

**Magyarok Erdélyben:
a minta kialakítása és az adatfelvétel**

Kiss Tamás – Kapitány Balázs

Kiss Tamás - Kapitány Balázs (2009): Magyarok Erdélyben: a minta kialakítása és az adatfelvétel. In Spéder Zsolt (szerk.): *Párhuzamok. Anyaországi és erdélyi magyarok a századfordulón*. KSH NKI Kutatási jelentések 86. Budapest, KSH Népeségtudományi Kutatóintézet, 31-54.

Tartalom

Az erdélyi magyarokat célzó survey-vizsgálatok háttéréről	35
Mérések és mintavételi technikák.....	36
A kvótás eljárás	37
A léptékes-kvótás kombinált módszer.....	38
Nemzetközi kitekintés	40
Az Életünk fordulópontjai – Erdély c. vizsgálatban alkalmazott mintavételi eljárás	40
A célpopuláció.....	40
Az adatfelvétel logikája.....	42
Elméleti és gyakorlati kontaktminta	43
Az gyakorlati kontaktmintától a tényleges kapcsolatfelvételig.....	46
A nyelvi szűrő	46
Sikeres kérdőívek	47
Az adatbázis súlyozása.....	47
Újdonságok és lehetőségek.....	49
Irodalom	53

Az erdélyi magyar népességre irányuló *survey*-vizsgálatok az 1990-es évek második felében kezdődtek.¹ Mielőtt az *Életünk fordulópontjai – Erdély c.* kutatás mintavételi eljárását bemutatnánk, röviden vázoljuk, hogy az eddigi munkák milyen mintavételi eljárást alkalmaztak. A saját megoldásunk bemutatását követően azzal is foglalkozunk, hogy az eltérő technikák milyen eredményre vezetnek a minta egyes jellemzőit illetően.

Az erdélyi magyarokat célzó survey-vizsgálatok háttéréről

Az erdélyi magyarokra reprezentatív mérés intézményesedéséhez leginkább két tényező járult hozzá. Egyrészt az 1994 és 1998 között Magyarországon hatalmon lévő Horn-, majd az 1998 és 2002 közötti Orbán-kormány saját (határon túli) magyarságpolitikájában, a döntéshozatalhoz (vagy ha máshoz nem, a döntések visszaigazolásához) igényelte a kisebbségi magyar népességekre fókuszáló felméréseket.² Ez motiválta a Magyarországgal szomszédos négy ország, Románia, Szerbia, Szlovákia és Ukrajna magyar népességét célzó összehasonlító vizsgálatokat, amelyek középpontjában egyrészt magyar kormány különböző kezdeményezéseinek megítélése, másrészt a magyarok közéleti, politikai és nemzeti értékrendje állt. A kutatásokat végző műhelyek közül a Balázs Ferenc Intézetet (BFI) kell kiemelni, amely 1997 és 2001 között hat, kisebbségben élő négy magyar közösségre kiterjedő felmérést végzett.

A budapesti adminisztráció mellett az RMDSZ részéről fogalmazódott meg erre vonatkozó igény. A szervezet politikai professzionalizálódásával párhuzamosan növekvő mértékben kezdte a politikai marketing szabványos eszköztárát alkalmazni saját kisebbségi célközönségére. Az 1996 és 2000 közötti ciklusban rendelte meg az erdélyi magyarokra reprezentatív első vizsgálatokat, amelyeket 2000-es választási kampányra való felkészülésben használt fel, a 2004-es és 2008-as választási kampányokban pedig már – részben mert versenyezni volt kénytelen a magyar szavazatokért – rutinszerűen alkalmazta. Ezekben az időszakokban a helyi (nagyobb városokra, megyékre) reprezentatív mérésekkel együtt tucatnyi vizsgálat készült. Ezek időközben eltolódtak a politikai marketing irányába, és kiszorultak belőlük a tudományos alapkutatáshoz közel álló tematikák (pl. migrációs potenciál, nyelvhasználat, etnikumközi viszonyok). Az utóbbi évek felmérései inkább piaci elvek szerint szerveződtek. A vizsgálatokat végző intézmények közül a Babes-Bolyai Tudományegyetem Szociológia tanszékének háttérintézeteit, az Etnikumközi Viszonyok Kutatóintézetét (CCRIT) és a Max Weber Társadalomkutató Központot, illetve kolozsvári szociológusok magáncégeit (Promoter Consult Center, TransObjective, Kvantum Research) emelhetjük ki.

¹ Tanulmányunkban kizárólag az erdélyi vizsgálat mintavételével és terepmunkájával foglalkozunk. A magyarországi *Életünk fordulópontjai c.* adatfelvétel iránt érdeklődőknek ajánljuk a kutatás korábban megjelent önálló módszertani kötetét: Kapitány 2003 (letölthető a www.demografia.hu honlapról).

² Ez 2002 után nagymértékben megváltozott. A Medgyessy-, majd a Gyurcsány-kormány, amennyiben élt valamilyen „magyarságpolitikai” kezdeményezéssel, nem a határon túli magyar közvéleményt szondázta, hanem a határon túli elitekkel konzultált.

Az erdélyi magyarokra reprezentatív mérések e két fő vonulata mellett – kisebb súllyal és jórészt a médiapiacra szorítkozva – harmadikként a kisebbségi piackutatás jelent meg. A kilencvenes évek végétől évente, két évente ismétlődő vizsgálatssorozatot indított a Duna Televízió, de kisebb rendszerességgel az erdélyi magyarokat célzó más sajtóorgánumok is rendeltek felméréseket.

A negyedik csoportot az akadémiai célú kutatások képezik. Tény, hogy (az utóbbi évekig) a politikai megrendelésre készített mérések jelentős része is az alapkutatás és az alkalmazott kutatás (közpolitika, politikai marketing) határmezsgyéjén mozgott. Az ezeket végző kutatók általában egyszerre voltak közvélemény- vagy piackutatók és kötődtek az akadémiai szférához. Ezek mellett azonban – igaz, kisebb és az erdélyi magyarokra reprezentatív mérések intézményesüléséhez önmagukban elégtelen számban – voltak akadémiai indíttatású vizsgálatok is. Ide sorolhatjuk az ELTE Szociológiai Intézet Minority Studies Tanszéke által végzett Kárpát Projektet, a Nemzeti Ifjúságkutató Intézet által koordinált Mozaik 2001-et, a CCRIT által 2000-ben elvégzett Etnobarométert, illetve a 2007-ben a MTA Kisebbségkutató Intézet és a Max Weber Társadalomkutató Központ Kárpát Panel vizsgálatát. Ezeknek a tematikája is nagyon közel állt a közvélemény-kutatásokéhoz. Elsősorban az etnikumközi viszonyokat, a közéleti értékrendet, nemzeti identitást állították középpontba. Az Életünk fordulópontjai – Erdély alapvetően szintén akadémiai indíttatású.

A vizsgálatok politikai és akadémiai dimenzióinak megkülönböztetésével nem akarunk a „politikamentes tudomány” mellett érvelni. A demográfia hagyományaival összhangban fontosnak tartjuk a társadalomkutatás közpolitikai hasznosíthatóságát. Mégis, egy kutatás megtervezését – a tematikától a mintavételig – alapvetően befolyásolja, hogy az milyen szempontoknak van alárendelve. Ezzel akkor is számolnunk kell, amikor a különböző mintavételi technikákat összevetjük. A mintavétel megtervezésekor az volt a szándékunk, hogy vizsgáljunk az elkövetkezőkben egyfajta módszertani normává váljon. A következőből azonban ki fog derülni, hogy a költség- és időkorlátok miatt az általunk javasolt megoldás marketing- vagy politikai közvélemény-kutatásokra való kiterjesztése miért problematikus.

Mérések és mintavételi technikák

Az eddigiekben említett vizsgálatok közös vonása, hogy többlépcsős kvótás, illetve léptékeskvótás eljárással dolgoztak. Az első lépcsőben a *településminta* kialakítására került sor; a kutatási pontok mellett ebben a fázisban határozták meg az egyes településeken kötelezően lekérdezendő kérdőívek számát. A nem önreprezentáló települések kiválasztása általában rétegenként történt. A vizsgálatok két vagy három rétegeképző változót használtak. Ezek különböző

kombinációkban a több megyéből álló régió,³ a településtípus (faluk, város), a magyarok településen belüli aránya,⁴ illetve a települméret voltak. A települméseknek jobb esetben az adott helység magyar lakosságának települmésrétegen belüli arányával megegyező, rosszabb esetben pedig (ettől a tényezőtől függetlenül) egyenlő esélyük volt a mintába kerülésre.

A települmésrétegeken (illetve az önreprezentáló települméseken) belül lekérdezendő kérdőívszámot az előző (vagyis az 1992, később pedig a 2002-es) népszámlálás települméssoros nemzetiségi megoszlásai alapján határozták meg.⁵ A vizsgálatok így a cenzus nemzetiségi adatainak megfelelően reprezentálták a magyar népesség térbeli eloszlását. A gyakorlati lehatárolásban – legalábbis konvencionálisan – a „*van-e ebben a háztartásban olyan személy, aki az elmúlt népszámlálás alkalmával magát magyarnak vallotta*” kérdés játszott szerepet.

