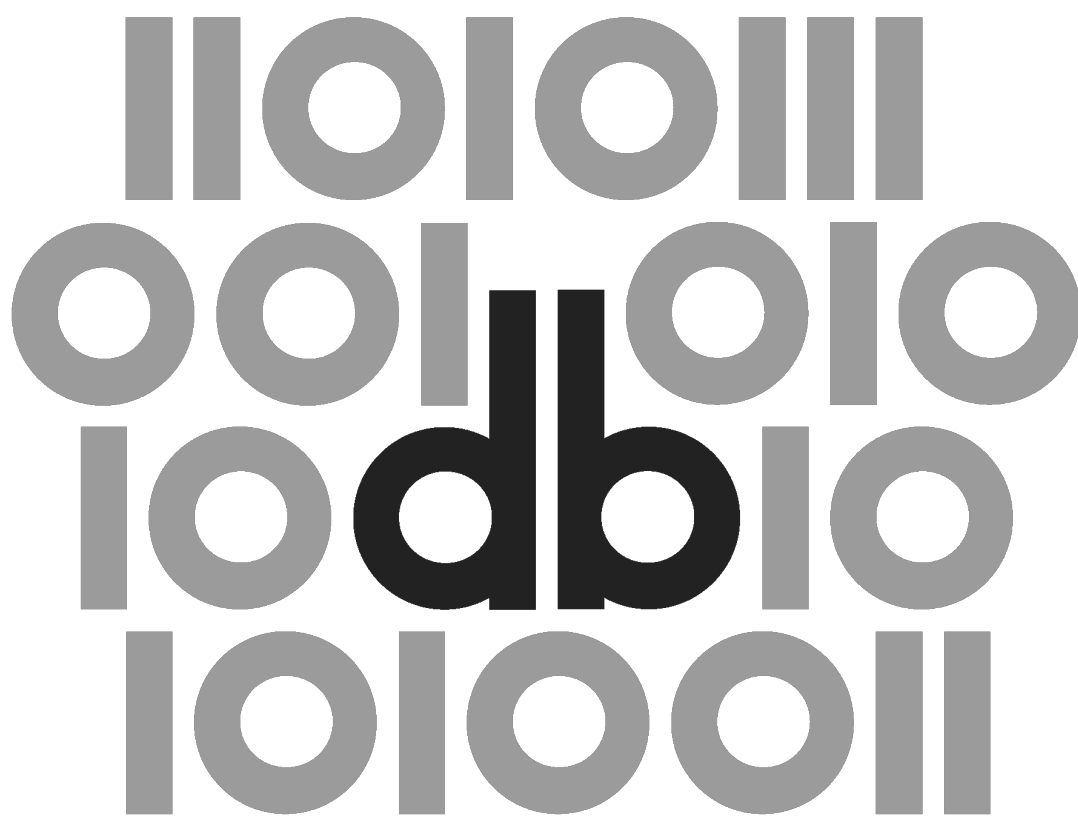


8 (2024)

<DIGITÁLIS BÖLCSÉSZET>



8 (2024)

</DIGITÁLIS BÖLCSÉSZET>

Digitális Bölcsészet
2024., nyolcadik szám

<DIGITÁLIS BÖLCSÉSZET>



8 (2024)

Felelős szerkesztő:

Maróthy Szilvia

Szerkesztőség:

Kokas Károly, Parádi Andrea

Rovatvezetők:

Tanulmányok: Kiss Margit

Műhely: Péter Róbert

Kritika: Almási Zsolt

Labor: Mártonfi Attila

Tanácsadó testület:

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

Korábbi munkatársaink:

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

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

ISSN 2630-9696

DOI 10.31400/dh-hun.2024.8

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

Felelős kiadó az ELTE BTK Régi Magyar Irodalom Tanszék vezetője.

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

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

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

Email cím: dbfolyoirat@gmail.com

Olvasószerkesztő: Bucsics Katalin

Tördelés: Mészáros Marcell

Grafika: Hegyi Gábor

<TANULMÁNYOK>

Szemes Botond  0000-0002-0637-6776

BTK Irodalomtudományi Intézet

szemes.botond@abtk.hu

Nagy Mihály  0000-0002-0577-8706

ELTE BTK Történettudományi Doktori Iskola

mihaly.nagy@btk.elte.hu

Ismétlés és innováció a drámai művekben

Kísérlet a szereplői megszólalások újszerűségének kiszámítására*

Tanulmányunkban egy olyan módszert ismertetünk, amellyel drámai szereplők megszólalásainak hírértéke mérhető össze kvantitatív módon. Innovatívnak tekinthető minden olyan szereplő, aki gyakran dolgoz ki a korábbiaktól eltérő jelentést; fenntartó vagy ismétlő funkciót lát el az a karakter, aki inkább a korábban kialakult diskurzus megismétlését (és ezáltal megerősítését) végzi el. Az összehasonlítás alapja a szereplői mondatokhoz az SBERT algoritmus segítségével hozzárendelt beágyazási érték. Ugyanakkor egy karakteren belül is megfigyelhető innováció és ismétlés kettőssége: a módszer ezt is vizsgálhatóvá teszi a mondat-szintű elemzés során. A szereplők páros összehasonlítása lehetővé teszi továbbá, hogy a közöttük lévő viszonyokat irányított gráfként ábrázoljuk, amivel a színművek hálózatos megjelenítésének kérdéséhez is hozzá kívánunk járulni. Az eljárást alkalmazva általános műfaji különbségeket azonosítunk Shakespeare tragédiái és komédiái között a bennük érvényesülő nemi szerepek alapján. Az általános mintázatok mellett a módszer az egyes színművek és szereplők újfajta értelmezése számára is alapot teremt, ami a szoros szövegolvasás és a statisztikai-nyelvtechnológiai módszerek együttes alkalmazhatóságára is rávilágít.

Kulcsszavak:

innováció, ismétlés, kvantitatív drámaelemzés, Shakespeare, SBERT



* A Kulturális és Innovációs Minisztérium ÚNKP-23-4 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

1. Bevezetés

Egy drámai műben megkülönböztetett jelentősége van az információ áramlásának, amelyet a szereplői megszólalások tartanak fent. Egyrészt a belső kommunikáció szempontjából, amennyiben a szereplők közötti információátadás (vagy -visszatartás) mozdítja előre a cselekményt,¹ másrészt a külső kommunikáció szempontjából, amennyiben a nézők/olvasók is legtöbbször e dialógusokon keresztül férhetnek hozzá a dráma világához.² Ennek felelően az elmúlt években egyre népszerűbb, közös jelenetek alapján létrehozott karakterhálózatokat is előszeretettel értelmezik a *belső* információs rendszer felől a kutatók, bár többnyire csak implicit módon, mint például a közöttség centralitás esetében, amelyből kiindulva karakterek közvetítő, akár „konspiratív” szerepére következtetnek.³ Benjamin Krautter azonban óvatosságra int, amikor rámutat, hogy az információ terjedésének ilyen vizsgálatában inkább a szereplők közötti ismeretátadást reprezentáló tudáshálózatok segíthetnek, amelyek viszont jó eséllyel más elrendezést mutatnak, mint a közös jelenetekben alapuló hálózatok.⁴

Jelen tanulmány ezekkel a kísérletekkel szemben a *külső* kommunikációs rendszer, azaz a befogadói tapasztalat felől elemzi a szereplői megszólalások információtartalmát Shakespeare műveiben. Andresen és munkatársai szintén figyelembe vették ezt a szempontot kutatásaikban, azonban kisebb mértékben, illetve csak egy specifikus tudásfajta átadása kapcsán.⁵ Továbbá az „információtartalom” meghatározásakor, velük szemben, nem követjük szorosan Manfred Pfister elméletét, vagyis nemcsak azokat a megszólalásokat vizsgáljuk, amikor egy szereplő konkrét ismereteket közvetít a befogadó számára;⁶ hanem minden megszólalást aszerint, hogy mennyiben adnak hozzá új jelentésrétegeket az ábrázolt világhoz. Amikor a *Hamlet*-ben Claudius felveti Hamlet száműzetésének tervét, a megszólalás információértékét növeli, hogy a darabban először kerül említésre Anglia és Dániához fűződő viszonya; a dráma világának horizontja szélesebb lesz. Azonban már korábban is volt szó Dánia külpolitikai kapcsolatairól (Norvégia esetében), így kisebb az eltérés a korábbi diskurzustól.

¹ Lásd: Melanie Andresen et al., „Who Knows What in German Drama? A Composite Annotation Scheme for Knowledge Transfer. Annotation, Evaluation, and Analysis”, *Journal of Computational Literary Studies* 1 (2022), <https://doi.org/10.48694/jcls.107>; Melanie Andresen et al., „Knowledge Distribution in German Drama”, *Journal of Open Humanities Data* 10, 1. sz. (2024): 1–7, <https://doi.org/doi:10.5334/johd.167>.

² A külső és belső kommunikációs rendszer fogalmához lásd: Manfred Pfister, *The Theory and Analysis of Drama*, ford. John Halliday (Cambridge: Cambridge University Press, 1988).

³ Pl. Mark Algee-Hewitt, „Distributed Character: Quantitative Models of the English Stage, 1550–1900”, *New Literary History* 48, 4. sz. (2017): 751–782, <https://doi.org/10.1353/nlh.2017.0038>; Szemes Botond és Vida Bence, „Tragikus és komikus hálózatok. Drámai műfajok csoportosítása szerkezeti tulajdonságok alapján”, *Helikon* 68, 2. sz. (2022): 345–368.

⁴ Benjamin Krautter, „Kopräsenz-, Koreferenz- und Wissens-Netzwerke. Kantenkriterien in dramatischen Figurennetzwerken am Beispiel von Kleists Die Familie Schroffenstein (1803)”, *Journal of Literary Theory* 17, 2. sz. (2023): 261–289, <https://doi.org/10.1515/jlt-2023-2012>. Lásd továbbá: Andresen et al., „Who Knows What in German Drama?”.

⁵ Andresen et al., „Who Knows What in German Drama?”.

⁶ Pfister példája Prospero Arielhez intézett szavai A *viharban*, amelyek inkább a nézők számára hordoznak információt a cselekmény előzményéről, hiszen Ariel mindennek tudatában volt már korábban is. Pfister, *The Theory and Analysis of Drama*, 40–41.

