

Szakmai zárójelentés

A magyar nyelvváltozatok geolingvisztikai kutatása

című projekthez

I. Az eredmények tudományos tartalma

A kutatás négy nagyobb feladat megvalósítását foglalta magában. Elsőként hagyományos dialektológiai adattárak informatizálását, másodikként a moldvai magyar nyelvjárások multimédiás diakrón nyelvátlaszának kialakítását, harmadik célkitűzésként A magyar nyelvjárások atlaszához kapcsolódó hangfelvételek feldolgozását foglalmaztuk meg, míg negyedik feladatként erdélyi névtörténeti anyagok informatizálására vállalkoztunk. E feladatok teljesítését a fenti sorrendben mutatjuk be, majd ezt követően ismertetjük a kutatás eredményeinek közzétételére vonatkozó eddigi tevékenységünket és jövőbeli terveinket. Ezek szorosan kapcsolódnak a Geolingvisztikai Műhely honlapjához, amelynek a kialakítását a kutatás során végeztük el. Eredményeinkről számos már megjelent vagy még kiadatlan, de már publikálásra elfogadott tanulmányban és CD-kiadványban számoltunk be; bibliográfiai adataikat a jelentés 1. mellékletében adjuk meg.

1. Hagományos dialektológiai adattárak informatizált változatának létrehozása

Célkitűzéseink négy részfeladatra oszthatók. Ezek a részfeladatok egy-egy hagyományos adathordozón (papíron) kiadott vagy még kiadás előtt álló nyelvátlasz(kötet) informatizált változatának létrehozását jelentették. Az egyes atlaszok vagy atlaszkötetek a következők voltak:

1. A magyar nyelvjárások atlasza IV–VI. kötete;
2. A romániai magyar nyelvjárások atlasza V–VIII. kötete.
3. a Székely Nyelvföldrajzi Szótár;
4. A moldvai csángó nyelvjárás atlasza.

Megjegyzendő, hogy a harmadikként említett Székely Nyelvföldrajzi Szótár nyelvátlaszadatokat ad közre szótárszerű formában.

A kiadványok informatizálása számos eltérő problémát vetett fel, ennél fogva érdemes áttekinteni, hogy milyen megoldásra váró nehézségek adódtak a részfeladatok során. Ezt megelőzően azonban röviden ismertetjük azt a nyelvészeti eljárást, amelyet a nyelvátlaszadatok informatizálásában felhasználtunk az adatfájlok formátumának és kódolásának meghatározásához.

1.1. Az adatfájlok formátuma és kódolása

A projekt keretében végzett informatizálási munka a Bihalbocs dialektológiai szoftverrel történt. A szaktudományos feladatok megoldásához speciális programok szükségesek, ugyanis a kereskedelmi forgalomban beszerezhető szoftverek (szövegszerkesztők, hangszerkesztők, adatbáziskezelők, grafikus programok, térinformatikai alkalmazások) csupán az igények töredékének kiszolgálására lennének alkalmasak. A

Bihalbocs olyan lejegyzésszerkesztő (speciális szövegszerkesztő), hangkezelő és térképgeneráló szoftver, amely egyrészt lehetővé teszi nyelvjárási adatok megfelelően kódolt rögzítését és adattárakba rendezését, másrészt az adatok sokoldalú felhasználását a kutatói igények szerint.

A megfelelően informatizált adattárak óriási előnye, hogy integrálhatók – ez a világos stuktúrából és a kompatibilis kódrendszerből következik. A nyelvjárási jelenségek kódolására olyan analitikus kódrendszert alkalmaztunk, amely a teljesség igényével optimálisan biztosítja az adatrögzítés pontosságát és az informatizált anyagok hatékony felhasználását. A szöveges anyagokat szövegfájlokban, illetve BXT-fájlokban tároljuk. Mivel a létrehozott anyagok használatát hosszú évtizedekre biztosítani szeretnénk, a legegyszerűbb kódolási és tárolási formátumokat alkalmazzuk, hogy a rendszeresen megújuló informatikai platformok közötti hordozhatóságot minél jobban elősegítsük. („Legegyszerűbb” megoldáson természetesen a konkrét feladat összetettségéhez viszonyított egyszerűséget értjük.)

A már kiadott, illetve korábbi gyűjtésekből származó anyagok kódolása és tárolása szövegformátumban történik. Az adatok szigorúan rögzített struktúrába rendeződnek, és fonetikai kódok sorozatából állnak. Az alapjeleknek (például *e*, *é*) vagy mellékjeleknek (például a félhosszúság jele, a palatalizáltság jele) egyaránt külön kódok felelnek meg.

Kutatásunk során vállalt feladatainknak megfelelően ebben a formátumban végeztük el A magyar nyelvjárások atlasza IV–VI. kötetének, A romániai magyar nyelvjárások atlasza V–VIII. kötetének, a Székely nyelvöldrajzi szótár anyagának, valamint A moldvai csángó nyelvjárás atlasza köteteinek informatizálását. A következőkben az egyes kiadványokhoz kapcsolódóan tekintjük át az informatizálás folyamatát.

1.2. A magyar nyelvjárások atlasza IV., V. és VI. kötetének informatizálása

A Magyar Tudományos Akadémia Nyelvtudományi Intézetében már korábban elkészült A magyar nyelvjárások atlasza (a továbbiakban: MNyA. vagy Nagyatlasz) első három kötete jelentősebb részének – pontosabban: az első 450 térképlapnak – a számítógépes változata. Ezt a munkát Balogh Lajos és Kiss Gábor irányította. Megemlítendő, hogy munkálataik kizárólag az atlaszadatok informatizálását tűzték ki célul, és ehhez nem kapcsolódott olyan – az általunk alkalmazotthoz hasonló – szoftver, amely a már informatizált adatok elemzését (statisztikai lekérdezések, keresés, analitikus térképek generálása stb.) is lehetővé tette volna. Balogh Lajos emellett részt vett abban a kutatásunkat megelőző folyamatban is, amely az atlasz soron következő, IV. kötetéhez kapcsolódott, és immár az előző pontban említett Bihalbocs lejegyzőszoftverrel történő informatizálásnak az előkészítését segítette elő.

Az előkészítést követően a IV. kötet közel 200 térképlapjának informatizálása már a Bihalbocs szoftverrel valósult meg projektünk első szakaszában. A második szakaszban az V. kötet informatizálását kezdtük meg, ebben periódusban fejeztük be az V. kötet adatainak rögzítését. Az adatbevitel után a projekt harmadik szakaszában végeztük el a kötet tételes ellenőrzését és a hibák javítását. A kötet adatainak teljes körű informatizálását (rögzítését, ellenőrzését és javítását) 2006 októberére fejeztük be. A VI. kötet informatizálását a már bevált „menetrendet” követve korábbi munkatársainkkal végeztük el 2007-ben, a kutatás utolsó szakaszában.

Az adatrögzítéshez kölcsönkértük az atlasznak a Nyelvtudományi Intézetben archivált kéziratát, és az adatokat ez alapján vittük számítógépre. Ezzel a kiadott térképlapok és az informatizált fájlok egyaránt a gépírási kéziratlapok alapján jöttek létre. Az informatizált fájlok az atlasz egy-egy térképlapjának felelnek meg, elnevezésük is utal a kiadott kötet lapjainak számozására (például a ma*0571.txt fájl a Nagyatlasz 571. térképlapjának adatait tartalmazza).

A kézirat és a kiadott térképlapok struktúrája egyetlen pontban tér el az informatizált változattól: a fő adatsorokat követő megjegyzésekben. A kiadott térképlapok terjedelmi megfontolásból egyes adatokat csak megjegyzésként adtak közre (ezeket „a térképről lezorzult adatok” címet követően a térképlapok legendájában tüntették fel). Az informatizált változatban ilyen terjedelmi megszorítások értelemszerűen nincsenek, ennél fogva ezeket az adatokat beillesztettük a fő adatsorba. Ugyancsak így jártunk el azokkal az adatokkal, amelyek azért kerültek csak a megjegyzések közé, mert a próbagyűjtés, a gyűjtés fő szakasza és az ellenőrző gyűjtés közül csak egyikben fordultak elő. Ezeket az adatokat azonban minden esetben lábjegyzettel láttuk el, a gyűjtésről szóló információt tehát más formában ugyan, de megőriztük.

A pályázatban vállaltaknak megfelelően legnagyobb nyelv atlaszunk három kötetének (IV–VI.) összesen közel 600 térképlapját (az 571.-től az 1162-ig) hoztuk korszerűen kutatható, elemezhető formába. Mivel a számítógépes rögzítés során a kéziratos eredeti példányokat használtuk, adatbázisunk a térképen megjelent kiadvánnyal összemérhető pontosságú adatokat tartalmaz.

1.3. A romániai magyar nyelvjárások atlasza V–VIII. kötetének informatizálása

A romániai atlasz (a továbbiakban: RMNyA.) informatizálási folyamatát projektünk első szakaszában az V. kötetel kezdük meg, majd a következő három pályázati szakaszban egy-egy kötet anyagát dolgoztuk fel, azaz a VI., VII. és VIII. kötetel folytattuk munkát. A kötetek anyagának informatizálása során megőriztük az adatokra vonatkozó minősítéseket, megjegyzéseket és az adattár nyomtatott változatában alkalmazott csoportosítást. Az így létrehozott adattár térképre vetíthető, kereshető és integrálható a munkánk során rögzített más adattárakkal. A négy kötet térképlapjait a Bihalbocs dialektológiai szoftverrel rögzítettük. (Az egyes fájlok elnevezése itt az rma1200.txt szerkezethez hasonlóan hordozza a kiadott atlaszkötet térképlapjainak számozását.)

A pályázatban vállaltaknak megfelelően a négy atlaszkötet összesen 1200 térképlapjának anyagát hoztuk korszerűen kutatható, elemezhető formába. Az ellenőrzés során az adatokat összevetettük Murádin László kéziratos térképeivel is, így lehetőségünk volt az adatbevitel alapjául szolgáló nyomtatásban megjelent változat esetleges hibáinak a javítására. A RMNyA. adataiból létrehozott adattár nem zárt rendszer: szükség szerint bármikor módosítható, tovább javítható, bővíthető. Az informatizált adatok kereshetők, tetszőleges szempontok szerint csoportosíthatók, integrálhatók más adattárakból származó, megfelelően rögzített nyelvjárási adatokkal.

