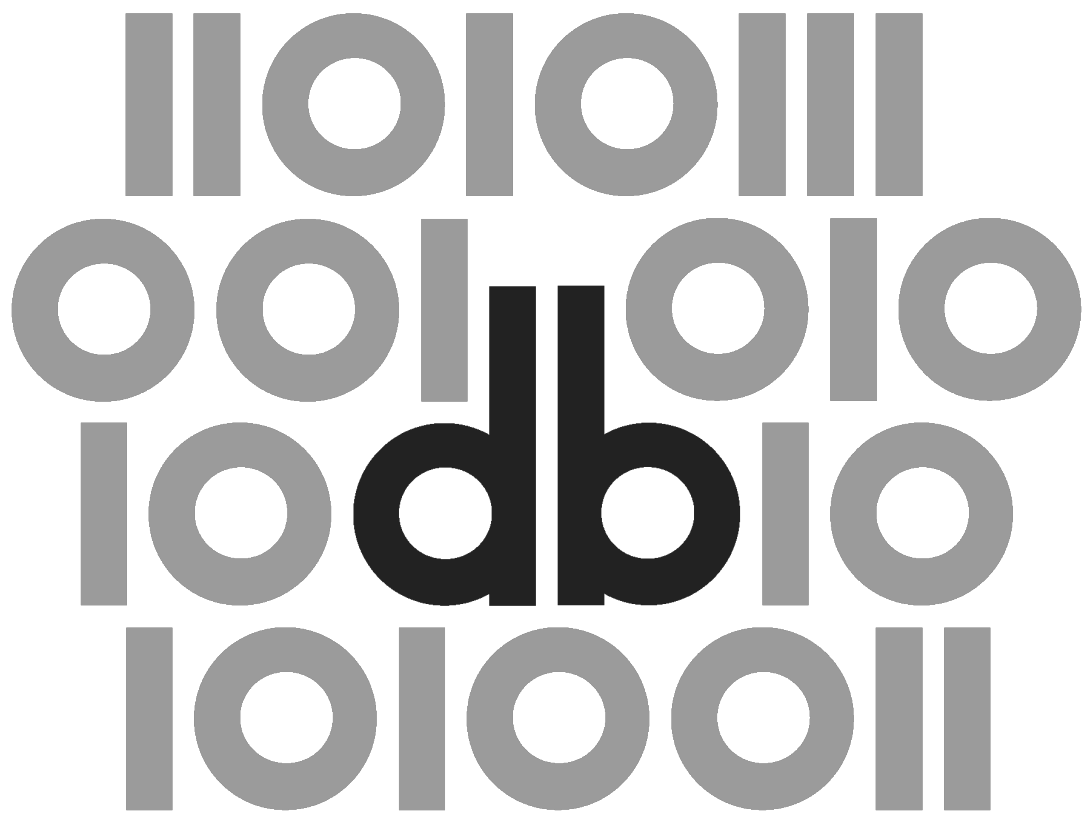


7 (2023)

<DIGITÁLIS BÖLCSÉSZET>



7 (2023)

</DIGITÁLIS BÖLCSÉSZET>

Digitális Bölcsészet
2023., hetedik szám

<DIGITÁLIS BÖLCSÉSZET>



7 (2023)

Felelős szerkesztő:

Maróthy Szilvia

Szerkesztőség:

Kokas Károly, Parádi Andrea

Rovatvezetők:

Tanulmányok: Kiss Margit

Műhely: Péter Róbert

Kritika: Almási Zsolt

Labor: Mártonfi Attila

Tanácsadó testület:

Bartók István, Fazekas István, Golden Dániel, Horváth Iván, Palkó Gábor, Pap Balázs, Sass Bálint, Seláf Levente

Korábbi munkatársaink:

Bartók Zsófia Ágnes (szerkesztő, rovatvezető), Fodor János (szerkesztő),

†Labádi Gergely (szerkesztő, rovatvezető), †Orlovsky Géza (tanácsadó testület)

ISSN 2630-9696

DOI 10.31400/dh-hun.2023.7

Kiadja a Bakonyi Géza Alapítvány és az ELTE BTK Régi Magyar Irodalom Tanszéke (1088 Budapest, Múzeum krt. 4/A).

Felelős kiadó az ELTE BTK Régi Magyar Irodalom Tanszék vezetője.

Megjelenik az Open Journal Systems (OJS) v. 3. platformon, melynek működtetését az ELTE Egyetemi Könyvtár- és Levéltár biztosítja.

Ez a mű a Creative Commons *Nevezd meg! – Ne add el! – Így add tovább! 2.5 Magyarország Licenc* (<http://creativecommons.org/licenses/by-nc-sa/2.5/hu/>) feltételeinek megfelelően felhasználható.

Honlap: <http://ojs.elte.hu/digitalisbolcseszett>

Email cím: dbfolyoirat@gmail.com

Olvasószerkesztő: Bucsecs Katalin

Tördelés: Hegedüs Béla

Grafika: Hegyi Gábor

<MŰHELY>

Ilyefalvi Emese  0000-0001-6677-5843

Eötvös Loránd Tudományegyetem

ilyefalvi.emese@btk.elte.hu

A kiterjesztett olvasás lehetőségei Elemzések a Voyant Tools segítségével: a magyar ráolvasáskorpuszon*

A tanulmányban a Voyant Tools online szövegelemző és adatvizualizációs szoftvercsomag segítségével különböző elemzéseket látunk a magyar nyelvű digitális ráolvasáskorpuszon. A vizsgálatok célja egyrészt különböző mintázatok keresése és azok értékelése, másrészt új típusú kérdések feltétele és a kiterjesztett olvasás lehetőségének megmutatása egy folklorisztikai korpuszon. A tanulmányban a gyűjtési helyszín, gyűjtő és gyűjtési módszertan alapján készülnek el a korpusz újrendezései.

Kulcsszavak:

digitális folklorisztika, folklórszövegek, kiterjesztett olvasás, ráolvasásadatbázis, számítógépes folklorisztika, Voyant Tools



1. Célkitűzés

Tanulmányomban *A magyar nyelvű ráolvasások digitális adatbázisához* előkészített szövegtörzsön végzett néhány elemzés tanulságát foglalom össze. Az adatvezérelt, számítógépes szövegelemzések és adatvizualizációk segítségével készített vizsgálataim elsődleges célja a különböző mintázatok keresése és azok értelmező feldolgozása volt: hogyan és mennyiben támasztják alá a magyar ráolvasáskutatás eddigi megállapításait a nagy léptékű korpusz algoritmikus elemzései? Mindemellett tanulmányomban implicite arra a kérdésre is reflektálok elemzéseimmel, hogy a különböző kiadási eljárások, a szövegek előkészítésének módja és az azok mögött meghúzódó elméleti koncepciók, azaz a textualizációs gyakorlat hogyan befolyásolja az interpretációt. Egyáltalán, a szövegek különböző digitális újrendezéseivel, a kiterjesztett olvasásra

* A tanulmány a 2020 januárjában megvédett *Léptékváltás a folklorisztikában. A számítógépes folklorisztika elméleti, módszertani és technikai kérdése a magyar ráolvasás-adatbázis alapján* című doktori disszertáció (ELTE BTK Irodalomtudományi Doktori Iskola) 6.2. alfejezetének átdolgozott és kiegészített változata. A tanulmány a Nemzeti Kutatási és Innovációs Hivatal – NKFI 132535 támogatásával és az Innovációs és Technológiai Minisztérium ÚNKP-23-4 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

(*enhanced reading*)¹ alkalmas eszközzel milyen új kérdéseket lehet feltenni a jelenlegi korpuszoknak?

2. A magyar nyelvű ráolvasások digitális adatbázisa

Az adatbázis-építés, a szövegek digitális előkészítése 2014 őszén indult meg,² jelen tanulmányban 6037 adatrekord anyagából generált különböző korpuszokat elemztek.³ Az adatbázishoz Excel-táblázatban vitték fel a kutatócsoport segédmunkatársai az egyes rekordok (ráolvasásadatok) metaadatait:

- tipológia (Pócs Éva 2014-ben megjelent tipológiai rendje alapján),⁴
- funkció (milyen célra használták a szöveget: pl. fejfájás gyógyítása, szerelmi mágia stb.),
- rítus-gesztus (a szöveghez kapcsolódó rítus vagy gesztus sztenderd magyar nyelvű leírása),
- forrástípus (szóbeliségből gyűjtött szöveg vagy különböző történeti forrástípus),⁵
- gyűjtés helyszíne,
- gyűjtő/lejegyző,
- adatközlő (neve, neme, életkora és további adatai, amennyiben ismertek),
- adat nyelve,
- adat lelőhelye
- és az eredeti megjelenés hivatkozása.⁶

A tabuláris megoldás mellett a viszonylag sok metaadattal rendelkező, ugyanakkor igen rövid és egyszerű szövegek miatt döntöttünk. A digitális előkészítés során a konkrét ráolvasásrekord is egy Excel-cellába került, a metaadatoktól elkülönítve, viszont az

¹ A kiterjesztett olvasáshoz lásd Geoffrey Rockwell and Stéfan Sinclair, *Hermeneutica: Computer-Assisted Interpretation in the Humanities* (Cambridge, MA: The MIT Press, 2016), 40. és 23. jegyzet, <https://doi.org/10.7551/mitpress/9522.001.0001>; valamint Stephen Wittek, „Distant Reading Conversion in Early Modernity,” *Journal of Cognitive Historiography* 1–2, 3. sz. (2016): 119–133, 131, <https://doi.org/10.1558/jch.31658>.

² Az adatbázis-építés a Népi vallás a keleti és nyugati kereszténység határán: folyamatosság, változások és kölcsönhatások 324214 sz. EKT-projekt keretén belül indult meg: <http://eastwest.btk.mta.hu/>.

³ Ez valamivel több ráolvasást tartalmaz, de hozzávetőlegesen ennyi szövegről beszélhetünk. A számszerűsítést nehezíti, hogy némely esetben igen nehéz volt meghúzni a határt egy-egy ráolvasás elejénél, végénél, néhol egy-egy adatrekord alatt akár több, 2-3 ráolvasásszöveg is megbújhat. A problémáról lásd részletesebben: Ilyefalvi Emese, „Textualizációs eljárások, tipológiai kísérletek, digitális adatbázisok: Az összehasonlító szövegfolklorisztikai kutatások jövője a ráolvasások tükrében,” *Ethnographia* 128, 3 sz. (2017): 383–415, illetve Ilyefalvi, *Léptékváltás a folklorisztikában*, 161–165.

⁴ Pócs Éva, *Ráolvasások: Gyűjtemény a legújabb korból, 1851–2012*, A magyar folklór szövegvilága, 2/A (Budapest: Balassi Kiadó, 2014).

⁵ Ilyefalvi Emese, *Ráolvasások: Gyűjtemény a történeti forrásokból, 1488–1850*, A magyar folklór szövegvilága, 2/B (Budapest: Balassi Kiadó, 2014).

