

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

3. Az (SF)-től való elmozdulás igénye a metatudományos reflexióban. – A) Az adatok. – 1. A nyelvészeti adatok sokfélesége. – Az adatokkal kapcsolatos metaelméleti reflexiónak mindenekelőtt – a 2. A)-ban és B)-ben vázolt tendenciákhoz kapcsolódva – meg kell haladnia az empíriavita előző stádiumát uraló introspektív adat vs. korpuszadat dichotómiát. Reflektálnia kell arra a tényre, hogy a nyelvészeti kutatások elmélyülése, valamint a kognitív tudomány egyes ágaival, a szociológiával, vagy a számítógéptudománnyal stb. való együttműködése nyomán új adattípusok és adatkezelési módszerek sora jelent meg. KEPSER–REIS (2005: 1) ennek megfelelően az adatok következő típusait sorolja fel példaként: introspektív adatok, korpuszadatok, kísérletekből származó adatok, szinkrón vs. diakrón adatok, tipológiai adatok, neurolingvisztikai adatok, az első és a második nyelv elsajátításából származó adatok stb. KEPSER–REIS (2005.), PENKE–ROSENBACH (2004.), BORSLEY (2005.) és SCHÜTZE (1996.) minden információt potenciális adatnak tekint, melyek a nyelvi viselkedés valamely aspektusával kapcsolatosak. Például PENKE–ROSENBACH (2004: 493) hangsúlyozza, hogy nincs a priori különbség az adatok különböző típusai között abban az értelemben, hogy az egyik eleve jobb evidenciát szolgáltathatna a nyelvi képesség vizsgálatához, mint a másik.

Noha LEHMANN a nyelvészeti adatok sokféleségét megragadó szempontrendszert javasol, megközelítése nem jelent egyértelmű szakítást az (SF)-fel. Míg egyfelől az introspektív és a korpuszadatok egyszerű szembeállításánál messzemenően differenciáltabb adattipológia felvázolására tesz kísérletet, addig másfelől éppen (SF) a) és b) szellemében utasítja el az introspektív adatok legitimitását: „[...] az introspekciónak ez az alkalmazása visszaélés az empirikus tudomány fogalmával és éthoszával. [...] az introspekció a nyelvészeti kutatómunka heurisztikus eszközeként szükséges és hasznos, de nem az empirikus metodológia része, és az ily módon létrehozott adatoknak nincs helyük az empirikus kutatásban azon túl, hogy illusztrálják azt, amire a nyelvész elméleti tevékenysége vonatkozik [...]” (2004: 196; kiemelés: K. A. – R. Cs.).

* L. MNy. 2008: 274–86.

Hasonlóan vélekedik GEERAERTS (2006: 27) is, amennyiben kizárólag a pszichológiában használt kísérleti és a szociológiában elfogadott adatgyűjtési technikákat, valamint a korpuszadatokat tartja legitimnek, és elveti az introspektív adatok használatát.

Összegzés. Mint látjuk, az említett álláspontok jelentősen túllépnek a korpuszadat vs. introspektív adat dichotómián. Míg azonban a jelenlegi szakirodalom egy része (legalábbis bizonyos feltételek mellett) legitimnek tartja az összes felsorolt adattípust, addig van olyan nézet is, amely – bár a korábbinál jóval differenciáltabb és gazdagabb adattaxonómiát dolgoz ki – továbbra is az (SF) a) és b) szellemében jár el, amennyiben kirekeszti az introspekcióval kapott eredményeket.

2. A nyelvészeti adatok forrásai. – Az egyes adattípusok tulajdonságainak feltárásakor releváns, de nem kizárólagos tényezőként kell figyelembe venni azokat, amelyek forrásaik különbözőségéből adódnak. PENKE–ROSENBACH (2004: 488–90) forrásuk alapján két nagy csoportba osztja az adatokat. A spontán adatok természetes módon előforduló nyelvi megnyilatkozások, melyek korpuszban rögzítve a nyelvészeti elemzés anyagául szolgálnak. Korpuszt létrehozhat maga a nyelvész speciálisan valamilyen nyelvi jelenség tanulmányozására, de dolgozhat nyilvánosan hozzáférhető, nem feltétlenül egy bizonyos konstrukció tanulmányozására szánt adatgyűjteménnyel is. A kiváltott („elicited”) adatok valamilyen kísérlet vagy megfigyelés révén, a nyelvész által létrehozott és irányított szituációban jönnek létre. PENKE és ROSENBACH úgy véli, hogy a kétfajta adattípusnak különböző előnyei és hátrányai vannak. A spontán adatok minőségét nyilvánvalóan befolyásolja a korpusz reprezentativitása és összeállításának szisztematikussága, a „kiváltott” adatok megbízhatósága pedig nagyban függ az adatfelvétel módjától, valamint az alkalmazott kísérleti technikától.

LEHMANN (2004: 196–8) hasonló csoportosítást alkalmaz, amikor – miután az introspekciót elutasítja – az adatgyűjtés két „komoly módját” különbözteti meg: a talált és a kutató által létrehozott adatokat, azonban egyértelműen az előbbi csoportot preferálja. Úgy véli, a talált adatok két szempontból is jobban képesek eleget tenni az empiricitás azon követelményének, hogy a vizsgált jelenségek tulajdonságainak függetlennek kell lenniük a kutatótól (vö. i. m. 192). Egyrészt álláspontja szerint a korpuszadatokra épülő kutatómunka objektivitása magasabb, mint a más adattípust felhasználóé, mert kevesebb lehetőség van az adatok manipulációjára. Másrészt nézete szerint egy olyan korpusz, amelyet nem a nyelvész maga konstruált, változatos, váratlan információkkal teli, ezért serkentheti az eredeti, az előzetes elképzelésekre rácafoló felismeréseket (i. m. 201). A kutató által létrehozott adatok közül egyedül a nem-nyelvi stimulusok által kiváltott nyelvi viselkedésből származó, valamint a többféle módszer kombinációjával nyert adatokat tartja legitimnek. A metanyelvi eszközökkel kiváltott adatokkal dolgozó, illetve az egyik nyelvről a másikra történő fordítást alkalmazó módszereket például nem empirikus, hanem hermeneutikai módszernek tekinti (vö. i. m. 201–4), mivel kontrollálhatatlannak ítéli azokat a tényezőket, amelyek a releváns adatok körének kiválasztását, megbízhatóságát és az eredmények értékelését befolyásolják. Ezért az ilyen forrásból származó adatokat csupán más típusú adatok mellett, azok kiegészítéseként tartja elfogadhatónak. Azonban, miközben rendkívül éles szemmel és ala-

posszággal tárja fel a talált adatok gyenge pontjait is, értékelésükkor jóval engedékenyebben jár el, mint a létrehozott adatok esetében, noha a problémák között szembevetendő hasonlóság tapasztalható (vö. pl. i. m. 200).

Összegezés. Egyrészt azt állapíthatjuk meg, hogy kétségkívül jelentős szemléletváltás történt az adatok forrásának megítélésével kapcsolatban. Míg az (SF) b) az adatok forrását még döntő tényezőnek tekintette, addig az itt vizsgált szakirodalom csupán a releváns tényezők egyikeként értékeli. Továbbá, az (SF) képviselői csupán egyetlen adatforrást tartottak legitimnek, ezzel szemben a mai álláspont források sokaságát ismerik el. Másrészt viszont ez a szemléletváltás nem jár együtt az adatforrások szisztematikus és megalapozott feltárásával, hanem következtetlenül megy végbe.

3. A nyelvészeti adatok komplexitása. – Az (SF)-ben bemutatott nézetrendszert nem csupán az adatok körének leszűkítése miatt tartják elégtelennek az említett szerzők, hanem azért is, mert – mint azt az (SF) c) tanúsítja – elhanyagolt egy másik, az adatok megbízhatóságát döntő módon befolyásoló tényezőt: az adatok komplexitását.

E komplexitás okát KEPSEK–REIS (2005: 1) abban látja, hogy a nyelvészeti kutatómunka tárgya nem áll közvetlenül a nyelvész rendelkezésére, hanem a nyelvi magatartás közvetlenül érzékelhető megnyilvánulásából kell rekonstruálni.

