

Kulturális evolúció: Nyelv és irodalom

A lapszám a kulturális evolúció elméleti és módszertani alapjait vázolja a nyelv és az irodalom témakörében. Míg a „kulturális evolúció” kifejezés leegyszerűsítő, teleologikus megközelítést sugallhat, a lapszám bemutatja, hogy ez az egyre népszerűbb keretrendszer távol áll az ideológiailag vezérelt elgondolásoktól: a tanulmányok a kognitív és társadalmi tényezők által formált kulturális változás fokozatos, gyakran véletlenszerű és decentralizált folyamata mellett érvelnek. Fő kérdésük, hogy az irodalmi és nyelvi formák – például a versformák, nyelvtani mintázatok, narratív konvenciók – hogyan jönnek létre az ismétlődő kulturális átadás révén. Központi kérdés továbbá, hogy miért válnak egyes formák és viselkedésmódok sikeressé, vagyis miért másolják és továbbítják őket nagy idő és térbeli távolságban az emberek, miközben mások hamar feledésbe merülnek.

A makroszintű történelmi tendenciákat és a mikroszintű interperszonális interakciókat egyaránt hangsúlyozva a lapszám először a nyelvi működés általános kérdéseit tárgyaló szöveget közöl, majd egy tanulmányt, amely az írás eredetének új elméletét vázolja. Ezt követően irodalmi szövegek kerülnek az előtérbe líra-, próza- és eszmetörténeti folyamatok azonosíthatóságának kérdéseivel.

Kiemelendő továbbá a bemutatott terület és a digitális bölcsészet diszciplínáinak közeledése az elmúlt években, amely közeledés természetesnek nevezhető, amennyiben mindkét megközelítési mód hosszan tartó kultúrtörténeti folyamatok azonosítását (is) célozza statisztikai modellek segítségével. Ezt az összefüggést a digitális bölcsészet felől különösen fontos hangsúlyozni, amely gyakorlatorientált területként sokszor nélkülözi az elméleti megalapozottságot. A lapszám létrehozása mögött meghúzódó egyik motiváció, hogy egy lehetséges elméleti keretrendszer körvonalai is kirajzolódjanak – még ha vázlatosan is – a hazai digitális bölcsészet számára.

Jelen lapszámunkat SZEMES BOTOND szerkesztette.

Cultural Evolution: Language and Literature

This issue presents the theoretical and methodological foundations of cultural evolution as applied to language and literature. Although the term *cultural evolution* may evoke a reductive or teleological approach, the contributions in this volume show that the increasingly popular framework is far removed from ideologically driven conceptions. The studies argue instead for a gradual, often random, and decentralized process of cultural change shaped by cognitive and social factors. Their central question is how literary and linguistic forms – such as poetic structures, grammatical patterns, and narrative conventions – emerge through repeated cycles of cultural transmission. Another key concern is why certain forms and behaviors prove successful – why people continue to copy and transmit them across long spans of time and space – while others are quickly forgotten.

In doing so, the issue considers both macro-level historical trends and micro-level interpersonal interactions. It begins with a text that addresses general questions about how language functions and how cultural transmission gives rise to the statistical characteristics of language. This is followed by a study that proposes a new theory of the origins of writing. Later contributions shift the focus to literary texts; they examine how large-scale trends within poetry, prose, and the history of ideas can be identified and modeled.

The issue also highlights the growing convergence between the field and the Digital Humanities in recent years – a natural development given that both approaches aim to identify long-term cultural and historical patterns using statistical models. From the perspective of DH, this connection is especially important to emphasize, because the field's practical focus often means it lacks a strong theoretical foundation. One of the motivations behind the issue was therefore to sketch – however provisionally – the contours of a potential theoretical framework for the DH community in Hungary.

This issue edited by BOTOND SZEMES.

Kulturelle Evolution: Sprache und Literatur

Dieses Heft stellt die theoretischen und methodischen Grundlagen der kulturellen Evolution in ihrer Anwendung auf Sprache und Literatur vor. Obwohl der Begriff *kulturelle Evolution* einen reduktiven oder teleologischen Ansatz nahelegen könnte, zeigen die Beiträge dieses Bandes, dass der zunehmend populäre Rahmen weit von ideologisch geprägten Vorstellungen entfernt ist. Stattdessen plädieren die Studien für einen allmählichen, oft zufälligen und dezentralisierten Prozess kulturellen Wandels, der von kognitiven und sozialen Faktoren geprägt wird. Im Mittelpunkt steht die Frage, wie literarische und sprachliche Formen – etwa poetische Strukturen, grammatische Muster oder narrative Konventionen – durch wiederholte Zyklen kultureller Weitergabe entstehen. Ein weiteres zentrales Anliegen besteht darin, warum bestimmte Formen und Verhaltensweisen erfolgreich werden – warum Menschen sie über große zeitliche und räumliche Distanzen hinweg weitergeben – während andere schnell in Vergessenheit geraten.

Dabei berücksichtigt das Heft sowohl makrohistorische Entwicklungen als auch mikrointeraktionale Prozesse. Es beginnt mit einem Beitrag, der allgemeine Fragen zur Funktionsweise der Sprache behandelt und untersucht, wie kulturelle Übertragung zu den statistischen Eigenschaften von Sprache führt. Darauf folgt eine Studie, die eine neue Theorie zur Entstehung der Schrift vorschlägt. In späteren Beiträgen richtet sich der Fokus auf literarische Texte; sie untersuchen, wie großräumige Entwicklungen in Lyrik, Prosa und Ideengeschichte identifiziert und modelliert werden können.

Das Heft hebt zudem die wachsende Annäherung zwischen diesem Forschungsfeld und den Digital Humanities in den letzten Jahren hervor – eine natürliche Entwicklung, da beide Ansätze darauf abzielen, langfristige kultur- und geistesgeschichtliche Muster mithilfe statistischer Modelle zu identifizieren. Aus der Perspektive der Digital Humanities ist es besonders wichtig, diese Verbindung zu betonen, da der praxisorientierte Charakter des Feldes häufig zulasten einer soliden theoretischen Fundierung geht. Eine der Motivationen hinter diesem Heft bestand daher darin, zumindest skizzenhaft die Konturen eines möglichen theoretischen Bezugsrahmens für die ungarische DH-Community zu entwerfen.

Diese Ausgabe wurde von BOTOND SZEMES redigiert.

TANULMÁNYOK

SZEMES BOTOND

Kultúra mint információátadás

Előszó a *Helikon Kulturális evolúció: Nyelv és irodalom* című
lapszámához

ORCID: 0000-0002-0637-6776
botond.szemes@ut.eee

—HELIKON—

Culture as information transmission
Foreword to the special issue *Cultural Evolution: Language and Literature*

Abstract

This introductory essay to the *Helikon* special issue on *Cultural Evolution: Language and Literature* outlines the theoretical and methodological foundations of cultural evolution as applied to language and literature. While the term “cultural evolution” may evoke outdated, teleological frameworks, the authors argue for a more nuanced understanding of long-term processes in cultural history: a gradual, often random and decentralized process of information transmission shaped by cognitive and social factors. The issue highlights how literary and linguistic forms – such as poetic meters, syntactic patterns, or style – evolve through repeated imitation within a population. Emphasizing both macro-level historical trends and micro-level interpersonal interactions, the volume brings together research on the origins of writing, the spread of literary forms, and the statistical modeling of linguistic change. It also explores the growing intersection between cultural evolution studies and digital humanities (DH), suggesting that DH could benefit from the theoretical foundations offered by cultural evolution.

Keywords: Cultural evolution, cultural information, cultural transmission, computational literary studies

A bölcsészettudomány belátásai felől ijesztőnek tűnhet a szókapcsolat: kulturális evolúció. Hiszen a kultúra evolucionista, fejlődés- és célelvű elképzelésének az érvénye a 19. század óta – amikor többnyire még e paradigma határozta meg a történeti haladás képzetét – kétségessé vált, amely kételyt tovább erősített a fajelmélet és az eugenetika térnyerése a 20. század elején.¹ A kulturális evolúció elméleti keretrendszere azonban távol áll az ilyen leegyszerűsítő, ideológiailag vezérelt elgondolásoktól: éppen, hogy lassú, véletlenszerű és diffúz folyamatként írja le a történeti változásokat. Célja időben kiterjedt kultúrtörténeti mintázatok és a mögöttük meghúzódó okok azonosítása.² Ezek az okok – nem kényszerítő erejű törvényszerűségek, inkább hatásösszefüggések – elsősorban a folyamatban résztvevők kognitív meghatározottságára, illetve a gyakorlatok és formák átadásának sajátosságaira tűnnek visszavezethetőnek ebben a keretrendszerben. Módszertanát tekintve kvantitatív fókuszú; legtöbbször statisztikai modellek és számszerűsíthető kísérletek igazolják/cáfolják a vélt összefüggéseket.

A kutatók gyakran felhívják a figyelmet arra, hogy a darwini evolúcióelmélet maga sem feleltethető meg a később elterjedt, vulgáris evolucionizmus-sal.³ Sőt arra is rámutattak, hogy az elmélet nem feltételezi gének jelenlétét sem,⁴ csupán három egyszerű összetevőre épít: „(1) egy populáción belüli változatosságra, (2) egy mechanizmusra, amely bizonyos variánsokat előnyben részesít másokkal szemben, és (3) a sikeres variánsok továbbításának módjára.”⁵ Mindez igaz a *kulturális információ* átörökítésére is, amely minden társas úton – azaz például utánzás, tanítás vagy nyelvi szimbólumok révén – átadott tartalomra vonatkozik (kulturális átadás, *cultural transmission*), „beleértve a hiedelmeket, ismereteket, attitűdöket, készségeket, normákat, nyelveket, in-

¹ Vö. Alex MESOUDI, „Cultural Evolution”, in *Open Encyclopedia of Cognitive Science* (Cambridge, Mass: MIT Press, 2024), <https://doi.org/10.21428/e2759450.304de239>.

² Míg a területen kezdetben az apró kapcsolatokból álló „mikroevolúció”, addig mostanra a „makroevolúció” vizsgálata került az előterébe: Alex MESOUDI, „Pursuing Darwin’s curious parallel: Prospects for a science of cultural evolution”, *Proceedings of the National Academy of Sciences* 114, 30. sz. (2017): 7853–7860, <https://doi.org/10.1073/pnas.1620741114>.

³ Lásd például Olivier Morin tanulmányát jelen lapszámban.

⁴ Peter J. RICHESON és Robert BOYD, *Not By Genes Alone: How Culture Transformed Human Evolution* (Chicago, IL: University of Chicago Press, 2010); Alex MESOUDI, *Cultural Evolution: How Darwinian Theory Can Explain Human Culture* (Chicago: University of Chicago Press, 2011), <https://doi.org/10.7208/chicago/9780226520452.001.0001>.

⁵ Oleg SOBCHUK, „Evolution of Modern Literature and Film”, in *The Oxford Handbook of Cultural Evolution*, szerk. Jamshid J. TEHRANI, Jeremy KENDAL és Rachel KENDAL (Oxford: Oxford University Press, 2023), <https://doi.org/10.1093/oxfordhb/9780198869252.013.45>.

tézményi szabályokat is”.⁶ Ezek ugyanis változatosak, amennyiben szüntelenül új tartalmakat hozunk létre, valamint vannak jobban elterjedt, sikeres változatok, amelyek generációkon keresztül képesek továbböröklődni. Ennyiben az evolúciós modell a kultúrára is alkalmazható. Ezáltal pedig olyan kérdések válnak vizsgálhatóvá, mint hogy miért lesz egy forma sikeresebb egy másiknál, miért ismétlünk meg korábbi mintázatokat, illetve hogyan terjed el és/vagy módosul egy gyakorlat az ismétlések révén?⁷

Eszerint tehát a kultúra nem más, mint társas kapcsolatok során átadott információk és ezek időbeli alakulása. Ennek megfelelően központi helyet foglal el a kulturális evolúció területén a gyakorlatok és formák, egyáltalán azok átadására való képesség *eredetének* a vizsgálata (lásd a lapszámban közölt tanulmányt az írás eredetéről). Ám ugyanennyire fontos az átadási láncok nyomon követése is. Például: milyen hatások mentén megy végbe nyelvek elkülönülése és eltérő fejlődése?⁸ Vagy a lapszám két költészeti példáját említve: hogyan terjed el a petrarcái versmérték Európában? Összefügg-e az időmértékes verselés visszaszorulásával a szintaktikai ismétlések térnyerése a magyar költészetben? Ekkor a jelenségek egyszerre makro- és mikroszintje is az elemzés tárgyává válik, hiszen nagy kiterjedésű, makroszkopikus folyamatok feltárása vonatkozik, ám ezeket csak az egyének közötti, apró kapcsolatok alapján tartja leírhatónak.

Ebből is látható, hogy a kulturális evolúció elméletének egy lényeges felismerése, hogy létezik az emberi kapcsolatoknak egy olyan, mindennapi szintje, amelyet a hagyományos történeti modelljeink és elbeszéléseink segítségével nem igazán tudunk megragadni. Létezik egy apró interakciókból álló, felülről vagy külső hatások által nem irányított hálózat, amely mégis képes markáns tendenciákat létrehozni. Gondoljunk például a gyereknevekben megfigyelhető divatokra. Persze állhatnak hírességek vagy akár hurrikánok nevei a(z amerikai) trendek háttérben, ahogyan a már népszerű nevekhez való hasonlóság is meghatározhatja egy forma felfutását (Anna–Hanna),⁹ ezek mégsem magya-

⁶ MESOUDI, „Cultural Evolution”.

⁷ Ez utóbbi kapcsán Ted Underwood és társai amellet érvelnek, hogy a generációk egymásra következése felelős az irodalmi változások többségéért: Ted UNDERWOOD, Kevin KILEY, Wenyi SHANG és Stephen VAISEY, „Cohort Succession Explains Most Change in Literary Culture”, *Sociological Science* 9 (2022): 184–205, <https://doi.org/10.15195/v9.a8>.

⁸ Balthasar BICKEL, Anne-Lise GIRAUD, Klaus ZUBERBÜHLER és Carel P. VAN SCHAIK, „Language Follows a Distinct Mode of Extra-Genomic Evolution”, *Physics of Life Reviews* 21, 50. sz. (2024): 211–225, <https://doi.org/10.1016/j.plrev.2024.08.003>.

⁹ Jonah BERGER, Eric T. BRADLOW, Alex BRAUNSTEIN és Yao ZHANG, „From Karen to Katie: Using Baby Names to Understand Cultural Evolution”, *Psychological Science* 23, 10. sz. (2012): 1067–1073, <https://doi.org/10.1177/0956797612443371>. Más kutatások ugyanakkor azt

rázzák általánosan, hogy miként lesz az egyik változat hirtelen (vagy fokozatosan) népszerű, vagy miért tűnik el egy másik. Egy szülő gondolhatja, hogy döntését nem befolyásolták mások, amikor gyermekének egyedi nevet kívánt adni – ám ekkor is kiderülhet, hogy egy általa még tudatosan nem érzékelt tendenciát követett. Nagy elbeszéléseink legtöbbször nem veszik figyelembe az átadások és személyközi kapcsolatok olyan kis lépésekből álló mechanikáját, amely például a névadásokban létrejött mintákat is magyarázhatják. A bölcsészet területén talán a nyelvtörténet tudatosította a leginkább ezt a jelenséget, hiszen szavak és szimbólumok jelentésének változásai is a mindennapi interakciókból adódnak össze.

Ugyanez igaz a művészeti formákra is: egyesek nagy tér- és időbeli távolságot felölelően képesek elterjedni (például a szonettforma az európai költészetben), míg másokat nem adják tovább ehhez szükséges gyakorisággal az emberek. Olivier Morin *How Traditions Live and Die* című könyvében például azt mutatja be, hogy bizonyos gyerekdalok és mondókák akár évszázadokon át is fennmaradhatnak külső (azaz felnőtt) beavatkozás nélkül, pusztán azért, mert kellő rendszerességgel megismétlik ezeket az amúgy gyakran változó közösségek.¹⁰ Oleg Sobchuk ehhez hasonlóan a narratív formák sikeres technikáit kutatja, amelyek különféle médiumokban és korszakokban is megjelennek.¹¹ Ekkor inkább arra összpontosít, hogy mi állhat egy forma elterjedésének hátterében (miért rövidül a hollywoodi filmekben a beállítások hossza? miért egyre gyakoribbak az időbeli előre- és visszaulások bennük?),¹² de érdemes azt is tanulmányozni, hogy miként (milyen ütemben, milyen közösségekben) történik az átadás. Az ilyen, legtöbbször statisztikai modelleken alapuló leírások figyelmesek az esetlegességekre és variációkra is, nem egy homogén, egyenes vonalú történet elbeszélését adják. Jó példa erre Nicholas D. Paige *Technologies of the Novel* című könyve, amely a francia és angol regény evolúcióját vázolja a 18. században kvantitatív módon, és amely mellett érvel, hogy ez az evolúció nem mesélhető el egy külső nézőpontból egyetlen ív mentén (például a kapitalizmus, a modernitás vagy a könyvnyomtatás történetébe

mutatják, hogy nem a hírességek befolyásolják a trendeket, sokkal inkább maguk is részesei azoknak: Stanley LIEBERSON, *A Matter of Taste: How Names, Fashions, and Culture Change* (New Haven London: Yale University Press, 2000).

¹⁰ Olivier MORRIN, *How Traditions Live and Die* (Oxford: Oxford University Press, 2015), 176–202.

¹¹ SOBCHUK, „Evolution of Modern Literature and Film”.

¹² Uo.

ágyazva), hanem csakis az egyes művek egymásra hatásának összetett, hálózatos folyamataként.¹³

Miért válik egy tárgy, gyakorlat vagy forma sikeresebbé másoknál, azaz miért másolják és utánozzák őket nagy idő- és térbeli távolságokban? A kulturális evolúció elmélete legtöbbször olyan hatásokat (vagy torzításokat, elfogultságokat, *bias*) azonosít ennek megválaszolásakor, amelyek csökkentik az egyének kognitív erőfeszítéseit választásaik során.¹⁴ Ilyen az alkalmazkodási torzítás (*conformity bias*), amely szerint hajlamosak vagyunk a többségi vélekedéssel vagy viselkedéssel tartani, hiszen ezek kipróbált és működőképes választásoknak tűnnek. Szintén gyakran emlegetett szempont a presztízsérték (*prestige bias*): eszerint érdemes a sikeres társakat utánoznunk, bízva abban, hogy ez minket is sikeressé tehet.¹⁵ Az ilyen kontextus-alapú szelekció (*context-based selection*) mellett szerepet játszhat a tartalmak kulturális vonzereje is (*content-based selection*), amely esetben valami azért öröklődik tovább, mert érdekes és figyelemfelkeltő a résztvevők számára („minden, amire érdemes, jobban emlékszik a DNS” – ahogyan a Sziámi zenekar is énekelte a 90-es években). Ezt okozhatja az érzelmi hatáskeltés sajátossága (a negatív tartalmak jobban terjednek¹⁶) vagy a tartalomnak a percepciók rendszerünkhöz való igazodása. A szörendi viszonyok például nagyban meghatározottnak tűnnek észlelési apparátusunk által.¹⁷ További példa lehet a nagyhatású, de sokat vitatott Berlin-Kay elmélet, amely értelmében a legkülönbözőbb kultúrákban is hasonló mintázatok figyelhetők meg a színnevek kapcsán: minden kultúra rendelkezik a sötét-világos (vagy fekete-fehér) ellentéppárral; ahol három színnév

¹³ Nicholas D. PAIGE, *Technologies of the Novel: Quantitative Data and the Evolution of Literary Systems* (Cambridge: Cambridge University Press, 2021).

¹⁴ Összefoglalását lásd Rachel L. KENDAL, „Social Learning Strategies: Bridge-Building between Fields”, *Trends in Cognitive Sciences* 22, 7. sz. (2018): 651–665, <https://doi.org/10.1016/j.tics.2018.04.003>.

¹⁵ Részletesebben lásd MESOUDI, „Pursuing Darwin’s curious parallel...”.

¹⁶ Kimmo Kimmo ERIKSSON és Julie C. COULTAS, „Corpses, Maggots, Poodles and Rats: Emotional Selection Operating in Three Phases of Cultural Transmission of Urban Legends”, *Journal of Cognition and Culture* 14, 1–2. sz. (2014): 1–26, <https://doi.org/10.1163/15685373-12342107>; Alberto ACERBI, „Cognitive Attraction and Online Misinformation”, *Palgrave Communications* 5, 1. sz. (2019): 15, <https://doi.org/10.1057/s41599-019-0224-y>; Joseph M. STUBBERSFIELD, Jamshid J. TEHRANI és Emma G. FLYNN, „Chicken Tumours and a Fishy Revenge: Evidence for Emotional Content Bias in the Cumulative Recall of Urban Legends”, *Journal of Cognition and Culture* 17, 1–2. sz. (2017): 12–26, <https://doi.org/10.1163/15685373-12342189>.

¹⁷ Arrate ISASI-ISASMENDI, „The Agent Preference in Visual Event Apprehension”, *Open Mind* 7 (2023): 240–282, https://doi.org/10.1162/opmi_a_00083.

létezik, a harmadik a piros; a negyedik a sárga vagy a zöld, és így tovább.¹⁸ Ekkor tehát a színnevek elterjedését és a kulturális átadás módjait antropológiai sajátosságok határozzák meg.

Többen rámutatnak azonban, hogy ezek a hatások csak korlátozott érvényre magyarázzák a kulturális információk terjedését. Olivier Morin szerint például fontosabb magának az átadási láncnak a szerkezete és pusztán az átadások száma, mintsem az átadás minősége, az átadott információ vonzereje vagy más kognitív tényezők. Az lesz az alapja a hosszantartó hagyományoknak, amit gyakran, és nem az, amit hűségesen ismételnek az egyének, mivel az ismétlések gyakorisága önszabályzó folyamatként korrigálja az átadások pontatlanságát.¹⁹ Ehhez hasonlóan Morin és Alexey Koshevoy egy különálló tanulmányban azt támasztják alá kvantitatív eszközökkel, hogy a Zipf-féle tömörségi törvény (*brevity law* vagy *law of abbreviation*: a gyakoribb szavak rövidebbek egy nyelvben), magyarázható egyszerűen azzal az általános tendenciával, hogy a szavak a használat során rövidülnek, és nem szükséges a „kommunikációs hatékonyság” kognitív elméletére hivatkoznunk (ez azt mondaná, hogy gazdaságosabb, kevesebb erőfeszítésbe kerül a rövid szavak használata, így érdemes a gyakoriakat rövid formába kódolni).²⁰ Tanulmányuk végén elővigyázatosságra intenek: nem biztos, hogy minden azért alakul ki, hogy hasznos és hatékony legyen – meglehet, hogy csupán az átadások mechanikája eredményez utólag ilyenképpen racionalizálható jelenségeket. Simon Kirby és munkatársai több kísérletben igazolták – ezek közül egy olvasható a lapszámban is –, hogy a nyelvi működés általános mintázatai szintén a gyakori másolásnak köszönhetők; az átadások láncolata önmagában létrehoz egy struktúrát – autopoetikus rendszert – az egyes nyelvelsajátítók szándékos tervezése nélkül.²¹ Az emögött meghúzódó legfőbb mozgatórugót a tanulha-

¹⁸ Mike DOWMAN, „Explaining Color Term Typology With an Evolutionary Model”, *Cognitive Science* 31, 1. sz. (2007): 99–132, <https://doi.org/10.1080/03640210709336986>.

¹⁹ Persze ekkor is szükséges a másoláshoz valamiféle motiváció. Részletesebben lásd a könyvről írt recenziót jelen lapszámban.

²⁰ Olivier MORIN és Alexey KOSHEVOY, „A Cultural Evolutionary Model for the Law of Abbreviation”, *Topics in Cognitive Science*, 2024. december 24., tops.12782, <https://doi.org/10.1111/tops.12782>. A szerzőpáros és Helena Milton nagymintás adatelemzés során bemutatják, hogy a tömörségi törvény az írásrendszerekben is létezik, azaz a gyakoribb írásjelek többnyire egyszerűbbek: Alexey KOSHEVOY, Helena MITON és Olivier MORIN, „Zipf’s Law of Abbreviation Holds for Individual Characters across a Broad Range of Writing Systems”, *Cognition* 238 (2023): 105527, <https://doi.org/10.1016/j.cognition.2023.105527>.

²¹ Ehhez lásd Simon Kirby és Iban Arnon tanulmányát a lapszámból, valamint Simon KIRBY, Hannah CORNISH és Kenny SMITH, „Cumulative Cultural Evolution in the Laboratory: An Experimental Approach to the Origins of Structure in Human Language”, *Proceedings*

tóság jelenti; egy strukturált rendszert ugyanis könnyebb elsajátítani és továbbadni, mint egy véletlenszerűt. Pontosabban kétirányú folyamatról van szó a szerzők szerint: az ismétlések során kialakuló struktúra könnyebben tanulható, amely így további ismétléseket motivál.

Szintén az átadást elősegítő tanulhatóság felől válik megkérdőjelezhetővé az önkényes nyelvi jel tételének kizárólagossága is. Az önkényesség mellett ugyanis megfigyelhető minden nyelvben bizonyos mértékű ikonicitás és szisztematicitás is. Az ikonikusság során forma és jelentés érzékszervi analógiák segítségével kapcsolódik egymáshoz. Közismert példák a hangutánzó szavak vagy a jelentéseket leképező grammatikai szerkezetek,²² de ilyen viszonyt mutatnak a geometriai formák és a szóképzés módja közötti összefüggések is.²³ Dingemanse és társai szerint az ikonikusság nemcsak a tanulhatóságot segíti elő (hiszen könnyebb emlékezni egy motivált kapcsolatra), hanem a hatékony kommunikációt is szolgálja (gondoljunk például a gesztusokra élőbeszéd közben, amelyek szintén értelmezhetők ikonokként).²⁴ A nyelvben érvényesülő szisztematikus kapcsolatok ezzel szemben a kategóriák felismerését támogatják. Ennek értelmében, habár egyetlen forma-jelentés kapcsolat a legtöbbször valóban önkényes, ám a különböző kategóriákba tartozó formák jobban hasonlítanak egymásra, mint más csoportok példányaira. Több nyelvben kimutatták a főnevek és igék alaki különbségét nagymintás kutatásokban; ahogyan szemléletes példa az a nyelveken átívelő szabályszerűség is, miszerint a közelre mutató névmásokat magas hangrendű szavak alkotják (*itt, ez; here, this* stb.), míg a távolra mutatók mély hangrendűek (*ott, az, there, that*).²⁵ Vagyis a nyelvi rendszer egymással versengő kívánalmaknak egyszerre kíván meg-

of the National Academy of Sciences 105, 31. sz. (2008): 10681–10686, <https://doi.org/10.1073/pnas.0707835105>.

²² Lásd: John HAIMAN, „The Iconicity of Grammar: Isomorphism and Motivation”, *Language* 56, 3. sz. (1980): 515, <https://doi.org/10.2307/414448>; Roman JAKOBSON, „Quest for the Essence of Language”, *Diogenes* 13, 51. sz. (1965): 21–37, <https://doi.org/10.1177/039219216501305103>. Jakobson egyik példája szemléletes: minden nyelvben a többes szám jelölt extra karakterrel az egyes számmal szemben, hiszen így morfológiailag is realizálódik a számbeli gyarapodás.

²³ „A legismertebb példák azokból a vizsgálatokból származnak, amelyek azt mutatják, hogy az emberek következetesen a lekerekített és a szögletes formákat olyan kitalált szavakhoz rendelik, mint a »baluma« és a »tukeetee«, illetve a »bouba« és a »kiki«.” Mark DINGEMANSE, Damián E. BLASI, Gary LUPYAN, Morten H. CHRISTIANSEN és Padraic MONAGHAN, „Arbitrariness, Iconicity, and Systematicity in Language”, *Trends in Cognitive Science* 19, 10. sz. (2015): 603–615, 610, <https://doi.org/10.1016/j.tics.2015.07.013>.

²⁴ Uo.

²⁵ Uo., 606–607.

felelni azért, hogy a forma-jelentés kapcsolatok különböző módjait dolgozza ki (ikonikus, önkényes, szisztematikus). A szisztematicitás a nyelvelsajátítás korai fázisában lehet kiemelten hasznos, mivel a különböző kategóriák különbségeinek felismerésében játszik szerepet, ám a szóalakok megszorodásával „megnő a zavar és a többértelműség lehetősége, ami szelekciós nyomást gyakorol az önkényesebb és jobban megkülönböztethető formák felé”.²⁶ A kapcsolattípusok aránya – ezúttal is – a kulturális átadás önszabályzó folyamatának eredménye.

Mindez jól mutatja a kulturális evolúció elméletének egy másik szemléletbeli sajátosságát is. Nevezetesen, hogy e megközelítésmód szerint a folyamatok gyakran több ellentétes elv közötti kompromisszumként jönnek létre, például az ikonicitás, a szisztematicitás és az önkényesség egyensúlyának kialakulásával. Hasonló működés figyelhető meg más szinteken is. A nyelvtörténet például egy olyan dinamika mentén alakul, amely egyaránt megengedi, hogy jelentősen eltérő variánsok (különböző nyelvek) jöjjenek létre a társas differenciálódás eredményeként, illetve hogy ezek az eltérések a faj egészére jellemző nyelvi képességek határain belül formálódjanak, azaz a különböző nyelvek univerzálisan elsajátíthatók maradjanak.²⁷ Röviden: egyaránt igény van az emberi közösségekben minél eltérőbb nyelvek kialakulására és az ezek közötti hasonlóságok fenntartására is. Egy másik példát említve: a nyelvi jelek és jelentések szintén két általános, de ellentétes nyomás alatt alakulnak. Az egyik a kifejezhetőség maximalizálása a kommunikáció érdekében (ez növeli az elemszámot és csökkenti a többértelműségeket, azaz arra törekszik, hogy minden jelenségre legyen külön szavunk), a másik a komplexitás minimalizálása a tanulhatóság értelmében (ez csökkenti az elemszámot és növeli a többértelműséget).²⁸ A két elvnek egyszerre kell érvényesülnie ahhoz, hogy jól használható és elsajátítható legyen egy jelrendszer; az első nyomás hatására maradnak a nyelvben erős különbségek, a második miatt tömöríthető. Továbbá e kettős hatás révén jöhetnek létre a nyelv legfontosabb strukturális tulajdonságai, a kombinatorikusság (véges számú diszkrét jelből gyakorlatilag végtelen sok kifejezés hozható létre – az átadási láncokat vizsgáló kutatások azt mutatják, hogy egy sorozat végére még az olyan folyamatos jelsorok, mint a

²⁶ Uo., 609.

²⁷ BICKEL, „Language Follows a Distinct Mode of Extra-Genomic Evolution”.

²⁸ SIMON KIRBY, MONICA TAMARIZ, HANNAH CORNISH és KENNY SMITH, „Compression and Communication in the Cultural Evolution of Linguistic Structure”, *Cognition* 45, 141. sz. (2015): 87–102, <https://doi.org/10.1016/j.cognition.2015.03.016>.

fütyty, is diszkrét, ismétlődő egységekre tagolódnak), valamint a kompozicionalitás (a különálló jelentések nagyobb értelmi egységekbe szervezhetők).²⁹

Az eddigi áttekintés még korántsem fedi le a kulturális evolúció gyorsan változó és sokszínű kutatási irányának egészét – és erre a lapszám sem vállalkozhat. Sok minden módosult ezen a területen az 1980-as évek alapítószövegei óta. Ezek közé tartozik mindenekelőtt Luigi L. Cavalli-Sforza és Marcus W. Feldman 1981-es *Cultural Transmission and Evolution: A Quantitative Approach*, valamint Peter J. Richerson és Robert Boyd *Culture and the Evolutionary Process* és *Not By Genes Alone: How Culture Transformed Human Evolution* című könyve. A lapszámban ezekről érintőlegesen lesz szó (a gyakori hivatkozások mellett Richerson és Boyd könyveiről recenziót közlünk),³⁰ és inkább az újabb fejleményekre összpontosítunk. Ennek egyike a terület közeledése a digitális bölcsészethez, amely közeledés természetesnek nevezhető, amennyiben mindkét megközelítési mód hosszan tartó kultúrtörténeti folyamatok azonosítását (is) célozza statisztikai modellek segítségével. Ezt az összefüggést a digitális bölcsészet felől különösen fontos hangsúlyozni, amely gyakorlatorientált diszciplínaként sokszornélkülözi az elméleti megalapozottságot. A lapszám létrehozása mögött meghúzódó egyik motiváció, hogy egy lehetséges elméleti keretrendszer körvonalai is kirajzolódjanak – még ha vázlatosan is – a hazai digitális bölcsészet számára. Ezért tartalmaz olyan írásokat is, amelyek nem kifejezetten a kulturális evolúció elméletébe ágyazva jöttek létre, ám a kulturális információk terjedését követik nyomon, illetve nagy időbeli folyamatok azonosítását és magyarázatát kívánják adni.

Egyébként is vékony a határ a két diszciplína között az utóbbi években. Több olyan kutatás említhető, amelyek mindkettő alá besorolható, vagy akár maga is hivatkozik mindkét keretre. A digitális irodalomtudomány követői számára például ismerősek lehetnek az olyan szövegbányászaton alapuló kutatások, mint például a már többször idézett Morin és Alberto Acerbi közös, a kulturális evolúció elméletét érvényesítő tanulmánya. Ebben azt mutatják be, hogy az angolszász irodalom elmúlt két évszázadában folyamatosan csökken az érzelmeket kifejező szavak aránya, amely elsősorban a pozitív szavak visszaszorulására vonatkozik (a negatív kifejezések aránya változatlan). A tanulmány számos lehetséges okot felsorakoztat a jelenség mögött, ám a statisztikai

²⁹ Monica TAMARIZ és Simon KIRBY, „The Cultural Evolution of Language”, *Current Opinion in Psychology* 8 (2016): 37–43, <https://doi.org/10.1016/j.copsyc.2015.09.003>.

³⁰ Meglehet, ezáltal mi is hozzájárulunk a terület „felhígulásához”. Vö. Richard McELREATH, „The Problem With Cultural Evolution”, *Elements of Evolutionary Anthropology*, hozzáférés: 2025.05.02, <https://eleventh.org/blog/2022/09/19/the-problem-with-cultural-evolution/>.

tesztek szerint egyik sem magyarázza kielégítően a tendenciát.³¹ Ugyanígy említhetők – a teljesség igénye nélkül – más példák is. Ilyen Kun Sun és Run Wang elemzése, amely egy másik, az érzelmesség visszaszorulásával párhuzamos folyamatot detektál: az angolszász irodalom egyre konkrétabb és jobban elképzelhetővé vált az évszázadok során. A szerzők szerint ez többek között gazdasági tényezőkkel magyarázható, amennyiben az időközben kiépült globális irodalmi piac egyre szélesebb, nem anyanyelvi közönséget foglal magában, akik számára könnyebben olvashatók a konkrét és elképzelhető szövegek.³² Egy harmadik kapcsolódó szempont a szavak entrópiájának, azaz annak vizsgálata, hogy mennyire váratlan egy kifejezés szereplése egy szövegben. Charlie Pilgrim és munkatársai ennek mérésére fejlesztettek ki statisztikai módszert, és úgy találták, hogy az angol nyelv entrópiája (vagy másként: információtartalma) fokozatosan nő 1900 óta.³³ Ezek a szövegbányászati eszközökkel feltárt irodalom- és nyelvtörténeti folyamatok egyúttal a kulturális átadás működését is vizsgálhatóvá teszik különböző, egymást kiegészítő szempontokból – ám inkább a mintázatok feltárása és kultúrtörténeti kontextusuk áll e munkák középpontjában, semmint a kulturális információ terjedésének a mikéntje és a résztvevők kognitív meghatározottságai.

Az eddigiekből és a lapszám címéből is látszik, hogy amellet, hogy az „új-donságokat” részesítjük előnyben, egy másik megszorítást is érvényesítünk: a kulturális evolúció tág területéről csak a nyelvre és az irodalomra vonatkozó tanulmányokat közlünk. E tárgyban viszont sokrétű összeállítást hoztunk létre. Először Simon Kirby és munkatársai szövege olvasható a nyelvi működés

³¹ Olivier MORIN és Alberto ACERBI, „Birth of the Cool: A Two-Centuries Decline in Emotional Expression in Anglophone Fiction”, *Cognition and Emotion* 31, 8. sz. (2017): 1663–1675, <https://doi.org/10.1080/02699931.2016.1260528>. Hasonló tendencia figyelhető meg a könnyűzenében (a pozitív érzelmek csökkenése mellett a negatív érzelmek erősödése): Charlotte O. BRAND, Alberto ACERBI és Alex MESOUDI, „Cultural Evolution of Emotional Expression in 50 Years of Song Lyrics”, *Evolutionary Human Sciences* 1 (2019): e11, <https://doi.org/10.1017/ehs.2019.11>. A mintázatok újfajta értelmezését kínálja: Xiaolu WANG, „Tug-of-War of Emotion: Measuring and Modeling Sentiment Cycles in Chinese-Language Pop Song Lyrics, 1967–2023”, *Journal of Cultural Analytics* 10, 2. sz. (2025), <https://doi.org/10.22148/001c.144809>.

³² Kun SUN és Rong WANG, „The Evolutionary Pattern of Language in English Fiction Over the Last Two Centuries: Insights From Linguistic Concreteness and Imageability”, *Sage Open* 12, 1. sz. (2022): 21582440211069386, <https://doi.org/10.1177/21582440211069386>.

³³ Charlie PILGRIM, Weisi GUO és Thomas T. HILLS, „The Rising Entropy of English in the Attention Economy”, *Communications Psychology* 2, 1. sz. (2024), <https://doi.org/10.1038/s44271-024-00117-1>. Az eredményt a médiakörnyezetben és a figyelem szerkezetében beállt változásokra vezetik vissza. Érvelésükben az élelemszerzés analógiájára (minél több zsákmány elérhető, annál válogatósabbak az állatok) minél több információ érhető el az emberek számára, annál többet válogatnak közöttük az emberek.

sajátosságairól. Ebben kísérleti úton bizonyítják, hogy a nyelv statisztikai jellegzetességei az átadások és ismétlések révén mintegy maguktól jönnek létre; szerepük az újabb átadások könnyebbé tétele. Minél jobban megjósolhatók ugyanis az egységek egymásra következése egy jelrendszerben, annál könnyebb azt elsajátítani és továbbadni – és minél gyakoribb az átadás, annál nagyobb az esély ilyen rendezettség kialakulására.

A második tanulmányban Morin, figyelmen kívül hagyva a dekonstrukciós hagyományt, az írás eredetének új elméletét vezeti elő. Összetett gondolatmenete jól példázza a kulturális evolúció elképzelését arról, hogy nagy folyamatok miként jönnek létre kis, akár véletlenszerű lépésekből. A szöveg a korábban hivatkozottaktól eltérően nem használ saját statisztikai eljárásokat, hanem történeti forrásokra hivatkozva építi fel érvelését – a lapszámba való szerepeltetését ez is motiválja (tudniillik könnyebb olvashatósága a folyóirat feltételezett olvasói számára).

Az első két szöveg tehát az eddig említett belátásokat dolgozzák ki részletesebben, konkrét esettanulmányok alapján. A nyelv, majd az írás vizsgálatát követően, tovább szűkítve a fókusz, irodalmi szövegek kerülnek a középpontba; a lapszám második felének írásai líra-, próza- és eszméletörténeti folyamatok azonosíthatóságát célozzák.

A Benjamin Nagy és munkatársai által kidolgozott módszer – a biológiai szervezetekben az átörökítő információ, a *genom* nevére rájátszó, a versek „metrikai DNS-ét feltáró” *Metronom* – lehetővé teszi különféle költészeti hagyományok és verselési rendszerek egy univerzális séma szerinti kódolását, ezáltal transznacionális költészettörténeti folyamatok nagymintás elemzését. A szerzők a részletes módszertani áttekintést követően három rövid esettanulmányt szemléltetnek az eljárás alkalmazhatóságát az ókortól a 19. századig.

Horváth Péter már eddig is fontos tanulmányokat közölt a magyar líratörténetről kvantitatív módszertant, valamint az általa létrehozott és szerkesztett *Verskorpuszt*³⁴ alkalmazva. Ezekben, jelentős irodalomtörténeti meglátásai mellett a vázolt nyelvelméleti keret szempontjából is értékes megfigyeléseket tett. Kimutatta például, hogy a versek hossza és formai összetettsége, valamint szemantikai komplexitása között ikonikus kapcsolat áll fenn: minél egyszerűbb, vagyis szabályosabb ismétlődésen alapul, illetve minél átlagosabb hosszúságú egy költemény, annál inkább gyakoriak benne a mentálisan köny-

³⁴ HORVÁTH Péter, „ELTE Verskorpusz – a magyar kanonikus költészet gépileg annotált adatbázisa”, in *XVIII. Magyar Számítógépes Nyelvészeti Konferencia*, szerk. BEREND Gábor, GOSZTOLYA Gábor és VINCZE Veronika, 375–388 (Szeged: Szegedi Tudományegyetem TTIK, Informatikai Intézet, 2022).

nyebben hozzáférhető kifejezések.³⁵ Ez az ikonikus kapcsolatnak a bevett nyelvészeti (a grammatikai-morfológiai és a szemantikai szintek megfelelésből kiinduló) értelmezéshez képest egy általánosabb érvényű, statisztikailag megragadható vizsgálatát tette lehetővé. Mivel a nyelvhasználat ilyen sajátosságait lírai művek kapcsán mutatta be, hozzájárult ahhoz, hogy a Roman Jakobson által leírt *poétikai funkciót*, azaz egy nyelvi közlés formai és tartalmi dimenziójának ikonikus megfelelését, ne kizárólag a költői lelemény eredményeként ismerjük fel, hanem mint alapvető, általában öntudatlanul működte-tett eljárást. A lapszámban közölt tanulmánya a *Metronom* módszer párdarabjaként a versekben nem a prozódiai, hanem a szintaktikai ismétlődések gépi azonosításának és elemzésének lehetőségeit ismerteti a kanonikus magyar költészetben a 20. század első feléig.

Szemes Botond szövege azt tárgyalja, hogy mit és mit nem tesznek elgondolhatóvá a (digitális) irodalomtudomány által előszeretettel használt diagramtípusok, elsősorban a statisztikai grafikonok és a leszármazási fák. Ezt követően új paradigmaként az irodalomtörténet hálózatos megjelenítése mellett érvel. Az ezt szemléltető hálózatait magyar regények stílárís hasonlósága alapján hozza létre: egyfelől több, mint 170 év irodalmának főbb tendenciáit mutatja be, másfelől a 20. század első harmadának stílusirányzatait igyekszik a korábbi leírásokhoz képest árnyaltabban láttatni.

Végezetül Péter Róbert ismertetése olvasható a *Reception Reader* eszközéről, amely lehetővé teszi, hogy szövegrészek vándorlását nyomon követhessük nagy időbeli távolságokban. Jellemzése jól mutatja, hogy a digitális/statisztikai eljárások miben járulhatnak hozzá a bölcsészeti kutatásokhoz: „Az [...] adatalapú eszközök és algoritmusok jelentősen előmozdították az eszmetörténet kutatását, elősegítve például a fogalmak megjelenésének és átalakulásának elemzését, a hosszú távú ideológiai és fogalmi változások, valamint az eszmék transznacionális áramlásának vizsgálatát, továbbá az intellektuális hálózatok rekonstrukcióját.” Az eszköz alapját képező módszertan megegyezik a *Metronom* központi eljárásával (ún. lokális szekvenciaillesztés), ami a tematikus és szemléletbeli kapcsolatok mellett a szövegek módszertani hasonlóságaira is rámutat.

A tanulmányblokkot követően – a szokásokhoz híven – a témában fontos könyvekről közlünk recenziókat.

³⁵ HORVÁTH Péter, „A lexikai választékosság és szemantikai specifikusság mintázatai a magyar kanonikus költészetben”, in *Korpusznyelvészeti módszerek a kognitív nyelvészeti kutatásban*, szerk. KUGLER Nóra, MODRIÁN-HORVÁTH Bernadett és SIMOND Gábor, 293–331 (Budapest: ELTE Eötvös Kiadó, 2024).

