

Tóth Valéria

Mai Magyar Digitális Helynévtár*

Egy kutatási program célja, eredményei, távlatai

Abban, hogy a Hungarológiai Kongresszuson megtartandó előadásom témájául éppen a „Magyar Digitális Helynévtár” programjának a bemutatását választottam, nemcsak az játszott szerepet, hogy úgy gondolom, ez a jellegzetesen magyarságtudományi program csakis Kárpát-medencei keretekben és együttműködéssel valósulhat meg eredményesen, hanem döntésemet az is ösztönözte, hogy a kongresszus helyszíne éppen az a Kolozsvár, ahol a helynévgyűjtő munkának igen komoly, évszázados hagyományai vannak. Mondandómmal — igaz, némielyest a modern kor igényeihez is igazítva azt — szorosan kapcsolódni kívánok Szabó T. Attila több ízben is hangoztatott állásfoglalásához, miszerint a magyar helynévkinccs lehető teljességének összegyűjtése az összmagyarág mielőbbi és alapvető érdeke. Ez a máig érvényes gondolat volt az alapja a „Magyar Digitális Helynévtár” létrehozásának is.

1. A „Magyar Digitális Helynévtár” című kutatási programról

1.1. Előzmények

A magyar helynévkutatásban a 19. század óta¹ él az a terv, amely a Kárpát-medence teljes helynévanyagának összeállítását tűzte ki célul. Az 1960-as években ez a program intézményi kereteket öltött azáltal, hogy a Magyar Tudományos Akadémia Nyelvtudományi Intézetének szakmai irányítása alatt országos helynévgyűjtés indult meg, amelynek a legeredményesebb évei az 1970-es évek és az 1980-as évek első fele voltak. Ezt követően azonban visszaesés következett be, és a munka mára lényegében elakadt. Az országos névgyűjtés eredményeként (az ország kétharmad részéről) hatalmas névtárak jelentek meg, amelyek kettős tanulságot hordoznak magukban. 1. A bennük megtalálható névanyag a magyarság és más, a Kárpát-medencében élő népek nyelvének és kultúrájának

* A publikáció elkészítését az OTKA K 100580 számú pályázat, valamint TÁMOP 4.2.1./B-09/1/KONV-2010-0007 számú projekt támogatta. A projekt az Új Magyarország Fejlesztési Terven keresztül az Európai Unió támogatásával, az Európai Regionális Fejlesztési Alap és az Európai Szociális Alap társfinanszírozásával valósult meg.

¹ Gyakorlatilag azóta, hogy 1821-ben gróf Teleki József, az Akadémia későbbi elnöke ennek fontosságát a tudományos élet figyelmébe ajánlotta.

hű lenyomata és kifogyhatatlan gazdagságú tárháza. 2. A meglévő névgyűjtemények jórészt kihasználatlanok maradtak: átfogó elméleti feldolgozások nemigen épültek rájuk. Pedig e munkák nemcsak a szűkebb értelemben vett nyelvtudomány számára adnak forrásanyagot, hanem széleskörűen hasznosíthatják a mai névállomány tanulságait más tudományágak is. Ha pedig a helynévgyűjtemények anyagát — a modern kor technikai vívmányait hasznosítva — adatbázisban rögzítjük, és térinformatikai programokkal kapcsoljuk össze, a lehetőségek köre még inkább kitágul, és a belőle adódó tanulságoknak nagy hasznát vehetik a szakemberek pl. a helynevek standardizálásának kérdésében vagy tájrekonstrukció tervezésében is.

A Kárpát-medencei helynévkutatásnak e téren tehát kettős feladata van: 1. a hatalmas forrásanyag nyelvészeti szempontú feldolgozása, amely aztán esélyt ad arra is, hogy 2. azonos alapelveket, módszereket alkalmazva újrainduljon a megrekedt helynévgyűjtő munka, immáron Kárpát-medencei keretekben. Alapvető igény továbbá az is, hogy a meglévő, ám sokszor még a kutatók előtt is rejtve maradó névanyagot minél célszerűbb és maradandó formában hozzáférhetővé tegyünk a nyelvészeti kutatás (és más tudományterületek) számára.