PENKE–ROSENBACH (2004: 487–8) ezzel összhangban különbséget tesz a nyelvtudást direkt, illetve indirekt módon visszatükröző adatok között. A korpuszadatok és az intuitív adatok szerintük közvetlenebbül tükrözik a nyelvtudás tulajdonságait, mint a kísérleti eredmények, melyek csupán indirekt módon adnak felvilágosítást egy-egy nyelvi konstrukció tulajdonságairól. Széles körű áttekintést nyújtanak az egyes adattípusok összetettségéből adódó, 2. C)-ben bemutatott problémákról, azonban nem tesznek kísérletet szisztematizálásukra és általánosítások megfogalmazására.

LEHMANN (2004.) az adatok szerkezetének sokkal több aspektusát tárja fel, gazdag szempontrendszert alkalmazva:

a) Az adatok nem eleve adóttak, hanem – legalábbis bizonyos mértékig – a kutató hozza őket létre. Bár mindig valamilyen, a kutatótól függetlenül létező entitásra vonatkoznak, még a legelemibb nyelvészeti adat identifikációja is olyan eljárás, amely absztrakcióra és szemiotikai folyamatokra épül.

b) Az elsődleges adatok a beszédesemények tér-idő koordinátákat is tartalmazó reprezentációi. A másodlagos adatok nem rendelkeznek történeti identitással, azaz nem köthetők konkrét helyhez és időhöz, következésképpen absztraktabbak. Közéjük tartoznak a tokennek tekinthető elsődleges adatok type-jai. Szintén másodlagos adatok a metanyelvi állítások, melyek a beszédesemények („speech events”) tulajdonságait vagy általánosításokat rögzítenek, mint például a grammatikalitási ítéleteket jelölő, csillaggal ellátott példamondatok vagy a negatív adatok, amelyek azt állítják, hogy egy konstrukció nem létezik. Az ilyen típusú adatokban a meta- és a tárgynyelv keveredik.

c) LEHMANN különbséget tesz nyers adatok és feldolgozott adatok között is. A nyers adatok a beszédesemények ikonikus leképezései, mint például egy magnófelvétel. A nyelvészek azonban szinte sohasem nyers adatokat vizs-

gálnak, hanem feldolgozzák őket oly módon, hogy speciális aspektusukat kiemelő, a vizsgálódás szempontjából érdektelen vonásaikat pedig elfedő, ezzel a beszédeseményt annak relevánsnak tekintett aspektusaira redukáló – és ily módon többé-kevésbé el is torzító – szimbolikus reprezentációt készítenek róluk. Ilyenek például a nyelvi megnyilatkozások transzkripciói vagy a szintaxisban használatos ágrajzok.

d) Az eredeti adatfelvételek közvetlenül a beszédesemény alapján készülnek, míg a levezetett reprezentációk más reprezentációkon, azaz adatokon alapulnak. Például egy beszélgetésről készített hangfelvétel az *A* nyelvész számára azokat az adatokat jelenti, amelyekre akusztikus fonetikai vizsgálatait építi. *B* nyelvész – szegmentális fonológusként – a hangfelvételnek az IPA segítségével előállított reprezentációját tekinti adatnak. *C* – aki posztgeneratív fonológus – ezt szupraszegmentális relációkat jelölő szimbólumokkal egészíti ki. *D* konverzáció-analitikus, aki például a németországi konverzáció-analitikai hagyományban ismert transzkripció-s rendszerek valamelyikét alkalmazva hozza létre a rögzített konverzáció szerkezetének bizonyos, többek között a „turn-taking”-et tartalmazó transzkripcióját. *E* szintakta, aki az utóbbit olyan adatnak tekinti, amely segítségével kidolgozhatja a Tübingen környéki sváb dialektus beszélt nyelvi grammatikájának egy fragmentumát. *F* e grammatika megállapításait tekinti olyan adatoknak, amelyekre hivatkozva a beszélt és az írott nyelv különbségeit kívánja feltárni. A példák sora természetesen hosszan folytatható, és jól szemléltet legalább két lényeges összefüggést. Egyrészt azt, hogy ami az egyik kutató számára az adatok elemzésének eredménye, az a másik számára adat lehet. Másrészt azt, hogy az adatok körének meghatározása nem azok inherens tulajdonságaitól, hanem a mindenkorin kutatók céljától, eszközeitől, fogalomrendszerétől, problémáitól függ.

Összegzés. A szerzők egyetértenek abban, hogy az (SF) c)-t fel kell adni, mert döntő módon akadályozza a 2. C)-ben leírt problémák megoldását. Az adatokkal kapcsolatos reflexió egyik legfontosabb és elengedhetetlen feladatának éppen ezért az adatok szerkezetének feltárását tekintik. Ugyanakkor az adatok komplexitásának megragadására tett kísérletek egészen eltérő kategorizációkat alkalmaznak, és eltérő alaposságúak.

4. Az adatok elfogadhatósága. – Az adatok elfogadhatósága a szakirodalom jelenlegi állása szerint – az előző szakaszban mondottakból kiindulva, szakítva az (SF) a)–c)-ben bemutatott elképzelésekkel és többé-kevésbé illeszkedve a 2. C)-ben kirajzolódó képhez – egy több tényezőtől függő folyamat alapján ítéltető meg, melynek figyelembe kell vennie az adatok forrását, szerkezetét és a szóban forgó elmélet bizonyos sajátosságait is.

KEPSER és REIS szerint a mindenkorin adatok elfogadhatóságára vonatkozó döntés nem korlátozódhat kizárólag egy bizonyos adat(típus) vizsgálatára, hanem az összes rendelkezésre álló adat(típus) és további, részben nem nyelvi tényezők figyelembevételével történik (2005: 1). Ebben a tekintetben lényegesen túllépnek az (SF) b) monolitikusságán.

PENKE és ROSENBACH megelégszik annak leszögezésével, hogy az adatgyűjtésnek szisztematikusnak kell lennie, noha az általuk idézett példák világosan jelzik e kritérium elégtelenségét.

BORSLEY is felveti, azonban nyitva hagyja az adatok megbízhatóságának problémáját. Kiemelendő ugyanakkor, hogy – CHOMSKYra hivatkozva – nem tekintti például a grammatikalitási ítéleteket megdönthetetlennek, hanem olyan kijelentéseknek, melyek megbízhatóságát számos, részben nem kontrollálható tényező befolyásolja (vö. 2005: 1476–7).

GEERAERTS (2006: 45) egyfelől a módszerek, és így az adadfajták sokféleségét hasznosnak tartja, mivel kiegészítve egymást megbízhatóbb eredmények elérését teszik lehetővé, mint külön-külön. Másfelől viszont úgy véli, mindenképpen szükség van arra, hogy ezeknek a módszereknek egy közös, szigorú metodológia szolgáltson alapul. A nyelvészet előtt álló feladatok egyik legfontosabbika ezért véleménye szerint egy elméletfüggetlen empirikus bázis megteremtése. A nyelvészeti elméletek hipotéziseit úgy kell megfogalmazni, hogy világosan szembesíthetők legyenek a „megfigyelésekkel” (vö. i. m. 23–7). Ehhez szerinte mindenekelőtt egy elméletfüggetlen megfigyelési nyelvre van szükség, amelyen az adatok megfogalmazhatók, és amelyre az elméletek hipotézisei lefordíthatók. Ezzel azonban visszacsúszik az analitikus tudományelmélet standard nézetét kritikátlanul elfogadó szemléletmódba. (A megfigyelés elméletfüggőségének tétele az 1960-as évektől vált a tudományelméletben általánosan ismertté HANSON (1958.), DUHEM (1974. [1906.]), és KUHN (2000. [1970.]) nyomán.)

LEHMANN (2004: 204–8) a 3. A) 3. szakaszban bemutatott adattipológia alapján azt az általánosítást fogalmazza meg, hogy a nyelvészeti adatok többnyire nem térben és időben lejátszódó beszédesemények, hanem azok különféle absztrakciós szinteken lokalizált, különféle típusú információkkal kiegészített reprezentációi. E megállapításból azt a következtetést vonja le, hogy az adatok elkerülhetetlenül rendelkeznek probléma- és elméletfüggő tulajdonságokkal. Egyrészt a vizsgálat tárgyának csupán bizonyos aspektusait reprezentálják, másrészt – mivel többszörös áttétellel jöttek létre – leegyszerűsítik és eltorzítják az eredetit. Úgy véli ugyanakkor, hogy rendelkezésünkre áll az a módszertan, amelynek alkalmazásával megbízható adatokhoz juthatunk. Ebben a tekintetben felfogása összhangban áll GEERAERTS nézetével (vö. LEHMANN 2004: 185–6).

