds. 2007.). Írásunkban az első közlésre hivatkozunk.

c) A „Lingua” c. folyóirat hasonló tematikus számot közölt az adatok problémájáról az elméleti nyelvészetben (BORSLEY ed. 2005.).

d) A Tübingeni Egyetemen 1999 óta folyik egy nagyszabású kutatás a nyelvészeti adatokról (REIS et al. 1998.). Az eddigi eredményeket többek között 2004-ben a „Linguistic Evidence. Empirical, Theoretical, and Computational Perspectives” c. konferencián mutatták be. A konferencia anyagához l. KEPSEK–REIS eds. (2005.).

e) A „The Linguistic Review” c. folyóirat 2004-ben tematikus számot szentelt a nyelvtudomány tudomány módszertani problémáinak, melynek egyik szekciója az adatokkal foglalkozott. E kötet releváns írásai LEHMANN (2004.), MEREU (2004.) és SIMONE (2004.).

f) Az Institut für Deutsche Sprache Mannheim 2005-ben szintén konferenciát szentelt az adatok problémakörének, melynek anyagához l. KALLMEYER–ZIFONUN Hrsg. (2007.).

g) A problémafelvetés a nyelvészet eddig említett irányzatain kívül újabbakra is kiterjed. Így pl. KRISTIANSEN et al. eds. (2006.) tanulmányai a kognitív nyelvészet empirikus alapjait tárgyalják, kiemelve többek között az adatproblematikát.

A közelmúltban a nyelvészeti adat-/evidenciavita két további vonulata bontakozott ki. Az egyik a „Corpus Linguistics and Linguistic Theory” c. folyóirat már említett, „Grammar without grammaticality” c. tematikus számában jelent meg (STEFANOWITSCH – GRIES eds. 2007.). A másik a „Theoretical Linguistics” c. folyóirat „Data in generative grammar” c. számában látott napvilágot (STERNEFELD

ed. 2007.). Az adat-/evidenciavita e két ága azonban – átfedések mellett – részlegesen különbözik a jelen írásunkban elemzettől, ezért itt nem térünk ki rájuk, hanem külön tanulmányokban tárgyaljuk őket (KERTÉSZ–RÁKOSI 2008a, b).

A fent felsorolt legújabb szakirodalom nagyon határozottan hívja fel a figyelmet a nyelvészeti adatokra és evidenciára vonatkozó problémák felvetésének jelentőségére. Vö. pl.: „Mindenekfelett nyilvánvaló, hogy az adatokra vonatkozó kérdések fontosabbak, mint olykor feltételezik” (BORSLEY 2005: 1479; kiemelés: K. A. – R. Cs.); „Röviden, a nyelvészeti evidencia különlegesen fontos téma, és egyúttal kihívást jelentő probléma minden nyelvészeti irányzat képviselői számára” (KEPSEK–REIS 2005: 1–2; kiemelés: K. A. – R. Cs.). Egyúttal több szerző is hangsúlyozza, hogy nem a korábbi viták mechanikus folytatásáról van szó, hanem új és a korábbiaknál relevánsabb kérdések kerültek reflektorfénybe: „[...] valószínűnek látszik, hogy az adatokra vonatkozó kérdések a jövőben némileg jelentősebbnek mutatkoznak majd, mint a múltban” (BORSLEY 2005: 1475; kiemelés: K. A. – R. Cs.); „A probléma alapvető jelentőségéhez képest a nyelvészeti evidencia figyelemreméltóan új témája a nyelvészeti diszkusszióknak” (KEPSEK–REIS 2005: 2; kiemelés: K. A. – R. Cs.). Az adat/evidenciaprobléma vitathatatlanul a jelenlegi nyelvtudomány érdeklődésének homlokterében áll. Nem köthető egyetlen irányzathoz, hanem a legkülönbözőbb részterületek és irányzatok művelőit foglalkoztatja és készletti reakcióra.

