

A DIGITÁLIS KORPUSZOKON ALAPULÓ NYELVTÖRTÉNETI VIZSGÁLATOK KORLÁTAI

KALCSÓ GYULA

Az adatközpontú nyelvtörténeti bűvárlatoknak mindig alapvető követelménye volt a felhasznált források lehető legnagyobb reprezentativitása. A nyelvészeti korpuszok számítógépes háttér hiányában azonban ritkán ölelhetek fel jelentős szövegmennyiséget, ha mégis, akkor egész munkaközösségekre hárult a feldolgozás fáradtságos munkája (mint pl. a TNyt. munkálatai során), vagy a kutató hosszú éveket áldozott hatalmas gyűjteményének kiaknázására (mint pl. Benkő Loránd a korai magyar felvilágosodás irodalmának vizsgálatakor). A digitális korpuszok nagyobb adatmennyiséget, a gyorsabb feldolgozás lehetőségét kínálják a kutatóknak, felhasználásukhoz azonban nagy körütekintésre, alapos előkészítő munkára van szükség. A szinkrón szövegegyüttesek építése és annotálása is sok nehézséget rejtő feladat, ám a diakrón korpuszok esetében a problémák még számosabbak (l. pl. Kiss – Kiss – Pajzs 2004). Talán ez az oka annak, hogy a magyar nyelv felvilágosodás előtti történetét egyelőre nem tanulmányozhatjuk mindenki által hozzáférhető, reprezentatív digitális gyűjtemény segítségével.

Jelen tanulmányban olyan problémákkal foglalkozom, amelyek egy saját készítésű, a magyar nyelvű könyvnyomtatás első fél századát (1527–1576) felölelő digitális korpusz feldolgozása során merültek föl. Ez a gyűjtemény 2001 és 2004 között jött létre, számítógépen rögzített, ún. plain ('egyszerű' v. korpusznyelvészeti értelemben 'nem annotált') szövegegyüttes, amely 238 877 szövegszövből (mintegy 1 200 000 n-ből) áll. A szövegminták rögzítése betűhív átiratban történt, ún. procedurális kódolású állományokban (azaz szövegszerkesztővel, nem nemzetközi szabványok szerint, hanem az egyedi kívánalmaknak megfelelő formai megoldásokat alkalmazva), ami ugyan a korpuszkészítés kevésbé szakszerű módját jelenti, és az elkészült korpusz terjesztését is korlátozottabban teszi lehetővé, azonban a rögzítésre és feldolgozásra rendelkezésre álló eszközök ezt a megoldást kívánták meg¹. A feldolgozás ún. számítógépes konkordanciaprogramok segítségével történt, amelyek közül jó néhány ingyenesen letölthető a világhálóról.

¹ A procedurális és (deskriptív) logikai kódolásról, valamint a korpuszkészítés legszakszerűbb módjáról l. az MTA Nyelvtudományi Intézete munkatársainak részvételével 2006. február 20-án tartott MANYE-ülés anyagát a világhálón: <http://corpus.nyud.hu/manye>.

A mennyiségi reprezentativitást a következők szerint biztosítottam. A korszakból a Régi magyarországi nyomtatványok (RMNy.) c. bibliográfia szerint 196 magyar nyelvű műről van tudomása a könyvtudománynak. Ebből 57 elveszett, vagy olyannyira töredékes formában maradt fenn, hogy érdemi nyelvészeti szempontú vizsgálatoknak aligha vehető alá. A hozzáférhető többiből 103 nyomtatvány² mintái kerültek a gyűjteménybe. A kihagyott művek többnyire olyanok, amelyeknek szerzőjétől ugyanazon nyomdából közel azonos időből, azonos műfajú szövegből már került be részlet (l. még alább, a minőségi reprezentativitásról szóló bekezdést). A 103 minta mindegyike vagy az adott szöveg 5%-a, azaz 20 oldalanként 1 oldalnyi, vagy a kisebb nyomtatványok esetében legalább ezer szövegszónyi részlet. Tíz rövidebb szöveg így teljes terjedelmében bekerült.

A minőségi reprezentativitást a következők biztosítják. Lehetőség szerint minden szerzőtől, minden nyomdából, minden kiadási évből, minden szövegtípusból (a kalendáriumtól a bibliafordításig) szerepel szövegrészlet. Ezen tényezők együttes reprezentációjára törekedtem (így pl. egy-egy szerzőnek akár minden műve szerepelhet a korpuszban, ha azok különböző helyeken és/vagy különböző kiadási évben és/vagy különböző műfajban jelentek meg). A szövegtípus (műfaj) szerint a következő csoportokat állapítottam meg: bibliafordítás (-részlet), énekeskönyv, hitvitázó (vallásos) próza, kalendárium, nyelvtan, prédikációgyűjtemény, törvénykönyv, világi próza, verses próza.

A korpusz feldolgozása a már említett konkordanciaprogramokkal történt. Ezek az egyszerű számítógépes alkalmazások a digitális formában tárolt szövegek különböző szempontú feldolgozására alkalmasak. A megfelelő utasításokat megadva összegyűjtik pl. egy-egy szó összes előfordulását egy vagy több szövegből, de alkalmasak ún. kollokációk (együttes előfordulások) keresésére is. Az alább ismertetendő vizsgálatban karakterláncrészletek keresését kellett végrehajtani. A feldolgozás technikai részleteinek ismertetésétől ez esetben eltekintek (l. Kalcsó 2005, 2007b), mivel jelen tanulmány középpontjában a kapott adatok elemzése során felmerült bizonyos problémák állnak.

² A nyomtatványok RMNy. I. szerinti sorszáma: 7., 8., 11., 12., 13., 14., 15., 16., 17., 18., 19., 20., 21., 22., 23., 24., 25., 26., 27., 39., 63., 47., 48., 49., 57., 58., 65., 64., 70., 74., 77., 78., 81., 85., 88., 80., 86., 88a, 91., 90., 92., 95., 98., 96., 99., 100., 101., 102., 103., 109., 108., 125., 137., 144., 151., 150., 154., 155., 156., 158., 159., 160., 161., 162., 165., 170., 164., 166., 169., 171., 172., 178., 185., 173., 181., 182., 183., 184., 186., 191., 191b, 192a, 193., 195., 192., 194., 196., 206., 205., 207., 208., 213., 218., 219., 220., 222., 230., 237., 238., 240., 229., 232., 233., 257., 240a, 241., 243., 242., 246., 253., 255., 259., 263., 266., 268., 260., 264., 265., 269., 273., 276a, 277., 281., 282., 283., 293., 276., 279., 280., 284., 286., 288., 289., 290., 294., 295., 296., 297., 298., 299., 301., 303., 304/1., 304/2., 307., 308., 308a, 308b, 311., 312., 314., 315., 318., 316., 321., 324., 331., 319., 320., 322., 323., 326., 327., 328., 333., 337a, 338., 339a, 334., 335., 337b, 340., 341., 342., 343., 344., 345., 339., 346., 347., 348., 349., 350., 355., 351., 352., 353., 357., 358., 359., 360., 362., 364., 367., 368., 370.