1.4. A moldvai csángó nyelvjárás atlasza informatizálása

A moldvai csángó nyelvjárás atlasza (a továbbiakban MCsNyA.) az 1950-es években kolozsvári kutatók által gyűjtött anyagot ad közre. Az atlasznak projektünk szempontjából különös jelentősége van, mivel a diakrón nyelvföldrajzi kutatásban ezt az anyagot hasonlítjuk össze újabb gyűjtésű adatokkal. Ennélfogva a többi atlasz informatizálási folyamatáról szóló beszámolókhöz képest itt részletesebben foglalkozunk az atlasz számítógépre vitelének lépéseivel. Emellett röviden ki kell térnünk az atlasz létrehozásának történetére is, mivel ez szorosan kapcsolódik projektünk egyes feladataihoz.

A MCsNyA. informatizálása szoros összefüggésben áll kiadásának történetével. Az atlasz terepmunkája 1949–1962 között ment végbe, de legjelentősebb része már 1957-re megvalósult, ezt követően, 1962-ben csak ellenőrző gyűjtéseket végeztek.

A terepmunka fő szakaszát követően a kutatás vezetője, Szabó T. Attila arról számolt be, hogy 1957-re elkészült nagyjából 900 térképlap sokszorosításra alkalmas tusrajzban. Ez akkor nem a tervezett atlasz teljes anyaga volt, az összesen 1100–1200 térképlapból állt volna, de ez még mindig csak része lett volna a Moldvában gyűjtött teljes nyelvföldrajzi anyagnak.

A már megrajzolt és kiadásra előkészített 900 térképlap kiadására hosszú időn át nem adódott lehetőség, csak az 1990-es évek fordulóján megkezdődhetett a Budapestre eljuttatott térképlapok megjelentetése. A MCsNyA. I. és II. kötete 1991-ben jelent meg Budapesten, több évtizeddel az atlasz elkészültét követően. Sajnálatos módon azonban a kiadás folyamatába hiba csúszott: a nyomdában olyan térképlapokat hoztak létre, amelyek az adatok mellett nem tüntették fel a kutatópontok számát. Azaz a megjelent kötetekben hiányzott a közreadott adatok mellett a lokalizálásukat elősegítő számozás. Így a főként a Szeret mentén sűrűn egymás mellé települt moldvai falvak adatbeírásos módszerrel megjelenített nyelvjárási adatait csak nehezen lehetett pontosan egy-egy kutatóponthoz kötni, különösen akkor, ha egy-egy térképlap nagyszámú adatot közölt.

A már megjelent kötetek informatizálásának az volt az egyik speciális indoka, hogy lehetővé váljon a kutatópontok és az adatok pontos egymáshoz rendelése, és ennélfogva az atlasz felhasználását ne nehezítsék meg a papíron kiadott változat hiányosságai.

Az informatizált változatban az adatok lokalizálása már egyértelmű, ezzel az atlasz használatának eddigi technikai akadályai elhárultak. (De ez azzal is együttjárt, hogy a rögzített adatok ellenőrzése során jóval nagyobb körültekintéssel kellett eljárunk, mint más atlaszoknál, amelyekben a kutatópontok és az adatok egymáshoz rendelése első pillantásra is áttekinthető.)

A papírkiadású atlaszhoz képest az informatizált változat további újdonsággal is szolgál, és ez jóval nagyobb horderejű, mint az előző. Ehhez ismét vissza kell térnünk az atlasz létrehozásának történetéhez. Mikor a budapesti kiadás lehetősége felvetődött, az atlaszt kiadásra előkészítő szerkesztői kollektívából (Gálffy Mózes, Márton Gyula és Szabó T. Attila) már senki nem volt életben. Így elsikkadt az a tény, hogy az atlasznak van egy harmadik kötete is, amelyről ugyan a kiadvány előszavát jegyző Péntek János említést tesz, de rögtön meg is jegyzi, hogy „ebben (...) valójában kevés az új információt tartalmazó nyelvi anyag“ (MCsNyA. I: 5). Péntek János itt nem említi meg, hogy ez a kötet ekkor lappangott, és csak néhány éve került elő. A III. kötet sorsáról Murádin László elmondása alapján Tánzos Vilmos a következőket írta: „a kiadás időpontjában (1991) a 3. kötet kéziratát a szerkesztők sehol nem találták. A kézirat az

1987-ben elhunyt Szabó T. Attila könyvtárában egy szekrény mögé volt elrejtve, és csak évekkel később került elő, akkor, amikor a nyelvészprofesszor örököse a könyvtár anyagát az Erdélyi Múzeum-Egyesületnek adományozta, és a könyvtár elszállítására is sor került“ (Tánczos Vilmos: A moldvai csángók nyelvészeti kutatása (1945–2004). In Kiss Jenő (szerk.), *Nyelv és nyelvhasználat a moldvai csángók körében*. Budapest: Magyar Nyelvtudományi Társaság. 208–285. 2004. 264).

Az így előkerült III. kötetet 2005 tavaszán vettük át Péntek Jánostól Kolozsváron, és kezdtük meg a már számítógépre vitt első két kötetből álló adatbázis kibővítését az újabb adatokkal. A III. kötet felépítése azonban több szempontból is eltért az előbbi kötetektől. A különbségek közül a leginkább szembevetendő az volt, hogy a kiadatlan térképlapok jelentős része ún. „összesítő térképlap“ – ma úgy mondanánk: jelenségtérkép – volt, amely több térképlap adatait egyetlen térképlapra sűrítve mutatta be. Ezt szemléltetendő tekintsük meg a következő mutatványt, amely a III. kötet első öt térképlapjának címszavait adja közre:

(1) Mutatvány az MCsNyA. III. kötetének kéziratát alkotó térképlapok címszóiból

609. **emberem (609)**; *fecske* (520); **vejem (610)** – az *e* hang ejtése hangsúlyos helyzetben;
610. **kilencven (611)**; *este* (566); *egyenes* (297) – az *e* ejtése hangsúlytalan helyzetben;
611. *ég* <ige> (33); *ég* <fn.> (461); *éjfélkor* (569) – az *é* ejtése hangsúlyos helyzetben;
612. *fehér* (272); *gané* (317: trágya); **ehetném (612)** – az *é* hangsúlytalan helyzetben;
613. *pepeny* 'sárgadinnye' (356: sárgadinnye); *veverica* 'mókus' (545: mókus); **berbécs** 'kos' (613) – az *e*-féle hangok ejtése román kölcsönszavakban.

A III. kötet kézírata a 609. térképlappal kezdődik. Ez rögtön három adatsort összesít (az *emberem*, a *fecske* és a *vejem* címszó megfelelőit), és jelenségtérképként a köznyelvi *e* nyelvjárási megfelelőit vizsgálja hangsúlyos helyzetben. Az utána következő térképlapok is három-három címszó nyelvjárási megfelelőit adják közre. A kiemelés nélkül álló címszók nem újak, már jelen voltak az atlasznak vagy az I., vagy a II. kötetében; a félkövér betűvel szedett címszók ellenben a megjelent kötetekhez képest új információkat jelenítenek meg. A címszók után zárójelben álló számok az atlasz informatizált változatához kialakított megfelelő térképlapok sorszámait. Az első két kötetben ezek a hagyományos és az informatizált változat térképlapjai esetében azonosak egymással. (A 613. kéziratot térképlap arra példa, hogy nem feltétlenül a címszók és a térképlapok azonosságáról van szó, ugyanis itt a köznyelvi *mókus* nyelvjárási megfelelői közül a jelenségtérkép csak azokat ábrázolja, amelyek a román *veverită* 'mókus' átvételei. Ekkor az első két kötet térképlapjának számát követően megadjuk az ott feltüntetett címszót is.)

A harmadik kötet kéziratot és számítógépes változatának térképlapjai eltérnek egymástól. Ez abból a döntésünkből adódik, hogy nem követtük a kézirat szerkezetét, és a jelenségtérképek adatait nem egyetlen fájlban rögzítettük, hanem külön-külön adatsorokat alakítottunk ki. Például a kézirat 609. térképlapjából két új fájl hoztunk létre, az egyiket a 609. sorszámot viselő *emberem* címszóval, a másikat 610. számmal *vejem* címszóval. A *fecske* adatainak rögzítésére természetesen már nem volt szükség, mivel azokat már korábban számítógépre vittük a II. kötet informatizálásának részeként.

A harmadik kötet anyaga azonban nem kizárólag az első két kötet adataiból készített jelenségtérképekből áll, hanem számos új térképlapot is tartalmaz. Ha

összevetjük a már megjelent kötetek anyagával, akkor ebben a kötetben a két kötet együttes anyagát megközelítő mennyiségű új térképlap található: az MCsNyA. I. kötetében 295, a II.-ban 303, a III.-ban 447 térképlap van. (A legutóbbi kötet papíron létrehozott változatában a térképlapok száma ennél kisebb, mivel számos új adat összesített térképlapokon jelenik meg.)

Ami a most elkészült atlaszból kimaradt, jelenleg a kolozsvári egyetem Magyar Nyelvészeti Tanszékén található meg kicédulázva, de más moldvai kutatások szóanyagával együtt – a készülő Csángó Szótár adatbázisának részeként. Ennek az anyagnak a nyelvföldrajzi hasznosítására jelen kutatásunkban nem vállalkozhattunk.

A MCsNyA. harmadik kötetét a projekt keretében adjuk közre; a kötet informatizált változata ilyen módon a nyelvatlaszok papíron kiadott és a számítógépre vitt változatai közötti hagyományos sorrendel szakítva hamarabb jött létre, mint a papírkiadású változat. (Egyben a számítógépes nyelvföldrajzi kutatások általánossá válásával feltehetőleg már most kiváltja az informatizált változat előállításához képest sokkal költségesebb nyomdai kivitelezéssel előálló hagyományos atlaszkiadványt.)