⁶ A metaadatokról lásd bővebben: Ilyefalvi, *Léptékváltás a folklorisztikában*, 167–174, valamint Ilyefalvi Emese, „A digitalizált magyar ráolvasáskorpusz metaadatainak távoli olvasata”, *Ethnographia* 133, 1. sz. (2022): 163–175.

egy-egy ráolvasásrekordokon belül további kódolást végeztünk, hiszen az nemcsak magát a ráolvasásszöveget tartalmazta. Textualizációs eljárásunkat az a szándék vezérelte, hogy az adatot a lehető legteljesebben, az esetleges kotextussal és kontextussal együtt, egyetlen adatként rögzítsük. Egy adaton belül annotáltuk a címet (< >), a kontextust (* *), a konkrét szöveget ({ }), a gyűjtői kérdést (# #) annak érdekében, hogy ezek később, az elemzések során akár külön-külön is vizsgálhatók legyenek.⁷

Az adatbázis építése jelenleg is folyamatban van, ezért a teljes állomány közzététele még nem lehetséges.⁸ A tanulmányban bemutatott elemzések alapjául szolgáló, 6037 adatrekordból generált különböző korpuszokat azonban elérhetővé tettem részletes dokumentációval. Az adatbázis forrásául szolgáló kötetek, publikációk és különböző kéziratok teljes bibliográfiája is ugyancsak itt érhető el.⁹

3. A *Voyant Tools*

A *Voyant Tools* online szövegelemző és adatvizualizációs szoftvercsomag sok-sok előd és béta-verzió után jelenlegi formájában 2016-ban vált elérhetővé Stéfan Sinclair (McGill University) és Geoffrey Rockwell (University of Alberta) által, kifejezetten azzal a céllal, hogy lehetővé tegyék főként a bölcsészettudományok, társadalomtudományok vagy könyvtártudományok művelői számára, hogy akár komolyabb programozói tudás nélkül is digitális szövegelemzéseket végezhesse, gyakorolják a kiterjesztett olvasás lehetőségét digitális környezetben.¹⁰ A program sokféle fájlformátumot tud kezelni, a felhasználók feltölthetnek dokumentumot, de korpuszokat is készíthetnek, mely által különböző szövegek és korpuszok összehasonlítására is lehetőségük nyílik. A szoftvercsomag több beépített szövegelemző és vizualizációs eszközt foglal magába, ezek a felhasználó igénye szerint a felületen szabadon kombinálhatók, változtathatók és dinamikusan együttesen is használhatók.¹¹ A *Voyant Tools* lehetővé teszi az elkészített vizualizációk többféle exportálását, és támogatja a hibrid publikációs formákat is (nyomtatott és hozzá tartozó online tartalmak, eszközök, interaktív vizualizációk), melyek egyre gyakoribbak a digitális bölcsészettudományi kutatásokban.

⁷ A textualizációs problémákról lásd bővebben: Ilyefalvi, *Léptékváltás a folklórisztikában*, 158–166.

⁸ Mindez a tanulmány eredményeit nem befolyásolja elsőrendűen, mert 2018 augusztusa óta zömében olyan adatok kerültek az adatbázisba, melyek ráolvasásszöveget nem tartalmaznak, csak utalnak a ráolvasásra. Bár a tanulmányban bemutatott elemzések menete követi az Ilyefalvi, *Léptékváltás a folklórisztikában*, 6.2. fejezetét, az elemzések alapjául szolgáló adatsor azóta tovább tisztult, a korpuszok generálását ezért újra elvégeztem, a vizualizációk a jelen tanulmány számára készültek.

⁹ <https://doi.org/10.14755/mtakik.2024.0001>.

¹⁰ A szoftvercsomag megálmodói mind a *Voyant Tools* oldalán (hozzáférés: 2024.02.14, <https://voyant-tools.org/docs/#!/guide/about>), valamint közösen jegyzet kötetükben is számos lehetőséget mutatják meg az eszköz használatának: Rockwell and Sinclair, *Hermeneutica: Computer-Assisted*.

¹¹ Az összesen 24 beépített digitális eszközt és azok leírását, a felhasználásra vonatkozó szemléletes példákkal lásd az alábbi linken, hozzáférés: 2024.02.14, <https://voyant-tools.org/docs/#!/guide/tools>.

4. Néhány elemzés a magyar ráolvasáskorpuszon

4. 1. Szógyakoriság, szófelhő, szóeloszlás

Nagy léptékű szöveghalmazok adatvezérelt elemzésekor a kutatás elején sok esetben irányított kérdésfelvetés nélkül először az a leginkább gondolatébresztő, hogy egy-egy számítógépes program a szövegállomány milyen jellegzetességeire és mintázataira hívja fel a figyelmet.¹² Stephen Ramsay a „*screwing around* hermeneutikájá”-nak nevezte a szövegek felfedezésének effajta játékos módját, s annak a tudományos kutatásokban betöltött fontos szerepéről értekezett.¹³ Mint amikor egy olvasó a könyvtár szabadpolcos kínálatát böngészve céltalanul adja át magát a véletlen rátalálás lehetőségének. Mit tudunk meg a ráolvasáskorpuszról speciális kérdésfelvetés nélkül? Egyszerű szógyakoriság-vizsgálat alapján a korpusz milyen tulajdonságai állapíthatók meg? Az alábbi képen (1. ábra) az összes, dátumhoz köthető teljes ráolvasásadat (ritus, kontextus és szöveg) szógyakoriság alapján generált szófelhőjét láthatjuk.¹⁴



1. ábra. A ráolvasáskorpusz szófelhő-vizualizációja

A vizualizáció a korpusz fontos tulajdonságait emelte ki. Például, hogy az állomány meghatározó része gyógyításhoz köthető szöveg (*gyógyítása*), vagy fontosak a számok (*három*, *háromszor*, *kilenc*). Feltűnő a keresztény vallásossághoz köthető főnevek gyakorisága is (*szent*, *Isten*, *Jézus*, *Mária*, *keresztet*), illetve kiemelkedtek az *árpa* és a *szem* főnevek is. A szövegelemzés első fázisában nem tettem szavakat tiltólistára, hiszen egyes vizsgálatok során a funkciósavak használata is elgondolkodtató

¹² Rockwell and Sinclair, *Hermeneutica: Computer-Assisted*, 282.

¹³ Vö. Stephen Ramsay, „The Hermeneutics of Screwing Around; or What You Do with a Million Books,” in Kevin Kee, *Pastplay: Teaching and Learning History with Technology*, 111–120 (Ann Arbor: University of Michigan Press, 2014). Rockwell and Sinclair, *Hermeneutica: Computer-Assisted*, 158–161.

¹⁴ A ráolvasás-adatbázis készítésekor betűhív átiratra törekedtünk, ezért fontos hogy tudatában legyünk: dialektológiai és helyesírási alakváltozatok miatt lehet, hogy bizonyos szavak csak azért türemkednek ki, mert a korpuszon belül egységesebb a helyesírásuk.

lehet.¹⁵ Esetünkben például beszédes az *akkor* és a *ha* viszonyszók kiemelkedése. Pócs Éva megállapítása szerint ugyanis a magyar ráolvasások esetében a párhuzamos kívánságok, parancsok „szintaktikailag jelölt formái közül a ráolvasó gyakorlatban legnagyobb jelentősége az »úgy..., mint« és »akkor..., amikor«, illetve »ha..., akkor« stb. kötőszavakkal összekapcsolt alárendelő szerkezeteknek van.”¹⁶

Az előző vizsgálatban nem a ráolvasásszövegek leggyakoribb szavait vizualizáltam, hanem a teljes szövegadat-állományt, azaz a teljes ráolvasásgyűjteményt minden elemével, a szöveg címével, esetleg rituális instrukciókkal, adatközlői, lejegyzői kiegészítésekkel. A *gyógyítása* például azért emelkedett ki a szógyakoriság tekintetében, mert ez általános címadási eljárási módja a ráolvasásszöveg-gyűjteményeknek, az adatrekord pedig tartalmazza a kutatói és/vagy adatközlői címet is. Tehát például: *árpa gyógyítása, fejfájás gyógyítása, torokgyík gyógyítása* stb. A megfelelő adat-előkészítésnek köszönhetően azonban könnyedén generálhatunk olyan állományt, amely csak a beszédaktusokat, tehát csak a ráolvasásszövegeket tartalmazza.

A funkciószavak kiszűrése után látható a beszédaktuskorpusz leggyakoribb szava-
inak listája (2. ábra). Az első tízből (*szent, jézus, szem, mária, szűz, isten, szentlélek, krisztus, szép, boldogságos*) egyedül a *szem* az, amely nem köthető a keresztény vallásossághoz, hiszen a *boldogságos* és *szép* valójában Szűz Mária jelzői, aki gyakori szereplője az európai, így a magyar ráolvasásoknak is.

<i>szent</i>	2491
<i>jézus</i>	1023
<i>szem</i>	904
<i>mária</i>	897
<i>szűz</i>	703
<i>isten</i>	687
<i>szentlélek</i>	608
<i>krisztus</i>	603
<i>szép</i>	578

2. ábra. A ráolvasásszöveg-korpusz leggyakoribb szavai, előfordulásuk számával

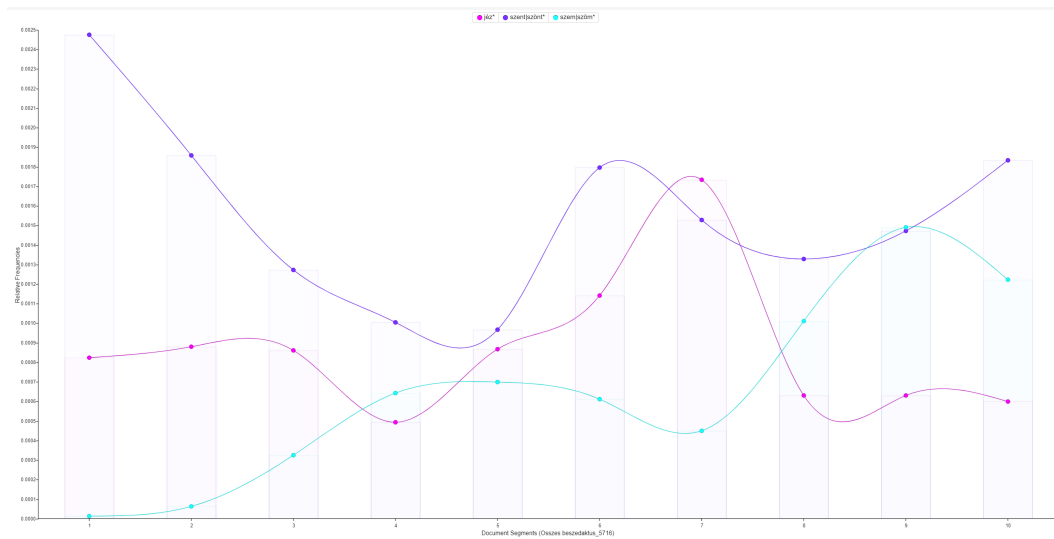
Mivel a korpusz adatait kronológiai sorrendbe állítottam, egy eloszlási diagram segíthet annak feltérképezésében, hogy vajon e szavak relatív szógyakoriság alapján a korpuszon belül milyen eloszlást mutatnak. Az alábbi diagramon (3. ábra) az első három leggyakoribb szó (lehetséges szóalakváltozataikkal együtt vizsgálva *jéz**, *szem*/szöm** és *szent*/szönt**) eloszlása látható.¹⁷ A *szem*/szöm** gyakorlatilag teljesen

¹⁵ Főként stilisztikai vizsgálatokhoz alkalmazzák. Vö. Labádi Gergely, „Az olvasó gép: Berzsenyi Dániel versei távolról,” *Digitális Bölcsészet*, 1. sz. (2018): 17–34, 17, 20–21, <https://doi.org/10.31400/dh-hun.2018.1.126>, valamint Rockwell and Sinclair, *Hermeneutica: Computer-Assisted*, 155. A szógyakoriság, szólisták alkalmazhatóságához lásd Dawn Archer, „Does Frequency Really Matter?” in Dawn Archer, ed. *What’s in a Word-List? Investigating Word Frequency and Keyword Extraction*, Digital Research in the Arts and Humanities, 1–15 (Farnham, England: Ashgate, 2009).