Az elkészült korpuszon végzett első vizsgálat (Kalcsó 2007a) anyaga a gyűjtemény névszóinflexiók alakjai közül a többféle változatban előfordulók köre volt. Hipotézisem szerint ezek korabeli nyelvi változók változóértékei. A több mint 49 000 adat elemzése során a következő kérdésekre kerestem a választ. Kimutatható-e az alakváltozatokról, hogy területi függést mutatnak (azaz nyelvjárási jelenségek)? Ha nem, akkor milyen tényezőktől függ az eloszlásuk? Kimutatható-e, hogy bizonyos alakváltozatok a (nyomatott) írásbeliség sajátjai? Ezzel összefüggésben kitapintható-e a sztenderdizáció folyamata, vagy akár a spontán sztenderdizálódás valamely fázisa? Tartható-e a nyomtatott könyvek nyelve autonóm lektusnak?

Ez utóbbi kérdéshez a nyelvi változók elkülönítése körüli nehézségek vezettek. A gyűjtemény adatai azonos nyelvváltozat produktumainak tarthatók-e? A diakrón korpuszok elemzésekor ez a kérdés kikerülhetetlen. Mivel a nyelvi változók per defínitíonem azonos nyelvváltozat különböző megoldásai ugyanazon funkció kifejezésére (vö. Kiss 1995: 62), ezért a kérdésre adott válasz kulcsfontosságú, hisz ha nem egy lektus termékeiből áll össze a korpusz, akkor még a funkcionálisan azonos alakváltozatai sem tarthatók egy nyelvi változó értékeinek. A felölelt időszak nyilvánvalóan nem tekinthető egy szinkrón metszetnek, ezért az alakváltozatok egy változóba sorolhatósága kérdéses. Az is nyilvánvaló, hogy az ország különböző részeiről származó szerzők bizonyíthatóan nyelvjárási eredetű nyelvi megoldásai különböző nyelvváltozatok megnyilvánulásai, tehát ugyancsak kétséges, hogy tarthatók-e egy változó értékeinek.

A vizsgálat diakrón jellege és perspektívája azonban ezeket a kétségeket eloszlatja. A nyomtatott könyv ugyanis új kommunikációs közeget jelent, és tudjuk, hogy később ennek a nyelvhasználati formának saját normái alakultak ki, autonóm lektussá vált. Ezek a normák lassú egységesülési folyamat eredményeképpen kristályosodtak ki, amelynek során az egyes szerzők, nyomdászok és egyéb literátorok egymás nyelvhasználati jelenségeire tekintettel voltak, és minden bizonnyal az átlagnak jobban megfelelő, azaz gyakoribb megoldásokat részesítették előnyben, az átlagostól jelentősebben eltérő, provinciálisabb formákat pedig kerülték. A követendő minták kiválasztása során bizonyára lényegesebb volt ez utóbbi kritérium (az átlagnak való megfelelés), mint a könyv kora vagy egyéb külső szempont. Ha a szerző normatívnak gondolta, akkor valószínűleg elfogadta akár egy félszázados szöveg megoldásait (és gyakran meg is eshetett, hogy ilyen döntéseket kellett hozni, pl. egyes régebbi könyvek újrakiadásakor), míg ha nem gondolta annak, akkor egyértelműen elutasíthatta akár egy kortárs szöveg alakjait is. Valószínűleg az sem volt lényeges szempont, hogy egy-egy szöveg szerzője honnan származott, sok esetben ezt nyilvánvalóan nem is tudhatták az olvasói. A leglényegesebb mozzanat a kommunikációs közeg azonossága lehetett, amely idővel azonos, vagy legalábbis konvergáló nyelvi megoldásokat követelt. Mindezek miatt a nyomtatott könyvek nyelvében megjelenő, azonos funkciót

ellátó alakváltozatokat diakrón vizsgálatban bizvást tarthatjuk egy nyelvi változó különböző értékeinek.

A korszak nyomtatott könyveinek nyelvére egyébiránt leginkább talán a Lanstyák (1998) által bevezetett közvelegesség fogalma illik. Eszerint közveleges az a megnyilatkozás, amelyben több, különböző lektusba tartozó nyelvi megoldás fordul elő (vö. Lanstyák 1998: 14). Szerinte az absztrakciónak tartható nyelvváltozatok átlagától a valódi kommunikációban alkalmazott formák minden esetben több-kevesebb eltérést mutatnak, más nyelvváltozatok jellegzetességeként számon tartott alakokkal keverednek. A korai könyvnyomtatás nyelve minden bizonnyal ilyen volt. Innen vezethetett az út a belső változatosságon át a megszilárduló autonóm lektus irányába (ezekre a fogalmakra nézve l. Trudgill 1997). A sztenderdizáció vagy sztenderdizálódás szempontjából tekintve azonban lényegtelen, hogy még csupán közveleges vagy már valódi, belső változóról beszélünk-e, mivel a normatív alak későbbi kiválasztódása során mindenképpen ezen formák jelentik a választási lehetőségeket.

A következő nehézséget a funkcióazonosság megállapítása jelentette, főként az esetragok alakváltozatainak megállapításakor. Ez több szempontból is problematikus: sokszor kérdéses, hogy egy-egy alak milyen esetben áll, továbbá bizonyos alakok több esetben is besorolhatók. Illusztrációképpen bemutatom az inessivusi és illativusi, valamint a belőlük származtatható esetek alakváltozatainak elkülönítése körüli nehézséget. Az inessivus és az illativus ragja helyhatározói viszonyt fejez ki, másodlagosan azonban kialakultak egyéb, elvontabb viszonyt kifejező funkcióik is. Az elsődleges és másodlagos funkciók közti különbségtételt a vizsgálatban nem tartottam fontosnak, mivel az alakváltozás mindegyikre egyformán jellemző lehet, nem az esetből következik egyik vagy másik alak előfordulása.

A TNyt. a vizsgált ragok használatáról az alábbiakat mondja el. A kései ómagyar korpusz adatainak „31,13%-a az alak és a funkció viszonyát tekintve »szabálytalan«-nak mondható” (Korompay 1992: 397). Ebből a nem hagyományos funkcióban (pl. illativusban) álló *-ban/-ben* az összes előfordulás 13,64%-a, míg az eredetitől eltérő funkcióban (pl. inessivusban) álló *-ba/-be* az összes adat 52,52%-a. Az alak és funkció megfelelésének három fő típusát különíti el: a) a *-ban/-ben* és a *-ba/-be* használatát világosan megkülönböztetőt, b) a két ragot közös *-ba/-be* alakban kiegyenlítő típust, c) a két ragot a *-ban/-ben* javára kiegyenlítő típust (vö. Korompay 1992: 398–9). A korai középmagyar kori nyomtatott írásbeliséget képviselő korpusz adatait is elsősorban abból a szempontból fogtam vallatóra, hogy mintái mely típust testesítik meg. A *-ba/-be*, valamint a *-ban/-ben* ragos alakok összegyűjtése után tehát el kellett különíteni az inessivusi (vagy belőle származtatható), továbbá az illativusi (vagy belőle származtatható) eseteket.

