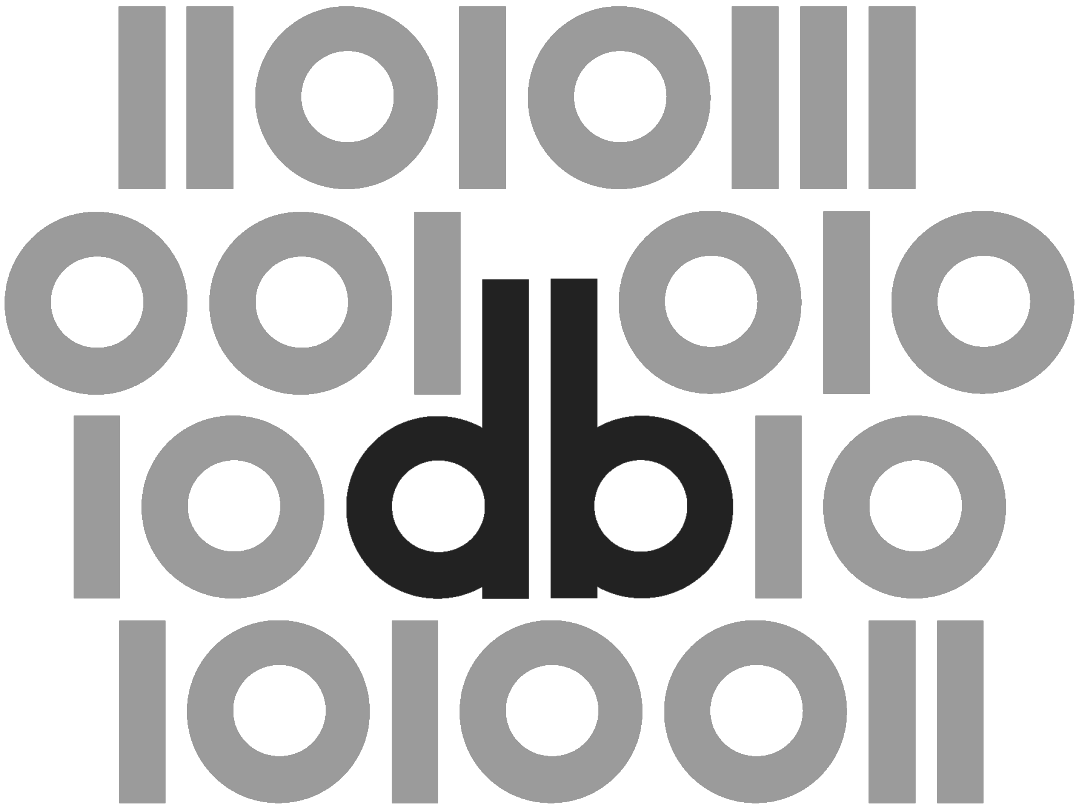
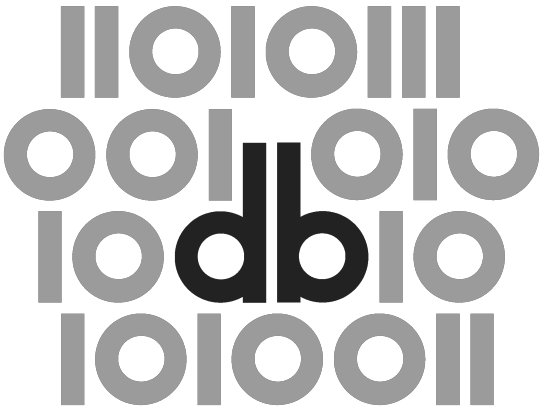
4 (2021) <DIGITÁLIS BÖLCSÉSZET>



4 (2021) </DIGITÁLIS BÖLCSÉSZET>

Digitális Bölcsészet
2021., negyedik szám

<DIGITÁLIS BÖLCSÉSZET>



4 (2021)

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), †Orlovszky Géza (tanácsadó testület)

ISSN 2630-9696
DOI: 10.31400/dh-hun.2021.4

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/digitalisbolcseszet`
Email cím: dbfolyoirat@gmail.com

Olvasószerkesztő: Bucsics Katalin
Tördelés: Hegedüs Béla
Grafika: Hegyi Gábor

<TANULMÁNYOK>

Szemes Botond

ELTE BTK Irodalomtudományi Doktori Iskola

boboszemes@gmail.com

Kidolgozott viszonyok

A tagmondatkapcsolatok automatikus azonosításának hasznosíthatósága a stilisztikában és az irodalomtörténet-írásban*

A tanulmány egy olyan módszer kidolgozására tesz kísérletet, amely lehetővé teszi a magyar írott nyelvben a tagmondatkapcsolatok különböző típusainak automatikus azonosítását a kötőszavak és a relatív névmások mondatban elfoglalt helyzetének elemzésén keresztül. Ez a módszer új távlatokat nyithat meg a stilometriai kutatások számára: a kötőszavak és a névmások egyfelől mint funkciószavak statisztikailag megfelelő mennyiségű adatot szolgáltatnak a szövegek mélystruktúráinak feltérképezéséhez, másfelől tartalmas – grammatikai – jelentést is hordoznak (a tagmondatok között létesülő kapcsolatokról), ezáltal az eloszlásukra vonatkozó eredmények könnyen értelmezhetők. Az egyes típusok relatív gyakoriságának vizsgálatával ugyanis lehetővé válik, hogy meghatározzuk az egy szövegre vagy szövegcsoportra legjellemzőbb/legkevésbé jellemző kapcsolattípusokat. Ennek köszönhetően a regények stílusa és poétikája a mondat szintjén válik megragadhatóvá, miközben a szövegeknek olyan topológiai-logikai karaktere is feltárulhat, amely az olvasás során általában reflektálatlan marad. A tanulmány 100 magyar regény vizsgálatán keresztül tesz kísérletet az egyes szövegek mondatstilisztikai elemzésére, valamint stílustörténeti tendenciák és poétikai hagyományok azonosítására a magyar regényirodalomban.

Kulcsszavak:
tagmondatkapcsolat, kötőszó, stilometria, irodalomtörténet, magyar regény

Bevezetés

A tanulmány 100 kanonikus, 1832 és 2005 között megjelent magyar regény mondatstilisztikai vizsgálatára vállalkozik. Ennek során kvantitatív és algoritmikus eljárásokat

* A kutatásban nyújtott segítségükért köszönettel tartozom a krakkói Computational Stylistic Group és az ELTE BTK Digitális Bölcsészet Tanszék munkatársainak (elsősorban Indig Balázsnak, Szlávich Eszternek és Bajzáth Tímeának); valamint Melhardt Gergőnek. Köszönöm továbbá a folyóirat bírálóinak a korábbi verziókhoz fűzött hasznos megjegyzéseit is!

részesít előnyben, hiszen ezektől a szövegek olyan jellemzőinek a feltárása remélhető, amelyek nem vagy csak kevéssé reflektáltak az olvasás folyamatában. Az alábbiakban ennek megfelelően egy olyan módszert ismeretetek, amely az összetett mondatok különböző típusainak gyakorisági vizsgálatát teszi lehetővé, kiindulópontként szolgálva az irodalomtörténeti, illetve az egyes regények nyelvi szerveződésére vonatkozó megállapítások számára.

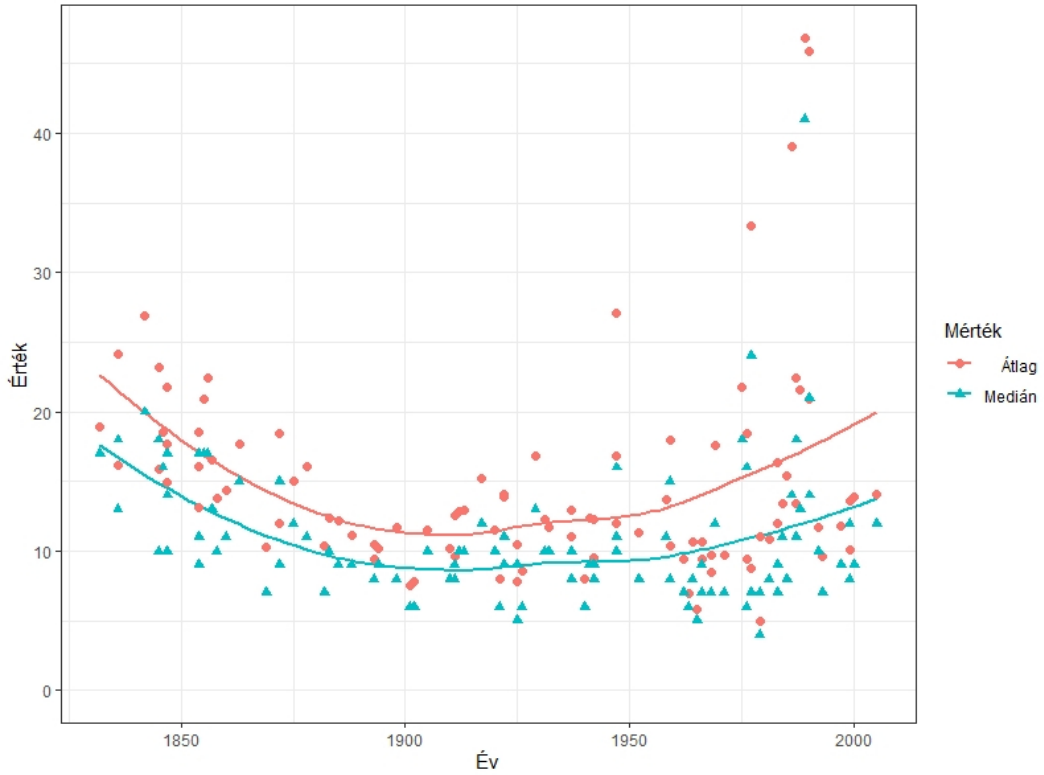
Az ezt megalapozó előzetes kutatás szintén hasonló megközelítést érvényesített, amikor a szövegeket mondatokra, majd a mondatokat szavakra bontva számította ki 100 kanonikus, 1832 és 2005 között keletkezett magyar regény[1] mondathosszúságának átlagát és mediánját. Az eredményeket bemutató ábrákon egyértelmű tendenciák rajzolódnak ki: már ezekben a tendenciákban is tetten érhetők a magyar regény stílustörténetének egyes csomópontjai és főbb változásai. Az *1.1. ábrá*n a korpusz egészét felölelő, az *1.2. ábrá*n pedig Jókai Mór 23 regényét tartalmazó, 1846 és 1894 közötti időszakban láthatjuk a szavak számában meghatározott mondathosszúságok átlagát és mediánját, valamint az adatpontokra illesztett regressziós görbét.[2] A 19. században kirajzolódó csökkenő tendencia – amely Jókai Mór életművében is megfigyelhető – a magyar prózahagyomány stiláris átalakulásán túl a sajtó szerepének megerősödésével, az új íróeszközök elterjedésével, valamint az írás és az olvasás oktatására vonatkozó iskolai reformokkal hozható összefüggésbe.[3] A huszadik század utolsó harmadában kevésbé egységes a megfigyelhető tendencia, inkább a különböző prózapoétikák együttes jelenlétéről beszélhetünk: az extrém értékeket mutató hosszúmondatos szerzők regényei (pl. Krasznahorkai László, Nádas Péter, Kertész Imre, Konrád György egyes művei) mellett kifejezetten rövidmondatos prózák (pl. Mándy Iván, Mészöly Miklós egyes művei) is megtalálhatók a korszakban.

[1] A korpuszok részletes leírását lásd az alábbi, *A korpusz* című részben.

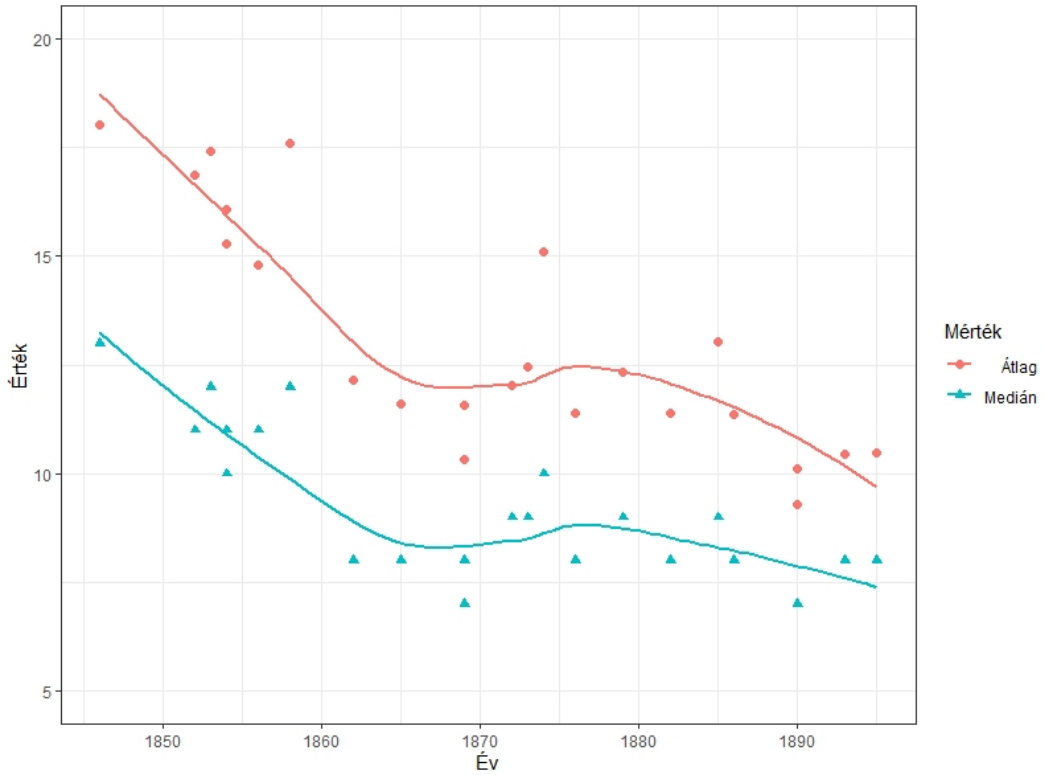
[2] Ez a görbe olyan becsült értékek alapján jön létre, amelyek y valószínűsített helyzetét adják meg x változóhoz (jelen esetben az évszámokhoz) képest úgy, hogy a becsült értékek és a valódi adatpontok közti különbség minimális legyen. Ennek a függvényszerű viszonynak a kiszámítását statisztikai egyenletek segítségével végezhetjük el, amely műveletek az R nyelvhez írt *ggplot* csomagba implementálva a vizualizáció során futnak le. Jelen esetben – a később bemutatott ábrákhoz hasonlóan – a használt kód alapbeállítása a *polinomiális regressziós illesztést* (a kódban: „loess") alkalmazta. Vö. Hadley Wickham, Danielle Navarro and Thomas Lin Pedersen, *ggplot2: Elegant Graphics for Data Analysis* (New York: Springer, 2016), `https://doi.org/10.1007/978-3-319-24277-4`.

[3] A mondathosszúságok vizsgálatát és a tendenciák elemzését lásd: Szemes Botond, „Mondathosszúság és irodalomtörténet: 100 magyar regény szövegstatisztikai elemzése," *Literatura* 46, 3. sz. (2020): 335–367.

 DOI: 10.31400/dh-hun.2021.4.2415



1.1. ábra. Mondathosszúság: átlag és medián, 1832–2005.



1.2. ábra. Mondathosszúság – Jókai Mór. Átlag és medián, 1846–1895.

Már tehát egy ilyen banális szempont (a mondatok átlagos hosszúsága) mentén létrehozott vizualizációk is képesek hírt adni irodalom- és stílustörténeti mozgásokról, és segíthetnek a rövid- és hosszúmondatos prózapoétikák elkülönítésében. Differenciáltabb kérdés azonban csak akkor merülhet fel, ha ezeknek a poétikáknak a hasonlóságaira és különbségeire, valamint a mondatok belső szerveződésére is rákérdezünk, ugyanis nem feltétlenül jelenti ugyanazt a „hosszúmondatosság" különböző alkotók vagy korszakok esetében. Annak érdekében, hogy részletesebb képet kapjunk e 100 regény stílusáról, vizsgálhatjuk többek között a szövegek szókincsgazdagságát vagy a különböző szófajok eloszlását is (pl. nominális és verbális stílusok megkülönböztetésekor),[4] ám jelen dolgozat elsősorban a mondatszerkezetek, pontosabban az összetett mondatok kötőszavakkal kidolgozott, így nyelvileg hozzáférhetővé tett fajtáinak statisztikai elemzésén keresztül kívánja felvázolni a magyar regény stílustörténetének egy szeletét. Azért a mondatstilisztikai szempontot részesítem előnyben, mert kiinduló hipotézisem szerint, amennyiben meg tudjuk adni egy szövegben vagy szövegcsoportban a tagmondatkapcsolatok relatív gyakoriságát, akkor azt is meg tudjuk határozni, hogy mely kapcsolattípusok jellemzők leginkább egy adott szövegre/szövegcsoportra, amin keresztül annak rejtett topológiai-logikai[5] karakterére is következtetni tudunk. Könnyen belátható ugyanis, hogy más annak a szövegnek a karaktere és a világhoz fűződő viszonya, amelyik például arányosan több feltételes tagmondatkapcsolatot dolgoz ki, mint amelyikben inkább az ellentétes kapcsolatok nagy száma figyelhető meg – ahogyan egy kapcsolattípus hiánya is sokat elárulhat a vizsgált műről.

Az ilyen következtetéseknek fontos előképei a magyar stilisztika kvantitatív módszereken nyugvó korábbi (elsősorban 70-es, 80-as évekbeli) munkái, amelyekben szintén erőteljesen érvényesül a mondatokra irányuló elemzési szempont:

> a legjellemzőbb képet a szövegről mondatainak minősítése adja, hiszen a mondatok közlésegységek, meglétük elengedhetetlen szövegsajátság, a szöveg alatt közvetlenül elhelyezkedő szinten állnak. Kvantitatív feldolgozásuk nem jelenti azt, hogy a szövegből kiragadott, attól elszigetelve szemlélt egyedi mondatok egyedi »stilisztikai szerepét« néznénk, hanem éppen egészében tekintjük át a szöveget, különféle mondatainak, mondatszerkezeti formáinak előfordulási arányaiban.[6]

Ezek a munkák bizonyították, hogy az összetett mondatok a szövegek legalapvetőbb típusai, és még a „rövidmondatosként" meghatározott prózastílusban is túlsúlyban

[4] Ilyen irányba tett kísérletekhez vö. Uo., 360–365.

[5] „Topológia" alatt egy hálózat elemeinek elrendezését értem. A „logika" kifejezést nem szigorú értelemben használom, amely így elsősorban nem a mellérendelő kapcsolatok formális logikai megfelelőire vonatkozik, csupán a tagmondatok között kidolgozott viszonyt jelöli. Éppen ezért pontosabb megfogalmazásnak tűnik a mondatok „diagrammatikus" karakteréről beszélni. Vö. Frederik Stjernfelt, „The Extension of Peircean Diagram Category: Charting the Implications of a Diagrammatical Revolution in Semiotic," in Olga Pombo and Alexander Gerner, eds., *Studies in Diagrammatology and Diagram Praxis,* 57–73 (London: College, 2010).

[6] V. Raisz Rózsa, „A kvantitatív módszer mondatstilisztikai alkalmazhatóságáról," *Magyar Nyelvőr* 110, 2. sz. (1986): 161–169, 164. Lásd még Deme László, *Mondatszerkezeti sajátságok gyakorisági vizsgálata* (Budapest: Akadémiai Kiadó, 1971); Nagy Ferenc, *Kvantitatív nyelvészet* (Budapest: Tankönyvkiadó, 1972); Zsilka Tibor, *Statisztika és stilisztika* (Budapest: Akadémiai Kiadó, 1974).

 DOI: 10.31400/dh-hun.2021.4.2415

szerepelnek.[7] Ugyanakkor ezeket a kvantitatív vizsgálatokat főként a stílustörténeti besorolás motiválta (például az impresszionista stílusjegyek felismerése), amely szempont ma kevésbé tűnik termékenynek az irodalomtörténet-írás horizontjából. Innen nézve a tanulmány következtetéseire a legnagyobb hatást a magyar stíluskutatás jelentős, de napjainkban alig idézett alakjának, Herczeg Gyulának a munkássága gyakorolta, aki szintén a mondatszerkezetek változásain keresztül tett kísérletet a magyar próza történetének megírására, és a nyelvi jellegzetességekből érvényesen következtetett egy szerző, vagy korszak világképére is, érdemben hozzájárulva a szövegek értelmezéséhez[8] – ám még digitális eszközök hiányában sokszor kis méretű korpusz rövid szemelvényeire és az olvasói intuícióra hagyatkozva. A mostani eljárás a Herczeg által kidolgozott mondatstilisztikai megközelítést nagy mennyiségű adat statisztikai feldolgozásával és vizualizációjával egészíti ki.

Ugyanakkor meg kell említeni a digitális kereséseken nyugvó kvantitatív elemzések egy sajátosságát is. Egy ilyen elemzés ugyanis más logikát kíván meg, mint a hagyományos mondatstilisztikai megközelítés. Jan Rybicki, napjaink számítógépes stíluskutatásának kiemelkedő művelője például azt vetette fel a kutatásról tartott egyik előadásom vitájában, hogy minek fáradozni annyit a tagmondathatárok (*A módszer* című részben ismertetett) azonosításával és a kidolgozott tagmondatkapcsolatok más szerkezetektől való elkülönítésével, amikor egyszerűbben lenne mérhető pusztán a kötőszavak gyakorisága. Ezáltal több adat állna rendelkezésre, amelyek segítségével statisztikailag is megalapozottabb következtetések tehetők, illetve a kutatás amúgy sem előre meghatározott grammatikai kategóriák azonosítását, hanem a szövegek említett topológiai-diagrammatikus karakterének feltárását tűzte ki céljaként, ez a kérdésfeltevés pedig nem kizárólag a tagmondatkapcsolatok gyakoriságára vonatkozik. A hozzászólás sokat alakított a kutatás előfeltevésein; mindazonáltal úgy gondolom, hogy egyszerűen a kötőszavak összeszámlálása a poliszém jelentések és a szemantikailag is jól elkülöníthető grammatikai szerkezetek miatt valójában nem vezetne értelmezhető eredményre. Az előre definiált nyelvészeti kategóriáktól esetenként elrugaszkodó digitális bölcsészeti logika ugyanakkor megjelent a keresések módszertanának véglegesítésében, és a tagmondatkapcsolatok azonosításakor megengedő módon jártam el, azaz a tagmondatkapcsolatokhoz hasonló grammatikai jelentéseket kidolgozó mondatrészek közötti kapcsolatok egy részét nem zártam ki a keresési találatok közül, vállalva ezzel azt, hogy a kialakított módszer csak korlátozásokkal alkalmazható kizárólag a tagmondatkapcsolatok azonosítására.

[7] „Kiderül a vizsgálatból, hogy az úgynevezett »rövidmondatos« stílust sem lehet egyoldalúan egyszerű mondatosnak, esetleg egyoldalúan mellérendelőnek elképzelni. Még Gárdonyi elvszerűen megalkotott rövid mondatos szövegeiben is csaknem 50%-nyi az összetett mondat, és kevés olyan összefüggő részt lehet találni a művekben, amelyben »nagyobb távon« csak egyszerű mondatot ismer fel a kereső." V. Raisz Rózsa, „Mondatgyakorisági vizsgálatok néhány századforduló kori szépprózai szövegben," *Magyar Nyelv* 72, 2. sz. (1976): 202–209, 208.

[8] Főbb művei e tekintetben: Herczeg Gyula, „A mondatstilisztikai kutatás mint módszer," in Telegdi Zsigmond, szerk., *Általános nyelvészeti tanulmányok XI: A szöveg megközelítései,* 127–142 (Budapest: Akadémiai Kiadó, 1976); Uő., *A régi magyar próza stílusformái* (Budapest: Tankönyvkiadó, 1985); Uő., *A XIX. századi magyar próza stílusformái* (Budapest: Tankönyvkiadó, 1981); Uő., *A modern magyar próza stílusformái* (Budapest: Tankönyvkiadó, 1975).