Ugyanígy lehet információtartalma annak is, ha egy szereplő az addigiaktól eltérő, új regiszterben szólal meg, hiszen ez mutatja, hogy az ábrázolt világban egyáltalán lehetségesek ilyen megszólalásmódok, illetve ezek mint kontextusok befolyásolják a többi megszólalás értelmezhetőségét is. Gondoljunk például a *Hamlet* második jelenét indító királyi beszéd és az első jelenetben Horatio és társai között elhangzó mondatok különbségeire vagy ezekhez képest a sírásók dialógusára az ötödik felvonásból. A királyi propaganda, valamint a baráti vagy tréfás megjegyzések feszültségei hozzák létre azt a keretet, amelyben a tragédia megvalósul. Ugyanakkor a sírásók mondatai, amelyek Hamlet száműzetéséről szólnak, kevésbé újszerűek, hiszen már korábban elhangzott a drámában, hogy Angliába küldik Hamletet vélt bolondsága miatt. A különböző típusú eltéréseket együttesen *szemantikai különbség*nek nevezzük, amelyet kísérleteink szerint jól meg tud ragadni egy BERT alapú nyelvi modell. A kifejezés egyszerre utal arra, hogy a dialógusok tartalmának vizsgálatát helyezzük az előtérbe, ám figyelembe vesszük a stílus szemantikai összetevőit is (például egy erősen metaforikus megszólalás általában jobban elkülönül kevésbé metaforikus mondatoktól).

Mindennek fényében arra vagyunk kíváncsiak, hogy milyen szerepet tölt be egy karakter az ábrázolt világ kialakításában. Két általános funkció különíthető el aszerint, hogy mennyiben járul hozzá új jelentésrétegek létrehozásához attól függően, hogy megszólalásai gyakran térnek-e el a korábban elhangzottaktól, vagy inkább a már kialakult diskurzus megisméltése és megerősítése jellemző rájuk. *Innovatív* szereplők az újabb jelentések kidolgozásáért felelősek, míg az *isméltők* vagy *fenntartók* a dráma központi témáinak és megszólalásmódjainak kialakulását erősítik. Természetesen egy karakteren belül is megfigyelhető innováció és ismétlés kettőssége. Módszerünk ezt is feltárhatóvá teszi. Ugyanis egy szereplő minden mondatának meghatározzuk információs értékét, és ezek alapján következtetünk az általános funkciójára, ami így lehetőséget teremt a különböző típusú mondatok arányának vizsgálatára is. A mondat szintű megközelítés továbbá abban is segítségünkre lehet, hogy megválaszoljuk, a tág meghatározáson túl a konkrét esetben mit jelent egy szereplő innovatív funkciója, és hogyan jön létre az annak háttéréül szolgáló általános diskurzus. A tanulmányban azt mutatjuk be, hogy Shakespeare innovatív karakterei két csoportra oszthatók: egyrészt azokra, akik valóban az ismeretek közvetítéséért felelősek, másrészt azokra, akik eltérő *módon* beszélnek a drámában uralkodó diskurzushoz képest, általában bizonytalanságot és/vagy érzelmeket kifejezve, illetve metaforikus nyelvhasználatot érvényesítve. Eredményeink továbbá komédiák és tragédiák különbségét is újszerűen teszik leírhatóvá, amennyiben a komédiákban a tragédiákhoz képest nagyobb esélye van női szereplőknek innovatív funkciót betölteni, illetve nagyobb eséllyel ismétlik őket a többi szereplő.

Végezetül a dolgozat a szereplői kapcsolatok hálózatos megjelenítésének kérdéséhez is hozzá kíván szólni. Benjamin Krautter arra hívta fel a figyelmet, hogy jelentősen befolyásolja a hálózatok értelmezhetőségét, hogy milyen típusú kapcsolatokat prezentálnak – a különböző módszerek különböző típusú értelmezéseket tesznek lehetővé.⁷ Az alábbiakban egy új módszert kívánunk bemutatni az eddigiek kiegészítéseként. Ennek alapja, hogy a megszólalások újdonságát páros összehasonlítások mentén határozzuk

⁷ Krautter, „Kopräsenz-, Koreferenz- und Wissens-Netzwerke.”

meg, vagyis külön-külön vetjük össze a szereplőket egymással. Ez egyfelől lehetővé teszi, hogy két szereplő megszólalásai mondat szinten is összemérhetőek legyenek. Másrészt ezáltal a kapcsolatokat egy irányított gráfon tudjuk ábrázolni, amely azt mutatja, hogy a páros összehasonlításban mely karakter ismétli inkább a másik megszólalásait. Hasonlóan Andresen és munkatársaihoz „a karakterhálózatok tartalomközpontúbb formáját kívánjuk létrehozni [...], ami a kvantitatív elemzés és az értelmező olvasás szorosabb integrációjához vezethet el.”⁸ Az így létrehozott hálózatokban egyszerre válik vizsgálhatóvá a dráma egészében betöltött szerep és két karakter viszonya is.

2. Korábbi kutatások

Tanulmányunk olyan korábbi kutatásokhoz kapcsolódik, amelyek szintén innováció és ismétlés mérhetőségére tettek kísérletet. Azonban ezek nemcsak módszereikben, hanem elméleti előfeltevéseikben is különböznek egymástól, és – kisebb-nagyobb mértékben – mást értenek „információ”, „újdonság” vagy „hírérték” kifejezések alatt.

Tobin South és társai például ismétlődő nyelvi elemek elemzéséből indultak ki, hogy ezáltal hírportálok Twitter-fiókjai között tárják fel az információ áramlását. Előfeltevésük, hogy minél több azonos sorrendbe rendezett szó található két szövegben, annál kisebb az újdonság mértéke közöttük, és fordítva: korábban nem használt kifejezések és új szórend a szövegek hírértékéhez járulnak hozzá.⁹ Ennek megfelelően módszerük szavak leghosszabb ismétlődő sorozatának azonosításán alapul. Ez az eljárás Twitter-bejegyzések és hírek esetében jól működik, ám kevésbé homogén és jóval poétikusabb drámai szövegeknél nem igazán alkalmazható, hiszen ekkor az ismétlődő sorozatok szinte kizárólag konvencionálizálódott, kevésbé tartalmas jelentést kidolgozó összetételeket jelölnek (pl. *ott van, jó reggelt*), így elsősorban nem szemantikai hasonlóságról értesítenek.

Szintén ismétlődő nyelvi elemek feltárását tűzte ki célul Matthew Sims és David Bamman, akik regénykarakterek szerepét kívánták meghatározni az adott mű társas-információs hálózatában.¹⁰ Ehhez nem pusztán szavak előfordulásait, hanem nyelvfeldolgozó (NLP) algoritmusok kimeneteként szófaji kategóriákat és grammatikai viszonyokat is vizsgáltak. A legfontosabb cselekményelemeket leíró igék kiválogatását követően ALANY–IGEI ÁLLÍTMÁNY–TÁRGY hármasokat (pl.: *Thomas – elhagyta – Bécset*) azonosították a karakterek szólamaiban: amennyiben egy hármas két szereplő megszólalásában is szerepel, azt mondhatjuk, hogy azonos eseményre utalnak, így a korábbi megszólaló információs hatást (informational impact) fejt ki a későbbire. Az eljárás nehézségei közé tartoznak a koreferenciaszámítás (amely során az egyes szólamokat a megfelelő karakterekhez rendeljük, ugyanakkor ez drámai műveknél jóval egyszerűbb művelet) és a függőségi elemzés pontatlanságai, valamint a figyelembe vett

⁸ Andresen et al., „Who Knows What in German Drama?”, saját fordítás - Sz. B. és N. M.

⁹ Tobin South et al., „Information flow estimation: A study of news on Twitter”, *Online Social Networks and Media* 31 (2022), <https://doi.org/10.1016/j.osnem.2022.100231>.

¹⁰ Matthew Sims és David Bamman, „Measuring Information Propagation in Literary Social Networks”, in Bonnie Webber et al., szerk., *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 642–652 (Association for Computational Linguistics, 2020), <https://aclanthology.org/2020.emnlp-main.47>. <https://doi.org/10.18653/v1/2020.emnlp-main.47>.

igék csoportjának némileg önkényes kijelölése. Sims és Bamman továbbá a karakterek közötti közvetlen hatást igyekezett feltárni, ezzel szemben mi az újdonságot az adott megszólalás és az azt megelőző teljes diskurzus viszonyában értelmezzük: még ha páros összevetést is végzünk, nem feltételezzük, hogy két szereplő mondatainak hasonlósága közvetlen oksági kapcsolatot jelöl, azaz, hogy közvetlen információátadást azonosítanánk, csupán azt vizsgáljuk, hogy mennyiben hasonlít egy megnyilatkozás tartalma a korábban elhangzottakhoz.

Hasonlóan Barron és társai kutatásához, amely azt a kérdést tette fel, hogy a Francia Forradalom parlamentjének felszólalói új témák felvetéséért vagy éppen korábbi napirenden tartásáért voltak-e felelősek.¹¹ Az általuk is használt Kullback-Leibler divergencia/eltérés (KLD) gyakran alkalmazott mérőszám a hírérték kiszámításakor információelméleti megalapozottsága miatt: a számokként reprezentált szövegek különbsége ennek segítségével nem a távolság térbeli metaforáján keresztül (milyen távol van egy szöveg egy másiktól egy vektortérben), hanem a váratlanság *tapasztalatának* mintájára lesz kiszámítható (mennyiben meglepő egy szöveg egy korábbi ismeretében)¹² Barron és társai először meghatározták, hogy a parlamenti megszólalások milyen arányban tartalmazznak különböző témákat, majd ezeknek az eloszlását hasonlították össze a KLD segítségével. Andrew Piper és társai ezzel szemben egyenlő részre osztott szövegek egyszerű szógyakoriságát vették alapul az eltérés (vagyis az újdonság) kiszámításakor, amin keresztül az információ adagolásának mértékét tudták ábrázolni különböző típusú szövegekben.¹³

Jelen kutatásban a szövegrészek összehasonlításakor szintén kiemelten fontos a szemantikai kapcsolat, így sem az ismétlődő karaktersorozatok, sem a szavak pusztán eloszlásának vizsgálata nem tűnt célravezetőnek. Ugyanígy nem tudtuk követni Barron és társai munkáját, akik a parlamenti beszédek témamodellezésére hagytak: egy dráma esetében a megszólalások legtöbbször túl rövidek ahhoz, hogy hatékonyan tudjunk bennük témákat azonosítani; ahogyan maga a drámaszöveg egésze sem biztosít elég adatot, hogy a szereplők jól elkülöníthetők legyenek egymástól a témák eloszlása szerint. Ezért fordultunk nagy nyelvmodellekhez, hogy segítségükkel megadhassuk egy dráma minden mondatának helyét a modell sokdimenziós vektortérben, azaz az adott nyelv szemantikai mezőjében. A beágyazási folyamatot az SBERT (Sentence-BERT) algoritmus vezérli, amely a korábbi BERT-modellek szószintű beágyazásainál nagyobb egységek, például mondatok jelentését is képes kvantitatív módon megragadni.¹⁴ A mondatokhoz rendelt értékek tehát lehetővé teszik,

¹¹ Alexander T. J. Barron et al., „Individuals, institutions, and innovation in the debates of the French Revolution”, PNAS 115, 18. sz. (2018): 4607–4612, <https://doi.org/10.1073/pnas.1717729115>.