A MCsNyA. három kötetének informatizált változata alapján Bodó Csanád és Vargha Fruzsina Sára készítette el a *Jelenségtérképek A moldvai csángó nyelvjárás atlaszából* című CD-kiadványt (bibliográfiai adatait lásd az 1. mellékletben). E kiadvány ötven olyan jelenségtérképet ad közre, amelyek valamilyen dialektológiai szempont alapján értelmezik a papíron megjelent térképek adatait, többnyire több térképlap adatainak integrálásával. A CD-ROM részeként a MCsNyA. három kötetének teljes anyagáról is közlünk információkat: közreadjuk a címszavak listáját a megjelenés sorrendjében és betűrendben, valamint az atlasz összes adatát összesítő Excel-táblázatokat. Mivel az atlasz mostanáig lappangó és általunk informatizált III. kötetének nem készült hagyományos adathordozón (papíron) hozzáférhető változata, ezzel a CD-kiadvánnyal tesszük közzé ezt az anyagot.

1.5. A Székely Nyelvföldrajzi Szótár informatizálása

A Székely Nyelvföldrajzi Szótárról tudjuk, hogy szótáraink és nyelvatlaszaink közötti átmeneti jellegű kiadvány, amely a kiadásra előkészített székely nyelvatlasz adatainak egy részét szótárformában adja közre. Eredetileg azonban a már kiadott anyag térképekre vetítve jelent volna meg, de az atlasz-formátumról anyagi és gyakorlati megfontolásokból le kellett mondani a kiadás során.

A szótár anyagának informatizálása az anyag atlasszá alakítását is jelenti. A számítógépre vitt adatok – a papíron kiadott változat struktúrájának megőrzése mellett – arra is alkalmasak, hogy nyelvatlasz-térképeket generáljunk belőlük. A szócikkfájlokat a kiadott szótár alapján ellenőriztük és javítottuk.

2. A moldvai diakrón nyelvátlasz multimédiás nyelvátlasz

A moldvai diakrón nyelvátlasz létrehozásának két fő feladata – a kutatás megtervezését és a próbagyűjtéseket követő szakaszban – a következő volt:

1. a diakrón nyelvátlasz terepmunkája;
2. a felvett anyag folyamatos feldolgozása.

2.1. A diakrón nyelvátlasz terepmunkája

Empirikus kutatásunk előkészítése során a diakrón nyelvátlasz elméleti megalapozásából adódóan a következő feladatokat kellett megvalósítanunk:

1. az adatközlők kiválasztásának szempontjai;
2. a felmérés kérdőívének kidolgozása;
3. a kutatópontok meghatározása;
4. a terepmunka megvalósítása.

E feladatok megvalósítása során a diakrón nyelvátlasz jellegéből adódó szempontokat követtünk, mint a továbbiakban részletesen is bemutatjuk.

2.1.1. Az adatközlők

Vizsgálatunk eltért a hagyományos nyelvátlaszok szociológiailag kevésbé rétegzett mintavételi technikáitól, így például a MCsNyA. eljárásától is, amelynek megfelelően a terepmunkások falvanként csak egy-két adatközlőt – ún. nyelvmentert – kérdeztek ki. Mivel a diakrón nyelvátlasz munkálatai a nyelvi változások vizsgálatát helyezik előtérbe, a korábbi atlaszadatokat olyan újabb gyűjtésű anyaggal vetjük össze, amely a jelenkori életkori rétegződés alapján teszi lehetővé a nyelvi mozgás, a folyamatban lévő nyelvi változás vizsgálatát. Ebből adódóan az adatközlők kiválasztásának elsődleges szempontja az életkor változója. Ehhez három életkori csoportot különítettünk el:

- A. idősebbek: 55 évnél idősebbek;
- B. középkorúak: 25–54 évesek;
- C. fiatalabbak: 24 évesnél fiatalabbak.

Az életkor mellett figyelembe vesszük a társadalmi nem változóját is, e szempont szerint mindhárom életkori csoportot egy-egy további alcsoportra oszthatjuk. Ezekben az alcsoportokban kutatópontonként 3–3 adatközlőt kérdezzük ki, azaz összesen egy településről 18 adatközlőnk van, ha a beszélők életkori csoportjai minden más feltételnek is megfelelnek.

Ezek a feltételek a napjainkban a moldvai magyar nyelvű közösséget érő társadalmi és nyelvhasználati változásokkal állnak összefüggésben. Ezek közül a leglényegesebb, hogy számos moldvai beszélőközösségben az utóbbi évtizedekben nyelvcsere van folyamatban, és a fiatalabb korosztályok már nem tanulnak meg magyarul. Az adatközlők kiválasztásának ebből adódóan fontos kritériumává vált a magyar nyelvismeret. Itt azt az eljárást követjük, hogy nem veszünk fel attól a korosztálytól adatokat, amelyben teljesen általános a román nyelv használata.

Kutatópontjaink egy részén a 18 adatközlő helyett csak 12 beszélőt kérdeztünk meg, ha a fiatalabbak között már nincsenek magyar nyelvűek, máshol pedig csak 6 adatközlőnk volt, ha 55 év alatt a beszélők már mind román egynyelvűek.

A nyelvtudás szempontja mellett figyelembe kellett vennünk a külföldre irányuló migráció hatásait is. Ez a változás az 1990-es évektől kezdődően jelentősen átalakította egy-egy helyi közösség életkori struktúráját. A külföldi migráció eredménye, hogy a fiatalabb korosztályokban sokan már akár egy évtizede is távol élnek a falujuktól, és ugyan rendszeresen hazajárnak, de közösségi szerepük már igen csekély. A külföldi munkavállalás intenzitásából és természetéből adódóan itt rendkívül sokféle változattal számolhatunk, így azt a megoldást választottuk, hogy azokat tekintettük potenciális adatközlőnek, akik a faluban élnek vagy rendszeresen hazajárnak. Olyan beszélőkkel azonban, akik más településen – akár külföldön, akár másutt – már letelepedtek, nem készítettünk interjút.

Összefoglalóan: a felmérésben kvótamintát alkalmazunk, azaz életkor és nem szerint elkülönített csoportokat vizsgálunk. Adatközlőink magyarul tudó, helyben felnőtt beszélők, akik a vizsgált településen élnek vagy oda rendszeresen visszajárnak.

2.1.2. A kérdőív

A kérdőív kialakításával kapcsolatban két kérdést érdemes megemlíteni: egyrészt a korábbi atlasz címszavaiból való válogatás szempontjait, másrészt a kérdőív felépítésének és a kérdések megfogalmazásának kritériumait.

A korábbi atlasz kérdőíve több mint ezer címszót tartalmazott. Ezeket az adatokat falvanként csak egy-két adatközlőtől kérdezték ki. Ha az adatközlők számát jelentősen megnöveljük, gyakorlati szempontból nem tartható fenn a kérdések eredeti száma, ugyanis a terepmunka és a feldolgozás időtartamát és egyben költségeit a nagy mennyiségű adat megsokszorozná (az idő- és anyagi ráfordítás pedig már meghaladná a projekt kereteit). Ebből következik, hogy az eredeti gyűjtés kérdéseinek körét csökkentenünk kellett, és végül a MCsNyA. címszavainak csak valamivel több mint egyötödét vettük át az újabb gyűjtés kérdőívébe.

A kiválasztás szempontjai háromfélék voltak: egyrészt nem vettük fel azokat a címszavakat, amelyek már kevésbé ismert fogalmakra vonatkoznak. A fogalom, tárgy, jelenség elavulttá válása ezekben az esetekben megjósolható módon jár együtt a megnevezésére szolgáló szó használatának visszaszorulásával, így a diakrón elemzés elsősorban nem a nyelv változásáról, hanem a tárgyi világ, a helyi kultúra stb. átalakulásáról, valamint annak külsődleges nyelvi vonatkozásairól tudósítana. Ezek a jelenségek például a kendermunka, a szövés és a fonás népi kismesterségéhez kapcsolódnak. Ez a regiszter a MCsNyA.-ban alaposan dokumentálnak tekinthető, az atlasz több mint tíz térképlappal reprezentálja az ide tartozó szakszavakat.

Másrészt csak kisebb számban vettünk fel olyan szavakat, amelyeknek a térképlapját a korábbi atlaszban nem vagy csak igen csekély mértékben jellemezte variabilitás.

Harmadrészt a főként a harmadik kötetben megjelenő grammatikai jelenségek közül is kiválogattunk néhány jellemző példányt. Például az igeragozási jelenségeket vizsgáló térképlapok címszavai közül csak egy produktív és egy terméketlen igei tőtípust vettünk fel a kérdőívbe, továbbá a paradigmájukat vizsgáló térképlapok közül is csak a fontosabbakat hagytuk meg az új felmérésben.

Az eddig tárgyalt szempontok alapján kialakított kérdőív két részből áll: az atlasz-címszavakra rákérdező rész mellett szerepel benne egy szociolingvisztikai modul is, amely az adatközlő kétnyelvűségére (nyelvválasztás és nyelvtudás) vonatkozó kérdéseket tartalmaz, továbbá szociológiai jellemzőit méri fel (kor, foglalkozás, iskolázottság stb.). A kérdőív szociolingvisztikai részének kialakításában támaszkodtunk „A magyar nyelv a Kárpát-medencében a XX. század végén“ című kutatási program kérdőívére is. Ezzel Moldvában gyűjtött adataink egy része összehasonlíthatóvá válik más magyar kisebbségi csoportok nyelvszociológiai jellemzőivel.

Mivel kutatásunk az 1940–1950-es években gyűjtött nyelvi anyagot hasonlít össze újabb gyűjtésű nyelvjárási anyagokkal, a felmérés módszerei értelemszerűen megegyezhettek volna a korábbi atlasz terepmunkája során használt eljárásokkal. A MCsNyA. adatainak gyűjtéséhez azonban a kutatók nem alakítottak ki kérdőívet, hanem kérdéseiket korábbi tapasztalataik alapján mindig helyben, az aktuális gyűjtési helyzetnek megfelelően fogalmazták meg. Kutatásunk során eltértünk ettől az eljárástól, és olyan standardizált kérdőívet készítettünk, amelyben különböző feladattípusok szerepelnek (képen látható tárgy megnevezése, rámutatás adott tárgyra, indirekt kérdés, fordítás románról magyarra).