A korpusz

Labádi Gergely a korai magyar regény adatbázisának létrehozásakor György Lajos korszakolását követve az 1730-tól 1830/40-ig tartó időszakban jelöli ki a magyar regény kialakulását vagy előtörténetét.[9] Vaderna Gábor és Szilágyi Márton irodalomtörténeti leírása szerint „az ekkor megalapozódó új [prózai] műfaj azonban eltér attól, amit később, az 1830-as évektől *regényként* határoz meg a korabeli kritika."[10] Éppen ezért a kiforrott magyar regényirodalom kezdetét, az 1830-as éveket tettem meg a korpusz kiindulópontjaként. A végpontot Nádas Péter *Párhuzamos történetek* című 2005-ös regénye jelenti, kizárva a kutatásból a kortárs magyar irodalom elmúlt évtizedeit, mivel nehéz lenne olyan szempontot találni, amelynek segítségével csupán pár kanonikus mű kiválasztható a jelenlegi magyar irodalomból. Kanonikus művekre azért van szükség, mert a 19. század első feléből jóformán csak ezek érhetők el digitalizált formában, ami így meghatározta a korpusz egészének szerkezetét. Ugyanígy az ebből a korszakból elérhető alkotók száma szabja meg azt is, hogy későbbi időszakokból hány szöveget vettem fel a kutatásba, mivel arányos időbeli eloszlásra törekedtem. Így minden évtizedből legkevesebb 3, legfeljebb 8 szöveg került a korpuszba. Ez összesen 100 regényt[11] és 173 évet, azaz átlagosan 1,7 évenként egy regényt jelent. Szintén az arányosság miatt egy írótól maximum 4, de inkább kevesebb regényt válogattam; ez összesen 58 szerzőt eredményezett, mindegyiknek átlagosan 1,7 regényét. Egy szerzőtől abban az esetben válogattam be 4 írást, ha azok között jelentős az időbeli távolság, hogy ezáltal vizsgálható legyen az is, hogy az egyes tendenciákat követi-e az adott alkotó életműve, vagy inkább időszaktól független szerzői „ujjlenyomatokról" beszélhetünk. Ugyanezért hoztam létre a Jókai Mór 23 regényét magagába foglaló alkorpuszt is. A kutatás egy későbbi szakaszában érdemesnek tűnik a vizsgált szempontokat más műfajú (kisprózai, nem fikciós, akár tudományos-értekező) szövegeket tartalmazó korpusz esetében is érvényesíteni.

A szövegek összetételét tehát az irodalomtörténeti konszenzuson túl a Magyar Elektronikus Könyvtár és a Digitális Irodalmi Akadémia adatbázisai határozták meg. Mivel egyik sem érvényesít határozott irodalomtörténeti koncepciót a digitalizáció során, ezért fordultam az ELTE BTK Digitális Bölcsészet Tanszék által egy nemzetközi projekt keretében összeállított regénykorpuszhoz, amely az 1850 és 1920 között megjelent magyar nyelvű prózairodalomból előre meghatározott szempontok szerint válogat,[12] valamint az Akadémia Kiadó *Magyar irodalom* című kézikönyvéhez és a magyar próza stílustörténetét tárgyaló korábbi monográfiákhoz (például Herczeg Gyula köteteihez). A korpuszt ezenkívül egy-két népszerű szerzővel egészítettem még

[9] Labádi Gergely, „A magyar regény adatbázisa," *Acta Historiae Litterarum Hungaricarum: Acta Universitatis Szegediensis* 32, 1. sz. (2016): 11–30. Lásd még: György Lajos, *A magyar regény előzményei* (Budapest: Akadémiai Kiadó, 1941).

[10] Szilágyi Márton és Vaderna Gábor, „A klasszikus magyar irodalom (kb. 1750-től kb. 1900-ig)," in *Magyar irodalom,* főszerk. Gintli Tibor, 313–641 (Budapest: Akadémiai Kiadó, 2010), 370.

[11] A regények listáját lásd a *Függelék* 1.1. pontjában.

[12] A projekthez hozzáférés: 2021.04.18, `https://www.distant-reading.net/eltec/`, amelynek magyar nyelvű adatbázisa elérhető az alábbi linken, hozzáférés: 2021.04.18, `https://github.com/COST-ELTeC/ELTeC-hun`. Az ebből az adatbázisból kinövő *Regénykorpusz* az összetett kereséseket lehetővé tevő keresőfelülettel elérhető az alábbi linken, hozzáférés: 2021.04.18, `https://regenykorpusz.elte-dh.hu/`.

 DOI: 10.31400/dh-hun.2021.4.2415

ki (például Rejtő Jenő, Gárdonyi Géza). Újabb szövegek digitalizációja és nagyobb, megbízható adatbázis kiépítése elsődleges feladata a hazai irodalomtudománynak; meglehet, a most bemutatott eredmények is árnyalhatók lennének egy szélesebb körű kutatás során.

A statisztika alapvető belátása, hogy minden eredmény a feldolgozott adatok közti viszonyt mutatja és csak bizonyos számú adat felett tekinthető ez a viszony reprezentatívnak. Az itt feldolgozott 100 regény semmiképpen sem elegendő ahhoz, hogy a magyar prózairodalom egészére nézve biztos következtetéseket vonhassunk le, inkább egy ma élő kánonra és a kanonikus szövegek történeti alakulására láthatunk rá általuk.

A módszer

A magyar helyesírásban a tagmondathatárok mondatközi írásjelekkel (vessző, pontosvessző, kettőspont, gondolatjel) jelöltek, így azonosításuk is könnyebben elvégezhető, mint más nyelvek esetében, amelyek nélkülözik az ilyen típusú jelölést. A tagmondatkapcsolatok azonosítása így a kutatásban a mondatközi írásjelek és kötőszavak vagy vonatkozó névmások kombinációjának azonosításán alapul: ezek a kombinációk reguláris kifejezésekkel könnyen detektálhatók. Felmerülhetnek más, összetettebb grammatikai jellemzők szerinti keresések is, amelyek szerkesztetlen szövegek (pl. lejegyzett beszéltnyelv), vagy más nyelvek esetében is használhatók lehetnek: ezek közül a legkifinomultabbnak a véges igék között elhelyezkedő kötőszavak azonosítása tűnik, hiszen egy tagmondat prototipikusan egy véges (lehorgonyzott)[13] igéből mint állítmányból és annak kidolgozásából áll. Ez a módszer ugyan általánosabb és kifinomultabb szempontot biztosít a tagmondatkapcsolatok felismeréséhez, ugyanakkor három nehézségbe ütközik. Egyrészt a véges igék felismerését a kutatásban használt *e-magyar* szóelemző[14] nem tudja hibátlanul elvégezni – ez különösen igaz a 19. század első felében keletkezett szövegekre –, másrészt több kötőszó gyakran nem a véges igék között helyezkedik el, hanem például a mondat elején, harmadrészt ez a szempont nem veszi figyelembe azokat az eseteket, amikor zéró kopulával alkotott névszói-igei állítmány építi fel a tagmondatot („Az a ház szép, amelyik magas"). Ez utóbbira, a zéró kopulás szerkezetek automatikus azonosítására ígéretes eljárások kifejlesztése kezdődött meg, amelyek a későbbiekben mind az *e-magyar,* mind a tagmondathatárok elemzésének hatékonyságát növelhetik.[15] A nehézségek miatt azonban a jelenlegi

[13] Episztemikus lehorgonyzás alatt a kognitív nyelvészetben azt a szemantikai műveletet értjük, amely működése révén a típusból példány jön létre (pl. a szótári igealakból teljes igei alak), tehát hozzáférhetővé válnak a referenciális jelenetre való vonatkozások az alaphoz viszonyítva a megnyilatkozó és a befogadó figyelemirányulásának és mentális feldolgozásának, valamint tudásának megfelelően. Vö. Ronald W. Langacker, *Foundations of Cognitive Grammar, Vol. I* (Stanford: Stanford UP, 1987). Az igék episztemikus lehorgonyzásához lásd Tolcsvai Nagy Gábor, *Nyelvtan: A magyar nyelv kézikönyvtára* (Budapest: Osiris Kiadó, 2017), 349–352.

[14] Váradi Tamás, Simon Eszter, Sass Bálint, Mittelholcz Iván, Novák Attila és Indig Balázs, „E-magyar – A Digital Language Processing System," in Nicoletta Calzolari et. al., eds., *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018),* 1307–1312 (Miyazaki: European Language Resources Association, 2018).

[15] Dömötör Andrea, Yang Zijian Győző és Novák Attila, *„Nesze semmi, fogd meg jól! Zéró kopulák automatikus felismerése neurális gépi fordítással,"* in *XVI. Magyar Számítógépes Nyelvészeti Konferencia,*

kutatás megmaradt a mondatközi írásjeleket követő, vagy a mondat elején álló kötőszavak/vonatkozó névmások reguláris kifejezésekkel történő azonosításánál, amely egy megbízhatóbb és pontosabb módszert jelent szerkesztett szövegeket elemezve.[16]

A reguláris kifejezéseknél az alábbi jellemzők meghatározása vált kiemelten fontossá: (1) a kötőszavak/névmások pozíciója, azaz az írásjeltől való lehetséges távolsága; (2) a poliszém kötőszók funkcióeloszlása és prototipikus jelentése, valamint (3) az adott kötőszó prototipikusan tagmondatok vagy mondatrészek között létesít-e inkább kapcsolatot. Ezeket a kérdéseket a korpuszból vett 100 mondatos minták kézi elemzésével igyekeztük megválaszolni Bajzát Tímea és Szlávich Eszter nyelvészek, a Digitális Bölcsészet Tanszék munkatársai segítségével.[17] Amennyiben egy kötőszó/vonatkozói névmás a 100 esetből legalább 60-szor egy típusú tagmondatkapcsolatot dolgozott ki, akkor az adott fajta alá soroltuk be, ellenkező esetben kizártuk a kutatásból. Például a *hogy* kötőszót vizsgálva a tagmondatkapcsolat típusa nem határozható meg egyértelműen, emiatt nem vettük fel a kutatásba. Egyedül abban az esetben kezeltük rugalmasan a kategorizációt, ha egy kötőszó egyértelműen beazonosítható grammatikai jelentést dolgoz ki, ugyanakkor formailag nem elkülöníthető, hogy mondarészek, vagy tagmondatok között szerepel mint kötőelem („[…] csak Lengyelországban lesz az emberből kiszolgáltatott, légüres térben kapkodó irodalmár, azaz semmi." [Spiró György, *Az Ikszek*]). Ekkor, ha az esetek több, mint 50%-ában (tehát 100-ból legalább 50-szer) tagmondatok közötti kapcsolatot dolgoz ki az adott kötőszó, akkor felvettük a kutatásba: ilyen volt például az *azaz* kötőszó a korpuszból vett 100-as mintát vizsgálva. A meglehetősen alacsony hibahatár kapcsán fontos itt is hangsúlyozni, hogy ugyan a kutatás a tagmondatkapcsolatok azonosítására irányul, ám valójában nem az előre meghatározott grammatikai kategóriák gyakoriságának, hanem a szövegek diagrammatikus szerveződésének elemzését tűzi ki célul, amelyhez elsősorban minél több adat előállítása szükséges. Így a hasonló grammatikai jelentést hordozó, de mondatrészek között létesülő viszonyokat megengedő módon kezeltem és nem minden esetben zártam ki a keresési eredmények közül, abból a hipotézisből kiindulva, hogy így a regények mondatainak domináns kapcsolattípusai jobban kidomborodhatnak a

MSZNY: Szeged, 2020. január 23–24, 385–398 (Szeged: Szegedi Tudományegyetem TTIK Informatikai Intézet, 2020).

[16] Az így kapott eredményeket ugyanakkor összevetettem a véges igék azonosításán alapuló módszer eredményeivel: ez utóbbi pontatlanabb, de arányait tekintve hasonló értékeket adott meg az egyes szövegekre vonatkozóan. Azaz egy-egy regény jellemzői ezen keresztül is kimutathatónak tűnnek. A véges igék azonosításán alapuló módszer eredményeit lásd a *Függelék* 1.2. pontjában.

[17] Az elemzéskor a tagmondatokat funkcionális keretben vizsgálva olyan egységként határoztuk meg, amelyek egy lehorgonyzott folyamattípus előhívásáért felelős, úgynevezett „magmondattal" rendelkeznek. Ehhez lásd Imrényi András és Kugler Nóra, „Mondattan," in *Nyelvtan,* szerk. Tolcsvai Nagy Gábor et. al., 663–889, (Budapest: Osiris Kiadó, 2018), 703, `https://doi.org/10.14232/JENY.2018.1.8`, valamint Imrényi András, *A magyar mondat viszonyhálózati modellje* (Budapest: Akadémiai Kiadó, 2013), `https://doi.org/10.1556/9789634544258`. A minták elemzésének eredményeit lásd a *Függelék* 2.1. pontjában. Az összetett mondat meghatározásához vö. „A több jelenetből, több elemi mondatból álló szerkezeteket összetett mondatoknak nevezi a szakirodalom. […] A tagmondatkapcsolat két (vagy több) jelenetnek nagyobb integráltságban megvalósuló közlését képviseli ahhoz képest, mintha a jelenetek elkülönülő mondatokban lennének megjelenítve. […] A tagmondatkapcsolatokban mindig bizonyos viszonytípusok valósulnak meg, a tagmondatkapcsolódási típusok grammatikalizálódott mintázatok." Imrényi és Kugler, „Mondattan," 808.

 DOI: 10.31400/dh-hun.2021.4.2415

keresések során. Fontos azonban, hogy a másik véglet, a megkötések nélküli, pusztán a kötőszókra vonatkozó és nagyobb adatbőséget ígérő keresések sem tűnnek célravezetőnek, hiszen a poliszém jelentések (az *így* mutatónévmási, a *bár* főnévi jelentése stb.) és a különböző szerkezetek közötti markánsabb eltérések miatt az így elért találatok már nem a szövegek egy jól körülhatárolható tulajdonságára vonatkoznának. A magyar helyesírás segítségünkre lehet ennek a megengedő, de korlátozásokat is érvényesítő szempontnak a gyakorlati alkalmazásában, amennyiben a mondatrészek között kidolgozott viszonyok csak abban az esetben jelöltek mondatközi írásjellel (az előbeszédben pedig szünettel), ha azok a tagmondatkapcsolatokhoz hasonló grammatikai jelentéssel bírnak („Hogy volt az embernek történelem előtti, azaz olyan korszaka, amelyről semmi, nemcsak írásban, de még szájról-szájra adott mondákban sem maradt fenn: az kétségen felül áll." [Móricz Zsigmond: *Légy jó mindhalálig*]). Kizárólag a tagmondatkapcsolatok azonosítása az itt bemutatott keresések szigorúbb korlátozásával hozható létre.

Fontos továbbá, hogy bizonyos kötőszavak csak közvetlenül az írásjelet követően dolgoznak ki tagmondatkapcsolatot *(és),* míg mások a szórendi rugalmasság miatt nagyobb távolságban szerepelhetnek (pl. *ugyanis, azonban*). Sass Bálint az ilyen eseteknél a vesszőtől való 0–3 szó távolságot határozza meg a kötőszavak lehetséges előfordulásaként,[18] a minták elemzése azonban azt mutatta, hogy ennél is nagyobb távolságban, tulajdonképpen a következő tagmondatokat elválasztó vagy mondatvégi írásjelig terjedő szakaszban bárhol előfordulhatnak ezek a kötőszavak. Az egyes szavakra vonatkozó megállapításainkat a típusok jellemzésekor ismertetem. A kutatás tárgya a teljes mondat, amely a benne kidolgozott kapcsolat szerint sorolódik valamely típus alá. A többszörösen összetett mondatok minden abban előforduló típusba besorolhatók, például a „Válaszában nem találok ellentmondást, s így ha jól meggondolom, igaza volt" mondat egyszerre tartozik a kapcsolatos, a következtető és a feltételes típusokhoz. Azokat az tagmondatkapcsolatokat, amelyek nem kidolgozottak kötőszavak/vonatkozói névmások által, a vizsgálat figyelmen kívül hagyta, például a következőt: „A mondatban nem találtunk kötőszót, [tehát] kihagytuk a kutatásból." Ez a mozzanat az eljárás alapvető jellemzőjére világít rá: a kutatás valójában nem általánosan az összetett mondatokra vonatkozik, hanem az olyan összetett mondatokra, ahol a tagmondatkapcsolatok a kötőszavak/vonatkozó névmások által kidolgozottak. Egy ilyen kidolgozott kapcsolat jobban hozzáférhetővé teszi a tagmondatok közötti viszonyt, azaz nem létrehozza, „hanem profilálja azt, hozzáférhetővé teszi az összefüggés jellegét, és ezzel irányítja az értelmezői folyamatot".[19] A kidolgozatlan kapcsolatoknál ezzel szemben többféle értelmezés is elképzelhető, illetve a tagmondatok közötti viszony kevésbé kerül a résztvevők figyelmének előterébe. Ezen túl a kötőszavak azért is lehetnek különösen fontosak egy kvantitatív elemzés számára, mert mint funkciószavak[20] nagyon gyakoriak egy szövegben, és így statisztikailag megfelelő

[18] Sass Bálint, *Igei szerkezetek gyakorisági szótára: Egy automatikus lexikai kinyerő eljárás és alkalmazása,* doktori disszertáció (Budapest: Pázmány Péter Katolikus Egyetem, 2011), 36.

[19] Imrényi és Kugler, „Mondattan," 811.

[20] A számítógépes stíluskutatás erősen hagyatkozik a szavak olyan felosztására, amely szerint léteznek nyitott nyelvosztályba tartozó, tartalmas jelentéssel bíró szavak (prototipikusan főnevek, melléknevek, igék) és zárt nyelvosztályba tartozó (azaz alig bővülő, meghatározott készletet alkotó), pontos

mennyiségű adatot szolgáltatnak a szövegek mélystruktúráinak feltérképezéséhez; ugyanakkor – több más funkcionális szóval ellentétben – tartalmas (grammatikai) jelentést is hordoznak (az általuk kidolgozott kapcsolatról),[21] ezáltal az eloszlásukra és gyakoriságukra vonatkozó eredmények könnyen értelmezhetők. A magyar regény kvantitatív módszerek segítségével felvázolt stílustörténete a kötőszavaknak ezen a kettős tulajdonságán alapul.

A minták kiértékelését követően a következő típusokat és szavakat különítettük el, amely típusok csak módszertani szempontok miatt alkotnak élesen elhatárolt kategóriákat, valójában inkább a családi hasonlóság mintájára szerveződnek. Ugyanígy az egyszerű és összetett mondatok elkülönítése sem magától értetődő, hiszen egy kötőszós szerkezet mind a kettőben szerepelhet. A kutatás ezt a nehézséget az alább felsorolt kritériumokkal igyekezte áthidalni, miközben a hibaszázalékok kapcsán említett megengedő szempontot is szem előtt tartotta.

1. Kapcsolatos mellérendelés. A mondatközi írásjelet közvetlenül követő *és, sőt, s, valamint, majd, aztán, azután* szavak valamelyike – az írásjelhez való közelség oka, hogy ezáltal kiszűrhető az *aztán* partikulaként való szereplésének nagy része, illetve a mondatközi írásjelet nem közvetlenül követő esetek is, amelyek nem a tagmondatok, hanem mondatrészek viszonyára vonatkoznak. Ezenkívül a tagmondatok közti írásjelet 0–2 szó után követő *ráadásul* és az *egyrészt … másrészt*, *hol … hol* összetételek valamelyike.
   A *majd* határozószó ugyan nem tekinthető kötőszónak (hiszen tartalmas jelentése van, és valódi kötőszavak mellett is szerepelhet, megjelenése voltaképp nem befolyásolja az összetett mondat típusát), ám 100 olyan esetből, amikor más kapcsolatos kötőszó nincs a tagmondatban, 63-szor kapcsolatos kronologikus jelentésű viszony kidolgozásában vett részt.[22] Ehhez hasonlóan az *aztán* is 80-szor szerepel kronologikus kapcsolatos mellérendelésben, és csak 14-szer diskurzusjelölőként a mondatközi írásjelet követően.
   A *nemcsak… hanem* többtagú kötőszó 100 esetből csak 18-szor jelölt tagmondatok közötti kapcsolatot, a többi esetben mondatrészek viszonyát dolgozta ki, így nem képezi a vizsgálat tárgyát.
2. Ellentétes mellérendelés. Egy mondatközi írásjel és egy másik írásjel között álló *azonban, ellenben, viszont, ámde, csakhogy, ellenkezőleg, mindazonáltal, ugyan-*

jelentést nélkülöző, erősen grammatikalizálódott szavak, azaz funkciószavak (pl. névelők, kötőszók, prepozíciók stb.) Vö. Daniel G. Morrow, „Grammatical morphemes and conceptual structure in discourse processing," *Cognitive Science* 10, 4. sz. (1986): 423–455, https://doi.org/10.1207/s15516709cog1004_2.

[21] A *hogy* kötőszó ebből a szempontból is kivételt jelent, hiszen „a tartalomkifejtő mellékmondat élén álló *hogy* kötőszó csak a főmondattól való függést jelöli, míg más kapcsolóelemek fogalmi szerkezete ennél általában kidolgozottabb (a kidolgozottság azonban mindig fokozat kérdése). Például a *mert,* a *mint* és egyéb, szófajuk szerint is kötőszói kifejezések úgy töltenek be kontextualizáló szerepet a mellékmondatban, hogy megnyitnak valamilyen fogalmi tartományt (mentális teret) az ábrázolt jelenet értelmezése számára: a *mert* az ok/magyarázat, a *mint* a hasonlítás tartományát jelöli ki ehhez." Imrényi és Kugler, „Mondattan," 832.

[22] A kronologikus egymásra következés a kapcsolatos mellérendelések prototipikus esetének tekinthető az MNSz adatai alapján: „Az és kötőszós tagmondatkapcsolatban az a leggyakoribb (30%), hogy J1 és J2 időben egymást követő jelenetek, J2 a cselekvésláncban feltételezi J1-et." Uo., 860.

 DOI: 10.31400/dh-hun.2021.4.2415

akkor, márpedig, mégse(m), mégis kötőszavak valamelyike, illetve a *csak hát* szerkezet – az írásjelet követő nagy lehetséges terjedelem a szórendi rugalmasság miatt fontos –, és az írásjelet közvetlenül követő *de, ám, pedig* kötőszavak. Az *ellenkezőleg* és a *csak hát* ugyan nem kötőszavak, de prototipikusan kapcsolóelemként működnek tagmondatok között; miközben az *ellenkezőleg* pontosítást és fokozást is kifejez. A reguláris kifejezés nem tesz különbséget a *nemcsak … hanem/de* típusú kapcsolatok második tagjában szereplő ellentétes jelentésű kötőszavak között.

3. Választó mellérendelés. A tagmondatok közti írásjelet közvetlenül követő *vagy, avagy* kötőszavak valamelyike. Az írásjelhez való közelséggel az olyan esetek nagy része kiszűrhető, amikor a *vagy* mondatrészek közötti kapcsolatot létesít, vagy egyes szám második személyű létigeként szerepel. A választó viszony esetenként nem tagmondatok, hanem mondatrészek között dolgoz ki kapcsolatot („Nem minden mutatványnak halál a vége, ahogy a mutatványoknak egyébként sok köze van a halálhoz és a pusztuláshoz, vagy legalábbis ahhoz a rettenethez, amely a halállal van összefüggésben." Darvasi László, *A könnymutatványosok legendája*), amit a kutatás nem különített el és tagmondatkapcsolatként azonosít.
4. Magyarázó mellérendelés. Egy mondatközi írásjel és egy másik írásjel között álló *azaz, azazhogy, ugyanis, hiszen, vagyishogy, tudniillik, miszerint, mármint, mégpedig, egyszóval, elvégre, utóvégre, illetve* kötőszavak valamelyike – az írásjelet követő lehetséges nagy terjedelem a szórendi rugalmasság miatt fontos (a korpuszt vizsgálva 5–6, vagy annál több szó távolságra egyébként elenyészőnek mondható a kötőszavak felbukkanása).
   Az *illetve* sajátosan viselkedő és poliszém kötőszó: magyarázó, választó és kapcsolatos jelentést is felvehet. A 100 összetett mondatból álló mintában azonban 70-szer magyarázó tagmondatkapcsolatban, legtöbbször a *pontosabban* szinonimájaként szerepel, sokszor kapcsolatos jelentésárnyalattal. (29 esetben is hasonló jelentést vesz fel, ám mondatrészek közötti kapcsolat kidolgozását végzi el.)
   Az *azaz* kötőszó 100 összetett mondatból 53-szor szerepelt tagmondatok közötti kapcsolóelemként, de a többi esetben is hasonló grammatikai jelentést hordoz, így a kutatás részét képezi, ugyanakkor a hasonló jelentésű és viselkedésű *vagyis* csupán 43-szor dolgozott ki tagmondatok közötti kapcsolatot, így nem vettük fel a kutatásba (szemben a *vagyishogy* kötőszóval, amely ugyan elenyésző számban szerepel a korpusz mondatai között, akkor viszont jellemzően tagmondatkapcsoló elemként).
   Ebbe a kategóriába tartoznak továbbá a mondatközi írásjelet közvetlenül, vagy a *mert, de, és, s* kötőszavakat követő *hisz, pontosabban.*
5. Következtető mellérendelés. Egy mondatközi írásjel és egy másik írásjel között álló *tehát, ennélfogva, eszerint, úgyhogy, szóval* kötőszavak valamelyike.
6. Megengedő alárendelés. Az alábbi kötőszavak, ha a mondat legalább egy mondatközi írásjelet tartalmaz: *bár, ha… is, habár, igaz, hogy, jóllehet, még ha, mégha, noha, ugyan, ámbár.* A reguláris kifejezés ekkor a leghatékonyabb, mivel figyelembe veszi a mondat elején álló kötőszavakat is, amelyek ebben az esetben prototipikusan fordított, mellékmondat-főmondat tagmondatsorrendet többször vezetnek be, és nem az előző mondatra vonatkoznak, például: „Bár szerelem

miatt kimondhatatlanul szenvedett, e reggelen se mulasztotta el mindennapi tisztelgését a Ferenciek terén…" (Krúdy Gyula, *Hét bagoly*). Ha előírjuk, hogy a mondat legalább egy mondatközi írásjelet tartalmazzon, akkor a vizsgált 100 mondatból 66-szor fordított tagmondatsorrendű megengedő kapcsolatot dolgozott ki a mondat elején álló kötőszó.
Ebbe a kategóriába tartozik továbbá a mondatközti írásjelet közvetlenül követő *holott* kötőszó, mivel a *holott* az előző mondatra vonatkozik prototipikusan a mondat elején, és nem fordított tagmondatsorrendű összetett mondatot vezet be.

