

fontos, régi és folyton visszatérő kérdése, hogy milyen paraméterek alapján különíthetők el egymástól a spontán beszédben a bekezdések, és milyen egységek jönnek így létre, hiszen a beszélt nyelv nem tartalmaz olyan egyértelmű határjelölőket, mint az írott nyelv. A beszéd feldolgozásának egyik alapfeltétele, hogy a hallgató képes legyen az elhangzottakat számára értelmes egységekre bontani és így a beszédet dekódolni. A kutató arra keresi a választ, hogy a hallgatók percepció mechanizmusa mennyire működik egyformán, illetve milyen akusztikai fonetikai, szintaktikai és pragmatikai jellemzők alapján szegmentálják az elhangzottakat. Eredményei azt mutatják, hogy a hallgatók ítéletei a határjelölésben nagymértékben megoszlanak, de a jelölések mintegy fele egyöntetű volt. Ez arra enged következtetni, hogy a spontán beszéd bekezdésekre tagolására különféle stratégiák léteznek. Elsősorban a szüneteknél és a lezárt szintaktikai egységek esetében jelöltek egyöntetűen bekezdéshatárt a hallgatók, minél hosszabb szünetet észleltek, annál többen vélték tagolópontra, és az összes jelölés túlnyomó többsége szintaktikai határvonalon található. A kutatás alapján elmondható, hogy a mondatnál nagyobb egységek meghatározása összetett folyamat, az egyes beszédjellemzők (témaváltás, szintaktikai szerkezetek, zöngemínőség, dallammenet) csak az egymással való összefüggésükben kapják meg határjelölő szerepüket. Gocsál Ákos tanulmánya is a beszéd kutatás egyik elengedhetetlen változójával foglalkozik, az alaphangmagasság mérésének módszertani kérdését járja körül. A tanulmány az F_0 -értékek számtani középértékét, móduszát és mediánját hasonlítja össze. A különböző adatok szignifikáns eltéréseket mutatnak, amely arra enged következtetni, hogy a magyar beszélők alaphangmagasság-megoszlása jellegzetesen aszimmetrikus, így a középérték és a szóráshoz nem minden esetben tükrözi az F_0 eloszlását.

A beszéd információtechnológiai vonatkozásai című fejezetben a beszélőfelismerés és a beszéd szintézis tárgyában olvashatunk egy-egy tanulmányt. Beke András a beszélőfelismerés kevert Gauss-modelleket alkalmazó módszerét mutatja be. Az algoritmus tanítását és tesztelését a BEA adatbázisból kiválasztott 80 adatközlő beszédén végezte. Az eredmények alapján jól látható, hogy a Gauss-modellek használata jelentősen javítja a szövegfüggetlen beszélőazonosítás eredményét. Olasz Gábor a beszéd szintézis múltját, jelenét és jövőbeni kilátásait ismerteti. A téma történelmi áttekintésével az olvasó megismerheti a különböző irányvonalakat, amelyen elindult a kutatás annak idején, képet kaphat a technikai-műszaki fejlődésről, a beszéd szintetizátorral szemben támasztott igények változásáról. Azt is láthatjuk, hogy jelenleg a beszéd szintézis még nem képes előállítani az emberi olvasási teljesítménnyel minden tekintetben összemérhető gépi felolvasót. Az optimális gépi szövegfelolvasó jellemzői a következők lennének: egyszerű használat, komplex, rugalmas produkciós modell, teljesen emberi hangminőség, kiváló érthetőség, a hibakeresés, hangolás biztosítása és a jó hangbiztonság. A jelenleg megvalósított, különböző technológiájú rendszerek mindegyike kompromisszumra kényszerül, hol a hangminőség, hol az egyéb jellemzők rovására, így csak szűk részterületeken alkalmazhatók. A szerző javaslatot tesz a hibák kiküszöbölésére, amelynek kulcsa véleménye szerint a hullámforma-elmőkiválasztásos módszer és a statisztikai, gépi tanulás technológiájának keresztezése lenne a megfelelő minőségbiztosítási rendszer mellett.

A tanulmányokat magyar és angol nyelvű összefoglalók követik, és egy tekintélyes, több mint háromszáz nevet felvonultató tabula gratulatoria zárja. Az olvasó tájékozódását segíti a kétféle tartalomjegyzék: az első a kötetben szereplő tematikus sorrendet követi, a második pedig lehetőséget ad a tudományterületek alapján történő keresésre. A kötetet tudományos értéke, szakmai színvonala, gazdagsága miatt nemcsak az ünnepelt fogja gyakran és örömmel forgatni, hanem a szűkebb és tágabb szakma – és (reméljük) az érdeklődők minél szélesebb köre.