Ez a feladat azonban gondot okozott akkor, ha

- az ige mellett mindkét bővítmény elfogadható, s nem ismerjük a korabeli bővítési lehetőségeket, nem tudjuk, hogy az elvileg szóba jöhetőek közül melyek voltak valóban használatosak, pl.: *eltemettetik vhol/vhová* – „El temettek kegijg *Cōstantinapolijba*” (RMNy. I. 16. P7 r.); *feltámad vhol/vhová* – „Ennec okaert fel nem tamadnac az hitetlenec az *itiletbe*: fem az bűnőfőc az igazaknac *giüleközētőkbe*.” (RMNy. I. 74. A r.–A v.); *szerez* ('létrehoz') *vmit vhol/vhová* – „Es minnyaian à bent Apaftaloc kiuáltképpen chac arra igyekőőstene, hogy berinberte az *orβagokba*, *tartománnyokba*, *várofockba* es *faluakba* plebanofságokat, az az, Lelkipaβtorfágokat, Bereβhtenénc.” (RMNy. I. 154. b2 r.); *írva van vmi vhol/vhová* – „végéhőz vinne a' tőruényec beβédit, mellyec írva valánac a' *kőnywbe*, mellyet meg találta vala a' Helcias pap” (RMNy. I. 208. Yy2 r.) stb.;
- ha az ige mellett mindkét alak előfordul, mint pl. a *száll* ('megszáll vhol' ~ 'beszáll vhová'): „Ő maga czac közel egy kis *házban* Balla” (RMNy. I. 351. K3 r.), de: „*Szēkes féyr várba* mingyaráft Balla” (RMNy. I. 368. E2 r.) – erre nézve l. még alább;
- többféle alaptag bővítményeként is azonosítható az adott alak, pl.: „az adō fel venni mentūnc vala Philep királyhoz *Macedoniában*” (RMNy. I. 324. I3 v.), ahol a *Macedoniában* a *mentünk* igealak inessivusi vagy illativusi bővítménye, valamint a király tartózkodási helyére utaló inessivusi esetben álló főnév egyaránt lehet;
- stb.

Az ilyen kétes eseteknek az automatizált annotálása is nehézséget jelent, legfeljebb statisztikai alapon valószínűsíthető egyik vagy másik bővítmény. Ha feltételezzük, hogy a korai nyomtatott írásbeliség már örökölt bizonyos normákat a kéziratos írásbeliség megszilárdulásának időszakából (vö. Korompay 1988), akkor – ha van ilyen – a szignifikánsan gyakoribb forma lehet a vizsgált nyelvváltozat sajátja. Bizonyos esetekben a korra vonatkozó pótkompetencia kialakításával (vö. Forgács 1994) eldönthetők a hasonló kérdések. Pl. ha ismerjük azt a tényt, hogy a *száll* ige a fenti jelentésben a korabeli szövegekben gyakrabban előfordul a *be* igeekötővel is (ritkábban a *meg* is szerepelhet mellette), akkor valószínűbbnek tarthatjuk az illativusi vonzatot. A *-ba/-be*, *-ban/-ben* vizsgálatok azonban még az utóbbi módszerek alkalmazása után is maradtak funkcionálisan azonosítatlan esetek, ezért összesen 87 alakot – jobb

híján – kihagytam az összevetésből. Ezek lényegesen nem torzíthatták az eredményt, hiszen az egyértelműen besorolható adatok száma több mint 6400 volt.³

A következőként említendő problémakör a korpusz reprezentativitásával kapcsolatos. A rendezésre álló fennmaradt nyomtatványok számát és terjedelmét tekintve az összeállított gyűjtemény mindenben megfelel a minőségi és mennyiségi reprezentativitás követelményeinek (l. fentebb). Kérdés, hogy ez elegendő-e a felvetett problémák megoldásához, az egyes változószabályok feltárásához. Bizonyos esetekben sajnos nemmel kellett válaszolnom.

A ritkább toldalékmorfémák például a korpusz reprezentatív mivolta ellenére nem tanulmányozhatók elegendő számú adatban ahhoz, hogy hiteles képet kapjunk alakváltozataik megoszlásáról. Illusztrációképpen álljon itt a főnévi essivus-modalisi rag zárt és középzárt alakjainak elemzése.

A primer ablativusi *-L*-ből keletkezett állapot- és módhatározói rag a kései ómagyarban felső és középső nyelvállású alakváltozatokból álló változót alkot (*-ul/-ül ~ -ol/-öl*), illetőleg a ragvégi *l* asszociatív nyújtó hatására kialakult hosszú magánhangzós változatok újabb változóértékeként jelentek meg (*-ül/-ül ~ -öl/-öl – vö*. Korompay 1991: 297, 1992: 375). A két szakirodalmi forrás nyelvjárási megoszlásról is tesz említést, de erről közelebbit nem mond. A fenti alakok a korpuszban is megjelennek, azonban a magánhangzók időtartamát még a korai könyvnyomtatás is csak elvétve különbözteti meg következetesen, ezért ez utóbbi változóérték csak korlátozottan vizsgálható.

A zárt alakváltozatok túlsúlyban vannak a korpusz adatai között: „az ű ann'a az Maria *iegeřül* adatott volna Iofefnek” (RMNy. I. 49. A3 r.), „Ki lehet ez földön kegyetlen Farahonal, ifmeg kit büntetet *kegyetlenbűl* nalánal” (RMNy. I. 353. h4 r.), „Az ő neuét *segitfegül* hija” (RMNy. I. 353. B2 r.), „azokath vezzy ftulaydonythya űnnűmmaganak *dichyretwl* ęs *dichyřfegwl*” (RMNy. I. 13. P3 r.), „*Világoffágul* adtalac tegődet az Pogányoknac” (RMNy. I. 334. Q7 r.), „Annac akoron is *nyauallyářful* vagyon dolga” (RMNy. I. 288. O2 r.), „Erről mind *Deakul Magyarul* fokat irtunc” (RMNy. I. 229. E2 v.) stb., összesen 74 mű⁴ 243 adatában.

A középzárt alakok 17 műben (RMNy. I. 13., 64., 77., 80., 125., 151., 156., 160., 194., 207., 264., 308b, 319., 334., 342., 346., 360.) mindössze 31 adattal képviseltetik magukat: „ahkit my magyarol egy haznak mondonk deiakul [!] ees *gđrđgđl*

³ Bár jelen dolgozatnak közvetlenül nem tárgya, a fentebb említett, kései ómagyar kori adatokkal való összehasonlíthatóság érdekében közlöm a vizsgálat százalékos adatait: inessivusi esetben álló *-ba/-be*: 42%, illativusi esetben álló *-ban/-ben*: 28%. Ha tisztában lennénk a korabeli normatudattal, akár azt is mondhatnánk, hogy a hiperkorrekció megerősödésének vagyunk tanúi.

