

- TÁTRAI SZILÁRD 2005. A nézőpont szerepe a narratív megértésben. *Általános Nyelvészeti Tanulmányok* 21: 207–29.
- TÁTRAI SZILÁRD 2011. Bevezetés a pragmatikába. Funkcionális kognitív megközelítés. Tinta Könyvkiadó, Bp.
- TOLCSVAI NAGY GÁBOR 1995. A stílus diszkurzív elmélete felé. *Helikon* 41: 219–31.
- TOLCSVAI NAGY GÁBOR 2005. Funkcionális nyelvtan: elmélet és gyakorlat. *Magyar Nyelvőr* 348–62.
- TOLCSVAI NAGY, GÁBOR 2005. A cognitive theory of style. Peter Lang, Frankfurt am Main.
- WARDHAUGH, RONALD 2005. Szociolingvisztika. Osiris Kiadó, Bp.
- WOLFRAM, WALT 1998. Dialect in society. In: COULMAS ed. 1998: 107–26.

About style definitions in sociolinguistics

This paper scrutinizes the various definitions of the notion of ‘style’ in sociolinguistics. In this examination, the vantage point is the Labovian notion of ‘contextual style’. The paper deals in detail with the definitions of style in the variationist paradigm, with the functionalist notion of ‘register’ as compared to the concept of style, with Bell’s model and with the approaches to style in Hudson’s and Coupland’s alternative sociolinguistic paradigms. As a result of this investigation, we recommend a comprehensive approach to style in sociolinguistics, one that is compatible with the functionalist approach to language.

Keywords: functional linguistics, contextual style, variational paradigm, register, audience design, referee design, item-based paradigm.

SIMON GÁBOR

A nyelvtörténeti adat: elvek, gyakorlat, lehetőségek*

A történeti mondatok kutatója számára valószínűleg a legfőbb hiány, hogy nincs, aki közölje vele: ezt így nem mondhatjuk.

(JAN TERJE FAARLUND)

1. „A magyar nyelv nagyszótára”-ban ekként szerepel az *adat* szó első jelentésének megadása: ‘valakinek, valaminek a megismeréséhez, jellemzéséhez, megvilágításához hozzásegítő, esetleg ilyen szándékkal közölt és nyilvántartott tény, információ, részlet, adalék’. A további jelentésmegadások sorában pedig többek között ez is olvasható: ‘(a tudományos nyelvben): valamely kísérlet, vizsgálat, mérés következtében megállapított, esetleg (mérő)műszerről leolvasott számszerű eredmény, tény’. A szó története alapján világos továbbá, hogy az etimológiája szerint eredendően ‘ami adott, ami van’ jelentésű szó a nyelvújítás korától létezik

* A cikk egy része elhangzott az MTA Nyelvtudományi Intézetében 2010. szeptember 30-án–október 1-jén rendezett „A nyelvtörténeti adat: érvény és értelmezés” című konferencián.

a latin > német *datum* szó analógiájára; a fentebb idézett mai jelentései az 1830-as évektől adathozhatóak. (Egyébként a nyelvújítás korában nemcsak ezekben a jelentésekben vezették be a szót, hanem egy ideig például az iratok keletkezését jelölő sor *dátum* nevét is *adat*-ra magyarították.) A „Nyelvújítási szótár” megjegyzi: 1864-ben BALLAGI MÓR már az *adathoz* (‘dokumentieren’) igét is szótárta „Új teljes magyar és német szótár”-ában.

Ha akár csak a bevezetőben idézett jelentéseket vesszük is szemügyre – valamint magától értetődőnek tekintjük, hogy a tudományos és a köznyelvi használat nem feltétlenül válik mereven szét –, látjuk, hogy adathoz, adatokra hivatkozni (akár a nyelvtudományban is) többféleképpen lehetséges. Éppen ez a többféleség és a belőle fakadó tudománymetodológiai, tudományfilozófiai kérdések azok, amik talaján az utóbbi két évtizedben feltűnően megszorodtak az olyan kutatások, tanulmányok, amelyek többek között a nyelvészeti adat státuszával, fajtáival, szerepeivel, forrásaival, legitimitásának kérdéseivel foglalkoznak.

2. Ezek a munkák azt is mutatják, hogy a vita, amely a különböző nyelvészeti iskolák egymástól eltérő adathozási módszereit érintette, újabb szakaszához érkezett. Az utóbbi éveknek az adathozás elveivel foglalkozó tanulmányaiban, úgy tűnik, nem feltétlenül a direkt vitapozíció az elsődleges, sokkal inkább a különböző indítatású iskolák szempontjainak egymással párhuzamos áttekintése, az ezeken belüli kérdések és válaszok adekvátságának vizsgálata, a nyelvi tények megragadásának módszertani vonatkozásai.

A) Bizonyára nem véletlen, hogy azt a tanulmányt, amely a legkörültekintőbben és legrészletesebben elemzi a nyelvészeti adatokkal kapcsolatos kérdéseket, egy olyan kutató írta, aki maga is aktív adathozó és -felhasználó. A grammatikalizáció korai monográfusa, paramétereinek felállítója, CHRISTIAN LEHMANN „Data in linguistics” című cikkében (LEHMANN 2004) egyszerre elvi igénnyel és szinte szakmai népszerűsítő elkötelezettséggel elemzi és mutatja be többek között az adat fajtáit, az adathozás módozatait és az adat érvényét érintő problémákat.