1.2. A program társadalmi és tudományos jelentősége

A „Magyar Digitális Helynévtár”-nak — amellet, hogy tudományos igényeket is szolgálni kíván — a társadalmi, sőt mondhatjuk így is: a nemzeti-kulturális jelentősége sem elhanyagolható. A neveknek mint sajátos nyelvi elemeknek ugyanis a mindennapi életünkben nagy szerepük van. A gyakorlati hasznukon kívül kifejeződik bennük a nemzethez, kultúrához való kapcsolatunk, de emellett igen erősen jelzik a közvetlen környezetünkhöz való lokális tartozásunkat is. A nevek mint a múltból ránk maradt elemek gyakran régebbi nyelvi és kulturális viszonyokat őriznek, nem véletlen tehát, hogy irántuk e szerepgazdagság miatt sokféle tudomány (a nyelvészet mellett a történettudomány, a történeti földrajz, a néprajztudomány) érdeklődik, s hogy tanulmányozásukra önálló tudományszak, a jellegzetesen interdiszciplináris névtudomány (onomasztika) alakult ki. E tudomány pedig igen sokat tett a Kárpát-medence évezredes kultúrájának és mai viszonyainak megismerése terén.

A nevek ismeretének nyilvánvalóan nemcsak tudományos haszna van, hanem a mai élet — a globalizációval is összefüggésben — mind élenkebben igényt tart a nevek ismeretére és az élet különböző szinterein való használatára. A közigazgatás, a posta, a közlekedés jó évszázada igényli a helyneveket, s ehhez ma már a tömegkommunikáció, a digitális technológiák (pl. GPS stb.) is kapcsolódnak, sőt az ezek segítségével tájékozódó mentők, tűzoltók, a katasztrófavédelem sem nélkülözheti a nevekhez kapcsolódó helyismeretet.

A mai társadalomnak ezt a növekvő névigényét egyre nehezebb kielégíteni, hiszen egy-egy ember a környezetében használt neveknek mindig csak a töredé-

két ismeri. Ezt a feszültséget a tudomány segítségével lehet feloldani, mégpedig a nevek élőszóbeli és történeti forrásokban való előfordulásainak összegyűjtésével, az adatállományok összerendezésével és hozzáférhetővé tételével: könyvek és online adatbázisok közrebocsátása révén.

Az interetnikus kapcsolatok múltjának feltárásában ugyancsak nagy segítségünkre lehet a helynévkincs tudományos igényű elemzése, de emellett a napi igényeket is sikerrel szolgálhatja. Az anyaországon kívüli régiók magyar nyelvének (vagyis a külső régiók magyar nyelvének) a nyelvpolitikai helyzetében a neveknek kitüntetett szerepük van (gondoljunk csak az ún. táblaháborúra, a hivatalos helynévhasználat körül azóta is zajló vitákra). E nyelvpolitikai helyzet szakmai oldalról való megtámogatását a nyelvtudomány a mai lehetőségei között (pl. helynévtárak, helynévgyűjtemények híján) alig tudja ellátni. Az, hogy ez a helyzet megváltozzon, a Kárpát-medencei magyarság egészének az alapvető érdeke. Az egyes államok nyelvpolitikájában, s ehhez is kapcsolódóan a sajtó, a közigazgatás, a térképészet stb. terén a nevek használatának fontos szerepe van, e szempontból is lényeges tehát, hogy a tudomány a társadalom számára megbízható névanyagot tegyen hozzáférhetővé.