A kulturális evolúció hozza létre a nyelv statisztikai szerkezetét

BEVEZETÉS

A nyelv két sajátossága biztosítja azt az alapot, amelyre a nyelvi struktúra egésze épül, meglepő azonban, hogy eredetük máig kevésbé ismert. Egyrészt a nyelv kisebb, szekvenciálisan (újra)kombinálható részekre tagolódik: a hangokat szavakká, a szavakat mondatokká lehet szervezni. Másrészt a szavak gyakorisági eloszlása rendkívül aránytalan. Míg a világ nyelvei több szempontból különböznek, abban szisztematikus hasonlóságot mutatnak, hogy meghatározott gyakorisági eloszlást követő részekből állnak.³ A tanulmányban – a kulturális evolúció kísérleti modelljére támaszkodva – bemutatjuk, hogy a nyelv ezen sajátosságok által tanulhatóbbá válik, továbbá hogy ezek generációról generációra, a kulturális átadás során alakulnak ki a nyelvhasználók ismétléses (iteratív) tanulása révén.

A nyelvet elsajátító csecsemők számára alapvető kihívás a megfelelő alegységek felismerése. A beszélt nyelvben ugyanis, számos írott nyelvvel ellentétben, a szóhatárok nincsenek egyértelműen jelezve. Ugyan többféle támpont segíthet a szóhatárok azonosításában (fonotaktika, a hangsúly jelölése, statisztikai információ stb.),⁴ egyik sem teljesen megbízható, és mindegyiket maguktól kell a csecsemőknek felfedezniük. A megfelelő egységek felismerése ugyanakkor kulcsfontosságú lépés ahhoz, hogy később eredményesen tudják azokat kombinálni. De hogyan fedezik fel a csecsemők a részeket anélkül, hogy tudnák, mit is kell keresniük?

Egy lehetséges válasz, hogy a gyermekek képesek felismerni azokat a statisztikai szabályszerűségeket, amelyek segítik a beszélt nyelv szóhatárainak kijelölését. Egy ilyen típusú, már széleskörűen tanulmányozott szabályszerű-

¹ A Jeruzsálemi Héber Egyetem Pszichológiai Tanszékének kutatója, Jeruzsálem, Izrael.

² Az Edinburghi Egyetem Filozófiai, Pszichológiai és Nyelvtudományi Intézetének kutatója, Edinburgh, Egyesült Királyság, e-mail: inbal.arnon@mail.huji.ac.il.

³ George ZIPF, *Human Behavior and the Principle of Least Effort* (Cambridge, MA: Addison-Wesley, 1949).

⁴ Peter W. JUSCZYK, „How infants begin to extract words from speech”, *Trends in Cognitive Sciences* 3, 9. sz. (1999): 323–328, [https://doi.org/10.1016/S1364-6613\(99\)01363-7](https://doi.org/10.1016/S1364-6613(99)01363-7).

ség a szótagok egymásra következési valószínűsége.⁵ Eszerint a gyermekek alapszintű statisztikai mintázatokra támaszkodhatnak abban, hogy megállapítsák, melyik szótag követ nagy valószínűséggel egy másikat egy szón belül, és hol vannak a szóhatárok. Ha a nyelv kombinálható részek összessége, akkor a részeken belüli következési valószínűség nagyobb lesz, mint az egységhatárok között. Ez a legtöbbször így is van: a szavakat alkotó szótagok következési valószínűsége nagyobb, mint azoké a szótagoké, amelyek túlnyúlnak a szóhatárokon.⁶ Sok különböző hang jelenhet meg egy szó után, mert sokféle szó követheti azt, míg egy szón belül kevesebb féle hang szerepelhet a kezdőhangot követően. Vegyük például a „aranyos kisbaba” (*pretty baby*) szekvenciát. Különböző szavak jelenhetnek meg az „aranyos” szó után (például autó, fiú, kalap, macska stb.), de csak kevés hang szerepelhet az „ara-” [*pre*] után egy értelmes szekvenciában (arabeszk, arabul stb.). Emiatt nagyobb a szó belsejében található szótag (példánkban a „-nyos” [*tty*]) következési valószínűsége, mint a szóhatáron túlnyúló szótagoké (a „kis-” [*ba*] a „-nyos” [*ty*] után). Bőséges empirikus bizonyíték támasztja alá, hogy a csecsemők, a gyermekek és a felnőttek is érzékelik ezeket a valószínűségeket, amit egy új, tagolatlan beszédflow részre bontásakor is hasznosítani tudnak.⁷ Továbbá mindez nem csak a nyelvre korlátozódik. A valószínűségek változatai más területeken is segítik a határkijelölést: a nagy valószínűség az elemek csoportosítását, míg az alacsony az elemek elkülönítését vonja maga után. A csecsemők például a következési valószínűséget vizuális alakzatok esetében is használni tudják, hogy ismétlődő elemeket fedezzenek fel alakzatok folytonos sorozatában.⁸

A statisztikai tanulásról szóló szakirodalom megoldást kínál arra, hogy a nyelvtanulók hogyan bonthatnak részekre egész megnyilatkozásokat, feltéve, hogy ezek a részek már eleve léteznek. Ha tehát a nyelv kombinálható részekből áll, a statisztikai tanulás kiszűrheti azokat. Marad azonban egy megválaszolandó kérdés. Hogyan válik a nyelv olyan részekből felépülő rendszerre, amelyek kombinációja egy ilyen típusú tanulási stratégiát hatékonyra tesz?

⁵ Jenny R. SAFFRAN, Richard N. ASLIN és Elissa L. NEWPORT, „Statistical learning by 8-month-old infants”, *Science* 274 (1996): 1926–1928, <https://doi.org/10.1126/science.274.5294.1926>.

⁶ Adele SAKSIDA, Alessandro LANGUS és Marina NESPOR, „Co-occurrence statistics as a language-dependent cue for speech segmentation”, *Developmental Science* 20, 3. sz. (2017): e12390, <https://doi.org/10.1111/desc.12390>.

⁷ Jenny R. SAFFRAN és Natasha Z. KIRKHAM, „Infant statistical learning”, *Annual Review of Psychology* 69 (2018): 181–203, <https://doi.org/10.1146/annurev-psych-122216-011805>.

⁸ Natasha Z. KIRKHAM, Jessica A. SLEMMER és Scott P. JOHNSON, „Visual statistical learning in infancy: Evidence for a domain general learning mechanism”, *Cognition* 83, 2. sz. (2002): B35–B42, [https://doi.org/10.1016/S0010-0277\(02\)00004-5](https://doi.org/10.1016/S0010-0277(02)00004-5).

Más szóval, honnan származnak a részek? Kézenfekvő magyarázat, hogy a nyelv azért áll részekből, mert a közvetíteni kívánt jelentés természetesen ugyancsak részekből áll. Azaz a nyelv szerkezete abból a késztetésből fakad, hogy kifejezően beszélhessünk a világról, amely maga is strukturált.⁹ Jelen tanulmányban ugyanakkor egy ennél erősebb hipotézist vizsgálunk: az egy-
ségek pusztán a tanulhatóság hatására jelennek meg, még akkor is, ha nincs közvetíteni kívánt jelentés. Ez egy csábító hipotézis annak megválaszolására, hogy a csecsemők miért kezdik részekre bontani a nyelvet, mielőtt megtanulnák a részek jelentését.¹⁰ Ezt erősíti továbbá, hogy más megtanulható rendszerek (például a zene) is jelentés nélküli részek kombinációjára épülnek.

A nyelv statisztikai szerkezetének egy másik figyelemre méltó sajátossága a részek gyakorisági eloszlására vonatkozik. Ha megnézzük bármely nyelv szó-készletét, felűnő, hogy nem minden szó fordul elő egyforma gyakorisággal. A szavak gyakorisági eloszlása ennek megfelelően nagyon ferde: csupán kevés szó fordul elő nagyon sokszor, míg a legtöbb kifejezetten ritkán. A gyakoriság továbbá nem lineárisan csökken. Egy nyelv szavainak az előfordulási eloszlása inkább a hatványtörvény szerinti eloszlást követi úgy, hogy egy szó gyakorisága fordítottan arányos a gyakorisági rangsorban betöltött helyével (tehát a leggyakoribb szó kétszer olyan gyakori, mint a második, és háromszor, mint a harmadik – és így tovább). Ezt a szabályt először George Kingsley Zipf mutatta ki, ezért szokás Zipf-eloszlásként hivatkozni rá.¹¹ A szavak követik a Zipf-eloszlást (vagy egy kvázi-Zipf-eloszlást¹²) a beszélt nyelvekben és jelnyel-

⁹ Simon KIRBY, Hannah CORNISH és Kenny SMITH, „Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language”, *Proceedings of the National Academy of Sciences* 105 (2008): 10681–10686, <https://doi.org/10.1073/pnas.0707835105>.

¹⁰ JUSCZYK, „How infants begin...”, 323–328; Anna FLÓ, Judit GERVAIN, Judit KISS, Judit KOVÁCS és Csaba PLÉH, „Newborns are sensitive to multiple cues for word segmentation in continuous speech”, *Developmental Science* 22 (2019): e12802, <https://doi.org/10.1111/desc.12802>.

¹¹ ZIPF, *Human Behavior*...

¹² Steven T. PIANTADOSI, „Zipf’s word frequency law in natural language: A critical review and future directions”, *Psychonomic Bulletin & Review* 21 (2014): 1112–1130, <https://doi.org/10.3758/s13423-014-0585-6>.

vekben,¹³ a gyermeknek szóló beszédben¹⁴ és a nyelvek különböző grammatikai kategóriáiban is.¹⁵ Azonban folyamatos vita zajlik az efféle eloszlások eredetéről és arról, hogy vajon ezek eredendő sajátosságai-e az emberi kogníciónak és kommunikációnak.¹⁶

Érdekes módon úgy tűnik, hogy a Zipf-eloszlás segíti a nyelvtanulást. A szavak szegmentálása egy mesterséges nyelvben javul, ha a tanulók ferde eloszlással szembesülnek egyenletes helyett (amelyben minden elem egyforma valószínűséggel jelenik meg).¹⁷ Sőt, a résztvevők akkor képesek a legjobban a szekvenciákat konstitutív elemeire bontani, amikor a részek a természetes nyelvekhez hasonló (az entrópiában megmutatkozó)¹⁸ ferde eloszlást követnek. A ferde eloszlásról az is kiderült, hogy elősegíti a vizuális statisztikai tanulást (ismétlődő vizuális hármassok felismerését egy összefüggő fo-

¹³ PIANTADOSI, „Zipf’s word frequency law in natural language...”, 1112–1130; Ilanit KIMCHI, Rachel STAMPS, Liam WOLTERS és Isaac ARNON, „Evidence of Zipfian distributions in three sign languages”, *Gesture* (elfogadva, megjelenés alatt); Ali MEHRI és Mohammad JAMAATI, „Variation of Zipf’s exponent in one hundred live languages: A study of the *Holy Bible* translations”, *Physics Letters A* 381, 31. sz. (2017): 2470–2477, <https://doi.org/10.1016/j.physleta.2017.05.061>; Ramon FERRER-I-CANCHO, „The variation of Zipf’s law in human language”, *European Physical Journal B* 44, 2. sz. (2005): 249–257, <https://doi.org/10.1140/epjb/e2005-00121-8>.

¹⁴ Ofir LAVI-ROTBAIN és Isaac ARNON, „Zipfian Distributions in Child-Directed Speech”, *Open Mind* 7 (2022): 1–30, https://doi.org/10.1162/opmi_a_00070.

¹⁵ PIANTADOSI, „Zipf’s word frequency law in natural language...”, 1112–1130; Carl BORSTELL, „Searching and utilizing corpora”, in *Signed Language Corpora*, szerk. Julie A. HOCHGESANG és Jordan FENLON, 90–127 (Washington, DC: Gallaudet University Press, 2022), <https://doi.org/10.2307/j.ctv2rcnfhc.9>.

¹⁶ PIANTADOSI, „Zipf’s word frequency law in natural language...”, 1112–1130; Ramon FERRER-I-CANCHO, Christian BENTZ és Cyrille SEGUIN, „Optimal coding and the origins of Zipfian laws”, *Journal of Quantitative Linguistics* 29, 2. sz. (2020): 165–194, <https://doi.org/10.1080/09296174.2020.1778387>; Edward GIBSON, Steven T. PIANTADOSI, Kyle MAHOWNEY, Evelina FEDORENKO, Roman FEDELENKO, Roger LEVY, Leon BERGEN, Martin T. SCHNUR és John T. HALE, „How efficiency shapes human language”, *Trends in Cognitive Sciences* 23, 5. sz. (2019): 389–407, <https://doi.org/10.1016/j.tics.2019.02.003>; Sarah SEMPLE, Ramon FERRER-I-CANCHO és Melissa L. GUSTISON, „Linguistic laws in biology”, *Trends in Ecology & Evolution* 37, 1. sz. (2022): 53–66, <https://doi.org/10.1016/j.tree.2021.08.012>.

¹⁷ Chigusa KURUMADA, Stephan C. MEYLAN és Michael C. FRANK, „Zipfian frequency distributions facilitate word segmentation in context”, *Cognition* 127, 3. sz. (2013): 439–453, <https://doi.org/10.1016/j.cognition.2013.02.002>; Ofir LAVI-ROTBAIN és Isaac ARNON, „The learnability consequences of Zipfian distributions in language”, *Cognition* 223 (2022): 105038, <https://doi.org/10.1016/j.cognition.2022.105038>.

¹⁸ Uo.

lyamban);¹⁹ a szituációkon átívelő szótanulást;²⁰ valamint az új grammatikai kategóriák²¹ és konstrukciók²² tanulását. A beszélőknek láthatólag kognitív preferenciája az efféle eloszlás: még kitalált főnevek egyenletes eloszlása is ferdévé válik, amikor a résztvevők egy történetet adnak tovább egymásnak.²³ Bár kevés tanulmány vizsgálja e mintázat hatását a nyelvtanulásra, ezek szintén a ferde eloszlás előnyeire világítanak rá számos nyelvi és nemnyelvi viszony esetében. Ezek az eredmények azt mutatják, hogy a Zipf-eloszlás elősegítheti a nyelvtanulást, mégis nyitva hagynak egy másik kérdést: miért alakul ki ez a tanulást megkönnyítő statisztikai törvény a nyelvben?

A nyelv mindkét visszatérő és szisztematikus sajátossága – az, hogy egységekből áll, és ezek meghatározott eloszlást követnek – a tanulhatóságban játszik fontos szerepet: segítik a tanulót a lényeges nyelvi egységek felismerésében és megjegyzésében. A tanulmányban azt mutatjuk be, hogy a nyelv statisztikai sajátosságai nemcsak elősegítik a tanulást, hanem ez utóbbi legfőbb oka létezésüknek. A nyelv statisztikai szerkezete tehát egyszerre oka és okozata a hatékony tanulásnak. Ez a látszólag körbenforgó érv egy olyan nyelvszemléletből ered, amely a kulturális átadás folyamatára épül. Ennek értelmében a generációk a korábbiak tanulásának eredményeként létrejött kimenetet sajátítják el. Más szóval a nyelv, amelyet a gyermek megpróbál szegmentálni, egy hasonló szegmentációs folyamat eredménye, amelyen nevelőik vagy a nyelvi közösség más tagjai már korábban keresztülmentek. És ugyanígy az ő tanulásuk során kialakult változat is nyelvi adattal fog szolgálni a jövőbeli generációk számára. Az ismétléses tanulás ilyen folyamata megfelelő magyarázatnak bizonyult a nyelv több szerkezeti sajátossága esetében

¹⁹ Ofir LAVI-ROTBAIN és Isaac ARNON, „Visual statistical learning is facilitated in Zipfian distributions”, *Cognition* 206 (2021): 104492, <https://doi.org/10.1016/j.cognition.2020.104492>.

²⁰ Anne T. HENDRICKSON és Amy PERFORs, „Cross-situational learning in a Zipfian environment”, *Cognition* 189 (2019): 11–22, <https://doi.org/10.1016/j.cognition.2019.03.005>.

²¹ Katherine D. SCHULER, Patricia A. REEDER, Elissa L. NEWPORT és Richard N. ASLIN, „The effect of Zipfian frequency variations on category formation in adult artificial language learning”, *Language Learning and Development* 13, 4. sz. (2017): 357–374, <https://doi.org/10.1080/15475441.2016.1263571>.

²² Jennifer K. BOYD és Adele E. GOLDBERG, „Input effects within a constructionist framework”, *The Modern Language Journal* 93, 3. sz. (2009): 418–429, <https://doi.org/10.1111/j.1540-4781.2009.00899.x>.

²³ Amir SHUFANIYA és Isaac ARNON, „A cognitive bias for Zipfian distributions?: Uniform distributions become more skewed via cultural transmission”, *Journal of Language Evolution* 7, 1. sz. (2022): 59–80, <https://doi.org/10.1093/jole/lzac005>.

is, úgy mint a kompozicionalitás,²⁴ a kettős tagolás,²⁵ a szabályosság²⁶ vagy a szavak szemantikája²⁷ kapcsán, valamint még az olyan nemnyelvi struktúrák esetében is, mint a dobminták²⁸ és a versformák kialakulása.²⁹ Általánosságban véve a több generáción át tartó ismétléses tanulás olyan nyelvek kialakulásához vezet, amelyek alkalmazkodnak a nyelvet továbbadók tanulási preferenciáihoz.³⁰

Érvelésünkben az ismétléses tanulás folyamata egy meghatározott tanulási stratégiával párosítva – amely az egésztől halad a részek felé – hozza létre a nyelv két statisztikai sajátosságát. Míg a nyelvelsajátítás tankönyvi leírása gyakran a résztől az egész felé haladó tanulást hangsúlyozza, amelyben a kisebb egységek nagyobb egészekké állnak össze (például a szótagok szavakká, a szavak mondatokká), egyre több bizonyíték van arra, hogy az egésztől a részek felé haladó stratégia – amelyben az egész később részekké bomlik fel – szintén fontos szerepet játszik,³¹ elsősorban a nyelvi struktúrák megtanulásában. A csecsemők ugyanis kezdetben nem tudják, hol vannak a szóhatárok, vagy hogy egyáltalán a szavak léteznek. Korai építőelemeiket szavak egyvelege és többszavas mondatok képezik, amelyek csak később bomlanak részekre (például: mi-ez [*what's-this*]). Ha a nyelvtanulást a nagyobb elemek felől kezdjük, amelyeket csak ezután szegmentálunk, könnyebb a szavak közötti önké-

²⁴ KIRBY, CORNISH és SMITH, „Cumulative cultural evolution in the laboratory...”, 10681–10686.

²⁵ Thomas VERHOEF, Simon KIRBY és Bart DE BOER, „Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals”, *Journal of Phonetics* 43 (2014): 57–68, <https://doi.org/10.1016/j.wocn.2014.02.005>.

²⁶ Uo.

²⁷ James W. CARR, Kenny SMITH, Jennifer CULBERTSON és Simon KIRBY, „Simplicity and informativeness in semantic category systems”, *Cognition* 202 (2020): 104289, <https://doi.org/10.1016/j.cognition.2020.104289>.

²⁸ Andrea RAVIGNANI, Teresa DELGADO és Simon KIRBY, „Musical evolution in the lab exhibits rhythmic universals”, *Nature Human Behaviour* 1, 1. sz. (2016): 0007, <https://doi.org/10.1038/s41562-016-0007>.

²⁹ Valeria DE CASTRO-ARRAZOLA és Simon KIRBY, „The emergence of verse templates through iterated learning”, *Journal of Language Evolution* 4, 1. sz. (2019): 28–43, <https://doi.org/10.1093/jole/lzy013>.

³⁰ Michael L. KALISH, Thomas L. GRIFFITHS és Simon KIRBY, „Iterated learning: Intergenerational knowledge transmission reveals inductive biases”, *Psychonomic Bulletin & Review* 14, 2. sz. (2007): 288–294, <https://doi.org/10.3758/BF03194066>; Simon KIRBY, Mike DOWMAN és Thomas L. GRIFFITHS, „Innateness and culture in the evolution of language”, *Proceedings of the National Academy of Sciences* 104 (2007): 5241–5245, <https://doi.org/10.1073/pnas.0608222104>.

³¹ Uo.; Isaac ARNON, *Starting Big – The Role of Multi-word Phrases in Language Learning and Use* (Stanford, CA: Stanford University, 2010).

nyes grammatikai viszonyok elsajátítása (ilyen például a nyelvtani nemmel rendelkező névelők és a főnevek közötti viszony egyes nyelvekben).³² Mivel a felnőttek már tudják, hogy léteznek szavak, és ismerik is anyanyelvük szó-készletét, kevésbé támaszkodnak a többszavas egységekre, ami hátrányos a grammatikai viszonyok tanulásakor. Empirikus bizonyíték szolgál arra, hogy a csecsemők már a korai szakaszban azonosítanak többszavas egységeket,³³ a felnőttek viszont ezt kevésbé teszik a nyelvtanulás során;³⁴ a nagyobb egységek előtérbe helyezése ugyanakkor elősegíti a nyelvtan elsajátítását.³⁵ Ezek a tanulmányok jól szemléltetik az egésztől a részek felé haladó tanulás jelenlétét a nyelvelsajátítás és a nyelvi struktúrák megismerése során.

A statisztikai tanulásról, az ismétléses tanulásról és az egésztől a részek felé haladó tanulásról szóló szakirodalmakra építve az a sejtésünk, hogy a nyelvi elemek eloszlásának mintázatai szintén a kulturális átadás során jönnek létre. Ha az egésztől a részek felé haladó tanulási stratégia generációkon át ismétlődik, akkor a kezdetben holisztikus nyelvből fokozatosan kialakulnak

³² Greville G. CORBETT, „Gender, grammatical”, in *Encyclopedia of Language & Linguistics*, szerk. Keith BROWN, 749–756 (Amsterdam: Elsevier, 2006²), <https://doi.org/10.1016/B0-08-044854-2/00191-7>.

³³ Isaac ARNON, Shane M. McCAULEY és Morten H. CHRISTIANSEN, „Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases”, *Journal of Memory and Language* 92 (2017): 265–280, <https://doi.org/10.1016/j.jml.2016.07.004>; Barbora SKARABELA, Mitsuhiko OTA, Rory O’CONNOR és Isaac ARNON, „‘Clap your hands’ or ‘take your hands’?: One-year-olds distinguish between frequent and infrequent multiword phrases”, *Cognition* 211 (2021): 104612, <https://doi.org/10.1016/j.cognition.2021.104612>.

³⁴ Shane M. McCAULEY és Morten H. CHRISTIANSEN, „Computational investigations of multiword chunks in language learning”, *Topics in Cognitive Science* 9, 3. sz. (2017): 637–652, <https://doi.org/10.1111/tops.12258>; Noam HAVRON és Isaac ARNON, „Reading between the words: The effect of literacy on second language lexical segmentation”, *Applied Psycholinguistics* 38 (2017): 127–153, <https://doi.org/10.1017/S0142716416000138>.

³⁵ Isaac ARNON és Eve V. CLARK, „Why ‘brush your teeth’ is better than ‘teeth’—Children’s word production is facilitated in familiar sentence-frames”, *Language Learning and Development* 7 (2011): 107–129, <https://doi.org/10.1080/15475441.2010.505489>; Isaac ARNON és Michael RAMSCAR, „Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned”, *Cognition* 122 (2012): 292–305, <https://doi.org/10.1016/j.cognition.2011.10.009>; Noam SIEGELMAN és Isaac ARNON, „The advantage of starting big: Learning from unsegmented input facilitates mastery of grammatical gender in an artificial language”, *Journal of Memory and Language* 85 (2015): 60–75, <https://doi.org/10.1016/j.jml.2015.07.003>; Noam HAVRON, Limor RAVIV és Isaac ARNON, „Literate and pre-literate children show different learning patterns in an artificial language learning task”, *Journal of Cultural and Cognitive Science* 2 (2018): 21–33, <https://doi.org/10.1007/s41809-018-0015-9>.

újrarendezhető részek.³⁶ Mindez egy idő után meg kell mutakozzon a létrejövő nyelv statisztikai sajátosságaiban, vagyis az egységeken belüli következési valószínűségeknek nagyobbaknak kell lenniük, mint az egységhatárokon. Ezen kívül, ha a részek ferde eloszlása egyszerre oka és okozata a hatékonyabb tanulásnak, azt várjuk, hogy az eloszlás idővel egyre ferdebb lesz, és még inkább illeszkedik a nyelv hatványtörvényszerű működéséhez.

A hipotézis ellenőrzésére a nyelv kulturális evolúciójának egy kísérleti modelljét hívjuk segítségül, amelyben a résztvevők „nyelvet” tanulnak. A soron következő résztvevőnek reprodukálnia kell egy olyan „nyelvet”, amelyet egy korábbi résztvevő generált; a kimenet tehát bemenetté válik a következő résztvevő számára. A kísérletben a közvetített tartalom nem nyelvi jellegű, hanem színszekvenciákból áll. Mindegyik szekvencia egy megnyilatkozásnak felel meg, összességük pedig maga a megtanulandó „nyelv”. Alapvető fontosságú, hogy a kezdeti szekvenciahalmaz nem strukturált: mindegyik színkombináció egyformán lehetséges, azaz a szekvenciák nem ismétlődő részekből állnak. Kérdés, hogy a kezdeti sorozatot a résztvevők felruházzák-e az átadás folyamán a nyelvben fellelhető két statisztikai sajátossággal: részekből fog-e állni, amelyek egy meghatározott gyakorisági eloszlást követnek. A részek azonosításához a statisztikai tanulásról szóló szakirodalomra támaszkodunk (a részleteket lásd lentebb).

Fontos megjegyezni, hogy a feladat explicit módon egy egész szekvencia másolására irányul, mintsem kombinálható részek megtalálására; és úgy alakítottuk ki, hogy ez a másolás nehéz legyen, ami nyomást gyakorol a rendszerre, hogy megváltozzon és tanulhatóbbá váljon. Két okból adunk nem nyelvi feladatot. Egyrészt azért, mert ekkor sem a részeknek (amelyeknek ki kell alakulniuk), sem az egésznek (a szekvencia) nincs jelentése. Ez lehetővé teszi, hogy ellenőrizzük hipotézisünket, miszerint a statisztikailag elkülönülő részek, amelyek meghatározott ferde eloszlásba rendeződnek, a jelentéstől függetlenül alakulnak ki, illetve elősegítik a rendszer további elsajátítását. Másrészt a nem nyelvi ingerek, főként a jelentés nélküli egyedi elemek (és azok kombinációi), minimalizálják annak kockázatát, hogy a résztvevők egyszerűen áthelyezzék a saját nyelvük statisztikai szerkezetének ismeretét a kulturálisan átadott tartalmakra. Így tehát a kombinálható részek létrejötté a

³⁶ Alison WRAY, „Protolanguage as a holistic system for social interaction”, *Language & Communication* 18, 1. sz. (1998): 47–67, [https://doi.org/10.1016/S0271-5309\(97\)00033-5](https://doi.org/10.1016/S0271-5309(97)00033-5); Simon KIRBY, „Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners”, in *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, szerk. Chris KNIGHT, 303–323 (Cambridge: Cambridge University Press, 2000), <https://doi.org/10.1017/CBO9780511606441.019>.

tanulók implicit preferenciáinak vagy elfogultságainak (*bias*) az eredménye lesz az ismétléses tanulás során, nem pedig az általuk bevezetett tudatos szabályokénak.

MÓDSZER

Kutatásunkban egy korábbi kísérlet eredményeit elemezzük újra,³⁷ amely szintén színszekvenciák ismétléses tanulását vizsgálta, hogy ellenőrizzük hipotézisünket, miszerint a nyelv statisztikai szerkezete a kulturális evolúció révén jön létre a nyelvtanulók ismétléses, az egésztől a részek felé haladó tanulása során. Az eredeti kísérletben azt vizsgáltuk, hogy vajon az absztrakt, ismétléses tanulás által közvetített szekvenciák generációnként tanulhatóbakká válnak-e. Nem elemeztük azonban a szekvenciák belső struktúráját, és a kísérletet sem ennek érdekében alakítottuk ki. Ennek az adathalmaznak az újraelemzésével tehát kifejezetten hatékonyan ellenőrizhetjük hipotézisünket, hiszen ez felszámolja a kutató elfogultságát, amely akkor keletkezik, ha egy kísérletet egy már előre remélt eredmény fényében tervez meg.

A kísérletet a „Simon” elnevezésű társasjáték online változatának mintájára alkottuk meg. A résztvevők felvillanó fényeket látnak egy kétszer kettes színtáblán (piros, zöld, sárga, kék). Közvetlenül ezután vissza kell idézniük ezt a fényszekvenciát a vonatkozó gombok megnyomásával, majd pontszám formájában visszajelzést kapnak arról, hogy milyen pontosan teljesítették a feladatot. Az eredeti játékhoz képest, amelyben a játékosok először csak egyetlen villanást látnak, majd kettőt, majd hármat stb., kísérletünkben a résztvevők mindig egy teljes szekvenciát látnak és idéznek fel.

A résztvevők átadási láncolatokba szerveződnek, ahol az n -edik résztvevő válaszait az $n + 1$ -ediknek kell lemásolnia (a szekvenciák sorrendje véletlenszerű). Minden résztvevő egy 60 elemből álló szekvenciahalmazt lát. A halmazt kétszer, két részletben mutatjuk meg, a szekvenciákat mindkétszer véletlenszerűen rendezve (a résztvevők nincsenek tudatában annak, hogy két részlet van, csak az egymást szünet nélkül követő szekvenciákat látják). A kiértékeléskor a második blokk válaszait rögzítjük és dolgozzuk fel. Minden láncolat első résztvevőjének egy olyan véletlenszerűen összerakott halmazt kell lemásolnia, amelyben egy szekvencia 12 villanásból áll, és mind a négy szín egyenletes el-

³⁷ Hannah CORNISH, Kenny SMITH és Simon KIRBY, „Systems from sequences: An iterated learning account of the emergence of systematic structure in a non-linguistic task”, in *Proceedings of the Annual Meeting of the Cognitive Science Society* 35 (2013): 340–345.

oszlással szerepel benne. A kezdeti véletlenszerű halmaz minden láncolatban más. A kiinduló szekvenciák hosszúsága azért 12 elem, hogy ne legyen se túl könnyű, se túl nehéz megjegyezni: szándékunk szerint a szekvenciák lemásolása elég nehéz ahhoz, hogy hibát is vétsenek a résztvevők, de nem túl nehéz ahhoz, hogy ne legyen reprodukálható. A lemásolt szekvenciák ezután a láncolat következő résztvevőjéhez kerülnek. Ha egy résztvevő 6 elemnél rövidebb szekvenciát hoz létre, akkor a forrásszekvenciát újra megmutatjuk az első blokk végén.

Minden láncolat 10 generáción keresztül folytatódik, azaz 660 szekvenciát elemzünk láncolatonként (egy véletlenszerű, 60 elemből álló kezdőhalmazt és 10 darab szintén 60 elemből álló sorozatot, amelyet a résztvevők generálnak). Hat ilyen láncolat eredményeit vizsgáltuk, mindegyikre ugyanazt a beállítást alkalmazva.³⁸ Ezáltal 3960 szekvenciát kapunk, generációk és láncolatok szerint rendezve ($60 \text{ szekvencia} \times 10 \text{ generáció} \times 6 \text{ láncolat} + 6 \text{ kezdeti } 60 \text{ elemű halmaz}$, összesen 360 szekvencia). Megjegyzendő, hogy a feladat elején a résztvevők nem kapnak utasítást arra, hogy mintákat másoljanak, vagy hogy figyeljenek a halmaz egészének jellemzőire. A szekvenciákban felmerülő bármely szisztematikus rendszer tehát egy implicit tanulási torzítás eredményeként jöhet csak létre, amely hatással van a halmazok egészének kialakítására is. Ahogyan már említettük, a feladatot nehéznek szántuk, főleg a láncolatok legelején: azt várjuk, hogy a résztvevők sok hibát vétsenek a szekvenciák reprodukálásakor, különösen a korai generációk esetében. Kérdés, hogy a hibák elvezetnek-e egy strukturáltabb (és tanulhatóbb) rendszerhez.

EREDMÉNYEK

A teljes adathalmaz, illetve az ábrák létrehozásához és statisztikai elemzéséhez szükséges kódok elérhetők az alábbi repozitóriumban: https://github.com/smkirby/cultural_evolution_of_statistical_structure. Az 1. ábra a hibaarány generációnkénti csökkenését mutatja. A hibaérték kiszámítása a forrásszekvencia és a résztvevő által előállított szekvencia közötti normalizált Levenshtein-távolság felhasználásával történt – az 1-es érték teljes eltérést jelent a forrásszekvenciától, míg a 0-ás érték tökéletes egyezést jelöl. Egy lineáris ve-

³⁸ Az első négy láncolat elemzésének eredményei: CORNISH, SMITH és KIRBY, „Systems from sequences...”; a másik kettő eredményei: SIMON KIRBY, HANNAH CORNISH és KENNY SMITH, „Systems emerge: The cultural evolution of interdependent sequential behaviours in the lab”, in *Evolution of Language: Proceedings of the 10th International Conference* (2014): 463–464, https://doi.org/10.1142/9789814603638_0095.

gyes hatású modellt³⁹ alkalmaztunk a hibák alakulásának előrejelzésére. A modell figyelembe vette a generáció hatását mint rögzített hatást,⁴⁰ a láncok közötti különbségeket mint véletlen eltéréseket,⁴¹ generáció hatásának láncolatonkénti eltérését,⁴² az eredeti szekvencia hatását (a 0. generációban) mint véletlen eltérést.⁴³ (Próbálkoztunk azzal is, hogy az eredeti szekvencia hatását láncenként külön-külön modellezzük, de ez a modell matematikailag instabillá vált [ún. „singular fit”], ezért leegyszerűsítettük úgy, hogy véletlen interceptet használtunk.)⁴⁴ Generációnkénti szignifikáns hibaérték-csökkenést észleltünk ($\beta = -0.023$, $SD = 0.0032$, $p < 0.001$). Más szóval, a szekvenciák aszerint változnak, hogy a résztvevők egyszerűbben tudják lemásolni azokat. Ez megszokott eredménye az ismétléses tanulási kísérleteknek, amelyekben pusztán a generációk közötti átadásnak köszönhetően tanulhatóbakká válnak a továbbadott minták. A 2. ábra egy láncolat egy példaszekvenciáját és annak generációnkénti változását mutatja (emellett bemutatja az alább kifejtett szegmentációs módszerrel azonosítható alegységek kialakulását is). A könnyebben másolható szekvenciák nagy valószínűséggel maradandóbbak lesznek, míg a nehezebben másolhatók változni fognak a hibák miatt. Ennek köszönhetően a szekvenciák az idő elteltével egyre könnyebben másolhatókká válnak.

³⁹ A lineáris vegyes hatású modellek olyan statisztikai módszerek, amelyek lehetővé teszik, hogy egy elemzés egyszerre vegye figyelembe az általános törvényszerűségeket (például egy változó átlagos hatását) és az egyedi eltéréseket (például azt, hogy egyes csoportok vagy megfigyelési egységek eltérően viselkednek az átlagtól). A modell tulajdonképpen egy egyenletként értelmezhető, amelyben a különféle – rögzített és véletlenszerű – hatások szerepelnek. Ez az egyenlet segít megbecsülni, hogyan befolyásolják ezek a tényezők az empirikusan megfigyelt eredményeket. Jelen esetben például azt, hogy a generációk egymásra gyakorolt befolyása-e a résztvevők által vétett hibák számát, ha erre potenciálisan más tényezők is hatnak – *a szerk.*

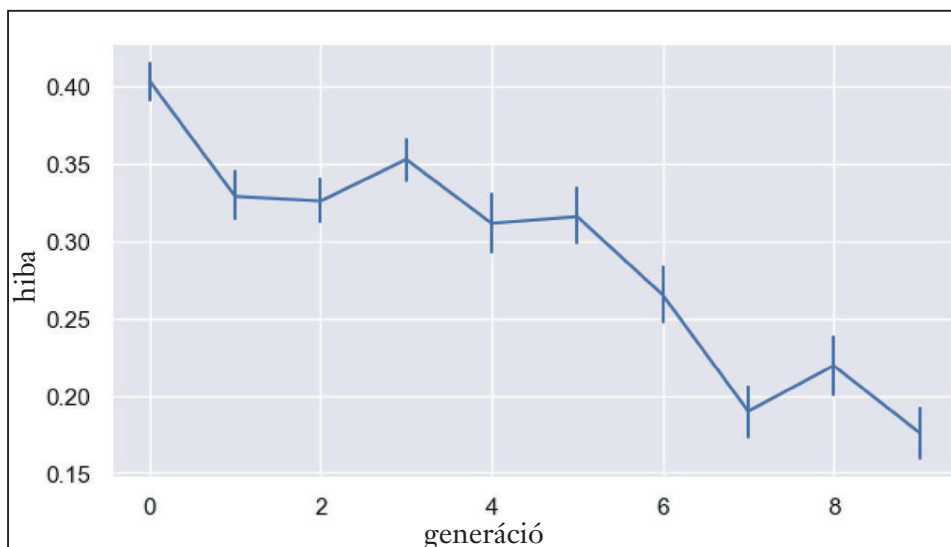
⁴⁰ Ez annak vizsgálatát jelenti, hogy hogyan befolyásolja a generációk száma a hibák mértékét – *a szerk.*

⁴¹ A modell megengedi, hogy minden lánc saját, kissé eltérő kiinduló hibaszinttel rendelkezzen – *a szerk.*

⁴² Tehát, hogy a generációk hatása ne legyen egyforma minden láncban – *a szerk.*

⁴³ Vagyis azt, hogy a kiindulópontként használt szekvencia önmagában is befolyásolhatja a hibákat – *a szerk.*

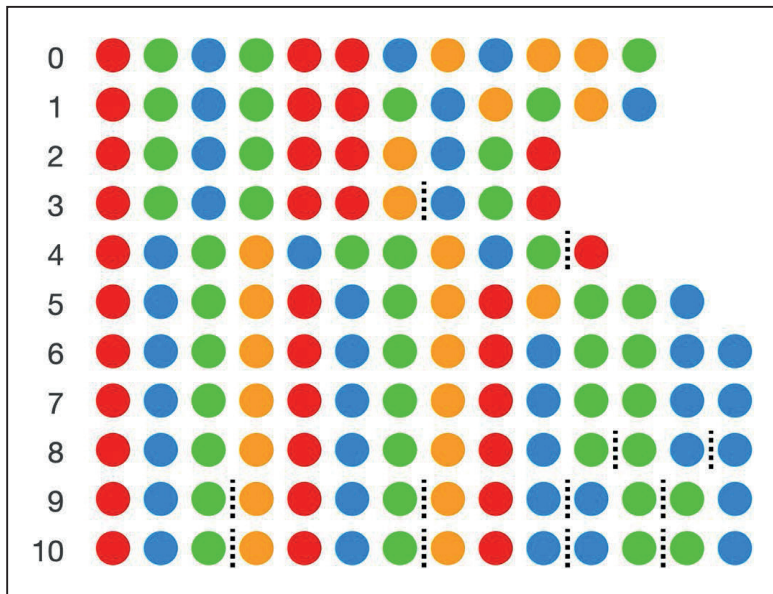
⁴⁴ Ez azt jelenti, hogy nem modellezték részletesen, hogyan viselkedik minden szekvencia minden láncban külön-külön, hanem csak azt vették figyelembe, hogy már az elején (a 0. generációban) lehettek különbségek a szekvenciák között abból a szempontból, hogy mennyire voltak hajlamosak hibákra – *a szerk.*



1. ábra. A hibaarány változása a kísérletben vizsgált generációk között. A hibaértéket átlagolva adjuk meg az egyes generációk különböző láncolataiban (a hibaérték minden szekvenciára külön-külön kiszámítva, majd az adott halmazhoz tartozó szekvenciák hibaértékei átlagolva, végül a hat láncolat értékei átlagolva). A hibaszávok 95%-os konfidenciaintervallumokat jelölnek, amelyeket bootstrap (újramintavételezésen alapuló) módszerrel számoltunk ki. Jól láthatóan csökkenő tendencia figyelhető meg a hibaértékekben, ami azt mutatja, hogy a szekvenciahalmazok a későbbi generációkban könnyebben másolhatók a korábbi generációkhoz képest. A szekvenciák tehát tanulhatóbbakká váltak.

A SZEGMENTÁCIÓ KIALAKULÁSA

A szekvenciákat annak megfigyelésére is felhasználtuk, hogy vajon megjelennek-e bennük alegységek egyik generációról a másikra. Ne feledjük, hogy a statisztikai tanulás elméletei szerint a szótagok közötti következési valószínűség a szóhatárok jelzéseként funkcionál. Akkor történik szegmentálás, amikor csökken egy szótagszekvencia következési valószínűségének értéke – ez a csökkenés jelzi a szóhatárok lehetőségét, hiszen a szavakon belül a következési valószínűség általában nagyobb, mint a szavak között. Mivel a nyelvnek visszatérő szavai vannak, a szavakon belüli követkeзések kevésbé váratlanok, mint a szavak közöttiek. Felvethetjük, hogy hasonló mintázatok jelennek meg a mi adatbázisunkban is. A „Simon” játékban négy szín képezi az alapeleme-



2. ábra. Mi történik a kísérlet során egy meghatározott szekvenciával (a 60-ból) egy meghatározott láncolatban (a 6-ból)? A számértékek a kísérletben vizsgált generációkra vonatkoznak, a 0. generáció a kezdeti véletlenszerű sorozatot jelöli, az utána következő generációk pedig azok a szekvenciák, amelyeket a résztvevők generáltak.

Kezdetben a véletlenszerű szekvenciákat nehéz lemásolni, ezért több változtatás figyelhető meg, de generációk után a változtatások könnyebbé teszik a szekvenciák pontos ismétlését. A szaggatott vonalak az alább bemutatott szegmentációs módszerünkkel azonosított részek (amelyeken belül kifejezetten alacsony a következősi valószínűség) közötti határokat mutatják. Annak mértéke, hogy egy ehhez hasonló szekvencia könnyen vagy nehezen másolható, a halmaz többi 59 szekvenciájától is függ. Ugyanúgy, hogy mi számít szegmentációs határnak, az a teljes halmazra vonatkoztatott következősi valószínűség függvénye.

ket, a következősi valószínűség tehát az adott szín megjelenésének valószínűségét jelöli egy másik után. Ha létrejönnek alegységek, akkor bizonyos színek nagyobb valószínűséggel következnek egymásra, mint mások. Mindezek alapján a statisztikai tanulás szakirodalmára támaszkodva kidolgoztunk egy eljárást, amely képes azonosítani a következősi valószínűség csökkenését egy adott színszekvenciában, ami alegységek kialakulását jelezheti.

Az egymás utáni következősi valószínűségek hányadosán keresztül ugyanis számszerűsíthető a szegmentáció. Ehhez meg kell határoznunk magukat a valószínűségeket, amit két korábbi elem alapján teszünk meg, vagyis egy szín

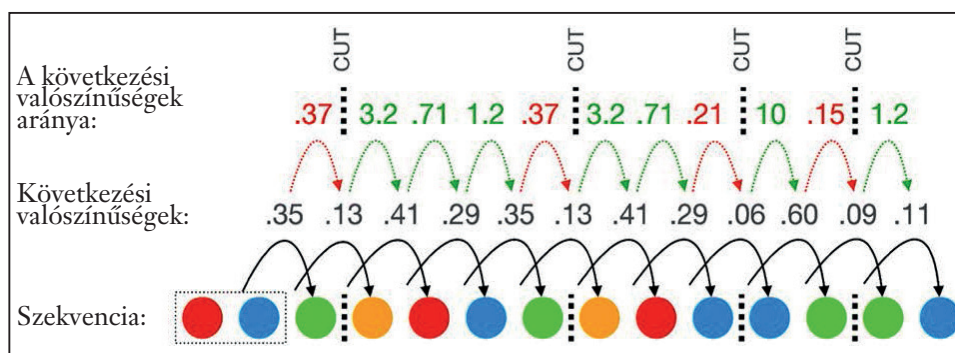
megjelenésének valószínűségét az azt megelőző két színhez viszonyítva számítjuk ki. Egy másik lehetőség az lett volna, ha csak az előző színt használjuk fel, de mivel csak négy színből áll a jelkészlet, ez az adatsor nagyon pontatlan becsléséhez vezetett volna. A teljes feltételes valószínűség elvét (vagyis az összes addigi színt a sorozatban) is használhattuk volna, ám az egyes generációkban rendelkezésre álló kis adatállomány (mindössze 60 szekvencia) miatt ez adatritkasági problémákhoz vezetne. Ezért a két megelőző szín figyelembevételével jelenthet kompromisszumot a két szélső érték között.