¹⁶ Pócs, *Ráolvasások*, 37.

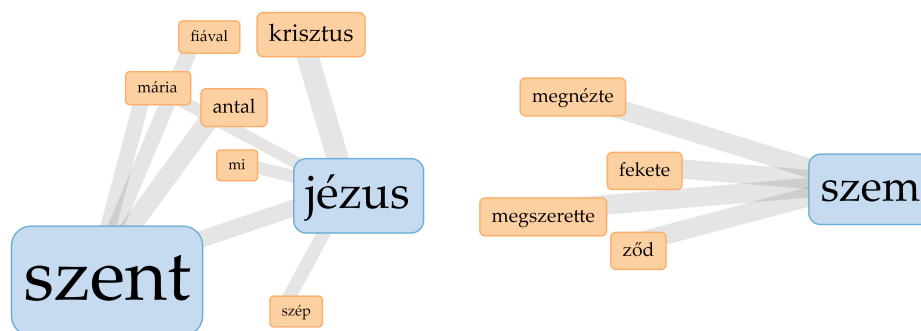
¹⁷ A *Voyant Tools* a szó alakjának többféle keresését teszi lehetővé. A *jéz** paranccsal például az összes *jéz-* kezdetű előfordulást egyetlen találatként értelmezi. Lehetőség van ennek mintájára

hiányzik a korpusz első egytizedéből, s ez arra hívja fel a figyelmet, hogy érdemes a leggyakoribb szavakat különböző időperiódusokban megvizsgálni, mert azok eloszlása közel sem egyenletes.



3. ábra. A ráolvasáskorpusz első három leggyakoribb szóalakjának (*jéz**, *szem/szöm**, *szent/szönt**) korpuszon belüli eloszlása

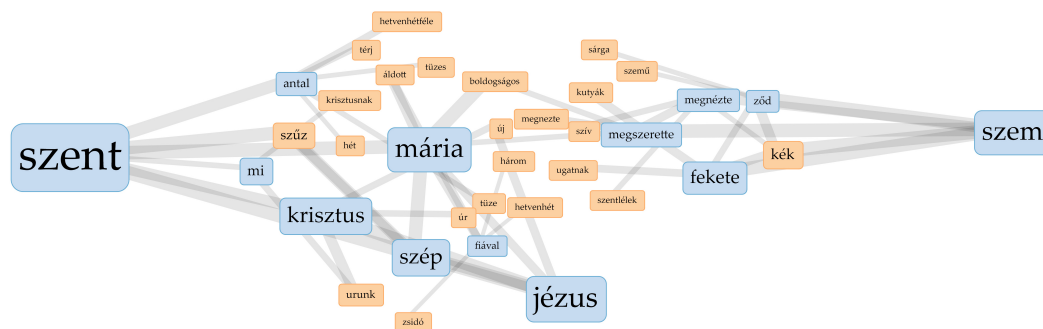
Kérdés persze, hogy a pusztán szógyakorisággal mit állapíthatunk meg a korpuszról. A szólisták és szógyakoriság-vizsgálatok például felhívhatják a figyelmünket olyan szavakra, amelyek jelezhetnek egy-egy témát, de önmagukban nem elegendők témák vagy esetünkben például egy-egy szövegtípus, motívumegyüttes, ráolvasásszöveg vizsgálatához.¹⁸ A leggyakoribb szavak közelében elhelyezkedő további szavak kollokációs hálózati gráfja azonban már a segítségünkre lehet. Az alábbiakban (4–5. ábra) elsőként az eddig vizsgált három leggyakoribb szó (*jézus*, *szent*, *szem*) hálózati gráfját látjuk, majd pedig a *szem* kulcsszó részletesebb gráfját (6. ábra).



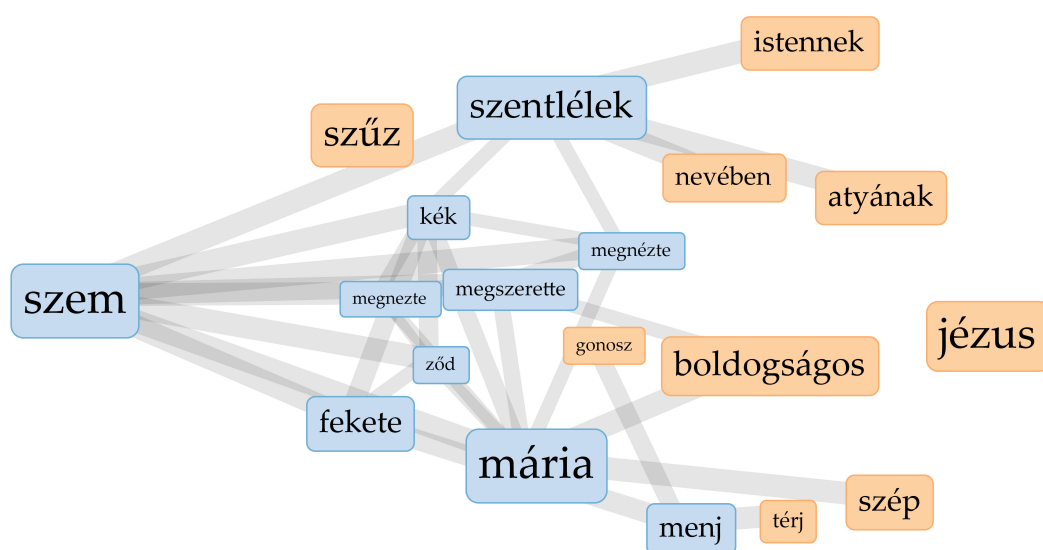
4. ábra. A *jézus*, *szent* és *szem* szóalakok kollokációs hálózati gráfja (3–3 szótávolság)

szóvégződésekre is rákeresni vagy az egymástól öt szó távolságra levő szavakat megtalálni, valamint különböző szavakat összevonni. Az összes lehetséges kombinációról és a jelölésekről lásd részletesen, hozzáférés: 2024.02.14, <https://voyant-tools.org/docs/#!/guide/search>.

¹⁸ Rockwell and Sinclair, *Hermeneutica: Computer-Assisted*, 175.



5. ábra. A *jézus*, *szent* és *szem* szóalakok kollokációs hálózati gráfja (5–5 szótávolság)



6. ábra. A *szem* szóalak kollokációs hálózati gráfja (17–17 szótávolság)

A *szem* környezetében szereplő szavakból (*megszerette*, *megnézte*, *megnézte*, *zöld*, *fekete*, *kék*), pontosabban a szavak együttállásából egyértelműen kirajzolódik egy jól ismert ráolvasástípus. A Pócs Éva tipológiája szerinti XI. főcsoport (az okozók megnevezése + óhaj, parancs, fohász, áldás, átok) szövegei fedezhetők fel a hálózati gráf kulcsszavaiban. A ráolvasástípust a nézéssel való rontás (*szemverés*, *igézés*, *megszólítás*) diagnosztizálására és elhárítására használták, ez a magyar népi kultúrában az egyik leggyakoribb olyan betegségtípus, amelyre ráolvastak.¹⁹ Az első hálózati gráf az egyszerűbb, pusztán az okozók felsorolását tartalmazó szövegeket reprezentálja, a második gráfból azonban már kitűnik a szövegtípus gyakori kontaminációja egyéb szövegekkel, például epikus ráolvasásokkal és/vagy vallásos fohászokkal (*boldogságos*, *mész*, *Mária* stb.), lásd:

¹⁹ Pócs, *Ráolvasások*, 577. A típusról és nemzetközi párhuzamairól lásd bővebben: 577–581. A típus példaszövegeit lásd továbbá: 577–690.

< Igézet gyógyítása > / * A gyógyító „elfűtta a fejit”, és háromszor ráolvasott: * / {Megnézte kék szem, / ződ szem, / fejr szem, / fertelmes szem, / rossz szem. / Boldogságos Szűz Mária szálljon bé fejibe, szemibe, / minden ize-csontjába, / legyen, hogy a Szűz Máriától van hozva születve! [a] / Ptü, ptü, ptü [köpköd]!}²⁰

< Igézés gyógyítása > / * Keresztet vetek a homlokára. Azt mondja: * / {Kék szem megnézte, / fekete szem megigízte, / vigasztaló szép Szűz Mária / vigasztald meg ezt a fehérnépet!} / * Háromszor el kell mondani, s akkor hezzányalintani a nyelvével a homlokához, s eltöpdösni. ^{*21}

< Igizet gyógyítása > / * Ugy beszélnek, hogy az az igizős fehérnép vagy férfi, akit a csicstól elválajszák, s aztán megint megszoptassák. Így hallom, kacagják, olyan temonda beszédbe. / S akkor hazajött, s úgy fájt a feje, hogy alig tudott bójönni, s letette, amit hozott s lefeküdt s mig el nem aludt olvastam rea: * / {Kék szem, ződ szem, tarka szem, sárga szem megnézte, szű megszerette, azokval az igékvel vetek vizet, kivel Jézusnak vetek [a] vót, mikor Jézust megnezték vót. Elküldte Mária édesanyánk Szent Ilona asszont, a Jordán viziből hozzon három csepp vizet. Elment Szent Ilona asszon, hozott három csepp vizet. Jézust avval megmosták, abba a helybe jobban leve. Ugy használjon Annának es, mint Jézuskának ahogy használt vót.} / * s elfujtam s mondtam, högy töpdösse el, csujszon tovább a helyről, ahol olvastam, azért, hogy a betegség es mennyen el róla, arról a helyről. Az éjen aztán magához jött. ^{*22}

< Ima a boldogságos Szűz Máriához > / * Igézés ellen vizvetésre / * {Elindult a boldogságos Szűz Mária egy igen / hosszú útra. Háram zsidó lányt elébe / talala, kitakargaták, megcsókolgaták, menyetek / el Szentek és szentasszonyok vessetek a Jordán / vizeben három szent szentet úgy használjon / az én vizvetésem ezen személynek amint / ahogy használt a boldogságos Szűz Mária / az ő áldott szent Fiának / Szem megnézte, szív megigízte, és a szentlélek / Isten gyógyítsa meg. Miatyánk / Dicsértessék a Jézus Krisztus.}²³

A számítógépes szövegelemzések egyik kihívása, hogy a programok nem tudnak különbséget tenni a poliszemantikus szavak között, ahogy a homonimákkal sem boldognak, hiszen nem jelentést olvasnak, hanem karaktersort. A probléma áthidalására

²⁰ Gajcsána, gyűjtés helye: Egyházaskozár, 1968, Dömötör Tekla és Pócs Éva gyűjtése. Pócs Éva, *Ráolvasások*, 663–664.