7. Hasonlító alárendelés. A mondatközi írásjelet közvetlenül, vagy 1 szó után követő, valamint a mondat elején, vagy az első szó után álló *mint, mintha, akár, akárha, akárcsak, minthogyha* kötőszavak valamelyike. Kivételt képeznek azok az esetek, amikor a *nem* vagy az *a* előzi meg a *mint* kötőszót, mert az előbbi esetben a *nem mint + főnév* szerkezetek, utóbbiban pedig a régiesen használt *amint (a mint)* megvalósulása a jellemző.
Az *akár* hasonlító jelentése csak akkor van túlsúlyban, ha nem *akár… akár…* sémájú konstrukcióban szerepel, így ennél a kötőszónál előzetesen biztosítani kell, hogy egy *akár* legyen a mondatban – ekkor 82-szer hasonlító és 18-szor partikula szerepben volt azonosítható a 100-as mintában. Ezt a kritériumot a reguláris kifejezésben is meghatározhatjuk, ám a teljesítmény optimalizálása miatt célszerűbb a reguláris kifejezést megelőzően a korpusz mondatait egy új változóhoz rendelni, és a *nem mint* és az *a mint* szószerkezetekkel együtt azokat a mondatokat is eltávolítani ebből a korpuszból, amelyekben több *akár* szerepel – majd ezen a szűkített korpuszon végezni a keresést.
Amennyiben egy szó áll a mondatközi írásjel és a kötőszó között, 60-szor hasonlító kapcsolat, 30-szor nem tagmondatkapcsolat állt fenn a vizsgált 90 esetből.
A hasonlító összetételnél figyelembe kell venni, hogy a főmondatrész gyakran csak odaértett módon jelenik meg a mondatban: „[Olyan vagy,] Mintha megváltoztál volna." A mondat elején álló kötőszó ebben az esetben 100-ból 31-szer jelentett tipikus hasonlatot, 16-szor nem tagmondatok közötti viszonyt írt le, és 53-szor, leginkább a *mintha* kötőszó esetében odaértett főmondatrészről volt szó. (A *mint* kötőszó a mondat elején 50-es mintában 37-szer vonatkozott kidolgozott hasonlítói összetételre, 7-szer nem tagmondatok közötti viszonyra, 7-szer pedig odaértett főmondatra). A jelenlegi kutatás nem tett különbséget a valóban kidolgozott és az odaértett hasonlatok között, ám a későbbiekben érdemes lehet a hasonlító összetett mondatok differenciáltabb tipológiáját megadni.
A reguláris kifejezés továbbá tévesen ismer fel hasonlító összetett mondatként mondatrészek közötti kapcsolatot is, például: „Találkozott Jánossal, kivel mint szakemberrel beszélhetett." Az ilyen hibák azonban statisztikailag elenyészők a hasonlító összetételek számához képest. Nem jelent hibát továbbá, hogy a minősítő minőségjelzői kapcsolatok közül csak a kötőszóval kidolgozott tagmondatkapcsolatokat sorolja be a keresés a hasonlító összetett mondatok alá („Egy olyan nyakkendőre lenne szükségem, mint [amilyen] az öné"), szemben azokkal, amelyeknél a viszonyt nem dolgozza ki kötőszó („Olyan vagyok, [mint]

 DOI: 10.31400/dh-hun.2021.4.2415

amilyen te"), hiszen a kutatás éppen a formai oldalon is kidolgozott és a figyelem előterében álló kapcsolatokra vonatkozik.

8. Ok-, célhatározói alárendelés. A tagmondatok közti írásjelet közvetlenül, vagy 1 szó után követő *mert, merthogy, minthogy, mivel, nehogy* kötőszavak valamelyike – az írásjeltől való nagyobb távolság többnyire téves eredményekre vezet, elsősorban a *mert* és a *mivel* homonim jelentései miatt. Ebbe a kategóriába tartozik továbbá a mondatközi írásjelet közvetlenül követő *különben* kötőszó, illetve az *amiatt, avégre, avégett, azért …, hogy* szerkezetek.
   Ezenkívül szintén ebbe a típusba tartoznak a mondat eleji, vagy az első szót követő *minthogy, mivel* kötőszavak valamelyike a legalább egy mondatközi írásjelet tartalmazó mondatokban, mivel ezek prototipikusan fordított tagmondatsorrendű összetett mondatot vezetnek be („Minthogy többségben vannak az ilyen esetek, felvettük őket is a kutatásba"). A korpuszból vett 100 mondatban 90-szer fordult elő fordított tagmondatsorrendű ok-/célhatározói összetett mondat, amennyiben a reguláris kifejezés előírta, hogy legalább egy mondatközi írásjelet tartalmazzon a mondat.
9. Időhatározói alárendelés. A mondatközi írásjelet közvetlenül követő *mire, közben,* valamint az alábbi kötőszavak/névmások, ha a mondat legalább egy mondatközi írásjelet tartalmaz: *alighogy, ameddig, amíg, amikor, amióta, attól fogva/kezdve …, hogy, amint, mígnem, amidőn, midőn, mielőtt, míg, mialatt, mihelyt, mikor, mióta, miután.* Ezek a kötőszavak ugyanis a mondat elején prototipikusan (100 esetből 94-szer) fordított tagmondatsorrendű időhatározói kapcsolatot vezetnek be, amennyiben a reguláris kifejezés előírja, hogy legalább egy mondatközi írásjelet tartalmazzon a mondat (kizárva az egyszerű kérdő mondatokat, például „Mikor jössz?"). A kifejezés – tévesen – olyan eseteket is felismer, mint a „Mikor jössz, he?", ám ezek száma statisztikailag elenyésző a valódi összetett mondatokhoz képest. A *mire* kötőszó 100 esetből 64-szer dolgozott ki időhatározói kapcsolatot, és csupán 14-szer vezet be prototipikus vonatkozói alárendelést.
   Ezenkívül kezelni kell a mentális állapotról beszámoló *hogy* kötőszós mellékmondatokat, amelyben időhatározói vonatkozó névmás is szerepel, mert ez a konstrukció nem időhatározói alárendelésre utal („Azt kérdezte, hogy mikor jön el a boldogság és a nyugalom ideje"). Ugyanígy ki kell szűrni azoknak a mellékmondatoknak a nagy részét is, amelyekből hiányzik a *hogy* kötőszó („Nem tudom, mikor oldják fel a korlátozásokat"). Ezek közül azokat lehet kis hibaszázalékkal azonosítani, amelyek esetében a mentális folyamatokat kidolgozó igék és ragozott formáik közvetlenül[23] előzik meg a mondatközi írásjelet (lásd az előző példát). Ennek érdekében a Magyar Nemzeti Szövegárból (*MNSz* v2.0.5) vételezett mintákon vizsgáltuk, hogy prototipikusan milyen igék szerepelnek a *hogy* kötőszós mellékmondatok előtti főmondatokban,[24] majd ezeknek az igék-

[23] Azokban az esetekben, amikor nem közvetlenül előzik meg az írásjelek a felsorolt igéket, az eredmények megbízhatatlanok voltak és nem csupán a *hogy* kötőszós mellékmondati konstrukciókra vonatkoznak.

[24] Oravecz Csaba, Váradi Tamás, and Sass Bálint, „The Hungarian Gigaword Corpus," in Nicoletta Calzolari et. al., eds., *Proceedings of LREC 2014,* 1719–1723 (Reykjavík: European Languages Resour-

nek a viselkedését a korpuszból vett mintákon elemeztük. A minták kiértékelését követően az alábbi töveket tudtuk meghatározni, amelyek mentális folyamatot kidolgozó igeként prototipikusan ilyen konstrukciókban szerepelnek: *főnévi igenév + akar, elárul, eldönt, elképzel, érdekel, tud, kérdez, kérd, képzel, kitalál, megállapít, megért, megmutat, sejt.*
Ezenkívül fontos a *mint* kötőszó után álló időhatározói vonatkozói névmások kiszűrése is („Olyan volt, mint amikor elfelejted a locsolóverset"). Ezeknek figyelmen kívül hagyását a *hogy* kötőszós mellékmondatokkal együtt a reguláris kifejezésben is definiálhatjuk, ám a teljesítmény optimalizálása miatt célszerűbb a reguláris kifejezést megelőzően a korpusz mondatait egy új változóhoz rendelni, és ezeket a konstrukciótípusokat eltávolítani ebből a korpuszból – majd ezen a szűkített korpuszon végezni a keresést.

10. Helyhatározói alárendelés. Az alábbi névmások, ha a mondat legalább egy mondatközi írásjelet tartalmaz: *ahol, ahonnan, ahova, ahová, ameddig, amerre* (a mondat elején álló kötőszó ekkor a 100-as mintából 77-szer helyhatározói alárendelést vezetett be), valamint a mondatközi írásjelet közvetlenül – vagy a régies írásmód miatt az *a* mutató névmás után – *hol, honnan, hová, hova, meddig, merről, merre* kötőszavak valamelyike – kiszűrve a nem a tagmondatok közti viszonyt jelölő kérdő névmások egy résztét („Megtántorodott, nem volt hová hátralépnie").
A reguláris kifejezés nem tesz különbséget a *hol …, hol* típusú kapcsolatos mellérendeléssel, amely azonban hordoz helyre vonatkozó jelentésárnyalatot.
Ezenkívül ennél a típusnál is kezelni kell a *hogy* kötőszós mellékmondatokat („Nem tudom, [hogy] hol vagyok."), amelyek egy részét a fent említett módszerrel szűrtük ki a korpuszból.

11. Prototipikus (nominatívuszi) vonatkozói névmással bevezetett mellékmondat. Az alábbi névmások és ragozott alakjaik (azaz azok a vonatkozó névmások, amelyeknek van személyjelölő párja), ha a mondat legalább egy mondatközi írásjelet tartalmaz: *ami, aki, amely, mely,* kivéve, ha a *mint* kötőszó áll előttük („Úgy viselkedtem, mint aki villanyszerelő"). A mondat elején az említett vonatkozói névmások közül az *aki* 100-ból 91-szer, az *ami/amely* 100-ból 70-szer vezetett be fordított tagmondatsorrendű vonatkozói alárendelést összetett mondatokban. Ahhoz, hogy csak a névmások ragozott alakjaira vonatkozzon a kifejezés, a szövegből (a fent ismertetett módon) előzetesen ki kell szűrni az *amint, amikor, amiként, amiképp, amiatt, amidőn, amióta, amialatt, amielőtt, amiután, amíg, melyik* (amennyiben nem az *a melyik* régies írásmód érvényesül) szavakat.
Ebbe a kategóriába tartoznak továbbá az *és, s, de* kötőszavak, (a régies írásmód miatt) az *a* mutató névmás, továbbá a mondatközi írásjelet közvetlenül követő *ki, a mit, kit, mik, kik, miben, kiben, mibe, kibe, a mire, kire, a miről, kiről, mitől, kitől, miből, kiből, kinél, kivel, a mivel, minek, kinek, min, kin, mihez, kihez, miket, kiket, mikben, kikben, mikbe, kikbe, mikre, kikre, mikről, kikről, miktől, kiktől, mikből, kikből, miknél, kiknél, mikkel, kikkel, miknek, kiknek, miken, kiken,*

ces Association, 2014), `https://doi.org/10.13140/2.1.3957.1840`. A konkordanciákat lekérő eljárás és az eredményekből létrehozott gyakorisági listák a *Függelék* 2.2. pontjában találhatók.

 DOI: 10.31400/dh-hun.2021.4.2415

mikhez, kikhez névmások (kiszűrve a nem a tagmondatkapcsolatot kidolgozó kérdő névmások egy részét [„Egyedül harcolok, hiszen kiben is bízhatnék."]), illetve a mondatközi írásjelet közvetlenül követő *mi* névmás (kiszűrve a többes szám első személyű személyes névmások java részét, bár a kifejezés így is tartalmaz olyan eseteket, mint például: „ők dolgoznak, mi beszélünk róla.")
A kifejezés nem tartalmazza a *mivel, mire, mit, miről* névmásokat (csak a régies írásmóddal írt *a mivel, a mire, a mit, a miről* alakokat), mert az első prototipikusan ok-/célhatározói alárendelésben, a második időhatározói kapcsolatban, utóbbi kettő pedig kérdő névmásként, vagy *hogy* kötőszós mellékmondatokban fordul elő; valamint nem különíti el a *ki* szó igekötői használatát, ám a mondatközi írásjel utáni távolság korlátozásával ezek nagy része kiszűrhető.
Ezenkívül ennél a típusnál is kezelni kell a *hogy* kötőszós mellékmondatokat („Nem értem, [hogy] mit mondasz."), amelyek egy részét a fent említett módszerrel szűrtük ki a korpuszból, azzal a kiegészítéssel, hogy ebben az esetben a mentális igékre vonatkozó tövek közé felvettük az *ért* tövet is; a *mi* vonatkozó névmás és ragozott alakjai előtt pedig nem előírás, hogy a mentális folyamatot kidolgozó ige közvetlenül a vessző előtt álljon.

12. Feltételes alárendelés. Az alábbi kötőszavak, ha a mondat legalább egy mondatközi írásjelet tartamaz: *ha, hacsak, amennyiben, hogyha.* Ezek a kötőszavak a mondat elején prototipikusan (100-as mintából 94-szer) fordított tagmondatsorrendű összetett mondatot vezetnek be összetett mondatokban.
A reguláris kifejezés nem tesz különbséget a *ha … is* típusú megengedő mellérendeléssel, ám az is tartalmaz feltételes jelentésárnyalatot.

Az R programnyelvben lefuttatott keresés ezeket a szempontokat érvényesítette, amikor a mondatokra felbontott regényekben (vagy azok szűkített változatában) számlálta meg az egyes típusokhoz tartozó kötőszavak számát. Ezt a számot az adott regény szavainak számával elosztva (és a kezelhetőség érdekében 1000-rel megszorozva) kaphatjuk meg a típusok relatív gyakoriságát. A mondatok számával való osztás nem lenne célravezető, hiszen abban az esetben a hosszúmondatos prózák magasabb értéket vennének fel a rövidmondatos regényekkel szemben. A véges igék száma viszont szintén jó osztó lehetne, de ehhez ismét szükség lenne egy olyan nyelvi elemzőre, amely a 19. századi szövegeken is megbízhatóan működik (és tudja kezelni például a múlt idők különböző fajtáit). A következő részben ismertetett számok és ábrák a regények egészén lefuttatott keresések eredményei – az összes érték (12 típus 100 regényre vonatkozóan) megtalálható a *Függelék* 1.1. pontjában. Célszerű ugyanakkor ezeket az értékeket olyan számításokkal is kiegészíteni, amelyek nem a regények egészén, hanem az azokból vett, lehetőleg dialógusmentes[25] és egyenlő hosszúságú mintákon (jelen esetben: 250 mondat, amely alacsony szám oka Esterházy Péter *Fuharosok* című

[25] A leíró és párbeszédes részek automatikus elkülönítése a MEK és a DIA felületein elérhető különböző kiadások miatt nehézkes (a DIA a könyvbeli forráskiadást imitálja, így a különbségek ott a szerzők és kiadók szokásaira vezethetők vissza); nem adható meg ehhez csupán egyetlen tipográfiai szempont. Az itt használt kódok a szólamváltást jelölő írásjelek mellett a mondást kifejező igéket vették figyelembe a dialógusmentes részletek elkülönítésekor. A kutatás során meghatározott mondást kifejező igék tövei: *mond, kérd, felel, szól, folytat, így, kiált, int, hökk, ord, üvölt, sik, súg, suttog, morm, vél, ismét, nevet, bólog, tett, tev, hall.*

művének rövidsége) alapulnak: míg az előbbi módszer a művek egészéről mond el többet, addig az utóbbi statisztikailag jobban összehasonlítható értékeket szolgáltat (bár ezek az értékek kevesebb adatra vonatkoznak, így esetenként nem is található egy adott kapcsolattípus a vizsgált 250 mondatban). Az eredmények megtalálhatók a *Függelék* 1.3. pontjában.

Ugyanilyen kereséseket végeztem a regények időben egymást követő 10-es csoportjain is a történeti tendenciák vizsgálatához; valamint külön kódok számították ki, hogy az egyes kapcsolattípusok egymáshoz képest milyen arányban találhatók meg a regényekben (lásd *Eredmények 3.*) – ekkor az egy regényben azonosított összes találat számával osztottam el az egyes típusokhoz tartozó értékeket és szoroztam meg 100-zal (esetenként két típust összevontam egy tágabb kategóriába a jobb átláthatóság kedvéért). Az ezeket bemutató értékek a *Függelék* 1.4. (évcsoportok) és 1.5. pontjában (belső arány) találhatók.

Eredmények 1.

A kiszámított értékeket három szempont szerint ábrázolhatjuk: (1) az idővonal mentén a történeti változásokra összpontosítva; (2) összehasonlító módon vizsgálva, hogy mely regényekre mennyire jellemzők az egyes kapcsolattípusok; (3) valamint aszerint, hogy egy regényen belül milyen arányban oszlik el a típusok gyakorisága.

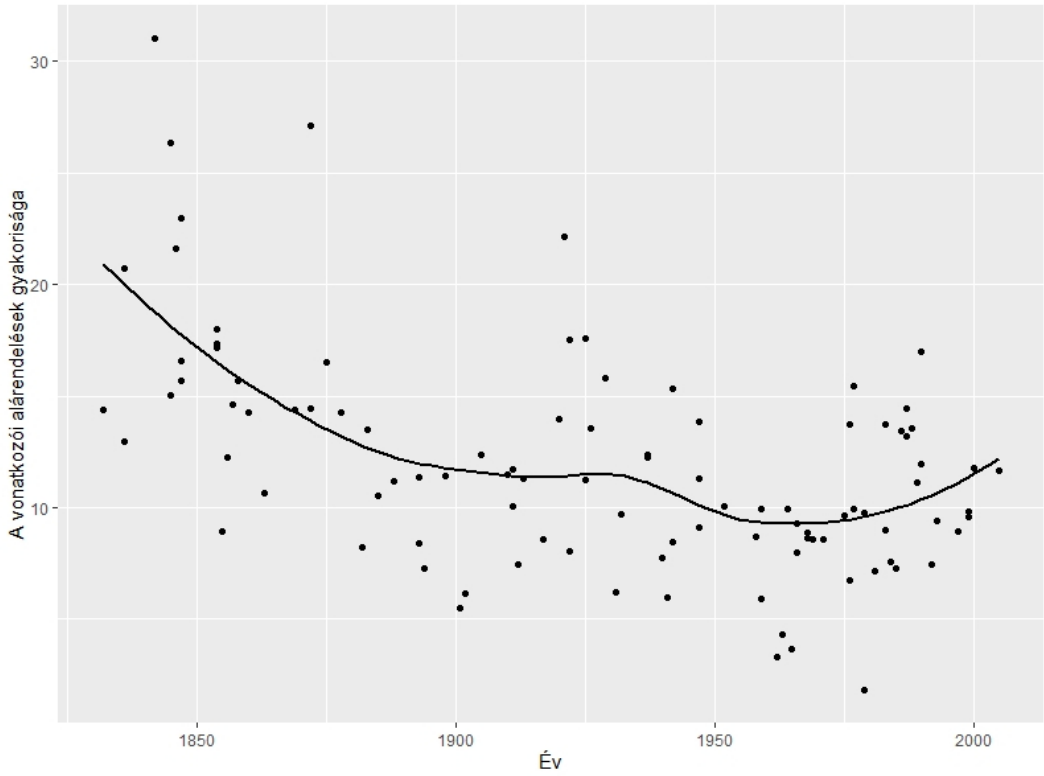
Az első szempont szerint kétféle grafikont hozhatunk létre. Egyfelől ábrázolhatjuk a regényekhez rendelt értékeket a megjelenési évek viszonyában, valamint az így felrajzolt adatpontokra illeszthető regressziós görbét, másfelől a 10 regényenkénti évcsoport értékeit oszlopdiagramok formájában is összehasonlíthatjuk.

A prototipikus vonatkozói alárendelések (időbeli változás)

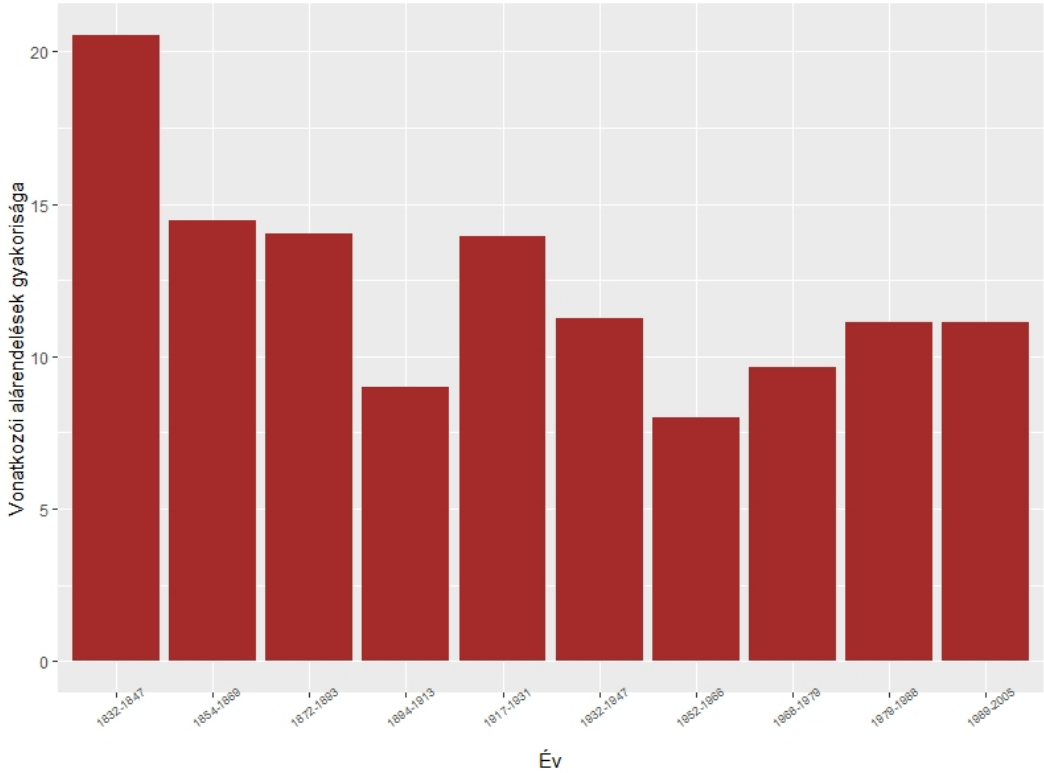
A *2.1.* és *2.2. ábrá*n a prototipikus (nominatívuszi) vonatkozói alárendelések változása figyelhető meg. A *2.1. ábrá*n megfigyelhető csökkenő tendenciát[26] egyszerre okozzák az 1840–50-es évek kiugró értékei, valamint az, hogy az 1870-es évekig alig található szöveg a grafikon alsó régióiban. Az ábra ezáltal arra a poétikai-történeti összefüggésre hívja fel a figyelmet, miszerint a korszak előszeretettel dolgozza ki a mondatok szereplőit részletekbe menően, külön tagmondatok segítségével.

[26] Az egyszempontos varianciaanalízis (ANOVA) alapján erős korreláció mutatható ki a vonatkozói kapcsolattípus gyakorisága és a megjelenési évek között (F-érték: 35.44), azaz a kirajzolódó tendencia által jelölt különbségek statisztikailag szignifikánsnak tarthatók.

 DOI: 10.31400/dh-hun.2021.4.2415



2.1. ábra. A prototipikus vonatkozói alárendelések változása, 1832–2005.