LEHMANN úgy véli, hogy a nyers adatok segítségével ellenőrizhető és korrigálható a reprezentálás helyessége (2004: 207). Ez a felfogás nem lép túl az (SF)-ben összefoglalt nézetrendszeren, hiszen metaelméleti háttérként az analitikus tudományelmélet standard nézetét előfeltételezi, és általános módszertani elvekkel biztosíthatónak véli azt, hogy az adatokat szilárd tényekként kezelhessük. Ez azonban azzal a következménnyel jár, hogy a metaelméleti reflexió egyik szintje ellentmondásba kerül a másikkal. Az adatok belső struktúrájának elemzésekor LEHMANN több megállapítása is arra utal, hogy az adatok – lévén reprezentációk – nem tekinthetők bizonyosan igaz, eleve adott „tényeknek”, hanem csupán olyan kiindulópontoknak, amelyek további finomításra szorulnak, és elégtelennek, nem eléggé megbízhatónak, sőt akár hasznavehetetlennek is bizonyulhatnak (vö. pl. i. h.). Ezzel szemben azok az eljárások, amelyeket javasol, az ellenkező irányba mutatnak, amennyiben nem csupán normaként előírják, hanem minden további nélkül megvalósíthatónak is tételezik, hogy maximálisan megbízható adatokhoz jussunk: „Általánosságban ahhoz, hogy egy adatot egy adott tudományágban valóban adatnak fo-

gadhassunk el, olyan eljárásokra van szükség, amelyek a másodlagos adatokat az elsődlegesekhez, az elsődleges adatokat pedig a végső szubsztrátumhoz kapcsolják. Az ilyen eljárások az adott tudományág módszertanának részét képezik, azaz olyan módszerek részei, amelyek lehetővé teszik a kutatók számára, hogy kontrollálják az elmélet és az adatok közötti viszonyt. [...] Ha nincsenek ilyen eljárások, akkor először is nincs alapja annak, hogy egy adatot bizonyosan adottnak tekintünk, ami azt jelenti, hogy definíciónk szerint nem tekinthető adatnak; másodszer pedig nincs lehetőség arra, hogy az elméletet kapcsolatba hozzuk egy érzékelhető episztemikus objektummal, ami azt jelenti, hogy nem empirikus elmélet” (i. m. 185–6; kiemelés: K. A. – R. Cs.).

SCHÜTZE viszont – noha ő is egyértelműen elutasítja a szigorú kontroll nélkül kapott, kísérletekkel alá nem támasztott adatokra épülő elméleteket (vö. 1996: 4–5) – nem tartja a természettudományokban, a pszichológiában vagy a szociológiában elterjedt módszertani elveket tévedhetetlennek. Rámutat arra, hogy az adatok megbízhatóságának megítélése rendkívül komplex feladat, mely nem szakítható el az elméletalkotás folyamatától. Csak akkor tudjuk kontrollálni vagy éppen kiküszöbölni az adataink megbízhatóságát befolyásoló tényezőket, ha szisztematikus módon feltárjuk azoknak a forrásoknak a tulajdonságait, amelyekből az adataink származnak – azaz ha hipotéziseket állítunk fel struktúrájukról és működésükről. Ez pedig azt jelenti, hogy az adatok megbízhatóságával kapcsolatos hipotéziseink is az általunk létrehozott nyelvészeti elmélet részét képezik (1996: 9). SCHÜTZE stratégiája azért tekinthető konstruktívnek és gyümölcsözőnek, mert arra ösztönöz, hogy aktívan kutassunk azok után a megkérdőjelezetlen háttérfeltevések után, amelyek hibásnak bizonyulhatnak, és ronthatják az adatok megbízhatóságát. Ez azt is jelenti, hogy az adatok megbízhatóságát befolyásoló tényezők feltárása, a hipotézisek felállítása és az adatokkal történő összevetése nem izolálható egymástól.

Összegzés. Mint láttuk, minden szerző felismeri az adatok elfogadhatóságával összefüggő problémák súlyát, és megkísérel kritériumokat felállítani, ami egyértelműen utal a metareflexív gondolkodás jelentős finomodására. Viszont a javasolt kritériumok között igen lényeges eltéréseket tapasztalhatunk, miközben egyik szerző sem vázol fel koherens és a felvetődött nehézségekkel valóban számot vető megoldási javaslatot. Mindenekelőtt azok az eszközök hiányoznak, amelyek lehetővé tennék az adatok komplexitásából fakadóan kiküszöbölhetetlen bizonytalansági tényezők kezelését. A bemutatott álláspontok egyike sem képes feloldani azt a feszültséget, amelynek egyik pólusa az a felismerés, hogy az adatok megbízhatóságát befolyásoló tényezők nem kontrollálhatók kimerítő módon, a másik pólusa pedig az analitikus tudományelmélet standard nézetétől örökölt intenció, hogy az adatoknak a hipotézisek teszteléséhez szilárd alapot kell biztosítaniuk (l. ehhez a 3. A) 5. és a 3. B) szakaszokat is).

5. Az adatok és az elmélet viszonya. – Az adat-/evidenciavita negyedik szakaszának metaelméleti irodalmában az adatok és az elmélet viszonyával kapcsolatos elképzeléseket továbbra is az indukcióra és a dedukcióra építő modellek uralják (LEHMANN 2004: 187, GEERAERTS 2006: 24). Ugyanakkor – a 2. D)-ben jelzett tendenciákkal összhangban – az egyes szerzők olyan meglátásokat is

megfogalmaznak, amelyek egyértelműen túlmutatnak az (SF) d)-n. Először is felvetődik, hogy az elméletalkotás folyamatában az adatok és a hipotézisek viszonya nem szűkíthető egyszerűen az induktív vagy a deduktív eljárás lineáris alkalmazására. GEERAERTS szerint „az empirikus kutatás magában foglal egy empirikus ciklust, amelyben az adatgyűjtésnek, a hipotézisek tesztelésének és az eredmények értelmezésének több fordulója követi egymást” (2006: 24–5; kiemelés: K. A. – R. Cs.).

Másodszor, szintén összeegyeztethetetlen az adatok és az elmélet hipotézisei közötti lineáris kapcsolattal az, hogy az egyes szerzők probléma- és elméletfüggőnek tekintik az adatok kiválasztását. Ha ugyanis az adatok kiválasztását jelentősen befolyásolják a vizsgált probléma sajátosságai (vö. LEHMANN 2004: 191; BORSLEY 2005: 1476; PENKE–ROSENBACH 2004: 480), valamint az adott elmélet konceptuális és metodológiai háttere (LEHMANN 2004: 181; KEPSEK–REIS 2005: 1; PENKE–ROSENBACH 2004: 493), akkor az adatok és a hipotézisek kapcsolata nem redukálható az elméletnek a tapasztalattal való szembesítésére, azaz a hipotéziseknek az elmélettől független adatok segítségével történő tesztelésére.

LEHMANN (2004: 191) az adatok és az elméletek viszonyát másokénál gazdagabb szempontrendszer alapján tematizálja. Három funkciót tulajdonít az adatoknak. Az első: az elmélet létrehozása során (az analitikus tudományelmélet standard nézetének terminológiájában ez a „felfedezés kontextusának” felel meg) a hipotézisekhez vezető úton az indukció bázisát alkotják. A második: a dedukcióval kapott hipotézisek tesztelésének eszközei (ami azt jelenti, hogy az „igazolás kontextusában” is szerepet kapnak). Végül, az elmélet prezentálásakor az argumentáció során az olvasó meggyőzését szolgáló evidencia szerepét is ellátják. Nem tisztázza azonban a három funkció egymáshoz való viszonyát, aminek következtében megközelítése több szempontból is nehézségekhez vezet.