Bata Sarolta

PhD-hallgató

ELTE BTK

Nyelvtudományi Doktori Iskola

Cs. Nagy Lajos: Medvesalja magyar nyelvjárási atlasza. Luminosus, n.o. Kiadó, Nagykapos, 2011. 46 lap + 750 térkép

A magyar nyelv földrajzi kutatás újabb regionális atlaszsal gazdagodott. Cs. Nagy Lajos munkája olyan felvidéki magyar nyelvjárás csoportot mutat be, amelyről eddig csak szórványos adataink voltak.

Már Benkő Loránd is felhívta a figyelmet 1992-ben a III. Dialektológiai Szimpozionon az itteni kutatások fontosságára: „Mintha a környező országok közül Csehszlovákia volna az a terület, ahol a magyar dialektológiának a gyűjtőmunkája, feldolgozó munkája nem érte volna el ezt a szintet, amelyet a környező országokban elért. Annál is inkább mondom ezt, mert ebben a körben nem kell különösképpen bizonygatni, hogy az északi peremvidéken milyen csodálatosan szép, becses nyelvi értékek, a kutatásban jobbra még felderítetlen területek vannak” (1998: 22). Szerencsére a szlovákiai magyar nyelvészek buzgalomból és az ott dolgozó magyarországi vendégtanárok kutatásaként azóta több kiváló nyelvjárási tanulmány is megjelent. 2004-ben Sándor Anna: A Nyitra-vidéki nyelvjárások atlasza című munkája, 2011 végén pedig Medvesalja magyar nyelvjárási atlasza (MedvA.) járult hozzá ahhoz, hogy erről az eddig alig ismert nyelvjárási vidékről korszerű, gazdag ismeretanyaghoz juthassunk.

A nyelvatlazok készítése nagyon időigényes, sok gonddal járó munka, Cs. Nagy Lajos kutatása is húsz évvel ezelőtti előzményekre nyúlik vissza: „Nyitrai vendégtanári kiküldetésem (1991–1996) lehetőséget adott arra, hogy egy felvidéki tájegység, a Nyitrától kb. 250–300 km-re fekvő festői Medvesalja nyelvkincséből meríthessek. A nagyatlaz alapján – családi és hallgatói összefogással – végeztem a kutatásomat megalapozó helyszíni gyűjtést Szlovákia egyre inkább elnéptelenedő, valamikor idegenforgalmilag is látogatottabb tájegységén. Kapcsolatom e tájegység palóc lakosságával később sem szakadt meg. Az OTKA támogatásának köszönhetően újabb gyűjtő utakat szervezhettünk 2007-ben és 2008-ban kutatótársammal, N. Császi Ildikóval. Ezekben az utakon újabb nyelvi adatokkal gazdagodtunk, s összesen több mint 180 ezres korpusz jött létre. Ebből az adatmennyiségből 100 ezret meghaladó került az atlasz térképeire” (3) – írja a szerző atlasza Előszavában.

A medvesalji falvak az egykori Gömör megyének Nógráddal határos délnyugati részén – a mai Szlovákia területén – a Gortva-patak völgyében fekszenek. A lakosság római katolikus vallású, kultúrájában, nyelvjárásában hagyományörző, sok tekintetben archaikus jellegű. A tizenkét települést (Sőreg, Ajnácskő, Csevice, Gömöralmágy, Dobfenek, Gömörpéterfala, Óbást, Újbást [korábban Egyházasbást], Bakóháza, Vecseklő, Medveshidegkút, Tajti) magában foglaló tájegység népessége rohamosan fogy. „Az összlakosság száma 1980-ban 5797 fő volt. Közülük 5500 volt magyar nemzetiségű, 297 pedig más nemzetiségű. 2001-re [...] a magyarok száma 1129 fővel, 20,53%-kal lett kevesebb két évtized alatt. A fogyás évtizedenként átlagosan 10–11% között van. A szlovák nemzetiségűek lélekszáma 1991-hez képest – tehát 10 év alatt – 105-tel gyarapodott, ez 39,77%” (6).