¹² Kent K. Chang és Simon DeDeo, „Divergence and the Complexity of Difference in Text and Culture”, *Journal of Cultural Analytics* 5, 2. sz. (2020), <https://doi.org/10.22148/001c.17585>.

¹³ Andrew Piper, Hao Xu és Eric D. Kolaczyk, „Modeling Narrative Revelation”, in Artjoms Šeļa, Fotis Jannidis és Iza Romanowska, szerk., *Proceedings of the Computational Humanities Research Conference 2023, Paris, France, December 6-8, 2023.*, 500–511 (CEUR-WS, 2023), <https://ceur-ws.org/Vol-3558/paper6166.pdf>.

¹⁴ Nils Reimers és Iryna Gurevych, „Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks”, in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 3982–3992 (Hong Kong: Association for Computational Linguistics, 2019). <https://doi.org/10.18653/v1/D19-1410>. A szóbe-

hogyan azok szemantikai hasonlóságát is kiszámíthatjuk – ami jelen esetben a drámai karakterek mondatainak összevetésekor kamatoztatható. A *szemantikai hasonlóság* nagyrészt témabeli egyezésekre vonatkozik, de részét képezi a mondatok stílusa is (az azonos stílushoz/regiszterhez tartozó szavak jelentésüket tekintve is hasonlóbbak). Ezek alapján azt mondhatjuk tehát, hogy szemantikailag minél kevésbé hasonlít egy mondat az azt megelőzőkhöz, annál nagyobb a benne közölt újdonság/információ mértéke; illetve minél jobban hasonlít a korábbiakhoz, annál inkább a már meglévő diskurzus ismétléséhez járul hozzá.

Ezt az eljárást alkalmazta Edgar Dubourg, Andrei Mogoutov és Nicolas Baumard mozifilmek cselekményének újszerűségét mérő tanulmánya is. Ebben a filmek cselekményleírásait alakították át a fenti módon, az SBERT algoritmus segítségével vektorokká, majd kiszámították e vektorok és az összes korábbi film értékei közötti koszinusztávolságot, végül ezeket átlagolva határozták meg egy film innovációs pontszámát, vagyis az aktuális beágyazásnak a korábbiaktól való átlagos távolságát.¹⁵ Hasonló módon járunk el mi is a drámaszereplők mondatainak összehasonlításakor. Ezt azért is fontos hangsúlyozni, mert Dubourg és társai a módszer értékelését is elvégezték: eredményeik magas szintű korrelációt mutatnak a nézői vélemények szövegbányászati alapuló mérésekkel.¹⁶ Ilyen összevetésre a drámai szereplők vizsgálatakor nincsen lehetőség korábbi eredmények hiányában, illetve mivel az eddig említett eljárások, ahogy láttuk, nem adaptálhatók problémátlanul a kutatási kérdés megválaszolására.

Egy ilyen, a szereplői szólamok váratlanságára vonatkozó kérdés eddig nem merült fel a kvantitatív drámaelemzés területén. Elsősorban a művek szerkezeti sajátosságai kerültek az előtérbe,¹⁷ míg a nyelvi megformáltságra leginkább nagyobb korpuszok téma-modelljezése (és ezen keresztül műfaji különbségek feltárása),¹⁸ Shakespeare életművében pedig a szerzőség kérdése és a drámák kulcsszavai,¹⁹ illetve nyelvi válto-

ágyazások (és minden nyelvmódel) abból az előfeltevésebből indulnak ki, miszerint egy szó jelentése megragadható a környezetében gyakran előforduló szavak jelentésén keresztül. Ennélfogva egy kellően nagy tanítókorpusz lehetővé teszi, hogy a szavak ilyen viszonyrendszere statisztikailag is feldolgozható legyen, ami által megadhatók szavak valószínűsíthető szinonimái (azaz olyan kifejezések, amelyek hasonló szöveggörnyezetben szoktak előfordulni), parafrazálhatóvá-összefoglalhatóvá válhatnak szövegek stb. A mondat-szintű beágyazások ezt az elvet terjesztik ki nagyobb szövegegyeségekre, akár egész bekezdésekre is.

¹⁵ Edgar Dubourg, Andrei Mogoutov és Nicolas Baumard, „Is Cinema Becoming Less and Less Innovative With Time? Using neural network text embedding model to measure cultural innovation”, in Artjoms Šeļa, Fotis Jannidis és Iza Romanowska, szerk., *Proceedings of the Computational Humanities Research Conference 2023 Paris, France, December 6-8, 2023*, 676–686 (CEUR-WS, 2023), <https://ceur-ws.org/Vol-3558/paper7806.pdf>.

¹⁶ Yingyue Luan és Yeun Joon Kim, „An integrative model of new product evaluation: A systematic investigation of perceived novelty and product evaluation in the movie industry”, *PLoS One* 17, 3. sz. (2022), <https://doi.org/10.1371/journal.pone.0265193>.

¹⁷ Áttekintéshez lásd Szemes és Vida, „Tragikus és komikus hálózatok.”

¹⁸ Christoph Schöch, „Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama”, *DHQ* 11, 2. sz. (2017), <https://www.digitalhumanities.org/dhq/vol/11/2/000291/000291.html>.

¹⁹ Hugh Craig és Arthur F. Kinney, *Shakespeare, Computers and the Mystery of Authorship* (New York: Cambridge University Press, 2009). <https://doi.org/10.1017/CBO9780511605437>.

zások feltárása²⁰ kapcsán irányult a figyelem. A kutatásunkhoz legközelebb Andresen és társai, valamint Krautter munkái állnak a *Bevezetésben* említett különbségekkel. Fontos megemlíteni még Artjoms Šeļa és társai kutatását is, amelynek során a szereplői szövegek különbségét kívánták kiszámítani szerzőazonosítási feladatokra kidolgozott stilometriai módszerekkel²¹ – ám ekkor sem a szövegek szemantikai tartalma és újdonságuk mértéke vált hangsúlyossá, hanem kizárólag stiláris különbségeik. Reményeink szerint jelen tanulmányunk segítségével új szempontokkal tudunk hozzájárulni ehhez a területhez, amin keresztül az egyes művek értelmezhetőségét is gazdagítani tudjuk.

3. Módszer

A kutatáshoz a Shakespeare-szövegeknek a Drama Corpus adatbázisban²² elérhető, a Folger Shakespeare Library gyűjteményén alapuló TEI XML verzióit használtuk fel (ShakeDraCor).²³ Első lépésként a dokumentumokat egy táblázattá alakítottuk oly módon, hogy minden mondathoz hozzárendeltük: 1. a megszólaló nevét, 2. a mondat sorszámát mint időbélyeget (1-től az utolsóig), 3. a felvonást, amelyben a mondat elhangzik, 4. a mondat szintű beágyazás értékét. Ez utóbbi kapcsán fontos a megfelelő modell kiválasztása. Ennek érdekében különböző, az aktuálisan legjobb teljesítményre képes²⁴ SBERT modelleket hasonlítottuk össze a korpuszból kiválasztott példamondatokon (lásd *Függelék*). A mondatok között található egymáshoz hasonló és eltérő jelentésűek is (a hasonlóság mértékét intuitíven állapítottuk meg); az összehasonlítások minden modellenél megadtuk a mondatok koszinusz hasonlóságát és ezen értékek szórását. Bár az eredmények nagyon hasonlóak voltak, azt a modellt választottuk ki ('all-MiniLM-L6-v2'), amelyik a legnagyobb szórást mutatta, azaz amelyik esetében a legnagyobbak a különbségek közeli és távoli jelentések között. Részletekhez lásd a kutatás repozitóriumát és a *Függelékben* közölt táblázatot, amely mutatja, hogy mely mondatokat milyen mértékben tart a modell hasonlónak. Ezek alapján úgy tűnik, hogy az emberi ítéletekhez közel esnek az SBERT eredményei, amelyek így jól használhatók a szemantikai hasonlóság meghatározásakor. A modell által azonosított leghasonlóbb mondatokat tekintve szembevetendő az is, hogy a szereplők nevei erősen befolyásolják a hasonlóság mértékét – ezek előzetesen kiszűrhetők lennének, de úgy döntöttünk, hogy ezeket is figyelembe vesszük a számításokkor, mivel a mondat jelentéséhez hozzátartoznak a megszólítások és megnevezések is. Kizártuk viszont a négynél kevesebb szót tartalmazó mondatokat (például: „Igen, uram”), hiszen ezek kevésbé dolgoznak ki önálló, tartalmas jelentést a hosszabbakhoz képest.

²⁰ Jonathan Hope és Michael Witmore, „Quantification and the language of later Shakespeare”, *Actes des congrès de la Société française Shakespeare* 31, La langue de Shakespeare (2014), <https://doi.org/10.4000/shakespeare.2830>.

²¹ Artjoms Šeļa et al., „From Stage to Page: Stylistic Variation in Fictional Speech”, in Melanie Andresen és Nils Reiter szerk., *Computational Drama Analysis*, 149–166 (Berlin: de Gruyter, 2024).

²² Frank Fischer et al., „Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama”, in *Proceedings of DH2019: „Complexities”*, 1–6 (Utrecht: Utrecht University, 2019), <https://dracor.org/>.