A következősi valószínűség legpontosabb becslést értékét egy adott generáció teljes, 60 elemből álló szekvenciahalmazában számítjuk ki egyetlen láncolat alapján. Tehát például a „piros-sárga-zöld-piros”, vagyis a „PSZP” szekvencia esetében megbecsüljük a „Z” elem következősi valószínűségét a „PS” után a teljes halmazhoz viszonyítva. Ehhez először megbecsüljük az „PSZ” elemcsoport előfordulási valószínűségét, majd elosztjuk a „PS” előfordulási valószínűségével. A valószínűségek mindig a teljes, 60 elemből álló szekvenciahalmaz függvényei. Aztán megbecsüljük a „P” következősi valószínűségét a „SZ” után. Ha az eredmény azt mutatja, hogy a „Z” nagy valószínűséggel következik „PS” után, de a „P” esetében ez nem mondható el, akkor jó okunk van rá, hogy a szekvenciát „PSZ–P” alegységekre szegmentáljuk.

Valahogyan el kellett döntenünk azt is, hogy a következősi valószínűség mikor elég alacsony ahhoz, hogy határt jelölhessünk ki. Egyik lehetőség, hogy a határt magára a következősi valószínűségre támaszkodva határozzuk meg, valahányszor egy bizonyos érték alá esik. Ezzel viszont nem lehet tetten érni a relatív csökkenést. Míg például egy 0,4-es következősi valószínűség nem túlságosan alacsony, egy 0,99-től 0,4-ig terjedő csökkenés szignifikáns lenne. Az ilyen típusú relatív csökkenések megragadására az egymást követő következősi valószínűségek közötti arányt (a két érték hányadosát) használjuk. Kiszámítjuk egy következősi valószínűség és az azt megelőző valószínűség arányát a szekvenciában (azaz a későbbit osztjuk el a korábbival), majd szegmentációs határt jelölünk ki, ha az arány szokatlanul alacsony (más szóval, amikor a következősi valószínűség hirtelen lecsökken). Az osztást követően 1 alatti értékek a következősi valószínűség csökkenését jelzik, míg az 1 fölötti arányok annak növekedését (vö. 3. ábra).

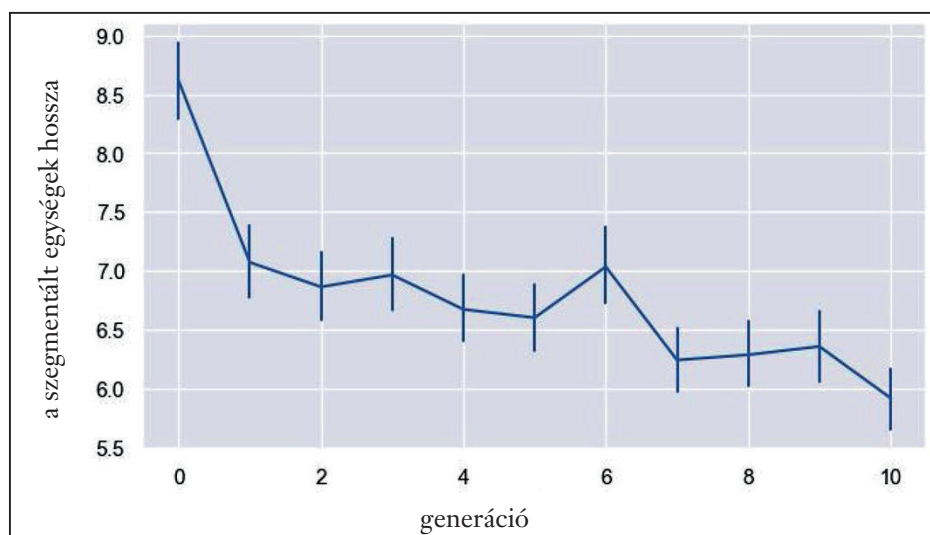
Ekkor már csak egy olyan módszerre van szükség, amellyel megítélhetjük, hogy egy következősi valószínűség-arány kellően alacsony-e ahhoz, hogy a sorozatot ezen a ponton meg lehessen szakítani. Ha túl alacsonyra szabjuk a küszöbértéket, akkor félő, hogy figyelmen kívül hagyjuk a tényleges szegmentációs határokat; ha pedig túl magasra, akkor szétválaszthatunk olyan ré-

szeket, amelyek inkább egybe tartoznának. Ennek megoldásaként a kezdeti véletlenszerű szekvenciákhoz fordultunk, mivel ezek nem rendelkeznek előre megadott struktúrával (azaz a színek egymásutánisága véletlenszerű és kiszámíthatatlan). Ezt a nemstrukturált halmazt használhatjuk annak megbecsülésére, hogy mennyire várható egy adott következési valószínűség csökkenése akkor, ha a szekvenciában nincs semmiféle belső rendszer. Ezáltal azonosíthatjuk a váratlanul nagy csökkenést jelző küszöbértéket. Ehhez kiszámítjuk a becsült következési valószínűség-arányok eloszlását mind a hat kezdeti, véletlenszerű sorozatban (minden láncolatból egyet), majd ezek összesített eloszlást hozzuk létre, amelynek 5%-os alsó régiója fogja képezni a küszöbértéket. Akkor szegmentálunk, ha a mért arány ennél (0,425) kisebb.



3. ábra. Példa a szekvenciák szegmentálására szolgáló eljárásra. A példa a kísérlet utolsó generációjának egyik szekvenciáját ábrázolja a megállapított szegmentációs határokkal. A szekvencia fölött azokat az értékeket tüntetjük fel, amelyek megmutatják, hogy egy adott szín milyen valószínűséggel következik az előtte lévő két szín után. Ezek az a láncolat teljes, 60 szekvenciából álló halmazára vonatkoztatva kerülnek kiszámításra. Egységhatárt tettünk oda, ahol jelentős csökkenés figyelhető meg a szekvencia következési valószínűségeiben. Ezt az egymásra következő valószínűségek kirívóan alacsony hányadosaként operacionalizáljuk. Az arányok (hányadosok) az ábra felső sorában találhatók. Például a következési valószínűség csökkenése 0,35-ről 0,13-ra kisebb a 0,425-ös küszöbértékünkénél, ami a szekvencia első zöld és első sárga egysége közötti határra utal. Piros színnel jelöltük, ha a következési valószínűségek hányadosa kisebb a küszöbértékünkénél, zölddel pedig, ha annál nagyobb. Megjegyzendő, hogy módszerünk egyik korlátja, hogy sosem helyezhetünk el határt a szekvencia harmadik színét megelőzve (mert a következési valószínűség az előző elempárra, míg az arányok mindig két egymás utáni következési valószínűségre épülnek).

Ezt az eljárást követve minden generáció esetében létrehozhatunk egy szegmentált részekből álló leltárat (lényegében egy lexikont). A 4. ábra bemutatja, hogyan csökken a szegmentált egységek hossza generációnként, azaz, hogy a szekvenciák egyre kisebb részekből tevődnek össze. Egy lineáris vegyes hatású modellt alkalmaztunk az egység-hossz alakulásának előrejelzésére. A modell tartalmazza a generáció hatását mint rögzített hatást,⁴⁵ valamint egy véletlen interceptet a láncokra,⁴⁶ és azt is megengedi, hogy a generáció hatása lánconként eltérő legyen (véletlen meredekség lánconként). Ez alapján szignifikáns generációnkénti csökkenés mutatható ki az egység-hosszokban ($\beta = -0.17$, $sd = 0.044$, $p = 0.012$). Ez még akkor is fennáll, ha nem vesszük számításba a kezdeti véletlenszerű szekvenciasorozatot ($\beta = -0.11$, $sd = 0.040$, $p = 0.041$).



4. ábra. Generációnkénti egység-hossz-változás (azaz az eljárásunk által szegmentált részek hosszának változása). A függőleges koordináták az összes egység átlagos hosszát jelölik egy meghatározott generációban, az összes láncolatra átlagolva. Az azonos hosszúságú egységeket többször is számítjuk az átlagolás során. A hibásávok 95%-os konfidenciaintervallumokat jelölnek, amelyeket bootstrap módszerrel számoltunk ki. Az egységek rövidebbek lesznek, ami generációnként egyre kisebb részek kiválasztását mutatja a nagyobb egészről.

⁴⁵ A modell azt vizsgálja, hogy általánosságban a generációk száma hogyan befolyásolja az egységek hosszát – *a szerk.*

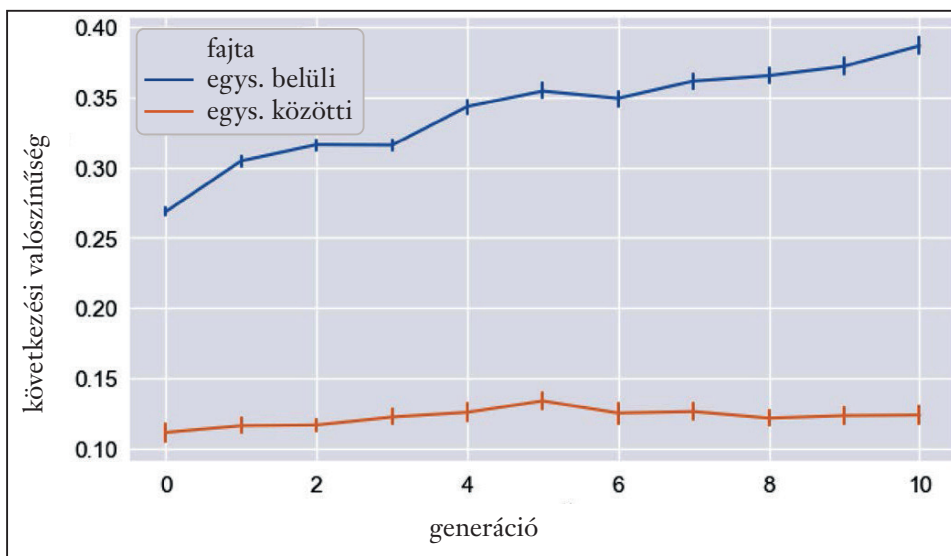
⁴⁶ Minden egyes lánc eltérő hosszúságú kiindulóegységekből állhat – *a szerk.*

Fontos továbbá, hogy az egységeken belüli következési valószínűségek generációnként nőnek, míg az egységek közöttiek kifejezetten alacsonyak maradnak. Ez azt jelenti, hogy a szekvenciahalmazok a kulturális evolúció során formálódnak, egyre egyértelműbb szegmentációs jelzéseket létrehozva (vö. 5. ábra). A következőképpen számoltuk ki ezt a két értéket: miután az eljárásunk szerint mindegyik halmazban szegmentáltuk a szekvenciákat, kategorizáltuk a következési valószínűségeket egységeken belüliekre és egységek közöttiekre. Például, ha szegmentáltuk a „PSZP”-t „PSZ–P”-re, akkor a „Z” következési valószínűsége „PS” után egyazon egységen belülinek számít, míg a „P” következési valószínűsége „SZ” után egységek közöttinek. Mivel a szekvenciákat a következési valószínűség csökkenésének függvényében szegmentáljuk, az egységek közötti követkeзések szükségszerűen kis valószínűségűek. Ezért fontosabb az, hogy az egységeken belüli következési valószínűségek viszont egyértelmű növekedést mutatnak, ami a szegmentált egységek statisztikai kohéziójának növekedését jelzi.

Ez esetben is egy lineáris vegyes hatású modellt alkalmaztunk a következési valószínűségek alakulásának becslésére. A modell rögzített hatásként tartalmazza a generációt, a következési valószínűség típusát (egységeken belül vagy között), valamint ezek kölcsönhatását. Emellett tartalmaz egy véletlen interceptet a láncokra,⁴⁷ és lehetővé teszi, hogy a generáció hatása lánconként eltérő legyen (véletlen meredekség). Az eredmény szignifikáns kölcsönhatást mutat a következési valószínűség típusa és a generációk között, azaz a generációnkénti valószínűségnövekedés nagyobb az egységeken belül, mint azok között. ($\beta = 0.010$, $sd = 0.00070$, $p < 0.001$).

Mindezek alapján azt mondhatjuk, hogy pusztán a szekvenciák generációnkénti átadásával megjelenik egy rendszerszintű statisztikai rendszer – annak ellenére, hogy a feladat nem egy struktúra megtanulása volt. A rendszer egyre erősebb szegmentációs jeleket fejleszt ki, ami az egységen belüli és az egységek közötti átmeneti valószínűségek növekvő különbségében mutatkozik meg. Más szavakkal, a részek az egészekből emelkednek ki.

⁴⁷ A láncok kiindulóegységeiben eltérők lehetnek a következési valószínűségek – *a szerk.*



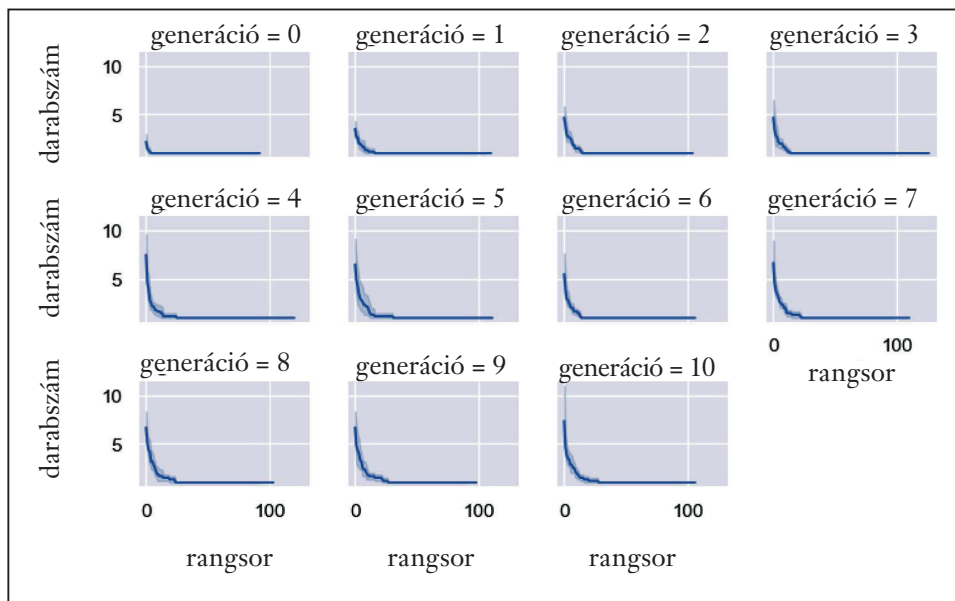
5. ábra. A következési valószínűség generációnkénti változása egységeken belül és egységek között. A hibasávok 95%-os konfidencia-intervallumokat jelölnek, amelyeket bootstrap módszerrel számoltunk ki. A szegmentációs módszerünk alkalmazása után mindegyik következési valószínűséget osztályozhatjuk aszerint, hogy egy egységen belül vagy két egység között helyezkedik el. A grafikon minden pontja azt az átlagot mutatja (egy generáció mindegyik láncolatára kiterjesztve), amely az egységen belüli és egységek közötti következési valószínűségekből lett kiszámítva.

A határok megállapításának módja miatt szükségszerűen kisebbek lesznek az egységek közötti következési valószínűségek, még a nulladik generációban is. Itt inkább az a lényeg, hogy az egységeken belüli következési valószínűségek növekednek.

A STATISZTIKAI STRUKTÚRA KIALAKULÁSA

Most, hogy kísérletünknek köszönhetően rendelkezésünkre áll minden generáció és láncolat szegmentált egységeinek leltára, feltehetjük a kérdést: hogyan változik az egységek gyakorisági eloszlása, amikor a szekvencialalmazokat ismétléses tanulás útján továbbítják? A kérdés megválaszolására a leltár összes egységének gyakoriságát kiszámítjuk minden generációra és láncolatra, majd a gyakoriságot az a gyakorisági rangsorban elfoglalt hely függvényében ábrázoljuk (vö. 6. ábra). Ez azt mutatja, hogy egy viszonylag konstans kezdeti állapothoz képest (majdnem az összes elem csak egyszer jelenik meg a kezdeti, véletlenszerű halmazban), egyre inkább ferde eloszlás alakul ki.

A leggyakoribb egységek 5–10-szer többször jelennek meg, mint a legkevésbé gyakoriak. Ferde eloszlás jön létre tehát pusztán a szekvenciák generációs átadásának eredményeképpen, az ismétléses tanulás során.

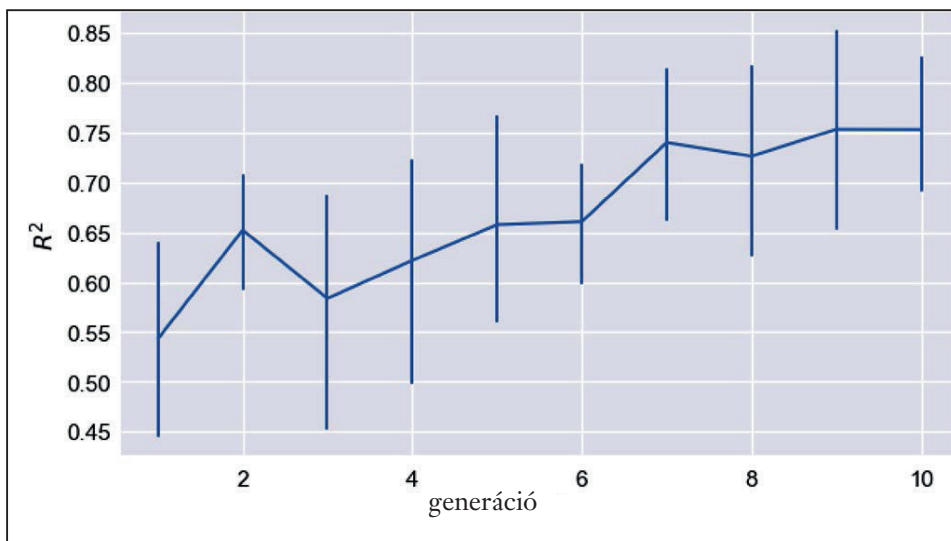


6. ábra. Az elemek eloszlása generációnként a szekvenciahalmazokban. A gyakoriságokat a gyakorisági rangsorban betöltött hely viszonyában ábrázoltuk. A láncolatok bootstrap módszerrel kiszámolt 95%-os konfidenciaintervallumai szürke mezőként láthatók az átlagok körül. A kezdeti viszonylag egységes eloszlásból egy ferde eloszlás alakul ki, kevés magas gyakoriságú és sok alacsony gyakoriságú elemmel.

Feltehetjük a kérdést, hogy az eloszlás mennyire áll közel a Zipf-féléhez, és hogy hogyan változik generációról generációra. Ehhez meg kell néznünk a gyakoriság és a rangsorban elfoglalt hely logaritmusának R^2 -ét.⁴⁸ Minden generáció minden láncolatához kiszámoltuk az R^2 -et. Látható, hogy a Zipf-eloszláshoz való illeszkedés generációnként nő (vö. 7. ábra). Ezt statisztikailag is alátámasztja egy lineáris vegyes hatású modell, amely az R^2 értéket jósolja

⁴⁸ Ha a gyakoriságokat és a rangsorban elfoglalt helyeket logaritmikus skálán ábrázoljuk, akkor a Zipf-eloszlás egy egyenes vonalat ad ki. Ezért egy egyenes vonalat illesztve az adatokra (lineáris regresszióval), kiszámítható, hogy az adatok mennyire szorosan követik ezt az egyenest. Ez az, amit az R^2 mutat meg: ha közel van az 1-hez, akkor az adatok szorosan követik a Zipf-féle mintázatot – *a szerk.*

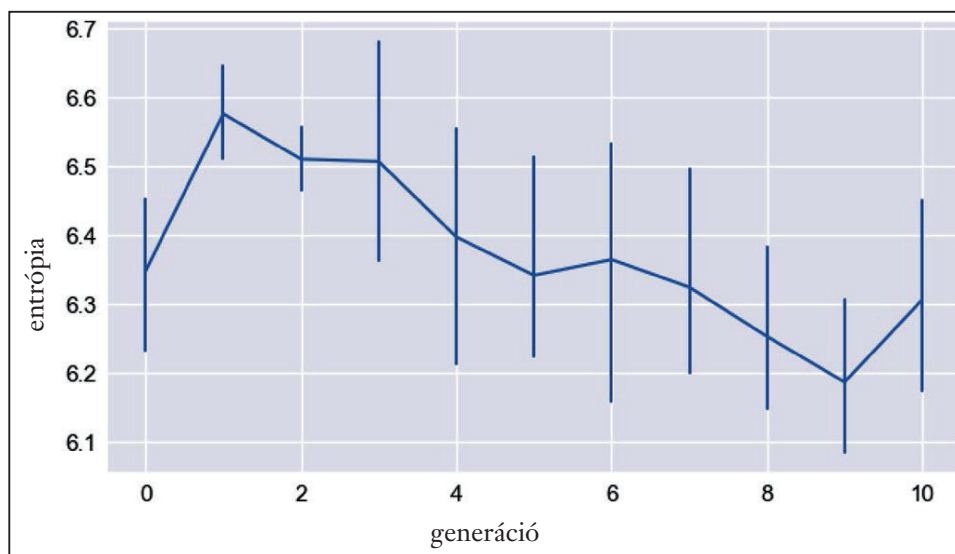
meg a generációk száma alapján. A modell tartalmaz egy rögzített hatást a generációra, valamint figyelembe veszi, hogy a láncok kiinduló szintjei különbözhetnek egymástól, és a generáció hatása is eltérhet lánconként. Ezek alapján az R^2 szignifikáns növekedését tapasztaltuk a generációk viszonylatában ($\beta = 0.022$, $sd = 0.0066$, $p = 0.020$).



7. ábra Az egységek logaritmizált gyakorisága és a logaritmizált rangsora közötti korreláció R^2 értékei (lásd a 6. ábrát). Az ábrán ezeket mutatjuk a generációk függvényében (a 0. generációt, azaz a kezdeti, véletlenszerű nyelvet kihagytuk, mivel több láncolat esetében az eloszlás teljesen lapos volt, így a korrelációs együttható nem számítható ki értelmesen). A hibasávok 95%-os konfidenciaintervallumokat jelölnek, amelyeket bootstrap módszerrel számoltunk ki. Idővel a szóeloszlások egyre inkább követik a Zipf-eloszlást. Ezt alátámasztja a növekvő lineáris kapcsolat a logaritmikus gyakoriság és a logaritmikus rangsor között.

Még jobban tudjuk számszerűsíteni a ferde eloszlás mértékét, ha a szekvencialmazok egységeinek entrópiáját számítjuk ki. Egyenletes eloszlás esetében ugyanis nagyobb az entrópia, mint a ferde eloszlásnál. Az entrópia generációnkénti csökkenését mutatja, még akkor is, ha figyelembe vesszük, hogy hány különböző egységet azonosított a szegmentálási eljárásunk. Statisztikailag ezt újfent egy lineáris vegyes hatású modellel teszteltük, amely az entrópiát jósolja a generációk és az eltérő egységek száma alapján. A modell

tartalmaz rögzített hatásokat a generációra és az egységek számára vonatkozóan, valamint számol a különböző láncok kiinduló szintjei közötti véletlenszerű eltérésekkel. Az egységek számát azért vontuk be, hogy biztosítsuk: a generáció hatása nem csak azért jelentkezik, mert kevesebb egység alacsonyabb entrópiát eredményez. (Egy bonyolultabb modell – amely lehetővé tenné, hogy a generáció hatása lánconként eltérjen – nem volt stabil, ezért egyszerűsített modellt használtunk.) Az eredmények szerint generációként szignifikánsan csökken az entrópia ($\beta = -0.033$, $sd = 0.0055$, $p < 0.0001$). Emellett, ahogyan várható volt, az egységek számának növekedésével az entrópia is nő ($\beta = 0.0028$, $sd = 0.00084$, $p = 0.0013$).



8. ábra. Az egységek entrópiája generációként ábrázolva. Az entrópiát mindegyik generációra és mindegyik láncolatra kiszámoltuk, majd az értékeket átlagoltuk.

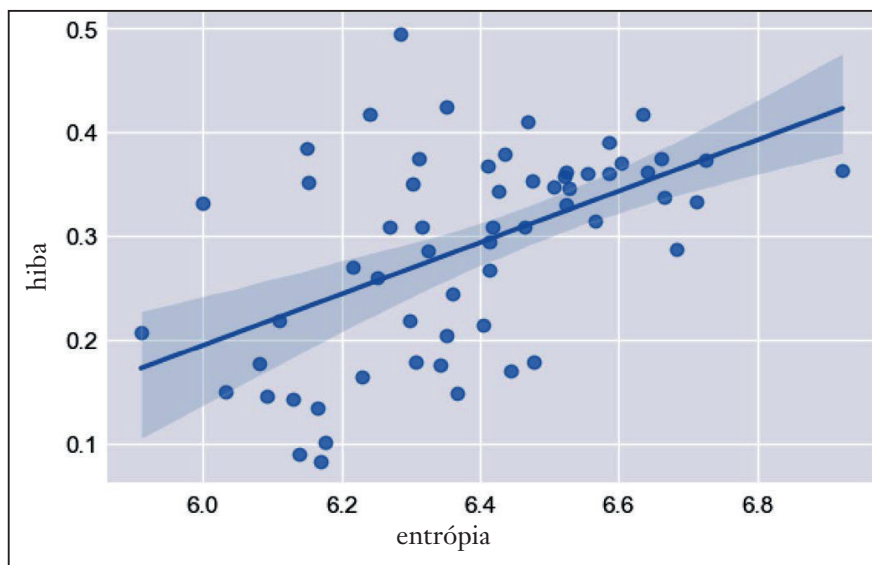
A hibásávok 95%-os konfidenciaintervallumokat jelölnek, amelyeket bootstrap módszerrel számoltunk ki. Megfigyelhető egy kezdeti entrópiánövekedés a halmaz méretének növekedése miatt (mivel a kezdeti véletlenszerű szekvenciahalmaznak nagyon kevés szegmentálható egysége van), de az egyre ferdébb eloszlásnak köszönhetően az egységek entrópiája is csökken generációról generációra.

A TANULHATÓSÁG MINT OK ÉS OKOZAT

Mindent összevetve az eredmények azt mutatják, hogy az ismétléses tanulás hozza létre a nyelv két szisztematikus sajátosságát: a részek elkülönülését és azok ferde eloszlását.

Bemutattuk, hogy ezeket a sajátosságokat az átadások eredményezik. De vajon be lehet-e bizonyítani, hogy a kialakuló statisztikai struktúra ténylegesen hozzájárul a rendszer tanulhatóságához? A 9. ábra azt mutatja, hogy milyen kapcsolat tételezhető a résztvevő által látott szókészlet entrópiája és az adott résztvevő által produkált hibaarány között. Eszerint világos kapcsolat áll fenn a statisztikai struktúra és a tanulhatóság között; a Person-féle korreláció értelmében a hiba és az entrópia pozitívan korrelál ($r = 0.52$, $p < 0.001$).

Az ismétléses tanulás során létre jövő statisztikai struktúra lehetővé teszi, hogy a megképződő rendszer egyre inkább tanulható legyen. A részek és azok ferde eloszlása tehát egyaránt oka a tanulhatóság növekedésének, és okozata a generációkon átívelő ismétléses tanulásnak.



9. ábra A hiba és az entrópia kapcsolata. Minden pont egy-egy résztvevőt jelöl; illetve ábrázoltuk az ezekre legjobban illeszkedő egyenest és a 95%-os konfidencia-intervallumot is. A kevésbé ferde eloszlású egységeket (azaz a nagyobb entrópiával rendelkező szekvenciahalmazokat) nehezebb megtanulni (azaz nagyobb hibaszámmal másolják le azokat).

AZ EREDMÉNYEK ÉRTELMEZÉSE

A nyelv kifejezőerejét – fajunk egyik legkülönlegesebb jellemvonását – kis részek nagyobb egészekké való kombinációja teszi lehetővé. A természet más kommunikációs rendszereihez képest a nyelv kombinálható részei azonban nem eleve adottak, hanem a gyermekek hallás után sajátítják el: ahhoz, hogy felfedezzék a részeket, el kell valahogyan különíteniük azokat egy tagolatlan beszédflowban. A részek nem egyforma eloszlásúak, hanem egy jellegzetes hatványtörvény szerinti eloszlást követnek kevés gyakori és sok ritka szóval (a gyakoriságok nem lineárisan csökkennek).

A tanulmány bemutatta, hogy a nyelv mindkét sajátossága – az, hogy részekből áll, amelyek meghatározott eloszlásúak – olyan kulturális átadás során is kialakulhat, amikor a tanulók egyszerűen sorozatok egészét reprodukálják. Egy nemnyelvi feladat segítségével kimutattuk, hogy a színszekvenciák az idő elteltével strukturáltabbá és tanulhatóbbá válnak. A strukturáltság és tanulhatóság növekedése a Zipf-eloszláshoz egyre jobban illeszkedő nyelvi részek kialakulásából adódik. Az elkülönülő részek meghatározása a színek közötti következási valószínűségek felhasználásával történt – ugyanazzal a statisztikai mutatóval, amellyel a csecsemők is felismerik a szóhatárokat. Az eredmények szerint generációnként nőnek az egységeken belüli következási valószínűségek; az eloszlásuk entrópiája időről időre csökken, és egyre inkább a Zipf-eloszlást követik; illetve az egységek entrópiája a hibaértékkel korrelál. Azt állítjuk, hogy a nyelv pontosan azért áll statisztikailag elkülöníthető részekből, amelyek egyre ferdebb eloszlásúak, mert a nyelvhasználók több generáción keresztül tanulják meg azt. A nyelv statisztikai struktúrája egyszerre oka és okozata a nagyobb tanulhatóságnak.

Hogyan jelenik meg mindez kutatásunkban? Eredményeink nem közvetlenül erre vonatkoznak, de felvázolnak egy lehetőséget. Kísérleti feladatunkban a résztvevők olyan szekvenciákat másolnak le, amelyeknek eleinte teljesen egynemű statisztikai struktúrája van: bármely szín követhet bármely más színt. A mintázatok másolásakor a résztvevők mégis bizonyos színeket gyakrabban párosítanak össze, mint másokat. Ez kezdetben pusztán hiba. Viszont emiatt a tanulók következő generációja számára a szekvenciák már statisztikai szempontból nem egyneműek. Így alakulnak ki a következási valószínűség mintái, amelyekben az alacsony valószínűségeket olykor nagyok követik. A tanulók ezt az esetet a szekvencia alrészei közötti határvonal jelzésének tekintik. Ily módon jön létre a lehetséges egységek leltára, amelynek száma idővel növekedhet, egyre nagyobb különbségeket hozva létre az egységek közötti és belüli következási valószínűségekben. Ez a gyakorisági előny generációról

generációra erősödik, létrehozva a nyelvtanulást elősegítő ferde eloszlást. Ebben a folyamatban „a gazdagabb még gazdagabb lesz” (*rich-gets-richer*), és az így kialakuló ferde eloszlás a következési valószínűségek hasznosíthatóságát növeli a határkijelölésben, ami az ismétlődő egységek még gyakoribb ismétléséhez, egyre ferdebb eloszláshoz, illetve pontosabb másolásához vezet. A folyamat ismétlődik és erősödik, az idő elteltével növekszik a ferde eloszlás mértéke, valamint az eloszlás korrelációja a nyelvtanulás sikerességével. Így tehát, még ha a nyelvtanulók olyan feladattal is állnak szemben, amelyben teljes szekvenciákat kell lemásolniuk (anélkül, hogy számítanának arra, hogy a szekvenciák alrészekből állnak, amelyek ráadásul a különböző szekvenciákban visszatérnek), egy idő után elkezdenek kialakulni elkülöníthető részek. Ezek a részek később újrendezhetőek, hogy a generációkon átívelő továbbadás szempontjából optimalizált szekvenciahalmaz jöjjön létre.

Az egésztől a részekig haladó stratégia látszólag kontrainuitív ahhoz a verzióhoz képest, amikor először különálló egységeket tanulnak meg a nyelvhasználók, és ezeket kombinálják újra. Maga a környezet indokolja azonban e stratégia létezését: a részeket ugyanis nem szegmentált egységként ismerjük meg. A nyelv sokféle ehhez hasonló példával szolgál: a fonémák a teljes szóalakból vezethetőek le,⁴⁹ a morféma többmorfémás szavakból,⁵⁰ a szavak pedig többszavas egységekből.⁵¹ Kutatásunk fényében a részek létezésének másik oka, hogy az egésztől a részek felé haladó stratégia valóban több szempontból megkönnyíti a nyelvtanulást, és elősegíti a részek gyakran önkényes viszonyainak megtanulását. A névelő-főnév illeszkedése például jobban tanulható, ha először egyetlen egységekkel kezdődik a tanulás, amelyeket csak aztán bontanak elemeire a nyelvhasználók, és nem külön-külön a névelőt és a főnevet tanulják meg a későbbi kombinálás érdekében.⁵² Az egésztől a részek felé haladó tanulás az egyén fejlődésének más aspektusaiban is segíthet. A csecsemők kezdeti gyenge látásélessége például, amely megakadályozza

⁴⁹ Shelley L. VELLEMAN és Marilyn M. VIHMAN, „Whole-word phonology and templates”, *Language and Speech* 32 (1989): 149–170, [https://doi.org/10.1044/0161-1461\(2002/002\)](https://doi.org/10.1044/0161-1461(2002/002)).

⁵⁰ Dorit RAVID és Alon MALENKY, „Awareness of linear and nonlinear morphology in Hebrew: A developmental study”, *First Language* 21, 61. sz. (2001): 25–56, <https://doi.org/10.1177/014272370102106102>.

⁵¹ Isaac ARNON, „The Starting Big approach to language learning”, *Journal of Child Language* 48, 5. sz. (2021): 937–958, <https://doi.org/10.1017/S0305000921000386>.

⁵² ARNON és RAMSCAR, „Granularity and the acquisition of grammatical gender...”, 292–305; SIEGELMAN és ARNON, „The advantage of starting big...”, 60–75.

őket a részletek észlelésében, a jelek szerint az átfogó térbeli elemzést támogatja, ami többek között az arcfeldolgozás szempontjából is kedvező.⁵³

Tekintve, hogy kísérletünkbe tudatosan nemnyelvi feladatot építettünk be, a memóriajátéknak nincs közvetlen köze a nyelvhez. A színszekvencia nem hasonlít semmilyen beszélt vagy írott nyelvhez, és a feladat nem úgy volt keretezve, hogy a résztvevők ehhez hasonló mintázatokat keressenek vagy ismerjenek fel. Továbbá sem a teljes szekvenciáknak, sem a kialakuló részeknek nincsen jelentésük. Ennek ellenére rendszerszintű statisztikai struktúra jön létre, amely látványosan hasonlít a nyelv szerkezetére. Ez arra utal, hogy a nyelvben lévő rendszer is ugyanígy, a kulturális evolúció és az ismétléses tanulás általános folyamatán keresztül jött létre, semmint a nyelvspecifikus tanulási preferenciák által. Várhatóan mindenütt ferde eloszlású szekvenciákat találunk, ha ezek szintén a kulturális átadás révén alakulnak, és ha ezeket a tanulók először egészként sajátítják el, majd ezután fedezik fel bennük a részeket. Például a zene is ilyen kulturálisan átadott mintázat, ezért itt is a Zipf-eloszláshoz való illeszkedést feltételezhetjük. Erre van is bizonyíték: a zenei elemek – többek között a hangmagasság, a dallamfutamok és a konszonancia – pontosan követik a Zipf-eloszlást.⁵⁴ Az ének is kulturálisan átadott mintázat, és szintén kitett az egésztől a részek felé haladó tanulásnak, hiszen a dalokat gyakran a maguk teljességében tanuljuk és értjük meg. Itt is igaz, hogy a világ különböző kultúráiban a ritmus- és dallamelemek eloszlása a hatványtörvényt követi.⁵⁵

Messzebbre tekintve feltehetjük a kérdést, hogy más fajok kommunikációs rendszerében is találhatók-e hasonló ferde eloszlások. Ha a tanulhatóság egyszerre oka és okozata a ferde eloszlásnak, ez várhatóan kialakulhat más kulturálisan átadott rendszerek esetében is, főként ha azok rész-egész viszonyokból állnak. Gondolhatunk néhány olyan faj kommunikációs módjára, amelyek rendelkeznek ezekkel a sajátosságokkal, ilyenek például a delfinek, a bálnák vagy az énekesmadarak. Érdekes módon a felnőtt palackorrú delfi-

⁵³ Lina VOGELSANG, Jürgen SCHMIDHUBER, Jonathan C. FLACK, Michael P. TKACH és Sarah B. HRDY, „Potential downside of high initial visual acuity”, *Proceedings of the National Academy of Sciences* 115, 44. sz. (2018): 11333–11338, <https://doi.org/10.1073/pnas.1800901115>.

⁵⁴ Bill MANARIS, Chaoyi ROMERO, Panayotis P. PANAYIDES, Dan McNAB és Joan M. KISER, „Zipf’s law, music classification, and aesthetics”, *Computer Music Journal* 29, 1. sz. (2005): 55–69, <https://doi.org/10.1162/comj.2005.29.1.55>.

⁵⁵ Samuel A. MEHR, Manvir SINGH, Dean KNOX, Daniel T. BOWLING, Cody M. MOSER és Patrick E. SAVAGE, „Universality and diversity in human song”, *Science* 366, 6468. sz. (2019): eaax0868, <https://doi.org/10.1126/science.aax0868>.

nek füttyrepertoárja követni látszik a hatványtörvényt,⁵⁶ ahogyan a törpeposzáta (egyfajta énekesmadár) dalrepertoárjának szótagjai is.⁵⁷ Ugyanígy a hosszúszárnýú bálna énekének elemeiről is kimutatták, hogy a nagyban követik a ferde eloszlást.⁵⁸ Ezek az eredmények arra utalnak, hogy a nem emberi fajok esetében is az emberi kommunikációhoz hasonló folyamatok mennek végbe. Ebből kiindulva további kutatások pontosabban feltárhatják majd a ferde eloszlás nyelvtanulást elősegítő hatását a nem emberi fajok esetében is.

Kutatásunk több különböző hagyományt fog össze a kulturális evolúció, a statisztikai tanulás, a fejlődépszichológia és a nyelvészet területeiről. Bemutattuk, hogyan lehet ezeket együttesen alkalmazni egy olyan kísérlet létrehozásához, amely hozzájárul az emberi nyelv alapvető sajátosságainak és ezek eredetének megértéséhez. Ez nem jöhetett volna létre a különböző nézőpontok vegyítése nélkül – ilyen az ismétléses tanulás a kulturális evolúció területéről vagy a következési valószínűség elemzése a statisztikai tanulás területéről. Úgy gondoljuk, hogy mindez bizonyítja az interdiszciplinaritás szükségességét a nyelv eredetének és evolúciójának empirikusan megalapozott elméleti kutatásában.

ADATBÁZIS ELÉRHETŐSÉGE

A teljes adathalmaz, illetve az ábrák létrehozásához és statisztikai elemzéséhez szükséges kódok elérhetők az alábbi repozitóriumban: https://github.com/smkirby/cultural_evolution_of_statistical_structure.

⁵⁶ Brenda McCOWAN, Sarah F. HANSER és Laurel R. DOYLE, „Quantitative tools for comparing animal communication systems: Information theory applied to bottlenose dolphin whistle repertoires”, *Animal Behaviour* 57, 2. sz. (1999): 409–419, <https://doi.org/10.1006/anbe.1998.1000>; Ryo SUZUKI, James R. BUCK és Peter L. TYACK, „The use of Zipf’s law in animal communication analysis”, *Animal Behaviour* 69, 1. sz. (2005): F9–F17, <https://doi.org/10.1016/j.anbehav.2004.08.004>.

⁵⁷ Antonio M. PALMERO, Javier ESPELOSÍN, Paolo LAIOLO és Juan C. ILLERA, „Information theory reveals that individual birds do not alter song complexity when varying song length”, *Animal Behaviour* 87 (2014): 153–163, <https://doi.org/10.1016/j.anbehav.2013.10.026>.

⁵⁸ Jessica A. ALLEN, Eleanor C. GARLAND, Rochelle A. DUNLOP és Michael J. NOAD, „Network analysis reveals underlying syntactic features in a vocally learnt mammalian display, humpback whale song”, *Proceedings of the Royal Society B* 286, 1917. sz. (2019): 20192014, <https://doi.org/10.1098/rspb.2019.2014>.

A fordítás az alábbi szöveg alapján készült: IbnaL ARNON és Simon KIRBY.
„Cultural evolution creates the statistical structure of language”. *Scientific Reports* 14 (2024): 5255 (article number).

Fordította: *László István*

Az írás fokozatos evolúciója

Vajon az írás fejlődés útján jött létre?¹

Vajon az írás fejlődés útján jött létre? Elcsépett, üres kérdésnek tűnhet; az írásnak nyilván vannak előfutárai, nem egyik napról a másikra jelent meg. Ugyanakkor, ha azt kérdezzük, hogy a grafikus szimbólumokból milyen fokozatokon keresztül alakult ki a nyelv hangalapú lejegyzése, már nem kapunk egyértelmű választ. Az írás eredetéről szóló könyvek egyik legtekintélyesebb darabja például egy tanulmánykötet, amely két szerző két, egymásnak ellentmondó válaszát is tartalmazza.² Egyfelől, Stephen Houston azt állítja, hogy az írás evolúciója archeológiai léptékben hirtelen, legfeljebb egy generáción belül következett be.³ Nem ő az egyetlen, aki ezen a véleményen van; maga is a kérdéskör egy másik fontos munkájára, DeFrancis könyvére hivatkozik.⁴ Másfelől, a kötet Bruce Trigger tanulmányát is közli, amelyben a szerző éppen amellet foglal állást, hogy az írás csak lassú folyamat eredményeként, fokozatosan válhatott képessé a nyelv jelölésére.⁵ Úgy tűnik, hogy a kiadvány óta eltelt közel húsz év sem tudta felszámolni ezt az ellentmondást.

A tudományos közegnek számos – jó és rossz – oka van arra, hogy óvakodjon az írás evolúcióelméleti megközelítésétől. Jó indok például annak az evolucionizmusnak az elutasítása, amely áthatotta az írásról szóló korai értekezéseket, például Diringer vagy Gelb munkáit.⁶ Az evolucionizmus értelmében – ahogyan e két szerző vallotta – a civilizáció egy lineáris fejlődés

¹ Köszönettel tartozom Silvia Ferrarának és Fabio Tamburininek, hogy felvetették ezt a különleges kérdést. A kutatómunkát a „Frontiers in Cognition” pályázat (ANR-17-EURE-0017 EUR) és a Max Planck Társaság támogatta.

² Stephen D. HOUSTON, szerk., *The First Writing: Script Invention as History and Process* (New York: Cambridge University Press, 2004).

³ Stephen D. HOUSTON, „Overture to *The First Writing*”, in HOUSTON, szerk., *The First Writing*..., 5–6.

⁴ John DEFRANCIS, *Visible Speech: The Diverse Oneness of Writing Systems* (Honolulu: University of Hawaii Press, 1989), 74, <https://doi.org/10.1515/9780824841621>.

⁵ Bruce G. TRIGGER, „Writing Systems: A Case Study in Cultural Evolution”, in HOUSTON, *The First Writing*..., 39–71.