⁴ RMNy. I. 8., 13., 16., 49., 63., 64., 74., 77., 78., 80., 86., 88a, 90., 91., 92., 95., 96., 100., 102., 108., 109., 125., 144., 151., 156., 160., 162., 164., 173., 181., 186., 194., 205., 207., 208., 213., 219., 229., 259., 264., 269., 280., 284., 288., 295., 296., 297., 299., 303., 304/2., 307., 308a, 308b, 311., 319., 320., 322., 324., 334., 337b, 339., 340., 344., 345., 346., 348., 349., 350., 351., 353., 360., 362., 367., 368.

ecclesianak mondatyk” (RMNy. I. 64. B6 r.), „Ez az Iftenec akarattia, hođ *βentől* éllietóc” (RMNy. I. 319. D6 v.), „az parazth *meltatlanol* meg werettetýk” (RMNy. I. 207. Dd4 v.), „Sion hegyéneç adtam Kiralyol az Atya Iften Bol a' Chriftufrol, az en népemneç erős *hadnagyol*, hogy ellenfégtől legyen *otalmol*” (RMNy. I. 160. P2 v.) stb. Ezek egyetlen műben ellenpélda nélküliek: az RMNy. I. 342. két adata mellett nincsenek zárt példák.

A vizsgálat adatai táblázat formájában:

RMNy.-szám	A zárt alakok száma	A középzárt alakok száma
13.	3	3
64.	2	2
77.	3	1
80.	1	1
125.	1	3
151.	2	2
156.	6	1
160.	1	4
194.	5	1
207.	3	2
264.	4	3
308b	3	1
319.	1	2
334.	6	1
342.	0	2
346.	1	1
360.	8	1

1. táblázat

A főnévi essivus-modalisi rag zárt és középzárt alakjainak megoszlása a korpuszban

A középzárt és zárt alakok értékei tehát erősen szóródnak. A kisszámú példa miatt (28 mintában⁵ egyáltalán nincs adat a ragra) nem lehet minden kétséget kizáróan eldönteni, hogy egy-egy mintában mely alakváltozat számít dominánsnak. A rendelkezésre álló adatok alapján négy szövegben szerepelnek a zártaknál valamivel nagyobb arányban (RMNy. I. 125., 160., 319., 342.). További öt mintában azonos vagy közel

⁵ RMNy. I. 15., 17., 98., 101., 154., 155., 158., 169., 171., 192., 218., 220., 222., 233., 243., 246., 260., 289., 290., 294., 304/1., 315., 326., 328., 341., 343., 358., 359.

azonos arányban fordulnak elő (RMNy. I. 13., 64., 151., 207., 264.). Mivel ezen szövegek szerzői az ország különböző területeiről származtak és tevékenykedtek, területi megoszlásra mindezek alapján nem lehet következtetni. A többi független változó elemzésével sem lehet az alakváltozatok megoszlásáról közelebbit mondani: a zárt példák időben egyenletes eloszlást mutatnak, többféle műfajba sorolható szövegekben, fordított és eredeti szövegekben, külföldi és hazai, keleti és nyugati nyomdákban jelennek meg.

Sajnos a nagy adatszámmal jellemezhető toldalékok esetében sem mindig áll rendelkezésünkre elegendő információ ahhoz, hogy a változószabályok megállapíthatók legyenek. Az elativusi, delativusi és ablativusai raghármasság egymástól nem független variálódásának vizsgálata (összesen 4176 adat elemzése) után is le kellett mondanom arról, hogy a változók megoszlásának okairól részletekbe menő leírást tudjak adni.

Benkő (1957) a következőt írja a szóban forgó elemekről: „Míg a *-ból/-ből* és *-ról/-ről* határozóragok hosszú magánhangzója nyíltabb irányú (*ó, ő*) monoftongizáció révén jött létre, addig a mai *-tól/-től* rag – a *tëü ~ töü* 'tő' szó diftongusának zártabb monoftongizációja következtében – jórészt *-tül/-tül* formájúvá alakult ki. Ennek a fejlődésnek megfelelően középkori nyelvemlékeink nagyobb részében a *-ból/-ből*; *-ról/-ről*; *-tül/-tül* ragsort, illetőleg ennek előzményeit találjuk [...]. Egyes nyelvjárástípusokban azonban már a XV. század folyamán analógiás kiegyenlítődésként történt: az *ó, ő*-s változatok lettek uralkodóvá a ragsor minden tagjában [...]. Az *ó, ő*-s kiegyenlítődésként középső és főként keleti nyelvjárástípusainkra volt jellemző. Tiszta *ú, ű*-s kiegyenlítődésként, vagyis *-bül/-bül*; *-rül/-rül*; *-tül/-tül* ragsorra a XV. századból nincs példánk. A következő században azonban ez a folyamat is megindul, elsősorban a nyugati magyar nyelvterületen” (Benkő 1957: 81; a térképmellékleteken a ragsor történeti változásai). Hasonló képet vázol föl Korompay (1988, 1992) is, azzal a különbséggel, hogy az újabb tanulmányában felteszi a *tëü ~ töü* nyíltabb irányú fejlődését is, amelynek eredményeképpen a kései ómagyarban a *-ból/-ből*, *-ról/-ről*, *-tól/-től* kombináció (a TNyt. alapkorpuszában főtípusként) is megjelenik, illetőleg a kései ómagyar korban a keleti-északkeleti vidékek esetében számol egy *-ból/-ből*, *-rül/-rül*, *-tül/-tül* ragsorral (Korompay 1992: 373–5). Vagyis lehetséges, hogy egyrészt bizonyos nyelvjárásokban már a kései ómagyarban a *-ból/-ből*, *-ról/-ről*, *-tól/-től* ragsor a kiinduló helyzet (ez esetben a zárt irányú kiegyenlítődésként természetesen nem lehet beszélni), másrészt a *-ból/-ből*, *-ról/-ről*, *-tül/-tül* kiinduló helyzetből kétféle irányban, és lehet, hogy a nyelvterület több részén is, egymástól függetlenül megindult az analógiás kiegyenlítődésként a kései ómagyar korban: egyfelől a középzárt alakok (a *-ból/-ből*, *-ról/-ről*, *-tól/-től* ragsor), másfelől a zárt alakok (a *-bül/-bül*, *-rül/-rül*, *-tül/-tül* ragsor) kizárólagos használata felé. A két szélső állapot közötti átmeneti vagy ingadozó kombinációk elvileg a következők lehetnek:

A középzárt irányban:

-ból/-ből, -ról/-ről, -tól/-től ~ -túl/-tűl

A zárt irányban:

-ból/-ből, -ról/-ről ~ -rúl/-rűl, -túl/-tűl

-ból/-ből, -rúl/-rűl, -túl/-tűl

-ból/-ből ~ -búl/-bűl, -rúl/-rűl, -túl/-tűl

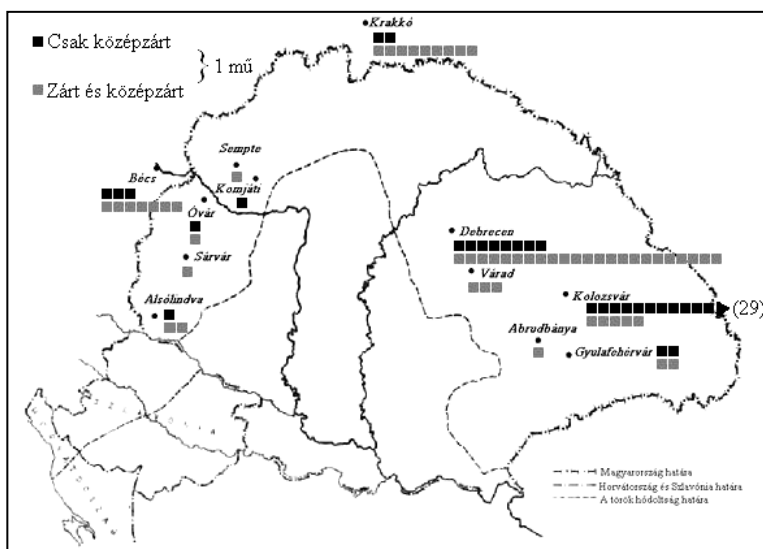
Előfordulhat, hogy egyszerre a ragsor delativusi és ablativusi, illetőleg elativusi és delativusi tagjában, esetleg mindháromban egyszerre van jelen az ingadozás:

-ból/-ből, -ról/-ről ~ -rúl/-rűl, -tól/-től ~ -túl/-tűl

-ból/-ből ~ -búl/-bűl, -ról/-ről ~ -rúl/-rűl, -túl/-tűl

-ból/-ből ~ -búl/-bűl, -ról/-ről ~ -rúl/-rűl, -tól/-től ~ -túl/-tűl

Mivel a hosszú-rövid szembenállás következetes jelöléséről a 16. századi nyomtatványok esetében még nem lehet beszélni, ezért a vizsgálat során a nyíltságra koncentráltam. A nyomtatványokban megjelenő főbb változóértékek a középzárt és a zárt alakváltozatok. 46 műben kizárólag középzárt alakok vannak, azaz a *-ból/-ből, -ról/-ről, -tól/-től* ragsor irányában kiegyenlített típust képviselik.⁶



1. ábra

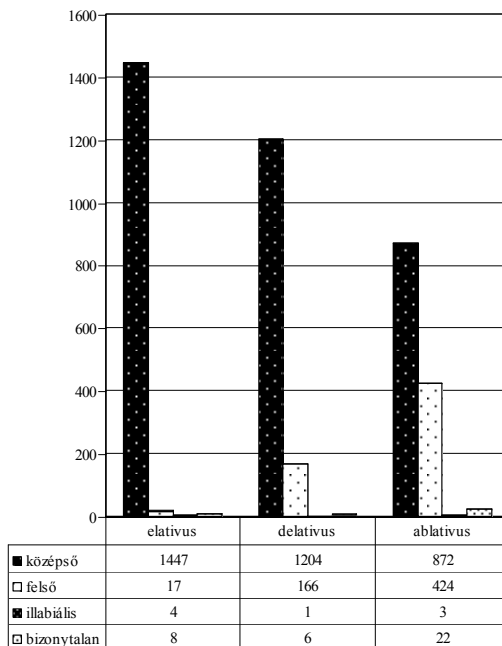
Az elativusi, delativusi és ablativusi eseteket jelölő ragok csak középzárt, ill. zárt alakjait is tartalmazó nyomtatványok megoszlása nyomdák szerint

⁶ RMNy. I. 16., 17., 63., 80., 86., 90., 91., 92., 95., 96., 98., 100., 101., 102., 108., 151., 154., 155., 156., 181., 186., 207., 208., 218., 219., 220., 233., 243., 264., 284., 288., 289., 290., 299., 304/1., 304/2., 307., 315., 319., 324., 326., 337b, 339., 351., 353., 358.

A nyomtatványokat áttekintve látható, hogy 29 a kolozsvári nyomda terméke. A vizsgált időszakban Kolozsvárott megjelent 34 mű zöme tehát ezt a típust képviseli. Meg kell jegyezni, hogy a további öt mű (RMNy. I. 109., 125., 162., 360., 368.) közül csak az RMNy. I. 162. köthető Heltai Gáspár nevéhez, a többi vagy közreműködésének szüneteltetése idején – Hoffgreff György munkája révén – vagy már halála után – a felesége gondozásában – jelent meg Kolozsvárott. A zsoltárfordításokat tartalmazó RMNy. I. 162. valószínűleg nem kizárólag az ő munkája, hanem a bibliafordítás-sorozata részeként munkatársaival közösen fordította. Nem tűnik tehát túlságosan elhamarkodottnak az a következtetés, hogy Heltai a nyomdájában megjelenő művek elativusi, delativusi és ablativusi ragjait egységesen középzárt alakúra formálta.

A kolozsvári nyomtatványokon kívül zömében külföldön (két esetben Bécsben és három esetben Krakkóban), illetőleg a nyelvterület keleti felén (nyolc esetben Debrecenben és két esetben Gyulafehérvárott) megjelentek találhatók. Mindössze három könyv látott napvilágot a Magyar Királyság nyugati felében: RMNy. I. 151., 319. és 353. Az RMNy. I. 151. Huszár Gál prédikációgyűjteménye, amely Óvárott jelent meg. A szerző korábban főként az észekkeleti vidéken tevékenykedett, így nyelvhasználatába bekerülhettek az ottani jellegzetességek is. Az RMNy. I. 153. egy Bornemisza Péter által összeállított énekeskönyv, több szerző műve, ezért egységes nyelvi képet hiába is várnánk tőle. Tehát valójában csak az RMNy. I. 319., azaz Kulcsár György *Az halálra való készölettről rövid tanosság* c. műve jelent kivételt, mert valóban a Nyugat-Dunántúlhoz köthető a szerzője. Kulcsárnak három műve jelent meg a vizsgált időszakban Alsólvándán, Hoffhalter Rudolf nyomdájában. Mindhárom különböző képet mutat az elativusi, delativusi és ablativusi ragok nyelvállását illetően. A másik két műben megjelennek zárt alakok, de nem egyforma arányban. Az RMNy. I. 320-ban csak az ablativusra találhatók zárt példák. Az RMNy. I. 334. Kulcsár prédikációgyűjteménye, azaz eredeti saját szöveg. Méretének megfelelően ebből került a legnagyobb minta a korpuszba: mintegy 8000 szónyi. Ez feltehetőleg híven tükrözheti a szerző nyelvét. Alapvetően a középzárt alakok szignifikáns dominanciája látszik ebben a műben, ám minden esetben állnak ellenpéldák is. A szöveget közelebbről megvizsgálva bebizonyosodik, hogy egy része egyszerűen plágium, korábbi bibliafordítások, prédikációgyűjtemények részleteinek szó szerinti átvétele. Vö. pl.: Székely István 1548-ban: „meg tekintem előzőrt az talent garaft, kit az menei íften en ream bizot, hog' amaz reft bolganac mogá berint el ne afnam, hanem vele keresködnem, es az íftenec bekinec eleibe vforual vinnem” (RMNy. I. 74. +2 r.), Kulcsár 1574-ben: „meg tekintém az talent garaft, kit az mennyei ISTEN en reiam bizot, hogy amaz reft bolganac módgya berint el ne áfnám, hanem vele kereskedném, es az ISTENNEC béki eleibe vforual vinném” (RMNy. I. 334. (:3 r.). A zárt példák a lemásolt szövegekből valók.