A kérdőívet először próbagyűjtésen teszteltük. Ehhez eltérő dialektológiai hátterű moldvai közösségeket kerestünk fel, és olyan moldvai székely és csángó nyelvjárási beszélőket kérdeztünk meg, akikkel az általunk kialakított kérdőív alkalmasságát többféle életkori csoportban is ellenőrizni tudtuk. A felkeresett települések Nagypatak, Klézse és Lészped voltak.

2.1.3. A kutatópontok

A kutatópontok kijelölésének fő szempontjai kétfélek: egyrészt törekszünk arra, hogy a korábbi atlasz kutatópontjai közül hálózatunkba minél többet felvegyünk, másrészt a nyelvi változások területi terjedésének elemzéséhez a korábbi atlasznál sűrűbb kutatópont-hálózatot alakítunk ki.

Az első szempont korlátja az lehet, hogy egyes településeken a magyarul beszélő lakosság száma már olyan csekély, hogy a felmérés adatközlők hiányában nem valósulhat meg. Ezt a veszélyt elkerülendő nem vettünk fel olyan településeket kutatópontjaink közé, amelyeken újabb becslések szerint 20 százaléknál kisebb a magyarul beszélő lakosság aránya. Ennélfogva a kutatást az MCsNyA. azon kutatópontjain kezdtük meg, ahol Tánczos Vilmos, kolozsvári néprajzkutató becslése szerint a magyarul beszélő lakosság nagyobb arányban van jelen (A moldvai csángók lélekszámáról. In Pozsony F. (szerk.), *Csángósors. Moldvai csángók a változó időkben*. H. n. [Budapest]: Teleki László Alapítvány. É. n. [1999.] 7–32).

A Tánczos Vilmos által felsoroltak között vannak olyan települések, ahol a magyarul beszélő lakosság arányának meghatározása újabb felmérést igényelt. Ezt a vizsgálatot a kutatópont-hálózat végleges meghatározásához a kutatás kezdeti szakaszaiban végeztük el.

A kutatópontok kiválasztásának másik szempontja a tömbben és a szórványban élő közösségek megkülönböztetése volt. Itt abból a feltevésből indulunk ki, hogy a tömbben élő közösségekben az azonnyelvi nyelvjárások közötti érintkezés hatásai jelentősebbek, mint a szórványban. Ennélfogva a korábbi hálózatot kiegészítettük olyan kutatópontokkal, amelyek részletes adatokat biztosíthatnak a nyelvjárási jelenségek

területi terjedéséről vagy visszaszorulásáról. Kutatásunk során két nagyobb tömbként a Szeret és a Tázló menti nyelvjárásokat vizsgáltuk, mivel az északi (Szabófalva környéki) és a Tatros menti településekhez képest az előbbieket magyarnyelv-használati szempontból viszonylag homogénebb tömböket alkotnak.

2.1.4. A terepmunka megvalósítása

Kutatásunk első szakaszában hat településen készítettük el 18–18 adatközlővel a nyelv földrajzi interjút. Ezek a települések Klézse, Pusztina, Frumósza, Újfalva (a Tatros mentén), Trunk és Bahána. Közülük a legutolsó település kivételével mindegyik falu a régebbi atlasz kutatópontja is volt. Csak 10 adatközlőnk volt Szolonckán, ahol a középkorú korosztályban sem teljes körű már a magyar nyelv használata, valamint 8 interjú készült Pokolpatakon. Egy további kutatóponton, Szőlőhegyen a felmérés megkezdődött, de csak négy beszélővel végeztünk adatfelvételt.

A projekt második szakaszában a terepmunka Somoskán és Szitáson folytatódott, mindkét településen 18–18 adatközlővel, valamint befejeződött Szőlőhegyen (14 további adatközlővel).

A projekt harmadik szakaszára esett a moldvai terepmunka második nagyobb szakasza. Ebben a kutatási periódusban a következő településeken folyt terepmunka: Szabófalva (itt 5 interjú készült el), Lészped, Lujzikalagor, Gajdár, Bogdánfalva, Gyoszény, Ketris, Szalánc (Ciresoaia), Onyest, Gorzafalva és Vizánta.

A kutatás utolsó szakaszában terepmunkát folytattunk Nagypatakon, Esztufujban, Lábnikon és Dózsában.

A projekt során összesen 26 települést kerestünk fel, és mindenhol a korábbi beszámolómban kifejtett elveknek megfelelő számú adatközlőt kérdeztünk ki (kivéve Szabófalvát, valamint Pokolpatakot, ahol csak 8 felvétel készült). Kutatópontjainkon összesen 408 adatközlővel készítettünk interjút.

2.2. A felvett anyag feldolgozása

A terepmunka alatt felvett anyag feldolgozása folyamatosan zajlott. Itt két részfeladatot kellett megoldani: egyrészt standardizáltuk az adatok és kísérő információik kódolását, másrészt pedig kialakítottuk azt az informatikai keretet, amelyben a nyelvjárási adatok hangzó és átírt formája egymással összekapcsolódik.

Az előbbi feladathoz a MCsNyA. eljárásával azonos módon a magyar egyezményes átírást választottuk, de az egyes adatok minősítését már a korábbi atlaszhoz képest részletesebben oldottuk meg. Ez azt jelenti, hogy egyrészt minden egyes adathoz hozzátartoznak az adatközlő szociológiai és nyelvészeti jellemzői (pl. életkora, neme, esetleges magyarországi tartózkodásának időtartama; valamint nyelv választási szokásai, nyelvtudása, a magyar nyelv helyi és más változataihoz fűződő attitűdjei), másrészt pedig azt, hogy az adatközlő által adott minősítések ugyancsak megjelennek a lejegyzett szövegben az adatot követően.

A projekt keretében folytatott terepmunkából származó adatokat olyan formátumban tároljuk, amely lehetővé tette, hogy a Bihálbocs dialektológiai szoftver további fejlesztésével kardinális multimédiás nyelvjárási atlaszt hozzunk létre – ebben a kutatóponthoz nem csak a lejegyzett adatot, hanem annak hangzó formáját is hozzárendeljük.

Az adatok lejegyzéséhez egy általunk kialakított fájlformátumot használtunk, amelyet BXT-nek nevezünk. A BXT-t abból a célból fejlesztettünk ki, hogy lehetővé tegyük lejegyzett nyelvjárási adatokat tartalmazó szövegfájlok és WAV formátumú hangfájlok szinkronizálását. Vagyis a lejegyzett szövegek nem csupán magukat a nyelvjárási adatokat tartalmazzák, hanem időzítési markereket is, amelyek a vonatkozó hangfájlok megfelelő időpillanataira utalnak a másodperc töredékének pontosságával. Az ebben a formátumban tárolt anyagok tehát lehetővé teszik a markerekkel megjelölt adat gombnyomásra történő meghallgatását az adott hangfájl megkeresése és megnyitása nélkül.

A BXT-fájlok szerkezete jól strukturált és viszonylag egyszerű, így az ebben a formátumban tárolt anyagokat a jövőben esetleg más, erre a célra kifejlesztett (dialektológiai) szoftverek is kezelni tudják majd.

A kutatás harmadik szakaszában készített felvételek teljes körű feldolgozása 2007 áprilisára fejeződött be. Ezt követően a felvételek ellenőrzése mellett a kutatás utolsó szakaszában felvett nagypataki, esztufuji, lábni és dózsai interjúk átírása következett.

A kutatásunk során gyűjtött anyag feldolgozásának lezárultával lehetővé vált a MCsNyA. adatainak és az újabb gyűjtésű anyagnak egyetlen közös adatbázisban való egyesítése. Ezzel létrejött a multimédiás diakrón moldvai nyelvátlasz. Az atlasz bemutatására Bodó Csanád vállalkozott az V. Dialektológiai Szimpozionon, Szombathelyen, 2007 augusztusában. Az előadás demonstrációs anyaga megtekinthető a http://geolingua.elte.hu/about/eload/bodocs_VDialSzimp.ppt.htm címen.

3. A magyar nyelvjárások atlasza munkálatai során készített hangfelvételek feldolgozása

A MNyA. ellenőrző gyűjtései során, az 1960-as évek első felében készített hangfelvételek feldolgozását két részfeladatra osztva végeztük el. Ezek a következők voltak:

1. A szövegfelvételek digitalizálása, archiválása és dokumentálása;
2. A kérdőíves adatok beillesztése az atlasz informatizált változatába; valamint szövegfelvétel-válogatások CD-n való megjelentetése.

A két részfeladat egyaránt folyamatosan, a projekt teljes időtartama alatt zajlott; az utóbbi részfeladatban egy CD-sorozat első három kiadványának vállaltuk évenként egy-egy kiadvánnyal. A sorozat harmadik darabja 2007 nyarára készült el.

3.1. A MNyA. szövegfelvételeinek digitalizálása, archiválása és dokumentálása

Az első részfeladat elvégzése során a beszámolási időszakban folytatódott és befejeződött A magyar nyelvjárások atlaszának az 1960-as években végzett ellenőrző gyűjtése során magnetofonszalagra felvett archív szövegfelvételei digitalizálása és CD-re írása. A munka sikeres elvégzését az MTA Nyelvtudományi Intézete és Zenetudományi Intézete közötti megállapodás tette lehetővé azért, hogy az elvégzendő feladathoz igénybe vehettük a Zenetudományi Intézet elektronikus berendezéseit és az ezek működtetéséhez szükséges infrastruktúrát.

A digitalizált hanganyagok fájlformátuma (WAV, 44100-as mintavételezés) a tudományterület legigényesebb elvárásainak is megfelel, és biztosítja, hogy a megismételhetetlen hangfelvételek minősége kicsit se csorbuljon. A viszonylag nagyobb tárhelyigény napjainkban – a nagy kapacitású adattároló eszközök árának egyre gyorsuló ütemű csökkenésével – már a legtöbb intézmény számára nem jelent megoldhatatlan problémát. További, legalább ennyire releváns szempont, hogy az akusztikai fonetikai elemzésekhez használatos szoftverek (például a Praat) nem olvasnak be tömörített (például mp3) formátumú fájlokat, a wav-formátumot ellenben minden elemző program támogatja. A veszteségesen tömörítő fájlformátumok (pl. mp3) alkalmazása nem ajánlatos a terepmunkával gyűjtött adatokat is felhasználó tudományterületeken (pl. fonetika, dialektológia, szociolingvisztika). Mi is a wav-változatot használjuk fel az interjúk lejegyzésének és hanganyagának összekapcsolása során.