⁶ David DIRINGER, *The Alphabet: A Key to the History of Mankind* (London: Philosophical Library, 1953) és Ignace J. GELB, *A Study of Writing* (Chicago: University of Chicago Press, 1963).

mentén jutott el a képirásos kommunikációtól a hieroglifáig, majd az írástudás csúcsáig, az alfabetikus írásig. Ez a fajta evolucionizmus közelebb áll Spencerhez, mint Darwinhoz; és az elméletet jogosan is bírálták amiatt, hogy önkényesen tart értékesebbnek bizonyos írásmódokat másokkal szemben, és túl nagy jelentőséget tulajdonít a kulturális változás folyamatos és egyirányú jellegének.

Egy másik ellenérv, hogy az írást nem lehet csak félig feltalálni. E nézet az írást általában hangjelölésként⁷ határozza meg, azaz a nyelv hangalapú lejegyzéseként, vagy olyan kódként, amely a betűket a lejegyzett nyelv hangjaihoz társítja. Innen nézve az ideografikus vagy piktografikus írásrendszerek, melyek szavak helyett direkt módon képeznek le gondolati vagy tárgyi tartalmakat, nem számítnak igazi írásnak.⁸ Ha ezt elfogadjuk, akkor az írás feltalálása egybeesik a fonematikus elv felfedezésével: azzal a gondolattal, hogy pusztán a nyelv hangjainak lejegyzése bármit leírhatóvá tesz, ami kimondható. Másképp megfogalmazva, a fonematikus elv felfedezése a legkülönbélebb tartalmak kódolását teszi lehetővé grafikus szimbólumok segítségével. Amióta az emberek rájöttek, hogy a képek hangokat közvetíthetnek az akrofónia vagy a rébusz-elv⁹ felhasználásával, elméletileg képessé váltak a fonematikus elv alkalmazására is bármely verbális üzenet esetében – még ha nem is minden esetben használták ezt a képességet.

Ez az érvelés ugyanakkor figyelmen kívül hagyja a mai értelemben vett írás másik két fontos jellemzőjét. Ezeket a fonematikus írás szigorú meghatározása nem foglalja magában; ami azonban nem jelenti azt, hogy eltekinthetnénk tőlük.

Az első tulajdonság az általánosság. A fonematikus írás, ahogyan ma használjuk, széles körben és sokféle célra alkalmazható. Bizonyos dolgokat az írásrendszerek segítségével nehezebben tudunk jelölni, ezért inkább más grafikus kódokhoz folyamodunk – emodzsikhoz, kottaíráshoz, repülőtéri jelekhez stb.

⁷ Ebben a tanulmányban a „fonematikus” és a „hangjelölő” melléneveket felcserélhetőnek tekintem. Ezáltal átsiklok egy finom különbségen a fonematikusság (a nyelv lejegyzése) és a hangjelölés (a hangok lejegyzése) között. Máshol már részletesen kifejtettem e különbséget (Olivier MORIN, „The Puzzle of Ideography”, *Behavioral and Brain Sciences* (2022): 1–69, <https://doi.org/10.31234/osf.io/7pd5m>), de jelen gondolatmenetből talán kihagyható nagyobb veszteség nélkül.

⁸ E gondolat ismertetéséhez vö. MORIN, „The Puzzle of Ideography”, 1–69.

⁹ Akrofónia: amikor egy szó első hangjának vagy mássalhangzójának jelévé maga a szó válik, például „A mint Aladár”. Rébusz-elv: az írás egyik módszere, amelyben a szavak vagy szótagok hangalakját más, azonos hangzású, de eltérő jelentésű szavak képi ábrázolásával jeleltik meg. Ezt a módszert különösen a szótagírás kialakulásának kezdeti szakaszában használták, amikor egy-egy képi jel nem a jelentésére, hanem a kiejtésére utalt – *a szerk.*

Ám a mai írástudó társadalmakban a grafikus kommunikáció (jelek konzisztens használata információ kódolására) túlnyomó többsége fonematikus írással történik. Azonban a fonematikus írás nem volt mindig ennyire általános. Persze lehet azt állítani, hogy a fonematikus írás mindig is magában hordozta a lehetőséget az általános használatra – *in nuce*: a hangjelölés általános célú kód. Ugyanakkor ez a potenciál kiaknázatlan maradhat: a tanulmány második fejezetében bizonyítékokat fogok felmutatni arra, hogy miután feltalálták a fonematikus írást, évszázadokon át majdnem kizárólag csupán a tulajdonnevek leírásának eszköze maradt. Ennek kapcsán röviden majd azon is eltűnődöm, vajon miért kapcsolódik össze a hangjelölés és a tulajdonnevek kódolása, bár jelezni fogom, hogy magyarázatom hiányos.

A második tulajdonság az aszinkronitás: a képesség új információ nem egyidejű átadására írott üzenet formájában valakinek, aki még nem ismerte azt. Fontos kiemelni, hogy létezik olyan íráshasználat, amely nélkülözi az aszinkronitást, például a mnemotechnikai használat, amikor a feljegyzés emlékeztetőül funkcionál olyan információ felidézésére, amelyet valaki már részlegesen megjegyzett. Azokból az üzenetekből is hiányzik az aszinkronia, amelyeknek nincs értelme a jelentésüket már ismerők szóbeli magyarázata nélkül. A harmadik fejezet azokat az érveket ismerteti, amelyek szerint a tisztán aszinkron kommunikáció egy késői és viszonylag ritka fejlemény a fonematikus írás történetében. Közvetlen bizonyítékokkal szolgálnak azok a szövegek, amelyek expliciten szólnak arról, hogy az írásbeliség elsősorban szövegek felolvasását jelentette. Közvetett bizonyíték két körülmény tanulmányozásából vonható le, amelyek az írás aszinkron használatának lehetőségét biztosítják. Az aszinkron kommunikációhoz ugyanis önmagában értelmezhető (kontextusból kiragadva is megérthető), könnyen olvasható és visszakereshető szövegek szükségesek. A harmadik fejezetben emellett érvelek, hogy az ókori szövegeket általában nehéz volt aszinkron módon használni, mivel a tartalmuk nem volt elég explicit; a tárolásuk és archiválásuk pedig meglehetősen esetleges volt.

A fenti állítások nyomán az írás olyan evolúciója rajzolódik ki, amelyben az írásbeliség fő jellemzői nem egyidőben jöttek létre. A hangjelölés az első lépés – ám az a lehetőség évszázadokig kiaknázatlan maradt, hogy általános kommunikációs eszköztár szerveződjön a hangjelölés köré. Az írás használata az aszinkron kommunikációban pedig még ennél is később vált ismert gyakorlattá.

1. MIÉRT VOLT KEZDETBEN AZ ÍRÁS PUSZTÁN A TULAJDONNEVEK JELÖLÉSÉNEK ESZKÖZE?

Az első nehézség az írás története kapcsán rögtön a fonematikus jelölés felfedezése utáni időszakra vonatkozik, legalábbis bizonyos helyeken. Néhány írástudó kultúra ugyanis a felfedezést követően csak tulajdonnevek leírására használta a fonematikus lejegyzést; generációkon keresztül, vagy akár öröklé.

1.1. A hangjelölést tulajdonnevek jelölésére találták ki

Szoros kapcsolat van a fonematikus írás és a tulajdonnevek – személy-, hely-, intézmény- vagy nemzetségnevek – jelölése között. Ez a kapcsolat döntő szerepet játszott az írásrendszerek megfejtésében, mivel egy tulajdonnevet több különböző írásrendszer is rögzíthetett, ami a helyes kiolvasás ellenőrzését szolgálhatja. Az összes dokumentált ókori írásrendszer esetében, azaz Egyiptomban, Kínában, a prekolumbiánus Közép-Amerikában, Mezopotámiában és (a másodlagos találmányokat [*secondary inventions*] is beleértve) Anatóliában fonematikusan jelölnek legalább néhány tulajdonnevet attól kezdve, hogy el-sajátították a hangok kódolásának képességét az akrofónia vagy a rébusz-elv segítségével.¹⁰ Miközben ez nem igaz a beszélt nyelv többi területére: az igékre, a határozószókra, a melléknevekre, a kötésszókra stb. Vagyis a fonematikus írásjelek mindig képesek voltak a beszélt nyelv azon részeit kódolni, amelyek személyeket, intézményeket vagy helyeket jelölnek, ám cselekvéseknek, dolgok tulajdonságainak és elhelyezkedésének vagy szintaktikai kapcsolatoknak a kódolása csak opcionálisan és később alakult ki – ha egyáltalán kialakult.

Egyiptom. A hangjelölés az egyiptomi írásban először a tulajdonnevek (személy-, hely- vagy intézménynevek) lejegyzésekor jelenik meg olyan forrásokban, mint az Abüdoszban lévő i. e. 3200 körüli „U-j” jelű sír.¹¹ Csak sokkal később kezdenek el összefüggő mondatokat írni; a III. vagy IV. dinasztia ide-

¹⁰ Miguel VALÉRIO és Silvia FERRARA, „Rebus and Acrophony in Invented Writing”, *Writing Systems Research* 11, 1. sz. (2019): 66–93, <https://doi.org/10.1080/17586801.2020.1724239>.

¹¹ Andréas STAUDER, „The Earliest Egyptian Writing”, in *Visible Language: Inventions of Writing in the Ancient Middle East and Beyond*, szerk. Geoff EMBERLING és Emily TEETER, 137–149 (Chicago: Oriental Museum Publications, 2010); Elise V. MACARTHUR, „The Conception and Development of the Egyptian Writing System”, in EMBERLING és TEETER, *Visible Language...*, 115–137; John BAINES, *Visual and Written Culture in Ancient Egypt* (Oxford: Oxford University Press, 2007), <https://doi.org/10.1093/oso/9780198152507.001.0001>.

jén, azaz i. e. 2600 után.¹² A levelekhez hasonló hosszabb szövegek ennél is később jelennek meg.¹³ Ez az eltolódás még szembetűnőbb, ha figyelembe vesszük, hogy a koradinasztikus tulajdonnevek világos jelentésű, rövid mondatokból vagy főnévi szókapcsolatokból állnak (például: „Hórusz, a harcos”, „A férfi, aki vizet hoz”). Más szóval, az írásrendszer használói már rendelkeztek azokkal az eszközökkel, amelyekkel rögzíthettek volna más nyelvi elemeket is a tulajdonneveken kívül, hosszú ideig mégsem tették.

Mezopotámia. A fonematikus elv felfedezése és tökéletesítése ebben az esetben is személynevek lejegyzésekor kezdődött i. e. 3000 körül,¹⁴ és itt is évszázadokig kiaknázatlan maradt a hangjelölés tulajdonneveken túli használatának lehetősége. Úgy tűnik, hogy például a toldalékokat nem rögzítették mintegy i. e. 2800-ig.¹⁵ Hosszú ideig váratott magára, amíg a teljes mondatokat rögzítő folyamatos írás elkezdett az ékírás helyébe lépni,¹⁶ és még később jelentek meg az olyan hosszú szövegek, mint a királyfeliratok vagy a levelek.¹⁷

A tulajdonnevek és a hangjelölés kapcsolatát a prekolumbiánus *Közép-Amerika* szintén egyértelműen szemlélteti. A kora közép-amerikai írás ritkán hangjelölő, de akkor is csak tulajdonnevek (ranggal ellátott személyek és istenségek nevei) és a naptári napok nevei esetében – már amennyire ezt meg tudjuk állapítani a dekódolhatóság hiányában. Példa erre az első maja feliratok a San Bartolo-lelőhelyen található falfestményeken (i. e. 400–200 körül) vagy az időszámításunk előtti első évezredbeli olmék feljegyzések.¹⁸ Az azték írásrendszer kapcsán nemcsak az mondható el, hogy a fonematikus írás a tulaj-

¹² BAINES, *Visual and Written Culture...*, 39.

¹³ Uo., 100, 128, 143.

¹⁴ GELB, *A Study of Writing*; Denise SCHMANDT-BESSERAT, *When Writing Met Art: From Symbol to Story* (Austin: University of Texas Press, 2007).

¹⁵ J. COOPER, „Babylonian Beginnings: The Origin of the Cuneiform Writing System in Comparative Perspective”, in HOUSTON, *The First Writing...*, 71–99.

¹⁶ Hans J. NISSEN, Peter DAMEROW, Robert K. ENGLUND és Robert K. ENGLUND, *Archaic Bookkeeping: Early Writing and Techniques of Economic Administration in the Ancient Near East* (Chicago: University of Chicago Press, 1993); SCHMANDT-BESSERAT, *When Writing Met Art...*; Massimo MAIOCCHI, „Writing in Early Mesopotamia: The Historical Interplay of Technology, Cognition, and Environment”, in *Beyond the Meme: Development and Structure in Cultural Evolution*, szerk. Alan LOVE és William WIMSATT, *Studies in the Philosophy of Science* 22, 395–424 (Minnesota: University of Minnesota Press, 2019).

¹⁷ COOPER, „Babylonian Beginnings...”, 71–99.

¹⁸ John S. JUSTESON és Peter MATHEWS, „Evolutionary Trends in Mesoamerican Hieroglyphic Writing”, *Visible Language* 24, 1. sz. (1990): 88–132; Joel PALKA, „The Development of Maya Writing”, in EMBERLING és TEETER, *Visible Language...*, 225–231.

donnevek rögzítésével vette kezdetét, hanem hogy a későbbiekben sem használták másra.¹⁹

Kína a legkevésbé világos eset a négy közül, mivel a régészek által feltárt kínai írás egy olyan korszakból származik, amikor a technológia már meglehetősen kifinomult és sokoldalú volt. Mindazonáltal lehetséges, hogy a kínai pecséteken lévő emblémák és a Sang-kori bronzedény-feliratok már fonematikus írásrendszerre utalnak. Ha ez igaz, akkor Kína lenne a negyedik példa arra, hogy a hangjelölés eredete a tulajdonnevek lejegyzésével függ össze.²⁰

Ezek a példák azt mutatják, hogy a hangjelölő írásnak állandó jellemzője a tulajdonnevek fonematikus rögzítése; ez minden hangjelölő írásra igaz, sőt némely esetében kizárólagosan. Megjegyzendő, hogy az az írástípus is teljes értékűnek tekinthető, amelyik a hangjelölést kizárólag a tulajdonnevek rögzítésére használja, mivel elméletileg megvan a képessége, hogy a célnyelv bármely kimondott szavát rögzítse. Sok nyelvben a nevek gyakran átalakított főnévi szókapcsolatokból vagy akár teljes mondatokból állnak (ahogy fentebb az egyiptomi példában is láthattuk). Az azték írásrendszer szerint átírt navatl nevek gyakran például egész mondatokat rejtenek magukban, mint a „Xihuitltemoc”, amelynek fordítása: „Egy meteor szállt le”.²¹ Könnyen lehetett volna tehát rögzíteni hasonló, nem tulajdonnév értékű kifejezéseket is, de ez mégsem történt meg. Az azték írás így annak az általános tendenciának egy szélsőséges példája, miszerint a kezdetleges írásrendszerek funkciója néha évszázadokon át csupán nevek rögzítésére korlátozódott.

1.2. Hogyan kapcsolódnak össze a tulajdonnevek a hangjelöléssel?

Miért van ilyen sajátos kapcsolat a hangjelölés és a tulajdonnevek között? A leggyakoribb válasz a következő: a tulajdonneveket a köznevekhez (mint például „fa”, „víz”, „egy öl búza” stb.) viszonyítva nehéz ikonikusan megjeleníteni, azaz olyan motivált jelekkel, amelyek hasonlítanak a jelöltre. Ez arra készítette az írás feltalálóit, hogy a tulajdonneveket fonematikusan jelöljék. A magyarázat azonban két okból hiányos. Egyrészt a beszélt nyelv sok más részét is nehéz ikonikusan rögzíteni: az elvont kifejezéseket, a kötőszókat, a prepozíciókat, a határozószókat stb. szintén elég körülményes piktogramok-

¹⁹ Gordon WHITTAKER, „Aztec Hieroglyphics: A Name-Based Writing System”, *Language & History* 61, 1–2. sz. (2018): 60–76, <https://doi.org/10.1080/17597536.2018.1441950>.

²⁰ William G. BOLTZ, *The Origin and Early Development of the Chinese Writing System* (New Haven–Conn: Amer Oriental Society, 1993).

²¹ WHITTAKER, „Aztec Hieroglyphics...”, 60–76.

kal ábrázolni – a hangjelölés mégsem ezek rögzítésére lett kitalálva. Másrészt a tulajdonnevek más, egyszerűbb módokon is megjeleníthetők grafikusan, a hangjelölés felhasználása nélkül. Sok példa közül a heraldika a legszembetűnőbb. A heraldikai kódok személyeket, nemzetségeket, klánokat vagy intézményeket (beleértve az országokat, városokat stb.) ábrázolnak absztrakt vagy figuratív piktogramok felhasználásával. A nemzeti zászlók vagy fogyasztási cikkek márkái ismerős példák; történetileg pedig az európai és a japán családi címerek esetei a legjobban dokumentáltak.²² Az ókori városállamok pénzérméin megjelenített városi emblémák is rendelkeztek ezekkel a tulajdonságokkal: gyakran ezek elegendők voltak a kibocsátó város azonosítására bármilyen más jelzés nélkül.²³ Az írás kialakulásához közelebb álló mezopotámiai és egyiptomi pecsétek vagy bélyegek szintén legtöbbször pusztán piktografikusak voltak, és egyáltalán nem tartalmaztak írott lejegyzéseket.²⁴ Ezek a jelrendszerek hangjelölés nélkül rögzítették egy személy, egy klán vagy egy hely identitását.

(Ellenvetésként felhozható, hogy a rébusz-elvet gyakran használták a heraldikában is, általában képi szójátékok formájában.²⁵ Ez főként az európai heraldikára igaz, azon belül is a családnevekre – például a Bowes-Lyon-család címerében íjak és egy oroszlán szerepel; ezt nevezzük „beszélő címernek”. A rébusz-elv ilyen használata nem azonosítható a tulajdonképpeni hangjelöléssel, mivel nem eléggé módszeres. Nincs kodifikált keret, amely megszabná a hangjelölés konstans és az adott esetre alkalmazható használatát, hiszen különböző képi szójátékok jelölhetik ugyanazt a hangot, amely ugyanakkor kódolható a rébusz-elv alkalmazása nélkül is.)

Összefoglalva, a tulajdonnevek grafikus rögzítése nem követeli meg a fonematikus jelölést; és még ha meg is követelné, fontos hozzátenni, hogy más szófajok számára is kézenfekvő az ilyen leképezés. Hogyan magyarázható te-

²² Olivier MORIN és Helena MITON, „Detecting Wholesale Copying in Cultural Evolution”, *Evolution & Human Behavior* 39, 4. sz. (2018): 392–401, <https://doi.org/10.1016/j.evolhumbehav.2018.03.004>.

²³ Barbara PAVLEK, James WINTERS és Olivier MORIN, „Ancient Coin Designs Encoded Increasing Amounts of Economic Information over Centuries”, *Journal of Anthropological Archaeology* 56 (2019): 101–103, <https://doi.org/10.1016/j.jaa.2019.101103>.

²⁴ Holly PITMAN, „Towards an Understanding of the Role of Glyptic Imagery in the Administrative Systems of Proto-Literate Greater Mesopotamia”, in *Archives before Writing: Proceedings of the International Colloquium Oriolo Romano, 1991*, 177–204 (Torino: Scriptorium, 1994); David WENGROW, „Prehistories of Commodity Branding”, *Current Anthropology* 49, 1. sz. (2008): 7–34, <https://doi.org/10.1086/523676>.

²⁵ Michel PASTOUREAU, *Traité d'héraldique* (Paris: Editions A & J Picard, 2007⁵).

hát, hogy a fonematikus írás mégis sokáig a tulajdonnevek rögzítésére korlátozódott?

Kedvelt válaszom, hogy más nyelvi elemekhez képest a tulajdonneveket nehéz alapvető morfológiai, szintaktikai vagy szemantikai egységekre bontani.²⁶ Egy olyan szókapcsolat, mint „a nő az égszínkék pulóverben” szintaktikai és morfológiai egységekre bontható, amelyek meghatározható jelentéssel bírnak. De a „Patricia Paige Smith” névből hiányzik a belső morfológiai és szemantikai struktúra azon túl, hogy vezetéknévből, középső névből és keresztnévből áll. Egy ideografikus jelölés talán sikeresen rögzítené „a nő a sötétzöld oldalzsebes nadrágban” szókapcsolatot egy sor piktogrammal (pl. [nő] + [sötét + zöld] + [oldalzsebes + nadrág]), amelyek közül mindegyik önmagában is jelentéses. A „Patricia Paige Smith” nevet nem lehet ugyanígy ideogrammákkal jelölni, mert a „Patricia” típusú szavakból hiányzik az a fajta belső struktúra, mely könnyen fellelhető a nyelv más részeiben.²⁷

Két alapvető megoldás lehetséges: vagy egy specifikus, idioszinkratikus grafikai jel megalkotása, amely személy szerint vonatkozik Patricia Paige Smithre, vagy a név fonémáinak rögzítése (/p/, /a/, /t/ stb.). Az egyik lehetőség mnemotechnikai nehézséggel jár: annyi személy-jelkép asszociációt szükséges megtanulni, ahány tulajdonnevet rögzíteni akarunk. A másik lehetőség,

²⁶ Willy van LANGENDONCK és Mark van de VELDE, „Names and Grammar”, in *Oxford Handbook of Names and Naming*, szerk. Carole HOUGH, 17–38 (Oxford: Oxford University Press, 2016), <https://doi.org/10.1093/oxfordhb/9780199656431.013.21>.

²⁷ Ennek a válasznak számot kell vetnie egy ellenérvvel, melyet e cikk egyik bírálója vett fel. A tulajdonnevek ugyanis sok nyelvben felfejthető szókapcsolatokat rejtnek magukban – ahogyan az előbb láthattuk, többek között az egyiptomi vagy navatl nyelvek kapcsán: „Egy meteor szállt le” (*Xibuitltemoc*), „Kis szandál” (a *Caligula* becenév), „Ülő Bika” (*Thathánka Íyotake*) vagy „Küzdjete a hit jó harcában” (egy vallásos szónoki név, melyet a New England-i puritánok előszeretettel használtak). Nyilvánvaló, hogy ezek olyan nyelvi szókapcsolatokból származnak, amelyek könnyen bonthatók egységekre. Úgy tűnhet, hogy ez ellentmond annak, amit az előbb írtam. Még egy látszólag homályos név is, mint a „Patricia Paige Smith” olyan szavakból származik, melyeknek pontos jelentése van (vagy volt): „Smith” valaki, aki fémmel dolgozik, a „Patricia” a „Patricius” („nemes”) női alakja. Ez azonban nem releváns a „Patricia Paige Smith” tulajdonnév jelentésében. A „Smith” szó a „Patricia Paige Smith” névben nem azt jelöli, „aki fémmel dolgozik”, ahogyan a „Caligula” sem „kis szandál”-t jelent. A „Smith” név és a „smith” főnév a poliszémia eseteként kezelendő: két, közös múlttal rendelkező, de külön értelemmel bíró szó. A „Smith” egyszerűen a „Smith” tulajdonnevet jelenti. E tekintetben ugyanolyan homályos, mint bármely más vezetéknév. Hasonlóképpen, ha a „Patriciá”-t a „nemesasszony” jelentéssel ruházzuk fel, ez nem mond semmit a szó jelentéséről a „Patricia Paige Smith” név kontextusában. Nem lehet a „Patricia Smith”-t „Nemesasszony, aki fémmel dolgozik”-ra fordítani – ez semmiképp sem segít a név jelentésének közvetítésében, mely annak a személynek az identitása, akire a név vonatkozik.

a fonémajelölés megszünteti ezt a korlátot. Hiszen ekkor csak limitált számú fonémák grafikus jeleire szükséges emlékezni. Ez elehetővé teszi a tulajdonnevek széles körű rögzítését korlátozott számú jelekkel. A fonematikus elv e kulcsfontosságú előnye messze túlmutat a tulajdonnevek jelölésén, de homályos jelentésük és a belső struktúra hiánya miatt főként ezekre alkalmazva bizonyult hasznosnak.

Ennek ellenére, ahogy az imént láttuk, az ideografikus megoldás korántsem lehetetlen fonematikus jelölés mellett, sőt széles körben elterjedt gyakorlat a piktografikus kommunikációban. Hogyan tartható ez fenn a mnemotechnikai korlát ellenére? Ezt két dolog magyarázhatja. Egyrészt bár a tulajdonnevek általában kisebb mértékű morfológiai szabályosságot mutatnak a lexikai készlet többi részéhez képest, ez esetenként mégis változó. A „gróf Alekszej Kirillovics Vronszkij” név két személyazonosítóra – (Alekszej és Kirilov), egy képzővel jelezve azt, hogy Alekszej Kirilov fia – és egy nemesi címre bontható fel. Az „-ovics” képzőhöz hasonló morfológiai elemeket jelölni lehet egyetlen állandó grafikai elemmel, amely más személyek esetében is használható. Másrészt az ideografikus jelek használatához nem szükséges különféle személyekre vonatkozó emblémák létrehozásának képessége. Az emblémát jellemzően az a személy (vagy az egy háztartásban élő tag) alkotja meg, akit az embléma jelöl, hogy iratokat írjon alá, megjelölje a tulajdonát stb. Másoknak csak fel kell ismernie azt, de nem kell tudnia megalkotni. Az összetett grafikus szimbólumok esetében a felismerési emlékezet (*recognition memory*) jobban működik, mint a produkciós emlékezet (*production memory*). Produkciós emlékezetünk sok hétköznapi szimbólum esetében köztudottan hiányos: a legtöbb amerikai nem tudná pontosan reprodukálni fejből az USA egycentesének kinézetét (melyik oldalra néz Lincoln?).²⁸ Ettől függetlenül még megjegyezhetők ezek a szimbólumok, sőt olyan információt is kódolhatnak, amely egy tulajdonnévből hiányzik. Egy címer rögzíthet például leszármazási vonalat vagy házassági kapcsolatot, amelyre a tulajdonnevek vagy címek nem feltétlenül utalnak; egy nemzet zászlója hordozhat az állam nevén túlmutató vallási vagy politikai jelentéseket. Mindez valamennyire megmagyarázza, hogy miért részesítik előnyben bizonyos célokra az emblémákat a fonematikus jelöléssel szemben még az írástudó társadalmakban is.

Várhatóan tehát az egyéni ideografikus emblémák a viszonylag decentralizált közösségekben játszanak fontos szerepet, ahol minden egyénnek csak

²⁸ Raymond S. NICKERSON és Marilyn Jager ADAMS, „Long-Term Memory for a Common Object”, *Cognitive Psychology* 11, 3. sz. (1979): 287–307, [https://doi.org/10.1016/0010-0285\(79\)90013-6](https://doi.org/10.1016/0010-0285(79)90013-6).

a saját emblémáját szükséges létrehozni, és csak azokét kell felismernie, akikhez érdek fűzi (ami nem feltétlenül sok ember). Ezzel szemben a fonematikus jelölés akkor kerül az előtérbe, amikor egy szűk csoport olyan iratokat hoz létre, amelyek sokféle nevet rögzítenek. Úgy tűnik, hogy a hangjelölés felfedezéséhez vezető kommunikációs igények meglehetősen sajátos körülmények között alakultak ki, és kifejezetten a tulajdonnevek adminisztratív rögzítésére vonatkoztak. A tulajdonnevek rögzítése nem minden körülmény között kívánja meg a fonematikus jelölést; ezt számos emblémarendszer tanúsítja. Ha érvelésem helyes, a hangjelölés egy specifikus igényre adott válaszként jött létre: arra, hogy emberek kis csoportja nagy számú emblémákat idézzenek fel és alkossanak meg. Azonban amint ez a funkció teljesült, évszázadokig semmilyen előrelépés sem történt.

2. ASZINKRON KOMMUNIKÁCIÓ: TÉREN ÉS IDŐN ÁTÍVELŐ ÜZENETÁTADÁS

Az írás másik meghatározó jellemzője, az aszinkron alkalmazás, valószínűleg szintén graduálisan és a hangjelölésnél később alakult ki. Az aszinkron használat nemcsak az írásnak, hanem általánosságban véve minden grafikus jelölésnek kulcsfontosságú tulajdonsága.²⁹ Ekkor időn vagy téren (gyakran mindkettőn) átívelő módon adunk át információt emberi közvetítő nélkül. Ennek elérésére a grafikus kódok voltak az egyedül megbízható eszközök – egészen a legújabb (az archeológusok számára irreleváns) találmányok megjelenéséig. A grafikus kódokat mégsem mindig így használják. Ellenkezőleg, néhány esetben szinte kizárólag szinkron kommunikációt szolgálnak: a varázsigéket leképező piktogramok például a sámánokat a már nagyrészt memorizált énekek recitálásában segítik.³⁰ A hallgatóság (akik nem mindig értik meg a szöveget, vagy nem mindig figyelnek rá) nem közvetlenül a piktogramokból szerez információt, hanem a sámántól; csak maga a sámán tud értelmet adni a piktogramoknak, mert már előzőleg megjegyezte a rögzített – az általa kitalált

²⁹ Olivier MORIN, Piers KELLY és James WINTERS, „Writing, Graphic Codes, and Asynchronous Communication”, *Topics in Cognitive Science* 12 (2020): 727–743, <https://doi.org/10.1111/tops.12386>.

³⁰ Carlo SEVERI, „Their Way of Memorizing: Mesoamerican Writings and Native American Picture-Writings”, *Res: Anthropology and Aesthetics* 71–72 (2019): 312–324, <https://doi.org/10.1086/706117>; W. J. HOFFMAN, „Pictography and Shamanistic Rites of the Ojibwa”, *American Anthropologist* 1, 3. sz. (1888): 209–230, <https://doi.org/10.1525/aa.1888.1.3.02a00010>; Pierre DÉLÉAGE, *Inventer l'écriture Rituels Prophétiques et Chamaniques Des Indiens d'Amérique Du Nord, XVII^e-XIX^e Siècles* (Párizs: Les Belles Lettres, 2013).

vagy szájhagyomány útján terjedő – énekeket. Ez csak egy példa egy nagyon általános jelenségre, a grafikus jelek más esetekben is inkább olyan mnemotechnikai eszközök, amelyek szájhagyomány útján terjedő szövegek emlékeztetőiként funkcionálnak ahelyett, hogy új információt közvetítenének valakinek, aki még nincs annak birtokában. Eltekintve a sámánok énekeitől, történeti feljegyzéseket,³¹ misszilis leveleket vagy akár bevásárlólistákat³² is lehet piktogramokkal kódolni, de csak akkor, ha a tartalmukat előzőleg már memorizálják.

Más szóval nincs szükségszerű kapcsolat egy adott kultúrában jelenlévő grafikus kód és a nagy időbeli vagy térbeli távolságokon átívelő kommunikáció fenntartásának képessége között. Egy komplex grafikus kódot elsajátító kultúra nem válik *ipso facto* képessé arra, hogy távkereskedelemben vegyen részt, levelezést folytasson közvetítők nélkül, vagy hogy gondosan kialakított archívumokat tartson fenn.

Mit változtat ezen a hangjelölés? Elméletileg sokat; gyakorlatilag talán semmit. Az írásban rejlő lehetőség az aszinkron kommunikációra ugyanis kiaknázatlan maradhat évszázadokig, vagy akár még annál is tovább. Csak manapság adja az az írás vitathatatlan értékét, hogy időt és teret átívelő aszinkron kommunikációt működtet, megszüntetve ezáltal a szóbeliség korlátait. De hogyan mérhető, hogy mennyiben használták korábbi kultúrák is a már kialakult írásbeliséget aszinkron kommunikációra? Egy távoli korszak grafikus kódjának szerepét nehéz feladat rekonstruálni, a bizonyítékok többsége közvetett vagy retrospektív. Mégis az a véleményem, hogy az üzenetek sokáig elsősorban szóbeli előadásokat segítettek, vagy az olvasót már memorizált tartalmakra emlékeztették, nem pedig új információt közvetítettek egy új személynek mindenféle szóbeli magyarázat nélkül. Kétféle bizonyíték hozható fel ennek igazolására.

2.1. Bizonyítékok, hogy a korai írásos kommunikáció egyidejű volt

Közvetlen bizonyítékkal szolgál, ha egy szöveg expliciten megemlíti, hogy szinkron kommunikációra írták meg. Kétféle szinkron használatot érdemes kiemelni. Egyrészt sok ókori irodalmi műfaj elválaszthatatlan a szóbeliségtől so-

³¹ Garrick MALLERY és Smithsonian Institution Bureau of American Ethnology, *Pictographs of the North American Indians: A Preliminary Paper* (Washington: Govt. Printing Office, 1886), <https://doi.org/10.5479/sil.789861.39088012443370>.

³² Kenneth CROFT, „Graphic Mechanisms of Communication in Native North America”, *Indiana Magazine of History* 45, 4. sz. (1949): 339–352.

rán elsajátított és fenntartott emlékezés művészetétől.³³ Ez az írás szinkron használatát jelenti, mert a szöveg önmagában nem közvetít új információt egy új befogadónak; inkább segédletként szolgál valaki számára, aki már korábban megismerte az üzenetet (és amelyet egyébként nehéz lenne felidézni az írás eszköze nélkül). Az írás ilyen típusú használatára Eric Havelock „recitációs írásbeliségként” hivatkozik az olyan társadalmakat jellemezve, melyek esetében

az írástudás inkább az emlékezés, mintsem a folyékony olvasás készsége. A szó írott használata rendkívül szűk körű, az írott szó olvasása pedig a kultúra központi funkciójához, a recitáláshoz és a memorizáláshoz képest járulékosnak tekinthető.³⁴

A recitációs írásbeliség közvetlen bizonyítéka a memorizálásban játszott szerep explicit említése, amelyet az írástudás legfontosabb jellemzőjeként tartottak számon. Rengeteg ilyen utalás található a mezopotámiai, zsidó és hellenisztikus művelődési hagyományban.³⁵ A görög epikus költészet vagy a bibliai szövegek szintén prototipikus példái az olyan irodalmi műfajoknak, amelyek nagy mértékben szóbeli hagyományon alapultak, még azután is, hogy már évszázadok óta lejegyezték őket.

Más műfajokra szintén szinkron használat jellemző, mert olyan szövegekből állnak, amelyeket az általuk leírt élőszóbeli interakció kíséretére terveztek. Világos példa erre egy levél, mely megemlíti, hogy egy hírnök olvassa fel (ez egy mindenütt jelenlévő, jól dokumentált gyakorlat).³⁶ Önmagában a felolvasás művelete nem feltétlenül jelenti azt, hogy a hírnök recitál egy részben már megjegyzett szöveget (noha ez valószínűsíthető egy olyan kultúrában, amelyben egyébként a recitációs írásbeliség domináns); azt viszont jelenti, hogy a hallgatóság nem közvetlenül az íráson, hanem a szóbeliség eszközein keresztül nyer hozzáférést a levél által közvetített információhoz. Egy másik,

³³ David M. CARR, *Writing on the Tablet of the Heart: Origins of Scripture and Literature* (Oxford: Oxford University Press, 2005), <https://doi.org/10.1093/oso/9780195172973.001.0001>.

³⁴ Eric A. HAVELOCK, „The Preliteracy of the Greeks”, *New Literary History* 8, 3. sz. (1977): 369–391, 372.

³⁵ CARR, *Writing on the Tablet...*

³⁶ Gareth WEARNE, „The Role of the Scribe in the Composition of Written Correspondence in Israel and Judah”, in *Observing the Scribe at Work: Scribal Practice in the Ancient World*, szerk. R. AST, M. CHOAT, J. CROMWELL, J. LOUGOVAYA és R. YUEN-COLLINGRIDGE, 21–43 (Leuven: Peeters Publishing, 2021); Markus STOCK, „Letter, Word, and Good Messengers: Towards an Archaeology of Remote Communication”, *Interdisciplinary Science Reviews* 37, 4. sz. (2012): 299–313, <https://doi.org/10.1179/0308018812Z.00000000024>.

a szinkron használat nyomát hordozó forráscsoport azokból a szövegekből áll, amelyek egy már lefolytatott vagy eltervezett rituális eseményt rögzítenek. Erre példa többek között a babilóniai és asszír királyok írnokai által írt szövegek. Ezeket olyan rituálék megtervezésére szánták, amelyeket uralkodók szövetségének megkötésekor mutattak be.³⁷ Az írott szó a legtöbb ilyen szövegben az eskü (általában vázlatos) lejegyzésére szolgált, amelyet nyilvánosan el kellett mondania az uralkodónak és a hírnökének (aki mintegy az uralkodó helyettesét alakította). Ezek a mozzanatok azt támasztják alá, hogy a szövegek leginkább csak rituális interakciók leirataként játszottak szerepet: a szövetségkötés esküjét nem kellett leírni ahhoz, hogy érvényes legyen; általában nem foglalták írásba magát a szövetséget sem, amelyet a két király „élőben”, egyazon helyen kötött meg; valamint az eskük átírása olyan tartós anyagra, amely alkalmas lett volna a közzétételre és a hosszú távú megőrzésre, messze nem volt szisztematikus és elterjedt eljárás.

Az írás szinkron használatának itt felsorolt közvetlen bizonyítékai csak azokra a szövegekre vonatkoznak, amelyek eléggé kidolgozottak ahhoz, hogy tudósítsanak saját használatuk körülményeiről. A korai írásos források esetében ez ritkán van így. Használatuk körülményeire tehát közvetett elsősorban nyomokból lehet következtetni.

2.2. Dekontextualizáció és archiválás: az aszinkron kommunikáció két előfeltétele

Úgy tűnik, legalább két dolog szükséges ahhoz, hogy egy szöveg létrehozásának kontextusán kívül is értelmezhető legyen egy olyan olvasó számára, aki még nem ismeri a tartalmát.

Dekontextualizáció. Az aszinkron kommunikáció első összetevője a szöveg által szolgáltatott információra vonatkozik: ennek megértéséhez ugyanis nincs szükség az olvasó memóriájára vagy szóbeli magyarázatra. Egy üzenet önmagában értelmezhető (*self-contained*), amennyiben kevés kontextuális információval is érthető. Ez persze inkább fokozatosság kérdése, és nem egy „mindent vagy semmit” tulajdonság. A nyelvészek szerint az üzenetben foglalt információ és a kontextus adta információ közötti arány a lényeges.³⁸ A be-

³⁷ Dominique CHARPIN, „*Tu es de mon sang*”: *Les alliances dans le Proche-Orient ancien* (Párizs: Les Belles Lettres, 2019), <https://doi.org/10.4000/books.lesbelleslettres.258>.

³⁸ Alison WRAY és George W. GRACE, „The Consequences of Talking to Strangers: Evolutionary Corollaries of Socio-Cultural Influences on Linguistic Form”, *Lingua* 117 (2007): 543–578, <https://doi.org/10.1016/j.lingua.2005.05.005>; James WINTERS, Simon KIRBY és

széd különböző regiszterei is ezen a skálán mozognak: a spontán beszéd kevesebb információt rögzít, és jobban hagyatkozik a kontextusra, mint a tervezett beszéd, mely a kifejtettségre és a részletességre törekszik. (Felvetették azt is, hogy egész nyelvek elhelyezhetők ezen a skálán, amennyiben bizonyos nyelvek másoknál alkalmasabbak a dekontextualizált kommunikációra.)³⁹ Az aszinkron kommunikáció lecsupaszított kommunikáció, mivel hiányoznak belőle a kontextuális jelzések, amelyek csak személyes találkozások alkalmával adóttak.⁴⁰ Az írott szövegek ezt azzal ellensúlyozzák, hogy jobban kifejtettek és inkább értelmezhetők önmagukban, mint az élőbeszéd. Miközben nem minden írott szöveg képes erre, sőt az élőbeszéd némely változata is lehet rendkívül dekontextualizált. Eszünkbe juthat a jogi vagy politikai diskurzus az összetett társadalmakban, de kisebb léptékű közösségekben is találunk viszonylag széles körű és személytelen hallgatóságnak szánt műfajokat vagy formális beszédeket.⁴¹ Mindazonáltal, ha egy üzenetet aszinkron módon akarunk továbbítani (és a többi körülmény változatlan), akkor minden bizonnyal nagyobb szükség van dekontextualizációra.

Tárolás. A második összetevő annak biztosítása, hogy az olvasók hozzáférhessenek a számukra releváns szövegekhez, távol azok létrehozásának idejétől vagy terétől. Attól fogva, hogy egy aszinkron üzenet létrejött, ki van téve az eltűnés veszélyének valamilyen fizikai roncsolódás, hibás elhelyezés következtében, vagy mert lelőhelye feledésbe merül. Az írás sikeres használata az aszinkron kommunikációban megköveteli egyrészt a szöveg hosszú távú túlélésének és hozzáférhetőségének biztosítását (tartósság), másrészt annak elérhető és megbízható helyeken történő tárolását, vagy alternatívaként a szövegek szállításának biztosítását az illetékes olvasóknak (visszakereshetőség). Ennek eléréséhez egy drága, de egyszerű módszer monumentumok létrehozása: egy vésett sztélé elég tartós és szembetűnő ahhoz, hogy bármely írástudó számára könnyen olvasható legyen. Egy másik út az írott szöveg aszinkron működtetésére archívumok létrehozása, azaz az iratok rendezett, tartós tárolása. Az ilyen típusú megoldások nélkül egy könnyen megsemmisíthető vagy rosszul elhelyezett írott üzenet nem használható sokáig új információ

Kenny SMITH, „Contextual Predictability Shapes Signal Autonomy”, *Cognition* 176 (2018): 15–30, <https://doi.org/10.1016/j.cognition.2018.03.002>.

³⁹ Peter TRUDGILL, *Sociolinguistic Typology* (Oxford: Oxford University Press, 2011); James WINTERS, Simon KIRBY és Kenny SMITH, „Languages Adapt to Their Contextual Niche”, *Language and Cognition* 7, 3. sz. (2015): 415–449, <https://doi.org/10.1017/langcog.2014.35>.

⁴⁰ MORIN, KELLY és WINTERS, „Writing, Graphic Codes...”, 727–743.

⁴¹ Maurice BLOCH, *Political Language and Oratory in Traditional Society* (London–New York: Academic Press, 1975).

átadására olyasvalakinek, aki még nem ismeri a tartalmát. Lehet ugyan használni információátadásra, de csak korlátozott számú alkalomkorán, szóbeli magyarázattal egybekötve vagy mnemotechnikai eszközként.

Önmagukban értelmezhető szövegek és rendezett archívumok: bár e két dolog előfeltétele az írás aszinkron használatának, hiányuk mégsem korlátozza az írás kialakulását, ha azt általános, fonematikus kódként értjük. Az önmagukban értelmezhető szövegek és a rendezett archívumok az írott szöveg autonómiájának mértékét határozzák meg, vagyis, hogy mennyire van ráutalva szóbeli magyarázat közvetítése, nagy mennyiségű kontextuális tudásra és a szöveg tartalmának előzetes ismeretére. Az írást azonban létrejöttékor általában nem így használták. Néhány szerző az írás létrejötte és a dekontextualizált információ archiválása közötti szerves viszony mellett érvelt;⁴² én azonban az ellenkező véleményen vagyok.