²¹ Gyimesközéplek, 1971–1972, Bosnyák Sándor gyűjtése. Bosnyák Sándor, „A gyimesvölgyi magyarok hitvilága,” in Hoppál Mihály, szerk., *Folklór Archívum 14.*, (Budapest: MTA Néprajzi Kutató Csoport, 1982), 120.

²² Gyimesközéplek, 1972, Salamon Anikó gyűjtése. Salamon Anikó, „Gyimesi mondák,” in Kákossy László, szerk., *Mai népi hiedelmek*, ELTE Ökortörténeti Tanszékeinek Kiadványai 10 (Budapest: ELTE, 1975), 138; Pócs, *Ráolvasások*, 859–860.

²³ Székelyudvarhely, 1974–1999, Daczó Árpád gyűjtése. Daczó Árpád, *Hosszú utak megszorodának...: Archaikus népi imádságok, ráolvasók, szentes énekek Erdélyből és Moldvából*, (Budapest: Magyar Napló, 2003), 277; Pócs, *Ráolvasások*, 849.

azonban jó lehetőséget biztosítanak a KWIC (*Keyword in Context*) típusú eszközök.²⁴ A *Voyant Tools*ban például a Context elnevezésű ablakban viszonylag gyorsan áttekinthető, hogy a *szem* vajon mindig az emberi szem jelentésében áll-e korpuszunkban, vagy sem (7–8. ábra).

Left	Term	Right
megyek bűnös Ádám földjére / termő	szem	tördelni, / áldott búza veszteni." / Ez
adj te nekem többet! Vak	szem	élő' légy, / Isten földet teremtett
lélekek nevében. — Föld, édes anyám! [...]	szem	megnézte, szív megverte, sz. háromság
a használjon az egész világnak.	szem	megnézte, szív megverte, szent háromság
Fiú roncsa, / Szentlélek eloszlassa! »Kék	szem	, fekete szem. Kézzel, vízzel lemosom
Szentlélek eloszlassa! »Kék szem, fekete	szem	. Kézzel, vízzel lemosom. Hogyha férfi
hét, ez nem hat, »Kék	szem	, fekete szem, sárga szem, ezer
nem hat, »Kék szem, fekete	szem	, sárga szem, ezer látta, száz
Kék szem, fekete szem, sárga	szem	, ezer látta, száz levágta, Jézus
nevében a kősziklához. Ámmen!" [...] »Kék	szem	igézte, fekete szem igézte, verje
Ámmen!" [...] »Kék szem igézte, fekete	szem	igézte, verje meg az Isten
nevében. Ámen. / Jézus, Mária szerelmiért. /	szem	megverte, szív szerette: / Vetem kalap
szív szerette: / Vetem kalap alá. /	szem	megverte, szív bágyasztja: / Úr jézus
szív bágyasztja: / Úr jézus orvosolja! /	szem	megverte, szív szerette: / Vetem fejkötő
szív szerette: / Vetem fejkötő alá. /	szem	megverte, szív bágyasztja: / Úr jézus
szív bágyasztja: / Úr jézus orvosolja! /	szem	megverte, szív szerette: / Vetem pántlika
szív szerette: / Vetem pántlika alá. /	szem	megverte, szív bágyasztja: / Úr jézus
legyen apád, anyád, minden rokonod! „	szem	Piétér iés Szem Pá, [c
minden rokonod! „Szem Piétér iés	szem	Pá, [c] mon még [d
Ujvesárnop napküötökör eméntem ién Paradicsomkerbe,	szem	mise hagatásáro. Főtekintek aranuómafáro, beleesett
e, mikor alfelemet idé ütöm!" „	szem	Piétér iés Szem Pá, [c
idé ütöm!" „Szem Piétér iés	szem	Pá, [c] mon még [d
Úrjézus bocssássa elébe a gonosztevéket!	szem	látott, / Szív vert meg! / Szent
istem megatta, ámmén. Isten elindult	szem	[a] Pétér után a palláson
mégattad, ámmén. Segics még, isteném,	szem	[a] Márija, szent anyám, hogy
pártájába, ez kis Bandikánknak a	szem	ne ártson." Elindula szűz Mária
alu gyüttie, pártá alá mönny! [...] »	szem	meglátto, szív megverte, Isten lasso
megkereszteltetett! [...] Ellen mondok a ragályos	szem	megnézésének! [a] Ugy ne álljon
igézetek: kékszemű, feketeszemű, sárgaszemű emberek.	szem	megnézte, szív megszerette, szem megigézte
emberek. Szem megnézte, szív megszerette,	szem	megigézte. Én is tiltom tűzzel
az Atya, Fiú és Szentlélekisten.	szem	meglátta, szív megszerette, vigasztald meg
meg, / János verte meg, Kék	szem	látta, barna szem látta, mindenféle
meg, Kék szem látta, barna	szem	látta, mindenféle szem látta. Boldogságos
látta, barna szem látta, mindenféle	szem	látta. Boldogságos Szűz Mária jöjj
férfi, akkor «kalap alá». «Valahány	szem	megnézte, annyi angyal szájjon melléje
meg, / hasadjon ki a csicse	szem	megnezte, / szű megszerette, / vigasztaló szent
szent lélek, / vigasztald meg Pétert.	szem	megnezte, / szű megszerette, / vigasztaló szent

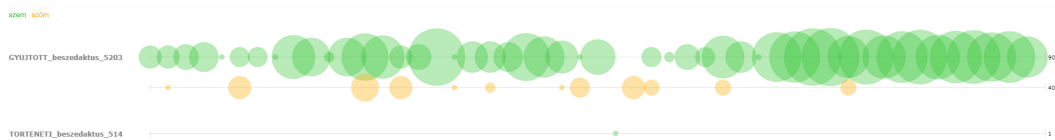
7. ábra. A *Voyant Tools* Contexts elnevezésű eszköze, keresés a *szem* és *szöm* alakokra

²⁴ KWIC típusú eszközök nyomtatott konkordanciákhoz való viszonyához lásd Rockwell and Sinclair, *Hermeneutica: Computer-Assisted*, 45–67.

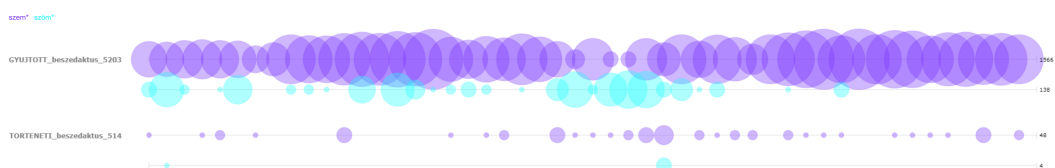
szem* (2536)	szöm* (301)
szem (947)	szöm (56)
szeme (161)	szömverés (19)
szemverés (161)	szömmel (18)
szembe (121)	szömverésbe (12)
szemű (108)	szömit (11)
szemmel (74)	szömibe (9)
szemet (68)	szömmē (8)
szemölcs (67)	szömöm (8)
szemít (36)	szömőcs (8)
szem]	szöm]

8. ábra. A Voyant Tools Contexts elnevezésű eszköze, a keresés közben felugró ablak a találatok számszerűsítésével és alakjaival

A fenti képen látható, hogy bár a *szem** és *szöm** kezdettel más, az elemzés szempontjából irreleváns szavakat (pl. *szömőcs*, *szemölcs*) is előhívhat a számítógépes keresés, esetünkben az adatok száma nem jelentős. Mindenesetre megvizsgálhatjuk a *szem* és *szöm*, de a *szem** és *szöm** kezdetű alakok eloszlását külön-külön is. A kérdés vizsgálatához két alkorpuszra (1. történeti szövegek, azaz minden 1851 előtti szöveg; 2. néprajzi gyűjtés, azaz minden 1851 utáni, folklorisztikai gyűjtéssel a szóbeliségből rögzített szöveg) bontottam a szövegállományt, és a *Bubblelines* elnevezésű eloszlási diagrammal vizsgáltam meg a szóalakok korpuszon belüli jelenlétét (9–10. ábra).



9. ábra. A *szem* és *szöm* korpuszon belüli eloszlása az 1851 előtt és utáni alkorpuszokban



10. ábra. A *szem** és *szöm** korpuszon belüli eloszlása az 1851 előtt és utáni alkorpuszokban

A szóeloszlási elemzés legfőbb eredményeként megerősítést nyert a ráolvasáskutatás korábbi megállapítása, miszerint a történeti, „egyházi vagy laikus írásbeliséghez kötődő forrásanyagban egyetlen adat sem található igézetgyógyító ráolvasásra”, melyből Pócs Éva szerint arra lehet következtetni, hogy a típus kizárólag a szájhagyományozott, népies gyakorlatban létezett.²⁵

²⁵ Pócs, *Ráolvasások*, 579.

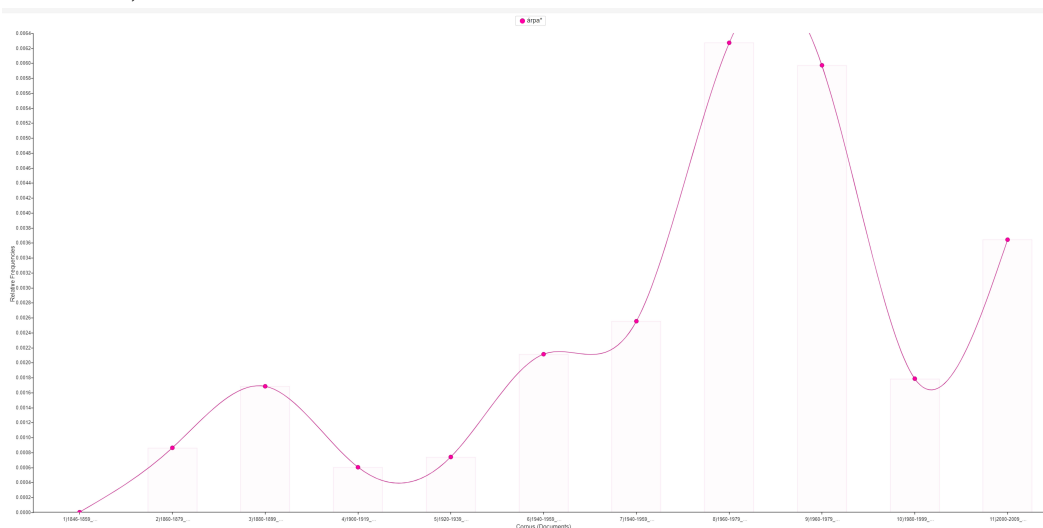
4. 2. Árpagyógyító és szemverés elleni szövegek

A két alkorpusz összehasonlításából kiderül, hogy a gyűjtött, néprajzi szövegek egyik jellegzetes szava az *árpa*, összesen 469 előfordulással. Az összesítés továbbra is egyezik a kutatás korábbi eredményeivel, hiszen Pócs Éva szerint a szemveréshez kötődő és az úgynevezett „számláló, kiolvasó” szövegek mellett a magyar ráolvasások ismert korpuszában az árpagyógyító szövegek csoportja a legjelentősebb, ráadásul a szemveréshez hasonlóan, az árpagyógyító ráolvasások sem ismertek történeti forrásainkban.²⁶

1851 előtti (történeti források)	A szóalak előfordulási száma	1851 utáni (folklorisztikai gyűjtés)	A szóalak előfordulási száma
<i>szent</i>	1823	<i>szent</i>	668
<i>szem</i>	903	<i>minden</i>	265
<i>mária</i>	854	<i>istennek</i>	220
<i>jézus</i>	840	<i>jézus</i>	183
<i>szűz</i>	631	<i>isten</i>	151
<i>szép</i>	554	<i>krisztus</i>	146
<i>boldogság</i>	553	<i>ámen</i>	125
<i>szentlélek</i>	536	<i>x</i>	124
<i>isten</i>	536	<i>szűz</i>	90

11. ábra. Az 1851 előtti és az 1851 utáni alkorpuszok, 10 leggyakoribb szóalak, a szóalakok előfordulásának számával

Ha megvizsgáljuk az *árpa** korpuszon belüli eloszlását, azt láthatjuk, hogy 1960-tól 1979-ig kiemelkedően magas a szó jelenléte még a relatív gyakoriságot tekintve is (11–12. ábra).