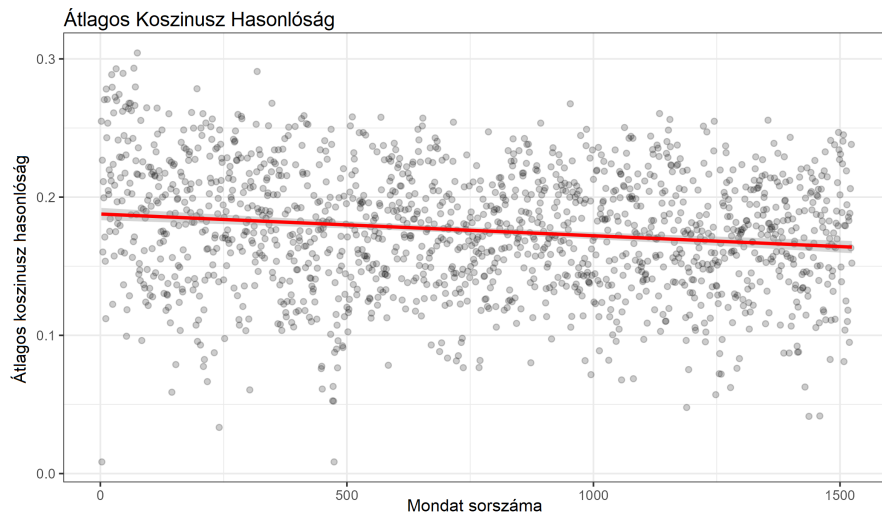
²³ Hozzáférés 2024.07.07. <https://github.com/dracor-org/shakedracor>.

²⁴ Hozzáférés 2024.07.07. https://www.sbert.net/docs/pretrained_models.html.

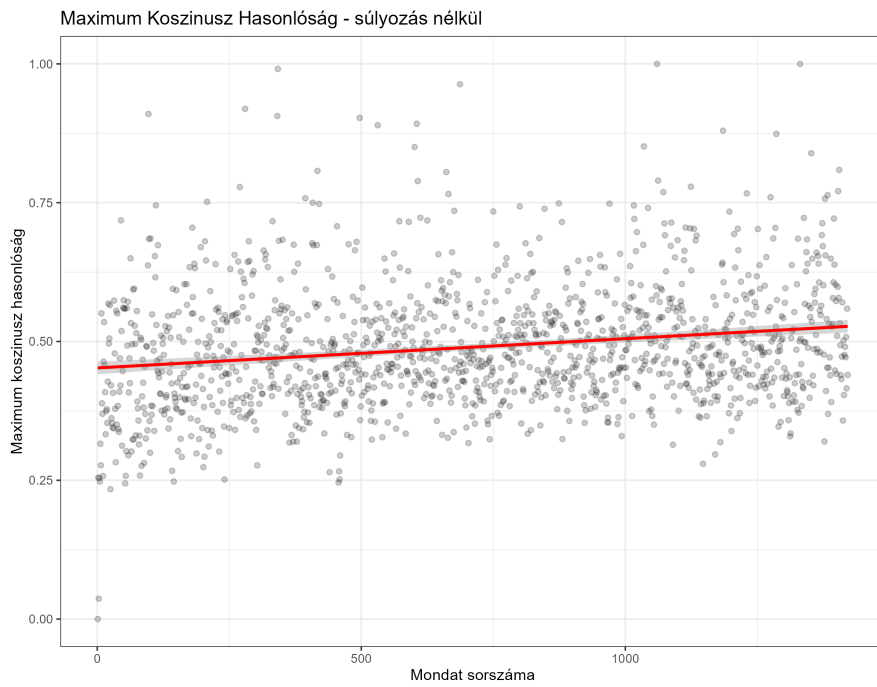
Ezt követően meghatározott sorrendű karakterpárokat hoztunk létre a leggyakoribb megszólalókból (azaz a főbb szereplőkből):²⁵ a páros első tagját tekintjük Forrásnak, a másodikat pedig Célkarakternek. Az összehasonlítás során megadtuk Célkarakter összes mondatának koszinusz-távolságát Forráskarakter minden, az aktuális mondatot megelőző mondatához képest (az időbeliség a mondatok sorszáma alapján állapítható meg). Szemben Duborg és társai módszerével, az egyes mondatokhoz tartozó hasonlóságnak nem az átlagát vettük, hanem közülük csupán a legnagyobbat választottuk ki a szemantikai közelség jellemzésére. Így Célkarakter minden mondatához egy számot rendeltünk, ami azt mutatja meg, hogy mennyire hasonlít Forráskarakter korábbi mondatai közül a szemantikailag leghasonlóbbra (maximum moszinusz hasonlóság – MCS). Minél nagyobb ez a szám, annál kevésbé tekinthető innovatívnak a mondat jelentése, hiszen annál inkább ismételi meg korábbi tartalmakat.

Több érv is szól a maximum koszinusz hasonlóság használata mellett az átlagos hasonlóság megadása helyett. Egyrészt, ha Forrásszereplő sokféle témát érintve, különböző regiszterekben is megszólal mielőtt Célkarakter aktuális mondata elhangzana, akkor átlagosan kevésbé lesz közel ezekhez Célkarakter mondata, még akkor is, ha Forráskaraktertől szó szerint ugyanaz elhangzott már korábban. Az MCS viszont csak a maximális értékekre összpontosít, így az eredményt más megszólalások nem befolyásolják – még ha emiatt arról nem is tud számot adni, hogy milyen gyakran hozott létre Forráskarakter hasonló jelentéseket. Másrészt az MCS értéke visszavezethető Cél- és Forráskarakter közötti leghasonlóbb mondatpárokra, ami nagyban hozzájárul az eredmények értelmezhetőségéhez (lásd *Eredmények*). Harmadrészt, az MCS értékei nagyobb szórást mutatnak az átlagos koszinusz-hasonlóság értékeihez képest (a *Hamlet*-ben a maximális pontszámok szórása 0.11, míg az átlagoké 0.04 – vö. az 1. ábrával), ami a mondatokat jobban megkülönböztethetővé teszi. Azt is fontos megjegyezni, hogy mindkét mérőszámot befolyásolja az időbeliség: az átlagos koszinusz-hasonlóság esetében minél korábbi a megszólalás, annál jobban hasonlít az előző diskurzushoz (lásd az 1a. ábrát), a maximális koszinusz-hasonlóság esetében pedig minél későbbi a megszólalás, annál inkább magas érték jellemzi (1b. ábra). Ez a hatás ugyanakkor az eredmények súlyozásával kiegyenlíthető. Ehhez először kiszámítottuk az átlagos MCS értéket az egyes felvonásokban és a dráma egészében is, majd a felvonások és az egészre vonatkozó értékek különbségét használtuk a súlyozáshoz, aszerint, hogy az adott mondat mely felvonásban hangzott el (tehát például az első felvonásban szereplő mondatokhoz hozzáadtuk az első felvonás és a dráma egészének átlagos MCS értékeinek különbségét). Látható, hogy ettől függetlenül nagyfokú variancia jellemzi az adatsort: ugyanúgy találhatók magas MCS értékű mondatok az első felvonásban, mint ahogy alacsonyak egy dráma végén. (1c. ábra)

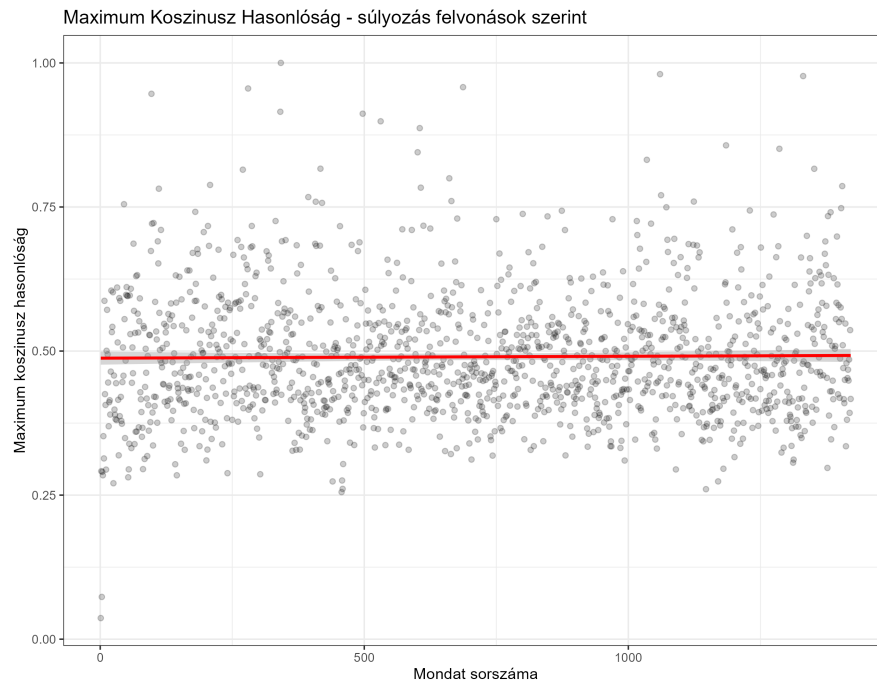
²⁵ Főszereplőnek tekintettük azokat, akiknek rövidebb színművek esetében több mint 30, közepes hosszúságú darabok esetében több mint 40, hosszú darabok esetében pedig több mint 50 hosszú (4 szónál több) mondatral rendelkeznek. Rövid drámának tekintettük az 1000-nél kevesebb, közepesnek az 1000-nél több, de 1700-nél kevesebb, hosszúnak pedig az ennél több hosszú mondatot tartalmazó szövegeket.



1a. ábra - Átlagos koszinusztávolság a szereplők mondatai között (*Hamlet*) – lineáris regresszió.



1b. ábra - Maximum koszinusztávolság a szereplők mondatai között súlyozás nélkül (*Hamlet*) – lineáris regresszió.



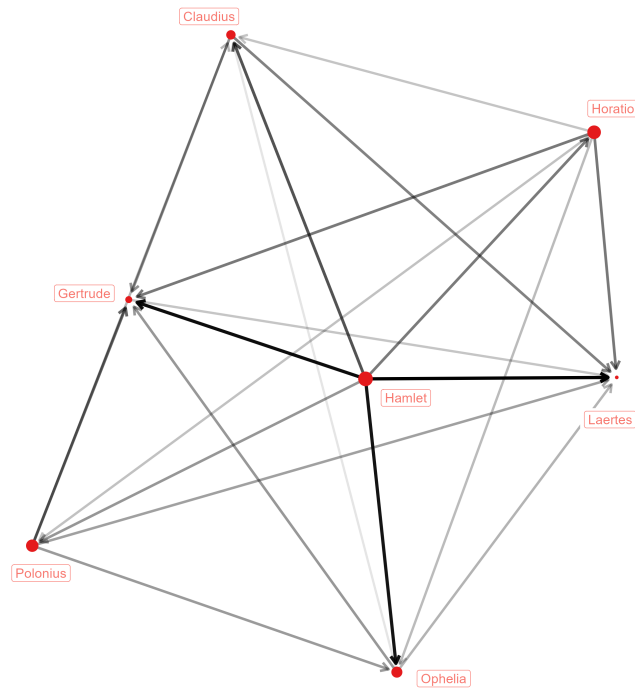
1c. ábra - Maximum koszinusztávolság a szereplők mondatai között, a felvonások átlagos értéke alapján súlyozva (*Hamlet*) – lineáris regresszió.