Először is, a hipotézisek felfedezésében játszott szerepük elkerülhetetlen hozadékaként az adatokat – ahogy említettük – elmélet- és problémafüggőnek tekinti, és ezzel összefüggésben megjegyzi, hogy „nem húzható éles határ adatok és [a nyelvész által absztrakcióval és más szemiotikai műveletek segítségével létrehozott – K. A. – R. Cs.] konstruktumok között” (i. m. 208). Ugyanakkor – mint azt 3. A) 4-ben is említettük – úgy véli, hogy az adatok akkor lehetnek alkalmasak második funkciójuk ellátására, azaz akkor tesztelhetjük velük az elmélet hipotéziseit, ha a relativisztikus elemek hatását ki tudjuk szűrni, és biztosítani tudjuk a legkevésbé elmélet- és problémafüggő adatfajtaéhoz, a nyers adatokhoz való visszatérést. Éppen ezért élesen elutasítja az introspektív adatokat, és – mint azt a 3. A) 1. szakaszban idézett passzusban láttuk – csupán heurisztikus szerepet hajlandó tulajdonítani nekik. Az elmélet igazolására alkalmatlannak tartja őket, mivel úgy véli: „az, hogy a nyelvész saját maga produkálja az adatait, majd ugyanezen adatokra hivatkozva alkotja meg elméletét, *circulus vitiosus* eredményez”. Ezért az introspektív adatoknak eleve nem lehet kontrollfunkciójuk, „és az eljárásnak semmi köze a tudományos módszerhez” (i. m. 196; kiemelés az eredetiben).

Azonban világosan kijelenti, hogy a nyers adatokhoz való visszatérés is csupán megbízhatóbbá teheti az adatainkat, de bizonyosságukat nem garantálhatja (i. m.

205). Ugyanakkor azt a kívánalmat fogalmazza meg, hogy az adatoknak – legalábbis bizonyos mértékig – függetlennek kell lenniük a kutatótól (i. m. 181). Mindez nyilvánvalóan az ellenkező irányba mutat, mint az adatok elmélet- és problémafüggő voltának hipotézise. Nem tematizálja azt a kérdést, hogy ez a parciális adat- és elméletfüggőség miért és mennyiben lehet elegendő ahhoz, hogy az adatok „betöltsek kontrollfunkciójukat”.

A második probléma, amely az adatok három funkciója közötti viszony tisztázatlanságából származik az, hogy az adatok argumentatív funkciójához egy, a másik két funkcióhoz kapcsolódóktól lényegesen eltérő, azokkal inkompatibilis kritériumot kapcsol: „[valami] adatként funkcionálhat egy kutatásban, amely a megkérdőjelezhetetlen evidencia szerepét tulajdonítja neki az érvelésben” (i. m. 181). Ez a megfogalmazás az elméletalkotás megkérdőjelezhetetlen alapjaként jellemzi az adatokat. A megkérdőjelezhetetlenséget azonban nem valamilyen objektívnek tekintett kritériumra vezeti vissza, hanem relativisztikusan értelmezi, és a szóban forgó elméletben az adatoknak tulajdonított argumentatív szerephez köti. Ebből viszont az adatok elméletfüggőségének kellene következni – ami ellentmondáshoz vezet (l. ehhez a 3. B) pontot is).

Összegzés. A bemutatott álláspontok plasztikusan mutatják mindazon nehézségeket, amelyek abból fakadnak, hogy az adatok probléma- és elméletfüggősége, komplexitása és bizonytalansága nem kompatibilis az (ATSN) normáival. Míg az (ATSN) egyértelműen előírja a „felfedezés kontextusának” és az „igazolás kontextusának” merev elválasztását, és ezzel egyirányú logikai kapcsolatot követel meg az adatok és az elmélet hipotézisei között, addig a bemutatott álláspontok e viszonyt – a tárgy tudományos gyakorlatot elfogadva és legitimálva – ennél jóval bonyolultabbnak tételezik. Átgondolt és kidolgozott metaelméleti háttér híján azonban e nézetek döntő pontokon hiányosak és problematikusak maradnak.

6. Az ellentmondások kezelése. – A nyelvészeti elméletekben használt adatok további igen lényeges tulajdonsága, hogy ellentmondások fellépésének forrásai lehetnek. A legtriviálisabb megnyilvánulási módja ennek az, hogy a nyelvész folyamatosan „kivételekkel”, „ellenpéldákkal” szembeáll, valahányszor hipotéziseket kíván felállítani.

PENKE–ROSENBAACH (2004: 484) CHOMSKYt (2002: 98–105) idézve kiemeli, hogy a nyelvészeti elméletalkotásban vannak olyan helyzetek, melyekben nem lát-szik feltétlenül szükségesnek egy ellentmondás azonnali feloldása. A hipotézisnek ellentmondó adatoknak nem kell az elmélet elvetéséhez vezetnie, hanem célszerű ignorálni őket és kívánni, míg a kutatás egy későbbi szakaszában megteremtődnek azok az eszközök, amelyek segítségével az elmélet magyarázatot tud adni viselkedésükre. Kommentár nélkül szembeállítják azonban e stratégiával a korpusznyelvészek és a generatív grammatikát elutasító további irányzatok képviselőinek álláspontját, amely nem fogadja el a hipotéziseknek ellentmondó adatoknak ezt a fajta figyelmen kívül hagyását. Nem tesznek kísérletet arra sem, hogy tisztázzák a fenti stratégia viszonyát a falszifikáció 2. D)-ben említett „erős”, illetve „gyenge” változatához.

Noha más összefüggésben, KEPSEK–REIS (2005: 3) is hangsúlyozza, hogy a nyelvészeti adatok sokféleségéből eredő ellentmondások adott esetben termékenyek lehetnek. A különböző adattípusok egymásnak ellentmondó hipotézisekhez vezethet-

nek, azonban ezek az ellentmondások álláspontjuk szerint feloldhatók, és így egyértelműen előreviszik a kutatást (i. m. 1).

Az alapvető probléma ugyanakkor az, hogy az idézett szerzők – a fenti fontos felismerések ellenére – nem tesznek kísérletet az adatokból származó ellentmondások feloldási stratégiáinak feltárására.

Ami az utóbbiakat illeti, SCHÜTZE (1996: 4) az adatok közötti ellentmondásoknak a természettudományban használatos technikáit hiányolja a nyelvészetből, azonban nem fejti ki, és hivatkozással sem utal arra, hogy pontosan milyen módszerekre gondol.

GEERAERTS (2006: 37) a korpuszok esetében statisztikai módszerek alkalmazásában látja az ellentmondások feloldásának módját. Ez a javaslat az eddigiekhez hasonlóan az ellentmondások kezelésének fontos aspektusait érinti, azonban nem szakít gyökeresen az (SF)-ben összefoglalt nézetrendszerrel, hanem az (ATSN)-en belül maradva keres árnyaltabb megoldásokat. Az esetek egy részében kritikátlanul elfogadja a falszifikációs modellt, bár bizonyos körülmények között nem tartja alkalmazhatónak. Ha ugyanis kevés előfordulását találjuk egy nyelvi jelenségnek egy korpuszban, akkor felvetődik a kérdés, vajon nem csupán marginális adatokkal van-e dolgunk, melyek nem tekinthetők cáfoló vagy éppen konfirmáló evidenciának. Ilyen esetekben meg kell vizsgálnunk a lehetséges alternatívákat is, és statisztikai módszerekkel össze kell vetnünk őket. GEERAERTS ezen javaslata azonban több kérdést nyitva hagy. Először is, nem oldja fel például azokat a helyzeteket, amikor két, különböző típusú adatok által körülbelül egyforma erősen alátámasztott hipotézis áll egymással szemben. Másodszor, nem minden adattípusnál alkalmazható. Harmadszor – SCHÜTZE-vel ellentétben, aki csupán a megbízhatóbb hipotézis kiválasztásának szempontjairól beszél – GEERAERTS tökéletesen objektívnek és egyértelmű végeredményt produkálónak tartja a statisztikai módszereket. Csakhogy az alkalmazott próba kiválasztása és eredményének interpretálása – legalábbis bizonyos határig – köztudottan a kutató döntésén alapul, továbbá más próba alkalmazása más eredményhez vezethet, azaz ellentmondást generálhat. A kísérletek esetében GEERAERTS az experimentum crucisra hivatkozik (vö. i. m. 26) – ezt azonban csupán az (ATSN) szellemében teheti.