Az oktatási és ismeretterjesztő szintű internetes anyagok esetében viszont egészen más szempontok érvényesek. Itt fel sem merül az akusztikai vizsgálat igénye, annál fontosabb viszont, hogy ezek az anyagok könnyen hozzáférhetők, megfelelő sávszélesség esetén pillanatok alatt letölthetők, meghallgathatók legyenek. Ennélfogva ezekben az anyagokban mi is mp3-formátumú hangfájlokat használunk.

A digitalizálási munkálatok lezárultával kifejezetten oktatási és népszerűsítő céllal olyan internetes felületet hoztunk létre, amelyen keresztül az érdeklődők részleteket hallgathatnak meg a MNyA. ellenőrző gyűjtése során készített interjúkból. A közzétett részleteket az egyes kutatópontokhoz kapcsolva, kattintható térkép segítségével lehet meghallgatni. Így az ellenőrző gyűjtés teljes kutatóponthálózatát alapul véve mind a 352 kutatópontról közlünk egy-egy hangfelvétel-részletet mp3-formátumban. Reményeink szerint az így létrehozott internetes nyelvjárási hangmintatár hatékonyan segíti majd az egyetemi és főiskolai hallgatókat (sőt, esetleg a közoktatásban résztvevőket is) a magyar nyelv területi sokszínűségének és az egyes nyelvjárások jellemzőinek megismerésében.

A hangminták megtekintéséhez a Geolingvisztikai Műhely honlapján a Térképek rovatot kell felkeresni, illetve közvetlen kattintással a következő címen található meg az adatokat: http://geolingua.elte.hu/maps/MNyA_hangosterk/index.html.

3.2. A kérdőíves adatok beillesztése az atlasz informatizált változatába és egy szövegfelvétel-válogatás megjelentetése

A részfeladat elvégzése során a restaurált és digitalizált hangfelvételeken rögzített kérdőíves kérdések bekerültek az atlasz informatizált változatába. Emellett elkészült három olyan szövegfelvételeket tartalmazó – oktatási és kutatási célokra egyaránt alkalmas – CD-ROM is, amelyek a Magyar nyelvjárási hangoskönyv című sorozatunk első kiadványai (bibliográfiai adataikat lásd az 1. mellékletben).

A kiadványok az eddig digitalizált anyagból olyan felvételeket adnak közre, amelyeknek a fonetikai átírása már megjelent a Hajdú Mihály és Kázmér Miklós által készített Magyar Nyelvjárási Olvasókönyvben (Bp., 1974.). Az Olvasókönyv átírt szövegeit és a hangfelvételeket együtt közreadó multimédiás CD-ROM-ok a magyar nyelvterület nyugati és északi régióiból származó adatokat tartalmaznak. A sorozat kiadványaihoz két változatot hoztunk létre: a kutatói változatot, valamint az abból a Bihalbocs szoftverrel részben automatikusan generált ismeretterjesztő változatot.

A sorozat első három darabja így összesen 14 kutatópontról ad közre nyelvjárási szövegfelvételeket. Mindhárom kiadvány internetes változata megtalálható a Geolingvisztikai Műhely honlapján (<http://geolingua.elte.hu>) a Hangoskönyv rovatban. Rögzítettük további 20 település anyagát is, szintén az eredeti hanganyaggal szinkronizálva, ezekhez a szövegekhez egyelőre kutatói változatban lehet hozzáférni.

A Hangoskönyv lényeges jellemzője, hogy az eredeti hangfelvételeket és fonetikus lejegyzésüket a szövegben elhelyezett időzítési markerekkel szinkronizáljuk. Hanganyag és lejegyzés kapcsolata az Interneten is közzétett oktatási és népszerűsítő változatban is megőrződött. Lényeges különbség van azonban a két változat között a struktúrát és a hanganyag minőségét, valamint már említett formátumát illetően. A kutatói változat interjúként egyetlen, speciálisan erre a célra kialakított, a lejegyzés mellett minden más szükséges információt is tartalmazó fájlból és a hozzá kapcsolódó (egy vagy több hosszú) hangfájlból áll, ahol a kapcsolódásokat a tárolt információk alapján a geolingvisztikai szoftver (Bihalbocs) biztosítja. A platformfüggetlen, internetes változat, amelyet a kutatói változathoz hozunk létre, ennél jóval bonyolultabb szerkezetű (az internetes böngészők lehetőségeinek függvényében), mivel itt a (rövid) hangrészletek meghallgatását a hangfájlok és az eredeti lejegyzés feldarabolásával tudjuk csak biztosítani. Így interjúként egyrészt közreadjuk a teljes lejegyzést és a hozzá tartozó, meghallgatható interjút (mp3-formátumban), másrészt annyi szövegrészletet tartalmazó HTML-fájlt és hozzájuk kapcsolódó, csak az adott részletet tartalmazó (mp3-formátumú) hangfájlt, ahány részletet korábban (a kutatói változat elkészítésekor) szinkronizációs pontok megadásával kijelöltünk. Így az internetes változat sok száz vagy akár sok ezer hangfájl és sok száz más szövegfájl rendszeréből áll. Kialakításához a geolingvisztikai szoftver – megfelelő beállítások után – automatikusan végzi el a feladatok tekintélyes részét a kutatói változat alapján (például a hangfájlok feldarabolását, a lejegyzések feldarabolását és html-formátumba való konvertálását). Nem speciális szoftverekkel (hangszerkesztővel, szövegszerkesztővel, html-szerkesztővel) szinte vállalhatatlanul nagy feladat lenne ilyen termékek előállítását.

A kutatói változat alapján elkészített népszerűsítő (oktatási és ismeretterjesztő) változatokat honlapunkon is hozzáférhetővé tettük, az egyes kiadványok közül például az első CD anyaga a http://geolingua.elte.hu/hkonyv/Hangoskonyv1_index.html címen érhető el. A mellékjeles lejegyzések megtekintéséhez a Firefox, Opera vagy Safari internetes böngésző használata szükséges.

A MNyA. ellenőrző gyűjtésének hanganyaga a szövegfelvételek mellett tartalmazza annak a 41 atlaszkérdésnek a válaszait is, amelyek a papíratlasz elkészítése során nem kerültek térképre, korszerű, a hangzó anyagot alapul vevő feldolgozásukra annak idején nem volt lehetőség. Az atlaszkérdésekre adott válaszok rögzítésére jelen pályázat keretein belül mi is csak részlegesen vállalkozhattunk. Az adatok feldolgozását a Magyar nyelvjárási hangoskönyv sorozat kialakításának szempontjaihoz is igazodva olyan településekről származó anyagokkal kezdtük meg, amelyek egyaránt kutatópontjai voltak a Magyar nyelvjárási olvasókönyvnek és a Nagyatlasznak is.

Az atlaszadatok rögzítése során a moldvai multimédiás atlasz interjúinak feldolgozásához kialakított technológiát alkalmaztuk. A feldolgozás alapegysége az interjú, így az egyes adatok nem közvetlenül kutatóponthoz, hanem adatközlőhöz kötődnek. Minden adat megtartja kapcsolatát eredeti hangzó formájával. Ezt úgy biztosítjuk, hogy szinkronizációs pontokat (időzítési markereket) helyezünk el a szövegben, amelyek egy adott hangfájl adott időpillanatára mutatnak. Az eredeti hangzó

forma elérhető marad az adatok adatbázisba rendezése után is, így valamennyi adat azonnal meghallgatható, a fonetikus lejegyzés bármikor ellenőrizhető és javítható.

Az eredeti szövegfájlokból automatikusan generált adatbázishoz kialakított keresőfelület segítségével általunk megadott szempontok szerint csoportosíthatjuk és térképezhetjük az adatokat. Lehetőségünk van az adatok és a keresendő kifejezés lejegyzési finomságát pontosan meghatározni. Ez azt jelenti, hogy bár a lejegyzés során a MNyA. gyakorlata szerint járunk el, az adatok feldolgozásakor általunk megadott szempontok szerint tekinthetünk azonosnak vagy eltérőnek például egy diftongust és egy hosszú monoftongust, egy adott hang nyíltabb vagy zártabb változatát. Mivel az adatok megőrzik kapcsolatukat az adatközlővel az adatbázisba rendezés után is, lehetőségünk van az adatközlők paramétereire szerinti keresésre. Egyetlen beállítás megváltoztatásával készíthetünk térképet az általunk vizsgált nyelvi jelenség nemek vagy kor szerinti megoszlásáról.

A MNyA. ellenőrző gyűjtésének atlaszadataiból kialakított adattárról és elemzési lehetőségeiről Vargha Fruzsina Sára tartott előadást az V. Dialektológiai Szimpozionon, Szombathelyen, 2007. augusztus 22-én. Az előadás HTML-változata megtekinthető a http://varghafr.web.elte.hu/VarghaFr_Szombathely2007.htm címen.

4. Szabó T. Attila háromszéki helynévgyűjtésének informatizálása

Szabó T. Attila erdélyi helynévtörténeti adattára háromszéki kötetének informatizálása a második kutatási periódusban kezdődött meg és a harmadikban fejeződött be. A következőkben áttekintjük Szabó T. Attila helynévgyűjtésének történetét, majd az informatizálás eddig lépéseit.

4.1. Szabó T. Attila erdélyi történeti helynévgyűjtéséről

Kelet-Magyarország és Erdély területéről a történeti földrajz és a helynevek kutatói viszonylag kevés adatot tudtak közzétenni, mivel a legkorábbi évszázadokból az ország más területeihez képest kevesebb írásos emlék maradt fenn. A történeti névkutatás legfontosabb forrásmunkái (Györffy, Csánki történeti földrajza, Pesty Frigyes helynévgyűjtése) mellett ezért égető fontosságú Szabó T. Attila adattárának közzététele.