2.2.1. *Önmagukban értelmezhető szövegek létrehozása*

Mennyire voltak önmagukban értelmezhetők az üzenetek? Mennyire voltak könnyen olvashatók? Az első kérdés megválaszolásának nehézségét az adja, hogy jelentést kell társítani egy létrejöttének kontextusából kiragadott szöveghez. A paleográfusok és archeológusok tudják a legjobban, ez mennyire nehéz. Azt is tudják, hogyan változik a kérdésre adott válasz helyszíntől, korszaktól és az anyag típusától függően. Mégis, az összehasonlítás jó kiindulópontja lehet a szövegek terjedelme. Elméletileg egy minden más szempontból azonos, de hosszabb szöveg több információt tartalmaz egy rövidebbnél; a terjedelem tehát – közvetett módon – az önmagában való értelmezhetőség mértékét jelezheti. A legjobb tudomásom szerint a szövegek terjedelméről még nem folytattak területek és időszakok szerinti általános összehasonlító vizsgálatot, és a dokumentumok egyesével történő összevetésekor minden bizonnyal komoly technikai nehézségek merülnének fel. Ha azonban a korai írásbeliség műfajait vizsgálva azt kérdezzük, milyen típusú szöveg alakult ki először, a válasz viszonylag világos. Néhány műfajnak inherens terjedelmi és tartalmi korlátai vannak, míg mások sokkal alkalmasabbak a széles körű, önmagában is értelmezhető üzenetek közvetítésére. Egy nevet tartalmazó kártus, egy pecsétfelirat, egy ívóedényen vagy egy cserépdarabon lévő szöveg eredendően rövid. Rendkívüli rövidségük azt sugallja, hogy az ilyen feljegy-

⁴² Jack GOODY, *The Domestication of the Savage Mind* (Cambridge England–New York: Cambridge University Press, 1977).

zések kevés információt hordoztak, és csak jól meghatározott kontextusban adhattak nekik értelmet azok, akik ismerték a háttérükben álló kommunikációs helyzetet. A teljes leveleket, krónikákat, eposzokat vagy törvénygyűjteményeket viszont sokkal könnyebb önmagukban megérteni, mert több és részletesebb információt tartalmaznak.

A négy terület közül, ahol az írás ősi találmánya létrejött, egyiknél sem egész levelek, krónikák vagy törvénygyűjtemények jelentek meg előbb, hanem inkább rövidebb műfajok, például kerámiafeliratok, pecsétek vagy azonosítójelek. Egyiptomban az azonosítójelek és a kerámiafeliratok időbeli elsőbbsége nyilvánvaló.⁴³ Mezopotámiában szintén, ahol a táblás könyvelési jegyzékek legalább évszázadokkal megelőzik az olyan forrásokat, mint a levelek, elbeszélések vagy jogi szövegek.⁴⁴ Csupán a szójegyzékek képeznek kivételt ebben a mintában; ezek ugyanis meglehetősen hosszúak lehetnek, és nem sokkal a könyvelési iratok után jelentek meg.⁴⁵ Ugyanakkor az is kétségtelen, hogy ezeket mnemotechnikai eszközként és nem új információ pusztán írás általi átadására használták. A prekolumbiánus közép-amerikai források ugyanezt a mintát követik, itt is kifejezetten rövid szövegek találhatók (minden esetben kevesebb mint pár száz jel hosszúságban, kivéve néhány későbbi maja feliratot).⁴⁶ Ami Kínát illeti, az írásos emlékek úgymond teljes fegyverzetben emelkednek ki az archeológiai leletek közül; a Sang-dinasztia korábók származó jóslási csontfeliratok már hosszú és önmagukban értelmezhető elbeszéléseket tartalmaznak. Ám ha igaz az állítás, miszerint a korábbi bronzedényeken talált emblémák a Sang-korabeli feliratok elődei voltak, akkor álláspontom beigazolódik, miközben az továbbra is homályos marad, hogy vajon ezek az emblémák hangjelölő írásnak minősülnek-e.⁴⁷ Mindenesetre azt láthatjuk, hogy az archeológiai leletek többsége alátámasztja a hosszú, önmagukban is értelmezhető szövegek megjelenésének gradualista modelljét.

⁴³ BAINES, *Visual and Written Culture...*

⁴⁴ NISSEN, DAMEROW, ENGLUND és ENGLUND, *Archaic Bookkeeping...*; MAIOCCHI, „Writing in Early Mesopotamia...”, 395–424.

⁴⁵ Rita WATSON, „Archaic Lists, Writing and Mind”, *Pragmatics & Cognition* 21, 3. sz. (2013): 484–504, <https://doi.org/10.1075/pc.21.3.04wat>; K. WAGENSONNER, „Early Lexical Lists Revisited. Structures and Classification as a Mnemonic Device”, in *Language in the Ancient Near East* 1, 1. sz. (2010): 285–312, <https://doi.org/10.1515/9781575066394-010>, *Proceedings of the 53e Rencontre Assyriologique Internationale*. Winona Lake: Eisenbrauns.

⁴⁶ JUSTESON és MATHEWS, „Evolutionary Trends...”, 88–132.

⁴⁷ Nicholas POSTGATE, Tao WANG és Toby WILKINSON, „The Evidence for Early Writing: Utilitarian or Ceremonial?”, *Antiquity* 69, 264. sz. (1995): 459–480, <https://doi.org/10.1017/S0003598X00081874>.

2.2.2. Az írott szöveg tárolása és visszakeresése

Az aszinkron kommunikáció második összetevője az iratok könnyű olvashatósága létrehozásuk kontextusán kívül, ami az iratok tartósságát és visszakereshetőségét kívánja meg. A korai írott dokumentumok tartósságának vizsgálatakor kézenfekvő problémával találjuk szembe magunkat: csak a tartós anyagokból lesznek régészeti leletek,⁴⁸ a kevésbé tartós szövegek pedig eltűnnek. Két dolgot mégis viszonylag biztosan meg lehet állapítani.

Először is azon a négy területen, ahol az írás létrejött, sokkal több sérülékeny (megőrzött vagy elveszett) irat keletkezett a tartós szövegekhez képest.⁴⁹ Haicheng Wang hosszú fejezetet szentel könyvében ennek tanulmányozására. Fő érve, hogy a tartós feliratok mind a négy területen gyakran utalnak olyan feljegyzésekre, amelyek teljesen elvesztek, de minden bizonnyal elterjedtek és gondosan őrzöttek voltak.⁵⁰ Másrészt léteznek közvetlen írott, képi vagy foszszilis bizonyítékok olyan sérülékeny hordozókra, mint például az egyiptomi fatáblák vagy a kínai bambusztekercesek. Harmadrészt Steinkeller szerint a mezopotámiai táblás könyvelési jegyzékeket a bennük leírt tranzakciót követően hozták létre, amelyeket vagy fejben tartottak, vagy olyan emlékeztető eszközök segítségével jegyeztek meg, mint például a számviteli pálcák vagy a számlálókövek.⁵¹ (Steinkeller hasonló érvet vázol fel az ókori egyiptomi fatáblákkal kapcsolatban is.)⁵² Az említett anyagok közül néhány csupán egy embe-
röltön át maradt tartós, archeológiai léptékben viszont sérülékeny volt; bár nem mindegyik. A viasz- és agyagtáblák például könnyen elkoptak. Lehetetlen felmérni, mennyi írott információ veszett el így (elégg csak arra gondolni, hogy a diákok is mindig letörölték a palatáblát). Az itt összefoglalt érvek min-

⁴⁸ Charles PERREAULT, *The Quality of the Archaeological Record* (Chicago, IL: University of Chicago Press, 2019), <https://doi.org/10.7208/chicago/9780226631011.001.0001>.

⁴⁹ COOPER, „Babylonian Beginnings...”, 71–99; POSTGATE, WANG és WILKINSON, „The Evidence for Early Writing...”, 459–480; Haicheng WANG, *Writing and the Ancient State: Early China in Comparative Perspective* (Cambridge: Cambridge University Press, 2014), <https://doi.org/10.1017/CBO9781139235709>.

⁵⁰ Lásd még BAINES, *Visual and Written Culture...*

⁵¹ P. STEINKELLER, „Archival Practices at Babylonia in the Third Millennium”, in *Ancient Archives and Archival Traditions: Concepts of Record-Keeping in the Ancient World*, szerk. Maria BROSIUS (Oxford: Oxford University Press, 2003), <https://doi.org/10.1093/oso/9780199252459.003.0003>; STEINKELLER, „The Function of Written Documentation in the Administrative Praxis of Early Babylonia”, in *Creating Economic Order: Record Keeping, Standardization, and the Development of Accounting in the Ancient Near East*, szerk. M. HUDSON and C. WUNSCH, 65–88 (Maryland: University Press of Maryland, 2004).

⁵² Uo., 16. lj.

denesetre arra sarkallnak, hogy a tartós feljegyzéseket inkább kivételnek, mintsem normának tekintsük. Még ha tartósak is, a monumentális feliratok nem feltétlenül értek össze az aszinkron kommunikációval, mert sok közülük olvashatatlan volt. (Az egyiptomi sírfeliratok például vagy túl magasan voltak ahhoz, hogy olvashatóak legyenek, vagy be voltak falazva egy sírkamrába, vagy egy talizmánba voltak zárva stb.)⁵³

A sérülékeny üzenetek csak kis mértékben segíthették volna az aszinkron kommunikációt, hacsak nem állították volna elő őket nagy számban, ami rendkívül nehéz és költséges lett volna ipusztriális vagy protoipusztriális eszközök nélkül. További nehézség, hogy egy szöveg csak akkor tud részt venni az aszinkron kommunikációban, ha elér a címzettekhez, ami csak akkor lehetséges, ha – a visszakereshetőség kívánalmának értelmében – a szövegek megfelelő módon vannak tárolva és rendezve.

A későbbi olvasás céljából tárolt és elrendezett írásos dokumentumok összességét közismerten archívumnak hívjuk; ugyanakkor az, hogy minden, amit a régészek archívumnak neveznek, valóban rendelkezik mindezekkel a tulajdonságokkal, már korántsem vitathatatlan. Úgy tűnik, az általánosan használt terminusnak⁵⁴ kevésbé szigorú a jelentése: alkalmazhatónak tűnik palotában vagy templomban található iratok bármilyen tárolására (legyen az írott szöveg vagy egyszerű számviteli irat). Így az a rejtély továbbra is fennáll, hogy e megőrző helyek közül melyek és milyen mértékben szolgáltak teljes értékű archívumként. A bizonytalanságot jól mutatja, hogy a kérdésben két, egymással szöges ellentétben álló álláspontot képvisel Christopher Eyre *The Use of Documents in Pharaonic Egypt*⁵⁵ című és Haicheng Wang már idézett, *Writing and the Ancient State* című munkájában. A két szerző teljesen eltérő következtetésre jut egymást átfedő korpuszt vizsgálva. (Eyre munkája Egyiptomra fókuszál, de gyakran utal Mezopotámiára, míg Wang egy összehasonlító kutatást végzett, amely magában foglalja az írás mind a négy ősi változatát, Egyiptommal együtt.)

Egyrészt fontos bizonyítékok szólnak amellet, hogy az ókori archívumok az írott információ tárolásának és hozzáférésének hatékony eszközeiként funkcionáltak. Legelőször maguk a megőrző helyek bizonyító erejük: fontos irattárakat találtak olyan helyeken, ahol ezeknek egyértelmű relevanciájuk lehetett a palota vagy a templom hivatalnokai számára. Még ha a közterületen

⁵³ BAINES, *Visual and Written Culture...*

⁵⁴ Például BROSIUS, szerk., *Ancient Archives and Archival Traditions...*

⁵⁵ Christopher EYRE, *The Use of Documents in Pharaonic Egypt* (Oxford: Oxford University Press, 2013), <https://doi.org/10.1093/acprof:oso/9780199673896.001.0001>; WANG, *Writing and the Ancient State...*

lévő archívumok vegyítették is a köz- és magániratokat,⁵⁶ jó részük hasznos lehetett az adminisztráció működtetésében. Vannak (területenként változó minőségű) eseti bizonyítékok indexelési technikákról is, mint például az El-bánál fennmaradt ékírásos agyagtáblákon elhelyezett címkék,⁵⁷ a hieroglif feliratok az Tell-El-Amarnában talált táblák oldalán⁵⁸ vagy a papirusztekercsek külsejére írt címek.⁵⁹ Az archiválási technikák minősége idővel javult, az iratgyűjtemények nagyobb és szisztematikusabbá váltak egészen a klasszikus könyvtárak koráig. Az iratok egyre könnyebben hozzáférhetőek lettek a tartalomjegyzék és az indexelés gyakorlatának lassú megjelenésével (ez még a klasszikus Rómában is kezdetleges és ritka volt).⁶⁰ Találhatunk néhány korai bizonyítékot arra is, hogy az archivált iratokat jóval az összerendezésük után is megtekintették.⁶¹ Ez alátámasztja azt az alapvető gondolatot, hogy néhány nagyon régi archívumot szándékosan a későbbi megtekintés érdekében tartottak fent, vagy legalábbis alkalmi szinten ezt a célt szolgálták a gyűjtemények.

Másfelől viszont a megfelelően gondozott feljegyzések rendszeres hozzáférhetőségét biztosító archívumok elterjedtségével szemben ellenvetések is felmerülhetnek, még az olyan, viszonylag késői korok kapcsán is, mint a hellenisztikus és a klasszikus Egyiptom vagy akár a késő középkori Anglia.⁶² Ugyanazok az érvek térnek vissza az ilyen ellenvetésekben. Először is hiányos bizonyítékaink vannak az iratok szisztematikus és folyamatos tárolásáról és megőrzéséről (ilyenek lennének az akkurátusan vezetett évkönyvek vagy krónikák, amelyek hosszú időtávú, szisztematikus feljegyzéseket tartalmaznak) az iratok *ad hoc* felhalmozásával szemben. Ez fontos téma Eyre könyvében, amely radikálisan újragondolja azt, hogy az archívumokban felhalmozott iratok mennyiben tekinthetők olyan feljegyzéseknek, amelyeket későbbi olvasásra szántak. Az iratokat szerinte jobb munkaanyagokként értelmezni, ame-

⁵⁶ EYRE, *The Use of Documents...*; Frederick Mario FALES, „Reflections on Neo-Assyrian Archives”, in BROSIUS, *Ancient Archives and Archival Traditions...*, 195–229.

⁵⁷ WANG, *Writing and the Ancient State...*, 245–247.

⁵⁸ Richard B. PARKINSON és Stephen QUIRKE, *Papyrus* (London: British Museum Press, 1995), 60.

⁵⁹ Uo., 60.

⁶⁰ Andrew RIGGSBY, *Mosaics of Knowledge: Representing Information in the Roman World: Classical Culture and Society* (Oxford–New York: Oxford University Press, 2019), <https://doi.org/10.1093/oso/9780190632502.001.0001>.

⁶¹ CHARPIN, „*Tu es de mon sang*...”, 120–121; EYRE, *The Use of Documents...*, 94.

⁶² EYRE, *The Use of Documents...*; Rosalind THOMAS, *Literacy and Orality in Ancient Greece: Key Themes in Ancient History* (Cambridge: Cambridge University Press, 1992), <https://doi.org/10.1017/CBO9780511620331>; M. T. CLANCHY, *From Memory to Written Record, England 1066–1307* (Oxford–Cambridge–USA: Blackwell, 1993).

lyek egy maroknyi írnonkon kívül nem szóltak senki másnak.⁶³ Kiemeli az iratok silány rendezését (később, a hellenisztikus korban is) a szövegek tartalmában (a témakörök elkülönítésének és a szövegrészek elemi felosztásának hiányára mutat rá) és azok tárolásában egyaránt. Thomas a klasszikus Hellásról és Clanchy a késő középkori Angliáról hasonló képet fest.⁶⁴ A végső érv, amely alapján kétségbe vonható, hogy a hivatali személyek rendszeresen és hatékonyan vették igénybe az archívumokat, az az iratok jogi státuszának viszonylagos gyengesége: az Eyre, Thomas és Clanchy által lefedett időszakokban és területeken a szóbeli források az írásbeliekénél nagyobb bizonyító erővel bírtak. Egy ügylettel kapcsolatos szóbeli tanúvallomásnak sokszor legalább ugyanakkora súlya volt, mint az írott feljegyzésnek, ami azt mutatja, hogy az írásos dokumentáció nem volt független a szóbeli és mnemotechnikai eszközöktől. A paloták archívumaiban a magán- és köziratok keveredése⁶⁵ és az olyan privát módon tárolt iratok, melyeknek nem volt nyilvánosan elérhető másolata, szintén azt jelzik, hogy milyen kis számban léteztek sok iratot szám-láló, megfelelően elrendezett közgyűjtemények.

Összességében tehát felhozható érv amellett, hogy az írás már feltalálásának pillanatában lehetővé tett kismértékű aszinkron kommunikációt, de ezt érdemes ellensúlyozni legalább ugyanolyan jelentőségű ellenérvekkel, miszerint két akadály is korlátozta az aszinkron használatot. Egyrészt a különböző szövegdarabokban foglalt információmennyiség gyakran elégtelen volt kontextustól független értelmezéshez, másrészt az iratok sokszor nem voltak tartósak vagy visszakereshetők, azaz csak néhány esetben tudták elősegíteni az aszinkron kommunikációt. E két akadály fokozatosan szűnt meg, és csak az írás néhány használata esetében. Manapság is az írásbeliség több formája majdnem kizárólag szinkron módon történik az aszinkron használatok mellett, gyakran egyazon kultúrán belül. Például a bibliai héber nyelv olyan fokú ismerete esetében, amelyre egy gyakorló zsidónak ahhoz van szüksége, hogy a Tórát a bar-micvón recitálhassa, az írott szó inkább emlékeztetőként szolgál; ugyanaz a személy angol nyelven aszinkron módon kommunikál, amikor e-mailt ír, könyvet olvas stb.

⁶³ Az újasszír archívumokról szóló hasonló véleményről lásd: FALES, „Reflections on...”.

⁶⁴ THOMAS, *Literacy and Orality...*; CLANCHY, *From Memory...*

⁶⁵ EYRE, *The Use of Documents...*; FALES, „Reflections on...”

3. KONKLÚZIÓ

A toll és a repülés egyesítése egyértelmű evolúciós sikertörténet: a tollak segítenek a madaraknak a repülés irányításában, és javítják az aerodinamikát. A toll és a repülés mégis különböző időben alakult ki (a dinoszauruszoknál, majd a madarak őseinél), eleinte teljesen más funkcionális kényszer nyomása alatt.⁶⁶ Bizonyos fajok nem rendelkeznek (vagy nem rendelkeztek) az egyikkel vagy a másikkal, míg mások rendelkeztek mindkettővel, de ezek teljesen más funkciót töltöttek be (a repülés a helyváltoztatásra, a toll a hőszabályozásra szolgált, vagy a külső megjelenés miatt volt fontos). Csak a madarak leszármazási vonalában volt kihasználva teljes mértékben a toll repüléssegítő potenciálja, más esetekben a toll nem ilyen szerepet töltött be. A tollal megsegített repülés evolúciója fokozatos volt. E felfedezés segített a zoológusoknak abban, hogy a toll evolúciójának tanulmányozásakor elkerüljenek bizonyos buktatókat. Nevezetesen arra jöttek rá, hogy nem ugyanaz az oka a tollal történő repülés jelenkori adaptivitásának és múltbeli kialakulásának. Ahogy Niko Tinbergen mondaná, egy-egy tulajdonság evolúciós történetének kérdését szigorúan el kellene különíteni a tulajdonság adaptív funkciójának vizsgálatától.⁶⁷

Ahogy a tollal megsegített repülés, az írás evolúciója is fokozatos volt. A mai írásbeliség, ahogyan mi is használjuk, legalább három összetevőből áll: a hangjelölésből (a nyelv hangalapú lejegyzéséből), a kód általánosságából, amely minden kimondott szót jelölni tud, és az aszinkron kommunikációból – az időn és téren átívelő üzenetátadás képességéből. Nézetem szerint a kapcsolat e három sajátosság között nem véletlenszerű. Máshol már megindokoltam, hogy a hangjelölés volt az egyetlen út (legalábbis a legutóbbi időig) egy általános grafikus kód eléréséhez,⁶⁸ és hogy az írás egyedülálló módon alkalmas az aszinkron kommunikációra.⁶⁹ De ez nem jelenti azt, hogy a három tulajdonság együttesen és egyazon okból alakult ki. A tanulmány amellet érvelt, hogy a hangjelölés, az általános jelleg és az aszinkron használat nem egyidőben jött létre. Két nagy időbeli eltolódás választja el egyrészt a hangjelölés felfedezését annak tulajdonneveken túli, általános használatától, másrészt az írás mint általános eszköz létezését annak rutinszerű, aszinkron használatától. Az írás feltalálói az egyes lépésekben különböző kihívásokra reagáltak. A hangjelölés

⁶⁶ Julia CLARKE, „Feathers Before Flight”, *Science* 340, 6133. sz. (2013): 690–692, <https://doi.org/10.1126/science.1235463>.

⁶⁷ Niko TINBERGEN, „On Aims and Methods of Ethology”, *Zeitschrift für Tierpsychologie* 20, 4. sz. (1963): 410–433, <https://doi.org/10.1111/j.1439-0310.1963.tb01161.x>.

⁶⁸ MORIN, „The Puzzle of Ideography”, 1–69.

⁶⁹ MORIN, KELLY és WINTERS, „Writing, Graphic Codes...”, 727–743.

létrehozói nem feltétlenül voltak tudatában, hogy találmányuk nem csak a tulajdonneveket képes jelölni. Akik az írást általános érvényű eszközzé fejlesztették, mely minden mondhatót rögzíteni tud, egyszerűen nem foglalkoztak a szövegek jövőjével. Aligha tudták megjósolni, hogy a kódot, amelyet szóbeli recitációk segédeszközeként vagy lebonyolított cserék emlékeztetőjeként használtak, többnyire egészen más módon fogják használni: szóbeli magyarázatot és mnemotechnikát nélkülöző információtárolásra térben és időben távoli befogadók számára, akinek az információ teljes mértékben újszerű.

Az írás fejlődés útján jött létre, és a közvélekedéssel ellentétben evolúciója hosszú és lépcsőzetes volt. Az írás mai evolúciós elméleteinek⁷⁰ semmi köze azokhoz az elavult megközelítésekhez, amelyek egyesek számára a naiv teleológia és az etnocentrizmus szinonimájává váltak. Épp ellenkezőleg, egy evolúciós perspektíva segíthet fölülírni azokat a funkcionista előítéleteket, amelyek befolyásolhatják az írás történetével kapcsolatos szemléletünket, és segíthet észrevenni, mennyire valószínűtlen és esetleges is ez a találmány.

Olivier MORIN. „The piecemeal evolution of writing”. *Lingue e Linguaggio* 26, 2. sz. (2023): 217–237.

Fordította: László István

⁷⁰ Andy LOCK és Matt GERS, „The Cultural Evolution of Written Language and Its Effects: A Darwinian Process from Prehistory to the Modern Day”, in *Writing: A Mosaic of New Perspectives*, szerk. Elsa GRIGORENKO, Elsa MAMBRINO és David PREISS, 11–36 (Sussex England: Psychology Press, 2012).

Metronom

Versmértékek változatainak azonosítása
a lokális szekvenciaillesztés módszerével

BEVEZETÉS

Minden versforma eredeztethető valahonnan, bár hajlamosak vagyunk azt gondolni – mint ismerős dolgok esetében általában –, hogy mindig is velünk voltak. Egy prozódiai sémát valóban évszázadokon vagy évezredekken át másolhatnak egy kultúrában, de meg is változtathatják azt az egyes szereplők, átvehetik idegen hagyományokból vagy népdalokból, ahogyan a nyelv alakulása is alapvető módosításokat eredményezhet. Ezeket a folyamatokat ugyanakkor közismerten nehéz nyomon követni nyelveken és korokon át: a legtöbb forrás túlságosan töredékes, illetve túlságosan az írásbeliségtől vagy a modern néprajz megfigyeléseitől függ. A nemzeti verstanok szintén jelentősen eltérnek egymástól és csak ritkán egyeztethetők össze; miközben az irodalomtudomány is egyre jobban eltávolodott a versmérték kérdésétől, amely manapság már inkább a nyelvészethez és a fonológiához kapcsolódik. A verselés történeti vizsgálata, amely Antoine Meillet,¹ Roman Jakobson² és később Mihail Gasparov³ munkáiban is a történeti nyelvészet eredményeire épült, sokáig a háttérbe szorult, ám manapság ismét előtérbe került az új szempontokat felvető számítógépes és kvantitatív módszerek elterjedésének köszönhetően.⁴

Tanulmányunk a versmértékek történeti vizsgálatának hagyományát folytatja, és elsősorban lírai szövegek *formai* kapcsolatának leírására törekszik.

¹ Antoine MEILLET, *Les origines indo-européennes des mètres grecs* (Paris: Les Presses universitaires de France, 1923).

² Roman JAKOBSON, „Zur vergleichende Forschung über die slavischen Zehnsilber”, *Slavistische Studien* 1, 5. sz. (1929): 7–20.

³ Michail Leonovič GASPÁROV, *A History of European Versification* (Oxford: Clarendon, 1996), <https://doi.org/10.1093/acprof:oso/9780198158790.001.0001>.

⁴ Vera POLILOVA, „Spanish Romancero in Russian and the semantization of verse form”, *Studia Metrica et Poetica* 5, 2. sz. (2019): 77–108, <https://doi.org/10.12697/smp.2018.5.2.04>; Artjoms ŠEĽA, Petr PLECHÁČ és Alie LASSCHE, „Semantics of European Poetry Is Shaped by Conservative Forces: The Relationship between Poetic Meter and Meaning in Accental-Syllabic Verse”, szerk. Marcus PERLMAN, *PLOS ONE* 17, 4. sz. (2022): e0266556, <https://doi.org/10.1371/journal.pone.0266556>.

A lokális szekvenciaillesztés módszerét alkalmazzuk, hogy versek részei között strukturális hasonlóságot mutathassunk ki: az eljárás lényege, hogy a szövegeket prozódiai jellemzők sorozataként kódoljuk egy négyelemű jelkészlet segítségével, amely bármely nyelv esetében alkalmazható függetlenül a verselési módtól (legyen az időmértékes, szótagszámláló vagy hangsúlyos), majd ezeknek a sorozatoknak a matematikai távolságát számítjuk ki a jelek egyezése/eltérése alapján. A feltárt összefüggések nem jelentenek szükségszerűen evolúciós kapcsolatot, de a strukturális hasonlóság önmagában tanulságos lehet: segíthet azonosítani és lokalizálni egy verselés változatait, feltárhatja formák folytonosságát korokon és hagyományokon át, illetve a történeti kutatások számára is hasznos összefüggésekre mutathat rá. Míg a versek prozódiai és szerkezeti jellemzőit egyre gyakrabban használják a szerzőség azonosítására,⁵ tanulmányunk a kultúrtörténet nagy, egyéni sajátosságokat meghaladó mintázataira összepontosít: a versmértékekre mint sokszor láthatatlan, de meghatározó erőre a költészeti hagyományok alakulásában.

A versmérték a beszédet ismétlődő egységekbe rendezi, a megszólalások között szisztematikus kapcsolatokat létrehozva.⁶ Ez különböző prozódiai rendszerek alapján írható le, amelyek közül a legismertebb a hangsúlyos és hangsúlytalan szótagok mintázata (szótagszámláló), a szótagok hosszának szabályos váltakozása (időmértékes, mint a klasszikus latin), vagy egyszerűen a szótagok számának a kötöttsége (izoszillabikus metrumok az újlatin nyelvekben).⁷ Ezenfelül a prozódia csak az egyik összetevője a versmértékeknek. Az óangol költészet a prozódiai minta mellett az alliterációk rendszerét is elő-

doi.org/10.1371/journal.pone.0266556; Mirella DE SISTO, „The Development of a Poetic Tradition. A Study of a Dutch Renaissance Poetry Corpus”, *Studia Metrica et Poetica* 10, 1. sz. (2023): 36–68, <https://doi.org/10.12697/smp.2023.10.1.02>.

⁵ Benjamin NAGY, „Metre as a Stylometric Feature in Latin Hexameter Poetry”, *Digital Scholarship in the Humanities* 36, 4. sz. (2021): 999–1012, <https://doi.org/10.1093/llc/fqaa043>; Petr PLECHÁČ és Artjoms ŠEĽA, *Versification and Authorship Attribution*, Czech Literary Studies (Prague: Karolinum Press, 2021), <https://doi.org/10.14712/9788024648903>.

⁶ Etunimetön FROG, szerk., „Metrical Entanglement: The Interface of Language and Meter”, in *Versification: Metrics in Practice*, 249–293 (SKS Finnish Literature Society, 2022), <https://doi.org/10.21435/sflit.12>.

⁷ Ahogy a tanulmány is említi, a nemzeti verselési hagyományok jelentősen különböznek egymástól. A latin-görög mintára létrejött és a magyarban is meghonosodott időmértékes verselés a szótagok hosszúsága szerint hoz létre bináris kódot (rövid-hosszú), míg például az angol és a legtöbb európai költészetben a hangsúlyos és hangsúlytalan szótagok váltakozása hoz létre hasonló rendszert. A kötött szótagú („szótagszámláló”) versek a verslábak és a szótagok számát is meghatározzák. – *A ford.*

írta, vagy ismertebb példaként említhetünk olyan rímes formákat, ahol a hangzásbeli összecsengés is meghatározott.

Egy versmérték tehát különböző fokú rugalmasságot megengedő szabályként működik, amelyet az emberi kogníció és a nyelv prozódiajának lehetőségei mindenképpen korlátoznak.⁸ Ezek a szabályok, még ha irodalmárok és költők elő is írják őket, változó társadalmi szokások függvényei, és minden esetben a nyelv adottságaihoz igazodnak.⁹ Mindez nagyfokú változatosságot eredményez az absztrakt minták megvalósulásában: a csonka láb, a kiegészítő szótagok és egyéb szabálytalanságok mind a metrikai rugalmasság és a közönség elvárásainak határait tesztelik. A mintákat a múltbeli formákat másoló egyéni használatok erősítik meg – ám a versek együttesen olyan tendenciákat hoznak létre, amelyeket az egyéni szeszély nem írhat felül egykönnyen. A versformáknak ez a fontos kulturális aspektusa arra készítette kutatókat, hogy az automatikus metrumfelismeréshez és a verselési szabályosság megállapításához alulról építkező megközelítéseket dolgozzanak ki,¹⁰ amelyek az adatokból kirajzolódó szerkezeti sémák mentén teszik lehetővé a csoportosítást. Ez a leíró megközelítés a generatív szemlélet ellenpontjának tekinthető, amely a metrikai variációkat a „mélystruktúrákból” és a nyelvspecifikus „megfelelési szabályokból” absztrahálja,¹¹ elvonatkoztatva a versmértékek történeti és kulturális aspektusaitól.

Módszerünket a metrumfelismerés nem triviális feladatán értékeljük ki egy annotált, többnyelvű korpusz segítségével, és megmutatjuk, hogy az eljá-

⁸ David C. RUBIN, *Memory in Oral Traditions: The Cognitive Psychology of Epic, Ballads, and Counting-Out Rhymes* (Oxford: Oxford University Press, 1995), <https://doi.org/10.1093/oso/9780195082111.001.0001>; Varun DECASTRO-ARRAZOLA, „Typological tendencies in verse and their cognitive grounding” (Disszertáció, LOT, 2018), <https://www.lotpublications.nl/typological-tendencies-in-verse-and-theircognitive-grounding>; Varun DECASTRO-ARRAZOLA, „Testing the robustness of final strictness in verse lines”, *Studia Metrica et Poetica* 5, 2. sz. (2019): 55–76, <https://doi.org/10.12697/smp.2018.5.2.03>.

⁹ Maxim SHAPIR, „Metrum et rhythmus sub specie semioticae”, in Maxim SHAPIR, *Universonum versus*, 1. köt., 91–128 (Moszkva: Iazyki russkoi kultury, 2000).

¹⁰ J. D. PORTER, „The Space of Poetic Meter — Stanford Literary Lab.”, 2018, <https://litlab.stanford.edu/hooddistance/>; Artjoms ŠEĻA és GRONAS, „Measuring Rhythm Regularity in Verse: Entropy of Inter-Stress Intervals”, in *CHR 2022: Computational Humanities Research Conference*, 231–242 (Antwerpen: CEUR-WS, 2022), https://ceur-ws.org/Vol-3290/short_paper5417.pdf; Petr PLECHÁČ és David J. BIRNBAUM, „Rhythm-based authorship recognition in syllabic and accentual-syllabic verse”, *Literatura: teoría, historia, crítica* 25, 2. sz. (2023), <https://doi.org/10.15446/lthc.v25n2.108502>.

¹¹ Nigel FABB, Carlos PIERA és Morris HALLE, *Meter in Poetry: A New Theory* (Cambridge, UK New York: Cambridge University Press, 2008), <https://doi.org/10.1017/CBO9780511755040>.

rás felismerhetővé tesz hasonló metrikus szerveződések egy egyszerre több nyelv esetében is (vagyis egy csoportba rendezi a különböző nyelveken írt, de hasonló metrumú verseket). Azonban nem az osztályozás az elsődleges célunk: a módszer lehetőségeit inkább a történeti kutatásban hangsúlyozzuk. Három esettanulmányt mutatunk be, amelyek különféle nyelveket, verselési rendszereket és korszakokat ölelnek fel (a klasszikus latin időmértékes verselésétől a szótagszámláló hangsúlyos romantikus költészetig). Végezetül, abban a reményben, hogy a módszertan hasznos lehet a tágabb közösség számára, létrehoztunk egy nyilvánosan elérhető Python-könyvtárat is *metronome* néven, amellyel minden számításunk elvégezhető.

MÓDSZER

A Metronom fejlesztésekor a genom metaforája vezetett bennünket, illetve a minden sejt szintű élet alapját képező négy DNS-molekula. Bár csak négy szimbólumból állnak, a genomok hossza jócskán eltérhet egymástól, és egyaránt képesek kisebb – például az emberi testvérek közötti – és jelentős különbségeket is kifejezni: ugyanazok a szimbólumok írhatnak le egy kaktuszt, egy delfint vagy egy kolibrit. A Metronom jelkészletét hasonlóképpen, a lehető legegyszerűbb módon alkottuk meg, így segítségével ugyanazokat a vizsgálatokat tudjuk elvégezni, amelyeket a DNS-ek elemzésére hoztak létre: jelentősen eltérő minták összehasonlítását, egy családhoz tartozó elemek klaszterezését, és időbeli változások felderítését vagy lokalizálását. Az alábbi szimbólumokat használjuk: az erős és gyenge szótagokra az 'S' és a 'w', a szóhatárokat a '.', a sorvégre pedig a 'l' jelet alkalmaztuk (a konkrét jeleknek csak emlékeztető funkciójuk van, bármilyen kódolás működjön). Például a „Once upon a midnight dreary, while I pondered, weak and weary”¹² verssort S. S.w.S.w.S.w.S.w.S.w.S.w.S.w.S.w.l formában fejezhetjük ki. A prozódia ilyen meghatározása (erős/gyenge a hosszú/rövid vagy a hangsúlyos/hangsúlytalan pár helyett) lehetővé teszi, hogy ugyanaz az jelrendszer kódolja az időmértékes hagyományokat, például a klasszikus latin költészetet, valamint a modern hangsúlyos verseléseket. Megjegyzendő, hogy az egyszerűség kedvéért a szavak közötti ortográfiai határokat vesszük alapul, nem pedig a fonológiai egységeket. Ez azonban nem feltétele annak, hogy a módszer működjön,

¹² Poe *A holló* című versének kezdősora. Tóth Árpád fordításában: „Egyszer egy bús éjfél táján, míg borongtam zsongva, fájón”, melynek Metronome-képlete: Sw.S.S.SSSS.S.wSw.Sw.SS.l – *A ford.*

csupán választás kérdése – olyan hagyományoknál, mint például a klasszikus vagy a neolatin, ahol jellemző a szavak összeolvadása a kiejtésben, elhagyható a szóhatár jelölése (lásd például a 26. lábjegyzetet arról, hogyan kezeltük ezt a kérdést a reneszánsz metrumok esetében).

A bioinformatikában különböző szekvenciaillesztő eljárások léteznek a DNS-láncok összehasonlítására az aktuális feladattól függően. A genomok (és a metronomok) esetében, amelyek eltérő hosszúak és nagyon különbözők lehetnek, a legjobb választás a lokális vagy helyi szekvenciaillesztés (*local sequence alignment*). Ez az eljárás két sorozat legjobban hasonlító közös szakaszát azonosítja. A globális szekvenciaillesztés ezzel szemben akkor jó választás, ha a sorozatok nagymértékben megegyeznek egymással. A lokális szekvenciaillesztésre a jól ismert Smith-Waterman algoritmust használtuk,¹³ amely egy rugalmas jelkészletből indul ki, és a hasonlóság meghatározásához jutalmakat vagy büntetéseket ad a szimbólumok egyezésekor/elérésekor. Az algoritmus közömbös a jelkészlet méretét illetően – gyakran alkalmazzák mind nukleinsav-szekvenciákra (négy szimbólum), mind fehérjékre (20 szimbólum) –, lényege, hogy az egyezéseket és eltéréseket mutató jutalom/büntetésmátrix alapján egyezési pontszámokat rendel a sorozatokhoz. Módszerünkben a legjobb helyi egyezési pontszámot a rövidebb sorozat saját pontszámával normalizáljuk, ezáltal összehasonlítható, 0 és 1 közötti távolságokat (vagy hasonlósági pontszámokat) létrehozva.

A kezdeti értékeket meghatározó jutalmakat és büntetéseket a versmértékek empirikus vizsgálatából kiindulva határoztuk meg: például a megegyező sorhosszúságokat jobban jutalmaztuk, mint az azonos erősségű szótagokat, a szó végi és a szó belseji szótagok közötti eltéréseket pedig egyáltalán nem büntettük (ugyanakkor a megegyező szóhatárokat jutalmaztuk, ami lehetővé teszi az algoritmus számára, hogy a cezúrát/sormetszetet tartalmazó sémákat egy csoportba rendezze – ehhez részletesebben lásd a *Függelék A* pontját.) A jutalom/büntetésmátrix optimalizálása érdekében egy sor osztályozást végeztünk a többnyelvű adathalmazon, majd az ábrákat elemezve értékeltük, hogy az egy metrikai sémához sorolható versek valóban egy csoportba rendeződnek-e.

¹³ T. F. SMITH és M. S. WATERMAN, „Identification of Common Molecular Subsequences”, *Journal of Molecular Biology* 147, 1. sz. (1981): 195–197, [https://doi.org/10.1016/0022-2836\(81\)90087-5](https://doi.org/10.1016/0022-2836(81)90087-5).

A Metronom-könyvtár a BioPython-könyvtár¹⁴ módszereit használja, amely a szekvenciaillesztésekhez jól használható rutinokat biztosít. A mi kódunk a mintavételezést egészíti ki egy párhuzamos módszerrel, amely egy $n \times n$ méretű távolságmátrixot állít elő a metronomként kódolt, n számú bemeneti mintából. Ezek a távolságmátrixok, egy „előre kiszámított” metrikát biztosítva, közvetlenül felhasználhatók hierarchikus klaszterezés (mint amilyen a *hclust* függvény az R-ben) vagy olyan térbeli csoportosítások bemeneteiként, mint az UMAP dimenziócsökkentő eljárás.¹⁵

ÉRTÉKELÉS

Az algoritmus teljesítményét egy nem triviális, többnyelvű feladaton értékeltük: négy nyelvből (cseh, német, orosz és klasszikus latin) és 12 metrumból származó versek mintáinak csoportosítását értékeltük, azaz hogy a különböző módszerek mennyiben képesek nyelveken átívelő metrikai csoportok azonosítására. Ezt a teljesítményt három alapmodellhez hasonlítjuk: egy egyszerű szerkesztési távolsághoz, egy lokális illesztéshez súlyozatlan mátrixszal (vagyis jutalmak és büntetések egyénileg meghatározott rendszere nélkül), valamint egy gépi tanuláson alapuló megközelítéssel, amely prozódiai n -grammok gyakoriságán alapul.

Adat és mintavétel

A modern nyelveket felölelő részhalmazunk (67 940 vers) a benne szereplő hat leggyakoribb versmértéket tartalmazza: jambikus és trochaikus pentametert, tetrametert és tetrameter-trimetert (ez utóbbi a tetrameter és trimeter változó soraiból áll, angolul néha közönséges metrumnak, *common meter*nek nevezik). A cseh részhalmaz (44 416 vers), a német részhalmaz (15 172 vers) és az orosz részhalmaz (8352 vers) is a PoëTree adatbázisból származik.¹⁶ A metrikus annotációt cseh nyelven gépi tanulással végezték, és a kimenetet kézzel

¹⁴ Peter J. A. COCK, „Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics”, *Bioinformatics* 25, 11. sz. (2009): 1422–1423, <https://doi.org/10.1093/bioinformatics/btp163>.

¹⁵ Elérhető: <https://github.com/bnagy/metronome>, installálható a pip paranccsal.

¹⁶ Petr PLECHÁČ, „PoëTree. Poetry Treebanks in Czech, English, French, German, Hungarian, Italian, Portuguese, Russian and Spanish” (Zenodo, 2023), <https://doi.org/10.5281/ZENODO.10008459>.

ellenőrizték;¹⁷ német nyelven szabályalapú algoritmus segítségével hozták létre,¹⁸ és orosz nyelven is hasonló megközelítést alkalmaztak, amikor szavak hangsúlyosztályait használták a ritmusra való következtetéshez.

A klasszikus latin részhalmoz (2588 vers) szintén e korpusz hat leggyakoribb versformáját tartalmazza: a disztichont, a hendekaszillabust, a nem strófákra bontott hexametert (ezekből kerül ki a minták és a szerzők többsége), valamint pár jambikus trimetert (másnéven *senarius*, a római komédiák verselése), choliambust (más néven sántikáló jambust vagy *scazont*) és alkaioszi strófát. A legtöbb vers egy nyelven belüli metrumból a cseh jambikus pentameter (több mint 17 000 mű), a legkevesebb pedig a latin alkaioszi verselés (38 vers, mind Horatiustól).

Mivel az egyes kategóriába eltérő számú vers tartozik, a mintavételezést 10 példányra korlátoztuk minden versmérték-nyelv kombináció esetében, ami összesen 240 darab véletlenszerű mintát eredményezett minden kiértékeléskor (a mintavételezést majd a klasszifikációt egymás után többször megisméltük). Ezután a mintákat a versmértékek alapján osztályoztuk (és azt néztük, hogy az osztályozás mennyiben felel meg a tizenkét metrumcímkének), ami körülbelül 8%-os véletlenszerű pontosságot jelent (*random baseline accuracy*).

Klasszifikációs módszer

A Metronomhoz hasonlóan működő, illesztésen alapuló módszerekhez szintén $n \times n$ nagyságú távolságmátrixokat hoztunk létre, amelyekre a k -legközelebbi szomszéd (k -Nearest Neighbours, kNN) osztályozót alkalmaztunk ($k=7$), az eredményeket pedig egy egyszerű pontossági metrikával jellemeztük („helyes”-e a csoportosítás). Ezek az illeszkedésen alapuló módszerek a következőképp működnek. A naiv szekvenciaillesztés esetében is a legjobb lokális megfeleltetést vesszük alapul, de anélkül, hogy megengednénk az úgynevezett indeleket (a megfeleltetett területeken belüli beszúrásokat vagy törléseket).¹⁹ Ez a lokális Hamming-távolsághoz hasonlatos, és egy nagyon könnyörtelen illesztés (mivel a szóhatárok különbségei általában ellehetetlenítik az identikus megfeleltetést).

¹⁷ Petr PLECHAČ, „Czech Verse Processing System KVĚTA – Phonetic and Metrical Components”, *Glottology* 7, 2. sz. (2016), <https://doi.org/10.1515/glot-2016-0013>.

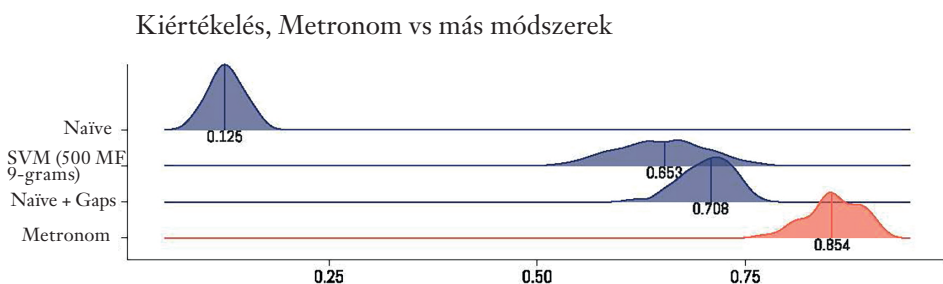
¹⁸ Klemens BOBENHAUSEN és Benjamin HAMMERICH, „Métrique littéraire, métrique linguistique et métrique algorithmique de l’allemand mises en jeu dans le programme Metricalizer 2:”, *Langages* 199, 3. sz. (2015): 67–88, <https://doi.org/10.3917/lang.199.0067>.