A következő lépésben minden Forrás-Cél párhoz hozzárendeltük a súlyozott MCS pontszámok átlagát. Ezután South és társai nyomán hálózati normalizációt végeztük az adatsoron.²⁶ Érdekes ugyanis figyelembe venni, hogy ha „A karakter” mondatainak tartalmát „B karakter” előszeretettel ismétli meg, de „A” karakter maga is a többiek szövegeit ismétli, akkor „B” karakter valójában szintén ezekkel áll kapcsolatban „A” karakter közvetítésén keresztül. A hálózati normalizációhoz megadtuk, hogy egy szereplő Célkarakterként átlagosan milyen hasonlóságot mutat a többi szereplővel, majd ezzel a számmal osztottuk el azokat a pontszámokat, amelyeknél a karakter mint Forrás szerepel.

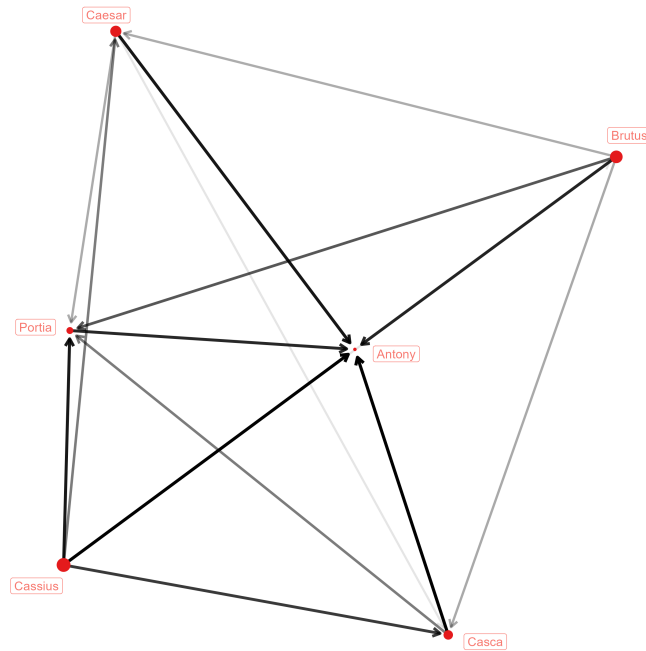
Végezetül kiszámítottuk az egy szereplőpárra vonatkozó pontszámok különbséget aszerint, hogy melyikük szerepel Forrás- illetve Célkarakterként (pl. Hamlet-Claudius vs. Claudius-Hamlet). Ha a különbség pozitív, akkor inkább Célkarakter mondatai dolgoznak ki Forráskarakter korábbi mondataihoz hasonló jelentést, mint fordítva – azaz Forráskarakter tekinthető kettejük viszonyában innovatívnak. Végző eredményként csak ezeket a pozitív értékeket tartottuk meg és használtuk fel a hálózati vizualizációhoz.

²⁶ South et al., „Information flow estimation”.

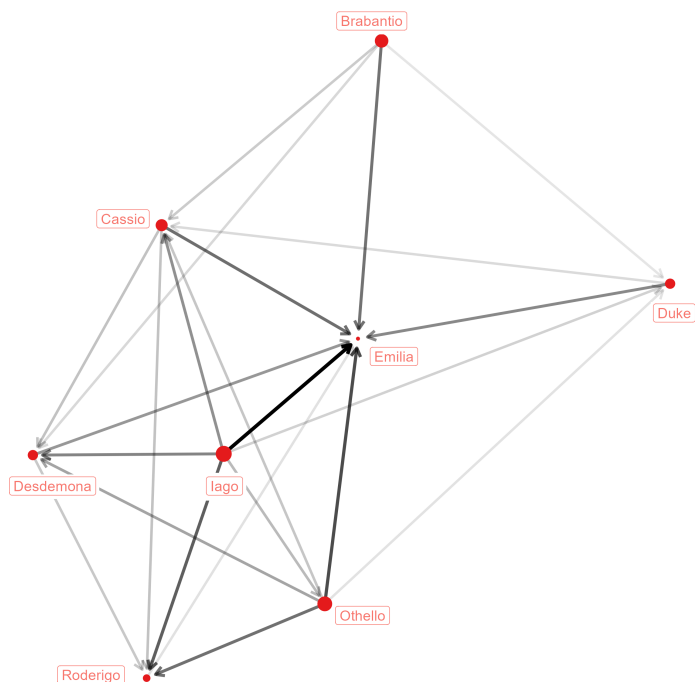
2(a-f). ábra: Shakespeare-drámák karakterhálózata innováció és ismétlés alapján.