2.2. ábra. A prototipikus vonatkozói alárendelések változása, évcsoportok.

A vonatkozói alárendelés a 19. század első felében nagy hatást kifejtő ritmikus-retorikus prózastílus esetében a leggyakoribb. Ez a stílus olyan többszörösen összetett mondatokból (az úgynevezett „romantikus körmondatokból") épül fel, amelyekben a mondatrészek és a tagmondatok azonos pozícióban és azonos szinten, szimmetrikusan helyezkednek el: ez jelentheti egyrészt azt, amikor ugyanolyan vagy hasonló jelentésű mellékmondatok dolgozzák ki a tagmondatokban ugyanazt a helyet elfoglaló (például a tagmondat elején álló), azonos funkciót betöltő mondatrészeket (pl. alanyokat), másrészt azt, amikor egyetlen mondatrészhez azonos fokú, egymással mellérendelő viszonyban álló tagmondatok felsorolásszerűen kapcsolódnak. Ennek az anaforisztikus stílusnak legfontosabb hazai képviselője Eötvös József,[27] akinek a korpuszba felvett mindhárom regénye az első négy legnagyobb értéket mutató szöveg között található – a 100 regény közül pedig *A karthauzi*ban a leggyakoribb a vonatkozói alárendelések használata. A szimmetrikusan elhelyezkedő vonatkozói alárendelésekre az alábbi példákat hozhatjuk életművéből:

> Ő, ki e házban született s növekedett, ki e falak közt szorítá apját utólszor kebeléhez, ki, mióta az meghalt s őt hivatalában követé – itt tölté életét, kit a faluban minden ház, a temetőn minden sír olyanokra inte, kiknek ő volt legjobb vigasztalója, s ki a pálya vége felé közelgve érzé, hogy emlékei közt más keserűek nincsenek, mint melyeket a halál hagyott: - miként álmodhatott más szerencséről […] (Eötvös József, *A falu jegyzője*)[28]

> […] ti bércek, melyeknek látásánál a gyermekszív egykor vágyakkal telt; te dörögve lezúgó hegypatak, melynek hatalmas szava túl nem fogja hangozni az öröm érzetét, mi e szívben felzajog; te virágzó mező, mely soha nem hazudtál még, s minden tavasszal meghozád virágidat! (Eötvös József, *A karthauzi*)

Az első példában öt, ugyanahhoz a főmondathoz tartozó, a *ki* névmás által bevezetett alanyi mellékmondat található, amelyek egy szinten, egymáshoz képest mellérendelői viszonyban találhatók. A második példában különböző alanyokat dolgoznak ki a *mely* névmással bevezetett mellékmondatok, ám ezek az alanyok és a tagmondatok is a mondat azonos szintjén, azonos pozícióban helyezkednek el. Ezáltal jól követhető, átlátható szerkezet jön létre, amely a mondatok hosszúságának ellenére is könnyen értelmezhető. Ez a szerkesztésmód egészen a 13. századi magánoklevelek és kancelláriai iratok stílusára vezethető vissza, amelyekben a párhuzamos részek az iratok áttekinthetőségét szavatolták: a jogi és retorikai műveltség összefonódása pedig megteremtette a ritmikus-retorikus próza alapjait, amely a kora reneszánsztól mint irodalmi nyelv is kifejtette hatását.[29] Ez az irodalmi nyelv a 19. század Magyarországán Herczeg

[27] Vö. Herczeg Gyula, „Eötvös körmondatai (Folytatás)," *Magyar Nyelvőr* 77, 1–2. sz. (1953): 165–179.

[28] A szépirodalmi idézetek hivatkozott kiadását a *Függelék* 4. pontjában ismertetem és nem jelölöm külön az egyes idézeteket követően.

[29] Herczeg, „A mondatstilisztikai kutatás mint módszer," 133–135. Érdemes elgondolkodni azon az összefüggésen is, amely az államszervezés, a retorika és a perspektivikus képábrázolás között áll fenn: ekkor a párhuzamos mondatszerkesztés nemcsak a táblázatos és egyéb formanyomtatványokba történő elrendezés nyelvi előképeként jelenik meg, hanem a látvány leképezésének mint a hely kiosztásának szintén a reneszánszban kifejlődött technikájával is összevethetővé válhat. Ennek

 DOI: 10.31400/dh-hun.2021.4.2415

Gyula megfogalmazásában „a politikai igények stilisztikai vetületeként" értelmezhető, amennyiben a felsorolásszerűen egymásra következő tagmondatok halmozásával „buzdít, mozgósít, patetikus hangvétellel bírál vagy magasztal".[30]

A ritmikus-retorikus próza a korpusz szövegei közül elsősorban Eötvös regényeiben fedezhető fel, ám találhatunk rá példát Kemény Zsigmond, Jósika Miklós vagy Toldy István műveiben is.[31] Esetükben azonban – ami a korszak prózairodalmára általánosan is igaz – gyakoribbak a beékelésszerű alárendelések, amelyek nem kerülnek a fent bemutatott módon párhuzamba más mellékmondatokkal; illetve a kevésbé szabályosan elrendezett, különböző szinteken elhelyezkedő vonatkozói névmásokkal bevezetett mellékmondatokból építkező összetettebb szerkezetek. Ez utóbbira az alábbi példákat említhetjük:

> Ali azon különös tekintetek egyikét veté most az alig tizenhat éves leányra, melyek szívtelen embereknél, mintegy átküzdvén kedélyök lemezén, inkább valami eszményi igazság érzetét a természetnek látszanak bebizonyítani, mint érzést, mely a szívből kelne – azon igazság érzetét, mely oly harsányul hangzik, a legelfásultabb kedélyek dacára, s akaratlanul kivíjja a méltánylást önkénytelen kifejezésében a vonásoknak, melyek néha többet tolmácsolnak, mint lelkünk érez vagy akar érezni. (Jósika Mikós, *Diamante*)

> […] és alig van társadalom, mely egy bukott nőt ritkábban kényszerítene pirulni, mint volt a párisi a császárság alatt, melyben a demi-monde úgy szólván bizonyos erkölcsi jogosultsággal bírt, melyet el nem ismerni annyi volt, mint „elmaradni a világtól." (Toldy István, *Anatole*)

az összevetésnek adhat alapot a tény, hogy Alberti is a retorikai mondatelemzés mintájára és – éppen a térbeli viszonyokból eredő – terminológiájával *(compositio, locus, circumscriptio)* írja le a képelemzés folyamatát. Vö. Bernhard Siegert, *Cultural Techniques: Grids, Filters, Doors, and Other Articulations of the Real,* trans. Geoffrey Winthrop-Young (Fordham: Fordham UP, 2015), 132, `https://doi.org/10.1515/9780823263783`.

[30] Herczeg, *A XIX. századi magyar próza stílusformái,* 84–85. Innen nézve talán nem véletlen, hogy Eötvös politikus is volt; stílusán érezhető országgyűlési beszédeinek és egyéb publicisztikáinak hatása is.

[31] Kemény Zsigmondnak inkább korai prózájában találhatunk ilyen szerkezeteket, de idézhető például *A rajongók* (1858–59) alábbi mondata is, ahol az *aki* mint általános alany szerepel és négy mellékmondatban vezeti fel a főmondatot: „*Aki* figyelemmel kísérte az ellentétet e férfiú szavai és kedélyhangulata közt, *aki* elleste a fájó benyomást, mely idegein átrezgett, midőn magasztalni hallá magát, *aki* látta, hogy máskor meleg, részvékeny és áradozó szíve mennyire elfeledte a hőbb dobogással együtt a lelkesülést és életörömet, *aki* fanyar mosolyát, sötét merengéseit, nyugtalan álmait és álmatlan éjeit aggódva követé szemeivel s gondjaival: az kénytelen volt valami mélyen ható, valami rejtélyes és bősz történetben keresni e rögtöni átalakulás okát." (Kemény Zsigmond, *A rajongók*) Jósika prózájából a következő idézet szolgálhat példaként, melynek bevezetése az azonos szinten és pozícióban lévő alanyok és a hozzájuk tartozó alárendelések felsorolásából áll: „A fez, *mely* fejét borítá; az ujjas, *mely* derekára szorult; a törökös csálvár, *mely* szárain folyt alá: minden fekete fustanella, *mely* derekán alul térdéig ért, volt fehér és tiszta, mint a tavaszi hó." (Jósika Miklós, *Diamante*) Toldy István 1872-es *Anatole* című regényéből pedig az alábbi példát említhetjük, amelyben egyetlen alanyhoz (a kereskedőhöz) kapcsolódnak a vonatkozói alárendelések: „A Flavigneux család minden ismerősei és rokonai közt alig volt ház, mely több tiszteletet érdemelt volna, mint e derék kereskedőé, *ki* ötven éves korában egy húsz éves ifjú gyöngédségével szerette nejét, *kinek* vágyait a negyven éves nő odaadó szerelme teljesen kielégíté s *ki* mindent elkövetett, hogy három gyermekét a lehető legjobb nevelésben részesítse." (Toldy István, *Anatole*)

Ezekben az esetekben a szimmetrikus szerkesztés hiányában és az egymásból kinövő alárendelések következtében a mondatok szerkezetének az átláthatósága, érthetősége is jelentősen csökken. Ami viszont közös az eddig említett szerzőkben – és a *2.1. ábra* alapján a korszak többi regényírójában is –, az a vonatkozói névmással bevezetett alárendelések nagy száma. Ezek az alárendelések a főmondat egy szereplőjét teszik megfigyelhetővé egy más jelenetben is, ezáltal többletinformációt szolgáltatnak róla, képesek annak meghatározására. A ritmikus-retorikus prózastílusban a mellékmondatok halmozása egyben az olyan jelenetek halmozását is jelenti, amelyben a korábbi tagmondat szereplőjét figyelhetjük meg, és ez a halmozás, az egymást erősítő (mert a korábbi szereplő hasonló tulajdonságait bemutató), és sokszor csattanószerű, rövid zárlatot előkészítő mondatrészek adják a szövegek retorikusságát.[32] Ez a retorikus elem csak helyenként található meg a század közepének prózájában, ám a tagmondatok egyes elemeinek vonatkozói névmással bevezetett mellékmondatokkal történő kidolgozása, azaz egy új jelenetben a tulajdonságok és viselkedésmódok meghatározása alapvetően jellemző a korszakra. Míg Eötvös regényeiben ez az eljárás az éles kritika és az ironikus hangvétel sokszínű eszközeként jelenik meg, addig például Toldy István 1872-es, amúgy nagyon gondosan megformált, *Anatole* című regényében néhol idejétmúltnak hat a társadalom helyes működése és a társadalmi szerepek meghatározásának ez az igénye, amelyet a vonatkozói alárendelések gyakori alkalmazásán keresztül kíván elérni.[33] Ennek megfelelően a század végére nemcsak a ritmikus-retorikus stílus szorul vissza, hanem a vonatkozói alárendelések gyakorisága is csökken (Eötvös három regényét tekintve is megfigyelhető ez a csökkenő tendencia), és inkább rövidmondatos próza veszi át a korábbi stílusok helyét, amelynek szerkezetei nem a szereplők új jelenetben történő megfigyeltetését részesítik előnyben.

A következtető és a magyarázó kapcsolatok (időbeli változás)

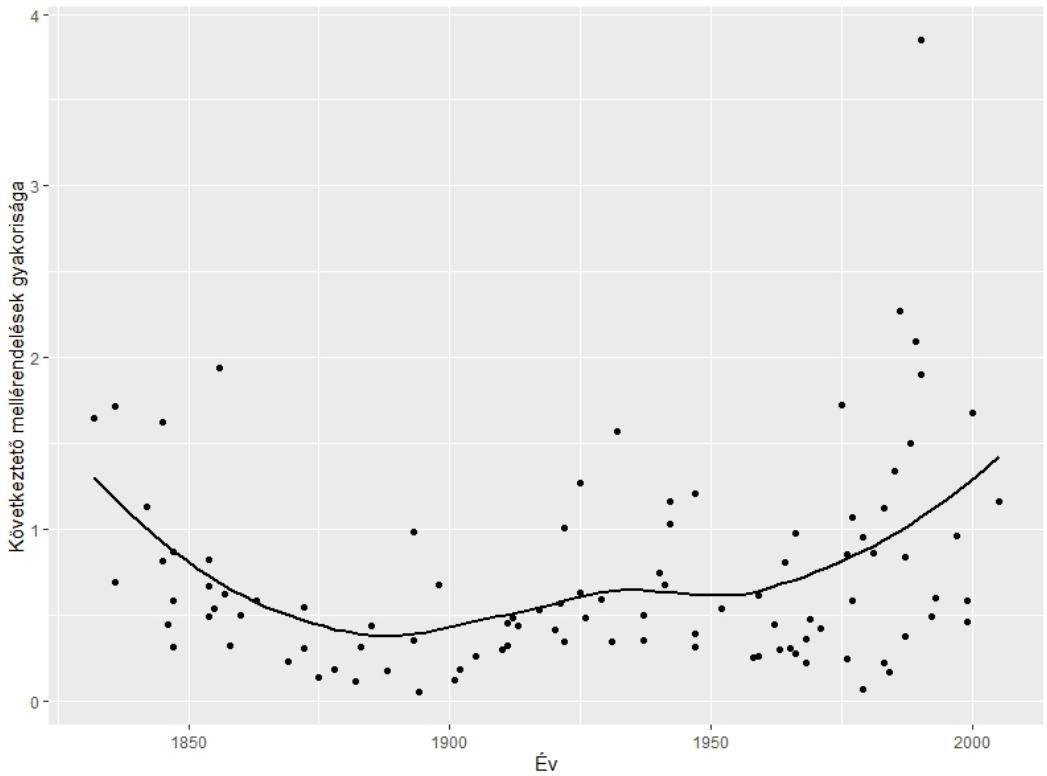
A mérési adatok alapján megfigyelhető egy másfajta prózastílus is a 19. század közepén, amely szintén többszörösen összetett mondatokat hoz létre, viszont kevésbé alárendelésekkel, hanem inkább mellérendelő kapcsolatokkal dolgozza ki a tagmon-

[32] A stílus fennköltsége ugyanakkor gyakran parodisztikus felhangokkal is társul, leginkább abban az esetben, amikor a honi viszonyok tarthatatlanságának leírásával párosul a fennkölt stílus, vagy amikor a vármegyei közigazgatást saját nyelvével gúnyolja ki a szerző. Vö. Herczeg, *Eötvös körmondatai (Folytatás),* 171.

[33] Szemléletes példa lehet ez utóbbira az alábbi részlet, amely ugyan azt állítja, hogy vannak annyira szép nők, akiket nem lehet szavakkal leírni, ám éppen hogy egy nagy ívű és szemléletes leírását adja ennek a szépségnek: „Vannak nők, *kikről* hívebb fogalmat adhatunk az olvasónak, ha nem is kisértjük meg leírásukat; ha fantáziájokra bízzuk elképzelni ama női eszményképet, *melyet* lelke egyszer megálmodott […] Mert e nő szépsége olyan volt, *melyet* nemcsak szemünkkel, elménkkel, hanem lelkünk, kedélyünk minden ízével élvezünk. (…) ha bámuljuk Vénust, *melynek* szépsége a vonások szabályos tökélyében rejlik, vajjon nem kell-e vele egy polczra helyeznünk az oly nőt, *kinek* szépsége nem a vonások anyagában, hanem az arcz szellemében, kifejezésében, a tekintet bájában s mosolya tiszta költészetében vakítólag lép elénkbe? […] E mosoly, *mely* szemében, ajkán s arczának minden izében honolt, ragyogott, mint a nap, s káprázatos glóriával vette körül az egész alakot, *melynek* fényében a körülötte sürgő fekete frakkok úgy tüntek föl előttem, mint egy csapat oktalan moly, *mely* meggondolatlan ösztönét követve, körülrepkedi a tüzet, hogy lángjában megperzselve szárnyait, lehulljon."

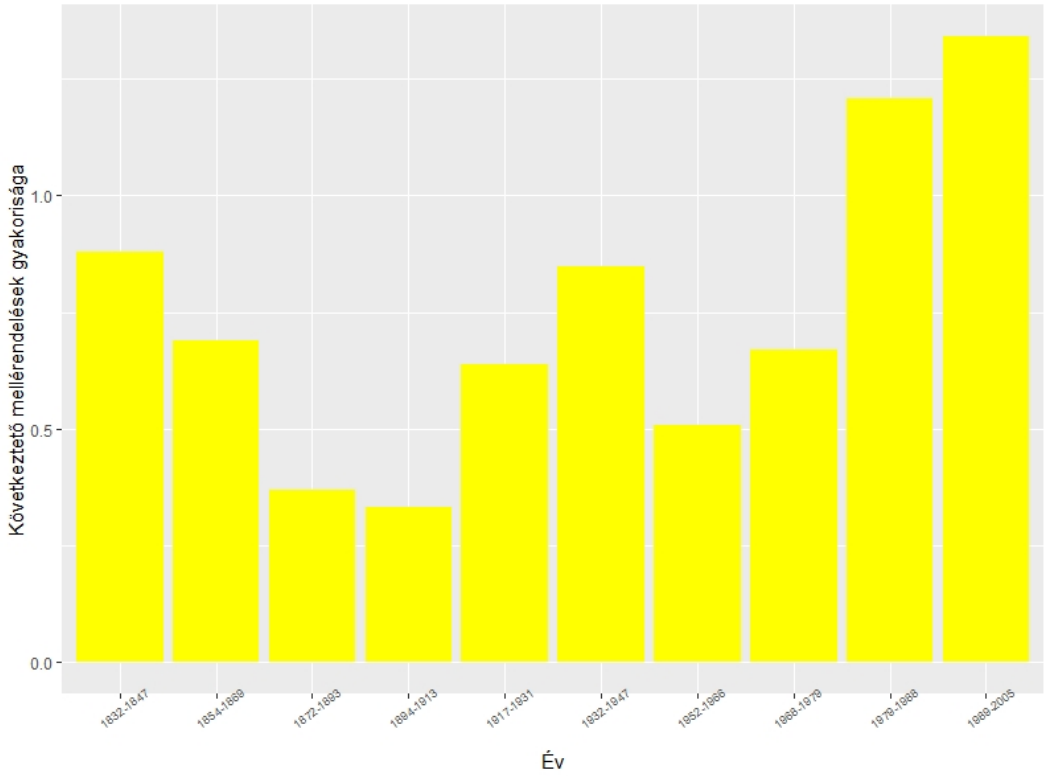
 DOI: 10.31400/dh-hun.2021.4.2415

datok közötti viszonyt. Ezek a tagmondatok akár önálló mondatokként is szerepelhetnének a szövegben, a kötőszavakkal történő összedolgozásuk a közöttük lévő logikai és/vagy ok-okozati összefüggést erősítik. Gaál József *Szirmay Ilona* című regényében az ellentétes, a következtető és megengedő, Nagy Ignác *Magyar titok* című regényében a következtető, a választó és a megengedő, Fáy András két regényében és Kuthy Lajos *Hazai rejtelmek* című regényében a választó, Vas Gereben *Nagy idők, nagy emberek* című regényében pedig a következtető mellérendelések nagy száma figyelhető meg a korpusz egészéhez képest. A tagmondatok ilyen összefűzése kevésbé szerkesztett és szabályos prózanyelvet hoz létre, ám lehetővé teszi az események dinamikus egymásra következését, valamint a motivációk és élethelyzetek részletes kifejtését. Olyan hosszúmondatos poétika ez, amely az 1970-es évektől jelenik meg ismét a magyar prózairodalomban – sőt még jellemzőbb lesz szerzők egy csoportjára, ahogy az a következtető mellérendelések gyakoriságának alakulását bemutató grafikonokon is látható. *(2.3.*[34] és *2.4. ábra)*



2.3. ábra. A következtető mellérendelés változása, 1832–2005.

[34] Az egyszempontos varianciaanalízis (ANOVA) alapján csak a három nagyobb időszak (1832–1860, 1861–1874 és 1975–2005) között mutatható ki korreláció a következtető kapcsolattípus gyakorisága és a megjelenési évek között (F-érték: 21.05), azaz a kirajzolódó tendencia által jelölt különbségek csak ezek között az időszakok között tarthatók statisztikailag szignifikánsnak.



2.4. ábra. A következtető mellérendelések gyakoriságának változása, évcsoportok.

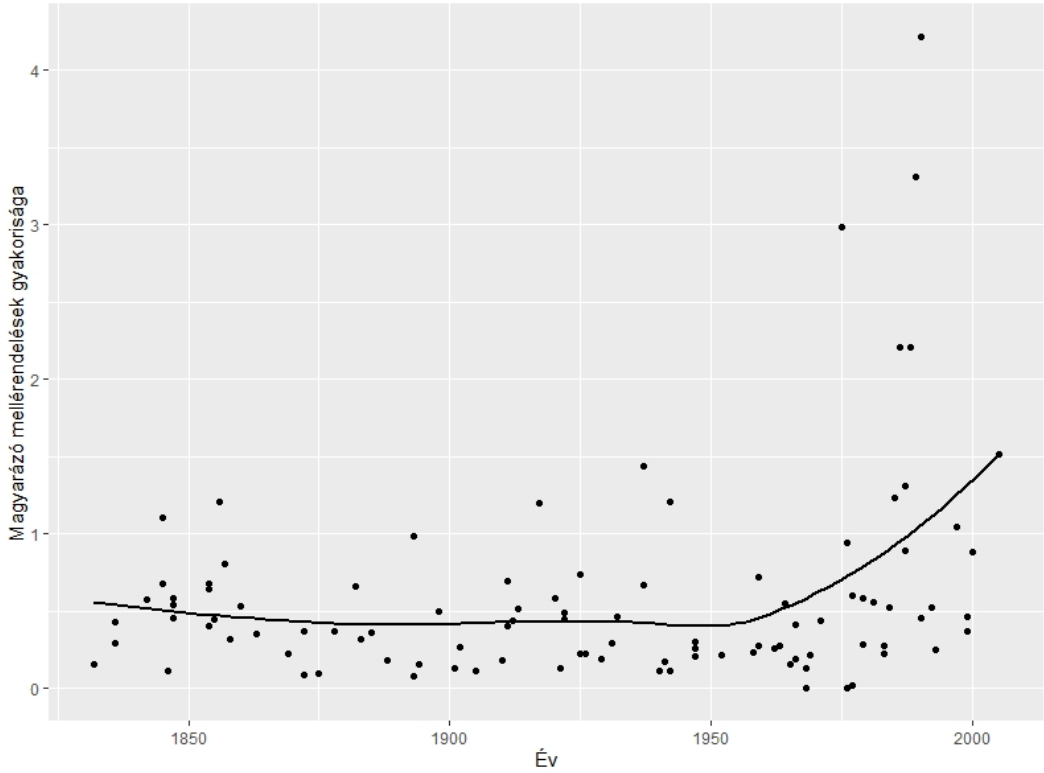
Ezek az ábrák és értékek kapcsolatot teremthetnek a két korszak regényirodalmának egyes képviselői között (és akár a stílustörténet ciklikusságára is felhívhatják a figyelmet), ám érdemes rávilágítani a köztük lévő különbségekre is. A következtető kötőszavak a 19. században legtöbbször az előzményekből következő eseménysort vezetik fel és egy cselekvő motivációját dolgozzák ki. Az egymásból következő események jelölése azért is szükséges, mert sokszor többszörösen összetett, bonyolult szerkezetek közötti viszony válik ezáltal könnyebben átláthatóvá.[35] Ugyanakkor már a 19. században megjelenik a következtetések ironikus-parodisztikus használata is, amelyek inkább meghökkentő, egymásból logikailag nem feltétlenül következő kapcsolatokat dolgoznak ki következtető mellérendelések vagy ok- és célhatározói alárendelések formájában.[36]

[35] Például: „A következő éjszaka, az után, melyet most leírtunk, fényesen volt a teli holdtól világítva, s már éjfél felé annyira haladtak a haramiák a várvíváshoz megkivántató szerek és eszközök elkészítésével, hogy ahhoz magához fogni lehetett; megparancsolá *tehát* Pintye Gregor, hogy az ágyúkat még az éji homályban a kitűzött helyekre vonják, oda hordván egyszersmind a már tetemes számmal öntött golyókat és puskaport, melyben a rablók hiányt nem szenvedtek." (Gaál József, *Szirmay Ilona*)

[36] Például: „Az árkokban rothadó víz és szemét igen ocsmány bűzt terjeszt és ez rendkívül hasznos, *mert* a faluról bejövő uraságokat arra figyelmezteti, hogy mielőbb siessenek ismét vissza egészséges falusi levegőjökre, hol a csődületi betegség nem érheti őket el oly könnyen, mint Pesten." (Nagy Ignác, *Magyar titok*)

 DOI: 10.31400/dh-hun.2021.4.2415

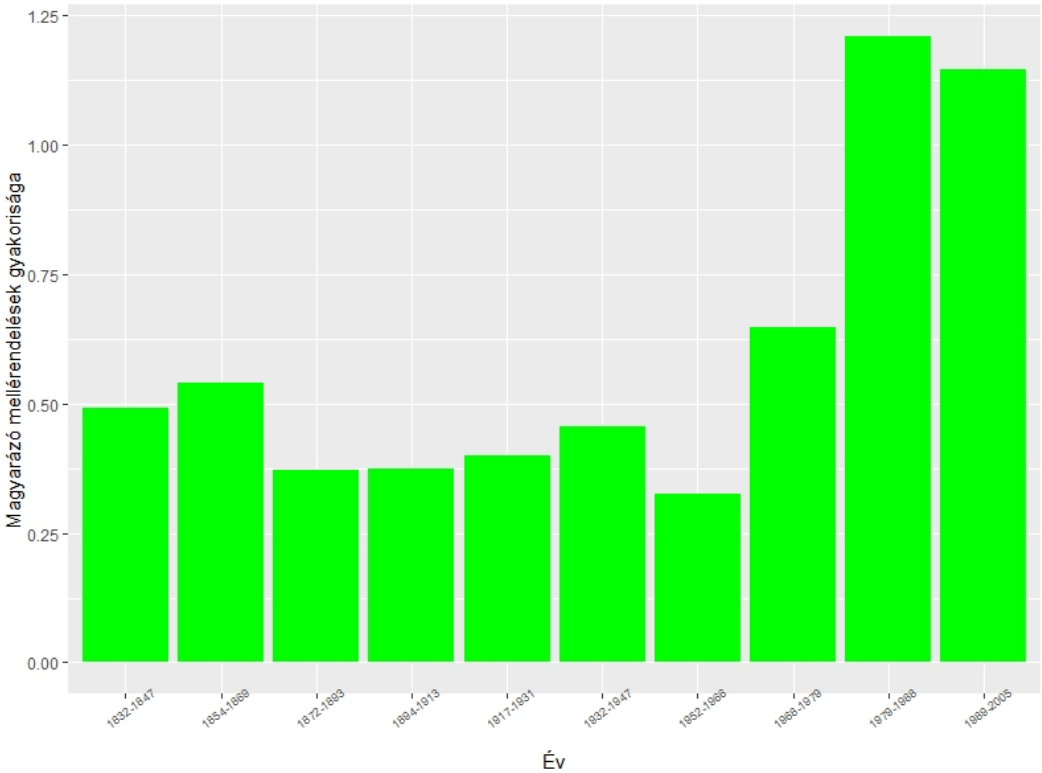
A vizsgálat fókuszát az időtengely másik végére helyezve megállapíthatjuk, hogy a következtetésekben rejlő ilyen potenciált aknázzák ki a 20. század utolsó harmadában jelentkező hosszúmondatos prózák is, amelyek, miközben részletes összefüggéseket képesek felvázolni a többszörösen összetett szerkezetekben, ezeknek az összefüggéseknek a megalkotottságára, esetleg abszurditására is felhívják a figyelmet. Ez a megalkotottság a 20. század második felének mondatszerkezeteiben összefügghet a kor irodalmának azzal a felismerésével, miszerint az eseményeket bemutató prózanyelv is egy beszédmód a sok közül, amely egy eleve valamiként értelmezett világot képes csak ábrázolni.[37] Ezt nem csak a következtető, hanem a magyarázó mellérendelések nagy számával érik el – olyan jellemző ez, amely a 19. századi regényirodalomból hiányzik. *(2.5.*[38] és *2.6 ábra)*



2.5. ábra. A magyarázó mellérendelés változása, 1832–2005.