Ugyanakkor a (P1) elválaszthatatlanul összefonódik olyan tudományelméleti háttérfeltevésekkel, amelyek tematizálása nélkül nem oldható meg. LEHMANN például így fogalmazza meg ezt a felismerést: „Az adatok [...] természetére vonatkozó kérdések [...] szorosan összefüggnek azzal a kérdéssel, hogy tulajdonképpen milyen tudomány a nyelvészet, és szemmel láthatóan feltételezik annak tisztázását. E két kérdés viszonya azonban nem értelmezhető egyoldalú előfeltételezésként, mert egy tudomány lényegét meghatározzák a tárgyára és metodológiájára vonatkozó döntések (LEHMANN 2007: 9; kiemelés: K. A. – R. Cs.).

Amennyiben tehát az adat-/evidenciaprobléma (P1)-gyel fémjelzett vonulatáról pontos képet kívánunk rajzolni, nem kerülhetjük meg annak vizsgálatát, hogy a (P1) megoldására tett javaslatokban milyen szerepet játszanak metatudományos megfontolások. Ezért jelen áttekintésünkben a következő kérdésre keresünk választ:

(P2) Milyen szerepet játszik a metatudományos reflexió a (P1) megoldására tett javaslatokban?

E) A gondolatmenet. – Bármily friss és aktuális is az adat-/evidencia-problémával foglalkozó vita említett vonulatának eddigi anyaga, máris szerteágazó és sokrétű. Több, egymással párhuzamosan futó, egymást részlegesen átfedő vagy éppenséggel egymással ellentétes irányba mutató tendenciát foglal magában. Jelen írásunkban a legújabb szakirodalomból azoknak a publikációknak az elemzésére szorítkozunk, amelyek progresszívitása abban rejlik, hogy középpontjukban explicit módon a (P1) áll. Elemzésük alapján választ adhatunk a (P2)-re. Ennek megfelelően nem térünk ki az említett kötetekben található esettanulmányokra, hanem a következő áttekintő írásokat tárgyaljuk: BORSLEY (2005.), GEERAERTS (2006.), KEPSEK–REIS (2005.), LEHMANN (2004.), PENKE–ROSENBAACH (2004.); néhány ponton utalunk SCHÜTZE (1996.)-ra is.

Gondolatmenetünk a következő lesz. A 2. fejezetben az imént idézett írások alapján bemutatjuk, hogy a nyelvészeti – azaz t á r g y t u d o m á n y o s szinten folyó – kutatómunka gyakorlata jól nyomon követhetően eltávolodott az (SF)-ben rögzített nézetrendszerektől. A 3. fejezetben azt vizsgáljuk meg, hogy a kutatás gyakorlatát ebben a tekintetben milyen módon követi a m e t a t u d o m á n y o s r e f l e x i ó, és milyen megoldásokat kínál a nyelvészeti adatokkal és evidenciával foglalkozó legújabb szakirodalom a (P1) problémára. Meg fogjuk mutatni, hogy a felvetett megoldási javaslatok egyike sem kielégítő. A szerzők mindegyike kísérletet tesz a nyelvészeti adatokkal és evidenciával kapcsolatos álláspontok áttekintésére és a helyzetértékelés tanulságainak levonására. Noha problémaérzékenységük, nyitottságuk és megoldási javaslataik számos aspektusa egyértelműen és jelentős módon túlmutat az (SF)-ben röviden jellemzett szemléletmódon, álláspontjuk számos releváns ponton hiányos marad, és az analitikus tudományelmélet standard nézetének maradványait tartalmazza. Ezért nem áll rendelkezésükre egy olyan szisztematikus és differenciált metaelméleti háttér, amely nélkülözhetetlen a (P1) árnyalt és átfogó kezeléséhez. A 4. fejezet ennek megfelelően összegzi a vizsgált írások kétarcú megoldását a (P1)-re és a (P2)-re, rajtuk túlmutató következtetéseket levonva. A következtetések a nyelvészeti adatok és evidencia egy új modelljének kidolgozását motiválják, melynek bemutatására azonban jelen írás terjedelmi korlátai között nem kerülhet sor.