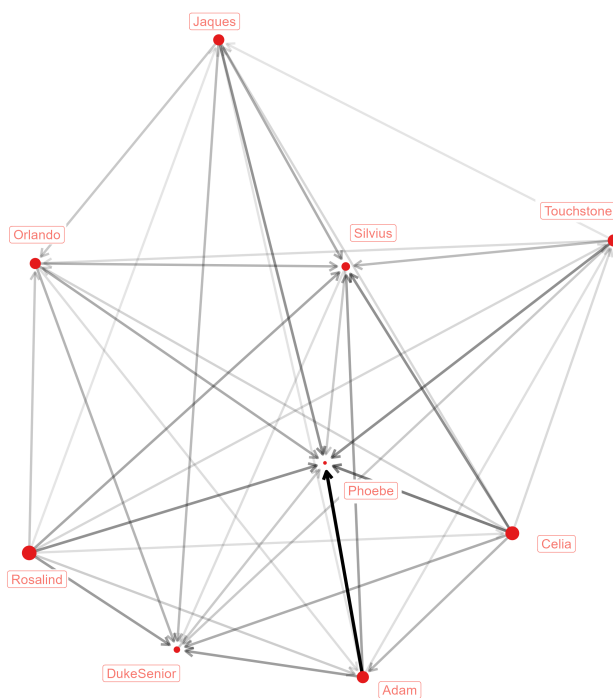
2a. ábra: *Hamlet*



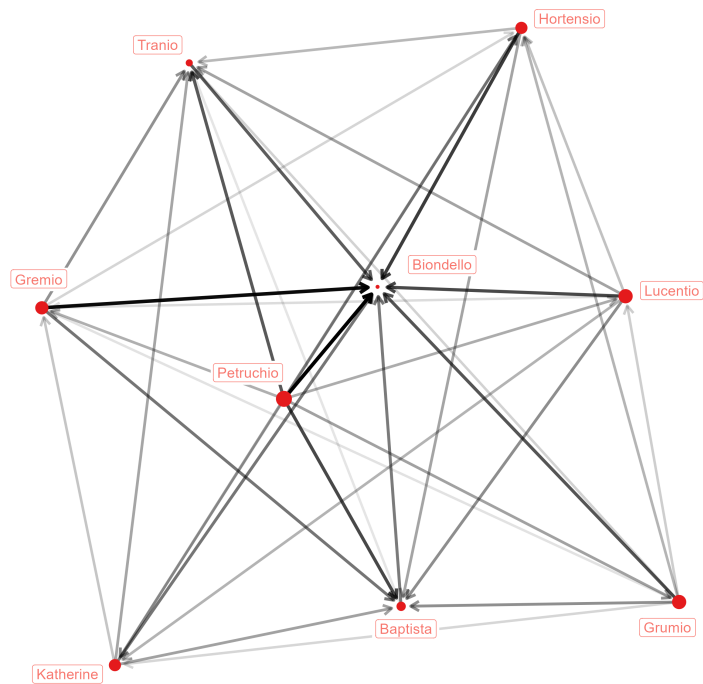
2b. ábra: *Julius Caesar*



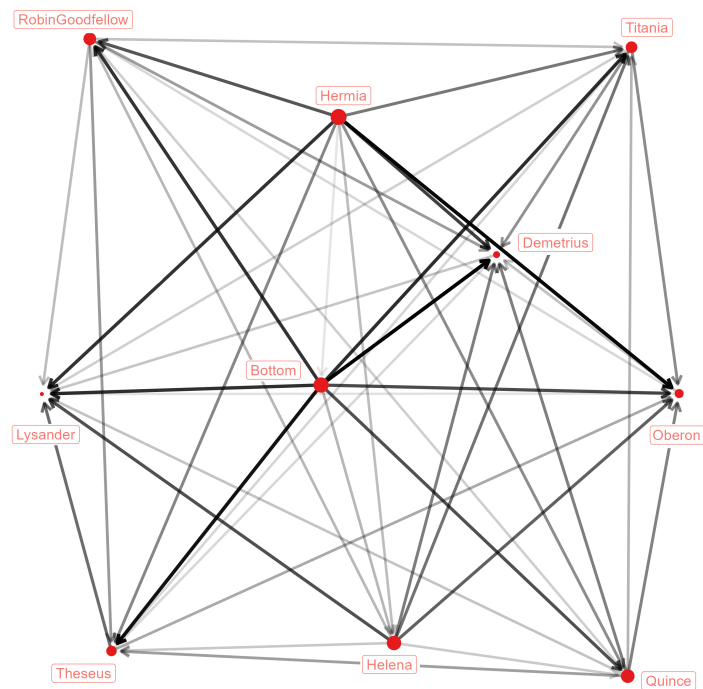
2c. ábra: *Othello*



2d. ábra: *Ahogy tetszik*



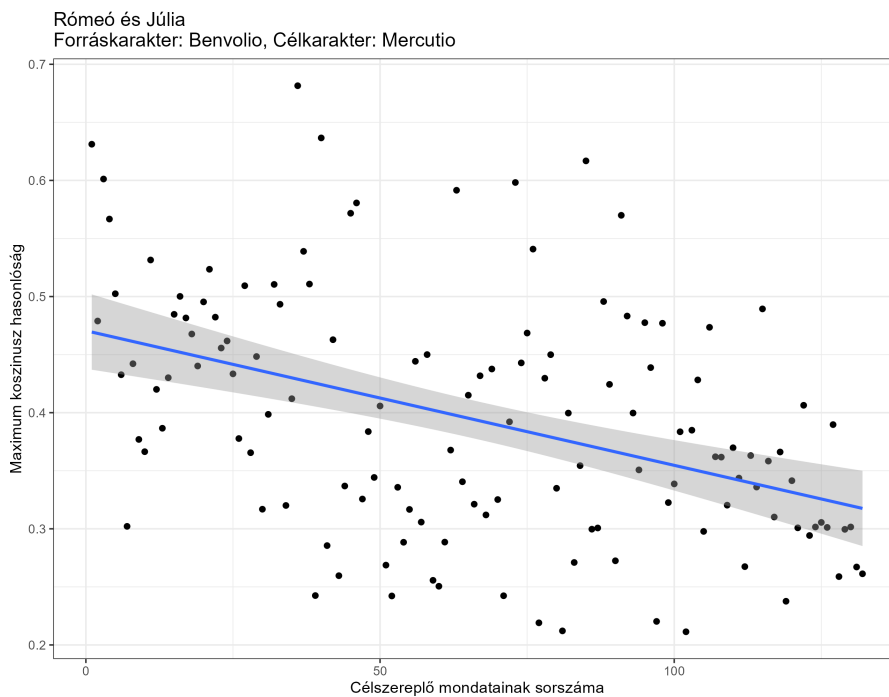
2e. ábra: *A makrancos hölgy*

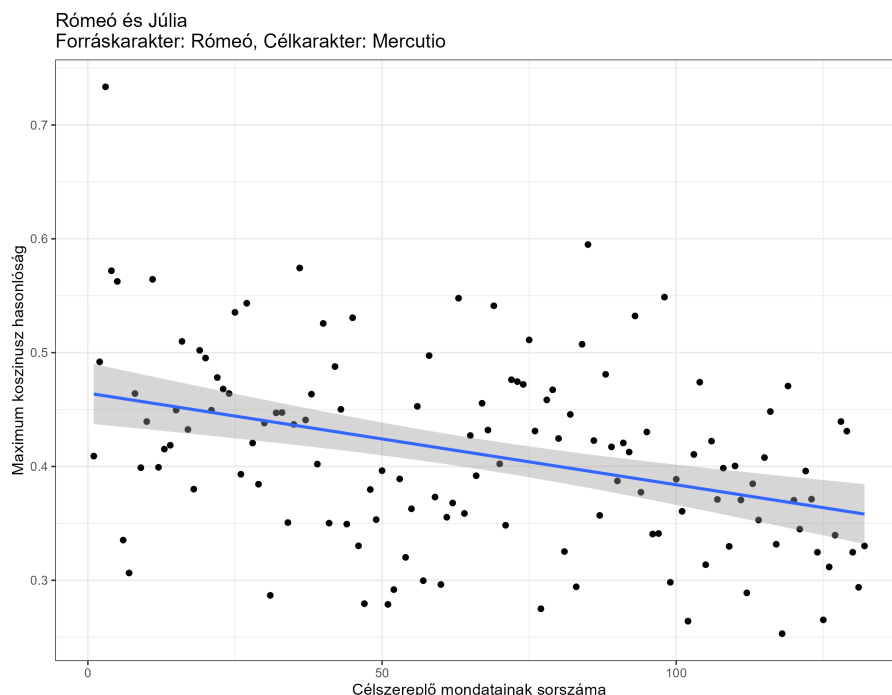


2f. ábra: *Szentivánéji álom*

4. Eredmények

Az eredmények lehetővé teszik, hogy a szereplők közötti kapcsolatrendszer hálózatként jelenítsük meg. A 2. ábrán látható gráfokon a nyilak Forrástól Célig tartanak (azaz azt jelzik, hogy mely karakter ismétli inkább a másikat), vastagságuk a hasonlóság/ismétlés mértékét jelöli, a csomópontok mérete pedig innovációs pontszámként azt mutatja, hogy az adott karakter milyen gyakran szerepel Forrásként, azaz milyen gyakran tekinthető innovatívnak a páros összehasonlításokban. Ez utóbbit egyszerre befolyásolja a vizsgált mondatok száma és a kimondás ideje: minél többször és korábban beszél egy szereplő, annál nagyobb az esélye újszerűnek lenni. Habár az időbeliség hatását általánosan kezeltük a fent említett súlyozással, mégis úgy tűnik, hogy azokhoz a szereplőkhöz, akik inkább a drámák második részében vagy a végén szólalnak meg, kisebb innovációs pont társul (például Antonius a *Julius Caesarban* vagy Emília az *Othellóban*). Akad példa azonban arra is, hogy egy szereplő egyre inkább eltér egy másiktól a cselekmény előrehaladtával – ilyen például Mercutio, Rómeóhoz és Benvolióhoz képest a legmagasabb innovációs pontszámmal rendelkező karakter a *Rómeó és Júliában*. (3. ábra)





3. ábra: Szereplők megszólalása közötti hasonlóság mértéke (*Rómeó és Júlia*). Fent: Forrás – Benvolio, Cél – Mercutio; lent: Forrás – Rómeó, Cél – Mercutio.

Shakespeare minden drámáját vizsgálva azt láthatjuk, hogy a szereplők közötti kapcsolat a legtöbb esetben hierarchikus (azaz a szereplők az innovációs pontszámok alapján hierarchikusan elrendezhetők). Ez különösen igaz a tragédiákra (vagy a nem komédiák tag kategóriájára²⁷), amelyeknél a legmagasabb innovációs pontszámmal rendelkező karakterek szinte mindig így rendezhetők el, és csak az alacsonyabb szinteken található azonos értékek. Az egyenlő pontszámok azt jelentik, hogy a drámákban bizonyos fokú körkörösség érvényesül: „A” szereplő „B”-t ismétli, „B” ismétli „C”-t, míg „A” „C”-t stb. Magasabb szinten ez főleg a vígjátékokban fordul elő (a nem komédiák közül a *Cymbeline*-ben, a *Macbeth*-ben és a sokat vitatott műfajú *Periklész*-ben). A *makrancos hölgy*-ben például Grumio és Gremio, valamint Lucentio és Katalin, az *Ahogy tetszik*-ben Orlando, Ádám és Próbakő; a *Szeget szeggel*-ben a Herceg, Lucio és Angelo vesznek fel megegyező értékeket. Ez a műfaji különbség összhangban van a korábbi, közös jelenetekre alapuló hálózatokból létrehozott eredményekkel. Ezek azt mutatják, hogy a vígjátékokat sűrűbb kapcsolatrendszer jellemzi, míg a tragédiákban gyakrabban szerepel egy-két, a társas kapcsolatokat irányító, összekötő funkciót betöltő szereplő (a csomópontok fokszámának hierarchikusabb az eloszlása). Ez azt is jelenti, hogy a vígjátékokban az interakciók során sok a félreértés és a párhuzamosság (két szereplő különböző útvonalakon is kapcsolódik egymáshoz), azonban éppen emiatt ezek a hálózatok „védettek” a széteséstől, ha egy útvonallal kiderül, hogy megbízhatatlan

²⁷ „Komédia” címkével láttuk el azokat a drámákat, amelyek az *Első fõlió*-ban (1623) is így szerepelnek. Az összes többi „nem komédia” megjelölést kapott – azonban a Nagy Tragédiákkal mutatott strukturális hasonlóságaik miatt, az egyszerűség kedvéért, esetenként tragédiaként hivatkozunk rájuk.

információt közvetít. Ezzel szemben a tragédiákban az információáramlás hatékony és gyors, de maguk a hálózatok törékenyek, hiszen egy összekötő szereplő kudarca az egész rendszer széteséséhez vezethet.²⁸

Mindezt tovább árnyalja a műfajok közötti újabb különbség a méréseink alapján. Szembetűnő, hogy a 23 nem komédiában a legtöbbször ismételt szereplők férfiak (kivéve Imogen a *Cymbeline*-ben és Lady Macbeth, aki ugyanannyira innovatív, mint Macbeth és Banquo), míg a vígjátékokban nagyobb eséllyel töltenek be női karakterek innovatív szerepet: 14-ből hatszor. Az *Ahogy tetszik*ben ilyen Rosalinda (és Celia a második helyen), a *Minden jó, ha vége jón*ban a Grófnő (és Heléna a második helyen), *A tévedések vígjátékában* Adriana; a *Szentivánéji álomban* Hermia (és Heléna a harmadik helyen, míg párjaik, Lysander és Demetrius a főszereplők közül a legalacsonyabb innovációs pontszámmal rendelkeznek); a *Sok hűhó semmiért*ben Beatrice, és talán a legmeglepőbb módon, *A vihar*ban Miranda, megelőzve Gonzalót és Prosperót is. Elmondhatjuk tehát, hogy a kétféle közösségben a diskurzust tematizáló (vagy legalábbis a mások által ismételt) szereplők nemét tekintve is különbség mutatkozik – még ha nem is kizárólagosan. A vígjátékok „védett” hálózataiban nagyobb esélye van nőknek betölteni ezt a szerepet, míg a hatékony, de kiszolgáltatott tragédiákban ez inkább férfiakra jellemző.

Mindezen túl érdemes az eredményeket részletesebben is szemügyre venni, és megadni az összehasonlítások alapján a leginkább és a legkevésbé hasonló mondatokat. Ez a módszer kvalitatív értékelése mellett a drámák szoros olvasásához és a szereplők jellemzéséhez is hozzájárulhat. A továbbiakban példaként olyan szöveget tárgyalunk Shakespeare életművéből, amelyek tanulságosak lehetnek az innováció és ismétlés kérdésekor. A *Hamlet* esetében például a címszereplő lényegi tulajdonságát, az önmagát meghatározni kívánó, ám önmagát szüntelen megkérdőjelező figura kettősségét ragadja meg a modell. Hamletnek ugyanis a többi szereplő korábbi megszólalásaihoz leginkább hasonló mondatai gyakran a saját és mások identitásának meghatározásáról szólnak; miközben a leginkább eltérő és innovatív mondatai kételyről, elbizonytalanodásról számolnak be, gyakran feltételes vagy kérdő modalitásban (1. táblázat; részletekhez lásd a kutatás repozitóriumát²⁹).

²⁸ Részletesebben lásd: Szemes és Vida, „Tragikus és komikus hálózatok.”

²⁹ Az itt közölt példamondatokat a páros összehasonlítások 10 legnagyobb és legkisebb koszinusz hasonlóságot mutató mondatai közül kézzel válogattuk ki az értelmezés céljából. A kiválogatás tehát némileg önkényes: ahhoz hasonlítható, amikor kulcsszóelemzés vagy téma modellezés kimenetét igyekszik egy kutató értelemmel felruházni. A teljes listát a *Függelék*ben megadott repozitóriumban közöljük.