Az adathoz mindenképp leválasztja a benne tükröződő tényről, kifejtve, hogy adat nem egyszerűen az, ami adott – ahogyan az angol (s tegyük hozzá a fentebbiek alapján: a magyar) szó eredeti jelentése alapján vélhetnénk –, hanem annak mentális reprezentációja (‘Darstellung’, azaz ‘ábrázolás’, pontosítja némely az angolul kétértelmű „reprezentáció” szó jelentését). Hiszen a tudományok tárgyerőlete nem a külvilágban adott fizikai valóság, hanem annak a tudományos kommunikációba beemelt adaptációja (i. m. 4–5). Adat nem létezik önmagában: egy adathoz a szerepe avat adathoz – valaki vagy valakik (a tudományos közösség) számára. S hogy mi ez a szerep? Induktív alapanyag empirikus kutatásokhoz és hipotézisek felállításához, illetve a deduktíve megalkotott szabály tesztelésére szolgáló bázis (i. m. 7–8). Ahhoz, hogy ezeket a funkcióit betölthesse, az elemzőtől függetlennek kell lennie, vagyis külső forrásból kell származnia, hiszen körköröséghez vezethet, ha az introspektíve megkonstruált adatok alapján felállított hipotézist ugyancsak introspektív adatokkal teszteljük (i. m. 23). Legmegbízhatóbbak – mivel ezek a legkevésbé manipulálhatóak – a „talált adatok”

(vagyis a gyűjtött adatok), ezeket azok a megnyilatkozások követik, amelyek olyan beszédeseményekből származnak, amelyeket a kutató nem nyelvi eszközökkel váltott ki. A többi adatforrást és adatolási módot az empirikus kutatásban a szerző csak kiegészítő eljárásként tartja elfogadhatónak (i. m. 24–6 stb.). Számolni kell ugyanakkor azzal is – vélekedik LEHMANN –, hogy mindenféle reprezentáció torzít az úgynevezett nyers adathoz képest, amennyiben például elvonatkoztat, és a nyers adatnak bizonyos, a kutató számára fontos vonásait kiemeli, míg más jellemzőit a háttérben hagyja (i. m. 35–6).

A szerző, amikor a nyelvészet filozófiai alapjai és tudománytörténete felől tekinti át az adatkezeléshez való viszonyulást, három ívet rajzol fel (i. m. 15–6). 1. Aki a nyelvészethez (Platóntól és Arisztoteléstől Chomskyig) mint logikai tudományhoz közelít, nincs szüksége adatra, hiszen azt externalizálja, amit a saját elméjében megalkot; legfeljebb illusztratív példákat használ. – 2. Aki a nyelvészethez (Dionysios Thraxtól és Appolonios Dyskolostól napjaink filológiájáig) a hermeneutika felől közelít, olyan adatokkal dolgozik, amelyek alapvetően tartalmi okokból fennmaradt nyelvemlékekből származnak, tehát nyelvi anyaguk sokban esetleges. – 3. A nyelvészetet empirikus tudományként megközelítők számára (Franz Bopptól és a történeti-összehasonlító iskolától Saussure-ig) a nyelvészet adatfüggő tudomány, mind az indukcióban, mind a dedukcióban. Az adatokkal kapcsolatos tudatosság először a terepmunkát alkalmazó kutatási ágazatokban fejlődött ki az előző századfordulón, melyhez majd sokban hozzájárult a szociolingvisztika és a pszicholingvisztika is. (Csak zárójelben megjegyezve: ha LEHMANN elfogadhatónak tartja, hogy a nyelvészeti adatok nem mindig reprodukálhatók – ahogy tanulmányában írja (i. m. 24) –, talán nem is tűnik olyan élesnek a határ az általa empirikusnak, illetve hermeneutikusnak nevezett közelítésmód alapanyaga kínálta lehetőségek között.)

Miközben LEHMANN az adatoláshoz kapcsolódó témákat itt még csak fel sem idézhető részletességgel tárgyalja elvi síkon, azt is érzékelteti: a nyelvész számára, amikor adott kutatásán dolgozik, nem ezek az elsődleges kérdések. A kutató fő gondja (és nemcsak a nyelvészetben, hanem több más empirikus tudományban is) általában az – írja cikkének egyik helyén (i. m. 19–21) –, hogy nem állnak rendelkezésére a kutatáshoz szükséges adatok, így ezekről neki kell gondoskodnia; s míg az elsődleges adatok terén ez mennyiségi kérdés, az ezekre épülő másodlagos adatok esetében már egyszerre mennyiség és minőség dolga is: például a tipológia és az univerzálékutatás számos kérdésére nem adható kielégítő válasz a meglévő grammatikai leírások alapján.

Amikor a szerző olyan példát mutat be, mint az alábbi – annak a kutatók nagy része számára magától értetődő dolognak a szemléltetésére, hogy ami az egyik vizsgálatban az adatelemzés eredménye, a másik vizsgálatban adatul szolgálhat, hogy tehát szintek egymásra épülnek, és az adat mibenléte az adott kutatás céljaitól függ –, indíttatása talán nem más, mint a gyűjtött adatok felhasználását elutasító irodalmakon iskolázottaknak szánt figyelemfelhívás. A nyelvész készít egy hangfelvételt – ez az ő adata; ortográfiai átíratát mint kutatási eredményét közli. B nyelvész ezt az írásos reprezentációt adatként használja, amikor az adott nyelv grammatikáját vizsgálja; s amire jut, azt saját kutatási eredményeként publi-

kálja. C nyelvész, aki tipológus, és adatforrása a grammatika, B leíró grammatikai állításait használja adatként, amikor beépíti őket adatbázisába, hogy segítségükkel általánosításokat fogalmazzon meg, amelyek már C kutatási adataiként jelennek meg a tudományos kommunikációban (i. m. 8).

B) Az adatolással foglalkozó tanulmányok sorában olyan munkák is születtek, amelyek szerzői egy-egy részfejezetben tudománytörténeti igénnyel tekintik át és elemzik az adatolással kapcsolatos korábbi eljárásokat, vélekedéseket. Ilyen például JAKUES DURAND egyik cikke (DURAND 2009), amely Chomsky 1968-ban napvilágot látott „Nyelv és elme” című munkájától kezdve nyújt összefoglalást a témáról. A chomskyánus háttérű, de korpuszadatokkal dolgozó DURAND a jelenhez érkezve így összegez (i. m. 17): ma már sok munka mellett érvel, hogy adatok nélkül nem alkotható elmélet. Továbbá: nem szükségszerű, hogy kibékíthetetlen ellentét legyen a racionalista és az empirikus nézőpont között; a korpuszokon alapuló kutatások egyáltalán nem feltétlenül antikognitivisták irányultságúak. Ebből a tanulmányból is érezhető némi meggyőző szándék, amikor például a szerző arról ír: ma már sokan tudják, hogy adaton nem a nyelvész noteszába beírt privát bejegyzést kell érteni, hanem azt a nyersanyagot, amelyen állnak vagy buknak az elméletek. – Ez a megállapítás több szempontból is eszünkbe juttathatja CSERESNYÉSI LÁSZLÓ egyik ironikusra hangolt megjegyzését: „Mindnyájan fel tudjuk idézni magunkban a tudományelőtti (azaz pregeneratív) korszakból itt ragadt, olcsó, műanyagkeretes szemüvege mögül pislogó, szövegekkel bíbelődő, cédulázgató úr, illetve a már a kognitív nyelvészeti agytudomány korszakába lépett, Toshiba dynabookon dolgozó, áramvonalas nyelvtudós imázsát. Mégis: talán nem túl szerencsés a nyelvészetet »preteoretikus« és »teoretikus« korszakra osztani.” (CSERESNYÉSI 2009: 406.)