A kvótás eljárás

Kezdetben (leginkább a BFI vizsgálataiban) a települméseken belül a megkérdezetek kiválasztása egyszerű kvótás eljárással történt. A kérdezőknek csupán az előre megadott, nem és életkor (esetleg iskolai végzettség) szerinti arányokat kellett figyelembe venniük.⁶ A kvóták meghatározásában szintén a népszámlálási adatok jelentették a kiindulópontot. Ez az eljárás természetesen csak azt garantálta, hogy a megadott két-három-négy szempont szerint pontos legyen az eloszlás. Mivel azonban a kiválasztás nem a véletlenül alapul, a módszer nem tekinthető (statisztikai értelemben) reprezentatívnak. Az egyes kvótacellákon (pl. 20–29 éves nagyvárosi érettségizett nő) belül ugyanis már semmi biztosíték nincs a kiválasztás véletlen voltára. Előfordulhat például, hogy egy román kérdezői csapat a nemzetiség mint kvótaszempont hiányában a szabályok legszigorúbb betartása mellett sem kérdezi meg egyetlen magyart sem. Ha mondjuk egy székelyföldi településen vonakodnak magyar háztartásokat felkeresni, az ott élő néhány román családból feltölthetik a kvótát. Így a módszertani előírásokat tökéletesen betartották, de például a nyelvhasználati kérdésekre adott válaszok hamisak lesznek. Az olyan magyar kérdezők tevékenységéről, akik a szórványban a református lelkésznel kezdik a

³ Az egyes vizsgálatok különbözőképpen határozták meg a régiókat. A Balázs Ferenc Intézet négy régióval dolgozott: Székelyfölddel (Hargita, Kovászna), Észak-Erdéllyel (Maros, Kolozs, Beszterce Naszód), Dél-Erdély és Bánság régióval (Brassó, Fehér, Szeben, Hunyad, Temes, Arad, Krassó-Szörény), valamint Partium és Máramaros régióval (Bihar, Szatmár, Szilágy, Máramaros megyék). A CCRIT régiói: Székelyföld (Hargita, Kovászna, Maros), Középső régió (Brassó, Fehér, Szeben, Kolozs, Hunyad), Bánság (Krassó-Szörény, Temes), Északi régió (Beszterce, Szilágy, Máramaros), Partium/Észak-nyugati régió (Arad, Bihar, Szatmár). A Max Weber, a TransObjective és a Kvantum Research kutatásai az Életünk fordulópontjai Erdély vizsgálattal azonos régiókat használt. Ezeket a későbbiekben bemutatjuk.

⁴ A magyarok aránya általában dichotóm (többség-kisebbség) változóként jelent meg, de esetenként több kategóriára volt bontva.

⁵ A 2002-es romániai népszámlálás magyarokra vonatkozó etnikai adatai – noha nyilván nem tökéletesek – alapvetően elfogadhatók, ez teszi lehetővé keretként való felhasználásukat. (Szemben például a 2001-es magyarországi népszámlálás adataival, amelyek a két legnépesebb magyarországi kisebbség – romák, németek – esetén nyilvánvalóan nem reálisak.)

⁶ A kvótás módszer legtöbbször alkalmazó Balázs Ferenc Intézet nem, életkor és iskolai végzettség szerinti kvótákat használt.

felmérést, lényegében ugyanez mondható el. Ők az egyházhoz, illetve más magyar intézményekhez kötődő személyek közül választanak,⁷ míg a magyar intézményekhez és hálózatokhoz nem kapcsolódó, de önmagukat magyarnak valló személyeknek nem lesz esélyük bekerülni a mintába. A következmény az, hogy a mintában a magyarok aránya nagyobb lesz mint, az alappopulációban, vagyis például a ténylegesnél kisebb arányt fognak képviselni a vegyes házasságban élő, a román nyelvű oktatási intézményekben végzett, a nem az RMDSZ-re szervező stb. magyarok.

Amellett, hogy az „interetnikus széleken” élők nem fognak a mintába kerülni, alapprobléma a társadalmi státus és a (településen belüli) lakhely szerinti szelekció. A kvótás eljárással dolgozó kérdezők eltekintenek a marginális helyzetűnek ítélt háztartásoktól, a rossz hírű vagy egyszerűen a központtól távol eső városrészekről. *„A kvótás mintavételi eljárás hiányossága úgy is megfogalmazható, hogy a kérdezőbiztosok által [akár tudatlanul – K.B.–K.T.] preferált társadalmi csoportok nagyobb mértékben kerülnek a mintába, mint az egyéb csoportok, még akkor is, ha a csoportképző szempontok egy része – az előírt kvóta miatt – nem is torzul.”* (Rudas 1998)

A fent jelzett problémákra tekintettel a kvótás mintavétel az elmúlt évtizedekben a nemzetközi kutatási gyakorlatban visszaszorult. A szakirodalomban szinte konszenzus van abban, hogy – néhány nagyon speciális esetet leszámítva – az így nyert adatok pontatlanabbak lesznek, illetve sokkal bizonytalanabbak a statisztikai hibák számításának módjai.

A léptékes-kvótás kombinált módszer

A későbbi vizsgálatok a település és a kiválasztott személy közötti szintre, a háztartásra összpontosítottak, ennek kiválasztási módját igyekeztek finomítani, mégpedig a rögzített kezdőpontról kiinduló léptékes módszer (véletlen séta) alkalmazásával. A háztartásokon belüli kiválasztás azonban továbbra is a kvóták segítségével történt.

Több vizsgálat (pl. a BFI 2000. és 2001. évi vizsgálatait vagy a MOZAIK 2001) a léptékes és az egyszerű kvótás eljárásokat kombinálta oly módon, hogy a többségében magyarok lakta településeken lépték (és a háztartáson belül kvóta), a szórványterületeken pedig kvóta szerint kérdeztek. Az ezredforduló óta azonban a kutatások többsége a léptékes-kvótás kombinált eljárás valamilyen válfaját használja.⁸

Az erdélyi mérések eljárási rutinja a léptékes módszeren belül döntően abban különbözött, hogy a kezdőpont véletlenszerűsége és a léptéktartás milyen mértékben ellenőrizhető.

⁷ Az erdélyi szociológusok ezért nevezték egymás között ezt az eljárást – némi öniróniával – presbitériumokra reprezentatív mintának.

⁸ A kvótás módszerről léptékesre való áttérés után kerültek bevezetésre a címkövetési listák, amelyek nem csupán a válaszolókra, hanem a meghiúsult kapcsolatfelvételekre vonatkozó adatokat is tartalmaztak, mint a sikertelenség oka (válaszmeztagadás, nincs magyar nemzetiségű a háztartásban, nem sikerült felvenni senkivel a kapcsolatot, nem volt a kvótáknak megfelelő személy). A címkövetési listák szisztematikus feldolgozása azonban még egyetlen kutatás keretén belül sem történt meg.

Egészen az utóbbi évekig a kezdőpont kiválasztását nagymértékben a kérdezőbiztosra bízta,⁹ az újabb kutatásokban azonban (Max Weber, TransObjective, Kvantum Research) a nem folyamatosan házszerű településeken a település és a háztartás közötti harmadik lépcsőként utcaminta készült, amelynek során az egyes utcáknak az ott található személyek/magánháztartások számával arányosan volt esélyük a mintába kerülni.¹⁰

A léptékes (vagy más szóval „véletlen sétás”) módszer elméletileg megfelel a véletlen kiválasztás szabályainak. Ennyiben ezen vizsgálatok léptékes almintája reprezentatívnek tekinthető. A módszert több országban (elsősorban az Egyesült Államokban, Dél-Amerikában) mértékadó, minden tudományos testület által elfogadott kutatásokban is alkalmazzák, bár főleg akkor, ha jogi okból nem lehet címlistát szerezni.

Európában a léptékes és a címlistas mintavétel évtizedek óta él együtt, jellemzően olyan munkamegosztásban, hogy marketing- és piackutatási célokra szánt „gyors és olcsó” vizsgálatok esetén léptékes, míg állami, szigorúan akadémiai vizsgálatok esetén a mintegy negyedével drágább regiszteralapú név- vagy címlistas módszer a preferált. Azonban jellemző, hogy a léptékes mintavétel esetén kialakul egy, a főbb szereplők által egységesen alkalmazott, írásban is lefektetett, rendszeresen aktualizált módszertani szöveg. (Németországban ilyen az úgynevezett ADM-mintavétel.) Összehasonlító módszertani kutatások azonban arra utalnak, hogy regiszteralapú mintavételek esetén az adatok minősége még az ilyen erősen szabályozott léptékes mintavételekénél is jobb. (A témához lásd Koch 1997.)

Magyarországon a terepmunkát végző cégek a léptékes módszerek használatáról általában úgy vélekednek, hogy afféle kompromisszumos megoldás, amelyet a „jobb híján”, „még elfogadhatónak” szoktak tartani,¹¹ amennyiben névlistán alapuló mintavételre nincs idő, pénz vagy lehetőség.