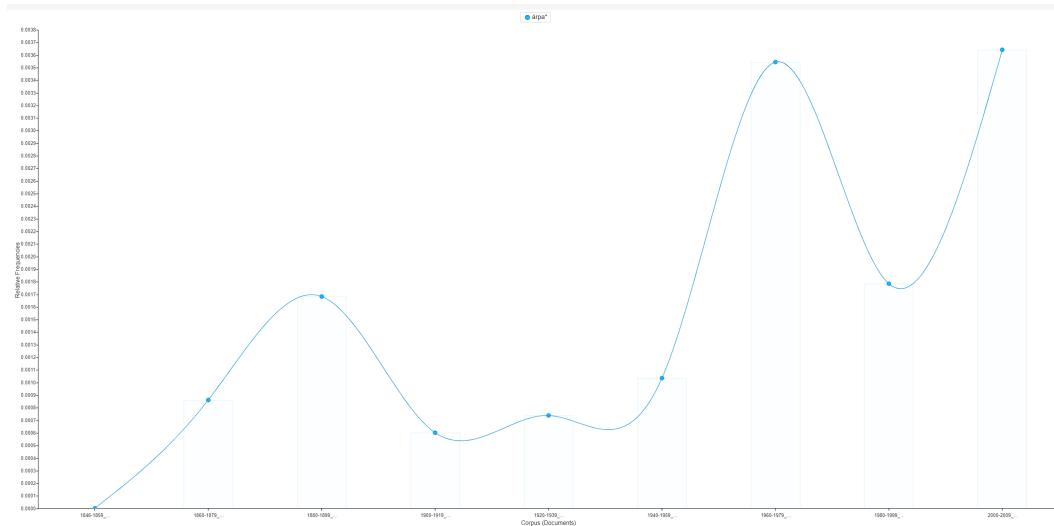
12. ábra. Az *árpa** eloszlása a 20 éves periódusokra bontott gyűjtési korpuszban

Minek köszönhető a kiemelkedés? A korpusz egy korábbi elemzéséből vált egyértelművé, hogy az adatmennyiséget tekintve „fénykorszaknak” számító 1960-as és

²⁶ Uo., 20, 257.

1970-es években a néprajzi atlaszokhoz kötődő gyűjtések aránytalanul torzíthatják a képet egy-egy kérdés esetében.²⁷ A szemén nőtt árpa gyógyításának módjára ráadásul külön kérdés vonatkozott a *Magyar néprajzi atlasz* IV. gyűjtői füzetében.²⁸

Ugyancsak az atlaszos gyűjtésekkel való egyértelmű kapcsolatot fedezhetjük fel akkor, ha gyűjtői korpuszokban vizsgáljuk meg az *árpa* főnév jelenlétét. A tíz legtöbb gyűjtő kutató korpuszát összehasonlítva az *árpa* összesen 181 előfordulással Pócs Éva gyűjtői korpuszában éri el a csúcspontot. Az *árpa** relatív gyakoriságát az atlaszos adatok nélkül megvizsgálva az alábbi eloszlási diagramot kapjuk (13. ábra):



13. ábra. Az *árpa** relatív gyakorisága a korpuszon belül az *Magyar néprajzi atlasz* (továbbiakban: MNA), *Magyar néphit topográfiaja* (továbbiakban: MNT)²⁹ és *Szolnok megye néprajzi atlasza* (továbbiakban: SZMNA) adatai nélkül

Ebből látható, hogy amennyiben az atlaszos adatokat nem vesszük figyelembe, relatív gyakoriságát tekintve a 2000–2009 időperiódusban is legalább olyan jelentős az *árpa** kifejezés előfordulása, mint az 1960 és 1979 közöttiben. A vizsgálat során azonban eddig csak a beszédaktuskorpuszal dolgoztam, viszont ismeretesek olyan árpagyógyító szövegek, amelyekben nem szerepel maga az *árpa* kifejezés, például a betegség tagadó szövegek vagy egyéb egyszerűbb típusú ráolvasások esetében. Az alábbi két példa jól szemlélteti ezt. A kapcsos zárójelben lévő tulajdonképpeni beszédaktus szövegében nem szerepel az *árpa* kifejezés:

²⁷ Ilyefalvi, „A digitalizált magyar,” 166–171.

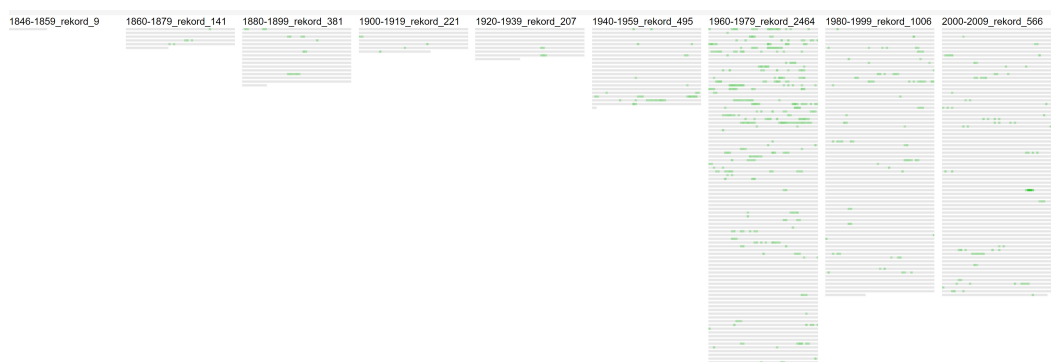
²⁸ 197. kérdés: Barabás Jenő, *A Magyar Néprajzi Atlasz kérdőíve I–IV.* (Budapest: Magyar Tudományos Akadémia, 1958), 110.

²⁹ A *Magyar néphit topográfiaja* című kutatásról, annak kéziratosszerű anyagáról lásd Pócs Éva, „A Magyar Néphit Archívum keletkezéstörténetéhez,” *Ethnographia* 130, 1 sz. (2019): 114–123.

< Árpa gyógyítása > / * Jaj! Hát azt babonáságból, így visszakézbül, így csinálták. Aszongya: * / {Nem az! Nem az! Nem az!} / * Oszt akkor a kezükkel le, a födre csaptak. Mög, hogy akkor a kutat kerülje meg, oszt egy szöm árpát dobjon a kútba. Majd elmúlik a szömirül az árpa. Dehát énnéköm nem múlt el. Csak az orvos gyógyította mög. *³⁰

< Árpa gyógyítása > / * Ha valakinek árpa lett a szemén, késsel „learatták”. Háromszor mondták, utána mindig köptek hármát-négyet: * / {Ma dagadjon, / holnap apadjon, / s holnapután már el se maradjon!}³¹

A vizsgálathoz tehát valójában a teljes ráolvasásadatok szükségesek, ahol szerepelnek a rítusleírások, de ami még fontosabb, szerepelnek benne a címek, azaz a funkciómegnevezések. Az alábbi, *MicroSearch*-vizualizáció (14. ábra) egyrészt szemlélteti a korpuszon belüli alkorpuszok egymáshoz viszonyított relatív arányait, másrészt az alkorpuszokon belül az *árpa** gyakoriságát.³²



14. ábra. Az *árpa** előfordulásai a teljes korpusz 20 éves periódusaira bontva

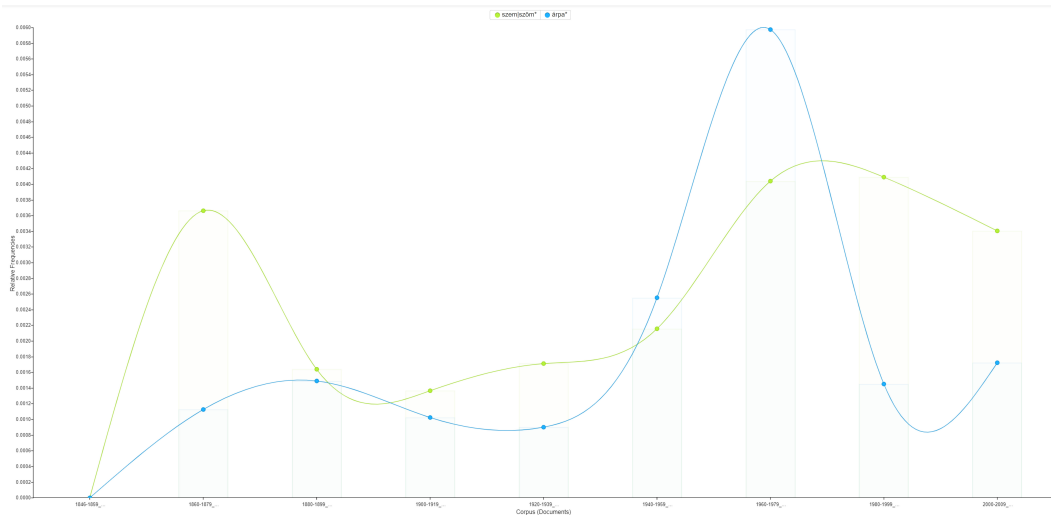
Ugyanakkor – az MNT kérdőívében – nemcsak az árpára kérdeztek rá az 1960-as és 1970-es években, hanem a korábban tárgyalt szemverés elleni szövegekre is.³³ A továbbiakban azt vizsgálom, hogy vajon az árpagyógyító és szemverés elleni szövegek mennyiségére hasonlóan hatottak-e az atlaszos gyűjtések vagy sem.

³⁰ Kiszombor, 1969, Ferenczi Imre gyűjtése. Imre Ferenczi, *Ráolvasások, archaikus népi gyógymódok a Maros mentéről* (Szeged: Szegedi Nyomda, 1977), 338; Pócs, *Ráolvasások*, 187.

³¹ Gyimesközéplok, 2005, Csörge Barnabás gyűjtése. Pócs, *Ráolvasások*, 350.

³² A *MicroSearch* eszköz leírását lásd, hozzáférés: 2024.02.14, <https://voyant-tools.org/docs/#!/guide/microsearch>.