Egy nyelvatlaz készítésekor a kutatóterület kijelölése mellett a másik legfontosabb feladat a vizsgált nyelvjárási jelenségek meghatározása, a leendő térképek címszójegyzékének összeállítása. Cs. Nagy Lajos természetesen a MNyA. anyagából indult ki, innen vette át a kérdéseket, sőt a „teljes palóc kérdőívet” is beépítette kérdőfüzetébe. Az anyaggyűjtés első szakaszában 1204 kérdésből álló, 69 oldal terjedelmű kérdőfüzettel dolgozott. Később ezt a kérdőívet átalakította: „A vidék táplálkozási, építészeti és népi szokásaiból, leíró munkákból, valamint Sándor Anna Nyitra-vidéki nyelvatlazából újabb kérdőívet állítottam össze 2006-ban, s ezt kérdeztük ki 2007-ben és 2008-ban. Az összes (1305) kérdés megoszlása a következő: MnyA.-kérdés 978 (617 lexikai, 356 morfológiai és 5 fonetikai), Palóc kérdőív 200 (65 lexikai, 44 morfológiai és 91 fonetikai) pótygyűjtés 120 (120 lexikai), NyvA. 7 (7 lexikai). A hangfelvételes kérdőíves gyűjtés mellett több szövegfelvételt is készítettünk” (8–9).

A MedA. a nyelvjárási alaprégét kutatását tűzte ki céljául, adatközlői az idősebb, a 60 éven felüli helybeli lakosokból kerültek ki. Ezért az atlaszban a térképlapokon egynemű adatokat találunk, nem közöl archaikusnak vagy neologizmusnak minősíthető példákat. Erre vonatkozóan a szerző az Előszóban a következőket írja: „Szociolingvisztikai jellemzők nem olvashatók ki atlaszomból. Ennek az az oka, hogy két lehetőség között választhattam: vagy sok adatot teszek közkinccsé, vagy kb. 100–150 címszónak az adatközlők szociológiai jellemzőitől meghatározott térbeli vizsgálati eredményeit vetítem térképre... Ekkora mennyiségű címszó esetében és ennyi idő alatt csak erre alkalmas számítógépes program, valamint több személyből álló kutatócsoporttal lett volna megvalósítható a többdimenziós atlasz” (4).

Ez az atlasz a tájegységi határokat követő kisatlazok közé tartozik, kutatóterülete Medvesalja és környéke 12 településére terjed ki; ha viszont az atlasz 750 térképlapját tekintjük, akkor az egyik leggazdagabb anyagú atlasznak számít, hiszen csak a nagy általános és regionális atlaszok

vettek fel ennél több címszót. Mivel A magyar nyelvjárások atlaszának ebben a kistérségben nincs egyetlen kutatópontja sem, a szerző arra törekedett, hogy kellő számú adatot szolgáltatson a hangtani, alaktani és lexikai sajátosságokra. A MedA. sem csupán nyelvtani érdekességekkel szolgál, más, rokon tudományok kutatói is csemezgethetnek belőle, erre a szerző kifejezetten törekedett a címszavak meghatározásakor, ugyanis a 750 címszó 90%-a lexikai természetű. Mivel mindegyik térképlapot adatbeírás előjárásával közli, ezek a lapok is tartalmazzák a területre jellemző fonetikai és morfológiai változatokat.

A nem kis megterhelést kívánó munkát a kiváló gyűjtő is csak önzetlen segítők és nem utolsósorban megértő adatközlők segítségével tudja sikerre vinni. „Nagy hálával gondolok adatközlőinkre, akik mindig türelmesen végighallgatták kérdéseinket, bátran nyilatkoztak a hangrögzítő készülék jelenlétében is. Köszönöm korábbi nyitrai hallgatóimnak és jelenlegi pázmányos hallgatóimnak is a gyűjtésben való aktív részvételét” – írja a szerző (4).

A térképek közlési sorrendje – a MnyA. gyakorlatának megfelelően – a kérdőfüzet tematikus felépítését követi, az ebben való tájékozódást kétféle Mutató is segíti. Mindegyik térkép adatbeírásos, közülük 32 térkép az egyes adattípusokat kiemelő és elkülönítő színes szemléltetést is tartalmaz. A térképek szerkesztése, a nyelvi variánsok bemutatása Vékás Domokos és Vargha Fruzsina Sára által fejlesztett Bihalbocs szoftverrel készült. Ez a szoftver lehetővé teszi, hogy a teljes korpusz informatizált változatából a kutató Macintosh gépeken igényeinek megfelelő új térképeket generálhasson, újabb és újabb összefüggéseket tárhasson föl, mutathasson be. „Ez az első olyan informatizált atlasz, melyben a gyűjtője saját anyagát teszi közzé, s amely digitális változatban (CD-lemezen) is megjelenhet” (3–4). A formátum lehetővé teszi, hogy a medvesalji korpusz integrálható legyen a magyar és az európai dialektológiai kutatási adatállományokba. Ebben a tekintetben szervezen kapcsolódik az ELTE Geolingvisztikai Műhelye által folytatott informatizálási munkálatokhoz.