Összegzés. Az elemzett szakirodalom fontos és progresszív felismeréseként értékelhetjük azt a meglátást, hogy a különböző adattípusok miatt fellépő ellentmondások a nyelvészeti elméletalkotás elkerülhetetlen velejárói. Ez világosan jelzi az (ATSN)-nel való szakítást, és a tudományelmélet újabb irányzataival kimutatható konvergenciát. (Az inkonzisztencia a tudományban a kortárs tudományelmélet egyik legfontosabb témája, vö. MEHEUS ed. 2002. Következésképpen az a tény, hogy az inkonzisztencia kérdése egyáltalán felvetődött a fent vizsgált nézetekben, egyértelműen progresszív lépés, mely hozzájárulhat az elméleti nyelvészet módszertanának korszerűbbé tételéhez. Azonban kevés mű vizsgálja szisztematikus módon az inkonzisztencia szerepét a nyelvészeti elméletalkotásban, vö. MORAVCSIK 1993., 2006., KERTÉSZ 2004., KERTÉSZ–RÁKOSI 2006., RÁKOSI 2005.) Az itt bemutatott írások arról tanúskodnak, hogy nincs konszenzus arról, milyen elvek alapján dönthető el, mikor tolerálható egy ellentmondás és mikor nem, és hogyan oldhatóak fel az ellentmondások. Az egyes szerzők állásfoglalásai hiányosak. Így

nem alkalmasak arra, hogy döntsenek a 2. D)-ben bemutatott stratégiák között, vagy hogy új módszereket kínáljanak.

B) A nyelvészeti evidencia. – Az újabb szakirodalom nem explikálja a nyelvészeti evidencia fogalmát. Az 1. szakaszban említett írásokban világosan kimutatható egyrészt az a z a v a r, amely e fogalom kezelését jelenleg is jellemzi, másrészt viszont az a tendencia is, amely alapján kitapintható néhány relevánsnak ítélt tulajdonsága. Abból, hogy egyik szerző sem definiálja, mit ért evidencián, még nem feltétlenül következne, hogy zavar van meghatározása körül, hiszen jelenthetné azt is, hogy közismert, pontosan körülhatárolt jelentéssel bír. Viszont, ha alaposabban szemügyre vesszük a szerzők szóhasználatát, kiderül, hogy különböző értelem-ben és következtetlenségül alkalmazzák e terminust.

PENKE–ROSENBACH (2004.) több helyen szinonimákként használja az „adat” és az „evidencia” kifejezést. Például az adattípusok felsorolásánál a 485. lapon „qualitative vs. quantitative data” szerepel, míg a következő oldalon „qualitative vs. quantitative evidence”. A direkt és az indirekt evidenciát is az adattípusok közé sorolják, ugyanakkor – teljesen más módon értelmezve és kapcsolatba hozva az ’adat’ és az ’evidencia’ kategóriáját, mint az iménti esetben – arról is beszélnek, hogy az adatok egyik funkciója az, hogy a hipotézisek mellett szóló evidenciaként szolgálhatnak (i. h.). Ily módon – az (ATSN) intencióit követve – az evidenciát az adatok halmazának egy speciális, kitüntetett részhalmazaként értelmezik, melynek feladata az elmélet hipotéziseinek tesztelése. Úgy vélik ugyanakkor, hogy nem dönthető el általános módszertani kritériumok alapján, hogy mely adatok tölthetik be ezt a funkciót, hanem elmélet- és problémafüggő aspektusokat is figyelembe kell vennünk – ezzel viszont meghaladják az (ATSN)-t. Utalnak továbbá az elmélet- és problémafüggőség közötti potenciális konfliktusokra is. Előfordulhat ugyanis, hogy olyan adattípusok tűnnek relevánsnak a kutatás egy pontján, amelyeket az adott elmélet módszertana vagy nem ismer, vagy esetleg megtilt. Úgy vélik, ilyen esetekben a nyelvész „nyitottsága” viheti előre a kutatást: dönthet úgy is, hogy a felvetett probléma hatékony megoldása érdekében az ilyen adatokat is figyelembe veszi, és evidenciának tekinti. Tisztázatlanok maradnak azonban azok a metodológiai normák, amelyek alapján az adatok megbízhatóságának ellenőrzése elhelyezhető volna az elméletalkotás folyamatában. Reflektálatlan marad az a viszonyrendszer, amely az elmélet hipotézisei, metodológiai háttérfeltevési, az adatok elfogadhatósági kritériumai, valamint a tárgytudományos tevékenység folyamata között fennáll. PENKE és ROSENBACH felveti ugyan, hogy az elmélet hipotéziseinek az adatokkal történő konfrontációja adott esetben a metodológiai elvek revíziójához (tehát például az elfogadott adattípusok listájának módosításához) vezethet, azonban nem integrálják az adatok és az elmélet hipotézisei közötti kölcsönhatást az elméletalkotás folyamatába.

KEPESER és REIS hasonló álláspontot foglal el. Úgy vélik, hogy az evidencia kiválasztása a rendelkezésünkre álló adatok halmazából összetett és távolról sem triviális feladat. Nem hajtható végre pusztán általános metodológiai normák alapján, a konkrét nyelvészeti kutatómunkától függetlenül (2005: 1). Csakis az elmélet, valamint a konkrét probléma sajátosságainak függvényében dönthető el az, hogy mi minősül evidenciának. Ez utóbbi tényezők határozzák meg ugyanis, hogy „relevánsnak” és „adekvátnak” tekinthető-e egy-egy adat(típus), azaz alkalmas-e arra, hogy

alátámassza vagy cáfolja az elmélet valamely hipotézisét (i. m. 3). Mivel úgy vélik, hogy az adatok adekvát voltát csak egyedi esetekben, mindig az adott problémától függően lehet eldönteni, ezért az adekvátság tárgyalására nem térnek ki. Hangsúlyozzák ugyanakkor, hogy az adatok csupán akkor szolgálhatnak „igaz evidenciaként”, ha megbízhatók és reprodukálhatók (i. h.). Ezzel feltételezik, hogy létrehozható egy olyan empirikus bázis, amely lehetővé teszi, hogy bizonyos adatokat „igaz evidenciának” tekintsünk. Ez pedig azt jelenti, hogy KEPSER és REIS elképzeléseiben is tetten érhető egy hasonló kettősség, mint PENKE és ROSENBACH esetében. Elismerik az adatok komplexitásából eredő bizonytalanságát, fallibilitását, ennek ellenére az (ATSN) szellemében megkísérlik eme adatok közül kiválasztani azokat, amelyek az elmélet tesztelésének elméletfüggetlen és biztos alapját jelentik. KEPSER és REIS azt feltételezik, hogy az elméletalkotást meg kell előznie a megbízható és reprodukálható evidencia kiválasztásának, és az elméletalkotási folyamatot lineárisnak kell tekintenünk. Csakhogy ez az eljárás sem képes biztosítani az elméletalkotási folyamat szilárd empirikus bázisát. A bázis tagjainak kiválasztásához, azaz az adatok megbízhatóságát befolyásoló nyelvi és nem-nyelvi tényezők kontrolljához egy vagy több további empirikus elméletre lenne szükség – ezek a további elméletek viszont maguk is adatokra támaszkodnak, melyek megbízhatóságának ellenőrzéséhez újabb elméletre lenne szükség, és így tovább ad infinitum. Ráadásul, mint LEHMANN (KEPSER és REIS állásfoglalásától független kontextusban) hangsúlyozza, a fizikától és egyes egyéb természettudományoktól eltérően bizonyos típusú nyelvészeti kutatásokban az adatok nem reprodukálhatók, hanem egyediék és egyszeriek (2004: 197).