KERTÉSZ ANDRÁS és RÁKOSI CSILLA „Megjegyzések a nyelvészeti adatok és evidencia problémájáról folyó vita jelenlegi állásához” című tanulmánya azokat a nézeteket összegzi, amelyek – a korábbi évtizedek talaján – az utóbbi körülbelül tizenöt évben az adattal kapcsolatos vitákban kialakultak. (KERTÉSZ–RÁKOSI 2008a: 274–86, 385–401; ugyanez a tanulmány kisebb változtatásokkal megtalálható a KERTÉSZ–RÁKOSI 2008b bevezető blokkjaként.) A kiterjedt szakirodalmi anyagot elemző munka a 60-as éveket megelőző korszaktól kezdve tekinti át az egyes elméleteknek az adathoz és az „evidenciához” való viszonyulását. (A szakirodalomban általában *evidencia*-nak nevezik – az angol *evidence* nyomán – az elmélet tesztelését szolgáló biztos adatok kategóriáját.¹) Az 50-es–60-as évektől kezdve szemben állt egymással az induktivitásra, illetve a deduktivitásra építő nyelvtudomány eszméje, s a kétféle meghatározó megközelítésmód vitája a következő évtizedeken is végigvonult (i. m. 278 stb.). Az egyik oldalon a generatív nyelvészek, akik az introspekció híveiként érdektelennek tartották a nyelvhasználatból vett, nagy mennyiségű adat gyűjtését és elemzését. A másik oldalon a korpusznyelvészek, akik ezzel szemben az adatokból kiinduló indukciónélkülözhetetlensége mellett érveltek. „Ha a nyelvészet valóban az intuíción alapulna,

¹ Vö.: Erős nyelvi adat, amely következtetések megtámogatására vagy egy másik elképzelhető következtetés opponálására alkalmas (TRASK 2000: 111).

nem lenne tudomány” – idézik a szerzők (i. m. 279) az utóbbiak egyik képviselőjét (SAMPSON 1975: 60).

KERTÉSZ és RÁKOSI a két, korábban élesen szembenálló nézetrendszer közös tulajdonságait a következőkben általánosítja (i. m. 279–80): 1. csupán bizonyos adattípust tekintenek legitimnek (introspekció/korpuszadat); 2. mindkét rendszerben kizárólag az adatok eredete dönti el, mi számít megfelelő adattípusnak; 3. saját adatkezelési eljárását mindkét tábor problémamentesnek ítéli; 4. az adatok és a hipotézisek közötti viszonyt mindkét megközelítésmód hívei egyirányúnak tételezik (dedukció/indukció); 5. az evidenciát mindkét oldalon kitüntetett szerepű adatnak tekintik.

A szerzők a továbbiakban megmutatják, hogy ma már a nyelvészeti kutatómunka gyakorlata egyre kevésbé követi mereven a két, előzőleg határozottan elkülönülő álláspont valamelyikét, és egyre inkább szándékot mutat a differenciált adatfelhasználásra. Ennek egyik példája, hogy ma már sok generatív nyelvész olyan kutatások eredményeit is felhasználja, amelyek nem introspektív adatok elemzésén alapulnak, és a korpuszal dolgozók is gyakran építenek intuícióikra (i. m. 284). Ezzel egyidejűleg a kutatók már nem tekintik problémátlannak a saját elméletük által preferált adattípusokat és eljárásokat. Ennek többféle következménye lehet, így például az is, hogy lazul a generatív grammatika szigorú falszifikációs kritériuma, mondván: a nyelvben sokkal inkább statisztikailag kimutatható tendenciák működnek, mintsem rigid előírások (i. m. 285).

De nemcsak a gyakorlatban, hanem a (meta)tudományos reflexióban is elmozdulás figyelhető meg a korábbi éles szembenálláshoz képest. A szerzők megállapítják, hogy az általuk elemzett újabb tanulmányok az elvek szintjén is túllépnek a fentebb összefoglalt merev nézőpontokon. (Így például valamennyi felsorolt pontot meghaladja a jelen cikkben fentebb többször hivatkozott LEHMANN is.) Mindezzel együtt „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” (i. m. 398).

A szerzők az elfogadott normák és a nyelvtudomány tényleges problémamegoldó praxisa közötti távolság csökkentésére könyvükben a plauzibilis érvelés elméletét dolgozzák ki (KERTÉSZ–RÁKOSI 2008b). Ebben az elméletben az adatok nem közvetlenül megfigyelhető entitásokként, hanem a plauzibilis érvelés folyamatába ágyazott kijelentéseként vannak jelen. S mint ilyenek, szerkezetük két komponensből áll: valamilyen információtartalmat hordozó kijelentés, valamint e kijelentés plauzibilitási értéke valamely közvetlen forrás alapján. Az adatokkal szemben itt tehát nem az a követelmény, hogy bizonyosan igaz tények legyenek, hanem hogy egyértelműen megbízhatók legyenek. Így sokféle adat tekinthető legitimnek, s ezek kombinálhatók is egymással; ha különböző adatfajták egyaránt evidenciaként szolgálnak egy hipotézis mellett, akkor együttesen nagyobb plauzibilitási értéket adnak a hipotézisnek, mint egyenként (i. m. 225–7). Az adatok funkciója egyrészt az, hogy plauzibilitási értéket tápláljanak az elméletbe, másrészt pedig az, hogy gyenge, erős vagy relatív evidenciaként szol-