Az atlasz térképlapjainak elemzése jobbra a későbbi kutatás feladata lesz, bár szerzőnk az anyaggyűjtés és a szerkesztés során már több előtanulmányt végzett és jelentetett meg (Cs. Nagy 1999, 2003 stb.). Az adatok egyfajta feldolgozásának tekinthető a színes térképek szemléltető megoldása, a térképlapok alsó, jobb sarkában, az úgynevezett „Legendában” lévő tipizálás. Ilyen jellegű, igen dicséretes újítás a lexikai változatok mutatójának az összeállítás is (28–39).

Az Adattár Szóföldrajzi térképek cím vezeti be, ez a korábbi terminológia használatának kissé megtevesztő, mert csak lexikai térképlapokat értünk rajta, ám itt – a tematikus közlési sorrendhez alkalmazkodva – fonetikai és morfológiai térképek is vannak. Ezek a nyelvtani térképek viszont csak eredeti szándékukban fonetikaiai vagy morfológiaiai, mert – esetenként – a megjelölt formájukban a szerző szerepelteti rajtuk a felbukkanó lexikai variánsokat is, például **460. kötény** Morf. adatai: *kötény, surc, kötiny*; a **730. hozzánk** Morf. adatai: *hozzánk, nálunk* stb. De ahogy a szerző *Az anyaggyűjtés* című bevezető fejezetben, Demét is idézve írja: „a különböző jelzetű kérdésekre kapott adatokat komplexen kell vizsgálni: az alaktani adatok egyben hangtaniak is, a lexikaik pedig egyben hang- és alaktani jelentőségűek is” (9). Ez a módszer – véleményem szerint – csak a kevés kutatópontot érintő, szerényebb mennyiségű lexikai változatok esetében alkalmazható. Medvesalja magyar nyelvjárás atlasza ezt a megoldást lehetővé tette, köszönhető ez a jól áttekinthető szerkesztésnek és annak, hogy a kiadó vállalta a teljes térképtár megjelentetését.

A vizsgált kistérségi nyelvjárás a palóc nyelvjárás régióba tartozik, ezt a teljes anyag dokumentálja: *árpf, búzf, s'sf* (a búza levele) stb. A **607. megver** Morf. térkép a palóc nyelvjárás területre jellemző, köznyelvtől eltérő archaikus teljes hasonulást példázza: *mévver*; a **608. sz. őrizte** Morf. lap a peremvidékekre jellemző (pl. Vas megye) szóvégi, alaktani természetű zárt *ë-zést* mutatja: *őrisztë, őrzöttë*. Egyébként az *e : ë* szembenállásban figyelhető meg a legtöbb különbség a dunántúli és az itteni regionális köznyelvi kiejtéshez képest, például **609., 610. sepersz** Morf.: *sëprësz, sëprëtem*; **722. titeket** Morf.: *titëkët, tyiktëkët*, további példák a térképlap idézése nélkül: *terhës, fülë, pislejë ~ pittÿë 'ajaka', gëgë, tÿnyër ~ tÿrÿny 'tenyér', gërinçë, gënnÿ, szëplôs, piszë, erdejë* stb. Nyilván ennek a jelentős különbségnek az okára egy későbbi tanulmányában a szerző is válaszolni fog. A Nyitra-vidéken jelentkező erős zárt *ë-zést* Sándor Anna például annak tulajdonítja, hogy az egyes, elszigetelődött magyar községekben igen erőteljesen érvényesül a szlovák nyelv hatása (2004: 68).

A lexikai térképlapok részben a vidék sajátos magyar néprajzát is tükrözik, gazdag adattárát adják például a hagyományos népi ételneveknek: **398. derelye**: *dërÿleye, bftushfluskf*; **400. fánk**: *pfmpuskf*, **401. forgácsfánk**: *heröce*, **404. krumplinudli**: *bëgyërö, bëgyëllö* stb. A térképlapok kö-

zött olyanokat is találunk, amelyek regionális tájszót tartalmaznak, a bizonyos tárgyat, ételt vagy munkafolyamatot csak ezen a vidéken ismerik, például: **438. virics** 'tavasszal a rügyező nyírfából baltavágással eresztett édes folyadék', **442. molnárkalács**, **443. gúnárnyak**: *gubf, gunárnyfk*, **444. páronfött**: *pārānfött*, **446. morvány**: *morvān, morvāny* stb. A figyelmes kutató több példát is találhat a szlovák nyelvi hatás érvényesülésére, ilyen például a **49. polyva** Lex. térképen a *mohëtka* szó. Ezen a téren is érdemes további vizsgálódást folytatni.