LEHMANN felfogása szintén plasztikusan szemlélteti az „evidencia” terminus használata körüli zavart. Először is – mint azt a 3. A) 5. szakaszban említettük – az adatok egyik funkcióját abban látja, hogy a „megkérdőjelezhetetlen evidencia” szerepét töltik be az érvelésben (vö. 2004: 181). Ezzel megfordítja az adat és az evidencia kategóriája közötti viszonyt: nem az evidenciát kell kiválasztanunk a rendelkezésünkre álló adatok halmazából, hanem az evidencia-státus lesz előfeltétele annak, hogy valami adat lehessen. Figyelemre méltó, hogy az említett idézetben LEHMANN az adatok evidencia voltát a mindenkor i é r v e l é s h e z viszonyítja. Ebből az következik, hogy nem az elméletek absztrakt logikai szerkezetének, hanem a r g u m e n t á c i ó s s z e r k e z e t é n e k kellene előtérbe kerülni, ami gyökeres fordulópontot jelenthetne a nyelvészeti elméletek metatudományos rekonstrukciójában. LEHMANN ezt a következtetést azonban nem vonja le, mivel az elméletalkotás folyamatában az adatokat az indukció, illetve a dedukció bázisának tekinti, és argumentációs szerepüket csupán az olvasó meggyőzését szolgáló retorikai eszközként értelmezi (vö. i. m. 191). Másodszor, az előző bekezdésben említett szöveghegyen az evidencia csupán az olvasó meggyőzését szolgáló retorikai eszközként jelenik meg. Ennek ellentmondva a 3. A) 5. szakaszban idézett passzus szerint pedig a „megkérdőjelezhetetlen evidencia szerepét” töltik be a nyelvész érvelésében. Harmadszor, LEHMANN az analitikus tudományelmélet standard nézetének felfogását követve az elméletek esetében a logikai szerkezetet tekinti relevánsnak, az argumentációs szerkezetet csupán az elmélet kifejtésében juttatja szerephez. Ennek viszont ellentmond az, hogy az adatok „megkérdőjelezhetetlen evidenciának” minő-

sülnek. Ez pedig azt jelenti, hogy LEHMANN felveti, de explicítanul hagyja a nyelvészeti elméletek logikai szerkezete és a nyelvészeti argumentáció viszonyát. Továbbra is az (SF)-beli nézetekhez közeli álláspontot foglal el, amennyiben az evidenciát – GEERAERTShez és KEPSEK–REISHOZ hasonlóan – elméletektől független, biztos empirikus bázisnak tartja.

Összegzés. Azt állapíthatjuk meg, hogy az evidencia fogalmának meghatározását illetően a szakirodalom távolról sem egységes. Egyrészt továbbra is jelen van az (SF) e)-ben összefoglalt szemlélet, csakúgy, mint az analitikus tudományelmélet standard nézetéből reflektáltanul átvett metodológiai elvek sora – mindenekelőtt az az elképzelés, amely szerint kiválasztható az adatok egy olyan csoportja, amelyet az elmélet tesztelése során biztos alapnak tekinthetünk. Másrészt jól kimutathatóan megjelennek az egyes felfogásokban olyan elemek, amelyek már nem egyeztethetők össze az (ATSN)-nel. Például releváns módon kerül előtérbe az a felismerés, hogy az evidencia elmélet- és problémafüggő, valamint az is, hogy a többé-kevésbé bizonytalan adatok és a hipotézisek közötti viszony jóval árnyaltabb annál, mint hogy az adatok igazolják vagy azonnal és véglegesen cáfolják a hipotéziseket. Teljes mértékben hiányoznak azonban azok a metodológiai elvek, amelyek az előtérbe állított új szempontok metatudományos explicációját lehetővé tennék.

4. Következtetések. – Az 1. szakaszban felsorolt szakirodalom elemzése többek között két tényre világított rá. Egyfelől az egyes jelenlegi álláspontokban továbbra is jelen vannak az (ATSN)-ből származó elemek. Így a nyelvészeti evidenciát változatlanul az adatok olyan kitüntetett részhalmazának tekintik, amely nem revideálható, hanem biztosan igaz kijelentésekből áll. Ezen kijelentések feladata a hipotézisek tesztelése (verifikációja, konfirmációja, falszifikációja). Másfelől ugyanezekben az álláspontokban világosan kimutathatók olyan árnyalt és progresszív felvetések, amelyek az (ATSN)-nel már nem egyeztethetők össze, és egyértelműen jelzik az attól való elszakadás igényét. Differenciálják, relativizálják és gyengítik a nyelvészeti adatokkal és evidenciával szemben felállított követelményeket.

A differenciálás azt jelenti, hogy nem egyetlen adatfajta tulajdonságaiból kiindulva kísérlük meg specifikálni a nyelvészeti adat fogalmát. Elismerik a nyelvészeti adatok/evidencia sokféleségét, azt, hogy nem kizárólag az adat/evidencia egyetlen típusa lehet legitim. Egy olyan gazdag szempontrendszer kidolgozására törekednek, amely alapján feltárható a különböző adatfajták belső szerkezete, és megalkotható a különböző nyelvészeti elméletekben alkalmazott adatok/evidencia tipológiája. A nyelvészeti adatokat komplex entitásoknak tekintik, melyek megbízhatóságát szerkezetük és forrásuk sajátosságai befolyásolják.

A relativizálás arra vonatkozik, hogy a nyelvészeti adatok alkalmazhatóságának megítélését nem tekintik episztemológiailag függetlennek az adott nyelvészeti elmélettől, hanem kiemelik az adatok elmélet- és problémafüggő voltát.

Gyengítésen azt értjük, hogy az (ATSN)-től eltérően nem csupán a tér- és időbeli koordinátákkal rendelkező, közvetlenül adott nyelvi megnyilatkozásokat rögzítő kijelentések tekinthetők empirikus adatnak, hanem a megnyilatkozások különböző szintű – a beszédeseményt elkerülhetetlenül idealizáló és redukáló – absztrakciói is. Továbbá megengedhető a falszifikációs kritérium enyhítése, és fel-

vetődik a különböző adattípusok alkalmazásából eredő ellentmondások ideiglenes tolerálhatósága is.

Következésképpen az említett kettősség fényében általános tanulságként csupán az (ATSN)-től és az (SF)-től való elszakadás igényét fogalmazhatjuk meg, de nem beszélhetünk egy ezeket felülíró, domináns szemléletről. Az (SF)-et nem egy új szemléletmód váltja fel, hanem egy *k é t a r c ú t e n d e n c i a*. E kétarcú tendencia egy sor különböző nézetet foglal magában, amelyek közös vonása részben az, hogy számos ponton kötődnek az (SF)-hez és az (ATSN)-hez, részben – olykor ellentmondásosan és következtelenül – releváns új felismerések megfogalmazásával egyúttal túl is kívánnak lépni az utóbbiakon.

(P1) bemutatott megoldási javaslatainak ezt a kettősségét (MP1)-ben összegezzük:

(MP1)

- a) A nyelvészeti evidenciát az (ATSN) és az (SF) szellemében továbbra is az adatok olyan kitüntetett részhalmazaként értelmezik, amely az adott elméletben nem revideálható, hanem biztos igazságként kezelendő, és amelynek feladata az elmélet hipotéziseinek tesztelése.
- b) Az (SF)-től eltérően a bemutatott nézetek egy sor progresszív és innovatív felismerést fogalmaznak meg. E felismeréseket nem minden szerző fogadja el, vagy ha közülük mégis többen, akkor eltérő módon és eltérő jelentőséget tulajdonítva nekik. Ennek következtében az új felismerések nem általánosíthatók. Ugyanakkor kiemelhetők a következő meglátások, amelyek egy vagy több írásban szerepelnek és – az (ATSN)-hez és az (SF)-hez viszonyítva – egyértelműen előremutató módon jelzik a differenciálás, relativizálás és gyengítés említett folyamatát:
 - α) Nem egyetlen adattípus preferálandó, hanem megengedhető a nyelvészeti adatok sokfélesége. Kívánatos több adattípus kombinációja egy adott kutatáson belül.
 - β) Az adatok eredete nem döntő jelentőségű, hanem csupán – az adatszerkezet, a komplexitás, a közvetlenség és az absztrakció stb. mellett – a releváns tényezők egyike.
 - γ) Minden adattípus problematikus, és felveti szerkezetük, valamint funkciójuk tisztázásának szükségességét.
 - δ) Az adatok és az elmélet közötti viszony nem lineáris, hanem ciklikus folyamatra épül.
 - ε) A nyelvészeti adatok elmélet- és problémafüggők. Az evidenciaként kezelt adatok rendszerint nem támasztják alá vagy cáfolják egyértelmű módon az elmélet hipotéziseit. Egyrészt nem fogadhatók el teljes bizonyossággal, másrészt ellentmondásokat generálhatnak.
- c) Reflektálatlan marad ugyanakkor a) és b) viszonya, valamint azok a kritériumok, amelyek a b)-ben összefoglalt álláspontok alkalmazhatóságát lehetővé tennék.