2. Az (SF)-től való elmozdulás igénye a tárgytudományos kutatómunka gyakorlatában. – A 80-as és 90-es évek egyre természetlenebbé váló tudomány módszertani vitái nem, vagy csupán elszórtan és rendszertelenül reflektáltak egy sor, az adatokkal és az evidenciával kapcsolatos, a mindennapos nyelvészeti kutatómunka gyakorlatában megfigyelhető fejleményre. E fejlemények tendenciaszerűen egy irányba mutatnak: arról tanúskodnak, hogy mély szakadék tátong a nyelvészeti kutatómunka g y a k o r l a t a és az (SF)-ben összefoglalt m e t a t u d o m á n y o s e l v e k között. Az alábbiakban – az előző szakaszban említett szakirodalom alapján – röviden áttekintjük azokat a tendenciákat, amelyek arra utalnak, hogy a nyelvészeti kutatómunka gyakorlata egyre kevésbé mereven követi az (SF)-ben rögzített két tudomány módszertani álláspont valamelyikét, és hogy egyre határozottabban lép fel a kutatásban felhasználhatónak tekintett adatok differenciálásának igénye.

A) A d (S F) a). – Olyan új módszerek és új adattípusok jelentek meg, amelyek szétfeszítik az (SF) a)-ban vázolt kategorizáció kereteit. Például a grammatikai kompetencia új értelmezését tették lehetővé probabilisztikus módszerekkel dolgozó megközelítésmódok (vö. PENKE–ROSENBACH 2004: 495–7). A számítógépes szimulációk olyan módszert testesítenek meg, amelyet mind a nyelvi intuíciókat adatként elismerő generatív nyelvészeknek, mind a nyelvi megnyilatkozásokat rögzítő korpuszokat felhasználó nyelvészeknek elvileg el kellene utasítaniuk. Azonban igen gyümölcsözőnek bizonyult felhasználásuk, mivel sikeresen tesztelhetők velük olyan nyelvelsajátítási modellek, amelyek a neurológia mai állása szerint közvetlenül nem ellenőrizhetők (vö. PENKE–ROSENBACH 2004: 491).

B) A d (SF) b). – Egyes generatív nyelvészek – figyelmen kívül hagyva (SF) b)-t – olyan kutatások eredményeit is felhasználják hipotéziseik alátámasztására, amelyek nem introspektív adatok elemzésén alapulnak, hanem valamilyen korpuszból, vagy kísérletekből származó adatokkal dolgoznak. Például az univerzális grammatika elveinek azonosításával foglalkozó generatív nyelvészek hasznosnak ítélték és integrálták tipológiai, nyelvtörténeti, neuro- és pszicholingvisztikai kutatások eredményeit (vö. PENKE–ROSENBACH 2004: 494–5). Ugyanakkor elemzéseik során a korpuszszal dolgozók támaszkodnak nyelvi intuíciójukra is. Egyrészt az adatok elemzésének egyik nélkülözhetetlen összetevője a vizsgált nyelvi elemet tartalmazó konstrukciók jelentésének összehasonlítása (vö. BORSLEY 2005: 1477). Másrészt a korpuszok feldolgozása során el kell különíteni a nyelvbtlásokat és más módon hibás megnyilatkozásokat a grammatikus mondatoktól. Az ilyen típusú döntések meghozatala során a nyelvész saját nyelvi intuícióját is használja (vö. SCHÜTZE 1996: 2; l. még ehhez LEHMANN 2004: 200, 207).