Végezetül egyetértően idézem Juhász Dezső véleményét, amelyet a büki élőnyelvi konferencián fogalmazott meg a *Merre tovább, magyar nyelvöldrajz* című előadásában: „Ideális megoldásnak azt tartom, ha azok a kutatók foglalkoznak először a nyelvöldrajzi anyaggal, akik a gyűjtést elindították és lebonyolították. Így sokoldalú tapasztalatok birtokában kezdődhet az analízis” (Juhász 2006). Ezt a további munkát bizvást várhatjuk Cs. Nagy Lajostól, és természetesen további kiváló nyelvöldrajzi tanulmányok, atlaszok készítését is.

A szép kiállítású könyvért köszönet illeti a nagykaposi Luminosus n.o. Kiadót. A kötet megjelenését a Szlovák Köztársaság Kormányhivatala támogatta.

SZAKIRODALOM

- Benkő Loránd 1998. Megnyitó beszéd. In: Szabó Géza–Molnár Zoltán (szerk.): *III. Dialektológiai Szimpozion. A Berzsényi Dániel Tanárképző Főiskola Magyar Nyelvészeti Tanszékének Kiadványai II.* Szombathely, 18–24.
- Cs. Nagy Lajos 1999. Néhány fogalom megnevezésének tagoltsága és szóföldrajzi realizációja a felvidéki Medvesalján. *Magyar Nyelvőr* 123: 35–45.
- Cs. Nagy Lajos 2003. *Lexikológiai vizsgálatok Medvesalján.* Fórum Kisebbségkutató Intézet–Lilium Aurum Könyvkiadó.
- Juhász Dezső 2006. Merre tovább, magyar nyelvöldrajz. In: Zelliger Erzsébet (szerk.): *Nyelv, területiség, társadalom.* A 14. élőnyelvi konferencia (Bük, 2006. okt. 9–11.) előadásai. MNyTK. 228. Budapest, 33–43.
- Sándor Anna 2004. *A Nyitra-vidéki magyar nyelvjárások atlasza* (NyvA.). Kalligram Könyvkiadó, Pozsony.

Király Lajos

ny. főiskolai tanár

Kaposvári Tanárképző Főiskola

Peters, Isabella: Folksonomies. Indexing and Retrieval in Web 2.0. Wolfgang G. Stock (ed.). (Knowledge & Information. Studies in Information Science). Berlin: De Gruyter Saur, 2009. 443 lap

„A digitális fogyasztó kifejezetten pragmatikus: véleménye szerint, ha egy információ nem található meg azonnal, egy helyen, akkor nem is érdemes keresni” – ez a 2008-as, az ismertett munká bevezetőjében szereplő Nicholas–Rowlands-idezet is lehetne a *Folksonomies. Indexing and Retrieval in Web 2.0* című kötet mottója. Hiszen a címbeli *folksonomy* éppen az információ létrehozásával, keresésével és elérésével foglalkozik.

Az információtudományi könyv címét magyarul pontosan, az angol nyelvi kontamináció megbontása nélkül nem lehet visszaadni. A *folksonomy* terminust 2004-ben használta először egy konferencián Thomas Vander Wal, majd Gene Smith tanulmánya alapján vált ismertté (Vander Wal 2004, Smith 2004; [154]¹), és a *folk* 'nép', illetve a *taxonomy* 'csoportosítás, osztályozás' szavak vegyítéséből származik [5], ily módon a „nép”, vagyis a nem professzionális indexelők által létrehozott taxonómiára utal.² A folksonomy szabadon választott, bármely információhoz köthető kulcsszavakból vagy címkékből (úgynevezett tagekből) áll – e címkéket a post-itelek elektronikus megfelelőinek is szokás nevezni (Müller-Prove 2008: 15; [153]). A kollaboratív tagelés az a folyamat, amelynek során a felhasználó a megosztott tartalomhoz szabadon választott kulcsszavak formájában

¹ Szögletes zárójelben az ismertett kötet oldalszámait adom meg.

² Az ismertetésben a könyv címétől eltérően az egyes számú *folksonomy* alakot használom, természetesen a fogalomhoz tartozó tartalmak összességét értve alatta.