gáljanak bizonyos hipotézisek mellett vagy ellen (i. m. 230–1). Az elméletalkotás ciklikus folyamatként értelmeződik, amelyben az adatok plauzibilitása nincs egyszer és mindenkorra rögzítve, menet közben változhat, például ha az érvelési folyamat során olyan problémák adódnak, amelyek megkérdőjelezhetik az adott forrás megbízhatóságát (i. m. 229–30). Az evidencia sem feltétlenül bizonyosan igaz állítás, hanem pozitív plauzibilitási értékkel rendelkező (azaz vagy plauzibilis, vagy bizonyosan igaz) kijelentés – következésképpen az evidencia sem tekinthető döntőbírónak (i. m. 234–5).

A nyelvtudomány eszközei prizmák – fogalmaznak metaforikusan a szerzők –, amelyek megtörik a fényt (vagyis beavatkoznak a láttatás folyamatába); arra kell törekednünk, hogy e prizmák torzító hatását kiszűrjük. „A plauzibilis érvelési folyamat lényegéből adódóan el kell fogadnunk a nyelvtudományban az egymástól eltérő és egymással össze nem egyeztethető nézőpontok pluralitását” (i. m. 238–9, 241).

3. Feltűnhet, hogy nyelvtörténeti adatról az eddig érintett munkákban nem esett szó, a LEHMANN-tanulmányban hermeneutikusnak besorolt filologizálás megemlítésén kívül.

A) A nyelvtörténeti adatolás problematikája valóban sokban eltér a szinkrón vizsgálatokétól, kezdve rögtön azzal, hogy a nyelvi folyamatok közvetlen tanulmányozására a nyelvemlékekből származó adatok nem adnak módot. Régi korok szövegei alapján csak közelítőleges nyelvi képet áll módunkban felrajzolni.

Az annak idején keveseknek szóló művek nyelvhasználatát sokban meghatározzák az adott írásos műfajok sajátosságai, miközben a sokaktól származó nagyszámú beszédeseményből a maguk valójában semmit nem áll módunkban vizsgálni (vö. LEHMANN i. m. 26–7; FAARLUND 1990: 18); más megfogalmazásban, más hangsúlyokkal: az írásos nyelvváltozat sokat elfed a változások szempontjából alapvető fontosságú élőnyelvi sajátosságokból, elsősorban a köznyelvet tükrözve pedig nem ad lehetőséget a többi nyelvváltozat kimerítő tanulmányozására (vö. KISS–PUSZTAI 2003: 62–3).

Mivel azonban a fennmaradt szövegek megőrzik a lezajlott változások nyomain, mégiscsak a belőlük kinyerhető elsődleges adatok jelentik a nyelvtörténetész számára a legfontosabb kapaszkodókat.

Nagyon különböző, hogy egy-egy nyelvtörténeti tankönyv, a nyelvi változásokkal foglalkozó áttekintő munka kitér-e az adatolás kérdéseire. Több kötetben nem szerepel ilyen részfejezet, ahol pedig igen, változatos, hogy a témának mely aspektusa(i) kap(nak) hangsúlyokat. Van, aki a nyelvtörténeti adatok fajtáira koncentrálnak azok érvényessége és a velük szemben alkalmazandó eljárásmodok szempontjából (FAARLUND i. m. 16–29). Régi szövegek tanulmányozásakor – írja a szerző – négyféle adatra számíthatunk: 1. grammatikus szerkezetek, amelyek összhangban vannak az adott korban érvényes szabályokkal; 2. az adott korban már elavultnak minősülő szerkezetek (amelyek korábban grammatikusnak számítottak) – ha ezeket vizsgálati anyagunkban az előbbiekkal egy szinten kezelnénk, a kronológiailag elkülönülő formákat szinkrón variánsokként tüntetnénk fel; 3. olyan szerkezetek, amelyek metrikai okokból térnek el a normától (főként

a szórend tekintetében) – ezek az adott kérdés szempontjából kizárandók a vizsgálati anyagból, más kérdések tanulmányozására azonban alkalmasak lehetnek; 4. más okokból agrammatikus szerkezetek – a negatív adat egy fajtájának tekinthető, ha kellően terjedelmes korpuszban nem ismétlődik meg a kérdéses forma. Az adat érvényességének megítélését, biztos fogódzó híján, az „elsajátított intuíció” („acquired intuition”) helyettesítheti (vö. „történeti nyelvérzék”: KISS–PUSZTAI 2003: 63). Van szerző, aki elsősorban a *filológia* terminus értelmezési lehetőségeire és a filológiai kutatásoknak a nyelvészettel való kapcsolatára, a filológia eredményeinek nyelvészeti felhasználhatóságára tér ki (CAMPBELL 1998: 327–8 stb.). A „Magyar nyelvtörténet” című tankönyvnek a nyelvtörténeti leírásról szóló fejezetében az adatok igazolt voltának alapkövetelményéről, majd a nyelvtörténeti kutatások adatokra utaltságáról, valamint az írásos adatforrások (fentebb már hivatkozott) korlátairól esik szó (KISS–PUSZTAI i. m. 60–3).