Az anyag létrejöttéről Szabó T. Attila leírja, hogy eleinte szülőföldjének, Désnek és vidékének helyneveit igyekezett minél nagyobb alaposággal összegyűjteni a levéltári forrásokból. A nagyszámú helynevet tartalmazó irattípusok (birtokösszeírások, határpercek) megismerése után már Erdély legkülönbözőbb vidékeiről álltak rendelkezésére adatok, így felvázolódott előtte egy nagyszabású, egész Erdélyre kiterjedő, történeti helynév-adattár ötlete. Az évtizedek alatt összegyűlt cédulaanyag végül nem egységesen oszlott el: a legtöbb történeti helynév a közép-erdélyi megyékből: a hajdani Kolozs, Szolnok-Doboka, Torda-Aranyos és Maros-Torda vármegyékből, valamivel kevesebb a Székelyföldről: Csík, Háromszék és Udvarhelyszék területéről, továbbá Alsó-Fehér vármegyéből került elő.

1958-as írásában Szabó T. Attila 650 000-re becsülte a névadatok számát. A cédulákat településenként rendezve 65 dobozban tárolta. Az 1980-as évek közepén adott interjújában már 700 000 történeti helynevet említett. Ennek a hatalmas anyagnak a kiadását évtizedeken át politikai okok nem tették lehetővé. Az Erdélyi Magyar

Szótörténeti Tár szócikkei közé azonban helynévi adatokat is fölvetett, ha azok egy-egy közszó története szempontjából fontos adalékkal szolgáltak.

Szabó T. Attila pontosan kifejtette egyik cikkében a Helynévtörténeti Adattárra vonatkozó gondolatait, később a kiadás előkészítésének tervét is összefoglalta. Bízott benne, hogy a folyamatos jelenkori gyűjtések is illeszkednek majd a történeti adatokat tartalmazó kötetekhez, és a létrejövő korpusz a magyar névtani vizsgálatok egyik reprezentatív, jól használható alapanyaga lesz. Fölvetette a névadatok térképre vetítésének, névföldrajzi elemzésének lehetőségét is.

Szabó T. Attila halála után a hatalmas cédulaanyagot tartalmazó dobozok az Országos Széchényi Könyvtárba kerültek. A gyűjtemény rendezéséhez, a kiadás előkészítéséhez Szabó T. Ádám fogott hozzá, de korai halála megakadályozta a munka befejezésében. A kiadás 1999-ben kapott új lendületet, Hajdú Mihály irányításával indult meg munka.

A kiadványsorozat Szabó T. Attila terveinek megfelelően megyénként kötetekbe rendezve közli a történeti helyneveket. A települések betűrendben követik egymást, a hozzájuk kapcsolódó helynévanyag időrendben áll. A helységnevek román és esetleges német nevét is közöljük. Az adatok legnagyobb része a 17–18. századból való, de akadnak 15. és 16. századi adatok is, ez utóbbiak főként az 1590-es évekből, a magyar nyelvű oklevelezés fölélnkülésének idejéből valók. Kisebb számban előfordulnak 19. századi adatok is. A gyűjteményből olykor előbukkannak az 1900-as évek első feléből gyűjtött szinkrón helynévgyűjtések is, melyek Szabó T. Attila szándékának megfelelően a történeti neveket kiegészítve belekerültek a kötetekbe. A névanyag döntő része a magyar névrendszerhez tartozik. Viszont magába foglalja a némely területen fölbukkanó más nyelvű, elsősorban román neveket is. Így alkalmassá válik román-magyar összehasonlító névvizsgálatok folytatására.

Az adatok kronologikus rendezése látványosan mutatja egy település helynévrendszerének változásait az idők során, sorról sorra követhetjük egy-egy név kisebb hangtani, morfológiai változásait, esetleg eltűnését, vagy új nevek keletkezését.

Hoffmann István tudománytörténeti munkájában a következőket jegyzi meg a gyűjteményről: „Nemrégiben megindulhatott Szabó T. Attila kivételes gazdagságú cédulaanyagának, az *Erdélyi Helynévtörténeti Adattár*-nak a feldolgozása és kiadása. [...] A szerkesztésben azt a módszert követik, amelyet Szabó T. Attila maga alakított ki a történeti helynévanyag közreadására: a nevek településenként, az egyes források időrendje szerint szerepelnek. Mutató nem készül a kötetekhez, ezek összeállítását ugyanis nagyon megnehezítené, hogy a neveket a szövegrészekben nem könnyű elkülöníteni. A kiadvány használatát jelentősen elősegítené azonban, ha elektronikus formában is hozzáférhetővé válna” (Magyar helynévkutatás 1958–2002. Debrecen, 2003. 58–59). Hoffmann István a mutató és a digitális változat hiányát már több fórumon joggal hangoztatta, véleménye a szaktudomány álláspontját tükrözi.

4.2. A háromszéki anyag informatizált változatának kialakítása

Szabó T. Attila helynévgyűjtéséből elsőként a háromszéki adatokat tartalmazó kötet informatizálására és részleges feldolgozására vállalkoztunk. Mivel korábban már informatizáltuk a Székely nyelvföldrajzi szótár teljes anyagát, ezért célszerűnek tartottuk, ha a helyneveket tartalmazó adattár feldolgozását egy székelyföldi kötettel kezdjük meg. Így lehetőségünk lesz a későbbiekben a nyelvjárási és helynévtári adatok összevetésére, integrálására.

A kötet informatizálása a betűhív forma megtartása mellett a sokoldalú felhasználásra (keresés, csoportosítás stb.) alkalmas kódolást és a megfelelő struktúrák kialakítását jelenti.

A névtani adattár történeti anyag lévén speciális karaktereket, ékezetes betűket tartalmaz, ezért feldolgozásához informatikailag korrekt, könnyen kezelhető kódrendszernek és hozzá kapcsolódó betűkészletnek a kialakítására volt szükség. A kötet teljes anyagát az általunk létrehozott kódrendszer alapján informatizáltuk, biztosítva, hogy minden adat megőrizze eredeti grafikus formáját.

Az informatizált és összesen 730 adatfájlra tagolt anyag struktúrája biztosítja az adatok helyhez (egy adott település névanyagához), időponthoz (az eredeti forrás kelte) és az eredeti nyomtatott változatban szereplő oldalszámhoz való kötődését. A kötött struktúra teszi lehetővé majd a későbbiekben az adatok megadott szempontok szerinti kereshetőségét, térképezését.

Szükségesnek tartjuk, hogy a helynévanyag részleges feldolgozását is elégezzük. Jelen kutatásunkban a névtestek kijelölésére, névfajták szerinti kódolására vállalkoztunk. Ennek megvalósításához olyan speciális szoftvert készítettünk – ez az Olló nevet viseli –, amely hatékonyan teszi lehetővé a történeti szövegekkel foglalkozó kutatók számára az anyagok feldolgozását, általuk megadott szempontok szerinti címkézését (manuális annotálását).

Munkánk nyomán (történeti és mai írásmód szerint egyaránt) kereshető és térképre vetíthető helynév- és névelemmutató jön létre, amely minden bizonnyal hatékonyan segíti majd a tulajdonnevekkel foglalkozó kutatók munkáját. Továbbá egy olyan új eljárást dolgozunk ki, amely új lehetőségeket nyit nyelvtörténeti anyagok számítógépen való rögzítéséhez és feldolgozásához.

4.3. A helynévgyűjtemény feldolgozása

Az informatizált változatnak lehetővé kell tennie, hogy a gyűjtemény adatai több szempontból kereshetők legyenek. A vármegye → település → évszám → forrás → helynévadat sorrend, amely a nyomtatott változat rendező elve, elsősorban azoknak könnyíti meg a munkáját, akik művelődéstörténeti, településtörténeti stb. kutatásokat folytatnak. A helynevek, illetve a nyelvtörténet kutatói számára az előzőnek éppen fordítottja hasznos: ha egy-egy nyelvi adat (helynév) is kereshető, de előhívhatók a hozzá kapcsolódó jellemzők is: az irat forrása, amelyben előfordult; a forrás datálása; a település, amelynek névrendszerébe tartozik; a vármegye, ahol a település található. A névrendszer egy nyelvközösség, egy táj, egy település vagy egy egyén által használt nevek összességét is jelentheti. A kutatás leggyakrabban egy település névrendszerével: helynevek körében a település külterületi és belterületi neveivel foglalkozik.

A feldolgozó munka három fő nehézsége: 1. a helynévgyűjteményben a névkialakulás különböző fokán lévő adatokkal találkozunk. A helynevek gyakran körülírással fokozatos leegyszerűsödésével alakulnak ki, a névvé válás momentuma nem mindig határozható meg. 2. Szabó T. Attila a pontos értelmezhetőség miatt olykor a szöveggörnyezettel együtt jegyezte ki a helynévadatokat az oklevelekből, így ki kell emelnünk ezekből a tulajdonnév-értékű részeket, pl. a *Horgos ér nevű helyben* adatból a *Horgos ér* egy kaszáló neve. 3. A helynevek több szóból is állhatnak, alkotóelemeik lehetnek más helynevek, pl. a *Szép mezőn Bacsoni útra jövőben* helynév a *Szép mező*, a *Bacsoni út* és a *Bacson* helynevet is magában foglalja.

Ha a tulajdonnevek elkülönítése megtörtént, az adatokhoz a névfajta meghatározását kapcsoljuk. A névfajta rendszerét Hoffmann István Helynevek nyelvi elemzése (Debrecen, 1993) című munkája alapján állítottuk össze.

Egy kijelölt névalakhoz az alábbi információkat rendeljük (valamennyi alapján kereshetjük, csoportosíthatjuk majd az adatokat). A program automatikusan elvégzi az eredeti cédulákon megadott információknak a névalakhoz rendelését. A névanyagban előforduló *Markos* adathoz például a következő információkat rendeli hozzá:

- Háromszék (2. kötet) anyagában található;
- 123. lapon;
- 1578-ból származik az adat;
- Forrás: EM. (Erdélyi Múzeum), Béli (család) levéltára;
- Márkus település névrendszeréhez tartozik (abban az anyagban található).

A további információkat manuálisan kapcsoljuk össze a névalakkal, például a fenti *Markos* alakhoz a „településnév” kategóriát.