Az erdélyi vizsgálatok kapcsán arra a következtetésre jutottunk, hogy a léptékes mintavétel esetén az adatok tényleges megbízhatósága összességében három tényezőtől múlik: (a) az eljárási rutintól; (b) a kérdés minőségétől; (c) az ellenőrzés rendszerétől, illetve az ellenőrzésre fordított energiától. Amennyiben ezek közül valamelyikkel probléma van, a címlistas módszer teljesen egyértelműen megbízhatóbb, mivel ennek során mind a kérdezőbiztos, mind a terepmunkát koordináló személyek feladata egyszerűbb.¹² Mi magunk egy címlistas mintavételi módszer mellett döntöttünk, de ettől függetlenül fontosnak éreznénk egy egységes erdélyi magyar léptékes mintavételi standard kialakítását.

⁹ Pl. „*Indulj el az templom jobb oldalán, majd fordulj balra*” vagy „*Válaszd ki az első B betűs utcát, majd miután elfogytak a házszerűk, fordulj jobbra*”.

¹⁰ Ezt, amennyiben rendelkezésre áll, a választói névjegyzékre hagyatkozva lehet meghatározni, ennek hiányában pedig vezetőket a telefonok száma alapján lehet becsülni.

¹¹ Lásd például a TÁRKI módszertani elveit, amelyek szerint a véletlen sétás eljárás „piackutatási és lakossági attitűdvizsgálati célokra még megfelelő.” <http://szochalo.hu/hireink/article/100395/3218/>

¹² A kérdezőnek azért, mert konkrét személyt keres, a szervezőknek pedig azért, mert sokkal könnyebben tudják a kérdezőket ellenőrizni.

Erdélyben tehát saját bevett, folyamatosan fejlődő módszertani rendje alakult ki a magyarokat reprezentáló survey-vizsgálatoknak, amelyek azonban nem hivatkoznak a nemzetközi eljárás-rendre. Ez azért feltűnő, mert az erdélyi magyarsághoz hasonló, saját teljes körű regiszterrel nem rendelkező részpopulációk („rare population”) körében történő reprezentatív mintavételeknek már az 1970–80-as évekre komoly nemzetközi módszertani standardjai alakultak ki, amelyek alapvetően eltérnek ez eddig alkalmazott erdélyi módszerektől (áttekintő jellegű cikkek: Kalton – Anderson 1986; Hedges 1979; Hughes et al. 1995; Sudman – Kalton 1986 stb.). Az erdélyi vizsgálatok ugyanis a célpopulációról rendelkezésre álló összesített népszámlálási keretszámokat próbálták „feltölteni” válaszadókkal. Az említett módszertani ajánlások viszont fordított logika alkalmazását javasolják: egy nagy országos vagy területileg szelektált mintából előzetes szűrési eljárás („screening”) segítségével kell összegyűjteni a megfelelő számú válaszadót. Hogy ezt szűrési eljárást milyen módon lehet a leginkább költségkímélő módon elvégezni, annak persze szintén külön szakirodalma van. Amennyiben a vizsgált populáció földrajzi eloszlásáról vannak megbízható becsléseink – és az erdélyi magyarság ennek tipikus példája –, ezek az információk természetesen hasznosíthatók (például a nagyobb sűrűségű területek felülreprezentálásával), de keretként – hiszen személyi lista nem állítható össze belőlük – nem használhatók.

Az Életünk fordulópontjai – Erdély c. vizsgálatban alkalmazott mintavételi eljárás

A célpopuláció

A vizsgálat során alkalmazott mintavételi mód több vonatkozásban eltért az Erdélyben eddig követett eljárástól. Az első különbség a célpopuláció meghatározásából adódott, ugyanis az adatfelvétel a magyar nyelven (is) jól beszélő 20–45 év közötti erdélyieket célozta meg. A minta megtervezésekor gyakorlati megfontolások alapján eleve kihagytuk Szeben és Krassó-Szörény megyét. A 2002-es népszámlálás adatai szerint a kutatás 14 megyét lefedő földrajzi hatókörében¹³ élt a romániai magyarok 97,3, az erdélyi magyarok 98,3, míg a teljes erdélyi népesség 83,4 százaléka.

Az erdélyi magyarokra fókuszáló kérdőíves vizsgálatok a nemzetiségi önbevallás alapján határozták meg az alappopulációt. A kérdés konvencionálisan abban az esetben volt elkezdhető, ha a válaszadó a „*Magyarnak vallotta-e Ön magát a legutóbbi népszámlálás alkalmával*” kérdésre igennel válaszolt. E megoldás legfőbb előnye, hogy a népszámlálási gya-

¹³ A vizsgálat a Fehér, Arad, Bihar, Beszterce-Naszód, Brassó, Kolozs, Kovászna, Hargita, Hunyad, Máramaros, Maros, Szatmár, Szilágy és Temes megyékre terjedt ki.

korlattal megegyező statisztikai konstrukcióra hagyatkozik, így súlyozás és az eredmények ellenőrzése során alkalmazhatók a vizsgált népességre vonatkozó standard táblázatok. Mi a nyelvismeretre alapozó szelekciós eljárást¹⁴ használtuk, amelynek egyik fő hátránya éppen abból adódik, hogy nem ismerjük a vizsgált népesség jellemzőit. A 2002-es népszámlálás adataiból létrehozható csoportok közül a magyar anyanyelvűek állnak legközelebb a mi vizsgálati populációnkhoz, azonban nem fedti azt le teljes mértékben, hiszen nem tudjuk, hogy a nem magyar anyanyelvűek közül hányan beszélnek magyarul

1. táblázat. A magyar nemzetiségűek és anyanyelvűek száma a vizsgált 14 megyében¹⁵

Megnevezés		Magyar nemzetiségű és anyanyelvű	Magyar nemzetiségű, de nem magyar anyanyelvű	Nem magyar nemzetiségű, de magyar anyanyelvű	Összesen
Teljes népesség	szám	1 366 785	27 765	43 807	1 438 357
	százalék	95,0	1,9	3,0	100,0
20–45 évesek	szám	440 030	9260	14 550	463 840
	százalék	94,9	2,0	3,1	100,0

Látható, hogy a magyar anyanyelvűek és nemzetiségűek között nagy az átfedés (összehasonlításképpen: jóval nagyobb, mint más romániai kisebbségek esetében). A vizsgált területen 1 438 357 személy vallotta magát legalább az egyik ismérv szerint, ebből 1 366 785 (94,5%) mindkét ismérv szerint magyarnak. Bizonytalanul azonosítható (a népszámlálás szerint) az a 71 572 személy, aki anyanyelvét és nemzetiségét inkonzisztens módon adta meg. A 20–45 év közötti népesség esetében a megoszlások hasonlóak.

Három, egymással összefüggő érv szól a nyelv szerinti szelekciós kritérium mellett:

(1) A magyarul (is) beszélőket megcélozva a gyakorlatban tágabban értelmezett alapsokasággal dolgoztunk. Mint az majd a mintavételi eljárás leírásából ki fog derülni, hogy bár a nem magyar nemzetiségű, de magyarul is beszélők megkérdezése többletköltséggel járt, a mintába került magyar nemzetiségűek számát nem befolyásolta. A nyelvi szelekciós kritérium alkalmazásával a mintába kerülhetnek: (a) magyar anyanyelvű¹⁶ és magyar nemzetiségű, (b) magyar anyanyelvű, de nem magyar nemzetiségű, (c) illetve nem magyar anyanyelvű és nem magyar nemzetiségű, de magyarul (is) jól beszélő személyek. Ez utóbbiakra vonatkozóan nincsenek népszámlálási adataink. A nemzetiségi önbesorolásra alapozó mintavétel ez utóbbi két kategóriát kizárja, viszont elvben beveszi azokat a magyar nemzetiségűeket, akik nem beszél-

¹⁴ A kérdezőbiztosok azt az utasítást kapták, hogy amennyiben a kérdés (teljes kérdések vagy kérdésblokkok fordítása nélkül) magyar nyelven megoldható, készítsék el az interjút attól függetlenül, hogy a megkérdezett milyen nemzetiségűnek vallja magát.

¹⁵ A teljes népességre vonatkozó számok teljes körű népszámlálási adatok, a 20–45 évesekre vonatkozó adatokat az IPUMSI 10 százalékos népszámlálási mintája alapján becsültük.

¹⁶ Itt természetesen látnunk kell, hogy az anyanyelv megadásának több dimenziója van. A funkcionális nyelvtudás (legjobban, leggyakrabban beszélt nyelv) mellett utalhat szimbolikus-érzelmi kötődésre vagy egyszerűen egy biográfiai tényre (ezt a nyelvet beszélte gyermekkorában a családban). Az anyanyelvi regisztrációról lásd Arel (2002).

nek magyarul.¹⁷ Azért csak elvben, mert a magyar mintán végzett országos felmérések egyikében sem alkalmaztak román nyelvű kérdőívet is, így a gyakorlatban a magyarul nem beszélőkkel nem készült interjú.¹⁸

(2) Az alkalmazott módszer lehetővé tette, hogy az elemzés során a magukat magyar-nak vallók szűkebb csoportját problémamentesen leválasszuk, ugyanakkor olyan nem magyar identitású csoportokról is információhoz juthattunk, amelyek valamilyen mértékben kapcsolódnak a magyar kultúrkörhöz (pl. a magyar nyelvű média vagy kulturális termékek potenciális fogyasztói, vagy migrációs pályájuk ugyanúgy vezethet Magyarországra felé, mint a magukat magyarnak vallóké).