Noha a 3. szakaszban bemutatott áttekintés jelzi az adat- és evidenciaprobléma jelenlegi megközelítéseiben a problémák és megoldási lehetőségeik differenciálására való törekvést, rávilágít egy alapvető nehézségre is: arra, hogy (P1) *s z i s z t e m a* -

tikus tudományelméleti reflexió nélkül nem oldható meg, miközben a tárgyalt írásokban továbbra is tisztázatlan a tudományelméleti reflexió szerepe a nyelvészeti elméletalkotásban.

Ugyanis azt láttuk, hogy a nyelvészeti adatok és evidencia mibenlétére vonatkozó feltevések jelentősen eltérnek egymástól. Nincs egyetlen helyes, minden nyelvész számára közvetlenül elfogadható jellemzése az adatok szerkezetének, típusainak, funkciójának. Amint a nyelvész elgondolkodik azon, hogy ő maga mit is tekint adatnak és evidenciának, azonnal egymástól eltérő, olykor egymásnak ellentmondó nézetek és megoldatlan problémák sokasága bontakozik ki. Ez a tény arra utal, hogy nem elegendő a nyelvész naiv önreflexiója, hanem szükség van a tudományelmélet eszközeinek tudatos és szisztematikus alkalmazásával elvégzett vizsgálatokra is. A nyelvészeti adatok szerkezete és funkciója ugyanúgy tudományos kutatás tárgyát kell, hogy képezze, mint a társadalom és a természet, vagy maga a nyelv. Az a kérdés, hogy mi a nyelvészeti adat, és milyen funkciói vannak, ugyanúgy tudományos kérdés, mint az, hogy mi a nyelvi rendszer, melyek egy adott nyelv grammatikájának szintaktikai szabályai, ige-e vagy melléknév egy bizonyos nyelv bizonyos szavainak egy csoportja. Ahogy nem elégséges az anyanyelvi beszélő naiv nyelvi kompetenciája a nyelv szabályainak feltárásához, hanem a nyelvtudomány módszereinek kidolgozása és alkalmazása is szükséges, ugyanúgy nem elégséges a nyelvész naiv önképe ahhoz, hogy megmondhassuk, a nyelvészetben mi lehet és mi nem lehet adat/evidencia, hanem szükség van tudományos – azaz ebben az esetben *metatudományos* – módszerek szisztematikus kidolgozására és alkalmazására is.

Ugyanakkor a bemutatott nézetek e kíváncsiaknak nem felelnek meg. Az adat/evidencia-problémával foglalkozó szakirodalom minden vizsgált elemén végigvonnul egy *második alapvető kettősség* is, amely a metatudományos reflexiónak tulajdonított szerepet jellemzi. Egyfelől a szerzők felismerik a metasztintú reflexió szükségességét. Másfelől azonban ezt a metareflexiót naiv, inkohérens és következtelen módon hajtják végre. A nyelvészeti kutatómunka gyakorlatából levont következtetések alapján belátják mind az (SF) tarthatatlanságát, mind pedig azt, hogy a nyelvészeti elméletalkotás egészen más módon működik, mint ahogy azt az analitikus tudományelmélet standard nézetének mechanikus nyelvészeti adaptációja sugallja. Azonban azt nem ismerik fel, hogy az (ATSN) tarthatatlanságából alternatív tudományelméleti támpontok keresésének a szükségessége következik. Ezért hol explicit deklarációk formájában, hol implicit háttérfeltevésekként folyamatosan vissza-visszatérnek az analitikus tudományelmélet dogmáihoz, szembevető ellentmondásokat generálva köztük és a nyelvészeti kutatómunka gyakorlatából levont következtetések között.

Ily módon eddigi elemzéseink végkövetkeztetéseként a következő választ kapjuk (P2)-re, mely tükrözi az imént vázolt második kettősséget:

(MP2)

- a) Az elemzett nézetek *felismerik* a nyelvészeti adatokra és evidenciára való metatudományos reflexió szükségességét.
- b) Ez a felismerés azonban nem lép túl a nyelvész naiv (ön)reflexióján. Noha az (ATSN) folyamatos, részben explicit, részben implicit konfrontációja a kutatói gyakorlattal utal az előbbi tarthatatlanságára, a be-

mutatott írások még sem ismerik fel alternatív tudományelméleti eszközök kidolgozásának szükségességét.

Ugyanakkor az a tény, hogy – mint azt elemzéseinkben kimutattuk és az (MP2) b)-ben összefoglaltuk – széles szakadék tátong a tudományos kutatómunka gyakorlata és az analitikus tudományelmélet tételei között, nem kizárólag a nyelvtudományra jellemző, hanem olyan általános probléma, amely legitimálja az adatok/evidencia új metatudományos modelljeinek kidolgozását. PETER ACHINSTEIN például az evidenciáról szóló, meghatározó jelentőségű művének bevezető fejezetében ezt írja: „[...] az evidenciával foglalkozó standard filozófiai elméleteket a tudósok figyelmen kívül hagyják (és figyelmen kívül is kell hagyniuk). Figyelmen kívül kell hagyniuk, mert ezen filozófiai elméletek evidencia-fogalmainak háttérfeltevései nem egyeztetetők össze a tudósok azon háttérfeltevéseivel, amelyekre támaszkodva a hipotézisek melletti evidenciáról beszélnek és a hipotéziseket evidenciákkal támasztják alá [...]” (2001: 3; kiemelés: K. A. – R. Cs.).

Az (MP1)-ben és az (MP2)-ben összefoglalt két ambivalencia kellőképpen motiválja egy olyan metatudományos modell szisztematikus kidolgozását, amely túllép az analitikus tudományelmélet standard nézetének korlátain, és a két ambivalenciát tudományelméleti szempontból igényesen, egyúttal pedig a nyelvészeti kutatómunka gyakorlatához igazodó módon oldja fel. E kíváncsnak megfelelően kezdeményez egy új programot KERTÉSZ–RÁKOSI (2008d) a jelen írás közvetlen folytatásaként.

A hivatkozott irodalom

- ACHINSTEIN, PETER 2001. *The book of evidence*. Oxford University Press, Oxford.
- ALLAN, KEITH 2003. Linguistic metatheory. *Language Sciences* 25: 533–60.
- ANDOR JÓZSEF 2004. A Mester és performanciája: Interjú Noam Chomsky-val. *Alkalmazott Nyelvtudomány* 4: 91–109.
- BORSLEY, ROBERT D. 2005. Introduction. In: BORSLEY ed. 2005: 1475–80.
- BORSLEY, ROBERT D. ed. 2005. Data in theoretical Linguistics [= *Lingua* 115/11: 1475–665].
- CHOMSKY, NOAM 1965. *Aspects of the theory of syntax*. MIT Press, Cambridge, MA.
- CHOMSKY, NOAM 1969. [1957.] *Syntactic structures*. Mouton, The Hague–Paris.
- CHOMSKY, NOAM 1969. Language and philosophy. In: HOOK, S. ed. *Language and philosophy: A symposium*. New York University Press, New York. 51–94.
- CHOMSKY, NOAM 2002. *On nature and language*. Cambridge University Press, Cambridge.
- DUHEM, PIERRE 1974. [1906.] *The aim and structure of physical theory*. Atheneum, New York.
- EISENBEIß, SONJA 1994. Elizitation von Nominalphrasen und Kasusmarkierungen. In: EISENBEIß, SONJA – BARTKE, SUSANNE – WEYERTS, HELGA – CLAHSSEN, HARALD eds., *Elizitationsverfahren in der Spracherwerbsforschung: Nominalphrasen, Kasus, Plural, Partizipien*. Arbeiten des Sondervorschungsbereichs 282, 57. Heinrich-Heine-Universität, Düsseldorf. 1–38.
- FEHÉR MÁRTA 2000. Thomas Kuhn tudományfilozófiai „paradigmája”. In: KUHN 2000 [1970]: 233–50.
- GOODMAN, NELSON 1983 [1955]. *Fact, fiction, and forecast*. Harvard University Press, Cambridge, MA.