[37] Vö. Kulcsár Szabó Ernő, *A magyar irodalom története 1945–1991* (Budapest: Argumentum Kiadó, 1994), 479–496.

[38] Az egyszempontos varianciaanalízis (ANOVA) alapján egy gyenge, de egyértelmű korreláció mutatható ki a vonatkozói kapcsolattípus gyakorisága és a megjelenési évek között (F-érték: 6.864), azaz a kirajzolódó tendencia által jelölt különbségek statisztikailag szignifikánsnak tarthatók.



2.6. ábra. A magyarázó mellérendelések gyakoriságának változása, évcsoportok.

A magyarázatok és következtetések halmozása összefüggést létesít a tagmondatok között, leginkább pontosítva a kapcsolatot megelőző tagmondat jelentését, ugyanakkor – éppen a következtetések és magyarázatok nagy száma miatt – ezeknek az összefüggéseknek a konstruáltsága, természetellenessége vagy pontatlansága (hiszen bármeddig pontosíthatók) is a figyelem előterébe kerül; „ez pedig nem más, mint magyarázat, a magyarázatok, mi több, a tényeket agyonmagyarázó, *vagyis* végső soron e tényeket megsemmisítő, de legalábbis elköd ösítő magyarázatok halmaza" – ahogyan Kertész Imre *Kaddis*ának elbeszélője fogalmaz. Kertész Imre mindhárom, a korpuszban szereplő regénye *(Sorstalanság, A kudarc, Kaddis a meg nem született gyermekért)* az első öt hely valamelyikén található a magyarázó mellérendelések gyakoriságát tekintve, és szintén magas értékeket vesznek fel a következtető kapcsolattípusok esetében. Ez a regények tematikus összefüggésén túl a köztük lévő stiláris hasonlóságra hívja fel a figyelmet, amelyet eddig kevésbé tárgyalt a szakirodalom.[39] A stiláris és tartalmi szintek azonban nem elválaszthatók egymástól: az említett Kertész-művek az egyes eseményeket a társadalmi rend működésének szükségszerű következményeként írják le, céljuk ennek a működésnek a minél pontosabb magyarázata – legyen szó a holokauszt történetéről, az 50-es évek vagy a Kádár-korszak Magyarországáról. A *Sorstalanság*ban Köves felismerése éppen az, hogy története nem érthető meg

[39] Hasonlóan magas értékek jellemzik a *Jegyzőkönyv, Az angol lobogó* című kisregényeket, sőt a *Gályanapló* című naplóregényt is, amely újabb – ezidáig feltáratlan – összefüggésekre világíthat rá az életművön belül.

 DOI: 10.31400/dh-hun.2021.4.2415

utólagos nézőpontból, hanem az események egymásból következő, logikus rendjét kell bemutatnia, ha saját élményeiről kíván beszélni. Az utólagosságot elutasító utólagos tudás paradox egysége szervezi az egyszerre naplószerű és múlt idejű-visszatekintő narrációt,[40] és ez érhető tetten a mondatszerkezetekben is, hiszen a magyarázatok nemcsak a dolgok természetes egymásra következését mutatják be, hanem azok hátborzongató abszurditását is: „Már egészen más dolog azonban – tették rögtön hozzá – az Arbeitslager, *vagyis* munkatábor: ott az élet könnyű, a viszonyok és az élelmezés, járt híre, hasonlíthatatlan, s ez természetes is, *hisz* ott a cél is más elvégre" (Kertész Imre, *Sorstalanság*). Ritkán hangoztatott jellemzője a Kertész-prózának, ám az okokozati kapcsolatok ilyen természetellenessége egyben – ha mégoly sötét és kínzó módon, de – humoros hatást is kiválthat. Ugyanígy a *Kaddis* elbeszélője is a világrend alapvető működéséről számol be, amely működés a legszörnyűbb pillanatokban is magyarázható és érthető (hiszen „amire tényleg nincs magyarázat, az nem a rossz, ellenkezőleg: a jó"); a hosszú magyarázatok egyszerre gyűrik maguk alá az olvasót és győzik meg igazukról, valamint egyszerre hatnak parodisztikusnak, helyenként humorosnak, sőt akár patologikusnak,[41] mániákusnak, vagy logikailag inkonzisztensnek,[42] de mindenképp magukban hordozzák a megértés, a magyarázat és az (ön)felszámolás ellentmondásos kapcsolatát:

> […] ez a kérdés te vagy, pontosabban én vagyok, de általad kérdésessé téve, még pontosabban (és ezzel nagyjából doktor Obláth is egyetértett): az én létezésem a te léted lehetőségeként szemlélve, vagyis én mint gyilkos, ha a pontosságot a végletekig, a képtelenig akarjuk fokozni, és némi önkínzással ez meg is engedhető, hiszen, hál' isten, késő, mindig is késő lesz már… (Kertész Imre, *Kaddis a meg nem született gyermekért*)

Hasonló figyelhető meg a 20. század utolsó harmadában íródott más regények esetében is, amelyek nagy számban dolgoznak ki magyarázó és következtető kapcsolatot, például Nádas Péter *Emlékiratok könyve* című munkája és Krasznahorkai László két regénye, elsősorban *Az ellenállás melankóliája* (de említhető a korpuszból Esterházy Péter *Harmonia cælestis* vagy Spiró György *A jövevény* című műve is). *Az ellenállás melankóliájá*ban a korpuszon belül a második leggyakoribbak a magyarázó és harmadik leggyakoribbak a következtető típusú kapcsolatok (és magas az ok- és célhatározói alárendelések száma is). Az ezekkel kidolgozott ok-okozati viszonyok leginkább szereplői tudattartalmakhoz kötődnek, azaz – a *Sorstalanság*hoz hasonlóan – az egyes szereplők alkalmazkodásáról értesítenek az egyre inkább fenyegető mindennapokhoz, ahogyan érthetővé és magyarázhatóvá kívánják tenni környezetük működését:[43] „Mindez már

[40] „A szöveg megkísérli csinálni is, amit mond, kétféleképpen beszéli el egyszerre a történetet, visszatekintő és szinkrón perspektívából." Vári György, *Kertész Imre:* Buchenwald fölött az ég (Budapest: Kijárat Kiadó, 2003) 57.

[41] Vö. Radnóti Sándor, „Auschwitz mint szellemi életforma," *Holmi* 3, 3. sz. (1991): 372–378, 377.

[42] Vö. Lengyel Imre Zsolt, „»egy lekerekített, üvegkemény tárgy.« Kísérlet a *Kaddis a meg nem született gyermekért* határainak olvasására," in Gintli Tibor, Kelemen Pál, Kulcsár-Szabó Zoltán és Szemes Botond, szerk., *Kertész másként* (Budapest: Kijárat Kiadó, 2021), megjelenés előtt.

[43] A Kertész esetében megállapított parodisztikus-humoros hatás Krasznahorkai stílusára is érvényes – ugyan bírálólag, de erre hívja fel a figyelmet például Horkay Hörcher Ferenc is, amikor megjegyzi,

senkit sem lepett meg igazán, *hiszen* az uralkodó állapotok természetesen ugyanúgy éreztették hatásukat a vasúti közlekedésben, mint bármi másban." Krasznahorkai korai prózájában „tudatkövető" technika érvényesül, amely a szereplők belső vagy kimondott beszédét dolgozza össze,[44] amely során a szereplők következtetései és motivációi felől ismerjük meg az ábrázolt világot – például: „[…] az »alapos szemügyre vételt« korántsem könnyítette meg, a tömeg nagyságára vonatkozó kérdéssel *ugyanis* (*hisz* mit figyeljen meg rajta, ha minden változatlan?) sehogy sem boldogult…" Ezáltal az ábrázolt világ összefüggései „belülről", a szereplői tudatokon keresztül bomlanak ki, aminek köszönhetően átélhetővé válnak az amúgy abszurdnak tűnő eseménysorok is.

Nádas Péter *Emlékiratok könyve* című regényében (amelyben arányosan az ötödik legtöbb magyarázó és második legtöbb következtető összetétel található a korpuszban) a következtetés és magyarázat halmozásának e kettős tulajdonsága, azaz az ok-okozati összefüggések feltárása és egyben természetellenességük, megalkotottságuk színre vitele főként abban érhető tetten, ahogyan az elbeszélők részletes pontosságra törekvő leírásai a helyzetek bármeddig pontosítható, így lezárhatatlan magyarázatának érzetét keltik, és látványosan bonyolult viszonyrendszereket hoznak létre:

> […] egyáltalán nem nagy baj, ha ezt gondolom róla, mert ő meg tudja rólam, amit szeretnék titokban tartani, hiszen Thea előtt Melchior igazán nem fegyelmezte a tekintetét, amit tehát titoknak szántunk, mármint, hogy nem egyszerűen jó barátok, hanem szerelmesek vagyunk, ez bizonyára előtte sem titok. (Nádas Péter, *Emlékiratok könyve*)

A hasonlító kapcsolatok (időbeli változás)

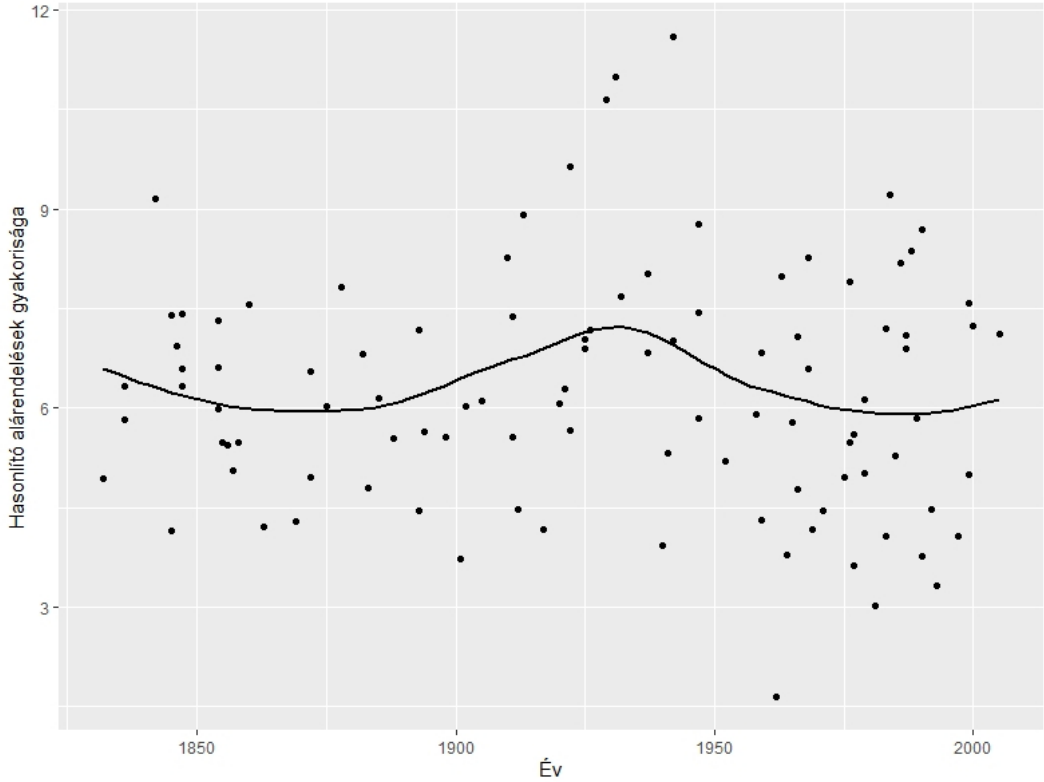
Az eddigi tendenciák főként a hosszúmondatos prózákat érintették, és az idővonal két végén mutattak kiugró értékeket. A hasonlító kapcsolatok relatív gyakorisága ezzel szemben a 20. század első felében éri el a csúcspontját. *(2.7.*[45] és *2.8. ábra)* A korpusz alapján elmondható, hogy a korszak prózája ritkábban dolgoz ki kapcsolatot a tagmondatok között, és a viszonyok létesítése helyett inkább az önálló jelenetek vagy képek egymás mellé helyezésével kívánja a helyzetek leírását adni. Ez alól azonban kivétel a hasonlatok alkalmazása, amelyek éppenhogy távoli területek összekapcsolására képesek – például: „[…] a gyeplőszár olyan egyhangúan rángatózik, *mint* a falusi temetőben reszket a nyárfa" (Krúdy Gyula, *A vörös postakocsi*). A korpuszból Márai Sándor *A gyertyák csonkig égnek,* Gelléri Andor Endre *A nagymosoda,* Szomory Dezső *A párizsi regény* című műveire és Krúdy Gyula regényeire jellemző leginkább ez a kapcsolattípus.

hogy Krasznahorkai körmondatai finoman szerkesztettek és szépek, „azonban ez az erősen artisztikus nyelv alkalmanként a paródia határára sodorja az írói elbeszélésmódot." Horkay Hörcher Ferenc, „A pusztulás anatómiája," *Holmi* 2, 5. sz. (1990): 576–580, 579.

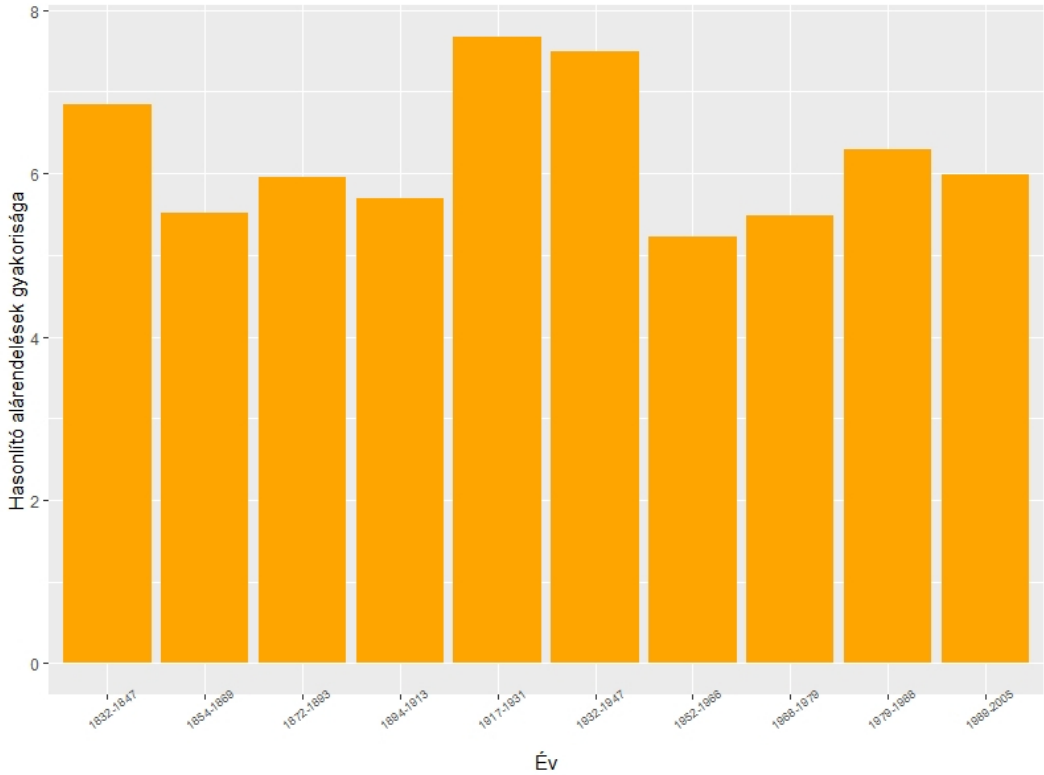
[44] Szirák Péter, „Látja, az egész nincsen: Krasznahorkai László prózájáról," *Alföld* 39, 7. sz. (1993): 41–51, 47.

[45] Az egyszempontos varianciaanalízis (ANOVA) alapján csak a három nagyobb időszak (1832–1909, 1910–1949 és 1950–2005) között mutatható ki korreláció a következtető kapcsolattípus gyakorisága és a megjelenési évek között (F-érték: 12.67), azaz a kirajzolódó tendencia által jelölt különbségek csak ezek között az időszakok között tarthatók statisztikailag szignifikánsnak.

 DOI: 10.31400/dh-hun.2021.4.2415



2.7. ábra. A hasonlító alárendelés változása, 1832–2005.



2.8. ábra. A hasonlító alárendelés gyakoriságának változása, évcsoportok.

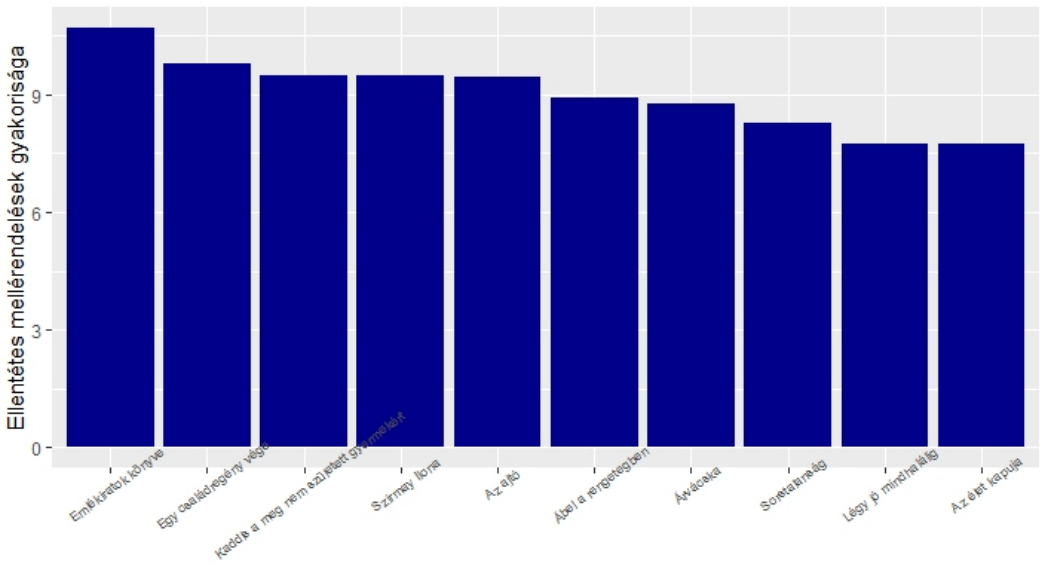
Ezek a grafikonok újabb, időben is elkülöníthető paradigmát jelölnek ki. Míg a vonatkozói alárendelések nagy száma a 19. század első felében, a magyarázó-következtető kapcsolattípusoké pedig a 20. század végén (illetve részben a 19. század első felében) figyelhető meg, addig a 20. század első felében a hasonlatok alkalmazása a jellemző.

Ezek a stiláris, mondatszerkesztési paradigmák egyúttal a világhoz fűződő, episztemológiai viszony történeti különbözőségeire is utalnak. A vonatkozói alárendelések a mondatok egy szereplőjének adják részletesebb meghatározását azáltal, hogy egy új jelenetben szerepeltetik, ami egyúttal a mondat jelentésének rögzítését is szolgálja. Ezzel szemben a folyamatos pontosítás és kifejtés állandóan új értelmezések felé nyitja meg a mondatot, így a világhoz kevésbé stabil hozzáférést tételez fel, ugyanakkor inkább képes azt belső perspektívából megfigyelhetővé tenni. A harmadik paradigmát jelentő hasonlatok két terület összehasonlítása révén tudnak új információt szolgáltatni – ebben az esetben nem a jellemzés vagy a kifejtés, hanem egy analógia létrehozása („Szép vagy, mint a rózsaszál”), vagy a különbségek artikulálása („Erősebb, mint a te apukád”) játszik központi szerepet.

Eredmények 2.

A regényekben az egyes kapcsolattípusok gyakoriságát nemcsak idővonal mentén, hanem összehasonlító módon, oszlopdiagramok formájában is ábrázolhatjuk (ahogyan eddig a regényenkénti évcsoportok esetében). Ekkor nem a történeti tendenciák, hanem az egyes művek sajátosságai válnak megfigyelhetővé. A csökkenő sorrendbe rendezett ábrák kezdeti és záró szakaszai bírnak a legnagyobb információértékkel, hiszen a legnagyobb és legalacsonyabb értékeket mutató regényekről mondható el, hogy az adott típus jobban/kevésbé jellemző rájuk a korpusz egészéhez képest.

Ellentétes kapcsolatok (legjellemzőbb regények)



3.1. ábra. Ellentétes mellérendelés – Az első 10 regény.