Nyelvtörténeti adatolással foglalkozó tanulmányokat, előadásvázlatokat inkább csak a legutóbbi évekből lehet találni. Közülük mindenképpen idekiváncozik egy előadás-melléklet megemlézése, amely a nyelvtörténeti adatok és eljárásmodok egymásra rétegződését szemlélteti – vélhetően az idézett LEHMANN-tanulmányból kiindulva (HIMMELMANN 2009: 1). A táblázat az anyanyelvi beszélőtől közvetlenül és közvetett módon származó adatokkal és kezelésükkel kapcsolatban a következőket emeli ki: 1. a nyelvtörténeti adat közvetett módon származik az anyanyelvi beszélőtől; 2. nyers adatok forrásaként például egy eredeti kézirat szolgál (az adatolás módszere: a filológiai kritika); 3. az elsődleges adatokat a kritikai kiadás szolgáltatja (az adatolás módszere: a történeti-összehasonlító elemzés és értelmezés); 4. a másodlagos adat (azaz a rendszeradat) pedig a nyelv- és kultúrtörténethez tartozó megállapítás. A nyelvtörténeti adat identitását szemléltető összegzésben fontos odafigyelnünk az *é r t e l m e z é s* („interpretation”) szóra. Mögötte ugyanis annak az állásfoglalásnak kell állnia (bizonyára még akkor is, ha a szerző időzjelbe tette is a szót), amit LEHMANN itt idézett cikkében (i. m. 28), illetve a grammatikalizációval kapcsolatban máshol is határozottan leszögez: a struktúra *f u n k c i ó t* szolgál; a nyelvben, a nyelvi változásban a forma és szerep szétválaszthatatlan egymástól.

B) A nyelvtörténeti adatok kezelése, érvényük minősítése tehát – mint ahogyan a 3. A) alatt idézettek is mutatják – számos olyan kérdést is felvet, amelyekkel a mai nyelv struktúrájának kutatója nem vagy nem feltétlenül szembesül.

Azok, akik a nyelvészetet természettudománynak tartják, a szinkrón rendszer vizsgálatakor nyilvánvalóan mást és máshogy tekintenek kutatásuk tárgyának, mint amire lehetőséget kínál a számos szempontból heterogén forrásokból kinyerhető és zárt nyelvtörténeti adatállomány (amely távol áll attól, hogy „az ideális nyelvhasználó” nyelvtenát képezze le). Nem működtethető az introspekció, az intuíció; gond van a kompetenciával éppúgy, mint a verifikációval és a falszifikációval, a deduktivitás igénye és a determinisztikusság feltételezése pedig komoly csapdákat rejt.

Azok, akik a nyelvészetet bölcsészettudományként művelik, a nyelvtörténethez is hasonló módon közelítenek, mint amikor szinkrón adatok alapján vizsgálnak nyelvi – vagy nyelvhasználati – kérdéseket. A lehetséges közelítésmódok-

ban, vizsgálati módszerekben számos eltérés mutatkozhat, egy azonban bizonyos: amennyiben elfogadjuk, hogy van tudományosan kutatható nyelvtörténet – tudjuk, erre a felvetésre sem csak igenlő válasz létezik –, nem kérdés, hogy fel kell használnunk az adatforrásként számításba jöhető nyelvemlékeket. Igaz, hogy a források nyelvészeti szempontból számos problémát vet(het)nek fel, s a források problémái továbbélnék az adatolásban. Minél régebbre nézünk, és minél kevesebb nyelvemlék áll rendelkezésünkre, annál több kétely adódik. De nem is kell feltétlenül az írásbeliség legkorábbi korszakáig visszatekintenünk ahhoz, hogy lássuk, mennyi teljes biztonsággal meg nem válaszolható kérdés merülhet fel – körülnézhetnénk például akár a könyvnyomtatás hőskorában is. Példaként most mégis az ómagyar kori kéziratos szövegemlékeket szeretném kiemelni (megjegyezve, hogy a szórványemlékek köz- és tulajdonneveinek, a glosszák, szógyűjtemények adatainak felhasználhatósága és problematikája is némileg másféle, mint a szövegemlékéé).

4. Egyes, szerencsésebb történelmű népek nyelvemlékirodalmához képest nem túl terjedelmes az a korpusz, amely a magyar nyelvtörténet kéziratos (ómagyar) korszakából fennmaradt. A korai ómagyar kor mindössze négy rövid szövegemléket számlál, a kései ómagyarból fennmaradt nagyobb terjedelmű anyagok – kódexek és kódextöredékek – száma a félszázat sem éri el. A szövegmenyiség behatárolt volta is vet föl kérdéseket – mindenekelőtt azt, hogy a tapasztalható hiányok esetlegesek-e vagy rendszerszerűek (ez a probléma egyébként tekintélyes méretű korpuszok vizsgálatakor is fennáll). A művek keletkezésének és hagyományozódásának sajátosságai azonban további kérdéseket hagynak nyitva.

A szövegek és szövegrészletek túlnyomó többségükben a korabeli vallásos irodalom műfajait képviselik (leszámítva elsősorban a kis mennyiségű levelet, hivatalos iratot és néhány világi verset), amelyek kulturális gyökereikkel a latin nyelvű irodalomba kapaszkodnak. Az írástudók és könyvhasználók csekély számát tekintve a nyelvhasználó közösség kis szegmensének dokumentumai ezek, még akkor is, ha felolvasás, elmondás útján többen találkoztak velük, mint ahányan kézbe vették őket. Az olvasni tudók számát a XV. század végén–XVI. század elején 10–20 ezerre szokás becsülni az ország 3–3,5 millió lakosából, az egyetemet végzetteket pedig jó pár százra. Egy adat szerint például a bécsi egyetemnek alapításától, 1365-től kezdve a mohácsi vészig körülbelül ötezer magyarországi hallgatója volt (SOLTÉSZ 2004: 23). Mindennek alapján szóba sem kerül, hogy a kor nyelvemlékei átfogóan és differenciáltan reprezentálnák a nyelvhasználat különböző rétegeit.

A fennmaradt szövegek szerzőit – tehát magukat az adatközlőket – legtöbbször nem ismerjük. A korszak irodalma zömmel fordításirodalom, amely másolatokban terjedt. Valaki lefordította a latin művet, amit azután tisztázati példányként vagy ő saját kezűleg vagy valaki más lemásolt. A művekből több példányra volt szükség; a mára megőrződött szövegek a nem ismert alappéldányoknak rendszerint már többszörös másolatai; az autográf vagy legalább feltehetőleg szerzői másolatban fennmaradt mű egészen kivételes. A scriptorok között nem sok akad,

akitől több szöveget is ismerünk, akinél tehát legalább óvatos következtetéseket levonhatunk arra nézve, hogy a másolás során hajlamos volt-e beavatkozni a szövegbe, vagy szándékai szerint hűségesen követte az előtte fekvő példányt. (A fennmaradt szövegek nyelvi állapota ugyanakkor arról tanúskodik: a másolók akaratuktól függetlenül is nyomot hagynak a műveken: nyelvjárási sajátosságai, saját szövegismeretük alapján, valamint tévesztései, hibáik által is.) Többnyire homályban maradnak a fordító és a scriptor nyelvhasználatát befolyásoló kontaktushatások is: a fordítók esetében hasznos volna tudni, esetleg milyen más nyelv hathatott rá a forrásnyelv latinján kívül (pl. a német?), a másolók esetében pedig legalább azt: vajon tudott-e latinul ő is?