¹⁹ Az indel a bioinformatikában használatos kifejezés, amely az inzerció (betoldás) és delécio (törlés) szavak összevonásával jött létre. – *A ford.*

tetést), amelyet csak alapként használunk a többi módszer értékelésekor. Ezután indeleket adtunk a naiv illesztőhöz, ami így inkább egy lokális Levenshtein-távolsághoz hasonlít, de egységes egyezés/eltérés alapú jutalmakkal és büntetésekkel – nagyjából így működnek a lokális szekvencia-illesztők a DNS-láncok esetében. Végül a teljes Metronom-konfigurációt használtuk az egyénileg meghatározott, változó büntetésmátrixszal. A negyedik módszer, és az egyetlen, amely nem illesztésen alapult, egy bevett gépi tanulós megközelítés, amelyben egy SVM (*Support Vector Machine*) osztályozót alkalmaztunk a négy szimbólum 500 leggyakoribb 9-gramjának gyakoriságára, 80/20 arányban elosztva a szövegeket tanító és tesztkorpuszra. Azért 9-grammokat választottunk, mert a jelkészlet nagyon kicsi, és így az osztályozó számára előnyös, ha szinte teljes sorokat képes vizsgálni. A tesztelés során a legjobb eredményt az 500 leggyakoribb 9-grammal kaptuk, amelyeket SVD-vel 50 dimenzióra redukáltunk, majd z-értékek formájában normalizáltunk. Itt kell megjegyezni, hogy az SVM osztályozót nagyban hátráltatja a szóhatárok jelölése. A legtöbb metrum esetében a szóhatárok nem részei a metrikai megkötéseknek, csak a stílus szempontjából és néhány olyan hagyományban fontosak, ahol bizonyos helyeken kötelező vagy megszokott a cezúra (nagyon fontos például a klasszikus latinban és a modern alexandrinusokban). Más szóval, a metrum felismerésében a szóhatárok legtöbbször zajt jelentenek. Ha a cél egyszerűen egy optimális osztályozó létrehozása lenne, a gépi tanulós megközelítések – a szóhatárokat egyszerűen figyelmen kívül hagyva – legalább olyan jól teljesítenének, mint a Metronom (és sokkal gyorsabbak), ám ez az árnyaltság és az értelmezhetőség rovására menne.

Eredmények

A teljes folyamatot (mintavételezés, csoportosítás, kiértékelés) ötvenszer ismételtük meg minden módszer esetében; a teljesítmény eloszlását és mediánját az 1. ábra mutatja. Ahogy említettük, az SVM-osztályozó (általában jó viszonyítási alap) váratlanul gyenge teljesítménye főként a szóhatárok jelölésének köszönhető, miközben az is világos, hogy a Metronomban használt egyénileg meghatározott büntetőmátrix a sztenderd illesztéshez képest határozott teljesítménynövekedést eredményez nyelveken és hagyományokon átívelően. Az osztályozási pontosság azonban csak közvetve jelzi a Metronomnak mint a versek közötti metrikai kapcsolatok megértését szolgáló eszköznek valódi hasznosíthatóságát. A következő részben az eszközben rejlő további lehetőségeket néhány kisebb, bemutató jellegű példán keresztül szemléltetjük – mintegy ízelítőként.



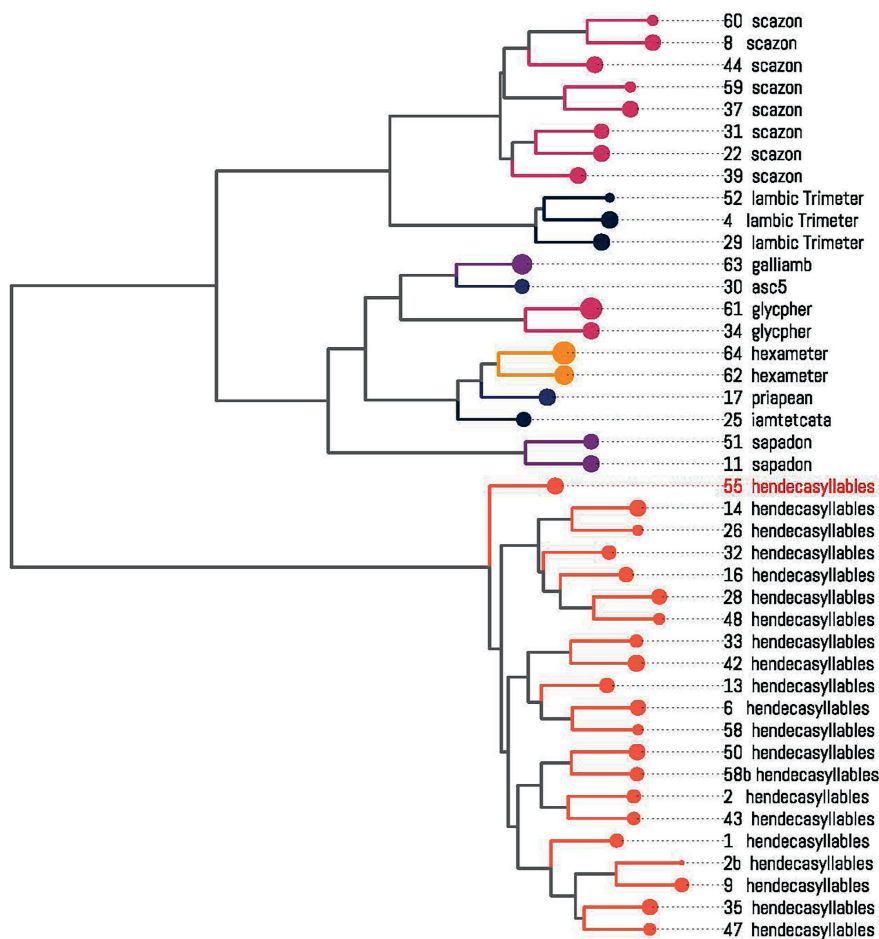
1. ábra. A Metronom osztályozási teljesítménye az SVM-mel és a másik két illesztésen alapuló módszerrel szemben. Pontossági értékek simított eloszlása 50 véletlenszerű mintavételezés után, medián pontszámokkal.

LEHETSÉGES FELHASZNÁLÁSOK

1. példa: a klasszikus latin verselés változatai

Az első példa azt szemlélteti, hogy a Metronom érzékeny a kisebb metrikai variációkra is. Catullus verseinek hierarchikus csoportosításakor vettük észre, hogy egy hendekaszillabusban írt vers, az 55. *carmen* jelentősen eltér az összes többi, ebben a metrumban írt verstől. Ez látható a 2. ábrán: az 55. *carmen* sokkal „korábban” (nagyobb eltérést jelezve) ágazik el a hendekaszillabus-csoport többi tagjától, mégis egyértelműen része ennek a csoportnak. Először azt gondoltuk, hogy ez csupán algoritmikus hiba lehet, ám közelebbről megvizsgálva, az 55. *carmen* valóban egy metrikai különlegesség. A versben Catullus a choriambusokban lévő két rövid szótagot (— ∪ ∪ —) gyakran egy hosszúra cseréli (*molossus*: — — —), ami 10 szótagos sorokat eredményez a szokásos 11 helyett (a hendekaszillabus szó szerinti jelentése ’11 szótag’). Ez látható a 3. ábrán, amely a vers elejét mutatja a Metronom átírással együtt, szembe állítva a 41. *carmen*nel, amely egy megszokott hendekaszillabus minden sorban egy standard choriambussal. Ilyen szabadságot csak kétszer enged meg magának Catullus az egész korpuszban (a másik példa az 58b), így az 55-ös vers valóban különös példány.²⁰ Pontosan ez a „laza”, családi hasonlóságból – vagyis abból, ahogyan a versmértékeket megírták és nem tankönyvi meghatározásokból – kiinduló csoportosítás mutathat rá a Metronomra

²⁰ Catullus verselésének közérthető leírása, beleértve ennek a kérdésnek a tárgyalását is, Peter Green kétnyelvű kiadásában olvasható: Peter GREEN, *The Poems of Catullus: A Bilingual Edition* (Berkeley: University of California Press, 2005), 32.



2. ábra. Catullus választott verseinek ágrajza. Az 55. *carmen*, miközben része a hendekaszillabusoknak, látványosan el is különül a csoport többi tagjától.

rejlő valódi lehetőségekre, amely ezáltal a versmértékek „távoli olvasásán” alapuló felfedezések eszköze lehet.

Ehhez a kis példához szükséges adatokat David Chamberlain biztosította, aki szabadon hozzáférhetővé tette a görög és latin költészet egy verstanilag elemzett, kiterjedt és digitalizált válogatását.²¹ Mindössze minimális feldolgozásra volt szükség ahhoz, hogy a szövegeket Metronom-sorokba konvertáljuk.

Oramus, si forte non molestum est,	SSS.S.Sw.S.wS.S.
demonstres ubi sint tuae tenebrae.	SSS.ww.S.wS.wSS.
Te Campo quaesivimus minore,	S.SS.SSwS.wSS.
te in Circo, te in omnibus libellis,	.S.SS.w.w.SwS.wSS.
te in templo summi Iovis sacrato.	.S.SS.SS.wS.wSS.
In Magni simul ambulatione	S.SS.ww.SwSwSS.
femellas omnes, amice, prendi,	SSS.SS.wSw.SS.
quas vultu vidi tamen sereno.	S.SS.SS.wS.wSS.
...	

Ameana puella defututa	SwSw.wSw.SwSS.
tota milia me decem poposcit,	Sw.Sww.S.wS.wSS.
ista turpiculo puella naso,	Sw.SwwS.wSw.SS.
decoctoris amica Formiani.	SSSw.wSw.SwSS.
Propinqui, quibus est puella curae,	wSS.ww.S.wSw.SS.
amicos medicosque convocate:	wSS.wwSw.SwSS.
non est sana puella, nec rogare	S.S.Sw.wSw.S.wSS.
qualis sit solet aes imaginosum.	SS.S.ww.S.wSwSS.

3. ábra. A sortöréssel ábrázolt Metronom-átiratok összehasonlítása: az 55. *carmen* (összevont choriambussal) és a 41. *carmen* (szabályos hendekaszillabussal) eleje.

2. példa: a „reneszánsz versmérték” elterjedése

A következő példa a reneszánsz versmérték különböző megvalósulásainak tipologikus összehasonlítására vonatkozik. A reneszánsz idején egy új szótag-számláló versmérték terjedt el Európában a korábbi formákat felváltva vagy átalakítva. Eredetileg 10 szótagos sorokból állt, amelyek az olasz költészetben 11 szótagosak lettek az utolsó előtti hangsúlyos szótag állandósulásával. Ez a hendekaszillabus Petrarca nyomán vált széles körben ismertté; és később

²¹ David CHAMBERLAIN, „Hypotactic: Greek and Latin Meter”, é. n., hozzáférés: 2025.05.02, <https://hypotactic.com/>.

ugyanebbből a metrumból – olasz és valószínűleg francia mintára – alakult ki a modern angol költészet alapformája is, a jambikus pentameter.

Minden hagyományban, amely átvette az újítást, a versmérték némi változáson ment keresztül, és alkalmazkodott a befogadó nyelvéhez és verselési stílusához, ami ugyanannak a költői formának a különböző megvalósításaihoz vezetett.²² Az új versmérték valószínűleg az okcitan hagyományból származik,²³ és onnan jutott el Itáliába. Petrarca és az *endecasillabo* európai hírneve az új metrum gyors felfutását eredményezte a többi neolatin költészetben, például a spanyolban, és az olasz félsziget olyan nyelvjárásaiban, mint a nápolyi, velencei és szicíliai. A francia költészet már korábban is használt tízszótagos versformát (például az *amour courtois* esetében),²⁴ ezt igazították később az *endecasillaboban* írt szonettek olasz trendjéhez,²⁵ mielőtt felváltotta volna az alexandrinus. Az okcitan verselés a katalán és portugál költészetre is hatással volt, amelyekben szintén létezett már hasonló, tíz szótagos verselés korábban.

A reneszánsz későbbi szakaszában az új metrikus forma elterjedt az angol és a holland költészetben is, egymástól függetlenül. Ezt követően a holland reneszánsz mérték a német és a fríz költőkre volt hatással. Az izoszillabikus (egyenlő szótagszámú sorokból álló) formákkal való érintkezésből valamilyen nyugat-germán átvétel egy szigorú jambikus metrumot hozott létre. Ez volt a modernitás kezdete az európai költészetben: az egyszerre szótagszámláló és hangsúlyos verselés (*accentual-syllabic verse*) megjelenése szinte mindegyütt Franciaországtól keletre.

A különféle reneszánsz hagyományokból származó kisméretű mintáink metrikai összehasonlítása (lásd a 4. ábrát) egyértelműen megkülönbözteti a forma átvételének imént leírt három „irányát”: azonos klaszterbe kerültek 1) az olasz félsziget költői és a spanyol verselés; 2) az okcitan, a katalán és a francia költészet; és végül 3) a germán hagyományok, amelyek a szabályos jambi-

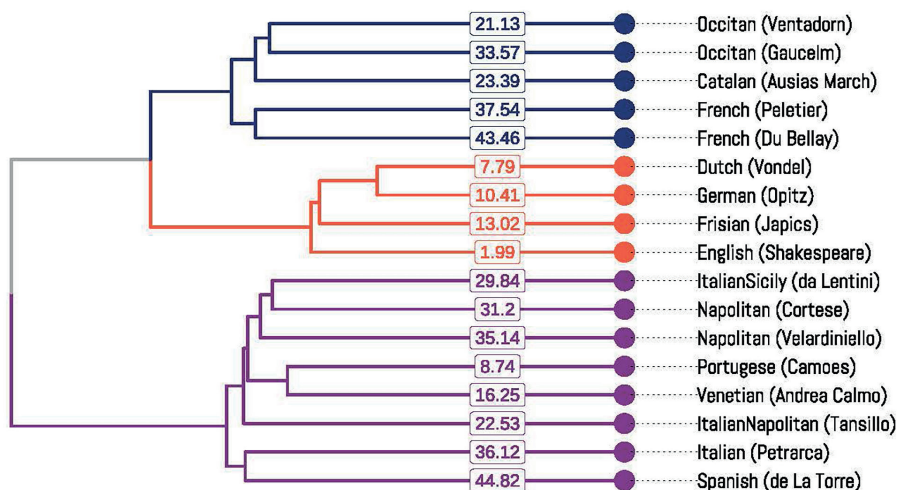
²² Mirella DE SISTO, „The interaction between phonology and metre: Approaches to Romance and West Germanic Renaissance metre” (Disszertáció, Radboud University, 2020).

²³ Pietro BELTRAMI G, „Cesura epica, lirica, italiana: Riflessioni sull’endecasillabo di Dante”, *Metrica* IV (1986): 67–107; Costanzo DI GIROLAMO és Aniello FRATTA, „I decenari con rima interna e la metrica dei Siciliani”, in *Dai Siciliani ai Siculo-toscani. Lingua, metro e stile, per la definizione del canone: Atti del convegno di Lecce (21-23 aprile 1998)*, szerk. Rosario COLUCCIA és Riccardo GUALDO (Galatina: Congedo, 1999); Dominique BILLY, „L’invention de l’endecasillabo”, in *Carmina semper et citrae cordi: Études de philologie et de métrique offertes à A. Menichetti*, szerk. Marie-Claire GÉRARD-ZAI, 31–46 (Genova: Slatkine, 2000).

²⁴ William Henry HUDSON, *A short history of French literature* (London: G. Bell and Sons LTD., 1919).

²⁵ Uo.; Sarah KEY, Terence CAVE és Malcolm BOWIE, *A short history of French literature* (New York: Oxford University Press, 2006).

kus metrumra tértek át (az ábrán szereplő értékek azt mutatják, hogy a minták mennyire „rugalmasak” a hangsúlyok elhelyezését illetően – figyeljük meg, hogy a germán minták jobban közelítenek a 0-hoz). Bár a hagyományok közötti részletesebb történeti kapcsolatok továbbra is homályosak maradhatnak a módszer számára (amely nagyban függ a mintavételezéstől és az olyan szóhatárokkal kapcsolatos technikai problémáktól, mint a *synalepha*²⁶ jelensége), összességében elmondható, hogy jól ragadja meg az azonos forma különféle megvalósulásainak szerkezeti és nyelvi hasonlóságait.



4. ábra. A reneszánsz versmérték különböző mintáinak Metronom-alapú csoportosítása. A mellékelt szám egy entrópia-alapú eltérést jelöl a szabályos metrikus formától:²⁷ Shakespeare a legszabályosabb, de La Torre a legkevésbé szabályos szerző.

²⁶ Számos neolatin költészeti hagyomány (például a spanyol és az olasz) nagymértékben alkalmazza a *synalephát* – két egymást követő szó határon lévő szótagjainak összeolvadása a kiejtésben –, ami bonyolítja a szótagszámlálást és a szóhatárok elhelyezését műveletét. Ezért kézzel annotáltuk a *synalepha* minden esetét, és az összeolvasztott szavakat egyetlen „fonetikus szónak” tekintettük, csupán utánuk elhelyezve szóhatárt. A Metronom elemzéséhez minden *synalephát* egyetlen szótagként annotáltunk: „erős”, ha az összevonás egyik oldalán volt erős (azaz hangsúlyos) szótag, és »gyenge«, ha mindkét szótag hangsúlytalan volt.

²⁷ Lásd: ŠEĽA és GRONAS, „Measuring Rhythm Regularity in Verse: Entropy of Inter-Stress Intervals”.

Ehhez a bemutatóhoz használt adatokat eredetileg Mirella De Sisto állította össze;²⁸ a minták olyan költők műveiből válogattuk, akik a reneszánsz hagyomány reprezentatív képviselőinek tekinthetők az adott nyelven. De Sisto a kézi annotálást a nyelvek fonológiájával és metrikájával foglalkozó szakértőkkel együtt végezte el.

3. példa: modern elterjedés – a szótagszámláló hangsúlyos verselés

A 18–19. században a verslábakból álló szótagszámláló hangsúlyos verselés, amely az izoszillabikus reneszánsz versemérték hatására jött létre, gyorsan elterjedt Németországból kiindulva észak és kelet felé,²⁹ és döntő szerepet játszott az újonnan kialakuló nemzeti irodalmak létrejöttében. A nagy hatású (és nagymértékben szerkesztett) népdalgyűjtemények,³⁰ például a *Volkslieder* (1786) révén a szabályos trochaikus verselés Európa-szerte a szóbeliséghez kapcsolódott, míg a jambikus verselés a német udvari költészettől a mintaképnek számító Shakespeare-drámáig a magas rangú világi irodalom ismertetőjegye lett. Ez képezte a jambikus és a trochaikus verselés hosszan tartó kulturális – és ennek következtében szemantikai – ellentétének alapját.³¹

Mint minden technológia esetében, a hódítás és a hatalom nagy szerepet játszott az új forma terjeszkedésében, és nem véletlen, hogy a „divatos” szótagszámláló hangsúlyos verselés gyorsan elterjedt az Osztrák–Magyar Monarchia és az Orosz Birodalom területén, közösségeiben és nyelveiben, többek között az irodalmi nyelvek és iskolarendszerek egységesítésének következtében.³² Az új verselés gyakran megbontotta a már meglévő formákat, vagy gyengítette a helyi alternatívákat (például a tisztán szótagszámláló verselést a lengyel, cseh és ukrán nyelvben).

²⁸ DE SISTO, „The interaction between phonology and metre. Approaches to Romance and West Germanic Renaissance metre”.

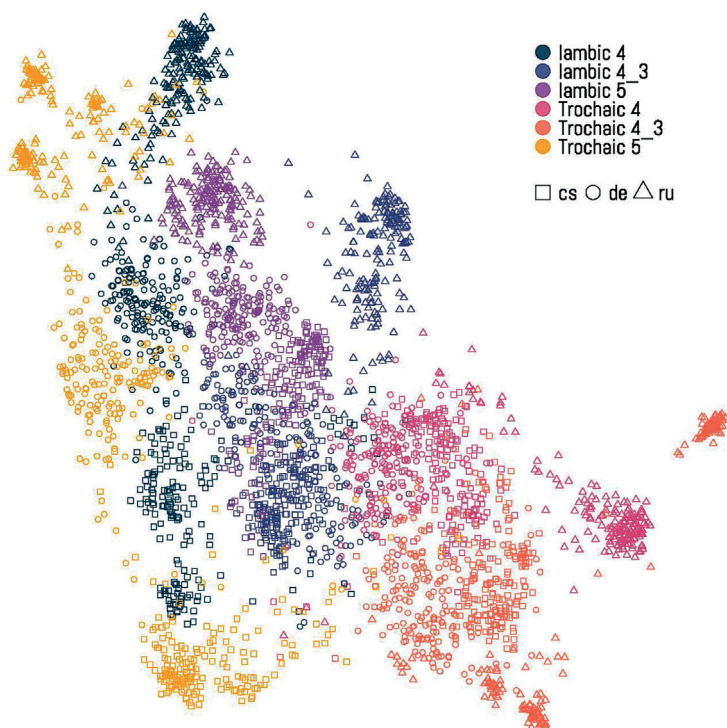
²⁹ GASPAROV, *A History of European Versification*, 207–209; Evgeny KAZARTSEV, „Iambic Verse in Different Literary Traditions”, *Glottottheory* 13, 1. sz. (2022): 3–19, <https://doi.org/10.1515/glot-2021-2004>.

³⁰ ROGER D. ABRAHAMS, „Phantoms of Romantic Nationalism in Folkloristics”, *The Journal of American Folklore* 106, 419. sz. (1993): 3, <https://doi.org/10.2307/541344>; JOEP LEERSSEN, „Oral epic: The nation finds a voice”, in *National Cultivation of Culture*, szerk. Joep LEERSSEN, Folklore and nationalism in Europe during the long nineteenth century 4, 11–26 (Leiden/Boston: Brill, 2012), https://doi.org/10.1163/9789004211834_003.

³¹ ŠEĽA, PLECHÁČ és LASSCHE, „Semantics of European Poetry Is Shaped by Conservative Forces”.

³² GASPAROV, *A History of European Versification*, 238–239.

Az utolsó példa (5. ábra) a modern versmértékek nyelvközi összehasonlításának lehetőségét szemlélteti a cseh, a német és az orosz költészetben. A módszerrel képesek vagyunk egyszerre felismerni az azonos metrumok szerkezeti egységét (színek), a trocheus és a jambus közötti jelentős különbséget (belső és külső régiók a grafikonon), ugyanakkor megőrizni a nyelvspecifikus variációkat (adatpontok formája) – ez utóbbi részben a ritmusbeli különbségekre vezethető vissza, részben pedig mert a szóhatárokkal a szavak hosszúságának nyelvspecifikus sajátosságait is kódoljuk. A német verselés, bár különböző korokban, de mindkét szláv hagyomány számára fontos forrás volt; a szerkezeti hasonlóság itt látható bizonyítékai elősegíthetik e metrikai leszármazások jobb megértését, valamint a formai kapcsolatok nyomon követését a globális irodalomtörténetben.



5. ábra. A PoëTree-korpusz 3222 cseh, német és orosz nyelvű versének UMAP-csoportjai a hat leggyakoribb európai versmértékben. A klaszterezés metrikájaként a Metronom által kiszámított távolságot használtuk.

Az 5. ábrán bemutatott adatok ugyanannak a gyűjteménynek a részhalma-zai, amelyet az értékeléskor használtunk (a nyelveken átívelő metrumfelismeréshez), és cseh, német és orosz versek annotált korpuszaiból származnak (mindegyik megtalálható a PoeTreen).

KITEKINTÉS

Végezetül a módszer korlátait és a további kutatás lehetőségeit vázoljuk. Mérőszámaink az azonos sorhosszúságtól és az ismétlődő prozódiai mintázatoktól függenek, ezért arra számítunk, hogy más elvek mentén szerveződő formák esetében a megközelítés nehézségekbe fog ütközni. Például az óangol és óészaki (*dróttkvætt*) verseket meghatározó alliterációs minták így nem azonosíthatók, és az illesztéshez egy gyökeresen eltérő jelkészletre lenne szükség. Az orosz tonikus verselésben csak a hangsúlyos szótagok száma és pozíciója kötött, a sorok hosszúsága viszont változhat, ami szintén megnehezíti az elemzést, hiszen az algoritmus érzékeny a sorok hosszára. Ugyanez mondható el az azonos versmértékű formákról, amelyek a versszakok és/vagy a rímek mintázata alapján különböztethetők meg. Például az *ottava rima* és az angol szonett egyaránt írható jambikus pentameterben, mégis különböző hagyományokat képviselnek. Ebben az esetben is a jelkészlet bővítése lehet célravezető: a strófahatárokat jelölő szimbólum hozzáadásával azonosítható lenne a versek szerkezete (de az eljárás kevésbé volna széles körben alkalmazható). Végezetül a Metronom általános módszertana (absztrakt sémák összehasonlítása) a ritmusszekvenciák helyett a rímképletekre is alkalmazható: elképzelhető például az európai szonettek történetének és tipológiájának rekonstrukciója a rímeket betűkre átíró klasszikus kódolás alapján (például abablababab|cdclcdc).

A jelenlegi megközelítésünk a cezúra – rögzített pozíciójú szóhatárok – tekintetében is hiányosnak mondható, még ha az jelentős szerepet is játszik az olyan versmértékek meghatározásában, mint például a francia alexandrinus (12 szótagú versszak, ahol a cezúra következetesen kettévágja a versszakot: 6+6). Ennek megoldására Metronom-elemzést végeztünk az alexandrinus szimulált változatain (lásd a *Függelék B* pontjában), hogy lássuk, a cezúrán alapuló formákat meg lehet-e különböztetni az egyszerű szótagoló formáktól a jelenlegi módszertannal. Az előzetes eredmények nagyon ígéretesek: a Metronom még konzervatív megszorítások mellett is hajlamos megkülönböztetni a cezúrát tartalmazó álverseket az azonos sorhosszúságú, de nem cezúrákra épülő álversektől.

Mindezen túl a módszer metrikai szempontból elemzett, nagy gyűjtemények elérhetőségétől is függ. Ekkor abba a nehézségbe ütközünk, hogy az elemzés gyakran a meglévő verstanok függvénye, amelyek csupán egy közösségen belül érvényesek, és a nemzeti hagyományok verstanai között ritkán van konszenzus. És habár verstantól független elemzés aligha lehetséges, az elméleti előfeltevések az automatizált elemző rendszerekben önbeteljesítő módon működhetnek: az empirikus mintákat arra kényszerítik, hogy illeszkedjenek néhány előre meghatározott sémához,³³ ami a formák változatosságát korlátozza vagy akár el is fedi.

Éppen ezért már korábbi munkáinkban is kiálltunk az alulról felfelé történő metrumfelismerés módszere mellett. Ugyanis minden elégszer megismételt minta, minden kellően erős szervezőelv tetten érhető előzetes meghatározottságok nélkül is; egy szisztematikus metrum természetéből adódóan minden megfigyelésben vagy mérésben megmutatja magát, vonatkozzon az a szöveg prozódiaja, morfológiája vagy szintaxisára.³⁴ A versmértékek lényegében a szabályismétlés meglehetősen egyszerű technológiái: minél többször ismétlünk meg egyet, annál jobban láthatóvá válik még a miénkhez hasonló durva eszközök számára is. Ez a tulajdonság teszi még a „szuboptimális” – a mögöttes elmélet és hagyomány szakértői ismerete nélkül készített – automatikus verselemzőket is használatóvá és hasznossá.

Összefoglalva, tanulmányunk azt mutatja meg, hogy a metrikus sorok összehasonlítása a legkülönbözőbb kontextusokban is fontos tanulságokkal szolgálhat – a bevett metrumok variációinak nyomon követésétől a nyelveken és korokon átívelő formák nagy csoportjainak felismeréséig. Az általunk javasolt módszer egyaránt jól működik hosszú és rövid szekvenciák esetében, és a költői hagyományok széles skáláján alkalmazható (példáinkban három fő hagyomány szerepel: időmértékes, szótagszámláló és szótagszámláló hangsúlyos). Mivel az eljárás a verselés mind nagyobb, mind kisebb szerkezeti különbségeire érzékeny, jól alkalmazhatónak tűnik a történeti és a nyelvek közötti összehasonlító kutatásban.

³³ Áttekintéshez lásd: Mirella DE SISTO, „Understanding Poetry Using Natural Language Processing Tools: A Survey”, *Digital Scholarship in the Humanities* 39, 2. sz. (2024): 500–521, <https://doi.org/10.1093/llc/fqae001>.

³⁴ Mikhail Leonovich GASPÁROV és Marina TARLINSKAJA, „The Linguistics of Verse”, *The Slavic and East European Journal* 52, 2. sz. (2008): 198–207, <https://doi.org/10.2307/20459662>.

ADATOK ELÉRHETŐSÉGE

A tanulmányhoz kapcsolódó repozitórium elérhető az alábbi címen: <https://github.com/bnagy/metronome-paper>. Minden kódra és adatra CC-BY licenz vonatkozik, kivéve, ha ezt korábbi licenszek korlátozzák. A repozitórium az értékeléshez szükséges adatokat és kódokat tartalmazza, illetve egyéb kiegészítő ábrák és magyarázatok is megtalálhatók benne.

FÜGGELÉK A

Röviden részletezzük a jutalom/büntetésmátrix létrehozásának elvét, amely megkülönbözteti a Metronomot az alapértelmezett Smith-Waterman illesztéstől (az algoritmus legtöbb leírásában ezt helyettesítési mátrixnak nevezik). DNS-szekvenciák esetén, ahol négy címke használatos (A, G, C, T), alapértelmezés szerint az egyezésért 1 pont jár, az eltérésért pedig –1 pont levonás. A fehérjeszekvenciák illesztésénél – mivel egyes fehérjék „közelebb állnak” egymáshoz – többféle mátrix is használatos, amelyek az aminosavak páronkénti egyezését különbözőképpen értékelik, de ezek itt most nem relevánsak. A négy Metronom-szimbólum alkotta sorok összehasonlításakor abból indultunk ki, hogy a sor végén lévő egyezés/eltérés a legfontosabb, ezért az algoritmus előnyben részesíti azokat a verspárokat, amelyek hosszukban és lezárásukban is megegyeznek. Ezután nagyobb jutalmat adtunk az erős szótagok (S) egyezéséért, mint a gyenge szótagokért (w), végül pedig úgy döntöttünk, hogy kis jutalmat adunk az egyező szóhatárokért (.) – de nem büntetjük az eltéréseket –, hogy segítsük azokat a versek közti illesztéseket, amelyek rögzített cezúrákat tartalmaznak. Ezt követően több próbát is lefuttattunk, és az algoritmust a versek csoportosításának kézi ellenőrzése során továbbhangoltuk. Ez egy alapvetően szubjektív folyamat, amely a saját szakértelmünkre és a használt korpuszokra épül: könnyen elképzelhető, hogy más elemzések pontosabb eredményt adnak, ha a helyettesítési mátrixot az adott kutatási célhoz igazítják – ami könnyedén megtehető a közzétett szoftvercsomagban.

FÜGGELÉK B

A modern versmértékek elemzéséhez ebben a tanulmányban nem volt szükség cezúrák – a sorokon belüli szisztematikus szóhatárok, és egyúttal gyakran szintaktikai szünetek – azonosítására, amelyek néhány hagyományban viszont

segíthetik a versmértékek meghatározását, mint például a szótagszámláló francia vagy a klasszikus latin költészetben. A bevett modern példa a francia 12 szótagos alexandrinus egy erős szótagot követő központi cezúrával. Ennek kapcsán fontos tudni, hogy a cezúrák jelenlétét vagy elhelyezését is figyelembe tudja-e venni a Metronom a különböző formák elkülönítésekor. Ennek tesztelésére ötféle pszeudo-verselést generáltunk, amelyek – bár meglehetősen naiv módon – a valódi versek prozódiaját utánozzák.

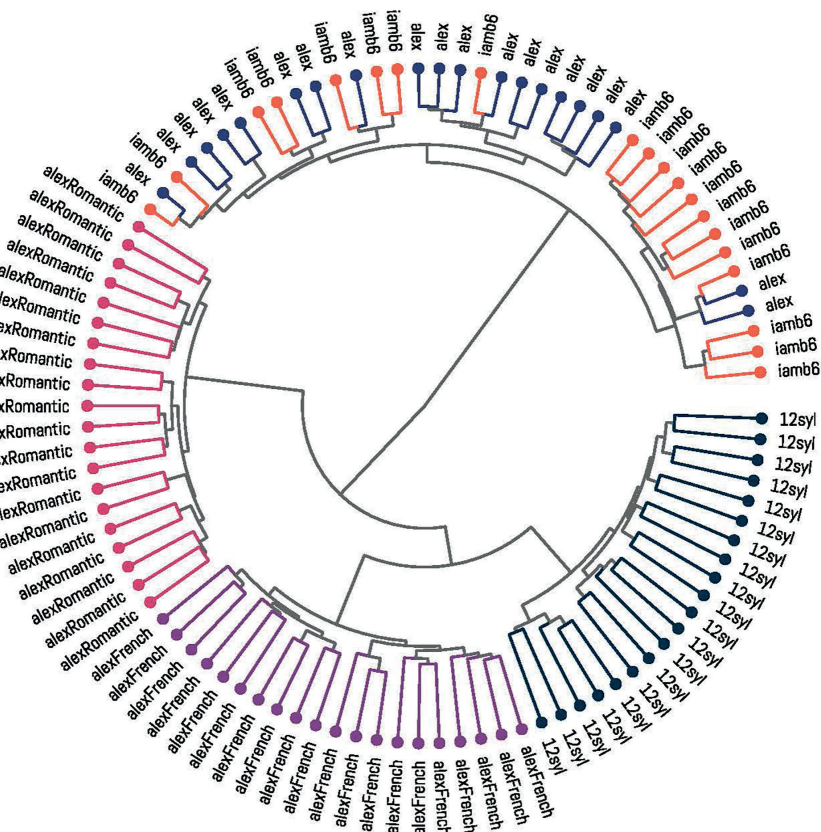
A generált versek hossza Poisson-eloszlást követ ($\lambda = 14$), így a mesterséges tesztünkben a versek átlagosan 14 sorosak, azaz szonettnyi hosszúak.

1. *Jambikus hexameter alexandrinus cezúrával (alex)*: minden sor szigorú mintát követ: „wSwSwS.wSwSwS”, a hatodik szótag után kötelező szóhatárral. Minden egyéb szóhatárt véletlenszerűen határoztunk meg a szóhosszúságok valószínűségi eloszlása alapján.
2. *Sima jambikus hexameter (Iamb-6)*: ugyanaz, mint az előző, de meghatározott cezúra nélkül; a szóhatárok véletlenszerű kijelölésével.
3. *Klasszikus francia alexandrinus (alexFrench)*: minden sor az alábbi általános mintát követi: „xxxxxS.xxxxxS”, ahol ‘x’ egy tetszőleges (‘w’ vagy ‘S’) szimbólum; a szóhatárok véletlenszerű kijelölésével mindegyik hemisztichionban (felsorban).
4. *Romantikus alexandrinus (alexRomantic)*: minden sor az alábbi általános mintát követi: „xxxS.xxxS.xxxS”; ez a forma Victor Hugóhoz és a klasszikus metrum romantikus továbbfejlesztéséhez köthető.
5. *Egyszerű 12 szótagos versmérték (12syl)*: egy elképzelt versmérték a sorhosszúságon kívül mindenféle megkötés nélkül.

A sorok létrehozásakor ezeket az egyszerű feltételeket tartottuk szem előtt, és nem vettük figyelembe semelyik valós nyelv szerkezetét, ahogyan nem próbáltuk a szavak hosszának eloszlását és a hangsúlyok helyét a szavakon belül sem hihetően reprodukálni (így például előfordulhatnak valószínűtlen szavak olyan prozódiai szerkezetekkel, mint a „ww” vagy az „SSS”). A szóhatárok kijelöléséhez 1-, 2-, 3- és 4-szótagos szavakat alkottunk 1:3:1:0.25 valószínűségi arányban. Ez az arány 45%-os valószínűséggel eredményez szóhatárt a hatodik szótag után a nem alexandrinus sorokban is, pusztán véletlenszerűen. Ez kellően magas érték ahhoz, hogy a Metronom igazolni tudja érzékenységét a szabályszerű különbségek azonosításában.

Bár ez a forgatókönyv nem reális, mégis átlátható, és lehetővé teszi, hogy a nyelvi szabályoktól függetlenül elemezhessük a versmértékek szervezőelvét. A 6. ábra a „versek” így kapott csoportosítását mutatja: jól látható, hogy a cezú-

rát alkalmazó formák megkülönböztethetők azoktól, amelyek nem alkalmaznak cezúrát. Pontosabban az elkülönítés a szótagszámláló verselések (alex-French, alexRomantic és 12syl) esetében működik jobban, mivel a jambikusokkal (alex, Iamb-6) ellentétben ezeknél van egy további megkülönböztető jegy: a két rögzített hangsúlyos pozíció a megkötés nélküli 12 szótagos sorokkal szemben.



6. ábra. Az alexandrinus öt változatát reprezentáló szimulált álversek csoportosítása: három szótagszámláló és két szótagszámláló hangsúlyos verselés.

További tesztek is megerősítik ezt az eredményt: 100 alkalommal, minden egyes formához 20 verset generálva, értékeltük a csoportosítás hatékonyságát a korrigált Rand index (ARI) segítségével. A medián ARI a jambikus versek-

nél 0.55, ami a cezúrán alapuló formák némileg pontatlan, de a véletlenszerű-nél hatékonyabb elkülönítését jelzi. A szótagszámláló versek esetében a medi-án $ARI = 0.95$, vagyis majdnem hibátlan a három variáció elkülönítése. Ezért gondoljuk azt, hogy a cezúrán alapuló formák szervezőelve még nagyon zajos környezetben is azonosítható módszerünkkel.

NAGY, Ben, Artjoms ŠELA, Mirella DE SISTO és Petr PLECHÁČ. „Metronome: tracing variation in poetic meters via local sequence alignment”. *Computational Humanities Research*, 2025. 1:e1, <https://doi.org/10.1017/chr.2025.1>.

Fordította: Szemes Botond

HORVÁTH PÉTER

Szintaktikai ismétlődés a magyar kanonikus költészetben¹

horvath.peter@btk.elte.hu
ORCID: 0000-0002-3517-5623

HELIKON

Syntactic Repetition in Canonical Hungarian Poetry

Abstract

This paper uses corpus linguistic methods to investigate syntactic repetition in canonical Hungarian poetry. The research corpus consists of grammatically and metrically annotated poems of 24 canonical Hungarian authors from the 19th and the first half of the 20th century. To measure the degree of syntactic repetition, the number of different part-of-speech 2-grams, 3-grams, and 4-grams is counted in different window sizes. By comparing the syntactic repetition scores of poems with the scores obtained for a reference corpus of Hungarian novels, the paper demonstrates that syntactic repetition is an important characteristic of canonical Hungarian poetry. The paper also investigates whether there is a correlation between syntactic repetition and phonological repetition in the poems. On the one hand, it argues that the increase in syntactic repetition in the 20th century compensates for the decrease in phonological repetition. On the other hand, it presents that qualitative meter poems based on bars have higher syntactic repetition scores than iambic poems based on feet. In addition, by calculating the overlap coefficient between the part of speech n-grams of lines, the paper shows that consecutive rhyming lines are more similar syntactically than consecutive non-rhyming lines.

Keywords: canonical Hungarian poetry, ELTE Poetry Corpus, ELTE Novel Corpus, syntactic repetition, poetic function, moving-average type-token ratio, part of speech n-gram, meter, rhyme

¹ A tanulmány az NKFIH K 137659 sz. pályázat (*A személyjelölés konstrukcióinak korpusz-alapú, kognitív poétikai vizsgálata*) keretében, illetve a Digitális Örökség Nemzeti Laboratórium támogatásával készült.

1. BEVEZETÉS

A ritmussal, a rímeléssel, az alliterációval és akár egyéb eszközökkel megvalósított fonológiai ismétlődés a költészet egyik jellegadó, bár nem feltétlenül szükséges tulajdonsága. A fonológiai ismétlődés mellett ugyanakkor az ismétlődés más típusai is fontos szerepet játszhatnak. A 18. század közepén, az óhéber költészetet vizsgálva már Robert Lowth rámutatott a párhuzamosságok egyéb típusaira.² Megfigyelte például, hogy az óhéber költészetben a párhuzamosság egyik gyakori megvalósítása egy korábbi, rendszerint az előző sor szintaktikai szerkezetének a megismétlése. Ahogyan arra Horváth Iván felhívta a figyelmet, a fonológiai ismétlődés mellett az ismétlődés egyéb típusait meghatározónak valló elképzelések a magyar verselméletben is már igen korán megjelentek.³ Ennek első példaként Földi János 1792-es írására utal, amely Lowth elgondolását mutatja be.⁴ Ebben Földi amellet érvel, hogy az óhéber költészet fogalmi és szintaktikai párhuzamosságokra épülő verselése az időmértékes és a rímes verselésnél alkalmasabb technika, illetve hogy a Lowth által megállapított párhuzamosságok egyik típusa a latin nyelvű költészetben is fellelhető.⁵ Horváth Arany Jánosnak *A magyar nemzeti vers-idomról* című verstani munkáját is a Lowth-féle elgondolás folytatásának tekinti. Az ismétlődés szisztematikusabb vizsgálata azonban csak a 20. századi strukturalista irodalomtudományban vált hangsúlyosabbá. Jakobson például több írásában is a grammatikai ismétlődést és variációt bizonyos versek, illetve költészeti hagyományok egyik legfontosabb jellemzőjének tekinti.⁶

A tanulmány egyrészt arra keresi a választ, hogy kvantitatív, korpusznyelvészeti eszközökkel igazolható-e a grammatikai ismétlődés egy fajtájának, a Lowth által is vizsgált szintaktikai ismétlődésnek a kitüntetett szerepe a magyar kanonikus költészetben. A tanulmány emellett azt is vizsgálja, hogy a versekben megjelenő szintaktikai ismétlődés mértéke összefügg-e a versek metrumával és a rímeléssel. Ez utóbbi vizsgálatok tehát arra az általános kér-

² Robert LOWTH, *Lectures on the Sacred Poetry of the Hebrews*, ford. G. Gregory (Boston–New York: Codman Press, 1829), 154–166.

³ HORVÁTH Iván, „A grammatikai szemlélet kezdetei a magyar verselméletben (Földitől Aranyig)”, *Irodalomtörténeti Közlemények* 76, 3. sz. (1972): 290–306; lásd még Horváth Iván, *A vers* (Budapest: Gondolat Kiadó, 1991), 50–88.

⁴ FÖLDI János, „Elmélkedés a ’Sidó Vers-írásról”, *Magyar Hírmondó* 1, 2. köt. (1792): 905–914.

⁵ Uo., 910–911.

⁶ Roman JAKOBSON, „A grammatika poétikája és a poétika grammatikája”, in Roman JAKOBSON, *Hang – Jel – Vers*, 258–277 (Budapest: Gondolat Kiadó, 1969); Roman JAKOBSON, „Grammatikai párhuzamosság a népköltészetben”, in JAKOBSON, *Hang – Jel – Vers*, 347–371.

désre keresik a választ, hogy a szintaktikai ismétlődés és a fonológiai ismétlődés megvalósulása egymástól függetlenül vagy pedig egymással összefüggésben jön-e létre. Vagy még általánosabban megfogalmazva, arra próbálnak választ adni, hogy van-e valamilyen motivációs viszony a hangzás és a grammatika szintjén megvalósuló ismétlődés között. Mindezek mellett a tanulmány a szintaktikai ismétlődés időbeli alakulását is vizsgálja, amit összefüggésbe hoz a metrumnak, azaz a versekben megjelenő fonológiai ismétlődés egyik fő típusának az időbeli alakulásával.