- GEERAERTS, DIRK 2006. Methodology in cognitive linguistics. In: KRISTIANSEN – ACHARD – DIRVEN – MENDOZA IBÁÑEZ eds. 2006: 21–49.
- HANSON, NORWOOD R. 1958. Patterns of discovery. Cambridge University Press, Cambridge.
- HEMPEL, CARL G. 1952. Fundamentals of concept formation in empirical science. University of Chicago Press, Chicago, IL.
- HOWSON, COLIN 2000. Evidence and confirmation. In: NEWTON-SMITH, WILLIAM H. ed., A companion to the philosophy of science. Blackwell, Oxford. 108–16.
- ITKONEN, ESA 1978. Grammatical theory and metascience. Benjamins, Amsterdam–Philadelphia.
- KALLMEYER, WERNER – ZIFONUN, GISELA eds. 2007. Sprachkorpora – Datenmengen und Erkenntnisfortschritt. De Gruyter, Berlin–New York.
- KELLY, THOMAS 2006. Evidence. Stanford Encyclopedia of Philosophy. <http://plato.stanford.edu/entries/evidence/>.
- KEPSEK, STEPHAN – REIS, MARGA 2005. Evidence in linguistics. In: KEPSEK–REIS eds. 2005: 1–6.
- KEPSEK, STEPHAN – REIS, MARGA eds. 2005. Linguistic evidence. Empirical, theoretical and computational perspectives. De Gruyter, Berlin–New York.
- KERTÉSZ, ANDRÁS 1991. Die Modularität der Wissenschaft. Konzeptuelle und soziale Prinzipien linguistischer Erkenntnis. Vieweg, Braunschweig.
- KERTÉSZ, ANDRÁS 2004. Philosophie der Linguistik. Studien zur naturalisierten Wissenschaftstheorie. Narr, Tübingen.
- KERTÉSZ, ANDRÁS – RÁKOSI, CSILLA 2006. Inconsistency and plausible reasoning in an analysis of German affricates. A case study in the philosophy of linguistics. Language Sciences 28: 386–423.
- KERTÉSZ, ANDRÁS – RÁKOSI, CSILLA 2008a. Conservatism vs. innovation in the (un)grammaticality debate. In: KERTÉSZ–RÁKOSI eds. 2008c: 61–84.
- KERTÉSZ, ANDRÁS – RÁKOSI, CSILLA 2008b. Conservatism vs. innovation in the debate on data in generative grammar. In: KERTÉSZ–RÁKOSI eds. 2008c: 85–118.
- KERTÉSZ, ANDRÁS – RÁKOSI, CSILLA eds. 2008c. New approaches to linguistic evidence. Pilot Studies. Lang, Frankfurt am Main – Bern – New York – Oxford – Wien.
- KERTÉSZ, ANDRÁS – RÁKOSI, CSILLA 2008d. Data, evidence and plausible argumentation in linguistics. Előkészületben.
- KRISTIANSEN, GITTE – ACHARD, MICHEL – DIRVEN, RENÉ – MENDOZA IBÁÑEZ, FRANCISCO J. RUIZ DE eds. 2006. Cognitive linguistics: Current applications and future perspectives. De Gruyter, Berlin–New York.
- KUHN, THOMAS S. 2000 [1970]. A tudományos forradalmak szerkezete. Osiris, Budapest.
- LAKI JÁNOS 1998. Empirikus adatok, metodológia, gondolkodás és nyelv a XX. századi tudományfilozófiában. In: LAKI JÁNOS szerk., Tudományfilozófia. Osiris – Láthatatlan Kollégium, Bp. 7–32.
- LAKI JÁNOS 2006. A tudomány természete. Thomas Kuhn és a tudományfilozófia történeti fordulata. Gondolat, Bp.
- LEHMANN, CHRISTIAN 2004. Data in linguistics. The Linguistic Review 21: 175–210.
- LEHMANN, CHRISTIAN 2007. Daten – Korpora – Dokumentation. In: KALLMEYER–ZIFONUN eds. 2007: 9–27.

- LEMNITZER, LOTHAR – ZINSMEISTER, HEIKE 2006. *Korpuslinguistik. Eine Einführung*. Narr, Tübingen.
- LÜDELING, ANKE 2006. Zum Geleit. In: LEMNITZER–ZINSMEISTER 2006: 4.
- MCENERY, TONY – WILSON, ANDREW 1996. *Corpus linguistics. An introduction*. Edinburgh University Press, Edinburgh.
- MEHEUS, JOKE ed. 2002. *Inconsistency in science*. Kluwer, Dordrecht.
- MEREU, LUNELLA 2004. Linguistic data as complex items. *The Linguistic Review* 21: 211–33.
- MORAVCSIK, EDITH 1993. Why is syntax complicated? In: EID, MUSHIRA – IVERSON, GREGORY K. eds. *Principles and prediction: The analysis of natural language*. Benjamins, Amsterdam–Philadelphia. 73–92.
- MORAVCSIK, EDITH 2006. *An introduction to syntactic theory*. Continuum, London–New York.
- PENKE, MARTINA – ROSENBACH, ANETTE 2004. What counts as evidence in linguistics? In: PENKE–ROSENBACH eds. 2004: 480–526.
- PENKE, MARTINA – ROSENBACH, ANETTE eds. 2004. What counts as evidence in linguistics? [= *Studies in Language* 28/3: 481–747].
- PENKE, MARTINA – ROSENBACH, ANETTE 2007. Preface. In: PENKE–ROSENBACH eds. 2007: vii–ix.
- PENKE, MARTINA – ROSENBACH, ANETTE eds. 2007. *What counts as evidence in linguistics?* Benjamins, Amsterdam–Philadelphia.
- PERRY, THOMAS A. ed. 1980. *Evidence and argumentation in linguistics*. De Gruyter, Berlin–New York.
- POPPER, KARL 1959. *The logic of scientific discovery*. Hutchinson, London.
- RÁKOSI CSILLA 2005. *Nyelvészet és argumentációelmélet*. DE Kossuth Egyetemi Kiadó, Debrecen.
- REICHENBACH, HANS 1938. *Experience and prediction: An analysis of the foundations and the structure of knowledge*. The University of Chicago Press, Chicago, IL.
- REIS, MARGA et al. 1998. *Linguistische Datenstrukturen: Theoretische und empirische Grundlagen der Grammatikforschung*. Tübingen, Universität Tübingen.
- RINGEN, JON D. 1975. Linguistic facts. A study of the empirical scientific status of transformational generative grammars. In: COHEN, DAVID – WIRTH, JESSICA R. eds., *Testing linguistic hypotheses*. Hemisphere, Washington. 1–41.
- SAMPSON, GEOFFREY 1975. *The form of language*. Weidenfeld and Nicholson, London.
- SAMPSON, GEOFFREY 2007. Reply. In: STEFANOWITSCH–GRIES eds. 2007: 111–29.
- SCHÜTZE, CARSON T. 1996. The empirical base of linguistics. Grammaticality judgments and linguistic methodology. The University of Chicago Press, Chicago–London.
- SIMONE, RAFFAELE 2004. The object, the method, and the ghosts. Remarks on a terra incognita. *The Linguistic Review* 21: 235–56.
- STEFANOWITSCH, ANATOL – GRIES, STEFAN TH. eds. 2007. Grammar without grammaticality [= *Corpus Linguistics and Linguistic Theory* 3/1: 1–129].
- STEGMÜLLER, WOLFGANG 1970. *Theorie und Erfahrung*. Springer, Berlin–Heidelberg–New York.
- STERNEFELD, WOLFGANG ed. 2007. Data in generative linguistics [= *Theoretical Linguistics* 33(3): 269–413].
- WUNDERLICH, DIETER ed. 1976. *Wissenschaftstheorie der Linguistik*. Scriptor, Kronberg.

Remarks on the current discussion on linguistic data and evidence

The twentieth century saw a series of closely related and inextricably intertwined methodological debates about the problem of the empiricalness of linguistic theories. From the second part of the nineties, the methodological debates on data and evidence were gradually extended both at an object- and at a meta-theoretical level, and were enriched with new elements. Today they affect several different research fields and branches of linguistics. The contemporary discussion no longer centres on the problem of whether linguistic theories should be empirical or not. Rather, the following questions are focused on: What types of data may be used, what data count as evidence, and what role can be attributed to them in different fields of linguistic theorising? The aim of the present paper is the systematization and the critical analysis of current answers to these questions. From the findings of the analyses conclusions will be drawn that are expected to pave the way for the future solution of the problems raised in the discussion at issue.

ANDRÁS KERTÉSZ – CSILLA RÁKOSI