Bár a fennmaradt szövegek egy részének a kora ismerhető, más részének pedig – legalább megközelítő pontossággal – kikövetkeztethető, valódi fogódzóhoz ezzel sem jutunk, lévén a datált szöveg általában egy korábbinak a hűségesebb/hűtlenebb kópiája, esetleg későbbiről, és más nyelvterületről származó másolótól. A keletkezés helye a szövegeknek csak kisebb része esetében ismert, a fent már említett okokból azonban messzemenő dialektológiai következtetésekre ezek az információk sem alkalmasak.

Mit lehet tenni ennyi tisztázatlan körülménnyel? Semmi mást, mint az adatokat is, az eredményeket is ezekkel együtt számításba venni. Élni mindazzal, ami felelősséggel kikövetkeztethető, lemondva az abszolút bizonyosság illúziójáról. A nyelvtörténet mítosz; a forrásokat értelmezzük, az adatokban bízunk kell – összegezzhetjük egy nyelvtörténeti előadás nyomán (NAGY 2007) ROGER LASS könyve egyik fejezetének („A történész mint mítoszeremtő”) tanulságát. Nem feltétlenül ismerjük a régi szövegek jelentését – s minél korábbiak, annál kevésbé –, mégis nyelvtörténetet alkotunk belőlük, amelyben azután hiszünk is (LASS 1997: 5).

A megannyi tisztáz(hat)atlanság ellenére sem tehet mást a kutató, aki a nyelv valamely írásos (de hangzó dokumentációval még nem rendelkező) korszakát vizsgálja, mint hogy a forrásokból induljon ki – nem csak a nyelvhasználatot, hanem a nyelvhasználat mögött meghúzódó rendszert illetően is.

5. A kései ómagyar kori szövegekkel kapcsolatban a közelmúltig elég sokszor megkerülhetetlen volt az adatgyűjtést megelőző filológiai munka is – elsősorban egy-egy nyelvemlék megbízható minőségű kiadása. Ennek eredményeként számos olyan vonatkozásra derült fény, amely differenciáltabbá teheti az adott forrásból kinyerhető adatokat: például több esetben újabb kezek ismertettek fel a korábban számon tartottakhoz képest; addig számon nem tartott azonoskezűségeket sikerült megállapítani; javításokat, átírásokat lehetett pontosítani. A szövegkiadás – akár csak a grammatikai kutatásokat megelőző, akkor még kézi módszerrel történő adatgyűjtés és -rendszerezés – roppant nagy idő- és energia-befektetést igényelt, amely munkák ráadásul nem is céloztak azonnali és közvetlen nyelvészeti eredményt (hiszen éppen ennek megalapozását végezték el). Így nem minden nyelvtudományi fórumon számítottak üdvözlendő tételnek a szakmai munkák sorában, ami egyben utalt a nyelvtörténet korábbi nem túl magas presztízsére is.

A ma folyó munkálatoknak köszönhetően a közeljövőben összehasonlítha-

tatlanul könnyebb lesz nyelvtörténeti adatokat gyűjteni. Számítógépes adatbázisok épülnek, amelyek a nyilvános hozzáférés biztosításával és az elektronikus adatkezelés felbecsülhetetlen előnyeivel egyszerű hozzájutást kínálnak majd (viszonylag) nagy tömegű adathoz. Ezek az újabb munkálatok azonban nemhogy nem teszik értelmetlenné a korábbi úgynevezett „filologizálgatást” – éppen ellenkezőleg, visszamenőleg is jelentősen megnövelik annak értékét, hiszen korpuszt építeni csak a már rendelkezésre álló, filológiaiilag maximálisan megbízható anyagokból lehet, illetve érdemes.

A Nyelvtudományi Intézet Finnugor és Nyelvtörténeti Osztályán a közel-múltban indult az a munkálat, amely 1772 előtti levelek és periratok szövegét készül közzétenni morfológiailag elemzett formában (a projektumot a jelen cikk szerzője vezeti).² Az adatbázis alapanyagául egyrészt az utóbbi évtizedekben megjelent perkiadások szolgálnak, másrészt a XIX. század második felétől napvilágot látott magánlevél-kiadások közül azok, amelyek a filológiai követelményeknek megfelelnek. A magyar nyelvtörténészek munkája ott rejlik egy másik, most készülő korpuszban is, amely a Magyar generatív történeti szintaxis részmunkálataként épül a Nyelvtudományi Intézet Elméleti és kísérletes nyelvészeti osztályán, hiszen ennek ómagyar kori anyagát nem jelentéktelen részben azok a kódexkiadások képezik, amelyek a Nyelvtudományi Intézet korábbi Nyelvtörténeti és Dialektológiai Osztályán, illetve az ELTE Magyar Nyelvtörténeti és Dialektológiai Tanszékén készültek.

A nyelvtörténet az utóbbi időben ismét az érdeklődés előterébe került. A talált adatok „népszerűségét” nagyban fokozhatja, hogy az elektronikus korpuszok segítségével kiküszöbölhető a keresgélés, gyűjtögetés időigényes munkafázisa. Az adatbázisokra épülő grammatikaírásnak pedig mintaadói lehetnek az olyan művek – szemléletmódjukat tekintve még akkor is, ha a nyelvtörténeti adatok (a fentebb vázolt okokból) hátrányos helyzetűek –, mint például a „Longman grammar of spoken and written english” (BIBER et al. 1999), amely az angol nyelvtan funkcionális, korpusz alapú megközelítését tartalmazza, nagy súlyt fektetve a nyelvhasználati rétegenkénti megoszlásra és a regionális, szociális és szituációs változók szerinti jellemzésekre (gyakorisági adatokat is közölve és elemelve).