Nem fonológiai jellemzők magyar nyelvű versekben való ismétlődését már vizsgálták automatizáltan, kvantitatív módszertannal: Seláf Levente és szerzőtársai 16. századi históriás énekekben karakter 2-gramok alapján vizsgálták a versekben megjelenő epikus formulákat, amelyek az értelmezésükben egy sornál nem hosszabb, szintaktikailag és lexikailag azonos vagy hasonló, sok esetben szó szerinti vagy közel szó szerinti ismétlődő szövegrészek.⁷ Az alább bemutatandó vizsgálatok az ismétlődés jelenségét absztraktabb szinten, a szintaxis szintjén próbálják megragadni. A vizsgálatok így eltérő módszertant kívánnak meg, hiszen az a cél, hogy a szintaktikai ismétlődésnek azokat az eseteit tegyük mérhetővé, amelyek nem a szó szerinti vagy közel szó szerinti ismétlődésből következnek. Ez utóbbiak itt inkább a mérést torzító, és ezért valamiképpen kiszűrendő tényezőkként jelentkeznek.

Az irodalomnak és más művészeti ágaknak is egyik meghatározó jellemzője az egyes alkotások között, illetve az egyes alkotásokon belül megvalósuló ismétlődés (az előbbire példa a műfaji konvenciók ismétlése).⁸ Ugyanakkor az ismétlődés nem csupán az irodalomnak, hanem magának a nyelvnek is az egyik alapjellemzője. Ugyanazok a szavak, szókapcsolatok akár egy rövidebb szövegen belül is gyakran többször megismétlődnek. Ha a szavaknál absztraktabb nyelvi kategóriák, például szófaji kategóriák szövegbeli előfordulását nézzük, akkor az ismétlődés értelemszerűen még szembetűnőbbé válik. Az ismétlődés nem csupán a nyelv lexikai és grammatikai szintjén, hanem a nyelv hangzásszintjén is megjelenik. A fonológiai ismétlődés olyan eljárásai, mint

⁷ SELÁF Levente, Petr PLECHÁČ, VIGYIKÁN Villő és KISS Margit, „A belső formulák a 16. századi históriás énekekben”, in *A históriás ének: Poétikai és filológiai kérdések*, szerk. SELÁF Levente (Budapest: Gépeskönyv, 2023), hozzáférés: 2025.10.14, <https://f-book.com/book/2023/A-historias-enek-kerdesek/index.php?chapter=8>. A cikkben szerepel egy részletes definíció arra vonatkozóan, hogy mit tekintenek a szerzők formulának.

⁸ Az ismétlődés fajtáiról és jelentőségéről például lásd SZEGEDY-MASZÁK Mihály, „A művészi ismétlődés néhány változata az irodalomban és a zenében”, in *Ismétlődés a művészetben: Tanulmányok*, szerk. HORVÁTH Iván és VERES András, 77–159 (Budapest: Akadémiai Kiadó, 1980).

a rím vagy az alliteráció, nemcsak irodalmi szövegekben, hanem a hétköznapi nyelvhasználatban is tetten érhetők.⁹ Mivel a nyelvre az ismétlődés, illetve annak itt vizsgált típusa, a szintaktikai ismétlődés általánosan jellemző, a versekben megjelenő szintaktikai ismétlődést önmagában nem vizsgálhatjuk, hiszen más típusú szövegekkel való összevetés nélkül nem tudjuk megmondani, hogy a versek esetében megállapított mértéke az ismétlődésnek a vers szöveg-típusához kapcsolható specifikusabb jellemző, vagy csupán a nyelv általános tulajdonságát tükrözi. Ezt a problémát egy regényeket tartalmazó referenciakorpusz használatával kezelem. A referenciakorpusz alapján már elhelyezhetjük a versek szintaktikai ismétlődésének a mértékét jelző értékeket.

A 2. rész röviden bemutatja a vizsgálatokhoz használt korpuszokat és a korpuszok felhasznált annotációs rétegeit. A 3. rész ismerteti a vizsgálatokban használt módszereket. A 4. rész mutatja be, hogy a vizsgálati korpuszban szereplő 24 kanonikus költőre milyen mértékben jellemző a szintaktikai ismétlődés jelensége. Az eredményekből kiderül, hogy a versekben általában nagyobb mértékben jelenik meg a szintaktikai ismétlődés, mint a regényekben, illetve az is kirajzolódik, hogy a 20. század első felének a szerzőinél többnyire nagyobb mértékben jelenik meg a szintaktikai ismétlődés, mint a 19. században alkotó szerzőknél. Ez utóbbi összefüggésbe hozható a versmetrum 20. századi fellazulásával: amellet érvelek, hogy a metrumban megnyilvánuló fonológiai ismétlődés mértékének a csökkenését a szintaktikai ismétlődés mértékének a növekedése kompenzálja. Az 5. és a 6. rész a fonológiai ismétlődés és a szintaktikai ismétlődés közötti további összefüggéseket vizsgálja. Az 5. rész bemutatja, hogy a szerzők többsége esetében az ütemhangsúlyos metrumú versekben a szintaktikai ismétlődés erőteljesebben jelenik meg, mint a jambikus metrumú versek esetében, ami az ütemek és a verslábak eltérő jellemzőivel magyarázható. A 6. rész pedig rámutat arra, hogy a közvetlenül egymást követő rímelő sorok szintaktikai szempontból jobban hasonlítanak egymásra, mint a közvetlenül egymás után álló nem rímelő sorok, ami a szintaxis szintjén érvényesülő absztrakt tartalmi jellemzők és a rímeléssel megvalósuló fonológiai jellemzők közötti ikonikus viszonyal magyarázható. A 7. rész röviden összegzi a vizsgálat eredményeit.

⁹ Például lásd Réka BENCZES, *Rhyme over Reason: Phonological Motivation in English* (Cambridge: Cambridge University Press, 2019), <https://doi.org/10.1017/9781108649131>.

2. A VIZSGÁLAT KORPUSZAI

A vizsgálat 24 magyar kanonikus költő összes versén alapul. A szerzők a 19. században és a 20. század első felében alkottak. A verseket a hozzájuk tartozó annotációkkal együtt az ELTE Verskorpusból szedtem ki.¹⁰ Az ELTE Verskorpuszban szereplő azon 19. és 20. századi szerzőknek a verseit vettem bele a vizsgálatba, akik legalább 100 verset írtak. A vizsgálati korpuszban szereplő költőket és a verseik számát az 1. táblázat mutatja be. A táblázatban külön oszlopokban feltüntettem a legalább 100, legalább 50 és legalább 25 tokenből álló versek számát is. A tokenszám a versekben található szavak és központosítások együttes számára utal. Azért tüntettem fel külön e három terjedelmi kategóriába tartozó versek számát, mert a szintaktikai ismétlődés mérését minimum 100, 50, illetve 25 tokenes ablakméretekkel végeztem el. (Ennek módszertanát a következő rész mutatja be.)

1. táblázat. Vizsgálati korpusz

Költő	Versek száma				Költő	Versek száma			
	Min. 100	Min. 50	Min. 25	Össz.		Min. 100	Min. 50	Min. 25	Össz.
Csokonai Vitéz Mihály	277	338	384	394	Komjáthy Jenő	205	220	236	246
Berzsenyi Dániel	104	122	130	136	Ady Endre	700	1075	1116	1116
Kisfaludy Károly	78	109	120	123	Kaffka Margit	94	101	102	102
Kölcsy Ferenc	79	129	145	150	Somlyó Zoltán	260	377	379	379
Vörösmarty Mihály	292	424	542	662	Juhász Gyula	571	1190	1251	1278
Arany János	252	315	375	417	Babits Mihály	431	502	512	514
Tómpa Mihály	436	474	490	491	Kosztolányi Dezső	405	582	615	630
Petőfi Sándor	529	738	823	839	Tóth Árpád	310	407	435	451
Madách Imre	208	245	283	319	Reményik Sándor	475	639	665	667
Gyulai Pál	114	152	156	156	József Attila	345	552	581	599
Vajda János	168	190	196	199	Dsida Jenő	583	987	1049	1060
Revczky Gyula	238	317	334	335	Radnóti Miklós	175	272	296	298

Az ELTE Verskorpusz tartalmazza a versek szövegeit és a szövegek automatikusan létrehozott annotációit. A korpuszban annotálva vannak a versek szerkezeti egységei, azaz a versszakok és a verssorok, a szavak grammatikai

¹⁰ HORVÁTH Péter, KUNDRÁTH Péter és INDIG Balázs, „ELTE Verskorpusz – a magyar kanonikus költészet gépileg annotált adatbázisa”, in *XVIII. Magyar Számítógépes Nyelvészeti Konferencia*, szerk. BEREND Gábor, GOSZTOLYA Gábor és VINCZE Veronika, 375–388 (Szeged: Szegedi Tudományegyetem TTIK, Informatikai Intézet, 2022). A korpusz innen letölthető: <https://github.com/ELTE-DH/poetry-corpus>.

jellemzői, valamint vershangzásra vonatkozó jellemzők. A szavak grammatikai annotációi az e-magyar programmal lettek létrehozva.¹¹ A korpusz többek között minden szó esetében szerepelteti a szavak lemmáját, azaz a szótári alakját, a szófaji címkéjét és a morfoszintaktikai jellemzőit a Universal Dependenciesnek a magyar nyelvre alkalmazott taxonómiája alapján.¹² A szintaktikai ismétlődés mérésének alapját ezek a szófaji és morfoszintaktikai jellemzők adják. A vershangzásra vonatkozó, úgyszintén automatizáltan létrehozott annotációkból a rímképletre¹³ és a metrumra¹⁴ vonatkozó annotációkat használja a vizsgálat. A versek négy időmértékes metrumtípus valamelyikébe, a daktilikus, az anapesztikus, a jambikus vagy a trochaikus metrumokba sorolódnak, és a metrumcímke mellett rendelkeznek egy 0 és 1 közötti szabályossági pontszámmal is, amely azt fejezi ki, hogy a vers ritmusa milyen mértékben realizálja a vers elvont metrumképletét. A vizsgálati kérdéstől függően 0,3 vagy 0,5 küszöbértéket használtam az időmértékes versek csoportjának kijelöléséhez. A korpuszban az időmértékes metrum mellett a versek esetleges ütemhangsúlyos metruma is fel van tüntetve.¹⁵

A vizsgálat referenciakorpuszát az ELTE Regénykorpusz adta, amely a vizsgálat idején 399 regényt tartalmazott a 19. századból és a 20. század első felé-

¹¹ VÁRADI Tamás, SIMON Eszter és SASS Bálint, „Az e-magyar digitális nyelvfeldolgozó rendszer,” in *XIII. Magyar Számítógépes Nyelvészeti Konferencia*, szerk. VINCZE Veronika, 49–60 (Szeged: Szegedi Tudományegyetem, Informatikai Intézet, 2017); NOVÁK Attila, REBRUS Péter és LUDÁNYI Zsófia, „Az emMorph morfológiai elemző annotációs formalizmusa,” in *XIII. Magyar Számítógépes Nyelvészeti Konferencia*, szerk. VINCZE Veronika, 70–78; INDIG Balázs, SASS Bálint és SIMON Eszter, „emtsv – Egy formátum mind felett,” in *XV. Magyar Számítógépes Nyelvészeti Konferencia*, szerk. BEREND Gábor, GOSZTOLYA Gábor és VINCZE Veronika, 235–247 (Szeged: Szegedi Tudományegyetem TTIK, Informatikai Intézet, 2019).

¹² VINCZE Veronika, SIMKÓ Katalin, SZÁNTÓ Zsolt és FARKAS Richárd, „Universal Dependencies and Morphology for Hungarian – and on the Price of Universality,” in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 356–365 (Valencia, Spain: Association for Computational Linguistics, 2017), <https://doi.org/10.18653/v1/E17-1034>.

¹³ HORVÁTH Péter, „Egy automatikusan generált rímszótár fejlesztése és a magyar kanonikus költészet rímszavainak néhány jellemzője,” in *Új technológiákkal, új tartalmakkal a jövő digitális transzformációja felé. 32 Workshop: országos konferencia*, szerk. TICK József, KOKAS Károly és HOLL András, 77–84 (Budapest: HUNGARNET Egyesület, 2023), <https://doi.org/10.31915/NWS.2023.12>.

¹⁴ HORVÁTH Péter, „Két eljárás magyar nyelvű versek metrumának gépi felismertetéséhez”, *Digitális Bölcsészet* 4 (2021): 79–103, <https://doi.org/10.31400/dh-hun.2021.4.2361>.

¹⁵ Az időmértékes és ütemhangsúlyos metrum felismertetésének, valamint az időmértékes metrumhoz kapcsolódó szabályossági pontszám kiszámításának részleteit lásd HORVÁTH Péter, „Két eljárás...”

ből.¹⁶ A korpuszban a regények szövegei és szerkezeti annotációi mellett szerepelnek a szavak lemmájára, szófajára és morfoszintaktikai jellemzőire vonatkozó annotációk is. Ezek az ELTE Verskorpuszhoz hasonlóan az e-magyar programmal lettek létrehozva, ugyanúgy a Universal Dependencies címkekészlete alapján. A két korpusz azonos programmal, illetve azonos címkekészlettel megvalósított annotálása lehetővé teszi ezek kvantitatív alapú összevetését.

3. A VIZSGÁLAT MÓDSZEREI

A szintaktikai ismétlődés mérésének alapja a szövegben egymást követő szavak szófajai által alkotott szófajcsoportok ismétlődésének a mérése. Például, ha a névelő-melléknév-főnév szófajcsoport vagy az ige-névelő-főnév szófajcsoport egy szövegrészen belül többször előfordul, akkor az növeli a szintaktikai ismétlődésnek a mért értékét. Ezt a megközelítést tehát valójában pontosabb lenne úgy definiálni, mint amely a szöveg szintagmatikus tengelyén méri szófajok szomszédsági viszonyainak az ismétlődését. Ugyanakkor a kutatás abból indul ki, hogy szófajok közötti szomszédsági viszonyok ismétlődése sok esetben mondattani értelemben vett szintaktikai viszonyok, azaz függőségi viszonyok ismétlődésére utal. Ha pedig bizonyos morfoszintaktikai jellemzőkkel, például az esetalakkal kiegészítjük a szófaji címkéket, akkor a szófajok közötti szomszédsági viszonyok és a függőségi viszonyok ismétlődése közötti kapcsolat még közvetlenebbé tehető. Mindazonáltal hangsúlyozandó, hogy a szófajok szomszédsági viszonyain alapuló eljárás azokban az esetekben, ahol a függőségi viszonyok különböző szórendben is megjelenhetnek (például az ige és különböző funkciójú főnévi bővítményei között), a szintaktikai ismétlődést akkor képes csak hatékonyan megragadni, ha nem csupán a függőségi viszonyok, hanem (valamilyen szinten) a szórend terén is megvalósul az ismétlődés. Ez ugyanakkor ebben a vizsgálatban nem feltétlenül hátrány, hiszen versek esetében a szintaktikai szerkezet ismétlésének poétikai eszköze általában nem csupán a függőségi viszonyok, azaz az absztrakt szintaktikai szerkezet megismétlését, hanem a szórend (bizonyos szintű) megismétlését is megkívánja, mivel ez utóbbi nélkül az olvasó jó eséllyel nem érzékelné az ismétlődést.

Az alábbiakban részletesen bemutatom a kutatásban használt, a szintaktikai ismétlődés mérésére kidolgozott módszereket. A kialakított módszerek alapján megvalósított méréseket automatizáltan, programkód segítségével végeztem el.

¹⁶ A korpusz innen letölthető: <https://github.com/ELTE-DH/regenykorpusz>.

3.1. A szintaktikai ismétlődés mértékének mérése

A vizsgálat fő módszerét a típus-token arány mozgó átlagát mérő eljárás módosított, szófaji címkékre és n-gramokra alkalmazott verziója adja. A típus-token arány mozgó átlagának a mérését a szókincs változatosságának a mérésére dolgozták ki. Egy szöveg szókinszének a változatosságát sokféle módszerrel próbálták már mérni. Ezek közül a legegyszerűbb a típus-token arány megadása,¹⁷ azaz a szövegben szereplő különböző szavak számának az elosztása a szövegben szereplő összes szó számával. Nem nehéz ugyanakkor belátni, hogy ezzel a módszerrel eltérő hosszúságú szövegeket nem lehet összevetni, hiszen minél hosszabb egy szöveg, annál kisebb eséllyel fognak benne új, a szöveg korábbi pontján még nem szereplő szavak szerepelni.

Ennek a problémának az egyik megoldása, hogy felosztják a szöveget egymást követő, egyenlő szószámú szegmensekre, a szegmenseknek külön kiszámolják a típus-token arányait, majd az így kapott értékeket átlagolják.¹⁸ Ennek az eljárásnak két problémája van. Egyrészt mivel a szövegek szószáma többnyire nem egész számú szorzata a szegmensek előre definiált szószámának, a szöveg vége kimarad a vizsgálatból. Másrészt a szöveg lexikailag eltérő szerkezeti részei és a szegmensek közötti véletlen megfelelés jelentősen torzíthatja az eredményt. E problémákat oldja meg a típus-token arány mozgó átlagának (*moving-average type-token ratio*) a kiszámítása.¹⁹ Ebben az esetben is ugyanolyan hosszú, azonban egymást átfedő szegmenseknek számolják ki a típus-token arányait, majd az így kapott értékeket átlagolják. Például 500 szavas szegmensek használata esetén kiszámolják a szöveg 1-től az 500. szóig, majd a 2-től az 501. szóig, majd a 3-tól az 502. szóig stb. terjedő szegmenseinek a típus-token arányait, és a kapott értékeket átlagolják. Mivel a szegmensek mindig csak egy szóval tolódnak el, a számításból nem marad ki a szöveg vége sem, illetve mivel az összes lehetséges szegmensen számításba veszi az eljárás, a szöveg szerkezeti egységei és a szegmensek esetleges egybeesésének a torzító hatása is kiküszöbölhető.

¹⁷ A magyar és más erőteljesebben ragozó nyelvek esetében a típus-token arány mérése helyett jellemzően a lemma-token arány mérését alkalmazzák.

¹⁸ Wendell JOHNSON, „Studies in Language Behavior: I. A Program of Research”, *Psychological Monographs* 56, 2 sz. (1944): 1–15, <https://doi.org/10.1037/h0093508>; Mike SCOTT, *Oxford WordSmith Tools: Version 4.0*. (Oxford: Oxford University Press, 2007).

¹⁹ Michael A. COVINGTON és Joe D. McFALL, „Cutting the Gordian Knot: The Moving-Average Type-Token Ratio (MATTR)”, *Journal of Quantitative Linguistics* 17, 2. sz. (2010): 94–100, <https://doi.org/10.1080/09296171003643098>.

A módszer azonban nem csupán a szókincs változatosságának, hanem – ahogyan arra a szerzők is utalnak²⁰ – kis ablakmérettel szavak egymáshoz közeli ismétlődésének a mérésére is alkalmas. Emellett nem csupán szavak ismétlődésének a mérését teszi lehetővé, hanem n-gramoknak, azaz szósoroknak vagy egyéb szavakhoz rendelt címkéknek, például szófajszorozatoknak az ismétlődését is képes mérni. A mérés elvégzéséhez két változót kell előre definiálnunk: az n-gramok méretét és az ablak méretét. Másként megfogalmazva: azt kell megadnunk, hogy a vizsgált n-gramoknak mekkora legyen az elemszáma, illetve, hogy az egyes szegmenseknek, amelyben az egymástól különböző n-gramok száma meghatározásra kerül, mekkora legyen az elemszáma. A szintaktikai ismétlődés mérésére kialakított eljárást az (1)-ben szereplő szófaji címkék alapján mutatom be. Tételizzük fel, hogy az (1) egy rövid, 14 szóból álló vers szavainak a szófajait és központozásait (két vessző és a végén egy pont) tünteti fel.

(1) ige névelő melléknév főnév , ige névelő melléknév főnév , határozószó ige névelő főnév .

Amennyiben az ismétlődést 12 elemszámú ablakokban vizsgáljuk, akkor az első ablakba a (2), a második ablakba a (3), a harmadik ablakba a (4), a negyedik ablakba pedig az (5) szófaji címkék, illetve írásjelek kerülnek.

(2) ige névelő melléknév főnév , ige névelő melléknév főnév , határozószó ige

(3) névelő melléknév főnév , ige névelő melléknév főnév , határozószó ige névelő

(4) melléknév főnév , ige névelő melléknév főnév , határozószó ige névelő főnév

(5) főnév , ige névelő melléknév főnév , határozószó ige névelő főnév .

Ha 3-gramok, azaz három elemből álló n-gramok alapján vizsgáljuk az ismétlődést, akkor a (2) 12-elemű ablak elemeit a (6)-ban felsorolt 3-gramokra bonthatjuk.

²⁰ Uo., 97–98.

(6) ige névelő melléknév
 névelő melléknév főnév
 melléknév főnév ,
 főnév , ige
 , ige névelő
 ige névelő melléknév
 névelő melléknév főnév
 melléknév főnév ,
 főnév , határozószó
 , határozószó ige

A (2) ablakban tehát összesen tíz darab 3-gram szerepel, amelyekből az ige-névelő-melléknév, a névelő-melléknév-főnév és a melléknév-főnév-vessző 3-gramok kétszer szerepelnek, vagyis a különböző 3-gramok száma hét. Ha az ablakban szereplő különböző 3-gramok számát elosztjuk az ablakban szereplő összes 3-grammal, akkor a típus-token arány mintájára kapunk egy 0 és 1 közötti értéket, amely kifejezi, hogy az ablakban szereplő n-gramok mennyire változatosak, illetve mennyire ismétlődnek. Minél közelebb van ez a szám a 0-hoz, annál kisebb az n-gramok változatossága, azaz annál nagyobb az n-gramok ismétlődésének a mértéke az ablakon belül. A (2) esetében ez az érték $7/10$, azaz $0,70$. Ha ezt az értéket az összes, azaz a (3), a (4) és az (5) ablakra is kiszámoljuk, akkor a kapott értékek átlagolásával adhatjuk meg a szöveg ismétlődési pontszámát, amely $0,85$. Mivel a típus-token arány mérésének az általános használatával szemben ebben az esetben nem a változatosság nagyságának, hanem az ismétlődés nagyságának mérése a cél, érdemes a kapott átlagot 1-ből kivonni annak érdekében, hogy a nagyobb érték ne az ismétlődés kisebb, hanem nagyobb mértékére utaljon, azaz az érték és az ismétlődés nagysága között ne fordított viszony legyen. A példa esetében tehát a szintaktikai ismétlődés mértéke $1 - 0,85$, azaz $0,15$.

A kapott pontszámokat befolyásolja az eljárás során használt n-gramok mérete. Ha 2-gramok alapján vizsgáljuk az ismétlődést, akkor a (2) ablak szó-faji címkéi, illetve írásjelei esetében a (7)-ben szereplő 2-gramokból álló listát kapjuk.

(7) ige névelő
 névelő melléknév
 melléknév főnév
 főnév ,

, ige
 ige névelő
 névelő melléknév
 melléknév főnév
 főnév ,
 , határozószó
 határozószó ige

A (7)-ben 11 2-gram szerepel, amelyekből négy kétszer jelenik meg, a különböző n-gramok száma tehát ebben az esetben is hét. A 7/11 osztást elvégezve viszont már nem 0,70-et, hanem két tizedesjegyre kerekítve 0,64-et kapunk, az egész szövegre pedig két tizedesjegyre kerekítve 0,70-et. A szintaktikai ismétlődés mértékének a nagysága tehát ebben az esetben 1 - 0,70, azaz 0,30. Vagyis, ha 2-gramok alapján végezzük el a számítást, akkor erőteljesebb ismétlődésre utaló értéket kapunk, mint ha 3-gramokkal végeznénk el azt.

A kapott értékeket ugyancsak befolyásolja az ablak mérete. Például, ha 3-gramok alapján végezzük el a vizsgálatot nyolc elemszámú ablakokkal, akkor az egész szövegre kapott ismétlődési érték kerekítve 1 - 0,94, azaz 0,06 lenne. De ha például hat elemszámú ablakokkal vizsgálnánk 3-gramokat, akkor mindegyik ablakra 1 értéket kapnánk, és így az egész vers átlagos értéke is 1 lenne, az ismétlődés mértéke pedig 1 - 1, azaz 0. Vagyis a túl kicsi ablak miatt az ismétlődő 3-gramok egyszer se kerülnének egy ablakon belülre, és ezért nem lehetne szintaktikai ismétlődést kimutatni.

A szintaktikai ismétlődés mérésekor problémaként jelentkezhet az, hogy szófaj n-gramok ismétlődésének oka lehet konkrét szövegrészek, például szókapcsolatok vagy akár teljes sorok megismétlése a versben (például refrének esetében). Mivel ez a kutatás kifejezetten csak a szövegek absztraktabb, szintaktikai szintjén megvalósuló ismétlések vizsgálatára irányul, a szövegrészek megismétlésével előálló szintaktikai ismétlődést valamiképpen figyelmen kívül kell hagyni. Ennek érdekében azt az eljárást követtem, hogy a fent bemutatott módszerrel kiszámítottam lemma n-gramok alapján is az ismétlődési értéket, majd ezt kivontam a szófaj n-gramokra kapott ismétlődési értékből. A lemma a szavak szótári alakja. Például az *asztal*, *asztalon*, *asztalokkal* szóalakok lemmája az *asztal* szóalak. Ezzel a módszerrel mind a szó szerinti, mind a lemma szerinti, azaz a közel szó szerinti ismétlések egyik típusa révén előálló szintaktikai ismétlődést figyelmen kívül lehetett hagyni a szintaktikai ismétlődés értékének megállapítása során. A módszer ugyanakkor azokat az ismét-

léseket nem hagyja figyelmen kívül, ahol az n -gramok szavai között legalább egy lemma szerinti eltérés van (például: *Kell a zsoltár*; *kell a fobász*; *Kell a bit*).²¹

3.2. Egymást követő sorok szintaktikai hasonlóságának mérése

A versekben megjelenő szintaktikai ismétlődés mértéke mellett a kutatás azt is vizsgálja, hogy a közvetlenül egymást követő rímelő sorok, illetve a közvetlenül egymást követő nem rímelő sorok között milyen mértékben valósul meg a szintaktikai ismétlődés jelensége. Magyar nyelvű históriás énekek esetében Seláf és szerzőtársai már tettek kísérletet a rímelő sorok grammatikai párhuzamosságának vizsgálatára az egymással rímelő sorokban szereplő utolsó négy szó grammatikai azonosságának automatizált meghatározásával.²² A szerzők beszámolója szerint azonban a morfológiai elemző a 16. századi szövegeken túl nagy hibaaránnal működött, ezért végül elvetették a módszer alkalmazását. Az általam követett módszer lényege, hogy a sorpárokat alkotó két sor n -gramjai közötti átfedési együtthatót (*overlap coefficient*) számítottam ki, majd az így kapott értékeket átlagoltam az egyes szerzők esetében. Az átfedési együtthatót két halmaz közötti hasonlóság számszerűsítésére használják. Értékét úgy kapjuk meg, hogy elosztjuk a két halmaz metszésének elemszámát a kisebb halmaz elemszámával. A kapott eredmény ebben az esetben is egy 0 és 1 közötti szám, ahol a 0 érték azt jelzi, hogy a két halmaznak nincs közös eleme, az 1 érték pedig arra utal, hogy a kisebb halmaz összes eleme megtalálható a nagyobb halmazban is. Például ha van két halmazunk, amelyből az egyik A, B, C és D elemeket tartalmazza, a másik pedig A, C, D, E és F elemeket, akkor a két halmaz metszése A, C és D elemeket tartalmazza. Az átfedési együttható értéke pedig $3/4$, azaz $0,75$.

Az itt bemutatandó vizsgálatban a két halmazt a két összehasonlítandó, egymást követő sor szófaj n -gramjai alkotják. Ugyanakkor mivel egy sorban ugyanaz a szófaj n -gram többször is előfordulhat, az átfedési együttható számításának azt a változatát használtam, amely multihalmazok hasonlóságát adja meg, azaz olyan halmazok hasonlóságát, amelyben ugyanaz az elem többször is megjelenhet. Ennek számítása során a két halmaz metszésének az elemszámát úgy kell megadni, hogy az elemeknek mindig abban a halmazban való előfordulási számát vesszük figyelembe, ahol azok kevesebbszer fordulnak elő. Például tételezzük fel, hogy szófaji címkék 3-gramjai alapján vizsgáljuk a (8)-ban és a (9)-ben szereplő sorok hasonlóságát multihalmazok átfedési együtthatója alapján.

²¹ Ady Endre: *A papok istene*.

²² SELÁF, „A belső formulák a 16. századi históriás énekekben”.

(8) melléknév főnév , melléknév főnév , melléknév főnév ,

(9) ige melléknév főnév , melléknév főnév ,

A (8)-ban és a (9)-ben szereplő szófaji címkék, illetve írásjelek a (10)-ben és a (11)-ben szereplő 3-gramokra oszthatók.

(10) melléknév főnév ,

főnév , melléknév

, melléknév főnév

melléknév főnév ,

főnév , melléknév

, melléknév főnév

melléknév főnév ,

(11) ige melléknév főnév

melléknév főnév ,

főnév , melléknév

, melléknév főnév

melléknév főnév ,

A (10) és (11)-ben szereplő 3-gramok közül a melléknév-főnév-vessző, a főnév-vessző-melléknév és a vessző-melléknév-főnév 3-gram szerepel mindkét sorban. A két multihalmaz metszésében lévő elemszám megadásához azt kell megnéznünk, hogy ez a három n-gram hányszor szerepel az egyes sorokban (multihalmazokban), és az alacsonyabb előfordulási számokat kell összeadnunk. A melléknév-főnév-vessző 3-gram háromszor szerepel az első sorban és kétszer a másodikban, így itt a 2 értéket tartjuk meg. A főnév-vessző-melléknév és a vessző-melléknév-főnév 3-gram egyaránt kétszer szerepel az első sorban és egyszer a második sorban, így itt mind a két n-gram esetében 1 értékkel számolunk. A két multihalmaz metszésében lévő 3-gramok száma tehát $2+1+1$, azaz 4. Az átfedési együttható kiszámításához ezt az összeget kell elosztanunk a kisebb multihalmaz elemszámával. A példa esetében a kisebb multihalmaz a második sor, amelyben a 3-gramok száma 5. A két sor átfedési együtthatója így $4/5$, azaz 0,8, amely a két sor 3-gramjainak kifejezetten nagymértékű egyezését jelzi.²³

²³ A leírt eljárásnak a képlete a következő (A és B a két multihalmaz): átfedési együttható:

$$= \frac{|A \cap B|}{\min(|A|, |B|)}$$

Itt is jelentkezik az a probléma, hogy ha a második sor szó szerint vagy közel szó szerint megismétli az első sort, akkor az a szófaj n-gramok esetében is magas átfedési együtthatót eredményez, miközben a vizsgálat a sorpárok-nak a szintaxis, azaz elsősorban a szófaj n-gramok szintjén megvalósuló hasonlóságát kívánja mérni. A problémát ebben az esetben is úgy oldottam meg, hogy a szófaj n-gramok alapján számolt átfedési együttható mellett kiszámítottam lemma n-gramok alapján is a sorpárok átfedési együtthatóját, majd ez utóbbi értéket kivontam a szófaj n-gramok alapján kapott értékből.

3.3. Kiterjesztett szófaji címkék használata

A vizsgálatokban a szófaj n-gramoknak egy olyan verzióját is használok, ahol a főnevek, a névmások és a melléknevek esetében nemcsak a szófaji címke, hanem az esetalak és az egyes szám harmadik személyű birtokos személyjel esetleges megléte is fel van tüntetve, az igék esetében pedig jelezve van, ha főnévi igenévről van szó.²⁴ Ezeket a továbbiakban kiterjesztett szófaji címkéknek nevezem. A (13)-ban szereplő 3-gram például egy nominatívusz esetalakú melléknévből, egy nominatívusz esetalakú főnévből és egy szublatívusz esetalakú, egyes szám harmadik személyű birtokos személyjellel ellátott főnévből áll (például *sötét gályánk szőnyegére*).²⁵ A (14) 3-gram pedig egy igét, egy akkuzatívusz esetalakú főnevet, valamint egy főnévi igenetet tartalmaz (például *segítettem koronát leütni*).²⁶

²⁴ Ez a megközelítés a mozaik-ngram módszerhez hasonlóan tehát abból indul ki, hogy a szintaktikai szerkezetek pontosabb megragadásához szükséges lehet az n-gramok egyes elemeinek eltérő specifikációs szinten történő definiálása. A mozaik n-gram módszerről lásd: Balázs INDIG, „Mosaic N-grams: Avoiding Combinatorial Explosion in Corpus Pattern Mining for Agglutinative Languages”, in *Human Language Technologies as a Challenge for Computer Science and Linguistics*, szerk. Zygmunt VETULANI, Patrick PAROUBEK és Marek KUBIS, 147–151 (Poznań: Adam Mickiewicz University, 2017); INDIG Balázs, LAKI László és PRÓSZÉKY Gábor, „Mozaik nyelvmodell az AnaGramma elemzőhöz”, in *XII. Magyar Számítógépes Nyelvészeti Konferencia*, szerk. TANÁCS Attila, VARGA Viktor és VINCZE Veronika, 260–270 (Szeged: Szegedi Tudományegyetem, TTIK, Informatikai Intézet, 2016); BAJZÁT Tímea és INDIG Balázs, „Személyjelölési mintázatok feltárása mozaik n-gram alapon három segédige + főnévi igenévi típus konstrukcióiban”, in *Személyjelölés és grammatika*, szerk. LACZKÓ Krisztina, TÁTRAI Szi-lárd és TOLCSVAI NAGY Gábor, 85–120 (Budapest: ELTE Eötvös Kiadó, 2024).

²⁵ Ady Endre, *Ima Baál istenhez*.

²⁶ Petőfi Sándor, *A Honvéd*.

(13) melléknév|nominatívusz főnév|nominatívusz főnév|szublatívusz|-
birtokviszony E/3

(14) ige főnév|akkuzatívusz igelinfinitívusz

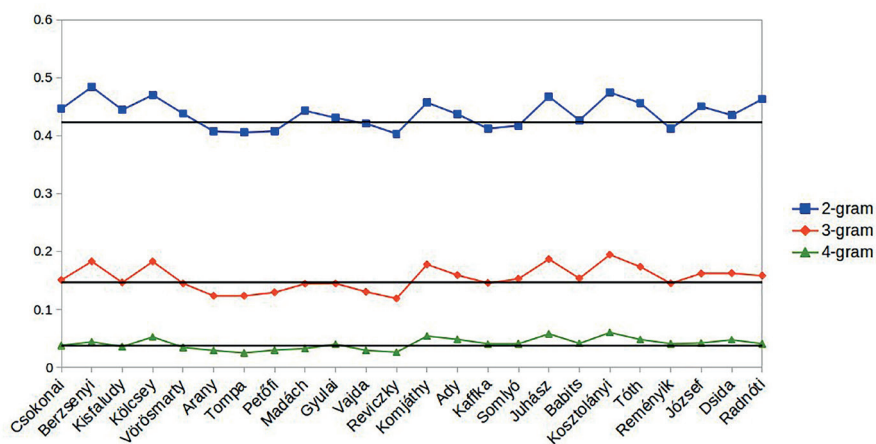
A főnevek, névmások és melléknevek esetalakjának a szerepeltetése lehetővé teszi, hogy a vizsgálat differenciáltabban vegye figyelembe a különböző szintaktikai szerkezeteket, hiszen az igei konstrukciók, illetve mondatkonstrukciók egyik lényegi tulajdonsága, hogy milyen esetalakú bővítményekkel jelennek meg. Ahogyan a (14) is példázza, a különböző számban és személyben megjelenő birtokos személyjelek közül csak az egyes szám harmadik személyű birtokos személyjelet vettem figyelembe. Ennek oka az, hogy a szintaktikai szempontból nagyobb relevanciával bíró birtokos jelzőszerkezeteknek ez a jelölője. A birtokszón első vagy második személyű birtokos személyjelek csak akkor fordulnak elő, ha a birtokos jelző első vagy második személyű személyes névmás (például az én cipőm, a te lakásod). Az ilyen névmási birtokos jelzők azonban jóval ritkábbak, mint a főnévi birtokos jelzők. Ugyanakkor versekben első és második személyű birtokos személyjelek számos esetben előfordulnak egyéb, a szintaktikai szerkezet szempontjából nem igazán releváns funkcióban, ami miatt az első és második személyű birtokos személyjelek figyelembevétele félrevezethetné a szintaktikai ismétlődésre fókuszáló vizsgálatokat. A főnévi igenév megkülönböztetése az igétől pedig azért indokolt, mert a főnévi igenév használatához több, kifejezetten gyakori szintaktikai konstrukció is kapcsolódik (például segédigés szerkezetek). Az esetalakhoz, birtokossághoz és főnévi igenévhez kapcsolódó morfoszintaktikai információk figyelembevétele tehát összességében érzékenyebbé teszi a kidolgozott n-gram alapú módszert a specifikusabb szintaktikai szerkezetekre.

4. SZINTAKTIKAI ISMÉTLŐDÉS A MAGYAR KANONIKUS KÖLTÉSZETBEN

Elsőként azt az általánosabb kérdést vizsgálom, hogy a magyar kanonikus költészetben nagyobb mértékben jelenik-e meg a szintaktikai ismétlődés jelensége, mint a magyar regényirodalomban. Amennyiben a kérdésre pozitív válasz adható, az arra utal, hogy a magyar kanonikus költészetben a fonológiai ismétlődés mellett a szintaktikai ismétlődésnek is kitüntetett szerepe van. Emellett azt is vizsgálom, hogy a magyar kanonikus költészetben milyen időbeli tendenciát mutat a szintaktikai ismétlődés mértékének az alakulása.

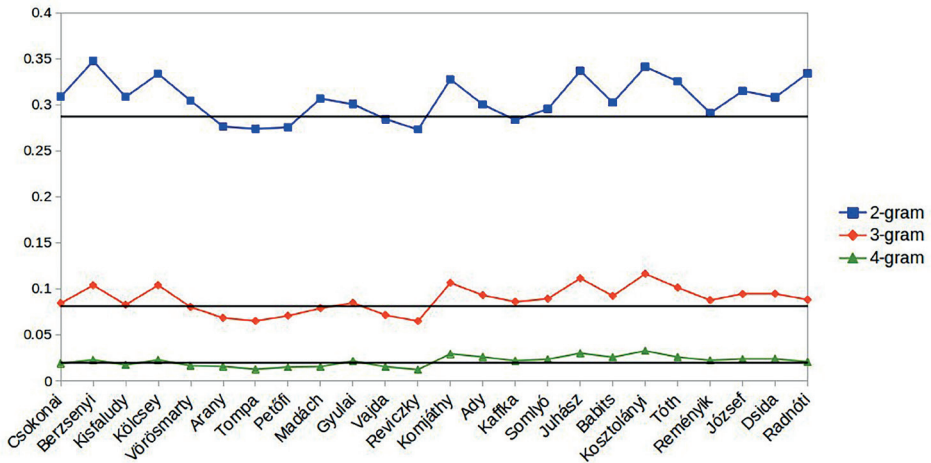
A két kérdés megválaszolásához első lépésben a kutatási korpusz 24 költőjének versei és a regénykorpusz esetében számítottam ki 100, 50 és 25 tokenes ablakbeállítással, kettő, három és négy elemszámú szófaj n-gramokra a szintaktikai ismétlődés mértékét. A tokenek száma az ablak elemszámára, azaz az ablakban található szavak és központozások együttes előfordulási számára utal. Az egyes költők esetében a számítást minden versre külön elvégeztem, majd a kapott értékeknek számítottam ki a mediánját.²⁷ A regénykorpusznál első lépésben a korpuszban szereplő 120 szerző esetében külön-külön kiszámítottam az egyes szerzők műveinél megállapított ismétlődési értékek mediánját, majd második lépésben az így kapott 120 mediánértéknek számítottam ki a mediánját. Erre az eljárásra azért volt szükség, hogy az egyes szerzők ugyanolyan súllyal vegyenek részt az egész regénykorpuszt reprezentáló mediánérték kiszámításában, függetlenül attól, hogy hány regénnyel szerepelnek a korpuszban. (Az egyes költőkre és a regénykorpuszra kapott mediánértékeket lásd a Függelék 7. táblázatában.)

A 100, 50 és 25 tokenes ablakméretekre kapott mediánértékeket az 1., 2. és 3. ábra mutatja be. Az egyes ablakméretekhez tartozó ábrák egyszerre jelenítik meg a 2-, 3- és 4-gramokra kapott értékeket a szerzők születési sorrendjében. A regényekre kapott mediánértékeket a vízszintes vonalak jelölik.

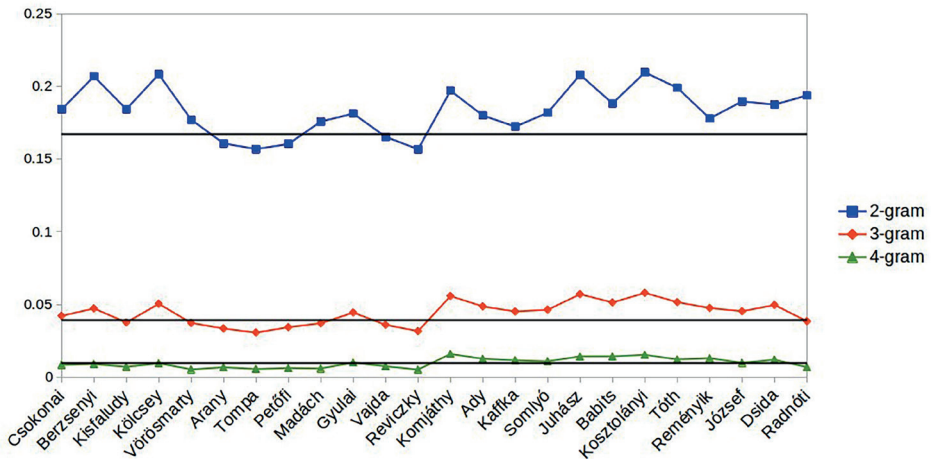


1. ábra. A szintaktikai ismétlődés mértékének mediánjai szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 100 tokenes ablakbeállítással

²⁷ A medián az átlaghoz hasonlóan egy számokból álló adatsor középértékét adja meg. Egy számokból álló adatsor mediánja a növekvő sorrendbe rendezett számokból a középső, páros számú adatsor esetén pedig a középső kettőnek az átlaga. A medián az átlagnál általában megbízhatóbb középérték, mivel nem érzékeny a kiugró adatokra.