(3) Nem utolsósorban olyan eljárást kerestünk, amellyel becsülni tudtuk a magyarul beszélő erdélyiek számát, amire az eddigi statisztikák, adatfelvételek nem szolgáltatnak semmilyen támponttal. (Erről lásd a kötetben *Horváth* tanulmányát.)

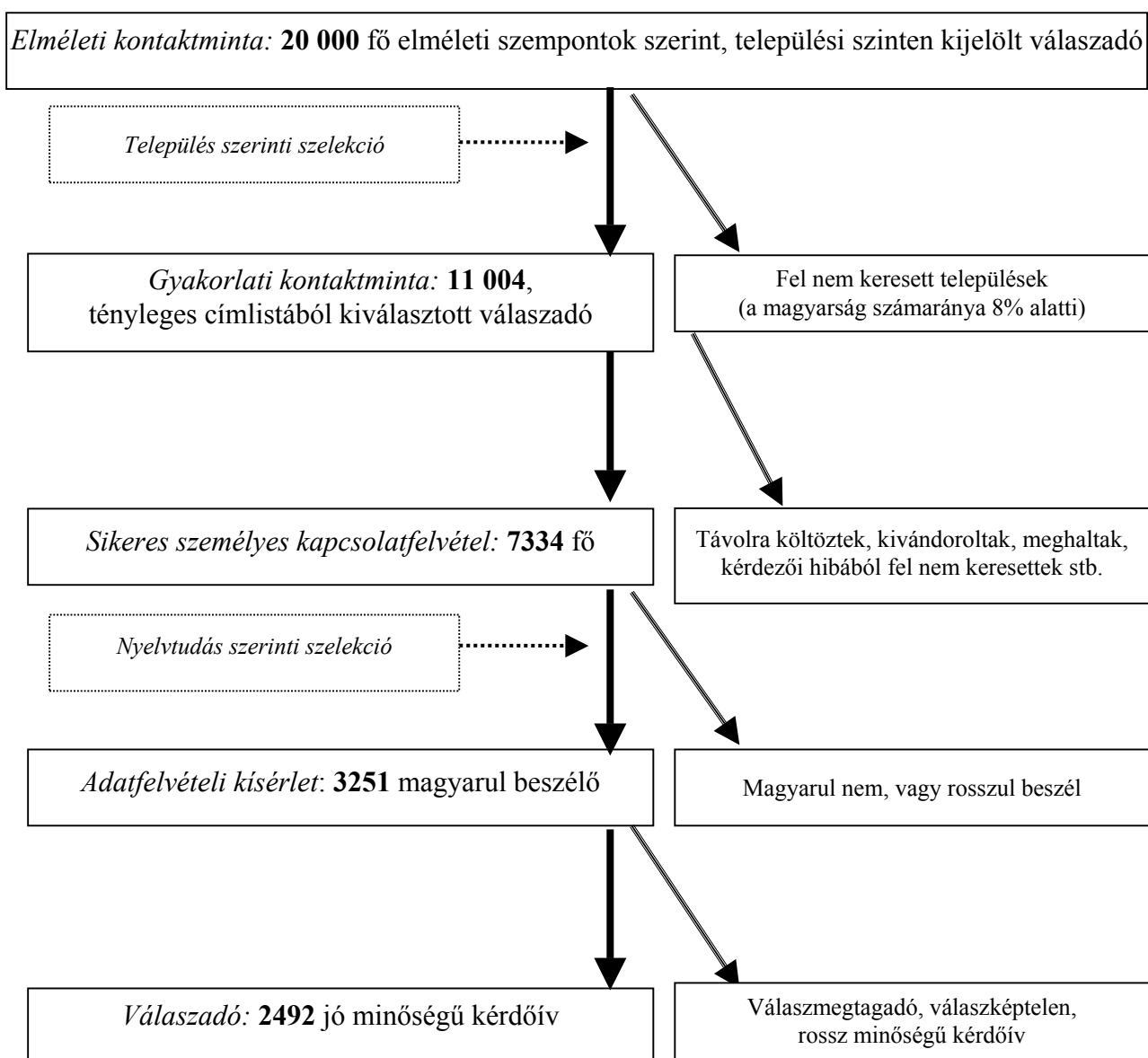
Az adatfelvétel logikája

A válaszadók egy négylépcsős szelekciós mechanizmuson keresztül kerültek a mintába. A négy szelekciós lépésből kettő (a települési és a nyelvi) a mi felvételünk specialitása volt, a másik kettő, amely a címlista pontatlanságára és a válaszadási hajlandóságra vezethető vissza, a regiszter alapján történő valamennyi adatfelvétel sajátja. A 1. ábra bemutatja az adatfelvétel általános logikáját, ezután külön-külön kitérünk az egyes szelekciós pontokra.

¹⁷ Természetesen a nem magyar anyanyelvű, de magyar nemzetiségű kategóriában elvben lehetnek olyanok, akikkel – bár nem magyar anyanyelvűek – a magyar nyelvű kérdés lefolytatható. Mivel azonban az anyanyelvi bevallásnak van egy szimbolikus-érzelmi dimenziója is, feltételezhetjük, hogy azok közül, akik magyar nemzetiségükhöz a népszámlálási regisztráció során ragaszkodnak, leginkább azok nem fogják anyanyelvüknek is a magyart vallani, akik azt nem beszélik társalgási szinten.

¹⁸ Ez alól a Max Weber Társadalomkutató Központ 2004. szeptemberi Temes megyei adatfelvétele képez kivételt. A magyar nemzetiségűekre fókuszáló kutatás során az interjúk 16 százalékát román nyelven folytatták le.

1. ábra. Az adatfelvétel logikája



Elméleti és gyakorlati kontaktminta

Az eddigi kutatási gyakorlattal összhangban vizsgálatunkban kétlépcsős, rétegzett mintavételi eljárást alkalmaztunk. Első lépésben kialakítottunk egy 20 ezer fős ún. elméleti kontaktmintát. Itt – más vizsgálatokhoz hasonlóan – kijelöltük az (elméleti) kutatási pontokat, mégpedig úgy, hogy az önreprezentáló települések mellett régió¹⁹ és településméret²⁰ szerinti rétegeket hoz-

¹⁹ Öt régióval dolgoztunk. Székelyföld: Hargita, Kovászna, Maros megye; Partium: Szilágy, Szatmár, Bihar megye; Észak-Erdély: Kolozs, Beszterce-Naszód, Máromaros megye; Dél-Erdély: Brassó, Hunyad, Fehér megye; Bánát: Arad és Temes megye.

²⁰ Tíz településkategóriával dolgoztunk: 250 ezer felett, 100 ezer és 250 ezer között, 30 ezer és 100 ezer között, 10 ezer és 30 ezer között, 4000–10 ezer között, 2000–4000 között, 1000–2000 között, 500–1000 között, 150–500 között, 150 alatt. Ezek közül az első négy kategória, vagyis a 10 ezer főnél nagyobb lélekszámú települések önreprezentálóak. A 150 alatti településeket gyakorlati okból kihagytuk a mintából és a nekik megfelelő esetszámokat a 150–500 közötti kategóriához rendeltük.

tunk létre. Az 10 000 fő alatti nem önreprezentáló települések esetén az egyes települési rétegekben belül a reprezentativitáshoz szükséges elemszámot egyenletesen osztottuk el, arra törekedve, hogy egy településen legalább 50 cím kerüljön kijelölésre. Az eddigi kutatási gyakorlattól eltérően az elméleti kutatási pontok meghatározásában nem volt szerepe a települések nemzetiségi (vagy anyanyelvi) megoszlásának, illetve a magyar népesség népszámlálás alapján feltételezhető földrajzi elhelyezkedésének.

A kiválasztott mintavételi pontokon belül kijelöltük a reprezentativitáshoz szükséges interjúalanyok számát, magukat a megkérdezetteket pedig a választói névjegyzékből véletlenszerűen választottuk ki.²¹ Értelemszerűen nem vettük figyelembe a potenciális válaszadók (feltételezett) nemzetiségét vagy magyarnyelvtudását. A 20 ezer fős elméleti kontaktminta tehát a 14 vizsgált erdélyi megyére reprezentatív névjegyzék alapján összeállított mintával egyezik meg.

Azontúl, hogy a településminta kialakításában nem hagyatkoztunk a census nemzetiségi megoszlásaira, az egyes településeken belül feltétlenül megkérdezendő interjúszámot sem határoztuk meg előre. Ez a másik alapvető különbség az eddigi felvételekhez képest. A kérdezőbiztosok, miután megkapták a célszemélyek nevét és címét, azt az utasítást kapták, hogy minden kontaktszeméllyel folytassák le az interjút, amennyiben azokat sikerül elérni és amennyiben a kérdés magyar nyelven elvégezhető.