C) A d (SF) c). – A kutatások elmélyülésének további igen lényeges hozadéka az, hogy a kutatók – (SF) c)-vel szakítva – immár nem tekintik problémamentesnek a saját elméletük által preferált adattípusokat sem, mivel egyre több olyan tényezőt azonosítanak, amelyek befolyásolhatják az adatok megbízhatóságát. Egyrészt folyamatosan nő a felhasznált adattípusok komplexitása és az ellenőrizendő tényezők száma. Másrészt sokszor azzal kell szembesülniük a kutatóknak, hogy e tényezők között akadnak olyanok, amelyeket nem tudnak kontrollálni. E problémákon azonban nem általános módszertani elvek segítségével próbálnak meg úrrá lenni, hanem az egyes adattípusok megbízhatóságát befolyásoló speciális tényezőket kísérlik meg minél pontosabban felderíteni. Megemlítünk néhány példát:

A korpuszokkal kapcsolatban például mind a 'pozitív', mind a 'negatív evidencia' komoly és a kutatók rendelkezésére álló metodológia eszköztárával nem vagy nehezen kezelhető problémákat vet fel. A pozitív evidencia bizonyítékot jelenthet valamely hipotézis mellett, azonban gyakran nehézséget okoz annak eldöntése, hogy hogyan különböztessük meg a releváns adatokat a lényegtelen, izolált, megbízhatatlan, gyanús információktól. Az adatok megbízhatóságának ellenőrzése elkerülhetetlenül további adatgyűjtéshez és az elmélet bonyolultabbá válásához vezet; ennek azonban nyilvánvalóan vannak gyakorlati korlátai. A negatív evidencia (azaz egy bizonyos forma/konstrukció hiánya) eleve gyengébb bizonyítékot jelent csupán. Ráadásul adott esetben nem feltétlenül nyújt felvilágosítást arról, hogy egy bizonyos grammatikai szerkezet hiányának oka törvényszerűnek tekinthető-e, vagy csupán bizonyos körülmények esetlegességében gyökerezik (például abban, hogy nem vizsgáltunk meg elég nagy adatmennyiséget, vagy véletlenül nem találtunk rá egy ritka szerkezet egyébként létező és dokumentálható előfordulására). A kvantitatív adatok esetében a kérdés nem az, hogy előfordul-e valamilyen jelenség, hanem az, hogy milyen mértékben (tehát például statisztikailag szignifikánsan) fordul-e elő (vö. PENKE–ROSENBACH 2004: 486; l. még ehhez a D) pontot is). A kérdőívek és interjúk szintén komoly metodológiai problémákat vetnek fel, melyek jól ismertek a szociológiából. Ilyen a „megfigyelő paradoxona”: a nyelvész azt akarja tanulmányozni, hogy hogyan viselkednek a beszélők, amikor nem figyelik meg őket – ezt azonban csak úgy teheti, ha megfigyeli őket. A paradoxon egyik jellegzetes

megnyilvánulási formája abban áll, hogy az adatközlők a kísérlet során explicit kérdésre válaszolva elhatárolják magukat a nem-standard alakoktól, miközben ők maguk a mindennapokban minden további nélkül használják őket (l. PENKE–ROSENBACH 2004: 490).

A kísérletek eredményét potenciálisan befolyásoló tényezők kontrolljára való törekvés ugyancsak komoly problémákhoz vezet. A változók kontrollja érdekében a kísérletek gyakorta olyan mesterséges szituációkat teremtenek, amelyek kevésbé természetes adatokat szolgáltatnak csupán. Ugyanakkor PENKE–ROSENBACH (2004: 490) – például EISENBEIß 1994-et idézve – olyan kísérleti technikákat említenek, amelyek (legalábbis szerzőik intenciói szerint) lehetővé teszik, hogy természetes szituációkból nyerjünk adatokat.

Végül, az introspektív adatokkal kapcsolatos problémák felvetették a generatív nyelvészeti adatkezelési technikáinak megváltoztatását is (vö. SCHÜTZE 1996: 1).