Magas hasonlóság, alacsony innováció	Alacsony hasonlóság, magas innováció
This is I,/ Hamlet the Dane. (Az én vagyok,/ Hamlet, a dán) ³⁰	I doubt some foul play. (Itt baj van. Valami gazzett.)
The King is a thing – (A király egy olyan –)	I would I had been there. (Ott kellett volna lennem.)
O God, Horatio, what a wounded name, / Things standing thus unknown, shall I leave behind me! (Horatio, folt/ lesz a nevemen, / ha minden így marad, ilyen titokban!)	Do they hold the same estimation they did when I was in the city? (Most is olyan népszerűek, mint amikor még bejártam a városba?)
If Hamlet from himself be ta'en away, /And when he's not himself does wrong Laertes,/ Then Hamlet does it not; Hamlet denies it. (Ha/ Hamlet kivetözik önmagából, és Laertesnek magán kívül árt, /azt nem Hamlet teszi, sohasem ő.)	The time is out of joint. (A világ szétesett;)
Here comes the King, The Queen, the courtiers. (Itt jön a király, a királyné, az udvar.)	These foils have all a length? (Mind egyforma hosszú?)

1. táblázat – Példák Hamlet legkevésbé és leginkább innovatív mondataiból (*Hamlet*)

Hamlet szövege tehát akkor hasonlít a leginkább az udvari diskurzushoz, amikor megnevez, azonosít, és leginkább akkor tér el, amikor kérdez, elbizonytalanodik. Mivel ő tekinthető a drámában a leginnovatívabbnak, azt mondhatjuk, hogy a kételkedésről szóló mondatai vannak többségben, ezek adják karakterének lényegét – de fontos látni az ellentétes irányú megszólalásait is. A drámában a második leginnovatívabb karakter Horatio, aki viszont a többiektől eltérő mondataiban sem fejez ki bizonytalanságot: ő valóban újdonságokat, híreket közöl. (2. táblázat) E mondatai jól jelzik az eseményeket és közösségeket összekötő dramaturgiai funkcióját,³¹ amennyiben ezekben szemtanúként számol be korábbi tapasztalatairól az érdeklődőknek (elsősorban Hamletnek), ami által kapcsolatot teremt különböző idők, helyszínek és szereplők között.

³⁰ A magyar szöveget Nádasdy Ádám fordításában adjuk meg: William Shakespeare, „Hamlet”, in *Uő, Drámák*, ford. Nádasdy Ádám, 339–351 (Budapest: Magvető, 2001).

³¹ Franco Moretti, „Network Theory, Plot Analysis”, *Stanford Literary Lab Pamphlets*, 2. sz. (2011), <https://litlab.stanford.edu/LiteraryLabPamphlet2.pdf>.

Alacsony hasonlóság, magas innováció
Not when I saw 't. (Mikor én láttam, nem.)
I will forestall their repair hither and say you are not fit (Szólok, hogy mégse fájradjanak ide, mert nem vagy formában.)
Indeed, I heard it not. (Nem vettem észre.)
It was as I have seen it in his life,/ A sable silvered. (Mint amilyenek életében láttam,/ ezüstös-fekete.)
It would have much amazed you. (Megzavartál volna.)

2. táblázat – Példák Horatio leginnovatívabb mondataiból (*Hamlet*)

A kétkedés, vagyis a belső tartalmakra vagy a külvilág érzékelésére/értelmezésére vonatkozó beszéd mint a közös diskurzustól való eltérés más drámák szereplőinél is megfigyelhető.³² Ilyen például Hermia szólama a *Szentivánéji álomban*, aki éppen amiatt lehet a dráma leginnovatívabb szereplője (még Zubolyhoz képest is, aki pedig a többségtől eltérő történet-számban jelenik meg, ezért gyakran tényleg *másról* beszél), mert előszeretettel kérdez rá a dolgok milyenségére (3. táblázat). A Hamletnél megfigyelt kettősség jellemzi továbbá Brutust is a *Julius Caesarban*. Az addigi diskurzushoz leginkább hasonló mondatai döntően a gyilkosságról szólnak, a legkevésbé hasonlókéiről és érzelmeiről számolnak be (4. táblázat). Érdekes ezt összevetni Caesar mondataival, aki csak rövid ideig bizonytalanodik el, hogy elmenjen-e a szenátusba (ezek a leginnovatívabb megszólalásai), és inkább vállalja halálát, csak hogy fentartsa a császárról kialakul egyezményes képet. Ennek érdekében gyakran beszél magáról egyes szám harmadik személyben: „Caesar shall forth.” („Caesar megy”³³); „Danger knows full well/ That Caesar is more dangerous than he.” („A veszély tudja jól,/ Hogy Caesar nála is veszélyesebb”) stb. Ám ezek mellett megjelennek olyan szereplők is, mint Horatio, akiknek újszerűségét konkrét eseményekre vonatkozó kijelentéseik és megállapításaik adják. A *Julius Caesarban* ilyen Cassius, aki Brutusnál is innovatívnak tekinthető. A korábbi diskurzushoz leginkább és legkevésbé hasonlító mondatai éppen ellenkező mintázatot mutatnak, mint Brutus esetében: akkor ismétli a többieket, amikor érzelmekre, belső tulajdonságokra utaló kifejezéseket használ, míg az eseményekre vonatkozó mondatai térnek el tőlük a legjobban (5. táblázat). Cassius ugyanis a cselekmény előremozdításáért felelős, híreket hoz, érvel; ő szervezi

³² Kapcsolatba hozható ezzel a következtetéssel Hauser Arnold megfigyelése, miszerint Shakespeare drámái az én elvesztését, a szilárd identitás elbizonytalanodását viszik színre. Ennek alátámasztásául – jelen dolgozat értelmezéseihez hasonlóan – szintén szereplői mondatokat listáz, pl.: „Önmagad se vagy.” (*Szeget szeggel*), „Nem az vagyok, ami.” (*Othello*), Uram már nem uram (*Othello*), „Cressida volt az? Ő az, s mégsem az.” (*Troilus és Cressida*) Idézi: Szőnyi György Endre, „Tragikomédia és románc”, in Szőnyi György Endre és Kiss Attila Attila, szerk., *Az angol irodalom története 2. A kora újkor*, 361–378 (Budapest: Kijárat Kiadó, 2020), 365.

³³ A magyar szöveget Fekete Ádám és Forgách András fordításában adjuk meg: William Shakespeare, *Julius Caesar*, ford. Fekete Ádám és Forgách András (Budapest: L'Harmattan, 2015).

be a bizonytalankodó Brutust is az összeesküvésbe. Amikor belső tartalmakról beszél, általában nem is magáról, hanem másokról szól. Brutus konkrét eseményeket jelölő mondatai viszont nem a szervezkedésre, hanem magára a gyilkosságra vonatkoznak; gyakran utólagosan, így kevésbé újszerűen. Amíg a gyilkosság meg nem történik, illetve amíg rá nem szánja magát, addig egy absztraktabb regiszterben mozog. Ezt írja le a Casear szólamától egyik legjobban eltérő mondata is: „Between the acting of a dreadful thing/ And the first motion, all the interim is/ Like a phantasma or a hideous dream.” („A szörnyű tett s az első gondolat/ Között eltelt idő: alattomos/ Lidércnyomás, csupa illúzió...”)

Alacsony hasonlóság, magas innováció
Who is 't that hinders you? (Ki akadályoz ebben? ³⁴)
Then I well perceive you are not nigh. (Úgy tűnik, hogy nem vagy a közelben.)
I understand not what you mean by this. (Én nem is értem, hogy miről beszélsz.)
Too high to be enthralled to low. (Ha alantassba szeretünk bele!)
Nothing but "low" and "little"? (Már megint „pici”? „bábu”? „alacsony”?)

3. táblázat – Példák Hermia leginnovatívabb mondataiból (*Szentivánéji álom*)

³⁴ A magyar szöveget Nádasdy Ádám fordításában adjuk meg: William Shakespeare, „Szentivánéji álom”, in Uő, *Drámák*, ford. Nádasdy Ádám (Budapest: Magvető, 2001).

Magas hasonlóság, alacsony innováció	Alacsony hasonlóság, magas innováció
Mark Antony, here, take you Caesar's body. (Vidd Caesar holttestét, Antonius.)	I would not, Cassius, yet I love him well. (Nem, Cassius, pedig őt kedvelem.)
And for Mark Antony, think not of him, For he can do no more than Caesar's arm/ When Caesar's head is off. (Túl véres munkának tűnne, Cassius,/ Feje után tagjait is levágni – / Ki örjöngve öl, gyűlöletre ébred./ Antonius csak Caesar egyik tagja.)	That you do love me, I am nothing jealous. (Hogy szeretsz, abban nem kételkedem...)
Vexèd I am /Of late with passions of some difference, Conceptions only proper to myself, Which give some soil, perhaps, to my behaviors. (Fölkaklatnak/ Saját viharos érzéseim, a bennem/ Zajló folytonos viták, amelyek / Viselkedésemet is indokolják:)	If I have veiled my look, I turn the trouble of my countenance/ Merely upon myself. (Ha elsötétül/ Pillantásom, a feltűnő rosszkedv / Csak magam miatt van.)
Hold, then, my sword, and turn away thy face/ While I do run upon it. (Tartsd a kardodat, s fordítsd el fejed, / Míg belefutok.)	But if these – /As I am sure they do - bear fire enough/ To kindle cowards and to steel with valor/ The melting spirits of women, then, countrymen, What need we any spur but our own cause/ To prick us to redress? (Ha ég bennünk annyi tűz – s ez biztos! – , / Mely gyávát lelkesít, s könnyen olvadó / Nőszívet acéloz: jó honfitársak, / Nem ösztökél saját ügyünk eléggé, / Hogy orvosoljuk a bajt?)
But, alas, Caesar must bleed for it. (Sajnos / Vérért kell ontanunk.)	Enjoy the honey-heavy dew of slumber. (Élvezd könnyű álmod mézharmatát...)