A korpusz zárt adatainak megoszlását az alábbi diagram szemlélteti.



2. ábra

Az elativusi, delativusi és ablativusi eset ragjainak változóértékei a korpuszban

A zárt adatok 56 nyomtatványban fordulnak elő. Ezek a három változó megjelenése szempontjából különböző típusokat képviselnek.

Nyolc műben van mindhárom eset ragjának zárt változata is (vagy csak az): RMNy. I. 15., 294., 308b, 322., 328., 334., 346., 349.

A kétféle változóban zárt alakban is (vagy csak abban) adatható, ám ugyanakkor egyben csak középzárt értékkel előforduló kombinációk közül a leggyakoribb (végig az elativus/delativus/ablativus sorrendben, a zárójel azt jelöli, hogy a középzárt alak nem mindenhol váltakozik a zárttal) a csak középzárt/zárt (~ középzárt)/zárt (~ középzárt) 20 műben⁷, ezt követi a zárt (~ középzárt)/csak középzárt/zárt (~ középzárt) egy műben: RMNy. I. 205.

⁷ RMNy. I. 49., 64., 77., 88a, 125., 173., 192., 213., 246., 259., 280., 295., 296., 297., 308a, 340., 344., 348., 350., 367.

A csak egyféle változóban zárt adatot is (vagy csak azt), ugyanakkor a többiben csak középzártat tartalmazó kombinációk közül csak kettő fordul elő: a csak középzárt/csak középzárt/zárt (~ középzárt) 24 műben⁸, és a csak középzárt/zárt (~ középzárt)/csak középzárt három műben: RMNy. I. 78., 162., 342.

A zárt alakokat is tartalmazó művek adatainak részletes vizsgálata során a következő típusokat lehetett elkülöníteni:

1. A *-ból/-ből, -ról/-ről, -tól/-től* típust képviselők (22 mű): RMNy. I. 15., 64., 74., 125., 158., 160., 162., 169., 171., 205., 222., 229., 246., 269., 303., 308b, 328., 334., 342., 348., 360., 367. Ezekben előfordulhat ugyan zárt ellenpélda, de különböző mértékben a középzárt alakok dominálnak. Vagy a csak ablativusban zárt ragsorból fejlődtek, a nyíltabb irányú analógiás kiegyenlítődésként útján, vagy a tiszta középzárt típusból alakultak úgy, hogy ablativusban elkezdtek ingadozni, de a középzárt forma még domináns bennük.

2. A *-ból/-ből, -ról/-ről, -tól/-től ~ -túl/-tül* típust képviselők (7 mű): RMNy. I. 13., 192., 259. 320., 341., 345., 368. Ezekben a *-túl/-tül* változatot tartalmazó ragsor kiegyenlítődése indult meg középzárt irányba, vagy esetleg a tiszta középzárt ragsor ablativusi ragjának a zárt alakváltozattal való váltakozása kezdődött el.

3. A *-ból/-ből, -ról/-ről, -túl/-tül* típust képviselők (12 mű): RMNy. I. 8., 109., 144., 164., 194., 260., 280., 311., 322., 343., 359., 362. Ezek vagy az eredeti, csak ablativusban zárt állapotot képviselik, vagy a tiszta középzárt állapotból kibillenve, ablativusban már zártak.

4. A *-ból/-ből, -ról/-ről ~ -rúl/-rül, -túl/-tül* típust képviselők (3 mű): RMNy. I. 297., 344., 346. Ezek a zárt irányú kiegyenlítődésként útján vannak. A delativus még egyformán előfordul bennük középzárt alakban is.

5. A *-ból/-ből, -rúl/-rül, -túl/-tül* típust képviselők (11 mű): RMNy. I. 49., 77., 78., 88a, 173., 213., 295., 296., 308a, 340., 350. Ezek már a delativusban is dominánsan zártak.

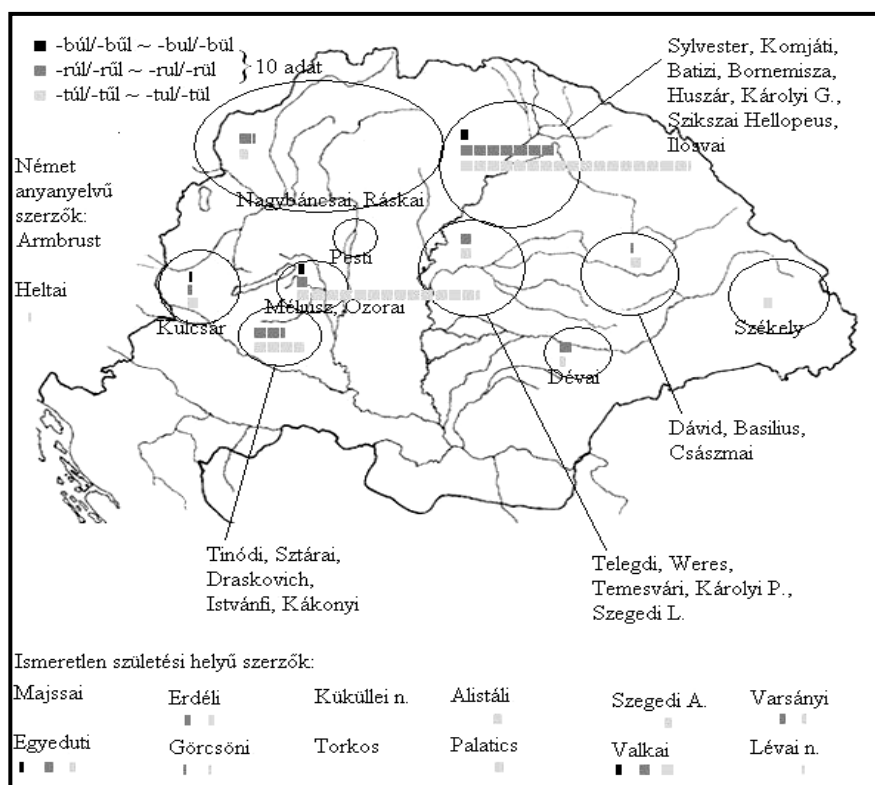
6. A *-ból/-ből ~ -búl/-bül, -rúl/-rül, -túl/-tül* típust képviselő (1 mű): RMNy. I. 349. Ebben már elativusban is megindult az ingadozás, de még tartják magukat a középzárt alakok.