A kérdőívszám előzetes meghatározásáról két tényező miatt mondtunk le. Egyrészt nem ismertük az alappopuláció (magyarul beszélők) megoszlásait, másrészt el akartuk kerülni a pótcímezt. Ez utóbbinak ugyanis „nagy hátránya: a tapasztalatok szerint a kérdező az alternatív megoldás tudatában nem tesz meg mindent a számára valamiért gondot jelentő személyek elérése és/vagy meggyőzése érdekében. Emiatt ez a megoldás a minta igen nagy rejtett, vagyis az elsődleges kemény demográfiai változók révén fel nem tárható torzulásával járhat.” (Kapitány 2003, 30)

Az elméleti kontaktminta kialakítása után azonban a célpopuláció népszámlálás alapján feltételezett területi megoszlását mégiscsak figyelembe vettük, a vizsgálat költségeit ugyanis nagymértékben, a sikeres interjúk számát pedig alig növelte volna, ha azokra a mintába került településekre is kérdezőket küldünk, ahol a 2002-es census szerint elenyésző arányban vagy egyáltalán nem élnek magyarok. Tehát az elméleti kontaktminta és a tényleges interjúk közé egy köztes fázist ékeltünk, amit gyakorlati kontaktmintának nevezünk. A gyakorlati kontaktminta a ténylegesen felkeresett kontaktszemélyek körét fedi. Praktikusán tehát azokra a településekre, ahol sikeres magyar nyelvű interjúra minimális volt az esély, nem küldtünk kérdezőbiztost. A határt az jelentette, hogy az adott településen a magyarok aránya a 2002-es népszámláláson elérte-e a 8 százalékot. Ennek hiányában úgy vettük, mintha az adott településen egyetlen sikeres kérdés sem valósult volna meg. A 20 ezer fős mintanagyság így 11 004 valójában felkeresendő személyre csökkent.

²¹ A választói névjegyzék a név és a lakcím mellett a személyi számot tartalmazza, amiből megállapítható az illető személy születési éve. A 20–45 éves minta kialakításában ezt használtuk fel.

2. táblázat. A kontaktszemélyek megoszlása a magyarok aránya szerint

<i>Minta</i>	<i>A magyarok aránya (%)</i>	<i>Kiválasztott mintaszemélyek</i>
Gyakorlati kontaktminta	75–100	2 152
	50–75	890
	30–50	1 445
	20–30	1 454
	10–20	3 561
	8–10	1 502
Összesen		11 004
Fel nem keresett településeken élő	8 alatt	8 964
Elméleti kontaktminta összesen		20 000

A döntésünk következményeként fellépő torzulás becsléséhez fel lehet használni azt az információt, hogy a vizsgált 14 megyében a 2002-es népszámlálás adatai szerint 75 391 magyar élt 8 százalékos alatti szórásban. A magyarul beszélők száma ezeken a településeken minden bizonnyal alatta marad a magyar nemzetiségűeknek, és a községsoros adatok alapján arra is következtethetünk, hogy az előrehaladott elöregedési folyamat miatt a 20–45 évesek aránya is jóval az erdélyi átlag alatti. Ezzel a szelekcióval tehát minden valószínűség szerint a célpopuláció kevesebb mint 5 százalékosról mondtunk le.

Így a gyakorlati kontaktmintába összesen 119, a népszámlálási adatok alapján legalább 8 százalékosban magyar nemzetiségűek által lakott település került.²² A 10 ezer főnél népesebb települések önreprezentálóak voltak, az az alatti településeken jellemzően 50–50 választót jelöltünk ki úgy, hogy az adott településkategórián belüli összes cím aránya megfeleljen az adott településkategória lakosságon belüli arányának.

²² *Arad* megye (5) : Angyalkút; Arad; Pécska; Zimandtelek; Zimandújfalu.

Bihar megye (12): Bályok; Belényes, Belenyésújfalva, Berettyószéplak, Diószeg, Érkörtvélyes, Érmihályfalva, Margitta, Nagyvárad, Szalacs, Nagyszalonta, Székelyhid.

Beszterce-Naszód megye (3): Bethlen, Radnaborberek, Teke.

Brassó megye (7): Alsórákos, Barcaújfalva, Bodola, Brassó, Kőhalom, Négyfalva, Sárkány.

Fehér megye (5): Búzásbocsárd, Magyarbecse, Marosnagylak, Marosújvár, Nagyenyed.

Hargita megye (18): Agyagfalva; Alsó és Felsőválya; Balánbánya; Csíkdelne; Csíkszentdomokos; Csíkszentmiklós; Csíkszenttamás; Csíkszereda; Gyergyóditró; Gyergyószárhegy; Gyergyószentmiklós; Homoródújfalva; Székelykeresztúr; Székelyszenterzsébet; Székelyudvarhely; Tusnád; Vasláb; Zetelaka.

Hunyad megye (2): Lupény, Petrozsény.

Kolozs megye (9): Aranyosgyéres; Bánffyhunad; Bogátpuszta; Dés; Kolozsvár; Sinfalva; Szamosújvár; Szucság; Torda.

Kovászna megye (9): Erdőfüle, Kézdivásárhely, Kovászna; Mikóújfalva; Ozsdola; Réty; Sepsibodok; Sepsiszentgyörgy; Torja.

Máramaros megye (7): Felsőbánya; Giródtótfalu; Koltó; Magyarláros; Máramarosziget; Nagybánya; Szinéváralja.

Maros megye (20): Ádamos; Albistelep; Bözöd; Dicsőszentmárton; Galambod; Körtvélyfája; Ludas; Magyarkirályfalva; Marosvásárhely; Marosszentgyörgy; Mezőcsávás; Nagysármás; Nyárádszentbenedek; Radnót; Segesvár; Szászrégen; Szolokma; Szováta; Tuson; Vámosgálfalva.

Szatmár megye (9): Mezőpetri; Mikola; Nagykároly; Pele; Piskolt; Szatmárnémeti; Szatmárpálfalva; Szatmáruddvari; Tasnádszántó.

Szilágys megye (6): Magyargoroszló; Szilágysámson; Szilágysomlyó; Varsolc; Zilah; Zsibó.

Temes megye (7): Aurélháza; Gátalja; Lugos; Nagycsanád; Nagyszentmiklós; Varjas; Zsombolya.

Az gyakorlati kontaktmintától a tényleges kapcsolatfelvételig

A gyakorlati kontaktmintába került 11 004, választói névjegyzékben szereplő személyből azonban csak 7334-gyel sikerült felvennünk a kapcsolatot, csaknem 3700-zal nem. Ez meglepően alacsony arány. Az ok sokrétű, de mintegy fele arányban (1660 személy esetén) arra vezethető vissza, hogy a kérdezett nem tartózkodott az adott településen vagy annak elérhető közelségében. Ez igen magas szám még akkor is, ha figyelembe vesszük, hogy itt az aktív (leginkább mobil) korosztályokról van szó. A tartósan távol lévők mintegy kétharmada (64%-a) tartósan külföldön tartózkodott, a többiek az országon belül laktak más településen.²³

A sikertelen kapcsolatfelvételek másik fele több alcsoportra oszlott: mintegy 800 esetben semmit sem sikerült kideríteni az adott címlistán szereplő potenciális válaszadóról, mintegy 400 esetben legalább annyi bizonyosan megállapítható volt, hogy az illető sohasem lakott a megadott hivatalos címén, 269 esetben a cím sem létezett (vagy nem sikerült megtalálni), 45 válaszadó elhunyt az utolsó választás óta, több mint 400 (jellemzően brassói, illetve aradi) címet pedig a kérdezőbiztosok nem is kerestek fel.

Összességében megállapítható, hogy a helyben (településeken) őrzött és jogszabály szerint nyilvános választási névjegyzékek nem képesek követni az erdélyi népesség intenzív migrációs mozgását. Még ha a tartósan külföldön tartózkodóktól eltekintünk is, 2006-ban (a választások után mintegy két évvel) a választási névjegyzéken szereplők mintegy 16 százaléka nem volt megtalálható. Itt azonban ismét hangsúlyozni kell, hogy különösen mobil korosztályról van szó.

A nyelvi szűrő

A 7334 fellelt válaszadóval akkor készült interjú, ha azok magyar nyelven képesek voltak válaszolni. Ezt a kérdezőbiztosnak kellett felmérnie. Azokkal, akik születési évüket és hónapjukat nem tudták megmondani magyarul, nem is tettek kísérletet .

A 7334 fő 47 százaléka a kérdezők megítélése szerint (szinte) egyáltalán nem beszélt magyarul, 6 százaléuk tudott valamennyire, de annyira nem, hogy ez a kérdőív kitöltéséhez elegendő lett volna. 44 százaléuk (3251 fő) beszélt jól magyarul, 3 százaléuk viszont annyira elutasító volt, hogy erről a tényről nem lehetett meggyőződni.

A kérdezők természetesen nem minden esetben tudták előzetesen felmérni a válaszadók nyelvtudását, a végül sikeresen felvett interjúk közül huszonötől nyilatkoztak utólag úgy, hogy „ha tudtam volna, hogy ennyire gyengén beszél magyarul, nem kérdezem le a kérdőívet”. Ilyen alacsony arány mellett azonban nem kell komoly torzítással számolni.

²³ A tartósan külföldön tartózkodókról kitöltött címkártyák elemzése a későbbiekben hozzájárulhat az erdélyi migrációs trendek elemzéséhez.