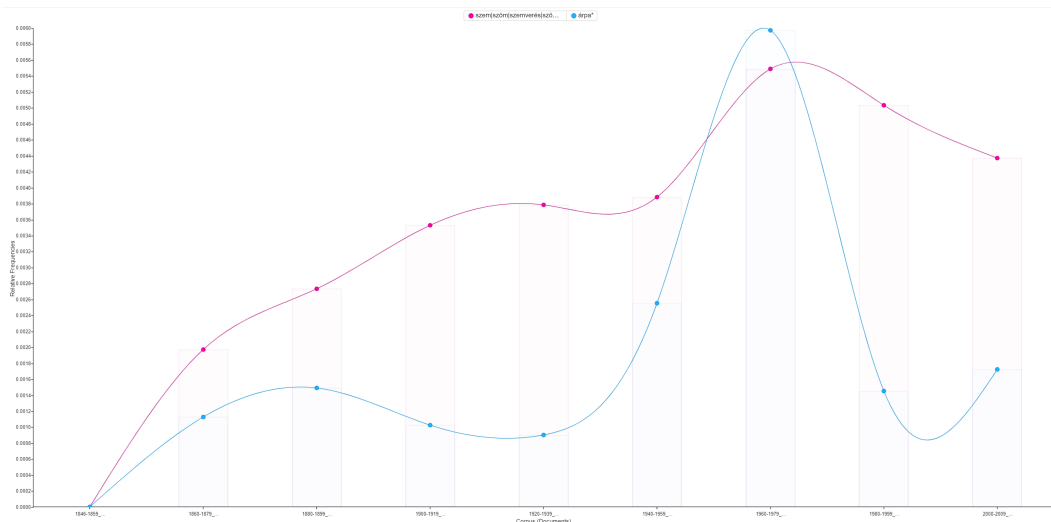
³³ Vö. MNT kérdőive, V.



15. ábra. A ráolvasáskorpusz összes gyűjtési adata 20 éves periódusokra bontva, az *árpa** és a *szem/szöm** alakok eloszlási diagramja

Mindkét diagramból (15–16. ábra) látható, hogy az *árpa** adatok sokkal inkább kiugranak az 1960-as években, mint a *szem/szöm** adatok. Ugyanakkor mivel a *szem/szöm** keresés esetében zavarhatják az eredményt egyéb, a kérdés szempontjából lényegtelen szóalakok (*szemölcs*, *szömölcs* stb.), a vizualizációt elkészítettem az alábbi szóalakok összevonásával is:

szem|szöm|szemverés|szömverés|szeme|szöme|szemmel|szömmel|szemű|
szömű|szemit|szömit.



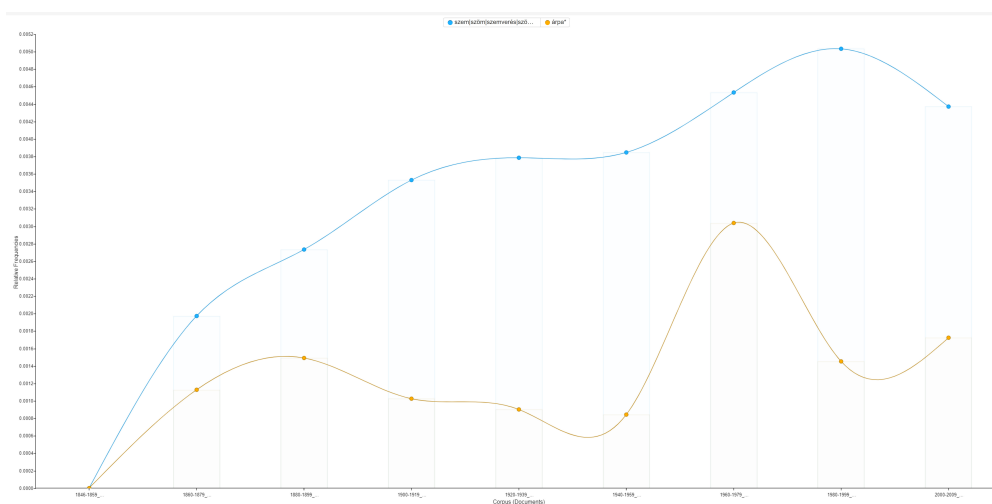
16. ábra. A ráolvasáskorpusz összes gyűjtési adata 20 éves periódusokra bontva, az *árpa** és a *szem/szöm|szemverés|szömverés|szeme|szöme|szemmel|szömmel|szemű|szömű|szemit|szömit* alakok eloszlási diagramja

árpa* (696)
árpa (668)
árpaszemet (10)
árpaszem (2)
árpaszemmel (2)
árpaszömet (2)
árpaszömmel (2)
árpaaratás (1)
árpafödemre (1)
árpagyógyítást (1)
árpa

17. ábra. A Voyant Tools Contexts elnevezésű eszköze, a keresés közben felugró ablak a találatok előfordulásainak számával és alakjaival (*árpa**)

Az *árpa* szerencsére minden esetben az elemzés szempontjából megfelelő jelentésű szóalakokkal szerepel a teljes korpuszon belül (17. ábra).³⁴

Az így kapott vizualizáció (18. ábra) még inkább azt erősíti, hogy a szemveréshez köthető szövegek mennyiségére kevésbé hatottak az 1960-as évek atlaszos kérdőíves gyűjtései.



18. ábra. A ráolvasáskorpusz összes gyűjtési adata az MNA, MNT és SZMNA adatai nélkül, 20 éves periódusokra bontva, az *árpa** és a *szem|szöm|szemverés|szömverés|szeme|szöme|szemmel|szömmel|szemű|szömű|szemit|szömit* alakok eloszlási diagramja

³⁴ Összesen két adat szerepel az adatbázisban moldvai gyűjtésből, ahol az árpa gyógyítására a *korsó* szót használják, az adatok kis száma miatt azonban az elemzésben eltekintek az *árpa** és *korsó** összevont alakok vizsgálatától. A *korsó*ra vonatkozóan lásd Pócs, *Ráolvasások*, 293; Takács György, *Elindula boldogságos szép Szűz Mária: Ráolvasók a régi Csíkszékről*, Fontes ethnologiae Hungaricae 10 (Budapest: L'Harmattan, 2015), 544–545.

Az eloszlási diagramokból (13, 15, 16. ábra) kiindulva két megállapítás fogalmazódhat meg bennünk. Egyrészt az, hogy a szemveréshez kötődő szövegek száma a korpuszban folyamatosan gyarapszik, és bár tetőpontját az 1960 és 1979 közötti időperiódusban éri el, az adatok száma nem csökken jelentősen az atlaszos kutatásokat követően sem. Ezzel szemben az árpagyógyító szövegek esetében azt figyelhetjük meg, hogy tulajdonképpen az atlaszos kutatások helyezték a típust az érdeklődés előterébe; sem az atlaszos kutatások előtt, sem azok után nem gyűjtötték intenzíven ezt a típust, csak az atlaszos munkálatok idejében. Ahogy korábban már többször volt róla szó a dolgozat folyamán, az árpagyógyító szövegek olykor a legegyszerűbb, sok esetben akár trágár kifejezéseket is tartalmazó egymondatos vagy párbeszédesszövegek, melyek országszerte elterjedt és mindenki által ismert szövegek voltak. A korpusz adatai tehát nem azt világítják meg, hogy mennyire voltak ismertek a különböző időperiódusokban az árpához vagy a szemveréshez kötődő szövegek, hanem sokkal inkább egy-egy szövegtípusra vonatkozóan a folklórisztikai érdeklődés hiányára, illetve annak meglétére hívják fel a figyelmet. Úgy tűnik tehát, hogy a túlságosan is közismert árpagyógyító szövegeket kevésbé tartották gyűjtendő és még kevésbé lejegyzendő szövegeknek a kutatók.

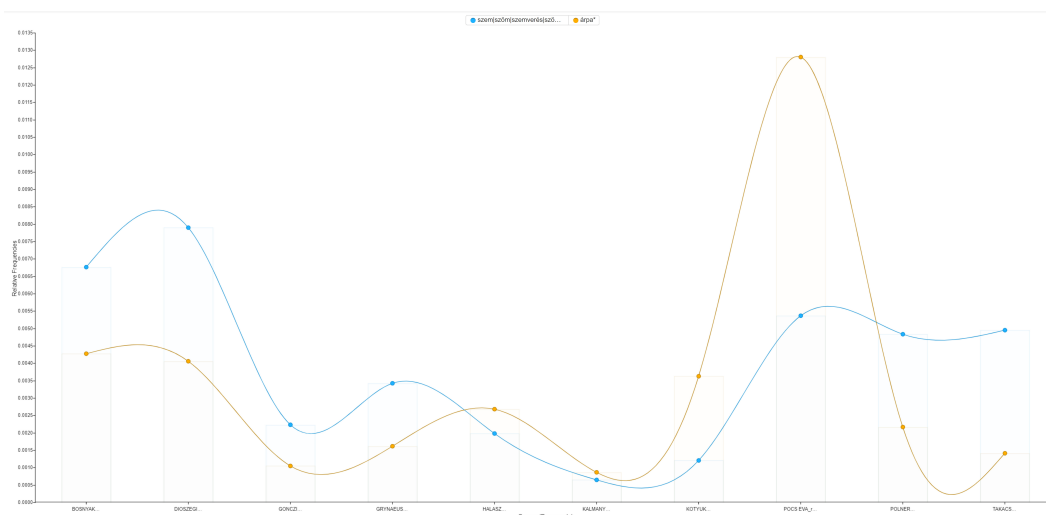
4. 3. Gyűjtői korpuszok

A folklórisztikai korpuszok tekintetében nemcsak a gyűjtési módszertan, a kutatás koncepciója és a tudománytörténeti korszak lehet rendkívül meghatározó, hanem a gyűjtő személye, érdeklődése is, mely szintén befolyásolja a végül lejegyzésre kerülő folklórszöveget.

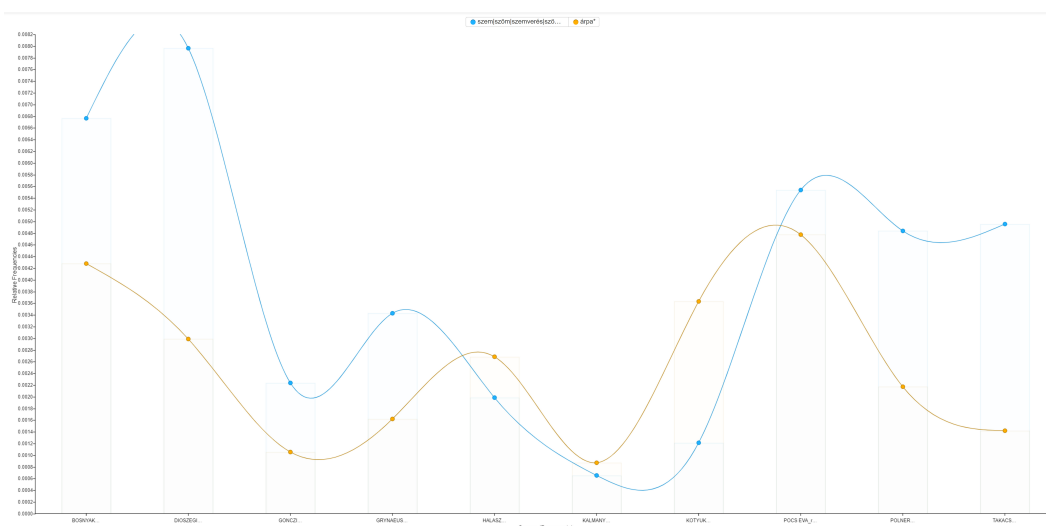
A gyűjtő szerinti szűrés lehetővé teszi, hogy gyűjtői alkorpuszokat hozzunk létre, így megvizsgálhatjuk például azt, hogy a gyűjtők egyéni preferenciája, érdeklődése, műfaji elképzelései, illetve az adott gyűjtési technika hogyan és mennyire érhető tetten korpuszaikon, és milyen viszonyban áll a gyűjtési területükre jellemző szövegállománnyal.