D) A d (SF) d). – Az adatok és az elmélet kapcsolatára vonatkozó, az analitikus tudományelmélet standard nézetének valamelyik irányzatára hivatkozó módszertani előírásokat is megsértik az egyes irányzatok képviselői, szembefordulva (SF) d)-vel. Egyrészt a korpusz-adatokat preferáló nyelvészek között is vannak olyanok, akik elfogadják a falszifikacionizmust, noha továbbra is jelentős – de immár korántsem kizárólagos vagy döntő – szerepet tulajdonítanak az indukciónak (vö. pl. GEERAERTS 2006: 24). Másrészt egyre többek számára vált világossá, hogy a POPPERi modell túl szigorú. Az egyik stratégia a *falszifikálhatóság kritériumának* *la zítása*. A kritérium „erős” változata (azaz az egyetlen ellenpéldát cáfoló erejűnek tekintő felfogás) mellett a kutatás gyakorlatában megjelent annak „gyenge” verziója is, amely szerint a nyelvi szabályok nem szigorú előírásokként értelmezhetők, hanem sokkal inkább statisztikailag kimutatható tendenciákként. Éppen ezért a statisztikailag ritka előfordulások még nem elegendőek egy elmélet vagy hipotézis megcáfolásához. Nem világos ugyanakkor, hogy hol húzódik a határ eme ritka előfordulások és a cáfoló erejűnek tekintendő ellenpéldák között (vö. PENKE–ROSENBACH 2004: 483). – Egy másik, elsősorban a generativisták által alkalmazott stratégia az *ellenpéldák időleges figyelmen kívül hagyása*. Ez bizonyos esetekben azért lehet CHOMSKY (2002: 98–105) szerint racionális döntés, mert az elmélet fejlődésének egy későbbi stádiumában megteremtődhetnek azok az eszközök, amelyek segítségével az adott ellentmondás feloldhatóvá válik (PENKE–ROSENBACH 2004: 484). Ezekben az esetekben elhamarkodott döntés lenne azonnal feladni a szóban forgó elméletet/hipotézist.

E) A d (SF) e). – Az ellenpéldák időleges figyelmen kívül hagyása egyúttal óvatos elmozdulás (SF) e)-hez képest is, mivel azt jelenti, hogy az ellenpéldát nem tekintjük cáfoló evidenciának, és az inkonzisztenciát nem az érintett hipotézis, hanem a szóban forgó adat – átmeneti – feladásával oldjuk fel.

Az a D)-ben már említett jelenség, hogy számos nyelvész az indukciónak és a dedukciónak egyaránt szerepet tulajdonít az elméletalkotás folyamatában, szintén elmozdulást jelent (SF) e)-től, és egyúttal a „felfedezés kontextusának”, illetve az „igazolás kontextusának” merev elválasztásától is (vö. GEERAERTS 2006: 24–5).

Összegzés. A fenti áttekintés összegzésképpen megállapíthatjuk, hogy a kutatás *gyakorlatának* változásaiban jól kitapinthatóan jelen van az (SF)-

ben összefoglalt nézetektől való elszakadás szándéka. A két nézet antagonizmusa feloldódni látszik. Állításaik jóval differenciáltabbá, rugalmasabbá, árnyaltabbá válnak. (Ezt a felismerést pl. LÜDELING így összegzi LEMNITZER–ZINSMEISTER 2006-hoz írott előszavában: „Megértettük, hogy eltérő kérdésfelvetések eltérő adatokat is igényelnek, hogy közös forrásokra, eljárásokra és standardokra van szükségünk, és ezeket éppen ezért létre kell hoznunk és értékelnünk kell” – LÜDELING 2006: 4).

A kérdés számunkra ezek után az, hogy a gyakorlat változásait követi-e, és ha igen, milyen módon és milyen mértékben a metatudományos reflexió változása. A következőkben áttekintjük az 1. szakaszban hivatkozott legújabb szakirodalomnak a nyelvészeti adatok mibenlétére vonatkozó felfogását. Noha a nézetek, mint majd látni fogjuk, számos tekintetben jelentősen eltérőek, mégis végigvonul rajtuk néhány olyan szempont, amelyek lehetővé teszik, hogy egységes szerkezetben, a konvergenciák kiemelésével, ugyanakkor a fennálló különbségek pontos bemutatásával értékeljük a (P1)-re adott megoldási kísérleteiket.

(Folytatjuk.)

KERTÉSZ ANDRÁS – RÁKOSI CSILLA