Természetesen feltehető, hogy a válaszadók nyelvi kompetenciájának feltérképezése nem volt független az interjú készítőjének személyétől. Kérdőbiztosaink – mint az erdélyi felvételek során általában – néhány kivétellel kolozsvári diákok, illetve fiatal helyi értelmiségiek voltak, jellemzően kérdőbiztosai tapasztalattal.²⁴ Az egyébként alkalmazottnál jobban kidolgozott, félnapos, kis csoportokban történt kérdőbiztosai oktatással és részletes kérdezői útmutató megszerkesztésével törekedtünk arra, hogy lehetőleg egységesítsük munkájuk színvonalát és csökkentsük az adatszolgáltatókra gyakorolt szubjektív hatásukat.

Sikeres kérdőívek

A 3251, magyarul jól beszélő potenciális válaszadó közül végül 2504-gyel sikerült interjút készíteni, 2492 esetében elfogadható minőségben. A sikertelenség oka túlnyomórészt, csaknem 700 esetben szigorú vagy lágy válaszmegtagadás volt. (Utóbbira példa, ha az interjúalany konkrétan nem utasítja vissza a válaszadást, de az időpont-egyeztetés nem jár sikerrel, mert sosincs rá ideje stb.) A válaszmegtagadókon túl 65 személyt minősítettek a kérdezők válaszképtelennek (súlyos betegek, idült alkoholisták stb.). A válasz-megtagadási arány ugyan nem alacsony, de megfelel a más európai országokban és Magyarországon megfigyelt mértéknek.

Az adatbázis súlyozása

Az adatbázis súlyozásának az a célja, hogy pontos alapsokasági eloszlások segítségével korrigálja azokat a kisebb torzulásokat, amelyek a terepmunka során fellépnek. Az eljárás abban áll, hogy bizonyos változók minta szerinti eloszlását hozzáillesztjük az alapsokaságban mutatkozó eloszláshoz. A kérdés mindig az, hogy honnan származtatjuk a bázist. A súlyozási keret rendszerint a népszámlálási adatállomány. Általában problémát jelent, hogy a népszámlálás óta eltelt időben (jelen esetben 4–5 év alatt) nem változik-e meg olyan mértékben a társadalom, hogy a súlyozás inkább torzítja adatainkat. Az „Életünk fordulópontjai – Erdély” esetén azonban ennél is komolyabb problémát jelentett, hogy a népszámlálásból alapsokasági adatok nem álltak rendelkezésre, hiszen abból a „magyarul beszélők” csoportja nem válogatható le. Ezért úgy döntöttünk, hogy két súlyt készítünk, egyet a teljes mintára (magyar nyelvű közösség), egyet pedig a magyar nemzetiségűek almintájára.

Ez utóbbi almintát a sztenderdeknek megfelelően a 2002-es népszámlálás magyar nemzetiségű csoportjához súlyoztuk nem, iskolai végzettség (max. 8 osztály; középfok; felsőfok), korcsoport és „etnikai kiséregió” szerint. Mivel az erdélyi magyarság körében 2002–2006 között igen jelentős *oktatási expanzió* zajlott le, a népszámlálási adatok ezen a területen ma

²⁴ A terepmunkát a Max Weber Társadalomkutató Alapítvány végezte, együttműködésben az Etnikumközi Kapcsolatok Kutató Intézetével (CCRIT).

már nem tekinthetőek reálisnak, ezért a 20–30 éves korcsoport esetén az iskolai végzettséget súlyozási szempontként nem vettük figyelembe

Az „etnikai kistrégió” olyan összevont csoportosítás, amely egyszerre próbálja figyelembe venni a településnagyságot, a regionális elhelyezkedést és a település etnikai jellegét. A kialakított 15 kategória a népszámlálási magyar populációt hozzávetőleg hasonló nagyságú, 100 000 fő körüli csoportokra tagolja. A 15 „etnikai kistrégió” a következő: Marosvásárhely; Kolozsvár; Szatmárnémeti és Nagyvárad; szórvány nagyváros; Csíkszereda és Sepsiszentgyörgy; egyéb székely város; vegyes kisváros; szórvány kis és középváros; határmenti magyar tömb;²⁵ partiumi kontaktzóna;²⁶ háromszéki tömb;²⁷ Hargita megyei tömb;²⁸ Marosszéki tömb;²⁹ erdélyi kontaktzóna;³⁰ szórvány.³¹

A súlycellák elkészítése során a gyakorlatban további kompromisszumot kellett kötni, néhány esetben a megfelelő súlycellák előállítására érdekében szükség volt további összevonásokra. Ebben az esetben nemcsak etnikai kistrégiókon belül a maximum 8 osztályt végzett, illetve középfokú végzettségű csoportokat egyesítettük.

A magyar nemzetiségű al minta súlyozását elvégezve készítettük el a teljes minta súlyváltozóját. Ennek során a nem magyar nemzetiségűek első lépésben ugyanakkora súlyt kaptak, mint a velük (nem, iskolai végzettség, korcsoport és etnikai kistrégió szerint) megegyező csoportba tartozó magyar nemzetiségűek, majd a teljes minta súlyait korrigáltuk a valós elemszámhoz. Összességében a súlyozatlan teljes mintában a 0,33 alatti súlyok aránya 1,2 százalékos, a 3 felettié 0,3 százalék. Ezek az arányok a súlyok nem ideális, de azért elfogadható szóródását jelzik.

²⁵ Észak-Bihar magyar többségű községei és a Nagykároly környéki sváb-magyar falvak.

²⁶ Észak-Bihar maradék része és Szalonta környéke: Szatmár megye maradék része, leszámítva az Avast és a megye keleti községeit; Szilágy megye északkeleti része; Arad Marostól északra lévő, magyar lakossággal rendelkező községei.

²⁷ Kovászna megye községei, leszámítva a Bodzafordulót és a megye nyugati vegyes, cigány-román-magyar községeit (Hidvég, Előpatak).

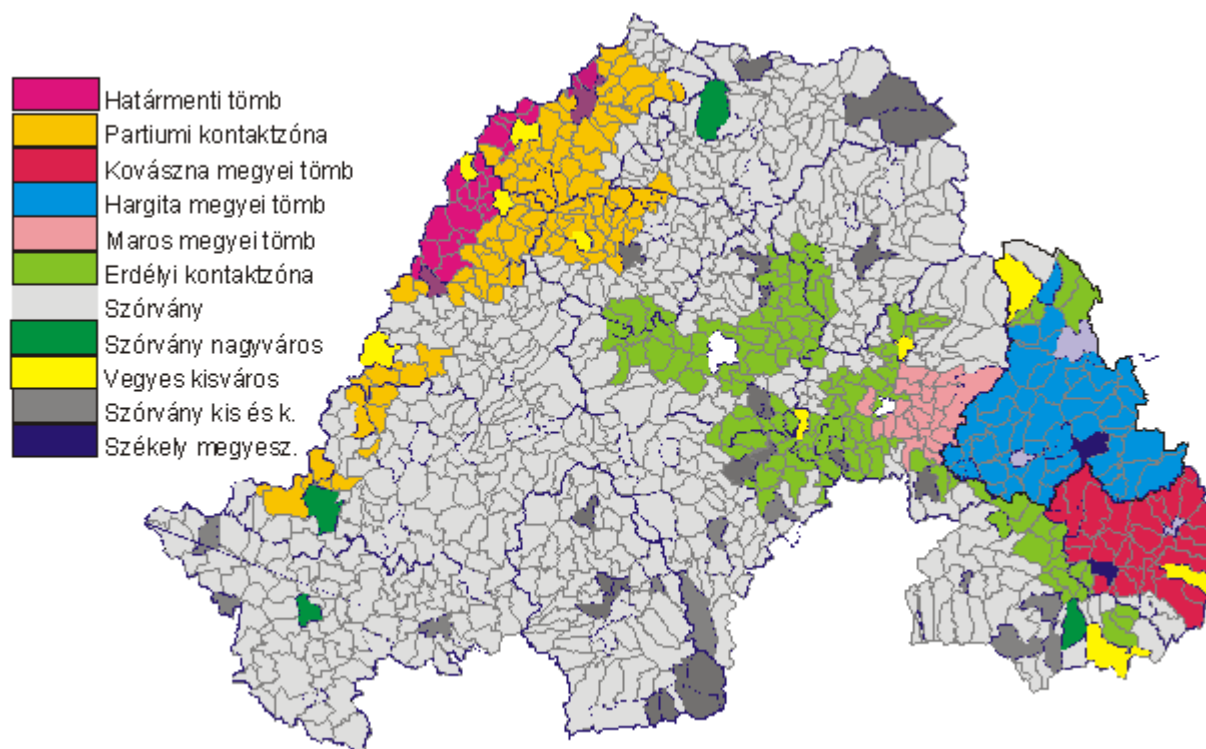
²⁸ Hargita megye községei, leszámítva a megye román többségű északi községeit.

²⁹ Az egykori Marosszék, leszámítva azokat a Marosvásárhely környéki (magyar-román-cigány) községeket, amelyek mára elveszítették abszolút magyar többségüket.

³⁰ Kolozs megye részben magyar lakta részei (Kalotaszeg, Mezőség, Aranyosmente), Beszterce-Naszód Déstől keletre eső néhány jelentős magyar lakosságú községe; Kalotaszeg Szilágy megyébe eső része; Fehér megyében Nagyenyed környéke; Maros megye nem marosszéki részének többsége (Küküllőmente, Mezőség, Vásárhely környéke, Marosmente Szászrégen és Magyaró között), de nem maradéktalanul (nem tartozik ide a Görgény-völgye, a Segesvártól nyugatra eső rész és a Maros megyei mezőség nyugati fele); Brassó megye keleti pereme (Apáca, Ürmös, Alsórákos és az Udvarhelytől elcsatolt falvakat tartalmazó községek, valamint Hétfalu falvai); Hargita északi része; a Kovászna megye nyugati részén levő vegyes községek.