 DOI: 10.31400/dh-hun.2021.4.2415

A *3.1. ábrá*n az ellentétes mellérendeléseket leggyakrabban kidolgozó regények láthatók. Feltűnő, hogy Nádas Péter két regénye található az első két helyen, azaz ezekben a szövegekben a legnagyobb az ellentétes tagmondatkapcsolatok relatív gyakorisága. Az *Egy családregény vége* – amelyben gyakoriak a feltételes és ok-/célhatározói kapcsolatok is, lásd *Függelék* 1.1. – a gyerek elbeszélő perspektívájából mutatja be az 1950-es évek Magyarországát, amely perspektívát elsősorban a felnőtt–gyerek, a képzelet–valóság, az egyén–családi/politikai közösség feszültsége jellemez a legjobban. Ezek az ellentétek szervezik az elbeszélő világát, aki azáltal mutat rá környezete ellentmondásaira, hogy megpróbál alkalmazkodni hozzájuk (ez a kettősség nagyon hatásosan képes kiaknázni a gyerekperspektívában rejlő poétikai lehetőségeket), mígnem a regény végére a teljes elutasításhoz jut el – ezt a szembehelyezkedést érzékelteti a szöveg híres zárómondata is: „Nem." Az ellentétek az eltérő horizontból megszólalók szólamainak különbségeiből is adódnak, ugyanis a gyerek elbeszélőével a bibliai hagyományokig visszanyúló nagypapa és a pártfunkcionárius apuka szólama is keveredik a szövegben. Mindennek megfelelően több értelmező is rámutatott a regényt szervező oppozíciók működésére;[46] amelyet Bazsányi Sándor a János evangéliumából vett mottóban megfogalmazott feszültségre vezet vissza, („És a világosság a sötétségben fénylik, / *de* a sötétség nem fogadta be azt."), miközben a nagypapa által elmesélt történetrészlet és a mottó szoros kapcsolatára is felhívja a figyelmet:[47]

> Egy pillanatra, míg valahová taszítják, lökik, látja a középső meztelen szemét; nem a szemet, csak a pillantás fényét rögzíti; és amikor sötét, forró szobájában imádkozik, s várja a hírt, hogyan vélekedik a történtekről a Szanhedrin és rabbi Abjatar, és imádkozik, azt kéri imáiban az Úrtól, hogy világosítsa meg elméjét a történtek felől, akkor a nagy sötétségben, amiben fuldokol, néha felsejlik ennek a pillantásnak a fénye, de nem érti a fényt; Simon nem érti, sötétsége a fényt nem tudja befogadni, pedig a fény a sötétségben világít; abba kéne kapaszkodnia, abba a fénybe, de ő nem érti és világosságért imádkozik az Úrhoz, de közben nem veszi észre a fényt, annak a tekintetnek a fényét, amit az Úr elküldött.

Az idézetben (ahogy a mottóban is) a fény és a sötétség ellentéte mellett a megértés/befogadás versus nem-értés/nem-befogadás ellentéte, illetve a cselekvések eredményeinek különbsége (például az imádkozás, amely nem talál meghallgatásra) válik hangsúlyossá, amely ellentéteket elsősorban ellentétes tagmondatkapcsolatok dolgoznak ki – ez a stiláris-tematikai összefüggés a szöveg alapvető jellemzőjének tekinthető.

Balassa Péter monográfiájában arról ír, hogy a *Családregény* igazi tétje „az egység és egyben látás megőrzésében létrejövő *vagy* elpusztuló *személyiség* (ezt vette észre a korabeli kritikák többsége, amikor többször az „igen" és a „nem" oppozíciójával írta le a könyvet),"[48] és maga is oppozíciópárok fontosságára hívja fel a figyelmet – például a regényben többször felbukkanó és a zárlatra motívummá váló fekete-fehér kövezetet is

[46] Kálmán C. György szerint például a különböző szólamok összeütközésének köszönhetően a „szerkezet szétfeszülése" figyelhető meg a Bazsányi Sándor által „tagadáselvű regény"-ként jellemzett szövegben. Kálmán C. György, „Csalárd regény: Újraolvasás", *Jelenkor* 44, 10. sz. (2002): 1051–1056, 1052; Bazsányi Sándor, *Nádas Péter* (Pécs: Jelenkor Kiadó, 2018), 62.

[47] Uo., 60.

[48] Balassa Péter, *Nádas Péter* (Pozsony: Kalligram Kiadó, 1997), 97, kiemelés az eredetiben.

mint a művön végigvonuló ellentétek metaforáját azonosítja: „részben a fekete-fehér egyértelműség és a véletlen kombinációjának játszmája ez, részben a két szín poláris ellentétével való játszadozás, mely a mű egész kontextusában az igen/nem, élet/halál bináris oppozíciójának, nyers kérdezz-felelek-jének a metaforája."[49] Elemzése az ezeket összefogó, végső, és az egész Nádas-prózát meghatározó ellentétre fut ki, azaz magában az ellentétben rejlő kettősségre, miszerint az *Családregény* mindenekelőtt a szembenállást felszámoló *dialógus* és a valódi kapcsolódást nélkülöző *dualitás* oppozícióját, dilemmáját viszi színre.[50] Másképpen megfogalmazva: Nádas első regényének alapkérdése a tőlünk elválasztott Másikkal való azonosulás lehetősége és a felszámolhatatlan különválasztottságunk bizonyossága közti feszültségre vonatkozik. Ez a dilemma az ellentétes tagmondatkapcsolatok gyakori alkalmazásával kerül kifejtésre, azaz ebben az esetben egyértelmű kapcsolat tételezhető a szöveg tematikus és stiláris szintjei, a történetet szervező ellentétek és az ellentétes tagmondatkapcsolatok nagy száma között. Sőt, ahogy már a kifejezésben rejlő oximoron (ellentétes kapcsolat) is tanúsítja, maga a grammatikai szerkezet is hordozni képes a Balassa által megfogalmazott dilemmát összekapcsolódás és különbözőség kettőségéről: az összetett mondat e fajtája ugyanis egyszerre állítja szembe és köti össze a tagmondatokat.

Az ellentétes kapcsolatok nagy száma az *Emlékiratok könyve* hasonló szerveződéséről értesít, topológiai-logikai karakterének szintén fontos összetevője az ellentétezés. A nagyregény hosszú, többszörösen összetett mondatait eddig inkább azok – tagadhatatlanul erős – hatása felől elemezték[51] vagy a világ leírásának egy sajátos módozataként azonosították,[52] ám ritkán vizsgálták tüzetesebben a mondatok szerkezetét. Ez alól kivétel Jolanta Jastrzębska nyelvészeti megközelítésből írt, fontos megfigyeléseket tartalmazó tanulmánya, amelynek általános megállapításai azonban nélkülözik a kvantitatív alapozást.[53]

Ezidáig a regényben nagyszámú következtető és magyarázó mellérendelések kerültek tárgyalásra, amely a részletes pontosságra és az összefüggések feltárására törekvő elbeszélő[54] stílusának fontos jellemzőjére világított rá. Mindezt az ellentétes kapcsolatok gyakoriságával egészíthetjük kis. Fontos látni, hogy az ellentétek és a következtetések/magyarázatok különböző funkciót látnak el a mondatokon belül. A mondatok szerkezete ugyanis olyan alapsémára vezethető vissza, amely nagyobb gondolati egységek részletes kifejtéseiből és ezeknek az egységeknek a szembeállításából jön létre. A részletes kifejtés a következtetések és magyarázatok mellett alárendelő kapcsolatokkal és mondatrészek halmozásával[55] történik. A halmozás az érzékek (nem értelmező

[49] Uo., 122.

[50] „Dialógus és dualitás ellentéte az igazi, általános oppozíció az egész életműben." Uo., 139.

[51] Például: Boros Gábor, „Nagy fúga: Adalék az *Emlékiratok könyve* értelmezéséhez," in Balassa Péter, szerk., *Diptychon: Elemzések Esterházy Péter és Nádas Péter műveiről,* 169–179 (Budapest: Magvető Kiadó, 1988).

[52] Bagi Zsolt, *A körülírás* (Pécs: Jelenkor Kiadó, 2005).

[53] Jolanta Jastrzębska, „Körmondatos stílus a modern magyar prózában. Nádas Péter: *Emlékiratok könyve* példáján," *Magyar Nyelv* 89, 1. sz. (1993): 26–40.

[54] Jelen elemzés nem tér ki az elbeszélők közötti különbségre és nem foglalkozik külön Krisztián mondatainak szerkezetével, amelyek bizonyosan torzítják a többi elbeszélő stílusáról mondottakat.

[55] A legtöbbször szinonim szavak halmozására többen felhívták már a figyelmet, vö. Jastrzębska, „Körondatos stílus," 29–30; Bernáth Árpád, „Mi hároman", in *Diptychon,* 137–157. A kutatás külön

 DOI: 10.31400/dh-hun.2021.4.2415

jellegű) működésével hozható kapcsolatba és inkább a szöveg érzéki dimenzióit erősíti, míg a magyarázatok és alárendelések az értelmi összefüggések feltárásáért felelnek.[56] Az ellentétes kapcsolatok többek között ezeknek a részletes kifejtést tartalmazó nagyobb egységeknek a szembeállításáért felelősek – az egységek elhatárolását gyakran pontosvesszők is erősítik. Terjedelmi okok miatt most csak egyetlen hosszú mondatot idézek, amelyben jelölöm a szembeállított gondolati egységeket (és amelyben a második egység éppen a „részletek részleteiben elvesző analízissel" szemben határozza meg magát).

> [1] Ám akkor beszélnem kell arról is, az igazság nevében és a teljesség kedvéért igyekezve eltüntetni azon látszatot, ami ezidáig keletkezhetett, miszerint egész akkori életem csupán nyúlós keservekből, szégyenteljes kegyetlenkedésekből, siralmas vereségekből, elviselhetetlen, igen, elviselhetetlen szenvedésekből állott volna, nem, beszámolóm kétségtelen egyoldalúságát ellensúlyozandó hangsúlyoznom, hogy nem és nem! mert az öröm és az élvezet természetesen legalább oly jellemző része volt, a szenvedés pedig talán csak azért hagy mélyebb nyomot, mert igénybe véve a tudat képességét a gondolkodásra, töprengéssel és rágódással megnyújtja az időt, míg az igazi öröm, mely minden tudatost megkerül, az érzékire szorítkozik, csupán annyi időt ad magának és nekünk, amíg létezik, miáltal végzetesen véletlenszerűnek és esetlegesnek tűnik, különállónak és elszakítottnak a szenvedéstől, s így a szenvedés hosszú és bonyolult történeteket hagy az emlékezésnek, míg az öröm percnyi felvillanásokat; [2] de félre a részletek részleteiben elvesző analízissel s a részletek értelmét kereső filozófiával is! ami szükséges ugyan, ha szemügyre akarjuk venni, hogy lelkiekben milyen gazdagok vagyunk, s miért ne vennénk szemügyre élvezettel? ám éppen mert e gazdagság oly végtelen, s a végtelen a leginkább megfoghatatlan dolgok közé tartozik, analitikus sietségünkben hajlamosak vagyunk oknak, sérüléseink, torzulásaink, szenvedéseink, lelki betegségeink, mondjuk ki, hogy nyomorékságunk kiváltó okának tartani egészen egyszerű és végső soron természetesnek mondható történéseket, mivel éppen e történések egészét veszítettük el bizonyos önkényesen kiválasztott részletek javára a szemünk elől, a részletek végtelen gazdagságától való rettenetünkben megálljt parancsoltunk magunknak ott, ahol még jócskán tovább lehetne menni, rettenetünk bűnbakot csinál, kis áldozati oltárokat emel, szertartásosan szúr levegőbe késeket, s így nagyobb zavart okoztunk, mintha egyáltalán nem is gondolkodtunk volna magunkról, bizony boldogok a lelki szegények! ne gondolkodjunk hát! adjuk át magunkat szabadon és minden elfogultságtól mentesen annak a kellemes

vizsgálta a 100 regényben a halmozások gyakoriságát, amely eljárás az *e-magyar* által elemzett szövegekben az egymást követő, vagy vesszővel elválasztott melléknevek (elsősorban jelzős szerkezetek) és főnevek hármas vagy négyes sorozatának azonosításán alapult. Tehát a bonyolultabb szerkezeteket (pl. jelzős határozók, mint az alábbi példamondat elején) ezzel a módszerrel nem lehet azonosítani, így az eredmények is csak az egyszerűbb felsorolásokra, halmozásokra vonatkoznak. A jelzős szerkezetek halmozásokban kiugróan magas értéket mutatnak Kaffka Margit regényei, de az *Emlékiratok könyvé*ben is gyakoriak az ilyen szerkezetek, lásd a *Függelék* 3. pontjában.

[56] Vö. Szemes Botond, „Borítójáról a könyvet: Szellemi, anyagi és történeti meghatározottságok az *Emlékiratok könyvé*ben," *Kalligram* 27, 5. sz. (2018): 70–71.

> érzetnek, hogy anyánk betegágya mellett ülünk a földön, fejünket úgy hajtva testét takaró hűvös selyempaplanára, hogy ajkunk a karjának csupasz bőrén pihenjen, hajunkban érezzük finom ujjait, a hajtövek kellemes borzongását, bizsergését, mert a másik kezével beletúrt zavarában, mintegy vigasztalón, kipróbálva, hogy szavainak élét tudja-e venni a mozdulattal, s bár e borzongás lassan testünk egész felületét birtokába veszi, [3] az elhangzottakat mégsem lehet ilyen egyszerűen visszavonni: bennünk is fölrémlett már a sejtés, hogy az apánk talán nem is apánk, s ha ama bizonyos két férfi között ő nem tudott dönteni, akkor e sejtés most csaknem bizonyossággá erősödhetik, [4] erről viszont, megértjük, nem lehet többet mondani, legyünk csendben hát, s akkor pontosan érezni fogjuk, hogy mindaz, mi szavaink hatására emlékként fölvillant és immár elült bennünk, bármennyire fontos, bármennyire meghatározó legyen, csupán érzelmeink és valódi érdeklődésünk hátterében áll, mert abban a térben, melyben mindezt megpróbáljuk megragadni, sajátunkként mérlegelni, melyben valódinak nevezhető történéseink zajlanak, teljesen egyedül vagyunk, ahhoz nekik nincsen és nem is lehet semmi közük.

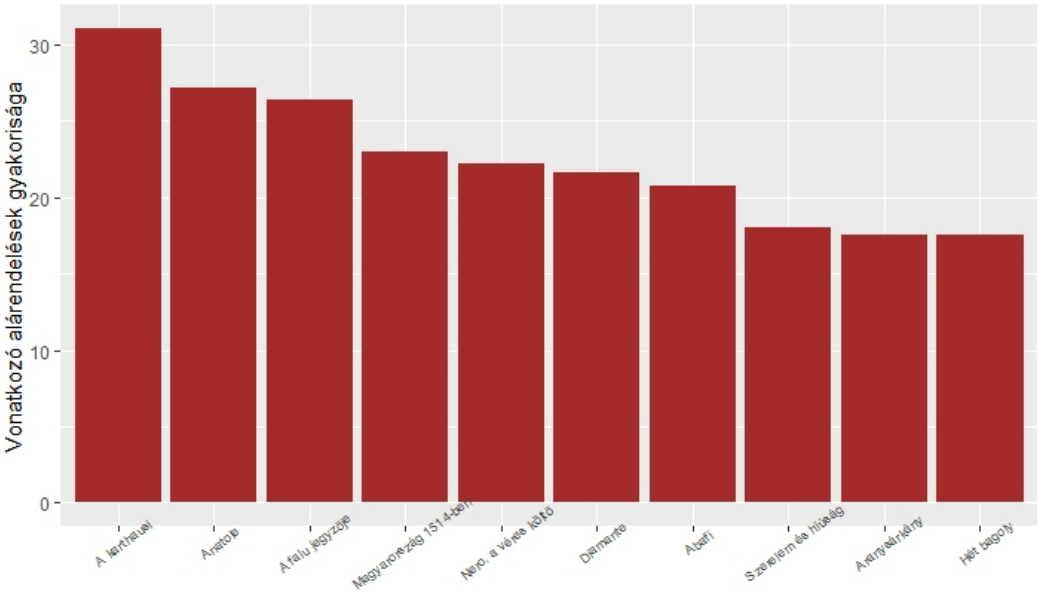
Az ehhez hasonló szembeállítások ezúttal is a regényt szervező oppozíciókkal hozhatók összefüggésbe. A látszat–valóság, a testi–akaratlagos/szellemi, az (el)képzelt–megvalósult ellentéteit dolgozzák ugyanis ki az ilyen ellentétes kapcsolatok, amelyek általában az elvárthoz, az érzethez, az elképzelthez vagy a korábbi emlékekhez képesti történéseket, gondolatokat vezetik fel. A regénynek ezt a szerkezetét erősítik a kisebb hatókörű ellentétes kapcsolatok is, amelyek a nagyobb gondolati egységeken belül létesítenek ellentétes viszonyt a tagmondatok között. Fontos látni, hogy az ellentétezések, ahogyan a *Családregény* esetében, itt sem a teljes elválasztásért felelősek, hanem a különbözőség és azonosság (dualizmus és dialógus) kettősségét hozzák játékba – ezáltal a regény központi problémáját[57] hordozzák magukban. Bagi Zsolt egy másik szöveghely elemzésekor hasonló következtetésre jut, amikor megjegyzi, hogy „egy »viszont« ellentételezése vezeti be a mondatba az új eseményt, az teremti meg benne helyét. Ez a »viszont«, amelynek rokonait oly gyakran fellelhetjük Nádas bekezdésnyi mondataiban, a mondatba mint test-írás alapegységébe egy olyan törést ír, amely a mondat felütésében adott szituációt anélkül borítja fel, hogy azzal való közösségét felmondaná. Ha egy mondat állna ellentétben az előzővel, akkor a törés elválasztaná a kettőt, és izolálná őket egymással szemben, így azonban bensőséges kapcsolatuk fennmarad; egymásba folynak, de megjelenik bennük egy vetemedés vagy gyűrődés, amely a folytonosságban a megszakítottságot képviseli."[58]

[57] Ez a probléma a személyes identitás és az interszubjektív világban átélhető közösségiség kapcsolatára vonatkozik, ahogyan azt Bagi Zsolt is megjegyzi a regény egy, a másikkal való azonosulás lehetőségét és lehetetlenségét ellentétes szerkezetek halmozásával körüljáró mondata kapcsán: Bagi, *A körülírás,* 63–65.

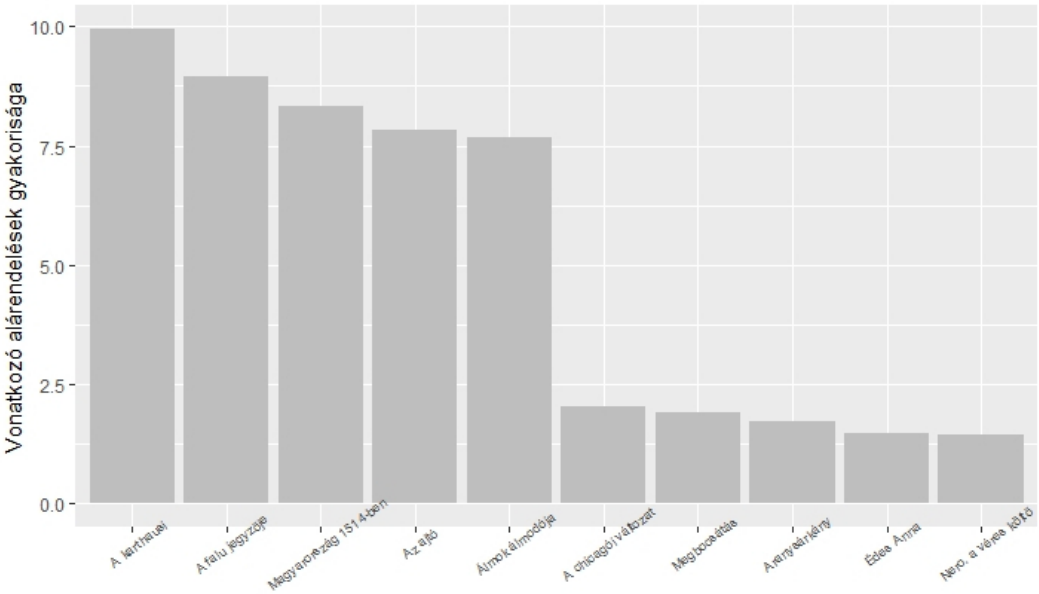
[58] Uo., 112.

 DOI: 10.31400/dh-hun.2021.4.2415

Prototipikus vonatkozói alárendelések és feltételes kapcsolatok (legjellemzőbb regények)



3.2. ábra. Prototipikus vonatkozói alárendelés – Az első 10 regény.



3.3. ábra. Feltételes alárendelés – Az első és az utolsó 5 regény.

A *3.2. ábra* a prototipikus vonatkozói, a *3.3. ábra* a feltételes alárendeléseket leggyakrabban, illetve legkevésbé gyakran alkalmazó regényeket mutatják. Eötvös József mindhárom regényében kifejezetten gyakorinak mondhatók ezek a kapcsolattípusok, ami a már tárgyalt anaforisztikus, ritmikus-retorikus prózastílusra vezethető vissza, hiszen az azonos pozícióban lévő mellékmondatok rendszeres halmozása nemcsak

a vonatkozói, hanem a feltételes mellékmondatokra is érvényes.[59] Ám korántsem szükségszerű a két típus ilyen összefüggése – ahogyan azt például Kosztolányi Dezső műveinél láthatjuk. Míg a *Nero, a véres költő* az ötödik, az *Aranysárkány* a kilencedik legnagyobb értéket veszi fel, és az *Édes Anna* is magas értéket mutat a vonatkozói alárendeléseket tekintve (*Függelék* 1.1.), addig e három szövegben a legkevésbé gyakoriak a feltételes tagmondatkapcsolatok. Fontos példa lehet ez arra, hogy nemcsak egy típus gyakorisága, hanem annak hiánya is beszédes lehet egy író stílusával kapcsolatban.

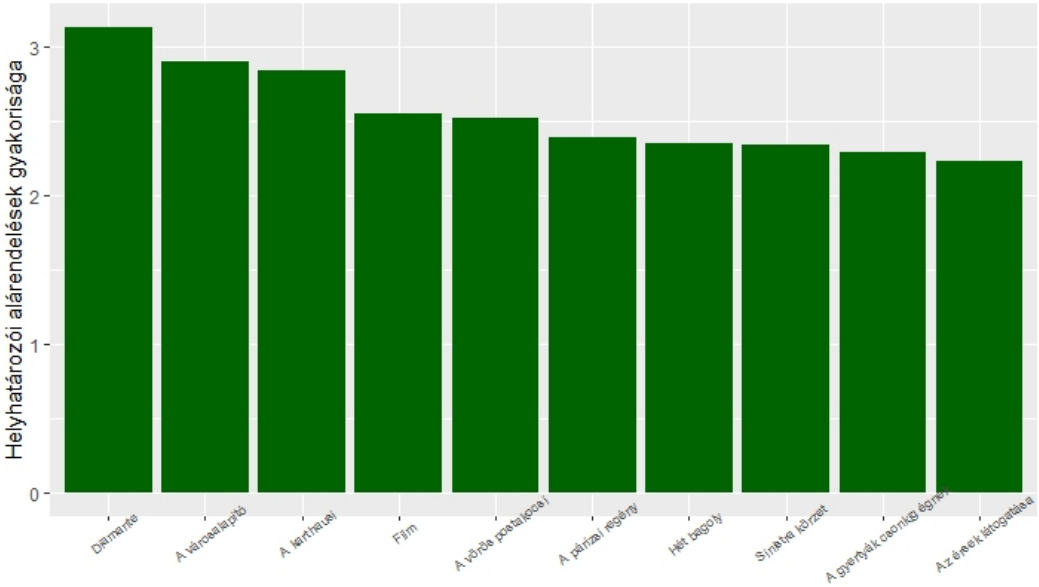
Kosztolányi rövidmondatos prózája mentes a tagmondatok halmozásától és a másod- vagy harmadfokú alárendelésektől, ahogyan a hiányos szerkezetek se jellemzők rá: éppen az egyszerű, hagyományos (névszói alanyok és igei állítmányok viszonyát kidolgozó) mondatok adják ennek a prózapoétikának az egyediségét.[60] A vonatkozói névmással bevezetett mellékmondatok, ahogy fentebb látható volt, a leírások pontosságát erősítik, hiszen azáltal, hogy egy szereplőt új jelenetben teszik megfigyelhetővé, annak jellemzéséhez és meghatározásához járulnak hozzá. Erre lehet példa a Senecától származó idézet a *Nero, a véres költő*ből, amelyben a szereplőket jellemző alárendelések egyben a regény gyújtópontjában lévő két, egymással nem összebékíthető szerepet (költő és császár) jelölik ki: „Hibát követtem el, tudom. Akkor a legnagyobbat, mikor odaadtam magamat Neronak, én, *aki* költő vagyok, őneki, *aki* csak császár." A vonatkozói névmásokkal kidolgozott mellékmondatok továbbá az elbeszélő külső, megfigyelő pozíciójához is hozzájárul; az átélt, szabad függő beszéd más kapcsolattípusokat is előnyben részesít, ám ezek nem jellemzők Kosztolányi vizsgált regényeire. Külön érdekes ebből a szempontból a feltételes alárendelések alacsony száma, amely a tárgyalt művek logikai karakterének alapvető jellemzőjének tekinthető: a regényekben a lehetőségek és feltételek számbavétele helyett a konkrét világ tárgyilagos leírására korlátozódik az elbeszélés. Ebben a konkrétságban jelölhetjük ki Kosztolányi prózastílusának alapját, amely stílus gazdagságát nem a mondatszerkezeti bonyolultság, hanem a pontos, érzékletes szóhasználat adja.