7. A *-búl/-bül, -rúl/-rül, -túl/-tül* típust képviselő (1 mű): RMNy. I. 294. A teljesen zárt változatok dominanciája a kései ómagyar korban még példátlan volt, a XVI. században azonban már nyomtatott könyvben is található rá példa: a Bécsben kiadott kalendárium bizonyos Egyeduti Gergely nevéhez köthető, akiről sajnos életrajzi adatokat nem ismerünk. Molnár (1963) dunántúli származásának tartja, mert az *-ít* képző középzárt alakban szerepel a műben. Ez az adalék is a zárt irányú kiegyenlítődésként nyugati gyökereinek a feltevését erősíti.

⁸ RMNy. I. 8., 13., 74., 109., 144., 158., 160., 164., 169., 171., 194., 222., 229., 260., 269., 311., 320., 341., 343., 345., 359., 360., 362., 368.

Az elativusi zárt alakok sohasem kivétel nélküliek, és csak néhány műre jellemzőek, ezért a továbbiakban csak a delativusi és ablativusi alakokra koncentrálok. Bizonyos szerzőkre jellemző a zárt alakok előnyben részesítése. Nyolc műben (RMNy. I. 88a, 173., 213., 296., 308a, 340., 349., 350.) kivétel nélkül zárt delativusi és ablativusi példák találhatók. Az RMNy. I. 49-ben egyetlen ellenpéldával ugyancsak felső nyelvállású alakok vannak, az RMNy. I. 295-ben nem szignifikáns adatmennyiségben szintén a zárt alakok dominálnak, továbbá két többkönyves szerző is van, akinek mindegyik művében a zártak vannak többségben: Dévai Bíró Mátyás és Sylvester János. A további többkönyves szerzők művei több csoportba sorolhatók, azaz még ugyanazon író sem mindig következetes a vizsgált ragsoport használata tekintetében. Méliusz Juhász Péter adatai pl. négy csoportban szerepelnek: található az első háromba és az ötödikbe tartozók is. Mivel az ő művei több nyomdában, több nyomdász közreműködésével jelentek meg, az adatok értékelésekor óvatossá kell lennünk. Az azonban világosan látszik, hogy korpuszbeli műveinek több mint a fele az első típusba sorolható (RMNy. I. 155., 169., 171., 181., 205., 229.), azaz a középzárt alakokat valószínűleg gyakrabban részesítette előnyben. Batizi András két műve közül az egyik (RMNy. I. 125.) Kolozsvárott, Heltainál jelent meg, ezért tartozik a másiknak ellentmondó, a középzárt alakok dominanciáját mutató típusba, a szerző sajátja azonban az RMNy. I. 88a nyelve lehet, amelyben az elativusi és ablativusi zárt alakok kivétel nélküliek.

A szerzők születési helye alapján (l. 3. ábra) valószínűnek látszik, hogy nem csupán a Benkó (1957) által említett (l. fentebb) nyugati, hanem esetleg bizonyos északkeleti nyelvjárásokban is jellemző lehetett a felső nyelvállás irányába megindult alaki kiegyenlítődés, hiszen Sylvester János és Dévai Bíró Mátyás esetében ez a vidék jön számításba az anyanyelvjárás szempontjából (Dévai korán elkerült Hunyad megyei szülőhelyéről, vö. Molnár 1963: 93), és Batizi András is innen származott, sőt egész életében itt tevékenykedett. Batizi műve (az RMNy. I. 88a) az északkeleti területen is elterjedt zárt alakok valódi bizonyítéka, hisz Sylvester és Dévai műveik megjelenése előtt éveken át munkálkodtak a dunántúli Nádasdy-birtokokon, tehát esetükben talán másodlagos nyelvjárási sajátosságról is lehet beszélni.

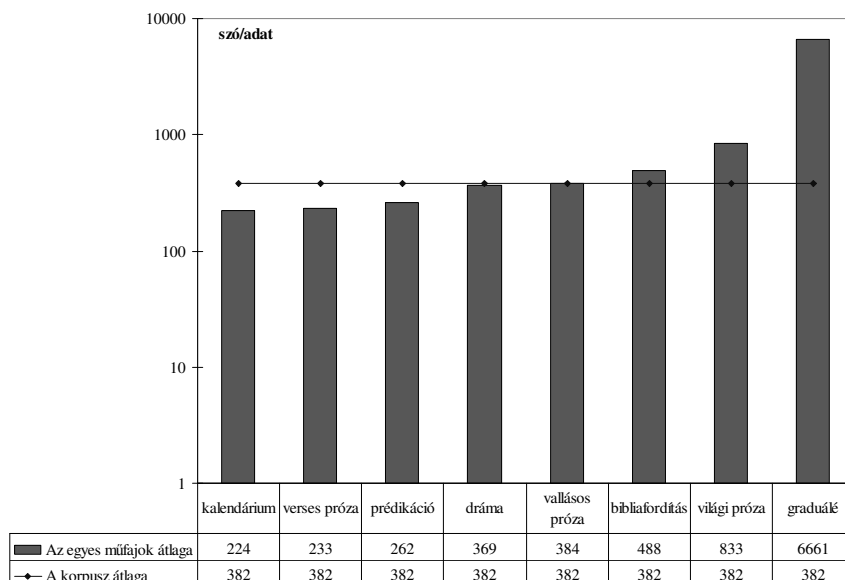


3. ábra

A zárt elatívusi, delatívusi és ablatívusi alakok megoszlása a szerzők születési helye szerint

A következő nyomdákban csak középzárt dominanciájú művek jelentek meg: Alsólintva, Gyulafehérvár, Komjáti, Óvár. Voltaképpen ebbe a csoportba tartozik Kolozsvár is, ahol Tinódi históriásének-gyűjteményén kívül ugyancsak nem jelent meg olyan nyomtatvány, amelyben a három közül bármelyik esetben a zárt alakok szerepelnének nagyobb számban, mivel Heltai ragaszkodott a középzárt változathoz (l. fentebb). Ez utóbbin kívül a többi nyomdában csupán néhány mű jelent meg, ezért a középső nyelvállású alakok gyakorisága azokban is egy-egy szerző vagy nyomdász nyelve miatt alakulhatott így. Vagyis a zárt alakok elterjedtsége nem a megjelenési hely függvénye, ilyeneket tartalmazó nyomtatványokat az ország egész területén nyomtatnak.

Az eredeti magyar nyelvű szövegekben valamelyest nagyobb arányban fordulnak elő a zárt alakok (52%), ám ez nem szignifikáns különbség. Határozottabban különbözik egymástól azonban a különböző műfajú szövegek raghasználata (l. 4. ábra).