³¹ Temes; Arad Maros alatti és egyéb maradék részei; Máramaros (leszámítva az egykor Szilágyhoz tartozó vegyes lakosságú Szamosardót; Hunyad; Beszterce (leszámítva a megye dél-nyugati peremén lévő néhány községet); Fehér, Nagyenyed környéke nélkül; Dél-Bihar, leszámítva Szalonta környékét (vagyis a Fekete-Körös völgye ide kerül); Szatmárból az Avas; Marosból a Maros és a Görgény-völgye, valamint a Mezőség; Kolozsból a Kisszamos mente, valamint a Mezőség és a havasi részek; Kovásznából a Bodza-forduló; Brassó maradék része; Hargitából Bélbor község.

1. térkép. A súlyozás során használt etnikai kistérségek



A 3. táblázatban néhány alapvető változó szerint mutatjuk be a súlyozás hatását, mind a teljes mintára, mind a magyar nemzetiségűekre szűkített almintára nézve. Az eredményekből látható, hogy az eljárás jelentősen, de nem alapvető mértékben módosított az eredményeken.

3. táblázat. Súlyozott és súlyozatlan eloszlások néhány alapvető arányszám szerint

Megnevezés	1. Teljes minta, súlyozatlan érték	2. Teljes minta, súlyozott érték	3. Nemzetiségi almintá, súlyozatlan érték	4. Nemzetiségi almintá, súlyozott érték
Férfiak aránya	47,6	51,5	47,2	51,4
Gyermektelenek aránya	39,2	43,1	38,8	43,3
Romának minősítettek aránya	7,9	9,5	5,6	6,2
Hargita és Kovászna megyeiek aránya	34,8	32,4	37,0	33,6
Házások aránya	61,8	58,0	62,5	58,6
Felsőfokú végzettségűek aránya	13,1	11,8	13,3	12,0
Esetszám	2492	2492	2216	2216

Újdonságok és lehetőségek

Az Életünk fordulópontjai – Erdély c. vizsgálat eredményei elméletileg három okból jelentenek újdonságot a hasonló tárgyú surveyek eredményeihez képest.

(1) **Eltérő mintavétel:** felvételünkben név- és címlistas mintavételt alkalmaztunk az Erdélyben általában használatos kvótás-léptékes vegyes mintavétel helyett.

(2) Az adatbázis súlyozása: súlyozott adatbázist készítettünk az általában használt súlyozatlan vagy nagyon egyszerűen súlyozott adatbázisok helyett.

(3) Eltérő alapsokaság: kutatásunk alapsokasága a magyar nyelvű lakosság volt, az általában alkalmazott magyarul beszélő magyar nemzetiségű lakosság helyett.

A következőkben néhány kiválasztott változón bemutatjuk a korábban végrehajtott erdélyi magyar surveyek és az Életünk fordulópontjai – Erdély c. vizsgálat eredményei közötti eltérést, illetve elkülönítjük azt a fent jelzett három komponens szerint. Ehhez egy hasonló alapsokaságra vonatkozó, megfelelő elemszámú összehasonlító adatbázist építettünk fel, két surveyből: a CCRIT által az RMDSZ felkérésére 2004 szeptemberében végzett, illetve a Max Weber Társadalomkutató Központ által 2004 decemberében végzett felmérésből. Ez utóbbi két adatállományból leválogattuk ugyanazokat a születési évjáratokat (1961–1988-as születésűek), amelyek a mi mintavételi sokaságunkat alkották.³²

Az összekapcsolt adatbázisba olyan változókat vontunk be, amelyek esetében előzetesen eltérést vártunk, és amelyek mindhárom kérdőívben (gyakorlatilag) azonosan szerepeltek. Ilyen volt a válaszadó iskolai végzettsége, a tanulmányok nyelve, a vegyes etnikai származás (édesapa, édesanya nemzetisége), a házastárs/partner nemzetisége, a nyelvhasználati szokások (otthon, munkahelyen stb.), a médiafogyasztás etnikai mintái, illetve etnikai nyelvhasználati sztereotípiákat mérő attitűdállítások.

Az etnikai identitáshoz kapcsolódó kérdésekre adott válaszok eltérése szinte minden esetben statisztikailag jelentős mértékű volt, mégpedig az előzetesen feltételezett szerint. Az Életünk fordulópontjai – Erdély adataiból egy etnikailag kevésbé zárt közösség képe rajzolódik ki, igazolva hipotézisünket. Abban bízunk ugyanis, hogy az új, költséges, de a nemzetközi standardokhoz közelebb álló adatbázis valamit csökkent majd az erdélyi magyar kérdőíves vizsgálatok jellegzetességén azon régóta megfigyelt, miszerint „*a nemzetiségileg tagolt mintavétel során kialakuló populáció a vegyes és szórványterületeken – az etnikai alapú nyilván tartások pontatlansága, a terepmunka nem megfelelő ellenőrizhetősége miatt – szinte törvényszerűen »magyarabb« lesz*” (Csata 2004).

A kapott eredményekből mutat be néhányat a 3. tábla (vesd össze az első és negyedik oszlopban található értékeket). Míg például a hagyományos erdélyi surveyek szerint a nem magyar etnikumú partnerek (házas és élettársak) aránya 8,9 százalék, az Életünk fordulópontjai – Erdély 19,5 százalékot mér; a dominánsan magyar nyelvű munkahelyi környezet aránya esetén ez az érték 57 vs. 47 százalék; a magyar nyelvű középiskolát végzettek aránya 27% vs. 46% stb.

Nyilvánvalóan adódik a kérdés, hogy melyik adat a „helyes” a „jó”. Ennek megválaszolásához azonban előzetesen azt kell tisztázni, hogy a jelzett három tényező közül mely(ek) ezeknek a jelentős eltéréseknek az oka(i)? Talán az, hogy máshogy választottuk ki a válaszadóinkat (a mintavételi módszer hatása), alkalmasint az, hogy a mi adatbázisunk súlyozott

³² Két kontrolladatbázis bevonására a megfelelő elemszám biztosítása érdekében volt szükség (a másik két survey teljes, az idősebbeket is tartalmazó elemszáma alig haladta meg az ezer-ezer főt).

(súlyozási hatás), vagy esetleg az eltérő alapsokaság hatásáról lehet szó? A feleletet az egyes oszlopok összevetéséből olvashatjuk le: általában, de egyáltalán nem kizárólag az alapsokaság eltérése játszik közre.

Ha azonban az alapsokaságok eltérésétől eltekintünk (vagyis az 1. oszlopot a 3.-kal vesszük össze), az esetek jelentős részében még mindig jelentős differenciákat találunk. Ezek közül néhány a mintavételi hatásra, néhány a súlyozásra, néhány pedig mindkettőre vezethető vissza.

A 4. táblázat az eredmények közötti eltéréseket a fenti tényezők szerinti felbontásban közli.

4. táblázat. *Eltérések kutatások és felvételek szerint, %*

Megnevezés	1. vs. 2.: a mintavételi módszer hatása	2. vs 3.: súlyozási hatás	3. vs. 4.: az eltérő alapso- kaság hatása	
	CCRIT – Max Weber összevont adatbázis	2. ÉF – „nép- számlálási ma- gyar” alminta, súlyozatlan	3. ÉF – „nép- számlálási ma- gyar” alminta, súlyozott	4. ÉF – teljes magyar nyelvi minta
Az 5–8. osztályt románul végezte (az érintettek körében)	14,6	14,9	16,5	23,2
A középiskolát románul végezte (az érintettek körében)	26,9	37,6	40,6	46,0
Házas(élet)társa nem magyar nemzetiségű	8,9	9,9	11,8	19,5
Édesanyja nem magyar nemzetiségű	4,5	3,4	3,6	11,6
Édesapja nem magyar nemzetiségű	4,3	2,9	3,2	11,7
Önbesorolása szerint akcentussal vagy rosszul beszéli a magyar nyelvet	3,3	5,5	7,0	12,1
Kizárólag magyarul beszél otthon	91,8	91,4	89,8	83,5
Csak magyar nyelvű tévét néz	34,6	18,9	18,0	17,5
Dominánsan magyar munkahelyi környe- zet (a munkahellyel rendelkezők körében)	57,1	53,0	49,2	46,8
„Helyesen cselekszenek azok a magyar szülők, akik gyermekeik érdekét nézve, azokat román tannyelvű iskolába adják.” – egyetért	25,7	25,5	25,7	29,6
A válaszadó				
iskolai	alapfok	7,4	11,6	18,7 ³³
végzettsége	„tág” középfok	76,1	75,2	69,5
	felsőfok	16,5	13,2	11,8

A 4. táblázat adatainak ismeretében érdemes visszatérni a „melyik adat a pontosabb?” kérdésre. Az eltérés három komponense közül az első kettő (a címlistas mintavétel és a súlyozás) meggyőződésünk szerint pozitív hatású, közelebb visz a „valósághoz”. E meggyőződésünket a népszámlálási, illetve az egyéb regiszteradatokkal való összehasonlítás támaszthatja alá.