2. Az MDH létrehozása, jelenlegi helyzete, informatikai háttere

2.1. A helynevek iránt megmutatkozó és egyre fokozódó igényeket látva a nyelvtudománynak időszerű volt e téren lépéseket tennie. A Debreceni Egyetem Magyar Nyelvtudományi Tanszékének két kutatócsoportja is részt vesz (Tóth Valéria, illetve Hoffmann István vezetésével) a TÁMOP által finanszírozott kutatóegyetemi programban, aminek keretében a „Magyar Digitális Helynévtár” (MDH) kialakításán dolgozunk. E programmal egy olyan hatalmas adatbank létrehozását vállaltuk, amely a magyar nyelv egész helynévállományának befogadására és kezelésére alkalmas. Ez az adatbank publikus, online formában elérhető mindenki számára, s információt nyújt a nagyközönségnek éppúgy, mint a szakembereknek. Ennek az adatbázisnak a programot illetően az alapjai készen állnak, nyelvi anyaggal, helynevekkel való feltöltése folyamatban van: első lépésként — mivel nagyobb területről innen állnak a rendelkezésünkre a leggazdagabb névanyagot tartalmazó gyűjteményes kötetek — a dél-dunántúli területek (Baranya, Somogy, Tolna megyék) feldolgozása indult meg.

Az adatbázis-kezelő program kiválasztása és az adatokat fogadó struktúra kialakítása ugyancsak megtörtént: a 4D adatbázis-kezelő rendszerbe importáljuk be a forrásanyagokon alapuló adatrögzítés során keletkezett adatállományokat. a) A térképek georeferálását, digitalizálást az ArcGIS és az ArcView térinformatikai programok segítségével végezzük el, b) a névadatokat pedig a legtöbb adatbázis-kezelővel kompatibilis, tehát általánosabb felhasználási lehetőséget is biztosító (Word, majd pedig) Excel táblázatkezelőben rögzítjük. E széles körben elterjedt programok általános ismertsége ugyanis lehetővé teszi egyetemi hallga-

tók bevonását az adatrögzítés első fázisában anélkül, hogy ez tőlük mélyebb számítástechnikai tudást várna el.

2.2. Mivel a munka több fázisban zajlik, az egyes területek névkincsének rögzítése, adatbázisba emelése különböző készültségi fokon áll. A munkálatok első fázisában a digitálisan elérhető (vagy digitalizált) helynévgyűjtemények névanyagának rögzítését végezzük el, ami annyit jelent, hogy járásonkénti, illetve azon belül településenkénti rendben haladva a helynévkötetek minden információját rögzítjük Word, illetve Excel-táblákban. Az adatfeltöltés fázisa jórészt hallgatók közreműködésével zajlik, és aktuálisan a következő készültségi fokot mutatja. Elkészült Tolna megye teljes helynévanyagának (40.000 rekordnak, azaz ugyanennyi névadatnak) a rögzítése. Baranya megye első kötete (tehát a Sásdi, a Szigetvári és a Pécsi járás) lényegében szintén elkészült. Ez 190 települést, és 85.000 rekordot jelent. Somogy megye 253 településéből elkészült 237, vagyis a Fonyódi, Siófoki, Marcali, Kaposvári és Csurgói járás teljes egészében, a Nagyatádi és a Barcsi járás pedig részben. Ez jelenleg mintegy 60.000 rekordnyi nevet tesz ki.

2010 decemberétől más intézmények oktatói és hallgatói is bekapcsolódtak a munkába, s ez az együttműködés további nagy számú névanyag adatbázisban rögzítését biztosította. A Pázmány Péter Katolikus Egyetem Magyar Nyelvtudományi Tanszékéről Hegedűs Attila vezetésével 9 hallgató vesz részt az adatbázis-építő munkában, s ők mintegy 25.000 rekordnyi névanyagot (98 település névállományát) rögzítették a Győr-Moson-Sopron megyei Győri járás és a Kaposvári járás, a Pest megyei Nagykátai járás, valamint Csallóköz területéről. A Veszprémi Egyetem Magyar Nyelvtudományi Tanszékéről Hári Gyula vezetésével 8 hallgató vesz részt a munkában. A veszprémi hallgatók a Veszprém megyei Veszprémi járás helynévanyagának adatbázisban rögzítésén dolgoznak, s a járás 71 településéből ezidáig 36 készült el, ami kb. 16.000 rekordnyi névanyagot ölel fel.