4. ábra
A zárt elativusi, delativusi és ablativusi alakok átlagos előfordulása
a különböző műfajú szövegrészekben

Az ábra vízszintes vonala a korpusz átlagát mutatja, amely átlagosan 382 szavanként egy zárt adat előfordulását jelenti. A vonal alatti oszlopok tehát a nagyobb gyakoriságot jelentik, hisz átlagosan kisebb mintában előfordul egy adatot jeleznek. Látható, hogy a kalendáriumok, a verses prózai művek és a prédikációgyűjtemények tartalmaznak az átlagosnál nagyobb számban zárt alakokat. A prédikációk nyelve ekkor is és később is meglehetősen közel állt a – nem is mindig a legválasztékosabb, hanem valamelyest köznapibb, a tömegek számára is könnyebben befogadható – beszélt nyelvhez, hiszen gyakran a templomokban valóban elhangzó szövegek voltak. A zárt alakok kiugróan magas aránya talán ennek a sajátosságnak tulajdonítható. A némileg az átlag feletti értéket mutató verses próza műfajába zömében históriás énekek tartoznak, amelyek ugyancsak élszóban előadott (énekelt) műfajnak számítanak.

A vizsgálat alapján kirajzolódó kép azt mutatja, hogy a könyvnyomtatás nyelvében uralkodó változóérték a középzárt. A zárt alakok dominanciája (a zárt irányú kiegyenlítőedés valamely fázisa) a fentebb felsorolt csoportok közül bizonyosan csak a 4–7-re jellemző, azaz összesen 16 műre. A csak ablativusban zárt típus nem számítható ide, hisz az esetleg az eredetileg is csak ablativusban zárt ragsort tartalmazó 12

művet jelenti (ezek talán a középzárt sorból ablativusban már zárttá fejlődött típust is képviselhetik, bár ez kevésbé valószínű), továbbá a csak ablativusban ingadozó kombinációt tartalmazó hét művet sem, mert azok már a középzárt irányba megindult ki-egyenlítődéss állapotát (esetleg a tiszta középzárt ragsorból az ablativusban már megindult zárt irányú fejlődés korai szakaszát) tükrözik.

A rendelkezésre álló adatok alapján tehát csupán azt lehet biztosra venni, hogy a korabeli Erdélyben inkább a középzárt alakokat kedvelték, és bizonyos, a beszélt nyelvhez közelebb álló műfajokra jellemzőbb volt a zárt alakok előnyben részesítése. Felmerülhet esetleg annak a lehetősége is, hogy a nyelvjárási helyett már ekkor is inkább beszélt nyelvi–írott nyelvi megoszlás jellemző erre a ragsoporra (a korabeli erdélyi nyomtatványok között nagy számban szerepelnek a beszélt nyelvihez közel álló műfajú szövegek). Látható, hogy a hatalmas adatmennyiség ellenére csak meglehetősen óvatos megállapításokat lehet tenni az alakváltozatok megoszlásáról.

A történeti korpuszok építésekor tehát bizonyos esetekben még a megfelelő mennyiségi és minőségi reprezentativitás biztosításával sem lehet elérni, hogy elegendő információknak legyen egyes jelenségek pontos leírásához. Rissanen (1989) már a nyolcvanas évek végén felhívta a figyelmet a diakrón korpuszok használata körüli hibalehetőségekre. Az általa említettek közül a jelen tanulmányban ismertetett vizsgálatban is tapasztaltam a „God’s truth fallacy” [kb. ’az isteni igazság buktatója’] és a „mystery of vanishing reliability” [kb. ’a szertefoszló megbízhatóság rejtélye’] (Rissanen 1989: *passim*.) nehézségeit. Az előbbi azt jelenti, hogy a hatalmas adatmennyiség miatt a kutató hajlamos lehet elhamarkodott következtetések levonására, abban a hitben ringathatja magát, hogy korpuszában minden kérdésre megtalálja a választ. Az utóbbi ezzel részben kapcsolatos: azt jelenti, hogy a gondosan összeállított korpusz végül nem adja meg a választ a kutató által megfogalmazott kérdésekre.

A fent elmondottak ellenére a gyűjtemény sok hasznos információval szolgált, és egy-egy toldalék alakváltozatainak megoszlását elég jól fel lehetett térképezni az összegyűjtött adatok segítségével. Összegzésképpen a vizsgálat legfontosabb tanulságai közül talán azt érdemes kiemelni, hogy a diakrón korpuszok nem is annyira az egyes jelenségkörök teljes, minden részletre kiterjedő leírásához szolgálnak adatokat, hanem elsősorban szubsztanciatesztek elvégzésére alkalmasak. Használatukkal választ kaphatunk azokra a kérdésekre, amelyek egy-egy jelenség meglétével vagy gyakoriságával kapcsolatosak.

HIVATKOZÁSOK

- Benkő Loránd 1957: *Magyar nyelvjárástörténet*, Budapest, Tankönyvkiadó.
- Forgács Tamás 1994: Zárt korpuszok és pótkompetencia, *Néprajz és Nyelvtudomány* **XXXV**, 17–23.
- Kalcsó Gyula 2005: A megszilárdulóban és a visszaszorulóban lévő névszórágok Tinódi Hoffgreff- és Heltai-féle kiadásában, in Zimányi Árpád szerk.: *Acta Academiae Paedagogicae Agriensis. Nova Series Tom. XXXII. Sectio Linguistica Hungarica*, Eger, EKF Líceum Kiadó, 91–104.
- Kalcsó Gyula 2007a: *A névszóinflexió nyelvi változói a magyar nyelvű nyomtatott írásbeliség első fél századában*, doktori értekezés, kézirat, Budapest, ELTE Magyar Nyelvtudományi Doktori Iskola.
- Kalcsó Gyula 2007b: Az elativusi, delativusi és ablativusi ragok vizsgálata a magyar nyelvű nyomtatott írásbeliség első fél századában, *Magyar Nyelv* **103**, 282–300.
- Kiss Jenő 1995: *Társadalom és nyelvhasználat*, Budapest, Nemzeti Tankönyvkiadó.
- Kiss Gabriella – Kiss Margit – Pajzs Júlia 2004: A Nagyszótár történeti korpuszának morfológiai elemzéséről, *Magyar Nyelv* **100**, 185–91.
- Korompay Klára 1988: Nyelvjárási megoszlás és normatív törekvések a kései ómagyar kor névszórágosításában, in Kiss Jenő – Szűts László szerk.: *A magyar nyelv rétegződése II*, Budapest, Akadémiai Kiadó, 605–10.
- Korompay Klára 1991: A névszórágosítás, in Benkő Loránd főszerk.: *A magyar nyelv történeti nyelvtana I/1: A korai ómagyar kor*, Budapest, Akadémiai Kiadó, 284–318.
- Korompay Klára 1992: A névszórágosítás, in Benkő Loránd főszerk.: *A magyar nyelv történeti nyelvtana II/1: A kései ómagyar kor. Morfematika*, Budapest, Akadémiai Kiadó, 355–410.
- Lanstyák István 1998: Közveleg, in *Nyelvünkben – otthon*, Dunaszerdahely, Nap Kiadó, 13–6.
- Molnár József 1963: *A könyvnyomtatás hatása a magyar irodalmi nyelv kialakulására 1527–1576 között*, Budapest, Akadémiai Kiadó.
- Rissanen, Matti 1989: Three problems connected with the use of diachronic corpora, *ICAME Journal* **13**, 16–20.
- Trudgill, Peter 1997: *Bevezetés a nyelv és a társadalom tanulmányozásába*, Szeged, JGYTF Kiadó.