Az előző alfejezet elemzését folytatva a tíz legtöbbet gyűjtő kutató (1. Takács György, 2. Polner Zoltán, 3. Pócs Éva, 4. Diószegi Vilmos, 5. Bosnyák Sándor, 6. Gönczi Ferenc, 7. Halász Péter, 8. Grynaeus Tamás, 9. Kótyuk Erzsébet, 10. Kálmány Lajos)³⁵ alkorpuszaiban az *árpagyógyítás* szövegei és a *szemverés, igézet* szövegei az alábbiak szerint oszlanak meg (19. ábra):

³⁵ Lásd részletesebben: Ilyefalvi, „A digitalizált magyar,” 170.



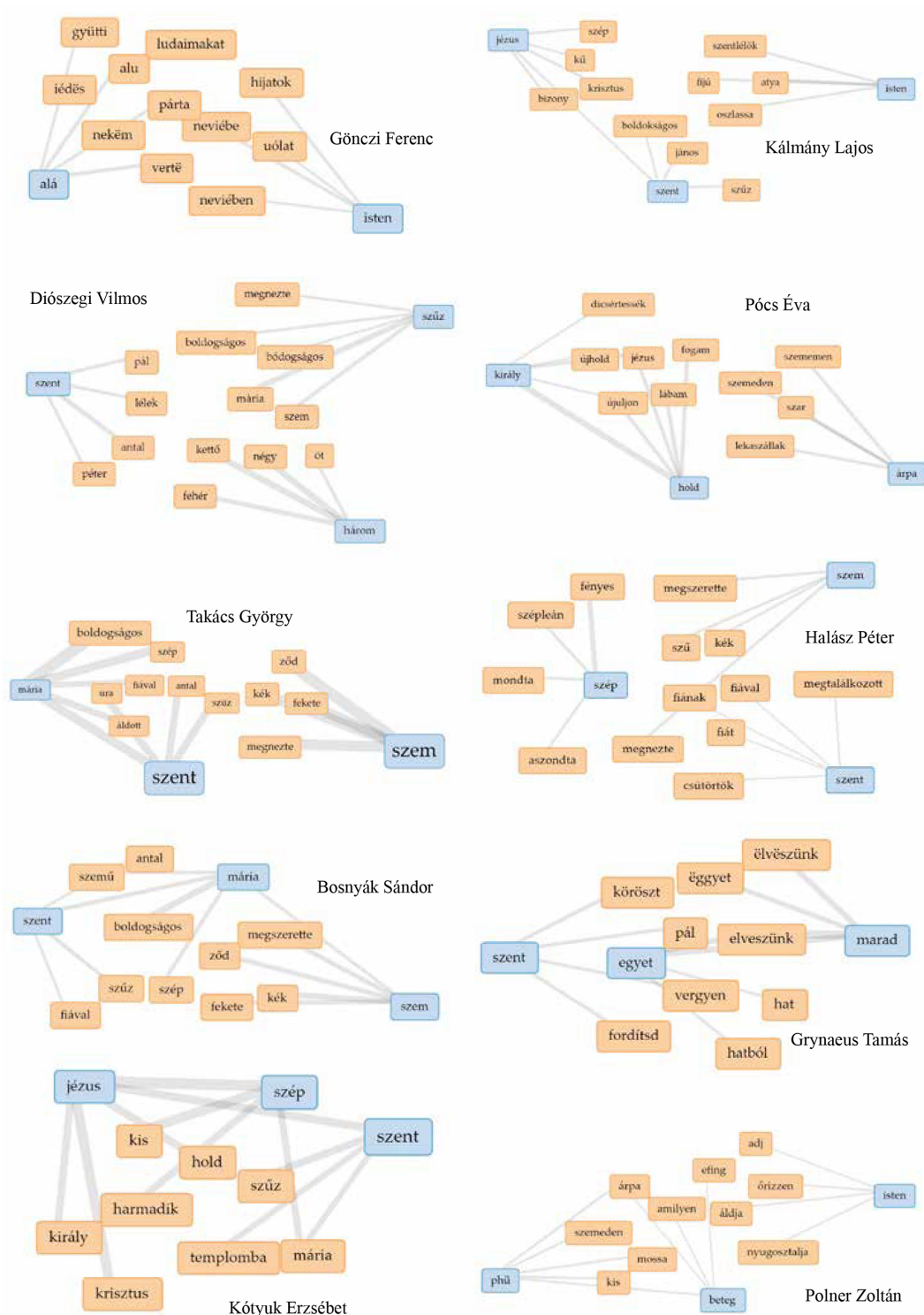
19. ábra. Az első tíz legtöbb ráolvasásszöveget gyűjtő kutató teljes adatainak korpusza, az *árpa** és a *szem/szöm/szemverés/szömverés/szeme/szöme/szemmel/szömmel/szemű/szömű/szemit/szömit* alakok eloszlási diagramja



20. ábra. Az első tíz legtöbb ráolvasásszöveget gyűjtő kutató teljes adatainak korpusza az MNA, az MNT és az SZMNA adatai nélkül, az *árpa** és a *szem/szöm/szemverés/szömverés/szeme/szöme/szemmel/szömmel/szemű/szömű/szemit/szömit* alakok eloszlási diagramja

A diagramon (19. ábra) jól látható, hogy az árpagyógyító szövegek számát illetően kiemelkedik Pócs Éva és Diószegi Vilmos – hiszen ők részt vettek az atlasz gyűjtőmunkálataiban. Ugyanakkor, összehasonlítva az MNA, az MNT és az SZMNA adatai nélküli diagrammal (20. ábra), az is kitűnik, hogy Pócs Éva az atlaszos gyűjtéseken kívül is figyelt e szövegtípus dokumentálására. A szemveréshez kötődő szövegek leggyakoribb szavainak kiemelkedései pedig egyértelműen összhangban vannak a gyűjtők gyűjtési helyszíneivel. Takács, Halász, Bosnyák, Diószegi ugyanis főként moldvai, bukovinai és gyimesi területeken gyűjtött (Diószegi esetében a Baranyába telepített moldvai és bukovinai csángókról van szó). Pócs Éva szerint a szemverés és a hozzá kapcsolódó

vízvetés rítusa „különösen intenzíven él a [...] magyarság keleti csoportjainál”, ezt jól tükrözik ezek szerint a gyűjtői korpuszok is.³⁶



21. ábra. A gyűjtői szövegtörzsek kollokációs hálózati gráfjai

³⁶ Pócs, *Ráolvasások*, 577–579.

A kérdést vizsgálhatjuk a gyűjtők beszédaktuskorpuszainak leggyakoribb szavaiból generálható kollokációs hálózati gráfokkal is (21. ábra). Az előzőkkel összehangban Takács, Halász, Bosnyák, Diószegi hálózati gráfjaiban kitűnnek a szemveréshez, igézet-hez tartozó szövegek kulcsszavai (*megnézte, kék, ződ, szem*), illetőleg azoknak hosszabb epikus szövegekkel való kapcsolódási pontjai (*boldogságos, Mária, megtalálkozott*).

Diószegi Vilmos esetében ezenkívül még a számláló ráolvasások kiemelkedését figyelhetjük meg, ezekkel Diószegi külön tanulmányokban is foglalkozott.³⁷ Pócs Éva korpuszában a korábban már említett árpagyógyító szövegek mellett az újhoddal kapcsolatos szövegek kulcsszavainak kiemelkedését látjuk, ezt szintén az MNA, az MNT és főként az SZMNA gyűjtéseivel hozhatjuk összefüggésbe, illetve azzal, hogy kutatói pályája során (a magyar ráolvasások tipológiájának megalkotásához is) végig fontosnak tartotta az egyszerűbb ráolvasásformák tanulmányozását.

4. 4. Takács György és a csiki, gyimesi, moldvai és Úz-völgyi adatok

A gyűjtői korpuszok közül messze kiemelkedik Takács György korpusza. Mivel Takács kizárólag csiki, gyimesi, moldvai és Úz-völgyi településeken gyűjtött, adódik a kérdés: gyűjtései által mennyire kapunk más képet a régió szöveghagyományát illetően? A következőkben ezért olyan elemzéseket végzek, amelyekben elsődlegesen azt vizsgálom, miként viszonyulnak egymáshoz az ugyanazonokon a településeken mások által gyűjtött adatok a Takács György által gyűjtött adatokhoz.

Elsőként Takács György teljes korpuszát egy olyan korpuszal hasonlítom össze, amely tartalmazza az összes szöveget azokról a településekről, ahonnan Takács György is gyűjtött.

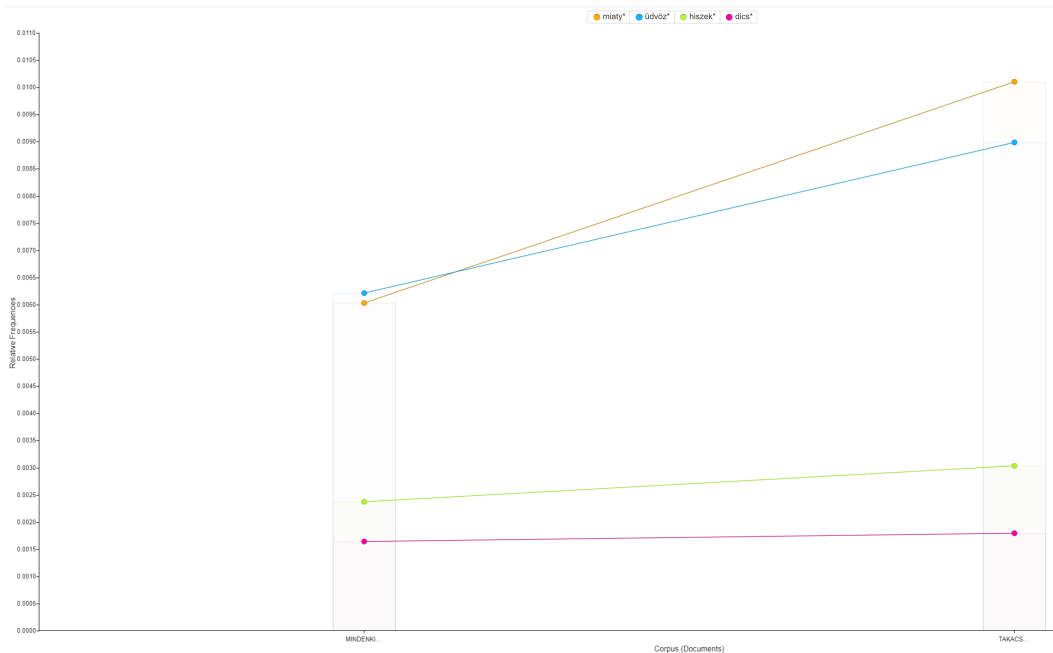
Takács György gyűjtői korpusza	A szóalak előfordulási száma	Takács György gyűjtési helyszínein	A szóalak előfordulási száma
<i>szent</i>	536	<i>szent</i>	318
<i>szem</i>	457	<i>szem</i>	196
<i>mária</i>	263	<i>jó</i>	161
<i>jó</i>	247	<i>fekete</i>	151
<i>szűz</i>	216	<i>mária</i>	136
<i>igízet</i>	215	<i>szűz</i>	130
<i>boldogság</i>	215	<i>boldogságos</i>	111
<i>szentlélek</i>	199	<i>hetvenhétféle</i>	109
<i>gonosz</i>	194	<i>gonosz</i>	106