A kész adattár integrálása az informatizált nyelvjárási adatbázisokkal olyan lehetőségeket teremt, amelyek nagyban megkönnyítik a nevek lokalizálását, térképre vetítését, névföldrajzi elemzését. Egy-egy névtípus, névalak vagy névalkotó földrajzi közszó (adott esetben tájszó) térképen ábrázolt elterjedése nem csak a névtani, nyelvészeti kutatás tárgya, hanem a migrációkutatás és a népesedéstörténet forrása is lehet.

A kialakított osztályozási rendszer alapján befejeződött a helynévanyag denotátumfajta szerinti manuális annotálása (címkézése) a Bihalbocs-alapokon kifejlesztett Olló nevű számítógépes alkalmazás használatával. A kijelöléseket és minősítéseket tartalmazó adatfájlokból automatikusan hoztunk létre adatbázist, valamint elektronikus mutatót, amely biztosítja a névanyagban való gyors, többszempontú keresést, miközben megmarad a kapcsolat az adat eredeti kontextusával, vagyis az eredeti cédulák szövegével, valamint a kötet nyomtatott változatának struktúrájával. (Adat és forrás ilyen szintű kapcsolatának megőrzése szokásos adatbáziskezelővel, célszoftver kifejlesztése nélkül megoldhatatlan lenne.)

A helynévanyag elemzésére speciális keresési rendszert és felhasználói felületet alakítottunk ki az Olló nevű programban, amely hatékonyan támogatja a tudományág szempontrendszer alapján a nyelvtörténeti anyag kutatását. A keresés során lehetőség van többek között az adatoknak települések szerint történő vagy a forrásdokumentum keltezését alapul vevő szűrésére, valamint a helynevek elemzése során kialakított osztályozási rendszer alapján az adatokhoz kapcsolt minősítések szerint megvalósuló keresésre.

Természetesen a program biztosítja azt is, hogy betűsorokra keressünk rá (ÉS- és VAGY-kapcsolatokat is meg lehet adni). Akár a speciális történeti karaktereket is egyszerűen, ugyanakkor betűhív módon lehet beírni a keresőmezőkbe (a kifejlesztett szimbólumrendszer és a hozzá tartozó történeti betűkészlet segítségével), illetve a rendkívül változatos speciális mellékjeleket akár figyelmen kívül is lehet hagyni a keresés során. (E funkciók használatára a szokásos adatbáziskezelőkkel egyáltalán nincs lehetőség, illetve a legutóbb említett funkcióra is csak akkor, ha nem betűhív kódolással lennének rögzítve az adatok).

A találatok csoportosításához további szempontokat is érvényesíthetünk. Emellett az adatok lokalizálása alapján a keresés és a csoportosítás eredményét térképes formában is azonnal meg lehet jeleníteni egyetlen kattintással. (A találati listák,

adatszoportok térképre vetítésének kutatói igénye szintén meghatározta a szoftverfejlesztés irányát.) Az adatelemzések alapján készített térképeket el lehet menteni kis tárigényű, vektorgrafikus képfórmátumban (pict), ezzel biztosítva a térképek használatát más programokban is. Az Ollóban létrehozott térképek egyszerűsített internetes (html + javascript + jpg/png) formába is átalakíthatóak oktatási és népszerűsítési célból.

Lehetőség van arra is, hogy a találati listákat strukturált txt-fájlokként mentjük el. Ezeket igény szerint a szokásos táblázatkezelő vagy adatbáziskezelő szoftverekkel (például MS Excellel) beolvashatjuk, statisztikailag elemezhetjük, formázhatjuk, kinyomtathatjuk stb.

A háromszéki helynevek feldolgozásáról és elemzési lehetőségeiről Báráth János számolt be 2006. augusztus 23-án Debrecenben, a VI. Nemzetközi Hungarológiai Kongresszuson *Háromszéki helynevek elemzése* című előadásában. Az előadáson bemutatott helynévtani térképekből elhelyeztünk három mintatérképet a Geolingvisztikai Műhely honlapján a Térképek című rovatban. A mintatérképek megtekintéséhez a http://geolingua.elte.hu/maps/sztahsz/sztahsz_hu.html linkre kell kattintani.

5. A kutatás eredményeinek közzététele

A projekt során elsődleges szempont volt, hogy a feladatok összetettségéhez képest minél egyszerűbb, s így szükség szerint bármely felhasználói rendszerbe is viszonylag egyszerűen konvertálható adatokat, adattárakat hozzunk létre. Elsődleges feladatunk tehát az adattárak kiépítése volt, a munka során kifejlesztett és alkalmazott nyelvészeti technológiák is főként ezt a célt szolgálják.

Az informatizált adatok létrehozásán túl megtettük az első lépéseket a nyelvészeti célú feldolgozás, elemzés irányába is, a kifejlesztett adatelemzési eljárások használata azonban egyelőre speciális felkészültséget igényel, feltételezi a fejlesztői környezet és az adattárak felépítésének alapos ismeretét. A létrejött adattárak feldolgozása jelen kutatásunk lezárásával teljesebbé válhat, hiszen mostanra álltak össze olyan mennyiségben és minőségben informatizált nyelvjárási adatok, hogy a hatékony adatrögzítés támogatása mellett az adatelemzés is előtérbe kerüljön. A tudományterület tényleges megújítása szempontjából elengedhetetlennek tartjuk, hogy az általunk létrehozott adattárakat ne csak azok használhassák, akik többéves tapasztalat birtokában maguk is képesek az informatizált nyelvjárási adatokat kezelő számítógépes nyelvészeti alkalmazások kifejlesztésére vagy módosítására, hanem mindenki, aki kutatni szeretné azokat. Ezt felhasználóbarát számítógépes alkalmazások kifejlesztésével és felhasználói kézikönyvek elkészítésével látjuk megoldhatónak. Ezeket az eszközöket az erőforrások függvényében hozzuk létre, és a kutatói igényekhez igazodva bocsátjuk az érdeklődők rendelkezésére az elkészült adattárakkal együtt. A kutatás teljes anyagát fokozatosan elérhetővé tesszük az ELTE Magyar Nyelvtudományi és Finnugor Intézetének könyvtárában is, amely így a magyar nyelvi geolingvisztikai adatbázisok referenciaközpontjának szerepét töltheti be a jövőben.

Szükségesnek tartjuk, hogy az általunk végzett kutatások, létrehozott nyelvjárási adattárak, nyelvészeti technológiák szélesebb körben is ismertté váljanak. Ennek érdekében kutatásunk első szakaszában internetes honlap kialakítását kezdtük meg, és a honlap tartalmi fejlesztését a kutatás teljes időtartama alatt folytattuk. A honlapot felkeresve az érdeklődőknek lehetőségük van belepillantani kutatócsoportunk

tevékenységébe, felvenni a kapcsolatot kutatócsoportunk tagjaival, az egyes részfeladatokat irányító kutatókkal. Az általunk létrehozott nyelvjárási anyagokat (hangoskönyveket, nyelvjárási térképeket) a Bihalbocs szoftverrel olyan formába konvertáljuk, amely alkalmas a világhálón való megjelenítésre.

Honlapunkon teljes terjedelmében elérhetővé tettük korábbi kutatási beszámolóinkat. A kutatás időtartamának végére a honlapon az általunk készített munkák közül a következő anyagokat helyeztük el: (a) a Magyar Nyelvjárási Hangoskönyv című sorozatunk I–III. kiadványának HTML-változatát, (b) néhány mintatérképet az általunk informatizált nyelvföldrajzi adattárak anyagából, (c) A magyar nyelvjárások atlasza ellenőrző gyűjtése során készített hangfelvételek listáját, (d) a Háromszék megyei történeti névtani anyag informatizált változatából készített mintatérképeket, (e) mutatványt a MNYA. ellenőrző gyűjtése során felvett szövegfelvételekből (teljes kutatóponthálózattal), (f) a kutatás időtartama alatt elhangzott előadások demonstrációs anyagai és a már publikált tanulmányok szövege, valamint (g) teljesebbé vált a honlap angol nyelvű változatának tartalma.

A kutatási anyagokból népszerűsítő és oktatási céllal a projekt lezárultát követően is folyamatosan helyezünk el mintákat a Geolingvisztikai Műhely honlapján (nyelvjárási hangoskönyvek, kattintható térképek, hanganyagok, előadások). Szándékunk szerint hosszabb távon is folytatjuk a szélesebb közönség számára is elérhető formátumú anyagok kialakítását és közzétételét a világhálón.

A honlap a következő címen érhető el: <http://geolingua.elte.hu>.

II. A költségterv és a pénzfelhasználás viszonya

A projekt tervezett és tényleges költségei				
	Terv (az eredeti szerződésben)		Tény	
	M Ft	%	M Ft	%
Alapkutatás	46,8	100	46,8	100
Alkalmazott kutatás	0	0	0	0
Kísérleti fejlesztés	0	0	0	0
Összesen	46,8	100	46,8	100

Projekt időtartama	Terv	Tény
(hónapokban kifejezve)	35 hónap	35 hónap

III. Az eredmények gazdasági és társadalmi hasznosíthatóságának bemutatása

A projekt eredményei elsősorban társadalmi hasznosíthatóságuk alapján értékelhetőek. Ebből a szempontból kétféleképp értékelhetjük kutatásunkat. Egyrészt a tudományterület nemzetközi színvonalú műhelyeiben is csak az utóbbi években kezdődött meg a hangzó nyelvi anyagok területi szempontból is értelmezhető

korpusznyelvészeti feldolgozása. Ezek a nemzetközi kísérletek az általunk alkalmazott nyelvészeti technológiákhoz képest kevésbé sokrétű és jóval korlátozottabb feldolgozási lehetőségeket hordozó geolingvisztikai megoldásokat alkalmaznak.

A nyelvészeti hasznosíthatóság mellett lényeges szempont a szélesebb, nem szakmai közönségben érvényesülő hatások kérdése is. A magyar nyelvközösségre jellemző erősen normatív tendenciák ellenében a magyar nyelvjárásokról szóló kutatások a nyelv területi változatairól való gondolkodás reflektáltabbá tételéhez és ezzel közvetetten a nyelvi sokszínűség fenntartásához járulhatnak hozzá a felhasznált modern technológiák tudománynépszerűsítő alkalmazásai révén. A Geolingvisztikai Műhely címmel kialakított honlap többek között ezeket a célokat is szolgálja.