Mindez pedig azt jelenti, hogy jelenleg összesen több mint 200.000 névadat esett át az adatfeltöltés, -rögzítés fázisán.

Ez az állomány azonban csak további munkálatok révén válik alkalmassá arra, hogy beemeljük az adatbázis-kezelő rendszerbe: a névadatok köznyelvi formájának a megállapítása és az egy objektumhoz tartozó névváltozatok közül a lemmanév (a fővariáns, az ún. táblané) kiválasztása ugyanis elengedhetetlen ahhoz, hogy bizonyos funkciók működhessenek. Minthogy azonban ezek a munkafázisok már jóval nagyobb szakértelmet követelnek meg, jobbra doktori hallgatók, tudományos segédmunkatársak végzik őket. Jelenleg a munkának ez a része egész Tolna megyére kiterjedően befejeződött, Baranya megyében a Sásdi járásban ugyancsak teljesnek mondható, ezeken a területeken kívül egyelőre a Szigetvári járás egy részében, valamint Somogy megye Kaposvári járásában végeztük el. Ez összesen mintegy 100.000 rekordot jelent. A névadatokat és a hoz-

zájuk tartozó filológiai információkat tekintve ez az a névadatmennyiség tehát, amely alkalmas arra, hogy az adatbázisba beemeljük. Erre ugyanakkor csak abban az esetben van mód, ha elkészült az adott járás térképes állománya is. A 4D adatbázis-kezelő rendszer ugyanis térinformatikai komponenssel is rendelkezik, azaz ennek köszönhetően a névadatok térképes megjelenítésére is lehetőségünk nyílik.

Mindezekből adódik, hogy jelenleg a helynévtár anyagának csupán egy része, közel 80 ezer modern kori helynévadat érhető el az interneten is a digitális helynévtár honlapján (<http://mdh.unideb.hu/>).

A 4D működtetésében, fejlesztésében német és magyar informatikusok, térinformatikai szakemberek vannak a segítségünkre, és egyes területek feldolgozása (azoké, ahol német nyelvű névanyag is található) szintén német–magyar együttműködés keretében zajlik.