4. táblázat – Példák Brutus legkevésbé és leginkább innovatív mondataiból (*Julius Caesar*)

Magas hasonlóság, alacsony innováció	Alacsony hasonlóság, magas innováció
Yet I fear him,/ For in the engrafted love he bears to Caesar – (Tartok tőle, Szereti Caesart, nagyon kötődik hozzá, és –)	The clock hath stricken three. (Hármat ütött.)
Well, Brutus, thou art noble. (Látom, Brutus, hogy nemes lelkű vagy...)	The morning comes upon 's. (Mindjárt itt a reggel.)
I blame you not for praising Caesar so. (Nem kárhoztatlak, hogy Caesart dicséred...)	And I do know by this they stay for me/ In Pompey's Porch. (És azt is tudom, hogy jelenleg mind / Pompeius csarnokában várnak rám.)
Caesar doth bear me hard, but he loves Brutus. (Caesar utál, de Brutust kedveli.)	When went there by an age, since the great flood,/ But it was famed with more than with one man? (Volt-e korszak az özönvíz után,/ Amelyben csak egyetlen név számított?)
I know that virtue to be in you, Brutus,/ As well as I do know your outward favor. (Jól ismerem benned ezt az erényt,/ Úgy, ahogyan külsődet ismerem.)	No, it is Casca, one incorporate/ To our attempts. (Nem, ő Casca. Ő is csatlakozott / Az ügryhöz.)

5. táblázat – Példák Cassius legkevésbé és leginkább innovatív mondataiból (*Julius Caesar*)

Végezetül érdemes még az *Othellót* megemlítenünk, amelyben Jagóhoz társul a legmagasabb innovációs pontszám. Ami nem meglepő, hiszen intrikusként a cselekmény előrehaladásával egyre sikeresebben tematizálja a diskurzust, sőt gyakran egész mondatokat ad más szereplők, főként Othello szájába (például: „Men should be what they seem” [„Az ember vagy legyen olyan, amilyennek látszik, vagy ne legyen olyan, amilyennek nem látszik.”³⁵ – Jago]; „Certain, men should be what they seem.” [„Így van. Az ember olyan legyen, amilyennek látszik!” – Othello]; „Or to be naked with her friend in bed An hour or more, not meaning any harm?” [„Vagy meztelenül feküdtek az ágyban, sokáig, s ez nem jelent semmi rosszat?” – Jago]; „Naked in bed, Jago, and not mean harm?” [„Ágyban, meztelenül, s nem tenni semmi rosszat?” – Othello]). Othellónak Jago korábbi megnyilatkozásaitól leginkább eltérő mondatai a dráma végéről származnak. Ezekben egy absztrakt, metaforikus nyelven írja le helyzetét; ami egyfelől jelezheti, hogy a cselekmény végére képes lesz kívülről,

³⁵ A magyar szöveget Márton László fordításában adjuk meg: William Shakespeare, „Othello, Velence négere”, ford. Márton László, *Színház* 41, 12. sz. (2009): melléklet, 3.

tágabb perspektívából az eseményekre tekinteni (Jago mesterkedése, hogy mindig a konkrét jelekre irányítsa a figyelmét.) Másrészt jelezheti azt is, hogy valójában nem képes új információt a diskurzusba bevezetni, és éppen akkor válik Jagóhoz képest innovatívvá, amikor nem nevezi meg a dolgokat, hiszen azt Jago teszi helyette. Ezt mutatja Othellónak Desdemonához címzett és Jago szólamától egyik leginkább eltérő mondata is: „Let me not name it to you, you chaste stars.” („Nem mondom el nektek, szűz csillagok!”)

5. Összegzés

A szereplői megszólalások mondatszintű beágyazásainak összehasonlítása konkrét drámák értelmezéséhez és általános mintázatok azonosításához is hasznos lehet. A dolgozatban javasolt eljárás értelmében azok a szereplők tekinthetők innovatívnak, akik mondatai szemantikailag a leginkább eltérnek a többi karakter korábbi mondataitól. Az eltérés mértékét ebben az esetben a vizsgált mondat és a hozzá leghasonlóbb viszonyában adjuk meg (Maximum Cosine Similarity), és nem a korábbiaktól való átlagos hasonlóság szerint. A páros összehasonlítások eredményeként létrejött hálózatok a szereplők közötti viszonyokat egyidejűleg teszik átláthatóvá a szöveg olvasásának vagy az előadás nézésének szükségszerűen időbeli folyamatához képest. Ezek a hálózatok továbbá új módon teszik leírhatóvá Shakespeare tragédiái/nem komédiái és komédiái közötti különbséget is. Míg a tragédiákban, amelyek az innovációs pontszámok elosztása szempontjából hierarchikusabbak, a férfi főszereplők szólamát idézik a leggyakrabban, addig a körkörösőbb komédiákban inkább a nők tematizálják a dráma diskurzusát.

Az összehasonlítások mondatpárjait néhány színműben közelebbről is szemügyre véve kétféle karaktertípus tűnik megkülönböztethetőnek Shakespeare drámáiban, amelyek egyaránt innovatívnak tekinthetők. Egyrészt vannak szereplők, akik valóban új információkat, híreket vezetnek be a diskurzusba. Ilyen például Horatio a *Hamlet*-ben, aki különböző események szemtanújaként összekötő, hírvivő funkciót is ellát; Cassio a *Julius Caesar*-ban, az összeesküvés fő szervezője; de említhető Zuboly is a *Szentivánéji álomból*, aki egy mellékszálát köt össze a fő cselekménnyel. Más karakterek esetében nem hagyományos értelemben vett hírekről van szó, azaz nem pusztán arról, hogy másról, hanem *máshogy* beszélnek, mint a többiek. Ez lehet a korábbiakban való elbizonytalanodás eredménye (például Hamlet az azonosság kérdésében, Hermia a külvilág érzékelésben és értelmezésben), az érzelmek előtrébe helyezése (Brutus) vagy szójátékok és virágnyelv használata (Mercutio). Ennek kapcsán általános mintázatnak tűnik az absztrakt és konkrét mondatok különbsége: minél poétikusabb és absztraktabb egy megszólalás, annál újszerűbbnek hat.

Repetition and Innovation in Drama

An attempt to calculate the novelty of actors' utterances

In our study, we present a method to measure the information value of characters' utterances in a dramatic work. All characters who frequently elaborate meanings different from previous ones are considered innovative; while

a character who tends to repeat (and thus reinforce) a previously established discourse performs a maintaining or repeating function. The comparison is based on embedding scores assigned to the character's sentences using an SBERT-based algorithm. The duality of innovation and repetition is also present within a single character – which the method can also investigate due to its sentence-level analysis. Furthermore, the pairwise comparison of characters allows us to represent the relations between them in a form of a directed graph, which we hope will contribute to the discussion of the network representation of plays. Using this procedure, we identify general differences between Shakespeare's tragedies and comedies – for example in terms of the gender roles within them. In addition to general patterns, the method provides a basis for new interpretations of individual plays and characters, which may highlight the joint applicability of close reading and statistical methods.

Keywords:

innovation, repetition, quantitative drama analysis, Shakespeare, SBERT

6. Függelék

1.

A kutatás során használt szövegek, kódok és eredmények elérhetők az alábbi repozitóriumban, Hozzáférés 2024.07.11.: <https://github.com/SzemesBotond/info-drama>

2. Hasonló és eltérő jelentésű mondatok a *Hamlet*ből az SBERT modellek összehasonlításához (a táblázat a kiválasztott modell alapján megadott koszinusz hasonlóságokat mutatja)

1. How now, what noise is that?
2. Alack, what noise is this?
3. Exchange forgiveness with me, noble Hamlet.
4. O Hamlet, speak no more!
5. To die, to sleep—No more—and by a sleep to say we end
The heartache and the thousand natural shocks
That flesh is heir to—'tis a consummation
Devoutly to be wished.
6. This gentle and unforced accord of Hamlet
Sits smiling to my heart, in grace
whereof No jocund health that Denmark drinks today
But the great cannon to the clouds shall tell,
And the King's rouse the heaven shall bruit again,
Respeaking earthly thunder.
7. To be or not to be, that is the question: Whether 'tis nobler in the mind to suffer
The slings and arrows of outrageous fortune,
Or to take arms against a sea of troubles
And, by opposing, end them.
8. Though yet of Hamlet our dear brother's death
The memory be green, and that it
us befitted To bear our hearts in grief, and our whole kingdom
To be contracted in one brow of woe,
Yet so far hath discretion fought with nature
That we with wisest sorrow think on him
Together with remembrance of ourselves.

9. Ay, truly, for the power of beauty will sooner transform honesty from what it is to a bawd than the force of honesty can translate beauty into his likeness.
10. Could beauty, my lord, have better commerce than with honesty?
11. Rest, rest, perturbed spirit!
12. Their residence, both in reputation and profit, was better both ways.

	1											
2	0,85	2										
3	0,04	0,04	3									
4	0,11	0,09	0,59	4								
5	0,05	0,09	0,36	0,34	5							
6	0,12	0,13	0,52	0,47	0,54	6						
7	-0,04	-0,01	0,39	0,33	0,40	0,32	7					
8	-0,03	-0,04	0,53	0,53	0,53	0,55	0,39	8				
9	-0,05	-0,07	0,26	0,19	0,30	0,31	0,22	0,25	9			
10	-0,06	-0,09	0,26	0,14	0,19	0,28	0,21	0,18	0,72	10		
11	0,10	0,09	0,23	0,18	0,42	0,36	0,19	0,27	0,20	0,14	11	
12	0,04	-0,03	0,16	0,01	-0,02	0,09	0,10	0,05	0,07	0,24	-0,03	

3. Hasonló és eltérő jelentésű mondatok a *Hamlet*ből – példamondatok a királyi beszédből, a sírásók dialógusából és az első jelenet megszólalásaiból (a táblázat a kiválasztott modell alapján megadott koszinusz hasonlóságokat mutatja)

1. He shall with speed to England For the demand of our neglected tribute.
2. It was that very day that young Hamlet was born – he that is mad, and sent into England.
3. Th' ambassadors from Norway, my good lord, Are joyfully returned.
4. Therefore our sometime sister, now our queen, Th' imperial jointress to this warlike state, Have we (as 'twere with a defeated joy, With an auspicious and a dropping eye, With mirth in funeral and with dirge in marriage, In equal scale weighing delight and dole) Taken to wife .
5. I think it be no other but e'en so.
6. Is not this something more than fantasy?
7. It harrows me with fear and wonder.
8. I like thy wit well, in good faith.
9. Cudgel thy brains no more about it, for your dull ass will not mend his pace with beating.

	1							
2	0,34	2						
3	0,27	0,22	3					
4	0,35	0,28	0,31	4				
5	0,10	0,12	0,15	0,19	5			
6	0,05	0,12	0,03	0,19	0,16	6.		
7	0,19	0,23	0,09	0,29	0,19	0,17	7	
8	0,06	0,17	0,23	0,21	0,14	0,09	0,18	8
9	0,26	0,23	0,08	0,20	0,10	0,10	0,23	0,20