[59] Például: „S ha végre a vész kitört, ha minden vágyait s érzelmeit egy tárgyba pontosítva, ifjú lelkének egész hatalmával egy kedves lényt fogott körül; ha egy asszonyt, nem – ha egy angyalt talált, kit határtalan képzetének kincseivel fölruházott, s leborulva bálványa előtt, minden gyönyört s bánatot, mindent, mi szörnyű s fenséges életünkben, egyszerre, egy pillanatban érezte; ha istennek s elátkozottnak gondolva magát, egy pillanatban bánat- s örömének terhe alatt lesüllyedett, s ami keblét tölti, több, mint hogy magába rejteni, nagyobb, mint hogy kimondani tudná; ki találna szót ezen érzelme leírására?!" (Eötvös József, *A falu jegyzője*) Idézi Herczeg, „Eötvös körmondatai (Folytatás)," 175.

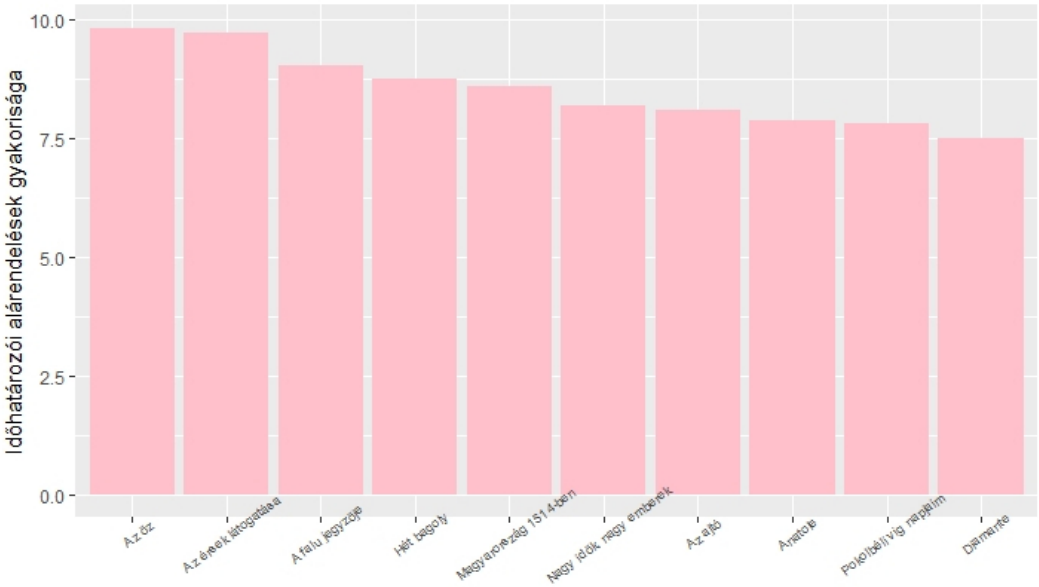
[60] Vö. Herczeg Gyula, „A prózaíró Kosztolányi," *Magyar Nyelvőr* 110, 4. sz. (1986): 432–449.

 DOI: 10.31400/dh-hun.2021.4.2415

Hely- és időhatározói kapcsolatok (legjellemzőbb regények)



3.4. ábra. Helyhatározói alárendelés – Az első 10 regény.



3.5. ábra. Időhatározói alárendelés – Az első 10 regény.

Röviden említhetjük még a *3.4.* és *3.5. ábrán* bemutatott hely-, illetve időhatározói kapcsolatok gyakoriságát is. Ezek a típusok a vonatkozói alárendelések aleseteinek tekinthetők, amelyek helybeli vagy időbeli viszonyok kidolgozásáért felelősek. A helyhatározói kapcsolatok nem meglepő módon azokban a regényekben a leggyakoribbak, amelyekben amúgy is magas a vonatkozói névmásokkal bevezetett mellékmondatok

száma, vagy amelyeknek középpontjában egy terület leírása áll: ilyen Konrád György *A városalapító,* Mészöly Miklós *Film* és Bodor Ádám *Sinistra körzet,* valamint *Az érsek látogatása* című regénye. Ez utóbbiakban az időbeliséget kidolgozó kapcsolatok is gyakorinak tekinthetők, hasonlóan Krúdy Gyula vizsgált regényeihez *(A vörös postakocsi, Hét bagoly),* amelyek szintén magas értékeket mutatnak mind a két kategóriában. Ez az események idő- és térbeli viszonyainak részletes kidolgozottságát jelöli, ami fontos szempontot kínálhat a regények stílusának leírásához.

Krúdy prózájában például a helyszínek és időpontok részletes jellemzése a különböző síkok egymás mellé rendezését és összehasonlítását szolgálják. Hely és idő gyakran nem is elválasztható regényeiben egymástól, amennyiben egy helyszín is a múltra vonatkozó emlékeket hordoz és az idősíkok összehasonlításának teremt alapot, például:

> A régi Arany Sas fogadóban szállott meg, ahol már egyetemi hallgató korában is lakott: akkor egy szobában, most kettőben, mert legényének is ott kellett aludnia, és éjjel levetkőztetni, lábát, kezét megdörzsölni, betakarni, és jó éjszakát mondani, mint otthon Ungban, az ősi házban, ahol a föld, a mező és a patak az első foglalásból maradott meg az Alvincziak kezén. (Krúdy Gyula, *A vörös postakocsi*).

Ahogyan az idézetben is látható, a pontos idő- és térbeli kidolgozottságot erősíti a hasonlatok gyakori alkalmazása is, amely egyfelől a különböző síkok összehasonlítását teszi lehetővé (a példában az Ungon töltött gyerekkort és az Arany Sas fogadóban telt egyetemista éveket), másfelől kevésbé az időpontok és helyszínek, hanem inkább a mentális tartományok között teremt kapcsolatot. A kötőszavak nélküli, felsorolás jellegű mellérendeléseken túl ezek a kapcsolattípusok tekinthetők gyakorinak Krúdy korpuszba felvett két regényében, amelyek sajátosságát így a különböző (valós, vagy mentális) területek és idősíkok részletes leírásában és a közöttük való mozgásban ragadhatjuk meg.

Eredmények 3.

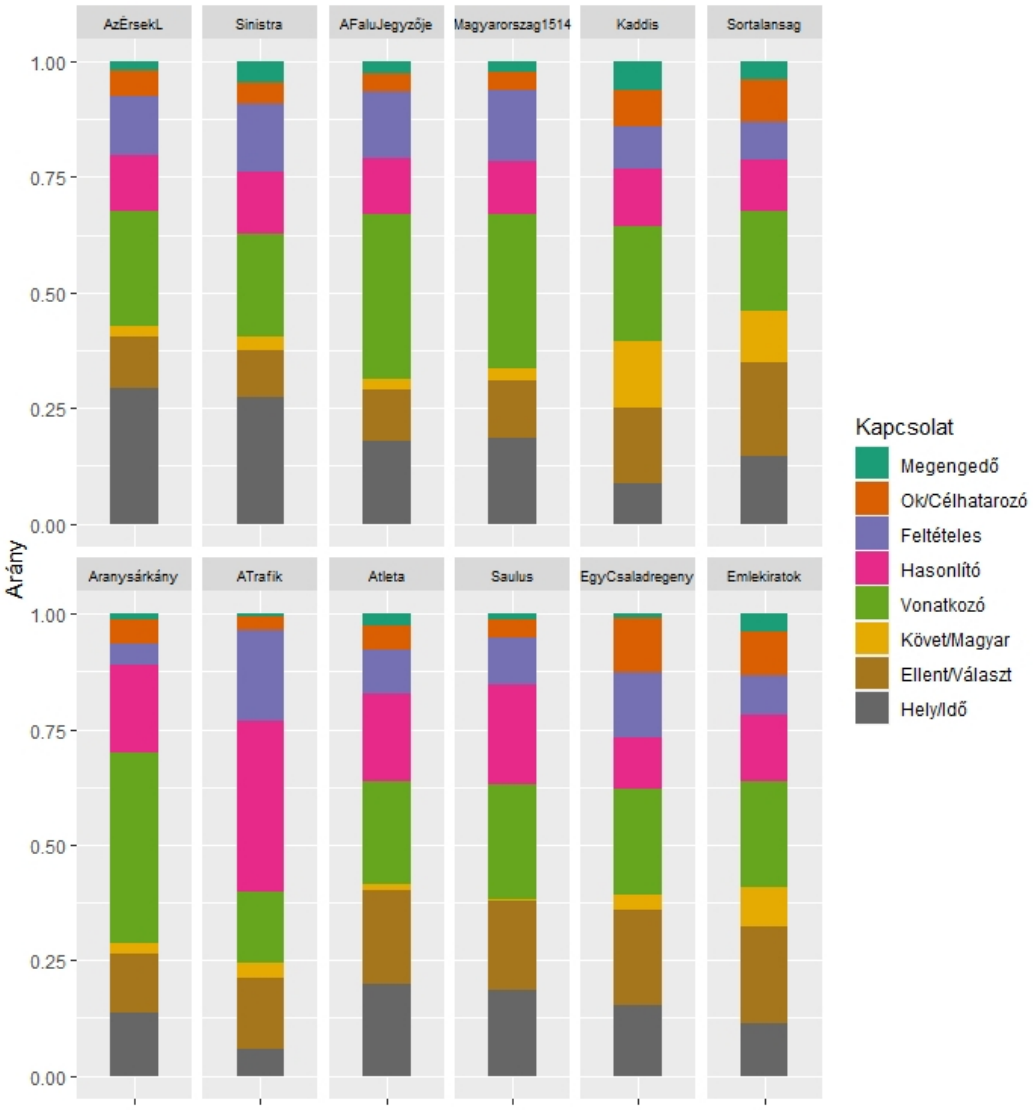
A harmadik lehetőség az eredmények ábrázolásában, amikor a regények belső szerveződésére összpontosítunk. Ebben az esetben nem a relatív gyakoriságokat, hanem csupán azt számítjuk ki, hogy összesen hány kidolgozott kapcsolat található egy szövegben, és ennek hány százalékát teszik ki az egyes típusok. Mindez természetesen csak a vizsgált típusokra vonatkozik és azoknak sem mindegyikére: azon túl, hogy az ábrák nem mutatják a kötőszóval nem kidolgozott, valamint a *hogy* kötőszós mellékmondatokat, nem látható a kapcsolatos mellérendelések aránya sem, mivel ezek bírnak a legkisebb magyarázó erővel egy szöveg stílusáról, illetve ezek esetében nagyon gyakori, hogy a kapcsolatos viszony nincs külön kötőszóval jelölve a szövegben, így a mért eredmények is csak megközelítőleg érvényesek. Mindezek miatt tehát nem mondható el, hogy a diagramok a regények egészének belső szerkezetét tükröznék – csupán a tagmondatkapcsolatok egy részének eloszlását ábrázolják. Továbbá a könnyebb átláthatóság kedvéért a *4.1.* és *4.2. ábrán* a választó és ellentétes mellérendeléseket, a magyarázó és következtető mellérendeléseket, valamint az idő-

 DOI: 10.31400/dh-hun.2021.4.2415

és helyhatározói alárendeléseket egy összevont típusba rendeztem, mivel az ezek által kidolgozott grammatikai jelentések viszonylag közel esnek egymáshoz. Könnyen létrehozható ugyanakkor egy részletesebb ábra is, ám csak az átláthatóság rovására – illetve ábrázolható az is, hogy csak az összevont kategóriákba tartozó típusok (pl. hely- és időhatározói alárendelések) hogyan aránylanak egymáshoz.

A relatív gyakoriságot mutató ábrák mellett több szempont miatt is érdemes ezeket a grafikonokat is szemügyre venni. Egyrészt ezáltal láthatóvá válik az eddig külön kezelt típusok gyakoriságának aránya a szövegeken belül. Mivel az egyes típusok (pl. a megengedő, vagy a magyarázó és következtető kapcsolatok) szinte mindig kevesebb számban fordulnak elő, mint mások (pl. vonatkozói alárendelések), az ábrák segítségével megválaszolható kérdés nem feltétlenül az, hogy melyik típus a leggyakoribb egy szövegben, hanem hogy milyen mértékűek a különbségek a típusok között. Másrészt így azoknak a regényeknek a sajátosságai is jobban kirajzolódhatnak, amelyek több korábbi kategóriában is kiugróan magas értéket mutattak (pl. Kertész Imre, *Kaddis*), vagy amelyek semelyik kategóriában nem mutattak kimagasló értéket, de belső szerveződésüket tekintve érdekes jellegzetességek figyelhetők meg (pl. Mány Iván, *A trafik*). Harmadrészt e diagramok segítségével a szövegek másképpen hasonlíthatók össze, mint eddig: mivel az ábrák egyszerre mutatják az összes kapcsolattípust, így mindegyiket figyelembe tudjuk venni az összehasonlítás során; a típusok arányai közti különbség pedig szinte azonnal érzékelhető.

A 12 regény belső szerveződését bemutató *4.1. ábrá*n például Mándy Iván *A trafik* című regényének jellegzetessége szembetűnő, hiszen esetében a hasonlatok egyértelmű dominanciája figyelhető meg. Azaz a korábbi (a relatív gyakoriságokat kiszámító) mérések alapján inkább az mondható el, hogy a regényre a tagmondatkapcsolatok hiánya és az egymástól elkülönült rövid benyomásszerű mondatok a jellemzők – a *4.1. ábra* mindezt azzal egészíti ki, hogy ha a szöveg mégis viszonyt dolgoz ki két tagmondat között, akkor azt leginkább hasonlítás útján teszi. Érdemes még megfigyelni, hogy a szövegben milyen kis számban fordulnak elő vonatkozói névmással bevezetett mellékmondatok a többi regényhez képest: gyakoribbnak tekinthetők ezeknél a szerkezeteknél a feltételes kapcsolatok, valamint a választó és ellentétes típusok összevont kategóriája is. Ezzel szemben az *Emlékiratok könyvé*ben vagy a *Kaddis*ban inkább a kapcsolatok arányos eloszlása figyelhető meg: más szövegekhez képes ezekben a regényekben például a megengedő és következtető/magyarázó kapcsolatok száma is jelentősnek tekinthető. Ez elsősorban az idő- és helybeli viszonyokat kidolgozó kapcsolatok rovására történik, legalábbis más regényekhez képest kisebb súllyal szerepelnek ezek a kapcsolattípusok az említett szövegekben. Az *Emlékiratok könyvé*ben például, habár nagyon fontosak az események idő és térbeli koordinátái és a különböző síkok viszonya, ez a viszony nem mindig kidolgozott nyelvileg, és inkább a történetszálak puszta egymás mellé rendeléséből jön létre. Gyakran csak több mondat után lehet megkülönböztetni például, hogy melyik fejezet hol és mikor játszódik, és csak egy-egy szereplő vagy történés kapcsán következtethetünk az elmesélt történet idejére/helyére – hiszen, amint az eddigiek során is láttuk, a kapcsolatok inkább egy konkrét jelenet összefüggéseit és nem a tágabb idő- és helybeli kontextust dolgozzák ki. Bodor Ádám regényeiben *(Sinistra körzet, Az érsek látogatása)* éppen ellenkezőleg, az ábrázolt kapcsolatoknak több mint a negyede dolgoz ki hely- és időbeli viszonyokat.



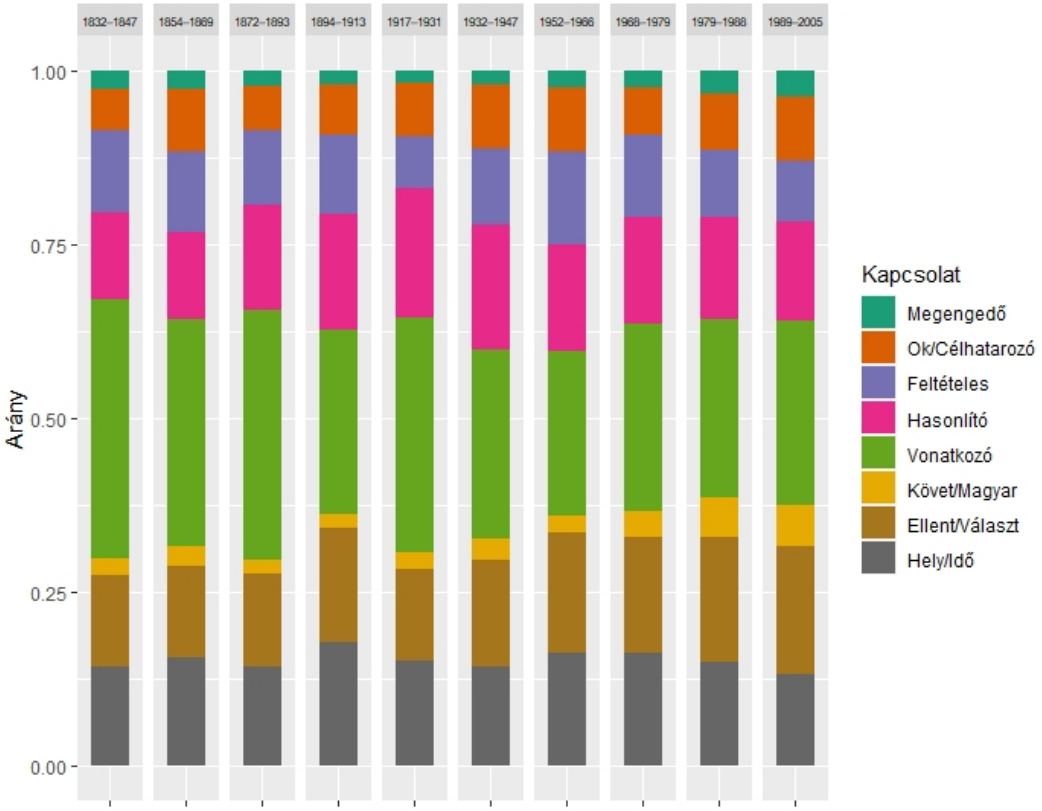
4.1. ábra. A kapcsolattípusok egymáshoz viszonyított aránya – 12 regény.

Szembetűnő az ábrán továbbá Mészöly Miklós *Saulus*ában és *Az atléta halálá*ban a magyarázó és következtető kapcsolatok elenyésző és az ok-/célhatározói kapcsolatok kis szerepe, ami a szövegek olyan szerveződésére hívja fel a figyelmet, amely a logikai kapcsolatok korlátozásán alapul.

Mindezek mellett lehetőség nyílik a 10 regényenkénti évcsoport belső szerveződését is hasonló módon vizualizálni. A *4.2. ábrá*n a korábban ismertetett történeti folyamatokat figyelhetjük meg az egyes oszlopok közti különbségeket vizsgálva (lásd *Eredmények 1.*). Egyedül a vonatkozói alárendelések arányának változásában nem rajzolódik ki egyértelműen a relatív gyakoriság alapján felfedezett tendencia, ami arra vezethető vissza, hogy míg a 19. század elejének hosszúmondatos regényei rendszerint másfajta kapcsolatokat is létesítenek a tagmondatok között, addig a 20. század első

 DOI: 10.31400/dh-hun.2021.4.2415

felében azok a regények, amelyekben gyakoriak a vonatkozói alárendelések, kisebb számban dolgoznak ki más típusokat (elsősorban Kosztolányi Dezső művei).



4.2. ábra. A kapcsolattípusok egymáshoz viszonyított aránya, évcsoportok.

Eredmények 4.

A regények belső szerveződését mutató diagram *(4.1. ábra)* segítségével könnyen felismerhetők az azonos mondatszerkezeteket előnyben részesítő művek, és szembetűnő az egy alkotóhoz (pl. Eötvös József) tartozó regények hasonlósága is. Ennek kapcsán merülhet fel a kérdés, hogy vajon a stílustörténeti tendenciák felrajzolásán és az egyes szövegek értelmezésén túl segíthet-e a mondatkapcsolatok kvantitatív vizsgálata az ismeretlen szerzőséget azonosító kutatások során is. Azaz milyen hatásfokon lehet elkülöníteni az egy szerzőhöz tartozó szövegeket a korpusz többi szerzőjétől a meglévő adatokra hagyatkozva. A szerzőségre irányuló kutatások több ízben bizonyították már, hogy a szövegek tematikus szintje alatt létezik egy úgynevezett „szerzői ujjlenyomat", amely a leggyakrabban – és ezért nem tudatosan – használt szavak, főként konkrét jelentés nélküli funkciószavak eloszlására vonatkozik, és amely eloszlás a szerzők különböző időszakban írt, különböző műfajú szövegeiben is hozzávetőlegesen állandó.[61] A leggyakrabban használt szavak relatív gyakorisága alapján tehát lehetséges

[61] Lásd például Harald Baayen, *Word Frequency Distributions* (Dordrecht: Kluwer, 2001), `https://doi.org/10.1007/978-94-010-0844-0`; John Burrows, „'Delta': A Measure of Stylistic

egy szerzőt (társadalmi-esztétikai kondícióitól függetlenül) más alkotóktól megkülönböztetni; kérdés, hogy ez igaz-e kizárólag a kötőszavakat és a tagmondatkapcsolatot létesítő vonatkozói névmásokat figyelembe véve is.

Az alábbi ábrák létrehozásakor tehát a kutatás során kiszámított értékeket vettem alapul. Ezek az értékek az egyes kapcsolattípusok relatív gyakoriságát vagy az egymáshoz viszonyított arányát mutatják: egy regényhez ezáltal több (12, illetve 8) érték is kapcsolódik, amelyek mentén egy többdimenziós térben/koordinátarendszerben lesznek a szövegek elhelyezhetők. Ez a pozíció a többdimenziós térben ugyanúgy határozható meg, mint ahogyan egy kétdimenziós koordinátarendszerben is jelölhető bármely pont, amelyhez megadjuk az x és az y tengelyre vonatkozó értékeket. Egy ilyen többdimenziós teret ugyanakkor sem ábrázolni, sem elképzelni nem tudunk – azonban létezik két lehetőség, hogy ebben a térben elfoglalt helyük alapján megadjuk a szövegek távolságát és közelségét, azaz (az értékek által leírt jellemzőkre vonatkozó) hasonlóságukat és különbözőségüket. Az egyik lehetőség, ha nem ábrázoljuk a pontokat ebben a térben, csupán a pozíciók közti távolságot számítjuk ki a két dimenzióban is érvényes metrikák szerint, majd a szövegeket az egymástól való közelségük alapján csoportosítjuk, végül ezeket a csoportokat egy dendrogram formájában jelenítjük meg. A kutatás során több távolságmetrika teljesítményét is összevetettem a regények ilyen csoportosításakor, amelyek közül a kétféle adatsor (a kapcsolattípusok relatív gyakorisága és egymáshoz viszonyított aránya – *5.1* és *5.2 ábra*) tekintetében a koszinusztávolság bizonyult a leghatékonyabbnak, míg az összevont adatsor alapján a Manhattan-távolság teljesített a legjobban (lásd a 65. és a 66. lábjegyzetet).[62] Az alábbiakban minden esetben a leghatékonyabb csoportosítást létrehozó módszereken alapuló eredményeket ismertetem, ami természetesen az eredmények kimazsolázását *(cherry-picking)* jelenti: ám ebben az esetben csupán a lehető legjobb csoportosítási eredményeket és azokon keresztül a kalasszifikáció korlátait kívánom bemutatni. A másik lehetőség a többdimenziós tér dimenzióinak a csökkentésére irányul, amelynek segítségével kétdimenziós felületen tudjuk ábrázolni az adatpontok közötti összefüggéseket. Jelen

Difference and a Guide to Likely Authorship," *Literary and Linguistic Computing* 17 (2002): 267–287, `https://doi.org/10.1093/llc/17.3.267`; Patrick Juola, „Authorship Attribution," *Foundations and Trends in Information Retrieval* 1, 3. sz. (2006): 233–334, `http://doi.org/10.1561/1500000005`. Stilometriai kutatásokat összefoglaló tanulmány: Maciej Eder, „Style-Markers in Authorship Attribution: A Cross-Language Study of the Authorial Fingerprint," *Studies in Polish Linguistic* 7, 1. sz. (2011), hozzáférés: 2021.04.28, `http://www.ejournals.eu/sj/index.php/SiPL/article/view/2261/0`.

[62] Mindazonáltal a szakirodalomban elfogadottnak tűnik az a megállapítás, miszerint a szerzősazonosítás esetében a koszinusztávolság a legmegbízhatóbb metrika, vö. Evert Stefan et. al., „Understanding and Explaining Delta Measures for Authorship Attribution," *Digital Scholarship in the Humanities* 32, 2. sz. (2017): 4–16, `https://doi.org/10.1093/llc/fqx023`. A Manhattan-távolság a kétdimenziós térben két ponthoz egy harmadikat jelöl, amellyel együtt egy olyan derékszögű háromszöget alkotnak, amelynek átfogója a két pont közötti egyenes; a távolság mértékét pedig a befogók hosszának összege határozza meg. A metrika elnevezését New York jellegzetes építészeti elrendezéséről kapta, hiszen az egymásra merőleges utcákon általában derékszöget bezáró útvonalon lehet eljutni egyik pontból a másikba. A koszinusztávolság ezzel szemben azt számítja ki, hogy ha két pontot egy koordinátarendszerben ábrázolunk, akkor mekkora a pontokat és az origót összekötő egyenesek által bezárt szög koszinusza.