22. ábra. A Takács György által és az ő gyűjtési helyszínein mások által gyűjtött korpusz összehasonlítása szógyakoriság alapján

A leggyakoribb szavak relatív szógyakoriságát tekintve a két korpusz között nem mutatkozik jelentős eltérés (22. ábra). Ugyanakkor ha a korpuszok rítusleírásait vizsgáljuk, a kanonikus imák előfordulásait tekintve már jelentékeny különbséget lát-

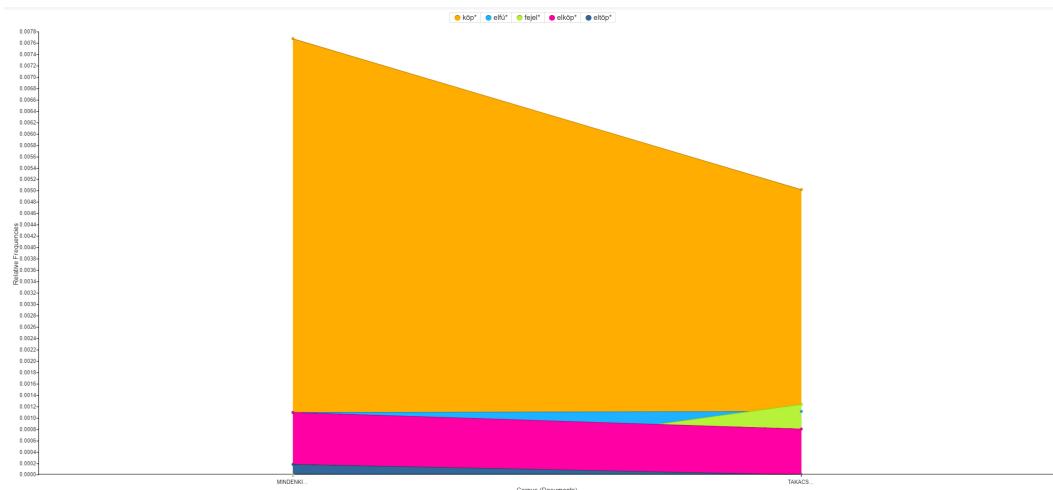
³⁷ Vö. például Diószegi Vilmos, „Visszafelé számláló ráolvasások,” *Néprajzi Közlemények* 2, 3–4 sz. (1957): 162–171.

hatunk. Takács György korpuszában jóval többször fordulnak elő a *Miatyánk* és az *Üdvözlégy* elmondására felszólító adatok (23. ábra):



23. ábra. A *miaty**, *üdvöz**, *hiszek** és *dics** alakok összehasonlítása a két korpuszban

A rítusleírások összehasonlító összegzéséből az is kitűnik, hogy Takács György korpuszában egyik jellemző szava a *fejelfújással*. *Fejelfújásra* az összes többi gyűjtő által lejegyzett korpuszban egyáltalán nincs adat. A *fejelfújás* rítusa ugyanakkor nem jelent mást, mint a szemveréshez kapcsolódó köpést.³⁸ A két korpuszon ezért megvizsgáltam a köpés alábbi lehetséges alakjait is: *köp**, *elfű**, *fejel**, *elköp** és *eltöp** (24. ábra):



24. ábra. A *köp**, *elfű**, *fejel**, *elköp** és *eltöp** alakok a két korpuszban

Ha megvizsgáljuk a *fejelfújás* megjelenéseit a Context segítségével, akkor láthatjuk, hogy azok minden esetben az adatközlők saját, émikus megnevezései a rítusra (25. ábra):

³⁸ Pócs, *Ráolvasások*, 610.

Left	Term	Right
megkapja a szemem, hogy itt	fejel	van ez a közkert ki
csiporba kilencszer ereszteti vizet, féligen	fejel	. S akkor a csipor tetejire
nagymamától tanoltam. * <Igzet gyógyítása> / * A	fejelfúvás	ugy van, hogy: * / {Kék szem
mondják. [...] Csak ezt mondd a	fejelfúvásra	es. [...] Az igizet es biztoson
vigasztald meg!) / * Ezt így mondták	fejelfúvásra	. Keresztet huzott a homlokára, elfujta
nem vettünk vizet. [...] S a	fejelfúvást	pedig úgy vettük, hogy keresztet
feje nagyon. Például ezt a	fejelfúvást	, ezt kisbabáknál szokták leginkább. Me
János vót, ő ittegyen született. * <	fejelfúvás	igézettől> / * Egyszer szépen vizet merek
fejét!) / * „Na, ennyi. Ez a	fejelfúvás	.” * <Zivatar, jégeső elhárítása> / {Az Atyának
tudomásán melyik fejét!) / * Ez a	fejelfúvás	. Vizet es csak így vetnek
nem tartják ezeket. * <Igzet gyógyítása> / *	fejelfúvással	: * / {Szem megnezte, / szív megszerette, / egye

25. ábra. A *köp**, *elfű**, *fejel**, *elköp** és *eltöp** alakok a két korpuszban

A fentiekből egyrészt azt szűrhetjük le, hogy a rítusleírásokat tartalmazó metaadat-sorunkat tovább kell egységesíteni – olyan módon, hogy például a *köpés*re rákeresve a *fejelfúvás* szövegeit is megtalálja a kutató. Másrészt azt a következtetést is levonhatjuk, hogy Takács bizonyára elődeinél jobban figyelte az énikus terminusok megtartására a rítusleírások tekintetében is.

5. Az elemzések tanulságai

Tanulmányomban a digitális magyar ráolvasáskorpuszt elemeztem számítógépes szövegelemző eszközökkel. A szövegelemzések során elsősorban olyan kérdéseket fogalmaztam meg a digitális korpusznak, amelyekre a korábbi nyomtatott szövegkiadások nem nyújtottak lehetőséget. A vizsgálatokban például külön elemzést lehetett lefuttatni a beszédaktusszövegekre és a hozzájuk kötődő rítusleírásokra.

A számítógépes vizsgálatok lehetőséget nyújtanak arra, hogy olyan korábbi tudományos sejtéseket, megállapításokat teszteljünk és kvantifikáljunk, azaz olyan állításokat igazoljunk vagy cáfoljunk, amelyeket (sok esetben intuitíve) már megfogalmazott a szakma, de nem voltak például (szám)adatokkal alátámaszthatók.³⁹

A tanulmányban végzett első, kísérleti elemzésekből többnyire a kutatás korábbi megállapításai igazolódtak vissza. Ugyanakkor számos olyan összefüggés kimutatása is lehetséges volt, amelyekre a digitális adatbázis és a számítógépes elemzés

³⁹ A digitalizált *Magyar néprajzi atlasz* adatain végzett számítógépes klaszteranalízis egyik legnagyobb eredménye Borsos Balázs szerint például éppen az, hogy matematikai módszerekkel igazolni lehetett, hogy a magyar népi kultúra négy nagyobb területi egységre tagolódik, s ezek nagyjából megegyeznek a korábban feltételezettekkel. Borsos Balázs szerint „az ötödik, a számítógép által meghatározott, a földrajzi értelemben vett Alföld északkeleti részére és Erdéllyel határos területeire kiterjedő, Átmenetinek nevezett nagyrégió is egy olyan vonal mentén válik el a Középső nagyrégiótól, amely a néprajzi kutatásokban is ismert volt, de az Alföld földrajzi egysége miatt nem kapott fontosságának megfelelő hangsúlyt”. Borsos Balázs, *A magyar népi kultúra regionális struktúrája: A Magyar Néprajzi Atlasz számítógépes feldolgozása fényében* (Budapest: MTA Néprajzi Kutatóintézet, 2011), 501–502. Korábbi kutatási eredmények alátámasztása, cáfolása számítógépes eszközökkel: Wittek, „Distant Reading Conversion,” 131; Folgert B. Karsdorp, *Retelling Stories: A Computational-Evolutionary Perspective* (PhD thesis, Radboud University, Nijmegen, 2016), 2–3; Labádi, „Az olvasó gép,” 17, 20.

és olvasás nélkül nem lett volna lehetőségünk, sőt a kérdés feltevését is a digitális eszköz implikálta. Sikerült kimutatni például azt, hogy ugyan a szemveréshez és az árpagyógyításhoz kötődő szövegek is érintettek voltak az 1960-as és 1970-es évek atlaszgyűjtéseiben, mindez azonban csak az árpagyógyító szövegek esetében okoz komoly aránytalanságot a korpuszban.

A kimutatások és elemzések az adatbázis félkész állapota miatt egyelőre kezdetlegesek. Az összes adat online közzététele, a tipologizáló munka befejezése, a funkciólista, a kulcsszavak és a digitális térkép használata jóval több elemzésre nyújt majd lehetőséget a jövőben. A szövegkorpusz mélyebb elemezhetősége érdekében előttünk áll még feladatként a szövegek normalizálása és szótövesítése is. Ezt követően pontosabb modellezésre, korpuszelemzésre nyílik majd lehetőség szofisztikáltabb számítógépes eszközök segítségével. A folklorisztika és azon belül a ráolvasáskutatás azon kívánalma, hogy az „előregyártott”, „merek” kutatói kategóriák helyett a számítógépes programok mutassák fel az egyes szövegek közötti összekapcsolódási pontokat és hasonlóságokat, tehát csak egy jövőbeli kutatás által valósulhat meg.

Úgy vélem, hogy a digitális magyar ráolvasáskorpuszon végzett elemzéseim által bebizonyosodott a digitális bölcsészet egyik fontos megállapítása, miszerint a távoli és szoros olvasás nem egymást kirekesztő vagy felváltó, hanem egymást kiegészítő olvasási módozatok, a digitális eszközök pedig nem válaszokat adnak elsősorban, hanem újabb kérdések feltevésére inspirálnak minket. A felhasznált számítógépes eszköz, a *Voyant Tools* a kiterjesztett olvasást (*enhanced reading*) vagy más megfogalmazásban a hiperolvasást (*hyper reading*)⁴⁰ tette lehetővé, biztosítva a folyamatos fókuszváltást a mikro- és a makroperspektíva között.

The Possibilities of Enhanced Reading:

Analysing the Hungarian Verbal Charm Corpus with the Voyant Tools

This study employs the Voyant Tools online text analysis and data visualization software package to conduct various analyses on the Digital Database of Hungarian Verbal Charms. The primary objectives are to identify and evaluate distinct patterns and to propose novel research questions, thereby illustrating the potential for enhanced reading within a folklore corpus. The digital collection is rearranged based on the site of the collections, the collector, and the collection methodology, providing a new framework for analysis

Keywords:

digital folklore studies, computational folklore, enhanced reading, verbal charm database, Voyant Tools

⁴⁰ N. Katherine Hayles, *How We Think: Digital Media and Contemporary Technogenesis* (Chicago, IL: University of Chicago Press, 2012), 12, <https://press.uchicago.edu/ucp/books/book/chicago/H/bo5437533.html>.