Ilyen összevetésre a fenti változók közül a vegyes házasság és az iskolai végzettség, illetve részben az iskolai oktatás választott nyelve esetén van lehetőség. A 2002-es népszám-

³³ Mivel az iskolai végzettség súlyozási szempont, a két oszlop adatai a választott súlyozási módszer miatt megegyeznek (lásd a súlyozás ismertetését).

lálási adatok szerint a párkapcsolatban (élettársi kapcsolatban vagy házasságban) élő 20–45 év közötti magyarok 16,5 százalékának volt más nemzetiségű párja. Népmozgalmi statisztikák azt jelzik, hogy az 1992 és 2001 között³⁴ házasságot kötött 160 887 magyar nemzetiségű közül 28 401 fő (17,6%) választott nem magyar nemzetiségű partnert (Horváth 2004). Persze, mivel a fenti adatok Románia egész területére és nem az általunk vizsgált területre vonatkoznak, az összehasonlítás mindét esetben többszörösen is korlátozott, de a teljes körű adatokhoz egyértelműen az Életünk fordulópontjai – Erdély adatai állnak közelebb.

Az iskolai végzettségről a 2002-es népszámlálás adatai állnak a rendelkezésünkre. Eszerint az adatbázis alapsokaságát jelentő 14 erdélyi megyében a 20–25 éves korcsoportban a magyar nemzetiségű lakosság képzettség szerinti megoszlása a következő: felsőfokú végzettségűek: 7,3%, középfokú végzettségűek (ideértve a szakmunkásképzést is): 65,9%; maximum alapfokú végzettségűek: 26,7%. Ha számolunk is a népszámlálás és az adatgyűjtések közötti időben bekövetkezett oktatási expanzióval, láthatjuk, hogy a súlyozatlan adatbázisok közül is az Életünk fordulópontjai – Erdély adatbázisa van közelebb a népszámlálási adatokhoz. A súlyozott állomány pedig (mivel az iskolai végzettség részleges súlyozási szempont volt) reményeink szerint már meglehetősen jól közelítheti a 2006-os állapotot.

Elméletileg mód lenne arra hogy, az oktatás nyelvére vonatkozóan is pontos teljes körű adatokat állítsunk elő (az oktatási és a népesség-nyilvántartási adatok egymásra vetítésével), azonban ez jelenleg nem oldható meg. Ennek hiányában kiindulópontként elfogadhatjuk Murvai László becslését (Murai 2000), aki szerint a magyar nemzetiségű gyermekek mintegy 20 százaléka járt alapfokú román iskolába. Ilyen feltételezés mellett a mi súlyozott adataink tűnnek ismét a teljes körű adatokhoz közelibbnek lévőknek.

Ugyanakkor 2005/06-os tanévre vonatkozó, nyilvánosságra nem hozott tanügyi adatok szerint a tanév kezdetén a magyar nemzetiségüként regisztrált gyermekek 14,3%-a végezte nem magyar nyelven általános iskolai tanulmányait, ami látszólag a korábbi surveyek eredményeit támasztja alá. Azonban ez az adat nyilván valamiféle alsó korlát: a román iskolába beíratott magyar nemzetiségű diákok jelentős része feltehetően románként regisztráltatta magát, míg ennek a fordítottja bajosan képzelhető el. Ráadásul itt egyetlen év adatáról van szó, míg a vizsgált kohorszok legalább egy, de jellemzően több évtizeddel ezelőtt jártak általános iskolába.

Végül is a teljes körű, illetve más forrásokból származó adatoknak ez az igen korlátozott – csak becslésekre alkalmas – összevetése arra utal, hogy a név- és címlistas mintavétel és az adatbázis súlyozása egyértelműen pontosabbá, megbízhatóbbá tette az adatokat, bár még a súlyozott név- és címlistas adatbázis is minden valószínűség szerint a magyar nemzetiség erősebb etnikai zártságát mutatja a valóságnál.

Az eddig elmondottakkal szemben nem egyértelmű a következménye, ha az alapsokaságot kiterjesztjük a nemzetiségi önbevallás által meghatározott csoporttól a nyelvi közösség

³⁴ A mintaanyagok jellemzően ebben a korban kötöttek házasságot.

irányába. Az utóbbi esetben ugyanis mindig a vizsgált kérdéstől függ, mi számít jobb megoldásnak.

Végül röviden térjünk vissza arra, hogy milyen okok vezetnek a kérdőíves felmérések eredményeinek eltéréseire, és foglaljuk össze, hogy ezekből milyen általánosítható következtetéseket vonhatunk le.

A mi eljárásunk módosító hatása irányában megfelel a várakozásnak, de feltételezett-nél kisebb a hozadéka. Mivel az ilyen típusú mintavételek igen komoly pénzügyi teherrel járnak, kérdéses, hogy az általunk alkalmazott módszer bevett gyakorlattá válhat-e az erdélyi magyar kérdőíves vizsgálatokban.

A súlyozás – mivel elsősorban módszertani felkészültség, nem pedig anyagiak kérdése – könnyen és gyorsan elterjeszhető az erdélyi magyar surveyekben, és igazoltan sok haszonnal járhat, ha megbízhatóbb adatokhoz szeretnénk jutni. Ugyanakkor számolni kell azzal, hogy az ennek révén elérhető hatás összefügg a mintavételi eljárással is. Az esetünkben elért látványos javulás annak tudható be, hogy az adatfelvétel területi torzulását sikerült kiküszöbölnünk. Ez utóbbi abból adódott, hogy a magyarok által kisebb arányban lakott településekről a vártnál kevesebb kérdőívet sikerült összegyűjteni. A súlyozással az innen beérkezett válaszok értékelődtek fel. A hagyományos („presbitériumra reprezentatív”) vizsgálatokon azonban a súlyozás sem segít. Ezekkel ugyanis nem az a baj, hogy a szórványtelepülésekről kevesebb kérdőívhez jutunk, hanem hogy ezek (nagyobb számuktól függetlenül) nem reprezentálják ténylegesen a magyar lakosságot.

Végül az alapsokaság tágabb lehatárolásával kapcsolatban nem hisszük, hogy általában szükséges az erdélyi magyar társadalom kutatásának horizontját kiterjeszteni a magát nem magyar nemzetiségűnek valló, de magyarul jól beszélő csoportokra.³⁵ Úgy véljük azonban, hogy bizonyos esetekben (pl. politikai marketing, piackutatás, nemzetközi eladhatóság miatt stb.) ez a lépés indokolt lehet. Úgy tűnik, hogy a nyelvi közösséget feltérképező adatbázis által mutatott etnikai zártság paradox módon többé-kevésbé megfelel annak a képnek, amelyet a teljes körű adatok alapján a nemzetiségi közösségről feltételezhetünk.

Irodalom

- Arel, D. (2002): Language categories in censuses: backward- or forward-looking? In: Kerzer, D.I. – Arel, D. (coord.), *Census and Identity. The Politics of Race, Ethnicity, and Language in National Censuses*. Cambridge: Cambridge University Press: 92–120.
- Csata Zsombor (2004): Az iskoláztatás nyelvének szociológiai háttere Erdélyben. In: Kiss Tamás (szerk.): *Népesedési folyamatok az ezredfordulón Erdélyben*. Kriterion–RMDSZ Ügyvezető Elnökség, Kolozsvár.
- Hedges, B.M. (1979): Sampling minority populations. In: Wilson, M.J. (ed.): *Social and Educational Research in Action*. London, Longman: 244–261.

³⁵ Eredményeinek szerint a legjelentősebb ilyen csoportok: magyar nyelvű romák, partiumi svábok, ve-gyes, asszimilálódott családban élő magyarok, dominánsan magyar nyelvi térben élő románok.

- Horváth István (2004): Az etnikailag vegyes házasságok az erdélyi magyar lakosság körében: 1992–2002. In: Kiss Tamás (szerk.): Népesedési folyamatok az ezredfordulón Erdélyben. Kriterion–RMDSZ Ügyvezető Elnökség, Kolozsvár.
- Hughes, A.O. – Fenton, S. – Hine, C. (1995): Strategies for sampling black and ethnic minority populations. *Journal of Public Health Medicine*, 17(2): 187–192.
- Kalton, G. – Anderson, D. W. (1986): Sampling Rare Populations. *Journal of the Royal Statistical Society Series A* vol 149. No. 1: 65–82.
- Kapitány Balázs (2003): Életünk fordulópontjai. Módszertan és dokumentáció. KSH–NKI. Budapest.
- Koch, A. (1997): ADM-Design und Einwohnermelderegister-Stichprobe. In: *Stichproben in der Umfragepraxis*. Westdeutscher, Opladen.
- Murvai László (2000): A számok hermeneutikája. A romániai magyar oktatás tíz éve 1990–2000. A Magyar Nyelv és Kultúra Nemzetközi Társasága, Budapest.
- Rudas Tamás (1998): *Hogyan olvassunk közvélemény-kutatásokat? Új Mandátum*, Budapest 1998.
- Sudman, S. – Kalton, G. (1986): Annual New Developments in the Sampling of Special Populations *Review of Sociology*, August 1986, Vol. 12: 401–429.