 DOI: 10.31400/dh-hun.2021.4.2415

esetben ehhez az úgynevezett főkomponens-analízis eljárását alkalmaztam, amely az adatpontok varianciáját megtartva redukálja a dimenziók számát.[63]

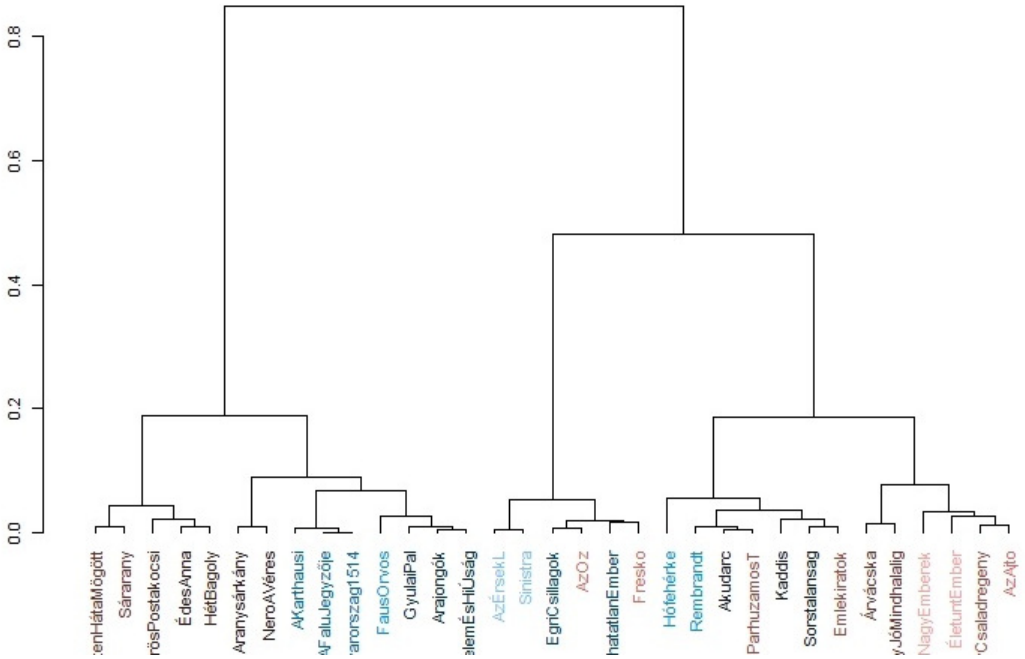
Az *5.1.* és *5.2. ábrán* a koszinusztávolság szerint és a Ward-módszer[64] alapján csoportosított 33 regény látható a korpuszból, amelyek közül legalább kettőnek azonos szerzője van (az azonos szerzőséget színkód jelöli), hogy ezáltal elemezhető legyen, hogy milyen hatékonysággal kerülnek ugyanahhoz az alkotóhoz tartozó szövegek egymás mellé a dendrogram ágrajzán. Az *5.1. ábra* a kapcsolatok belső arányát, a *5.2. ábra* a kapcsolatok relatív gyakoriságát vette figyelembe (a kapcsolatos mellérendelést kivéve) mint a szövegek helyzetét meghatározó értékeket. Mindkettőn helyesen került egymás mellé Bodor Ádám és Gárdonyi Géza két-két, valamint Eötvös József és Kemény Zsigmond három-három regénye, ahogyan Krúdy Gyula két és Kosztolányi Dezső három műve is közel, egy fürtön helyezkedik el az ábrákon. Ugyanakkor mindkét csoportosításban kisebb-nagyobb hibák is mutatkoznak: Szabó Magda és Bródy Sándor egyes regényei meglehetősen távolra kerültek egymástól, Móricz Zsigmond négy regénye pedig két csoportra oszlik *(Az Isten háta mögött* és *Sárarany,* illetve *Légy jó mindhalálig* és *Árvácska).* A regények szerzőség alapján történő elkülönítése így csupán közepesen sikeresnek mondható a kapcsolattípusok gyakoriságát vizsgálva.[65] Az eredmények kis mértékben javíthatók,[66] ha mind a két szempontot (relatív gyakoriság és belső arány) egyszerre érvényesítjük, ám az azonos szerzőségű művek elkülönítése így sem tökéletes *(5.3. ábra).*

[63] A főkomponens-analízis eljárása az alábbi módon képzelhető el: a módszer első lépésben az értékek átlaga által kijelölt pontot teszi meg a koordinátarendszer középpontjául, majd egy olyan egyenest húz ezen az origón keresztül, amelyhez a lehető legközelebb esnek az egyes adatpontok. Ezt követően erre az egyenesre további merőlegeseket állít a változók számának megfelelően. A végső lépésben ezek közül a merőlegesek közül azt a kettőt teszi meg az eredményül megadott kétdimenziós ábra két egyenesének (PC1 és PC2), amelyekre vetítve az adatpontok origótól való távolságának négyzetösszege a legnagyobb. A létrehozott diagram mutatja, hogy PC1 és PC2 hány százalékban játszik szerepet az adatpontok közötti különbség meghatározásában: minél nagyobb ez a szám, annál jelentősebb az egységnyi távolság e tengely mentén. Vö. Zakaria Jaadi, *A Step-by-Step Explanation of Principal Component Analysis (PCA),* hozzáférés: 2021.04.28, `https://builtin.com/data-science/step-step-explanation-principal-component-analysis`.

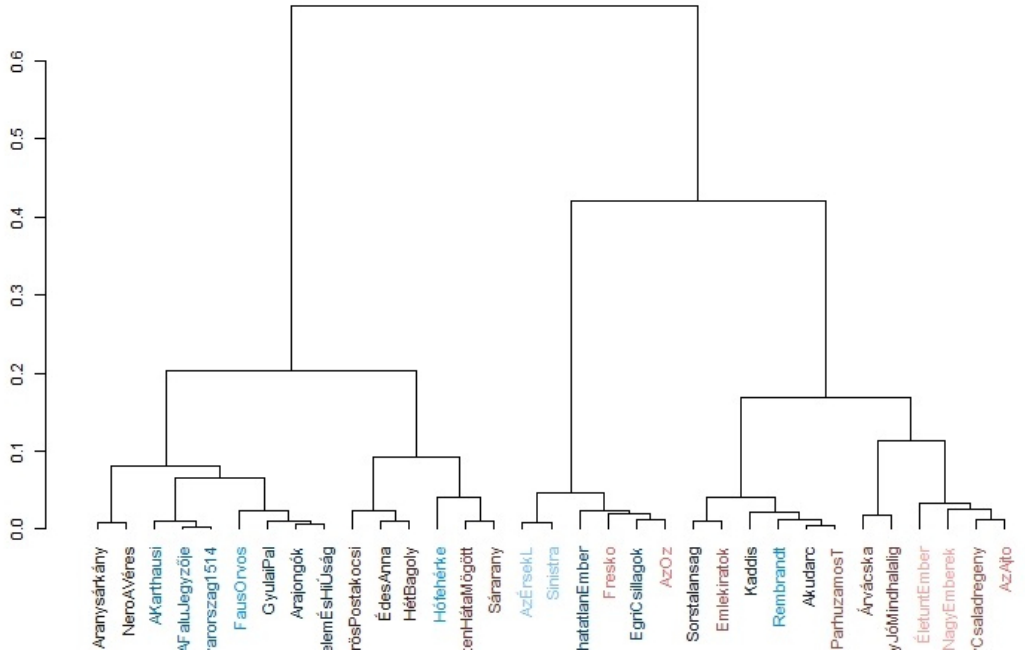
[64] A Ward-módszer arra törekszik, hogy az egy csoportba rendezett elemek varianciája, azaz a csoportátlaghoz képesti eltérések négyzetösszege a lehető legkevesebb legyen.

[65] A kapcsolattípusok egymáshoz viszonyított aránya alapján létrehozott csoportok a 12 szerző helyes elrendezéséhez képest 0,43-as értéket mutat a két csoportosítást összehasonlító korrigált Rand index (ARI) szerint (Manhattan-távolsággal: ARI = 0,43); a relatív gyakoriságok alapján létrehozott csoportosítás esetében ARI = 0,42 (Manhattan-távolsággal: ARI = 0,37). Ez a mérőszám minél inkább 0-hoz közelít, annál inkább egyezik meg az eredmény egy pusztán véletlenszerű csoportosítás eredményével; míg minél inkább 1-hez közelít, annál inkább egyezik meg a szerzőség szerinti helyes csoportosítással. (Ugyanakkor a statisztikailag kevésbé megbízható, mert a véletlenszerű eloszlás hatását nem, csupán az egy fürtre kerülő elempárokat vizsgáló nem-korrigált Rand index alapján mindkét csoportosítás magas, 0,9 feletti értéket mutat a szerzők mentén elrendezett csoportokhoz képest.)

[66] Manhattan-távolsággal: ARI = 0,54. Koszinusztávolsággal: ARI = 0,45.

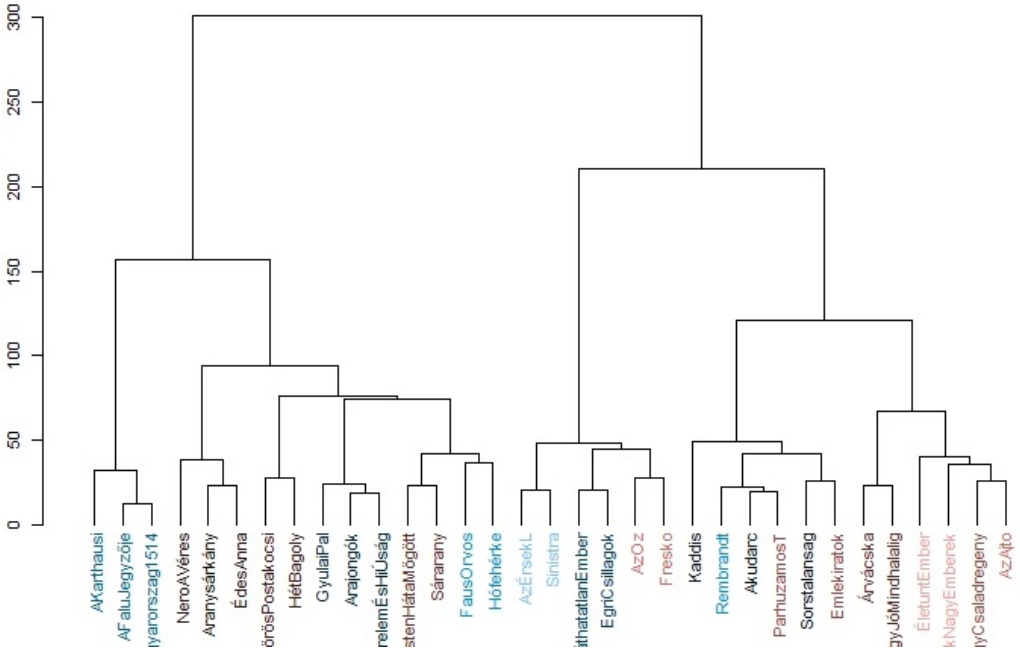


5.1. ábra. Koszinusztávolság. 33 regény csoportosítása a kapcsolattípusok egymáshoz viszonyított aránya alapján.



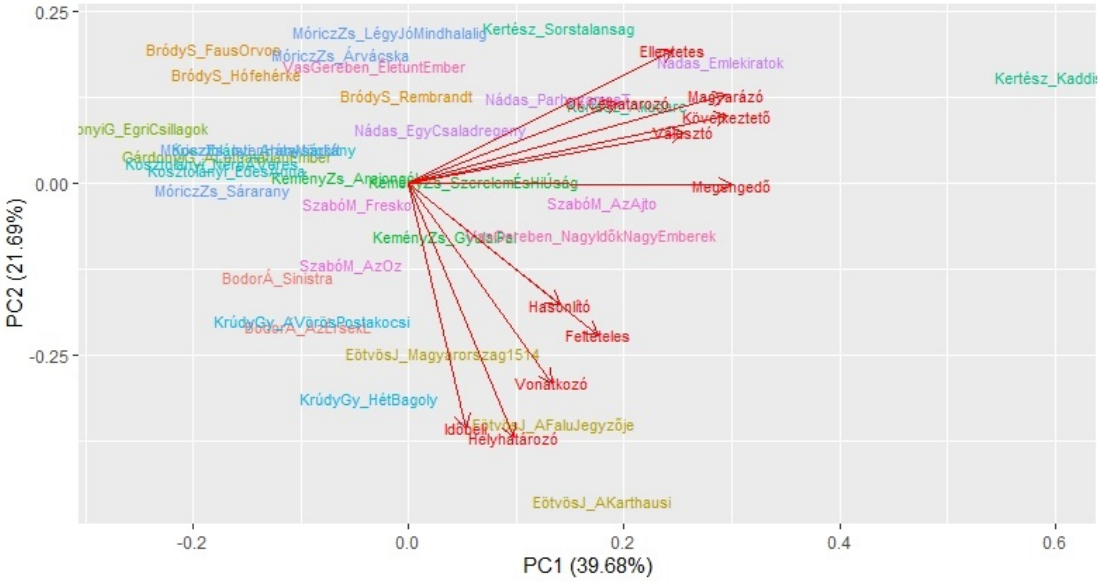
5.2. ábra. Koszinusztávolság. 33 regény csoportosítása a kapcsolattípusok relatív gyakorisága alapján.

 DOI: 10.31400/dh-hun.2021.4.2415



5.3. ábra. Manhattan-távolság. 33 regény csoportosítása a kapcsolattípusok egymáshoz viszonyított aránya és relatív gyakorisága alapján.

A relatív gyakoriságokra vonatkozó főkomponens-analízis eredménye látható az *5.4. ábrán*. Itt a szövegek közti különbség és hasonlóság a térben elfoglalt pozíció, valamint a tengelyekhez (PC1 és PC2) rendelt százalék szerint alakul – ez a szám mutatja, hogy milyen mértékben játszik szerepet a tengelyeken megfigyelhető távolság az adatpontok elkülönítésében. Tehát a kiválasztott szempont szerint hasonló szövegek egymáshoz közel helyezkednek el; míg az azonos szerzőségű művek azonos színnel szerepelnek az ábrán. Ezenkívül mint vektorok jelöltek az egyes kapcsolattípusok is az ábrán, amelyeknek iránya mutatja, hogy ezek a típusok hogyan befolyásolják a szövegek elhelyezkedését.



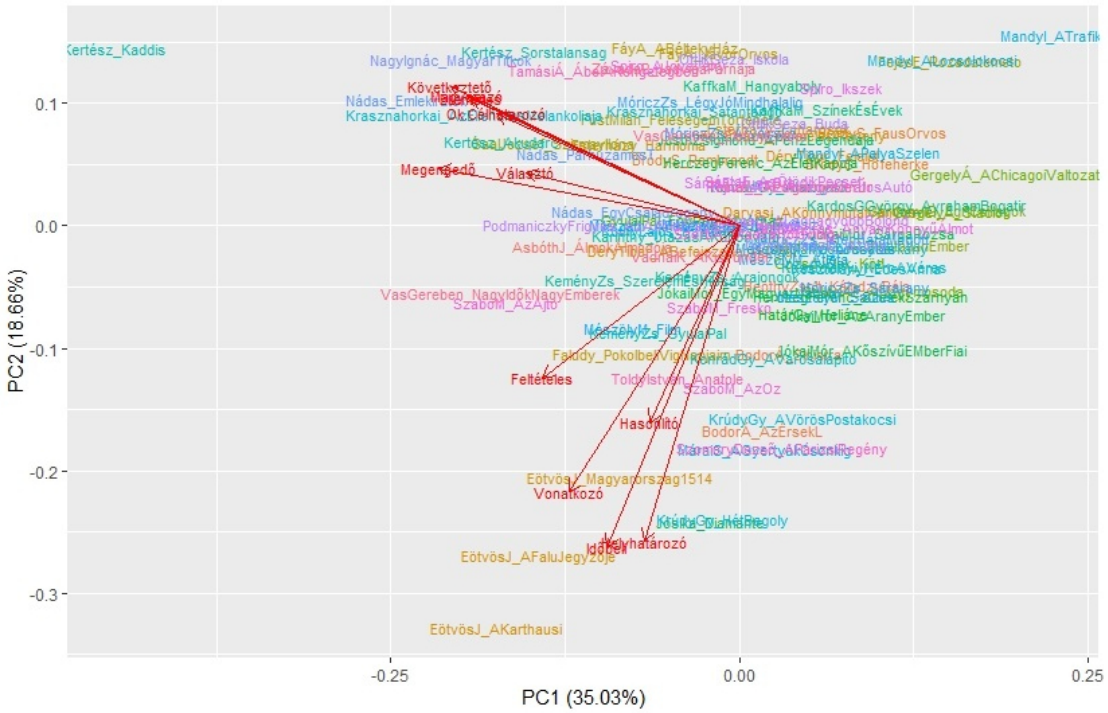
5.4. ábra. Főkomponens-analízis a kapcsolattípusok relatív gyakorisága szerint, 33 regény.

Az azonos szerzőségű regények elkülönítése ezúttal is egyszerre nevezhető sikeresnek és sikertelennek: a legtöbb esetben hasonló helyzetűek, de nem különíthetők el egyértelműen más szövegektől vagy szövegcsoportoktól. Mindez magukra a vizsgált jellemzőkre vezethető vissza. A kötőszavak és a tagmondatkapcsolatot létesítő vonatkozó névmások használata ugyanis a szöveg *félig tudatos* szintjére vonatkozik; hiszen míg más funkciószavak (pl. névelők) eloszlása teljesen független az alkotói szándéktól – gyakoriságuk ezért is képes a szerző stílusának „ujjlenyomatát" hordozni –, addig a tagmondatkapcsolatok kidolgozása nem teljesen független az alkotó esztétikai és tartalmi megfontolásaitól. A tudatos és a nem tudatos szintek között fejtik ki a hatásukat, amennyiben használatuk (a tartalmas jelentéssel bíró szavakhoz képest) nem kontrollált az alkotói folyamat során, ugyanakkor a tudatos írói döntések befolyásolhatják gyakoriságukat, illetve az általuk kidolgozott kapcsolatok gyakoriságát. Ezért tekinthető a kísérlet közepesen sikeresnek, azaz emiatt lehet több azonos szerzőhöz tartozó szöveget egy csoportba rendezni a vázolt módszerekkel, és ugyanígy emiatt nem felismerhető a más mondatszerkezeteket előnyben részesítő regények közös szerzősége. A kutatás eredményei azt mutatják, hogy nem is a szerzőazonosítás kérdésének megválaszolására alkalmas a tagmondatkapcsolatok gyakoriságának elemzése, hanem a különböző hagyományokhoz tartozó prózapoétikák különíthetők el általa.

Az *5.4. ábrán* világosan látható, hogy a relatív gyakoriság szerint a kapcsolattípusok két csoportja egymástól élesen elválik, azaz két irány különíthető el a piros vektorokat figyelembe véve: egyfelől a mellérendelő (választó, ellentétes, következtető, magyarázó és megengedő) kapcsolatok és az ok-/célhatározói alárendelések, másfelől a hasonlító, feltételes, vonatkozói, valamint hely- és időhatározói alárendelések fejtik ki hatásukat megegyező irányban. Mind a 100 regényt vizsgálva hasonló eredményre juthatunk *(5.5. ábra)*. Ami alapján arra a következtethetünk, hogy a mondatszerkezetek kvantitatív vizsgálatán keresztül a kötőszó- és vonatkozónévmás-használat szem-

 DOI: 10.31400/dh-hun.2021.4.2415

pontjából három prózastílus határozható meg a magyar regényirodalom kanonikus szövegeit tekintve. Egyrészt elkülönül egy, a tagmondatok közötti, elsősorban logikai viszonyt kidolgozó stílus – abban az értelemben, ahogyan a formális logika is főként a mellérendelő kapcsolatokat ruházza fel logikai értékkel; illetve az ok-/célhatározói kapcsolatok is ilyen összefüggések kidolgozásában játszanak szerepet, nagyon hasonló módon a következtető és magyarázó mellérendelésekhez, amelyektől való elhatárolása esetenként mesterségesnek is tűnhet. Másrészt egy csoportba rendeződnek az inkább a mondatok szereplőit vagy az ábrázolt jelenetek körülményeit részletesebben kidolgozó stílushoz tartozó regények; ezekben a vonatkozói, feltételes és hasonlító mellékmondatok száma tekinthető gyakorinak. A harmadik stílust képviselik azok a szövegek, amelyekre egyik típus sem jellemző, hanem inkább rövidmondatos,[67] tagmondatkapcsolatokat ritkábban kidolgozó prózanyelvet érvényesítenek. Természetesen ezek a stílusok inkább mint ideáltípusok írhatók le, és ritkán valósulnak meg tiszta formában. Mindazonáltal – ahogyan az *5.5. ábra* is mutatja – az első hagyományra Kertész Imre *Kaddis a meg nem született gyermekért,* a második hagyományra Eötvös József *A karthauzi,* míg a harmadikra Mándy Iván *A trafik* című regénye tekinthető a korpuszból a legjellemzőbbnek: a többi szöveg az ezek által határolt térben oszlik el.



5.5. ábra. Főkomponens-analízis a kapcsolattípusok relatív gyakorisága szerint, 100 regény

[67] A mondathosszúságokhoz lásd Szemes, *Mondathosszúság és irodalomtörténet* idézett tanulmányát és annak függelékét.

Összegzés

A tagmondatkapcsolatok eloszlásának vizsgálata nemcsak a szövegek nyelvi szerveződésének jobb megismeréséhez vezethet, hanem diagrammatikus, topológiai-logikai karakterüket is képes megvilágítani. Automatikus azonosításuk és gyakoriságuk mérése így a stilisztikai leíráson túl egyaránt lehetőséget teremt irodalomtörténeti tendenciák felvázolására, az egyes regények értelmezésére és különböző prózapoétikai hagyományok elkülönítésére. A dolgozat három hagyományt különböztetett meg egymástól: egy inkább alárendeléseket alkalmazó, a mondatok egy-egy szereplőjét részletesebben kidolgozó, ezáltal a jelentést jobban stabilizáló stílust; egy, a tagmondatok közötti logikai kapcsolatot létesítő stílust, amely állandóan új értelmezések felé nyitja meg a mondatot; valamint egy rövidmondatos stílust, amely inkább egyszerű mondatokból, valamint nem kidolgozott tagmondatkapcsolatokból (illetve a kidolgozott kapcsolatokat tekintve elsősorban hasonlatokból) épül fel. Az ilyen következtetésekhez járulnak hozzá az eredmények különböző vizualizációi, amelyek különböző összefüggéseket mutató elrendezésekben képesek ábrázolni a mérési eredményeket. A továbbiakban érdemesnek tűnik az egyes kapcsolatok részletesebb, alkategóriákra bontott elemzését is elvégezni, valamint a mondatok között kidolgozott kapcsolatok vizsgálatával kiegészíteni az eddigi kutatást.

Elaborated Relations. On the usability of automatic identification of clause relationships in stylistics and literary history

The paper aims to develop a method that allows an automatic identification of the different types of complex and compound sentences in the Hungarian written language through the analysis of conjunction words, relative pronouns and their positions in the sentence. This method opens up new perspectives in the stylometric research: on the one hand, conjunctions and pronouns as function words provide a large amount of data for statistical analyses, and on the other hand they also carry meaning – about the relationships between the clauses (eg. opposition, conditionality). By examining the relative frequency of each type, it becomes possible to identify the most characteristic relationship of the clauses in a given text or corpora. Thus, the style and poetics of the novels can be grasped at the scale of the sentence, while also revealing the topological-logical structure of the texts, which is not usually reflected during the reading process. The paper analyzes the frequency of each type of relation in a corpus of 100 Hungarian novels. This provides an opportunity to (1) register stylistic tendencies of the Hungarian novel in the period between 1832–2005); (2) examine which texts are the most typical of each type compared to other texts; (3) to visualize the 'internal structure' of a text, based on the relative proportions of types to each other; and (4) to distinguish different stylistic traditions.

Keywords:
clause relation, conjunctions, stylometry, literary history, Hungarian novel

 DOI: 10.31400/dh-hun.2021.4.2415

Függelék

A kutatás során használt kódok, valamint az egyes regényekre vonatkozó eredmények elérhetők:

`https://github.com/SzemesBotond/tagmondatkapcsolat`

A tanulmány függelékei és ábrái a cikk mellékleteként is elérhetők annak adatlapjáról:

`https://doi.org/10.31400/dh-hun.2021.4.2415`