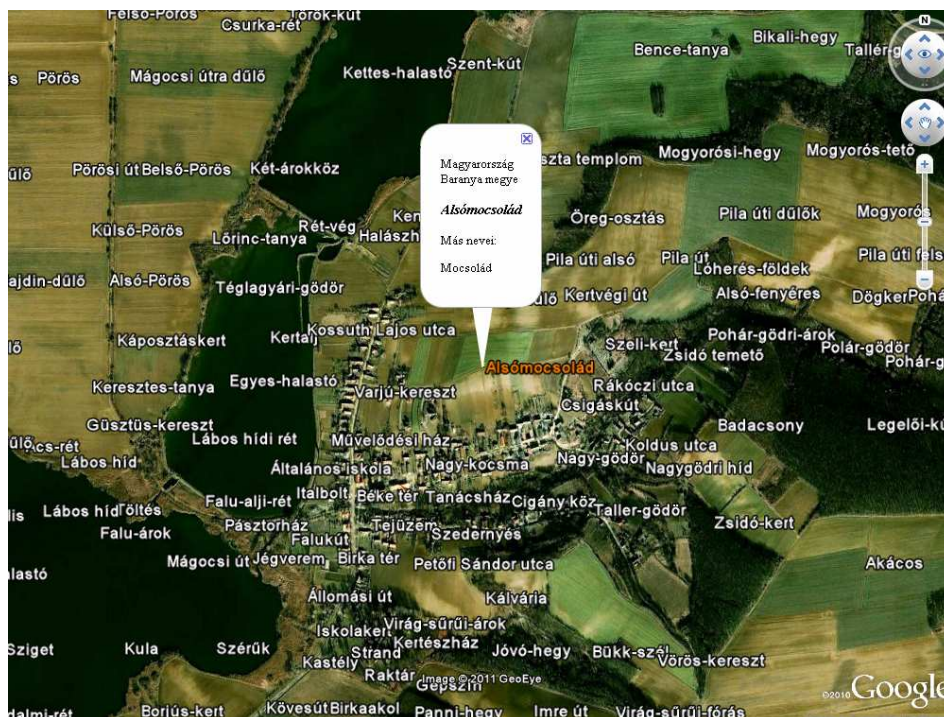
3. Az MDH nyújtotta lehetőségek

Az MDH jelenlegi állapotában a kiválasztási és keresési lehetőségek csupán az általános felhasználói felületen működnek, a tervek szerint összetettebb kereséseket és kiválasztásokat is lehetővé tevő kutatói verzió ugyanis még kidolgozás alatt áll.

3.1. Kiválasztás

Az általános felhasználók számára biztosított kiválasztás funkció arra szolgál, hogy egy adott, az adatbázisban már szereplő település valamennyi helynevét együtt láthassuk. Ehhez két úton juthatunk el: a térképes kiválasztással, illetőleg úgy, hogy a településnevet a kiválasztás mezőben egy településnév-listából kiválasztjuk. Az első esetben a nyitólap térkép ikonjának megfelelő részére kattintva a Google Föld segítségével megkapjuk az adott terület településneveit, s ezek közül a kívánt névre kattintva megjeleníthetjük a település valamennyi helynevét. (Minden helyhez tartozóan azonban csak egy névváltozat, az ún. táblanév szerepel a térképen ilyen esetekben.) A térképen megjelölt helynevek bármelyikéről néhány alapvető információt ugyancsak megkaphatunk akkor, ha az adott névre kattintunk. A kiválasztás mező használatával egy listához jutunk a kiválasztott település valamennyi helynévi adatával. (Azt is láthatjuk ezt a funkciót alkalmazva, hogy az adott településen hány helynévi forma van használatban.) A listában ugyancsak bármelyik névalakra rákattinthatunk, s mellette ekkor megjelenik a név bővebb adatlapja, számos, a helynévre és a helyre vonatkozó lényeges információval. Az adatlapon található földgömb ikonnal az adatok térképi megjelenítését is kérhetjük.

Az alábbi térkép a Magyar Digitális Helynévtár adatbázisából Alsómocsolád település valamennyi denotátumát (a hozzá tartozó táblanévvel) ábrázolja a Google Föld segítségével.



3.2. Keresés

A keresés funkció a kiválasztástól eltérően egy-egy konkrét helynévre vagy helynévdarabra (a nevekben szereplő lexémákra) történő keresést szolgálja. A keresés eredménye két listában jelenik meg: az első lista azokat a neveket tartalmazza, amelyek a keresett névvel szerkezetileg teljesen azonosak. A második listában pedig azokat a neveket találjuk meg, amelyek részeként a keresett név előfordul. E két lista segítségével az összes olyan névadatot megkaphatjuk, amelyekben a kívánt névforma vagy önmagában vagy valamely név részeként felbukkan.

Ugyanilyen eredményre jutunk akkor is, ha egy lexémára vagy szerkezetre keresünk rá (pl. egy-egy fanévre, földrajzi köznévre, névutóra stb.).

Mindkét esetben a listában látható névalakok bármelyikére rákattintva annak az adatlapja megnyílik a mellette lévő ablakban, a korábban már említett információkkal, illetve a térképi megjelenítés lehetőségével (ezt illusztrálja a következő ábra).