A) KERTÉSZ ANDRÁS és RÁKOS CSILLA a számítógépes módszerek elterjedésének és a nyomában kialakuló korpusznyelvészet kiépülésének és meggyőző erejének tulajdonítja, hogy az egymástól eleinte mereven elhatárolódó nyelvészeti iskoláknak sokban közeledett az adatokról, az adatolásról való véleménye (KERTÉSZ–RÁKOSI 2008a: 278–9). Így tehát azt is, hogy a generatív grammatika számos képviselője előtt növekvőben van a gyűjtött adatok presztízse, változóban a felhasználhatóságukról / felhasználandóságukról alkotott kép. Ez valóban nagyon jelentős lépésnek tűnik, amelyhez azonban több más nyelvészeti terület is hozzájárulhatott. LEHMANN kifejezetten a veszélyeztetett nyelvek dokumentálásának tapasztalataihoz, eredményeihez köti az adatgyűjtés, adatfelhasználás – leegyszerűsítve fogalmazva – generatív nyelvészeti fordulatát (i. m. 16 és

² Az OTKA 81189 sz. projektum címe: Morfológiailag elemzett nyelvtörténeti korpusz a magánéleti nyelvhasználat köréből.

passim). Magam egy harmadik területet is kiemelnék: a grammatikalizáció-kutatást, amelynek eredményeit, magyarázó elveit a generatív (indíttatású) grammatikusok is gyakran tekintetbe veszik, amikor egy-egy nyelvi változás lefolyását kívánják megragadni. Ez többek között azért érdemel külön figyelmet, mert a grammatikalizáció-kutatás egyértelműen a forma és funkció egymásra épülésére alapoz, míg a generatív grammatika – klasszikus elvei szerint – a forma és jelentés szigorú szétválasztását követeli meg.

B) É. KISS KATALIN, a magyarországi generatív grammatikai kutatások egyik vezető képviselője magától értetődőnek tekinti a változásokat: „A generatív nyelvelmélet evolúciója ma is tart. (...) Nincsenek az elmélet keretei között feloldhatatlan ellentmondások. Az időről időre felszínre kerülő problémák az elmélet eszközeivel is kezelhetők. Például a bírálatok kereszttüzebe került az a – gyakran csupán introspekcióra vagy esetlegesen kiválasztott informánsokra támaszkodó – módszer, mellyel az elmélet nyelvi adatokhoz jut. Az elmélet tanult a bírálatokból, és számos képviselője erőfeszítéseket tesz, hogy megalapozottabb, megbízhatóbb adatgyűjtési módszertant dolgozzon ki” (É. KISS 2008: 7).

Tudománytörténészek a későbbiekben bizonyára vizsgálni fogják: az „időről időre felszínre kerülő problémák” valóban kezelhetők maradtak-e az elmélet eszközeivel. Hiszen – visszautalva a fentebb már említettekre – adott egy elmélet, amely az egyén nyelvi képességének modellálását célozza meg, ezért az introspekcióra (és esetleg grammatikalitási tesztekre) épít; modellalkotása a funkciótól leválasztott formára irányul; a nyelvi jelenségekről azt tartja, hogy bizonyos feltételek között éppoly szükségszerűen bekövetkeznek, mint a természettudományokban. Az elmélet „evolúciója” pedig oda vezet, hogy képviselői – történeti kérdéseket vizsgálva – elfogadják a heterogén eredetű, gyűjtött adatokat, a nyelvi változásokat elemezve a formához kapcsolódóan a funkcióval is számolnak; és szembesülniük kell a ténnyel: a nyelvi változások nem determinisztikusak.

C) „Az elméletek változnak – az anyag szerencsére mindig ugyanaz marad”, hallani gyakran. A szigorúan adatközpontú, induktív szemléletű nyelvtörténeti kutatásokból ad absurdum akár valamiféle, a hiányzó paradigmához közelítő zárványszerűség is válhatott volna, amely a meghatározó kutatások számára egyszerűen kikerülhető (a „zárványállapot”-ra l. BÉKÉS 1997: 78.). Ennek azonban manapság – a nyelvtörténészek és a nyelvtörténeti kutatások vitathatatlan javára – sok szempontból az ellenkezője tapasztalható. A korábbi eredmények (és hiányok) beépülnek az újabb kutatásokba. Ez utóbbiak viszont – az univerzális grammatika és az annak megismeréséhez kapcsolódó elméleti vonatkozások felől közelítve a lezajlott nyelvi változásokhoz – friss, megtermékenyítő szempontokat vonnak be a nyelvtörténeti vizsgálatokba, korábban fel nem vetett, releváns kérdésekkel és megoldási javaslatokkal állnak elő.

A nyelvi anyagra való minél alaposabb ráközelítés megfogalmazása – mint a kutatásoktól elvárható alapigény – igen didaktikusan hathat. Eszünkbe idéződik azonban tanulmányok, kötetek – s ez a jelenség a legkevésbé sem iskolafüggő –, amelyek nem építenek az adatokra, legfeljebb csak felhasználják a példákat, illusztrációul helyezve őket a megállapítások mögé. Ez az eljárás azt jelenti, hogy a szerző eleve lemond a gyűjtött adatokkal való termékeny szembesü-

lés lehetőségéről. Ha az adatok vizsgálata képezi a kiindulópontot, és a megállapítások ezen a talajon születnek meg, figyelembe vehetők az anyagból kirajzolódó, de előre nem kalkulált jelenségek, új szempontok is. Az egyik út tehát: feltárni, megismerni az anyagot a maga sokrétűségében; a másik: kikeresni belőle olyan adatokat, amelyek a deduktív elemzés adott szempontjaihoz illeszkednek.

A nyelvészeti iskolák adatkezelési módszereinek közeledése (amely ténnyel részletesen foglalkozik a hivatkozott tanulmányok közül KERTÉSZ–RÁKOSI i. m. és DURAND is), az anyag- és adatközpontúság presztízsének növekedése, szerencsés módon, mára már a nyelvtörténetben is egyre inkább érzékelhető.