IV. A projekt monitoring mutatói

Lásd a 2. mellékletet.

V. Összefoglaló értékelés

A projekt során főbb célkitűzéseink teljesültek: a magyar dialektológia hagyományos adathordozón kiadott két legjelentősebb, valamint két eddig kevésbé használható kisebb atlaszának anyagát informatizáltuk; létrehoztunk egy nemzetközi szempontból is újszerű nyelvészeti technológiákon alapuló multimédiás (beszélő) nyelvatlaszt a moldvai nyelvföldrajzi régió nyelvjárásaiból, amely egyben követéses vizsgálattal tárja fel az 1950-es évek nyelvi állapotának és napjaink nyelvhasználatának különbségeit és azonosságait; digitalizáltuk és részlegesen feldolgoztuk A magyar nyelvjárások atlasza 1960-as években gyűjtött, ma már nyelvtörténeti jelentőségű adatbázisát, amely 352 településről származó nyelvi szövegeket és jelenségeket vizsgál; valamint Szabó T. Attila erdélyi helynévtörténeti adattára háromszéki kötetének informatizálását is elvégeztük.

A projekt eredményei hasznosíthatók további dialektológiai vizsgálatokban, amelyek az alkalmazott technológiák átvételével nemzetközileg is új dimenziót jeleníthetnek meg a nyelv területiségének kutatásában. A konzorcium résztvevőinek együttműködése gördülékeny volt, a kutatás kezdetén kialakított világos munkamegosztás lehetővé tette a partnerek hatékony együttműködését a feladatok megoldásában. A kisebb feladatokba a projekt során nagy számban vontunk be egyetemi hallgatókat és doktorandusokat, részvételük költségkímélő hatásai tették lehetővé például, hogy A moldvai csángó nyelvjárás atlasza korábban lappangó III. kötetének informatizálását is elvégezhettük a pályázatban vállalt első két kötet mellett. A magyar nyelvjárások atlasza hanganyagainak tervezett további feldolgozásához a konzorcium tagjainak együttműködése szükséges.

MELLÉKLETEK

1. melléklet: A projekthez kapcsolódó publikációk és demonstrációk

1. Önálló kiadványok

- Magyar Nyelvjárási Hangoskönyv I. Szövegek Hegykőről, Velemből, Csöngéről és Rábagyarmatról.* Multimédiás CD-ROM. Szerk. Hajdú Mihály és mtsai. Budapest: ELTE Magyar Nyelvtörténeti, Szociolingvisztikai, Dialektológiai Tanszék. 2005.
- Magyar Nyelvjárási Hangoskönyv II. Szövegek Bödeházáról, Ságodról, Szentáráról, Nagyhindről és Lipótról.* Multimédiás CD-ROM. Szerk. Hajdú Mihály és mtsai. Budapest: ELTE, Magyar Nyelvtörténeti, Szociolingvisztikai, Dialektológiai Tanszék. 2006.
- Magyar Nyelvjárási Hangoskönyv III. Szövegek Banáról, Bakonyszentlászlóról, Somlóvásárhelyről, Patakról és Nagyövedről.* Multimédiás CD-ROM. Szerk. Hajdú Mihály és mtsai. Budapest: ELTE Magyar Nyelvtörténeti, Szociolingvisztikai, Dialektológiai Tanszék. 2007.
- Bodó Csanád – Vargha Fruzsina Sára: *Jelenségtérképek A moldvai csángó nyelvjárás atlaszából.* Budapest: ELTE Magyar Nyelvtörténeti, Szociolingvisztikai, Dialektológiai Tanszék. 2007.

2. Tanulmányok, cikkek

2.1. Már publikált dolgozatok

- Bárth M. János: Szabó T. Attila Erdélyi Helynévtörténeti Adattára. In: Uő. (szerk.), *Emlékkönyv Szabó T. Attila születésének 100. évfordulójára.* Bp. 2006. 63–66.
- Bárth M. János: Háromszéki helynevek nyelvészeti elemzése informatikai módszerekkel. In: Hoffmann István – Tóth Valéria (szerk.), *Helynévtörténeti tanulmányok 2.* Debrecen. 2006.
- Heltai János Imre: Új lehetőségek a moldvai magyar beszélőközösségek nyelvi viselkedésének vizsgálatában. In: Heltai Pál (szerk.), *Nyelvi modernizáció. Szaknyelv, fordítás, terminológia.* I. kötet. MANYE – Szent István Egyetem, Pécs – Gödöllő. 2007. 72–77.
- Kiss Jenő: Egy új magyar nemzeti nyelvátlasz szükséges voltáról. *Magyar Nyelv* 102 [2006]: 129–142.
- Vargha Fruzsina Sára: Állatok kicsinyeinek megnevezése a keleti magyar nyelvjárásokban. In: Hoffmann István – Juhász Dezső szerk. *Nyelvi identitás és a nyelv dimenziói.* Nemzetközi Magyarágtudományi Társaság, Debrecen–Budapest, 2007. 237–248.

2.2. Publikálásra elfogadott, megjelenés előtt álló tanulmányok

- Bárth M. János: Székelyföldi történeti helynevek névföldrajzi vizsgálata. In: *VI. Névtudományi Kongresszus.* Budapest.
- Bodó Csanád: Követéses geolingvisztikai vizsgálat Moldvában In: Molnár Zoltán (szerk.), *V. Dialektológiai Szimpozion.* BDF: Szombathely.
- Bodó Csanád: A moldvai magyar nyelvjárások román kölesönszórétegének területisége. In: Benő Attila, Fazakas Emese és Szilágyi N. Sándor (szerk.), *Nyelvek és nyelvváltozatok. Köszöntő könyv Péntek János tiszteletére.* Kolozsvár: Anyanyelvápolók Erdélyi Szövetsége. (A Szabó T. Attila Nyelvi Intézet Kiadványai 4.)
- Bodó Csanád – Vargha Fruzsina Sára: Régi nyelvátlaszok – új módszerek. *Magyar Nyelv* 104. [2008.]

- Heltai János Imre: Vélekedések a magyar nyelv területi változatairól Moldvában. In: Molnár Zoltán (szerk.), *V. Dialektológiai Szimpozion*. BDF: Szombathely.
- Vargha Fruzsina Sára Nyelvi változók A magyar nyelvjárások atlasza hangfelvételeiben. In: Molnár Zoltán (szerk.), *V. Dialektológiai Szimpozion*. BDF: Szombathely.
- Vargha Fruzsina Sára: Nyelvjárási és helynévtörténeti anyagok számítógépes feldolgozása. In: *Kontextus–Filológia–Kultúra II*. Besztercebánya.

3. Demonstrációk

- Bodó Csanád és Vargha Fruzsina Sára: *Régi nyelvatlaszok – új módszerek*. Előadás a Magyar Nyelvtudományi Társaság 2005. október 25-i felolvasóülésén, Budapesten. (Az előadás demonstrációs anyaga a <http://geolingua.elte.hu/about/eload/regiterk.htm> címen érhető el.)
- Heltai János Imre: *Új lehetőségek a moldvai magyar beszélőközösségek nyelvi viselkedésének vizsgálatában*. Előadás a Magyar Alkalmazott Nyelvészek és Nyelvtanárok Egyesülete Kongresszusán, Gödöllő, 2006. április 10–12.
- Bárth M. János: *Háromszéki helynevek elemzése*. Előadás a VI. Nemzetközi Hungarológiai Kongresszuson, Debrecen, 2006. augusztus 23.
- Vargha Fruzsina Sára: *Az állatok kicsinyeinek megnevezése a keleti magyar nyelvjárásokban*. Előadás a VI. Nemzetközi Hungarológiai Kongresszuson, Debrecen, 2006. augusztus 23. (Az előadás prezentációja a http://varghafr.web.elte.hu/VarghaF_kicsinyek3.ppt.htm címen tekinthető meg.)
- Vargha Fruzsina Sára: *A Magyar nyelvjárási hangoskönyv és elemzési lehetőségei*. Előadás a XIV. Élőnyelvi Konferencián, Bük, 2006. október 4.
- Bodó Csanád: *A magyar nyelv vitalitásának területisége Moldvában*. Előadás a XIV. Élőnyelvi Konferencián, Bük, 2006. október 5. (Az előadás demonstrációs anyaga megtekinthető a következő címen: http://geolingua.elte.hu/about/eload/bodocs_buk.ppt.htm.)
- Vargha Fruzsina Sára: *Nyelvjárási és helynévtörténeti anyagok számítógépes feldolgozása*. Előadás a Kontextus–Filológia–Kultúra II. nemzetközi konferencián, Besztercebányán, 2007. május 30-án.
- Bárth M. János: *Székelyföldi történeti helynevek névföldrajzi vizsgálata*. Előadás a VI. Névtudományi Kongresszuson. Balatonszárszó, 2007. június 23.
- Bodó Csanád: *Követéses geolingvisztikai vizsgálat Moldvában*. Előadás az V. Dialektológiai Szimpozionon, Szombathely, 2007. augusztus 21–23. (Az előadás demonstrációs anyaga a http://geolingua.elte.hu/about/eload/bodocs_VDialSzimp.ppt.htm címen érhető el.)
- Heltai János Imre: *Vélekedések a magyar nyelv területi változatairól Moldvában*. Előadás az V. Dialektológiai Szimpozionon, Szombathely, 2007. augusztus 21–23.
- Vargha Fruzsina Sára: *Nyelvi változók A magyar nyelvjárások atlasza hangfelvételeiben*. Előadás az V. Dialektológiai Szimpozionon, Szombathely, 2007. augusztus 21–23. (Az előadás demonstrációs anyagának megtekintéséhez a következő címre kell kattintani: http://varghafr.web.elte.hu/VarghaFr_Szombathely2007.htm.)
- Bárth M. János: *Vápa ~ lápa*. Előadás a 2. Helynévtörténeti Szemináriumon, Síkfőkút, 2007. október.