The screenshot shows the website interface for 'MAGYAR DIGITÁLIS HELYNÉVTÁR'. At the top right, there are language options: magyar, English, and Deutsch. Below the header, there are navigation buttons: Impresszum, Általános felhasználók, Kutatók, Krónika, and Társhonlapok. The main content area is titled 'Kenderföld' and includes a search bar with the text 'A keresett név Kenderföld'. Below the search bar, there are three tables of search results:

A keresett név 88 objektum nevéként fordul elő.

megye	település	standard
Tolna megye		Kenderföld
Baranya megye	Abaliget	Kenderföld
Baranya megye	Alsómocsolád	Kenderföld
Somogy megye	Attala	Kenderföld

A keresett név 0 objektum névadataként fordul elő.

megye	település	standard

A keresett név 97 objektum nevének részeként fordul elő.

megye	település	standard
Somogy megye	Gölle	Alsó-kenderföld
Tolna megye	Sárszentlőrinc	Alsó-kenderföld
Baranya		Kenderföld a

The right side of the page shows details for 'Kenderföld', including its location in Somogy megye, Attala településén, and a list of other names: Kenderfölek, Tabodi első, and Tabodi második. There are also sections for 'Információk a helyről' and 'Névmagyarázat'.

4. Távlati célok, tervek

a) A helynévi adatokat és a térképi ábrázolásukat látva világosan megmutatkoznak egyúttal a helynévtár hiányosságai is. A feldolgozott helynévgyűjtemények — tekintve, hogy a gyűjtések akkor zajlottak — jórészt az 1970-es évekbeli állapotokat mutatják, és bizonyos, hogy azóta a helyekben magukban és az elnevezéseikben is nagy változások mentek végbe (volt közben például egy rendszerváltás is, amely a földterületek birtokviszonyait alapjaiban alakította át, és különböző okokból eredően számos név megváltozását is magával hozta). Ahhoz, hogy az adott területek névanyagát tekintve az adatbázis naprakész legyen, újabb helyszíni gyűjtésekre: kiegészítő és ellenőrző, aktualizáló gyűjtésekre van szükség. Ráadásul a névgyűjtemények térképei az adott kor technikai lehetőségeinek korlátjai miatt igen pontatlanok, ami a Google Földön való ábrázolás alkalmával azonnal ki is ütközik. A lokális jellemzők pontosítására ugyancsak a helyszínen van lehetőség.

Ezeket a (nyilvánvalóan hosszabb távú) feladatokat látva 2011 augusztusában kísérleti jelleggel és hat hallgató részvételével a Dél-Dunántúlon helynév-aktualizáló gyűjtést végeztünk, melynek keretében 16 település névanyagát „frissítettük” fel (azaz egészítettük ki 2011-re vonatkozóan érvényes információkkal, névadatokkal), és tettük rendbe a lokalizálásukat. E kísérlettel a célunk egyrészt az volt, hogy kidolgozzuk e munkafolyamat módszertanát, és másrészt felmérjük

azt is, hogy az ilyen feladatok megvalósításának milyen személyi, időbeli és finansziális vonatkozásai vannak.

b) Azt is fontosnak gondoljuk egyúttal, hogy azokon a területeken is előmozdítsuk a helynévgyűjtő munkát, ahol ezidáig nem (vagy csak korlátozottan) jelentek meg helynévgyűjtemények. E téren elsőként Hajdú-Bihar megyében tettünk lépéseket, egy szűkebb régió, a Hajdúhát településeit kiválasztva. Itt ugyancsak egyetemi hallgatók és doktoranduszok közreműködésével szerveztük meg a munkát. A helyszínen végzett gyűjtés fázisát követően az így nyert adatállomány természetesen beépül az MDH adatbázisába, illetőleg helynévgyűjtemény formájában könyvként is megjelentetjük: az első kötet (a Hajdúhát településeivel) 2012 ősztől kerülhet a nyomdába. A megye további régióit ugyanilyen módon kívánjuk feldolgozni.