Kulcsszók: adatgyűjtés, filológia, ómagyar kor, történeti nyelvérzék, intuíció, introspekció.

A hivatkozott irodalom

- BÉKÉS VERA 1997. A hiányzó paradigma. Latin Betűk, Debrecen.
- BIBER, DOUGLAS et alii 1999. BIBER, DOUGLAS – JOHANSSON, STIG – LEECH, GEOFFREY – CONRAD, SUSAN – FINEGAN, EDWARD, Longman grammar of spoken and written english. Pearson Education Limited, Harlow.
- CAMPBELL, LYLE 1998. Historical linguistics. Edinburgh University Press.
- CSERESNYÉSI LÁSZLÓ 2009. Még egyszer a kútról és a békáról. Magyar Nyelv 105: 401–11.
- DURAND, JACQUES 2009. On the scope of linguistics: data, intuitions, corpora. In: Yuji Kawaguchi – Makoto Minegishi – Jacques Durand eds., Corpus analysis and variation in linguistics. John Benjamins Publishing Company. 25–52. és <http://w3.erss.univ-tlse2.fr:8080/index.jsp?perso=jdurand&subURL=DurandTokyoPaper.pdf> 1–24. (A hivatkozások ez utóbbi lapszámai alapján.)
- FAARLUND, JAN TERJE 1990. Syntactic change – Toward a theory of historical syntax. Mouton de Gruyter, Berlin – New York.
- É. KISS KATALIN 2008. A 80 éves Noam Chomsky és a chomskyánus nyelvészeti forradalom. Magyar Nyelv 104: 1–8.
- HIMMELMANN, NIKOLAUS P. 2009. Linguistic data types and documentary linguistics. A University of Hawaii-n elhangzott előadás melléklete: scholarspace.manoa.hawaii.edu/bitstream/10125/5162/3/5162.pdf 1–2.
- KERTÉSZ ANDRÁS – RÁKOSI CSILLA 2008a. Megjegyzések a nyelvészeti adatok és evidencia problémájáról folyó vita jelenlegi állásához. Magyar Nyelv 104: 274–86; 385–401.
- KERTÉSZ ANDRÁS – RÁKOSI CSILLA 2008b. Adatok és plauzibilis érvelés a nyelvészetben. DE Kossuth Egyetemi Kiadó, Debrecen.
- KISS JENŐ – PUSZTAI FERENC szerk. 2003. Magyar nyelvtörténet. Osiris Kiadó, Budapest.
- LASS, ROGER 1997. Historical linguistics and language change. Cambridge University Press.
- LEHMANN, CHRISTIAN 2004. Data in linguistics. The Linguistic Review 21: 175–210. és http://www.christianlehmann.eu/publ/data_in_linguistics.pdf 1–40. (A hivatkozások ez utóbbi lapszámai alapján.)

- NAGY KATALIN 2007. Adatok és evidencia egy történeti pragmatikai kutatásban. www.nyfk.unideb.hu/konfPecsNagy.ppt
- SAMPSON, GEOFFREY 1975. The form of language. Weidenfeld and Nicholson, London. (Idézve: KERTÉSZ–RÁKOSI 2008b: 279.)
- SOLTÉSZ ZOLTÁNNÉ 2004. A grafika és a könyvnyomtatás elterjedése Magyarországon. In: V. ECSEDY JUDIT, A régi magyarországi nyomdák betűi és díszei. Balassi Kiadó – Országos Széchényi Könyvtár, Bp. 19–27. (Újraközlés innen: A magyarországi könyvdíszítés a XVI. században. Akadémiai Kiadó, Bp. 1961.)
- TRASK, ROBERT LAWRECE 2000. The dictionary of historical and comparative linguistics. Edinburgh University Press, Edinburgh.

Data in historical linguistics: principles, practice, and possibilities

The literature discusses the place and significance of data in linguistic inquiry, the possibilities of using data and ways of accessing them, primarily in the context of synchronic analysis. The issue of data attestation in historical linguistics is usually given less attention. This paper discusses the latter issue, based on experiences in investigations of Old Hungarian and summarizing the problems related to the study of linguistic material coming from earlier periods. The author furthermore surveys some more general issues pertaining to the relationship between linguistic research and linguistic data, as well as the changes of that relationship.

Keywords: data collection, philology, Old Hungarian, historical linguistic insights, intuition, introspection.

DÖMÖTÖR ADRIENNE

Az 1109. évi veszprémvölgyi ítéletlevél néhány alapkérdéséről

1. Bevezetés. – A Magyar Nyelv korábbi számában megjelent, „Géza vagy István idejében alapították-e a veszprémvölgyi monostort?” című publikációmban igyekeztem meggyőzően kimutatni, hogy a monostor alapításának 1018 körüli, MORAVCSIK–GYÖRFFY-féle datálása teljesen megalapozatlan, s megindokolni, hogy miért tartom helyesnek BALOGH ALBINNAK az alapítást Géza fejedelemhez kapcsoló nézetét. A monostor alapítólevelének kiállítását a 975–985 közötti évtizedre, legvalószínűbben 980 körülre datáltam. (HOLLER 2011).

E korábbi dolgozatban elsősorban a görög nyelvű alapítólevél volt vizsgálatunk tárgya. Ennek eredményeire támaszkodva az alábbiakban szeretném ismertetni elgondolásaimat az 1109. évi latin nyelvű renovációs oklevél néhány alapkérdésével kapcsolatban.

Az oklevél levéltári jelzete: MOL DL 11; regesztája: RegArp. 42a; a latin szöveg átírása: DHA. 366–7, ÁrpOkl. 36–7; magyar fordítása: ÁrpOkl. 124–5; szakirodalma: DHA. 365–6. Nr. 136/I, I. még DHA. 368–78. Nr. 136/II. Az oklevél jól nagyítható fotója elérhető az interneten (RÁCZ 2009).

Az oklevelet a jelen dolgozatban is a renovációs, illetve megújító