c) A helynevekkel való tudományos foglalkozást a legkorábbi időszaktól kezdődően jellemezte a történeti névadatok iránti fokozott érdeklődés, s ezen belül is különösen nagy figyelem fordult minden korban a korai okleveles helynévi adatok felé. Ebben a tágabb kontextusban (és a debreceni névkutató műhely alapvetően történeti irányultságát valamelyest itt is érvényesítve) döntöttünk úgy, hogy a „Magyar Digitális Helynévtár”-ban külön modult tervezünk a korai ómagyar kor (vagyis az 1350 előtti időszak) helyneveinek. Az MDH-nak ilyen módon két fő modulja különíthető el: az egyik az újkori-mai helynévállomány (főképpen 18–21. századi névadatokkal), a másik pedig a korai (leginkább 1350 előtti) névkinccs kezelésére használatos, de a két modul természetesen szervesen össze is kapcsolódik egymással. Minthogy a történeti nevek iránt megmutatózó igények kielégítését napjainkban számos kézikönyv szolgálja (a forráskiadványoktól kezdve a kompendiumokon át a szótárakig), az MDH korai helynévi moduljához az elsődleges forrásaink lényegében önként kínálkoztak. A helynévadatok térképes megjelenítését ebben a modulban rekonstruált középkori térképeken valósítjuk meg. A korai helynévtár kialakítása — kísérleti jelleggel az Árpád-kori Abaúj megye anyagán — megkezdődött, és reményeink szerint rövidesen ez is működésbe léphet.

d) Amint azt a korábbiakban már említettem, a digitális helynévtárnak egyelőre az általános felhasználói felületét készítettük el, és tettük elérhetővé az interneten is. Az is nyilvánvaló ugyanakkor, hogy az MDH mellett, hogy a nagyközönség igényeit is igyekszik kielégíteni, elsődlegesen mégiscsak tudományos célokat szolgáló adatbázis kíván lenni (noha a jelenlegi változat is szolgál már a kutatást közvetlenül segítő lehetőségekkel). Ezért szerepel a rövidtávú tervek között a kutatói verzió kidolgozása, amelyben összetett kiválasztási és keresési funkciókat biztosítunk a kutatóknak, lehetőséget adunk a nyelvi elemzések szempontjai szerinti leválogatásokra és térképi megjelenítésekre. Ehhez természetesen arra van szükség, hogy a filológiai információk mezője mellé a nyelvi elemzések mezőjével (a névszerkezetre, a helyfajta vonatkozó elemzésekkel, a

jelentéstani kapcsolatokkal, az etimológiával stb.) is kiegészítsük és feltöltsük az adatbázist.

5. Végezetül ide kívánczok még néhány általános, záró megjegyzés a program első évének tanulságai alapján. Természetesen tisztában vagyunk azzal, hogy mint minden nagyszabású vállalkozásnak, a „Magyar Digitális Helynévtár”-nak is megvannak a maga buktatói: a túlzott elvárások megfogalmazásával például az egész programot igen könnyen ellehetetleníthetjük. Ezért van nagy jelentősége annak, hogy a kezdetektől jól átgondolt keretek között dolgozzunk, ahol az egyes munkafázisok világosak, és szervesen egymásra épülnek, ugyanakkor ezek a keretek akár módosulhatnak is a munka tapasztalatainak fényében. Az MDH személyi és infrastrukturális feltételei alapvetően adva vannak, azonban a megvalósításához folyamatos pénzügyi támogatásra van szükség. Ezt továbbra is elsősorban különböző pályázati forrásokból igyekszünk biztosítani. A sikeres munkának emellett azonban a minél szélesebb körű, a magyar nyelvterület egészére kiterjedő, Kárpát-medencei tudományos együttműködés a leglényegesebb feltétele.

A „Magyar Digitális Helynévtár”-ra vonatkozó további információk olvashatók a helynévtár honlapján (lásd előbb), illetve az alábbi publikációkban:

TÓTH VALÉRIA (2011), A térinformatika alkalmazási lehetőségei a nyelvészetben.

In: *Az elmélet és a gyakorlat találkozása a térinformatikában II.* Szerk. LÓKI JÓZSEF. Debrecen. 143–50.

TÓTH VALÉRIA–ZICHAR MARIANNA (2011), Informatika a nyelvtudomány szolgálatában. In: *Informatika a felsőoktatásban 2011 konferencia. Konferenciakiadvány.* Szerk. CSER LÁSZLÓ–HERDON MIKLÓS. Debrecen. 843–8.