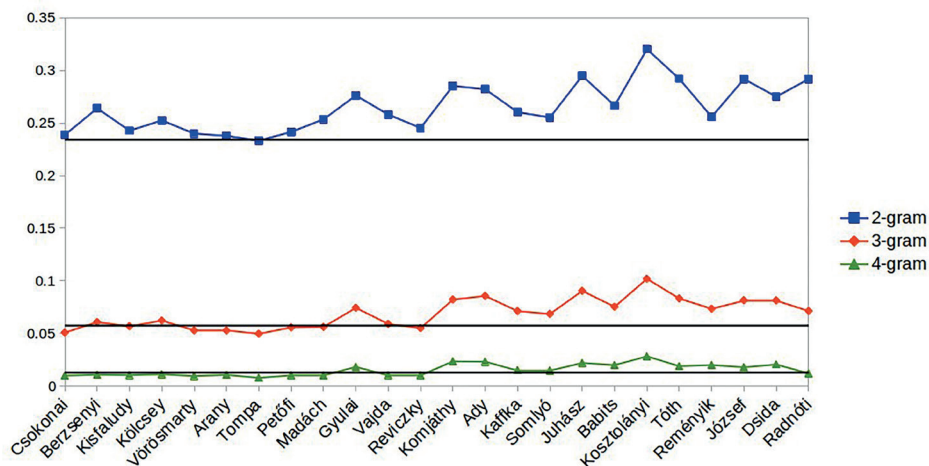
2. ábra. A szintaktikai ismétlődés mértékének mediánjai szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 50 tokenes ablakbeállítással



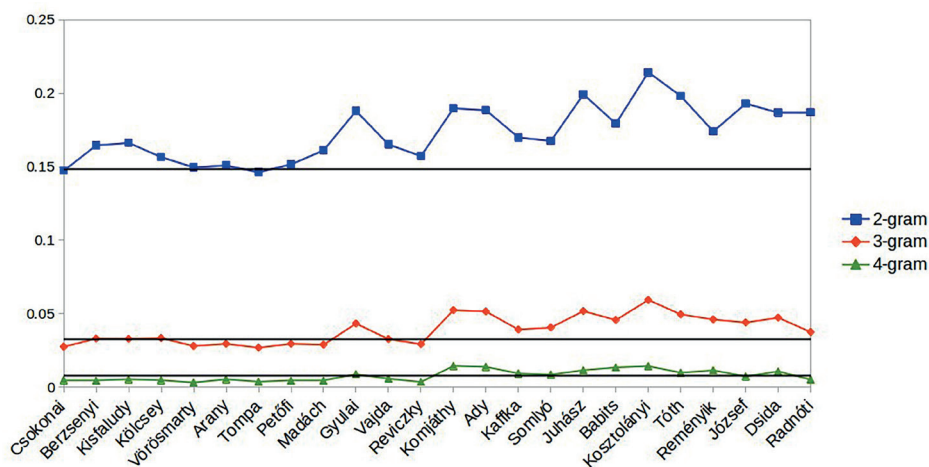
3. ábra. A szintaktikai ismétlődés mértékének mediánjai szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 25 tokenes ablakbeállítással

Ugyanezeket a méréseket elvégeztem kiterjesztett szófaji címkék alapján is. (Az egyes költőkre és a regénykorpuszra kapott mediánértékeket lásd a *Függelék* 8. táblázatában.) Ezek a mérések tehát a szófaji címke mellett az esetalakot, az egyes szám harmadik személyű birtokos személyjelet, illetve igéknél

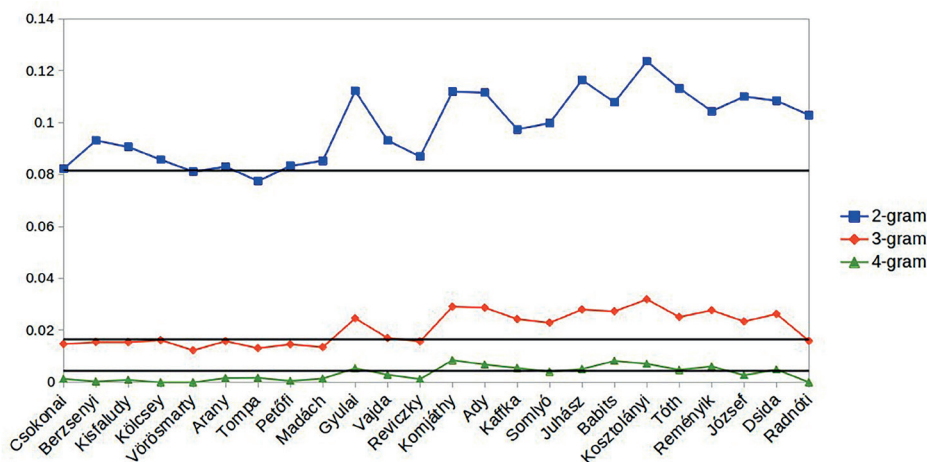
a főnévi igenévi jelet is figyelembe veszik, ami a szintaktikai viszonyok differenciáltabb megragadását teszi lehetővé. A mérések eredményeit a 4., 5. és 6. ábra mutatja be.



4. ábra. A szintaktikai ismétlődés mértékének mediánjai kiterjesztett szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 100 tokenes ablakbeállítással



5. ábra. A szintaktikai ismétlődés mértékének mediánjai kiterjesztett szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 50 tokenes ablakbeállítással



6. ábra. A szintaktikai ismétlődés mértékének mediánjai kiterjesztett szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 25 tokenes ablakbeállítással

A hat ábrából látható, hogy a legtöbb mérési beállítás esetén a költők többsége a regényekre kapott ismétlődési értéket meghaladja. A 2. táblázat mutatja be, hogy az egyes mérési beállításoknál a 24-ből hány szerző haladja meg a regényekre kapott értéket. A regényekre kapott értéket meghaladó legtöbb költőt kiterjesztett szófaji címkék 2-gramjainál találjuk. 50 és 25 tokenes ablakméretnél a 24-ből 22 szerző, 100 tokenes ablakméretnél pedig 23 szerző haladja meg a regényeknél mért értéket. Szürke cellaszínezéssel emeltem ki azokat a mérési beállításokat, ahol a regényekre kapott ismétlődési értéket csupán a vizsgált költők fele vagy kevesebb mint a fele haladja meg. Látható, hogy mind a négy ilyen eset a 4-gramos beállításhoz kapcsolódik. Ezekből három a kiterjesztett szófaji címkék használatánál jön elő.

2. táblázat. A regényekre kapott szintaktikai ismétlődési értékek mediánját meghaladó költők száma az egyes mérési beállításoknál (összes költő = 24)

	Szófaji címkék			Kiterjesztett szófaji címkék		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
100 token	16	13	15	23	16	12
50 token	18	17	15	22	17	11
25 token	19	15	12	22	13	10

A 3. táblázat a költőkre kapott mediánértékek regények mediánjától való eltéréseinek (pozitív vagy negatív szám) összegét mutatja be. Egy darab, szűr-kével jelölt mérési beállítást leszámítva mindenhol pozitív érték szerepel, ami úgyszintén azt jelzi, hogy a vizsgált költők verseiben a szintaktikai ismétlődés mértéke jellemzően nagyobb, mint a magyar regényirodalomban.

3. táblázat. A költőkre kapott mediánok regények mediánjától való eltéréseinek összege az egyes mérési beállításoknál

	Szófaji címkék			Kiterjesztett szófaji címkék		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
100 token	0.3461	0.1681	0.0691	0.7393	0.2670	0.0557
50 token	0.4632	0.1727	0.0296	0.5967	0.1647	0.0047
25 token	0.3927	0.1175	0.0065	0.4053	0.1006	-0.0246

A szófaji, illetve kiterjesztett szófaji címkék alapján végzett mérésekből az derül ki, hogy a magyar kanonikus költészetben általánosságban nagyobb mértékben jelenik meg a szintaktikai ismétlődés jelensége, mint a magyar regényirodalomban. Úgy tűnik tehát, hogy a fonológiai ismétlődés mellett a verseket valóban jellemzi egy másik típusú, a szintaktikai szerkezetek használatában testet öltő ismétlődés is. Ahogyan arra korábban utaltam, az ismétlődés, beleértve a szintaktikai ismétlődést is, a nyelv alapjellemezője. Az ismétlődés nyelvbeli fontosságát jól mutatja, hogy egymással szemben álló nyelvészeti irányzatokban is elsődleges, a nyelv működésében nélkülözhetetlen mechanizmusként tűnik fel. A konstrukciós nyelvtani megközelítések például a grammatikát olyan konstrukciók összességeként ragadják meg, amelyek az ismétlődő, egymáshoz fonológiaiilag és szemantikailag hasonló nyelvi megnyilatkozásokból absztrahálódnak, illetve amelyek egyúttal ezeknek az egymáshoz hasonló megnyilatkozásoknak az ismételt létrehozását motiválják.²⁸ De a konstrukciós megközelítésekkel szemben álló generatív nyelvészet is a nyelvi képesség alapjának tekinti az ismétlés egy típusát, a rekurziót, amely lehetővé teszi, hogy ugyanazt a műveletet újra és újra elvégezve egy nyelvi szerkezetbe ismételten beágyazzuk ugyanazt

²⁸ Lásd például Adele E. GOLDBERG, *Constructions: A Construction Grammar Approach to Argument Structure* (Chicago, London: The University of Chicago Press, 1995); Adele E. GOLDBERG, *Constructions at Work: The Nature of Generalization in Language* (Oxford: Oxford University Press, 2005), <https://doi.org/10.1093/acprof:oso/9780199268511.001.0001>; William CROFT és Alan D. CRUSE, *Cognitive Linguistics* (Cambridge: Cambridge University Press, 2004), 257–290, <https://doi.org/10.1017/CBO9780511803864>.

a szerkezetet.²⁹ A versekben megjelenő fonológiai és szintaktikai ismétlődés tehát nem egy új, a hétköznapi beszédben nem létező minőséget ad hozzá a nyelvhez, hanem a nyelv működésének, lényegének egyik alapjellemezőjét teszi láthatóbbá. A szintaktikai ismétlődés erőteljesebb megjelenését így a fonológiai ismétlődés erőteljesebb megjelenéséhez hasonlóan magyarázhatjuk a Jakobson által leírt poétikai funkcióval.³⁰ Mindkettő a versszöveg nyelviségére, a szöveg referenciája helyett magára a nyelvi mechanizmusra hívja fel a figyelmet azáltal, hogy a nyelvet lehetővé tevő, ugyanakkor a hétköznapi megnyilatkozásainkban kevésbé szembeűnő eljárást, az ismétlést állítja a figyelem középpontjába. A poétikai funkció szempontjából a felfokozott szintaktikai ismétlődés csupán annyiban különbözik a rímtől, a ritmustól vagy éppen az alliterációtól, hogy nem a hangzás, hanem a szerkezet szintjén mutatja fel az ismétlődést, mint a nyelv egyik jellegadó tulajdonságát.

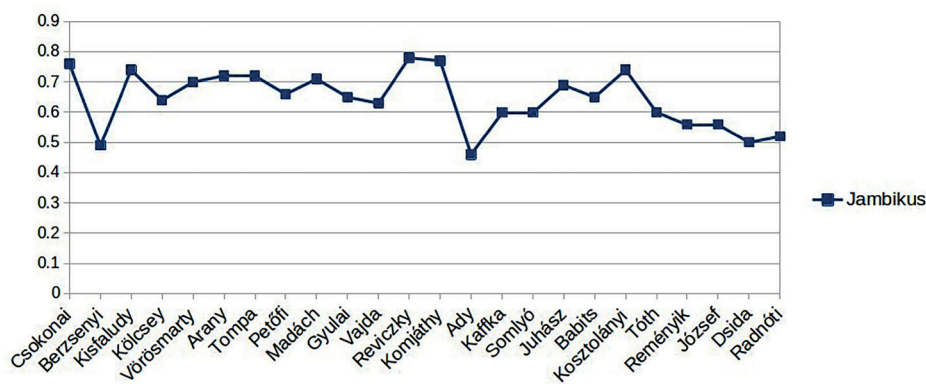
Az ábrákból az is jól látszik, hogy a szintaktikai ismétlődés mértéke időben változik. A 19. század utolsó harmadában alkotó Komjáthy-nál, illetve a 20. században alkotó szerzők esetében általánosan megnő a szintaktikai ismétlődés mértéke a korábbi időszakhoz képest. Egy kevésbé erőteljes mintázat is megfigyelhető: a 19. század elején, Berzsenyi, Kisfaludy, Kölcsey, illetve a sima szófaji címkéken alapuló beállításoknál Csokonai esetében a szintaktikai ismétlődés erőteljesebben jelenik meg, mint az őket követő, a 19. század közepén alkotó költőknél. Sima szófaji címkék esetén a 19. század második felében alkotó költők nagy része is alatta marad az említett 19. század eleji költőknek. A kiterjesztett szófaji címkék használatánál érdekes módon Gyulai esetében a 19. századi értékekhez képest kifejezetten magas az ismétlődés mértéke. Ennek pontos okait egy a szövegek szoros olvasását elvégző kvalitatív vizsgálat tárhatná fel.

Felvethető a kérdés, hogy a 19. század végén, illetve a 20. század elején a szintaktikai ismétlődés mértékének a megnövekedése a magyar költészet egyéb poétikai dimenzióitól függetlenül megy-e végbe, vagy pedig összefüggésbe hozható valamilyen más poétikai változással. Máshogyan megfogalmazva a kérdés az, hogy a szintaktikai ismétlődés erőteljesebbé válása pusztán többletminőségként jelenik meg a magyar költészetben, ami érintetlenül hagyva a már meglévő poétikai eszköztárat mintegy hozzáadódik ahhoz, vagy pedig a vers egyéb poétikai dimenzióira is kiható rendszerváltozás eredményeként áll

²⁹ Lásd például Marc D. HAUSER, Noam CHOMSKY és W. Tecumseh FITCH, „The Faculty of Language: What is It, Who Has It, and How Did It Evolve?”, *Science* 298, 5598 sz. (2002): 1569–1579, <https://doi.org/10.1126/science.298.5598.1569>.

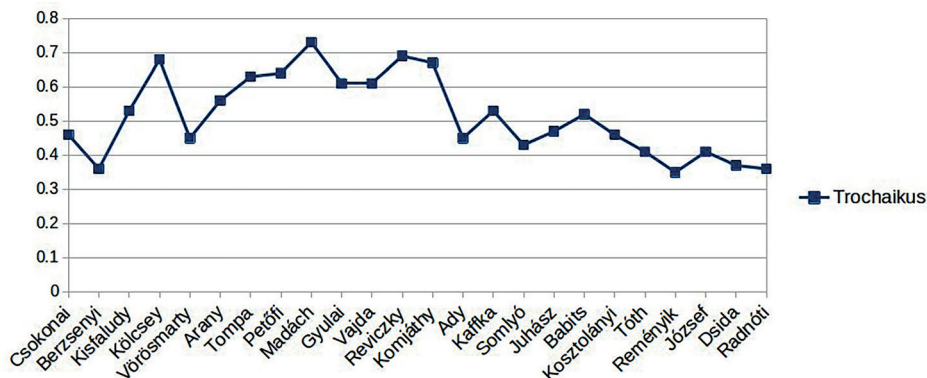
³⁰ Roman JAKOBSON, „Nyelvészet és poétika”, in JAKOBSON, *Hang – Jel – Vers*, 211–257.

elő. Amennyiben Lowth-nak, Földinek és Jakobsonnak a tanulmány elején hivatkozott elgondolásai alapján abból indulunk ki, hogy a versszerűség érzetét nemcsak a fonológiai ismétlődés, hanem egyéb típusú ismétlődések is megerősíthetik, akkor megfogalmazható az a hipotézis, hogy a szintaktikai ismétlődés erőteljesebbé válása a fonológiai ismétlődés visszaesését kompenzálja.³¹ Azaz a metrumnak a 20. századra bekövetkező fellazulása eredményezte egy másik típusú ismétlődés, a szintaktikai ismétlődés erőteljesebbé válását. A poétikai rendszerben tehát az egyik típusú ismétlődés gyengülését egy másik típusú ismétlődés megerősödése egyenlíti ki. A metrumnak a 20. század elejére bekövetkező fellazulását mind a jambikus, mind a trochaikus metrumot megvalósító versek esetében megfigyelhetjük. A 7. ábra a jambikus, a 8. ábra a trochaikus metrumú versek szabályossági pontszámainak a mediánját mutatja be az egyes szerzők esetében (a pontos mediánértékeket lásd a *Függelék* 9. táblázatában). A következő részben használt 0,5 helyett ezúttal 0,3 szabályossági pontszám felett vettem bele az egyes verseket a vizsgálatba annak érdekében, hogy a mediánértékek ne csak a jambikus és trochaikus metrumot szigorúbban megvalósító verseken alapuljanak. A jambikus és trochaikus versekbe nem vettem bele a szimultán metrumú verseket, azaz azokat, amelyekben ütemhangsúlyos metrum is megjelenik az időmérték mellett.



7. ábra. A jambikus metrumú versek szabályossági pontszámainak mediánjai szerzőnként (szabályossági küszöbérték: 0,3)

³¹ Köszönettel tartozom Szemes Botondnak, aki a tanulmány első változatát olvasva felvetette az ötletet, hogy a szintaktikai ismétlődés növekedése összefügghet a fonológiai ismétlődés fellazulásával.



8. ábra. A trochaikus metrumú versek szabályossági pontszámainak mediánjai szerzőnként (szabályossági küszöbérték: 0,3)

A 7. és 8. ábrából láthatjuk, hogy míg a 19. század költőinél többnyire szabályosabban valósul meg a jambikus és a trochaikus metrum, addig a 20. század első felében alkotó szerzőknél a szabályosság mértéke visszaesik. Ez pont fordítottja a szintaktikai ismétlődés esetében megfigyelt trendnek, ami megerősíti a hipotézist, hogy olyan rendszerszintű változásról van szó, ahol a metrumhoz kapcsolódó fonológiai ismétlődés mértékének a visszaesését kompenzálja a szintaktikai ismétlődés mértékének a növekedése. Megjegyzendő persze, hogy nem feltétlenül érdemes a fonológiai ismétlődés mértékét független változónak, a szintaktikai ismétlődést pedig ettől függő változónak tekinteni, azaz okra és okozatra felosztani a két jelenséget. Az ismétlődés e két típusának a változásai minden bizonnyal kölcsönösen hatottak egymásra. Természetesen egyéni kivételek vannak. Például Komjáthy és Kosztolányi esetében a metrum megvalósításának a szabályossága és a szintaktikai ismétlődés mértéke is nagy fokú.

Érdemes kiemelni, hogy a szófaj n-gramokkal megragadott szintaktikai jellemzők valójában szemantikai jellemzőknek az absztrakt kifejeződései. A szófajok megfeleltethetők sematikus szemantikai kategóriáknak,³² a szófajok közötti viszonyok pedig sok esetben sematikus szemantikai viszonyoknak.

³² Lásd: Ronald W. LANGACKER, *Cognitive Grammar: A Basic Introduction* (Oxford: Oxford University Press, 2008), 93–160, <https://doi.org/10.1093/acprof:oso/9780195331967.001.0001>; TOLCSVAI Nagy Gábor, „Jelentés”, in TOLCSVAI Nagy Gábor, szerk., *Nyelvtan*, 310–390 (Budapest: Osiris Kiadó, 2017).

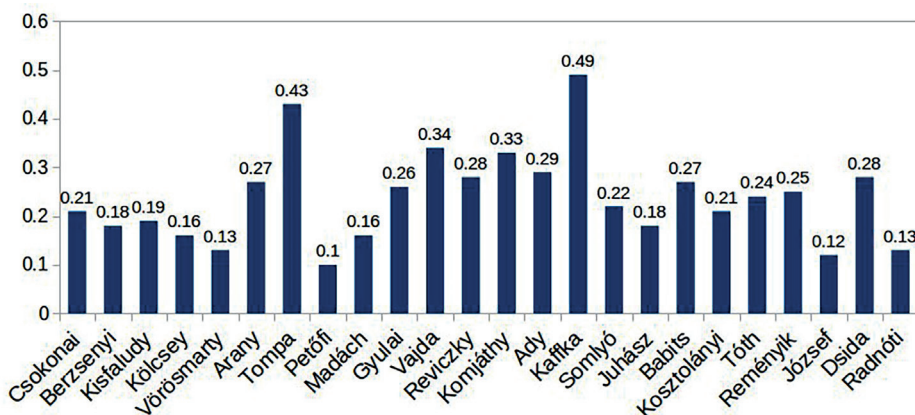
Például a melléknév-főnév kapcsolat általában a következő sematikus jelentéssel írható le: entitások egy adott tulajdonsággal rendelkező típusa. A szintaxis szintjén megragadott ismétlődés jelensége tehát nem más, mint absztraktabb szinten megvalósuló szemantikai ismétlődés. Amennyiben a szintaktikai ismétlődést absztrakt szinten érvényesülő tartalmi ismétlődéssel azonosítjuk, akkor levonhatjuk a következtetést, hogy a 20. század elejének a kanonikus szerzői nagyobb mértékben hoztak létre elvont, azaz a szintaxis szintjén tetten érhető tartalmi párhuzamokat, mint a 19. század szerzőinek a nagy része.

4.1. Szintaktikai ismétlődés és halmozás

A vizsgálatokhoz használt, átfedő n-gramokon alapuló módszer akkor is szintaktikai ismétlődést állapít meg, hogyha legalább háromszor szerepel egymás után ugyanaz a szófaji címke, központosítással elválasztva vagy központosítások nélkül. Ezek az esetek általában halmozásnál jelennek meg, vagyis amikor ugyanolyan szintaktikai funkcióban szereplő szavak fordulnak elő egymást követően. Ennek legtipikusabb esete a jelzőhalmozás (például *A ködből hulltak a szívek, / Csúnya, kicsiny, kemény szívek*).³³ A melléknév-vessző-melléknév-vessző-melléknév címkesor 2-gramjai a következők: melléknév-vessző, vessző-melléknév, melléknév-vessző, vessző-melléknév. Látható tehát, hogy a melléknév-vessző és a vessző-melléknév esetén is ismétlődő n-gramokat kapunk. Ha 3-gramokra bontjuk a címkesort, akkor csak a melléknév-vessző-melléknév n-gram fordul elő kétszer. Természetesen az azonos szintaktikai szerepű szavak ismétlése úgyszintén tekinthető a szintaktikai ismétlődés egy fajtájának. Ugyanakkor felmerülhet a kérdés, hogy a verseknek a regényekhez képest mért nagyobb mértékű szintaktikai ismétlődését, illetve a szintaktikai ismétlődés mértékének a növekedését a 20. század elején nem pusztán csak a halmozás gyakoribb használata okozza. A 9. ábra a vizsgált költők esetében mutatja be azoknak a verseknek az arányát, amelyekben legalább egy darab, minimum három elemből álló melléknév-, főnév- vagy ige-halmozás szerepel egy verssoron belül. Főneveknél és melléknéveknél azokat az eseteket tekintetem halmozásnak, ahol egy verssoron belül egymást követően vesszővel elválasztva vagy vessző nélkül minimum három főnév vagy melléknév szerepel ugyanolyan esetben úgy, hogy vagy egyik szónál sincs birtokos jelölve, vagy pedig mindegyik szó esetében ugyanolyan számú és személyű birtokos van jelölve. Igéknél csak azokat az

³³ Ady Endre, *A Léda szíve*.

eseteket tekintettem halmozásnak, ahol egy verssoron belül legalább három ige követi egymást teljesen megegyező morfoszintaktikai jellemzőkkel.

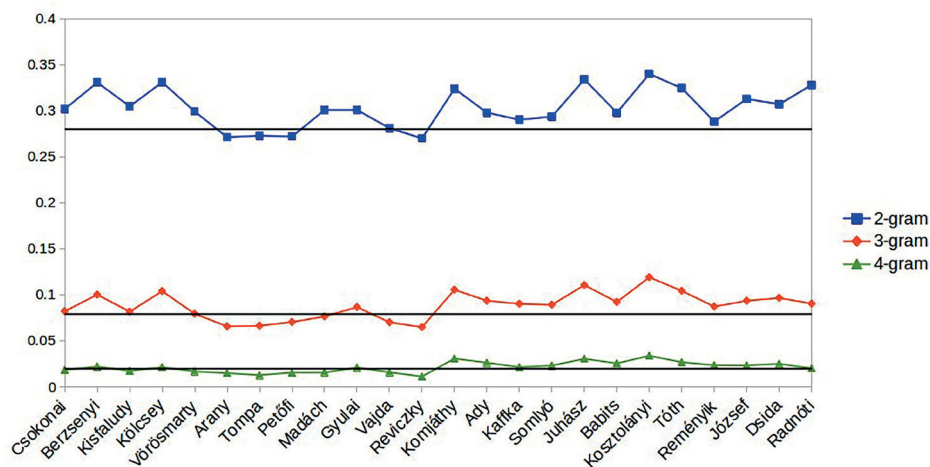


9. ábra. Minimum három elemből álló főnév-, melléknév- vagy igehalmozást tartalmazó versek aránya

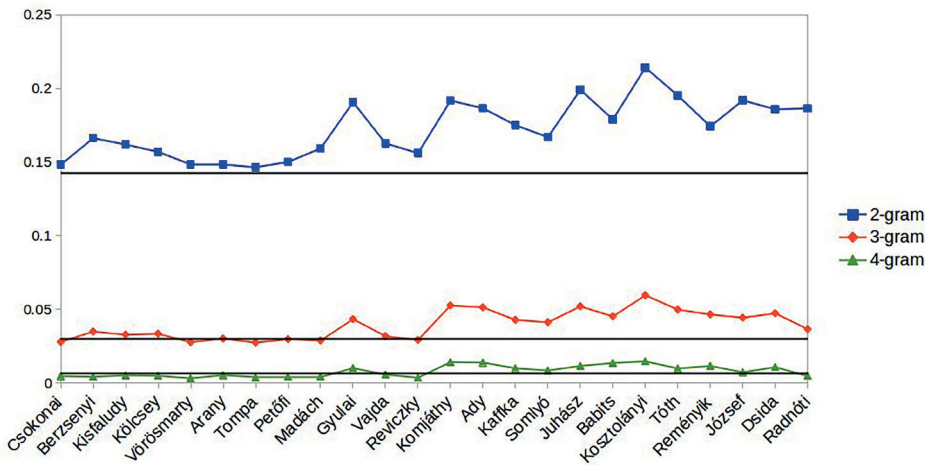
A 9. ábra oszlopdiagramjából láthatjuk, hogy viszonylag nagy arányban fordulnak elő legalább három elemből álló főnév-, melléknév- vagy igehalmozást tartalmazó versek az egyes életművekben. Tompa és Kaffka esetében ezeknek a verseknek az aránya megközelíti a verseik felét, de Arany, Gyulai, Vajda, Reviczky, Komjáthy, Ady, Babits és Dsida esetében is 25% felett van ezeknek a verseknek az előfordulása. Ebben a tanulmányban nincs mód a jelenség részletesebb vizsgálatára, a jövőben azonban mindenképpen érdemes lenne a halmozás különböző típusaival és ezek poétikai funkciójával behatóbban is foglalkozni.

Annak érdekében, hogy megbizonyosodhassunk arról, hogy a szintaktikai ismétlődésnek az előző részben megállapított, regényeknél magasabb értékeit, illetve 20. századi megemelkedését nem csupán a halmozás jelensége okozza, a méréseket elvégeztem úgy is, hogy a szintaktikai ismétlődés számításakor csak azokat az ablakokat vettem figyelembe, amelyekben nem szerepel egymást követően központozással elválasztva vagy központozás nélkül három vagy több ugyanolyan szófaji címke. Ez tehát azt jelenti, hogy szemben a 9. ábra előállításához használt módszerrel, csak a szófaji címkéket vettem figyelembe, illetve nemcsak a melléknév-, főnév- és igehalmozásokat tartalmazó

ablakokat zártam ki a vizsgálatból, hanem azokat az ablakokat is, amelyekben bármilyen más szófajú szóból szerepel egymást követően legalább három. Emellett az azonos szófajú szavakat elválasztó írásjelet nem korlátoztam a vesszőre, illetve nem vettem figyelembe a sorhatárokat sem (a regények esetében erre lehetőség sem lenne). Az így kizárt ablakokban szereplő halmozások jelentős része természetesen grammatikai értelemben nem tekinthető halmozásnak, a vizsgálatok eredményeit ugyanakkor a tényleges halmozásokhoz hasonlóan torzíthatják ezek az esetek is. A 10. ábra az 50 tokenes ablakmérettel futtatott mérések eredményeit mutatja be szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján. A 11. ábra ugyanezt mutatja be kiterjesztett szófaji címkék alapján. Ebben az esetben a halmozást tartalmazó ablakok kizárását is a kiterjesztett szófaji címkék alapján végeztem el. A 4. táblázat azt mutatja be, hogy az egyes mérési beállításoknál a 24-ből hány szerző haladja meg a regényekre kapott értéket. Az 5. táblázat pedig a költőkre kapott mediánértékek regények mediánjától való eltéréseinek összegét mutatja be a különböző mérési beállításoknál. Az egyes mérési beállításoknál nem kerültek bele a vizsgálatba azok a versek, amelyek esetében egy olyan ablak sem volt, amely nem tartalmazott halmozást.



10. ábra. A szintaktikai ismétlődés mértékének mediánjai szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 50 tokenes ablakbeállítással, a halmozást tartalmazó ablakok kizárásával



11. ábra. A szintaktikai ismétlődés mértékének mediánjai kiterjesztett szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 50 tokenes ablakbeállítással, a halmozást tartalmazó ablakok kizárásával

4. táblázat. A regényekre kapott szintaktikai ismétlődési értékek mediánját meghaladó költők száma az egyes mérési beállításoknál, a halmozást tartalmazó ablakok kizárásával (összes költő = 24)

	Szófaji címkék			Kiterjesztett szófaji címkék		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
100 token	17	16	15	24	20	14
50 token	20	18	15	24	18	12
25 token	20	16	13	24	19	11

5. táblázat. A költőkre kapott mediánok regények mediánjától való eltéréseinek összege az egyes mérési beállításoknál, a halmozást tartalmazó ablakok kizárásával

	Szófaji címkék			Kiterjesztett szófaji címkék		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
100 token	0.4595	0.2165	0.0560	0.9003	0.3397	0.0815
50 token	0.5607	0.2287	0.0441	0.7133	0.2284	0.0270
25 token	0.4611	0.1570	0.0125	0.5103	0.1425	-0.0109

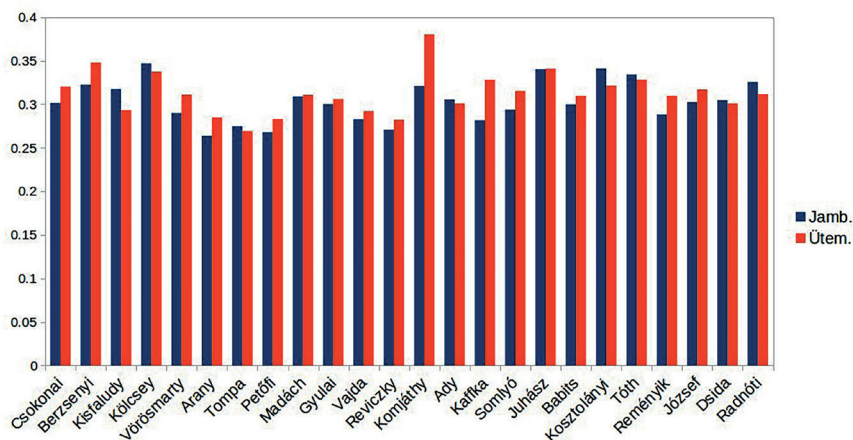
A kapott eredmények érdekes módon azt mutatják, hogy a halmazásokat tartalmazó ablakok kizárásával egyáltalán nem csökkent a regények és a versek között talált különbség a szintaktikai ismétlődés mértékében, sőt az még markánsabbá vált. Kiterjesztett szófaji címkék 2-gramjai esetén például az összes költő meghaladja a regényeknél mért értéket. A 10. és a 11. ábra alapján láthatjuk azt is, hogy a halmazást tartalmazó ablakok kizárásával kapott értékek alapján is megállapítható, hogy a szintaktikai ismétlődés mértéke megemelkedik a 20. század eleji költészetben a korábbi időszakhoz képest.

5. A METRUM TÍPUSA ÉS A SZINTAKTIKAI ISMÉTLŐDÉS ÖSSZEFÜGGÉSEI A MAGYAR KANONIKUS KÖLTÉSZETBEN

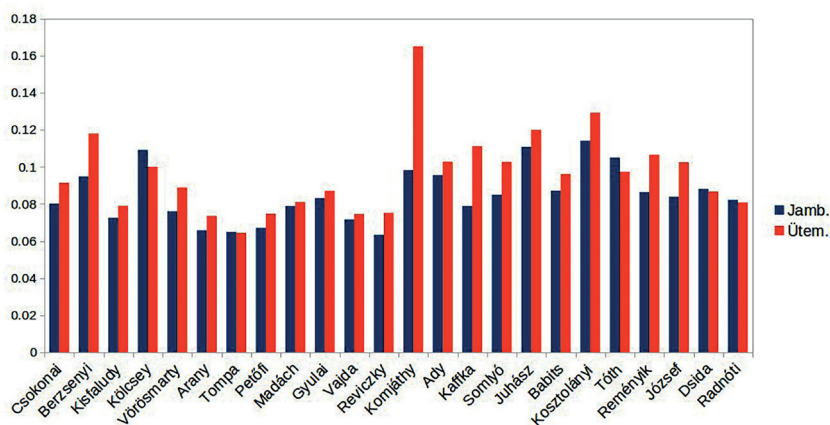
Az előző részben láthattuk, hogy a szintaktikai ismétlődés és a fonológiai ismétlődés időbeli alakulása feltehetően nem független egymástól. Ugyanakkor felvethető a kérdés, hogy a fonológiai ismétlődés különböző típusai, például a versben realizálódó metrumtípusok nem befolyásolják-e eltérő módon a szintaktikai ismétlődés megvalósulását. Az alábbiakban azt vizsgálom, hogy a metrum típusa és a szintaktikai ismétlődés mértéke között felfedezhető-e valamilyen összefüggés. A vizsgálat során a magyar költészet két leggyakoribb metrumtípusába, a jambikus és az ütemhangsúlyos metrumtípusba sorolható versek szintaktikai ismétlődését vetettem össze. A vizsgálatban azokat a verseket tekintettem jambikusnak, amelyeknek a szabályossági pontszáma meghaladja a 0,5 értéket. Ezekben a versekben a vers ritmusa jól érzékelhetően valósítja meg a jambikus metrumot, ami nagy valószínűséggel a szerző tudatos törekvésének az eredménye. A *Függelék* 10. táblázata mutatja be szerzőnként, hogy a metrumelemző program hány verset sorolt 0,5 küszöbértékkel a jambikus és a trochaikus, valamint az ütemhangsúlyos, a szimultán jambikus és a szimultán trochaikus versek közé. Ebben a vizsgálatban az ütemhangsúlyos kategória részének tekintettem a szimultán trochaikus verseket is, azaz azokat a verseket, amelyek egyszerre trochaikusak és ütemhangsúlyosak. Ugyanakkor a szimultán jambikus, azaz egyszerre jambikus és ütemhangsúlyos verseket kiszűrtem a vizsgálatból, vagyis ezeket se a jambikus, se az ütemhangsúlyos csoport részeként nem vettem számításba.³⁴ A *Függelék* 11. táblázata mutatja be a vizsgált szerzők esetében a jambikus és az ütemhangsúlyos metrumú versekre kapott szintaktikai ismétlődés értékei-

³⁴ A metrumelemző program működését részletesen bemutatja HORVÁTH Péter, „Két eljárás magyar nyelvű versek metrumának gépi felismertetéséhez”. A szimultán metrum Szepes és Szerdahelyi értelmezését követem, amely szerint azokban az esetekben beszélhetünk

nek a mediánját szófaji címkék 2-gramjai, 3-gramjai és 4-gramjai alapján, 50 tokenes ablakmérettel. A 12., 13. és 14. ábra oszlopdiagramjai a három n-gram-méretre kapott mediánértékeket tüntetik fel.

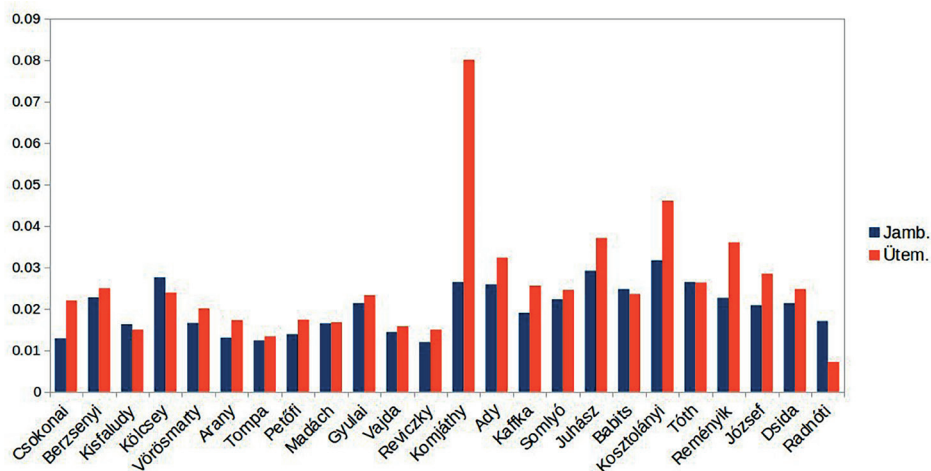


12. ábra. A szintaktikai ismétlődés mértékének mediánjai jambikus és ütemhangsúlyos verseknél, szófaji címkék 2-gramjai alapján, 50 tokenes ablakbeállítás



13. ábra. A szintaktikai ismétlődés mértékének mediánjai jambikus és ütemhangsúlyos verseknél, szófaji címkék 3-gramjai alapján, 50 tokenes ablakbeállítás

szimultaneitásról, „amikor ugyanazok a költemények egyszerre felelnek meg valamilyen idő-mértékes és ütemhangsúlyos képletnek, tehát kétféle ritmusúak”. SZEPES Erika és SZERDAHELYI István, *Verstan* (Budapest: Gondolat Kiadó, 1981), 510.



14. ábra. A szintaktikai ismétlődés mértékének mediánjai jambikus és ütemhangsúlyos verseknél, szófaji címkék 4-gramjai alapján, 50 tokenes ablakbeállítással

Ugyanezeket a méréseket lefuttattam kiterjesztett szófaji címkék alapján is. A kapott értékek úgyszintén a *Függelék* 11. táblázatában szerepelnek. A 6. táblázat mutatja be, hogy az egyes mérési beállításoknál a 24 szerzőből hány esetben nagyobb az ütemhangsúlyos versek szintaktikai ismétlődésének a mediánja a jambikus versek szintaktikai ismétlődésének a mediánjánál. A táblázatból láthatjuk, hogy minden mérési beállítás esetén a költők többségénél ez a mintázat jelenik meg.

6. táblázat. Az egyes mérési beállításoknál azon költők száma, akiknél az ütemhangsúlyos versekre kapott szintaktikai ismétlődés mediánja meghaladja a jambikus versekre kapott szintaktikai ismétlődés mediánját (összes költő = 24)

	Szófaji címkék			Kiterjesztett szófaji címkék		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
50 token	16	19	19	13	15	18

Az ütemhangsúlyos versek nagyobb mértékű szintaktikai ismétlődéséből arra következtethetünk, hogy az ütemhangsúlyos verselés nagyobb mértékben motiválja a szintaktikailag megegyező, illetve hasonló szerkezetek alkalmazását, mint a jambikus verselés. Ennek az lehet az oka, hogy míg a jambi-

kus verselés verslábai általában két szótagosak és nem feltétlenül szó elején kezdődnek, addig az ütemhangsúlyos versek ütemei jellemzően hosszabbak (például felező nyolcas esetében négy szótagosak, felező tizenkettes esetében pedig hat szótagosak), illetve az ütemek első szótagja általában szóeleji szótagra esik. Az ütemek nagyobb méretüknél fogva, illetve szavakat nem kérésztbe metsző jellegüknél fogva egy teljes szintaktikai frázist is magukban foglalhatnak. A sok esetben szintaktikai frázisoknak megfeleltethető ütemeknek egy meghatározott mintázat szerint való ismétlődése így motiválhatja a szintaktikai frázisok ismétlődését. Mindez a jambikus metrumról sokkal kevésbé mondható el, hiszen a jellemzően két szótagos és sok esetben nem szó eleji szótagnál kezdődő verslábak általában nem feleltethetők meg szintaktikai frázisoknak. Vagyis a verslábak valamilyen rendszer szerint megvalósuló ismétlődéséből kevésbé fog következni a szintaktikai frázisok valamilyen rendszer szerint megvalósuló ismétlődése.

Megjegyzendő, hogy az ütemek és a szintaktikai frázisok közötti kapcsolatra más terminológiát használva vagy éppen implicit módon régóta utalnak a magyar verstani művek. Már Arany János megfogalmazza *A magyar nemzeti vers-idomról* című művében, „hogy nem elég a verssort, úgy nevezett lábak és metszetek szerint, külsőleg feltagolni, azaz nem elég, hogy bizonyos helyeken végződjék a szó: hanem szükséges, hogy e külső felosztásnak megfelelően a benső: hogy a mondatnak (s így a gondolatnak is) azon részei csoportuljanak egy ütembe, egy, a sormetszet által elkülönzött ízbe, melyek egymással legszorosab kapcsolatban vannak, mert csak így lesz tökélyes a ritmusus”.³⁵ Horváth János úgyszintén felhívja a figyelmet arra, hogy „ha félsorból félsorba, vagy éppen sorból sorba törik át egy gondolati egység, kivált, ha az egyszersmind hangtani és mondattani egység (szólam, szószerkezet): a verstani határnak ezt az áthágását (az ú. n. *enjambement*-t) szabálysértésnek érezzük”.³⁶ Az ütemhangsúlyos metrum leírása során Vargyas Lajos is a szólamtagolás elvéből indul ki.³⁷ Megközelítésében a szólamok „a mondatnak, illetve a beszédnek szorosan összetartozó részei”,³⁸ amelyek lényegében szintaktikai frázisoknak feleltethetők meg. Preskriptívnek ható leírásában úgy fogalmaz, hogy „[ö]sszetartozó szólamokat széttagolhat [...] a vers, mégis jó lehet a ritmus,

³⁵ ARANY János, „A magyar nemzeti vers-idomról,” in ARANY János, *Tanulmányok és kritikák I.*, 288–320 (Debrecen: Debreceni Egyetemi Kiadó, 2012²), 296.

³⁶ HORVÁTH János, *Rendszeres magyar verstan* (Budapest: Akadémiai Kiadó, 1969 [1951]), 25–26.

³⁷ VARGYAS Lajos, *Magyar vers – magyar nyelv: Verstani tanulmány* (Budapest: Szépirodalmi Könyvkiadó, 1966).

³⁸ Uo., 11.

csak össze nem tartozókat ne kényszerítsen egy ütembe; vagyis két különböző szólamba tartozó szó találkozása egy ütemen belül hibás.”³⁹ Szuromi Lajos úgyszintén hangsúlyozza, hogy a több szóból álló ütemek általában szólam-egészt alkotnak és „fokozott értelmi tagoló” funkcióval rendelkeznek.⁴⁰ A szintaktikai megközelítés egyik legkövetkezetesebb alkalmazója Lotz János volt, aki a hangsúly szerepét másodlagosnak tekinti, és az ütemek határait mondattani határokkal hozza összefüggésbe.⁴¹ Hangsúlyozandó, hogy mindez nem jelenti azt, hogy az ütemeknek és a szintaktikai frázisoknak feltétlenül egybe kellene esniük. Az ütemhatárok keresztbe metszhetnek szintaktikai frázisokat anélkül, hogy azt valamiféle metrikai hibának kéne felfogni. Simon Gábor kognitív nyelvészeti kiindulópontú elemzésében részletesen bemutatja, hogy a szintaktikai frázisok és az ütemek közötti eltérés hogyan járulhat hozzá a vers komplex jelentésének a megképződéséhez.⁴²

6. A RÍMELES ÉS A SZINTAKTIKAI ISMÉTLŐDÉS ÖSSZEFÜGGÉSEI A MAGYAR KANONIKUS KÖLTÉSZETBEN

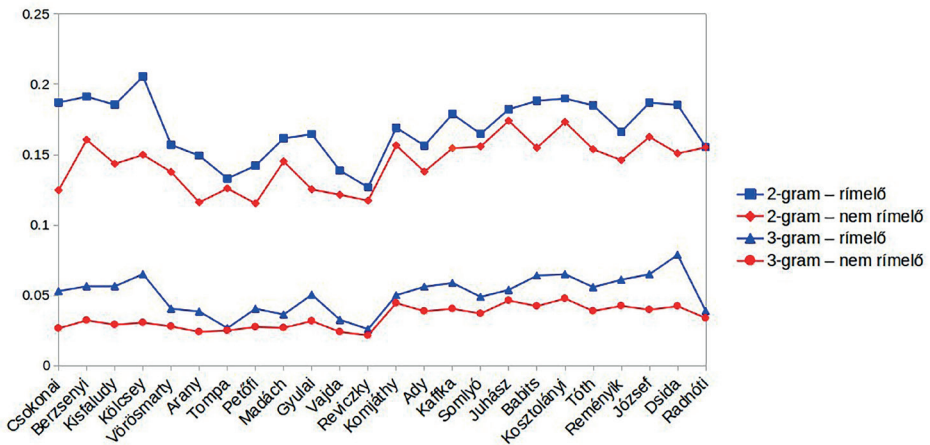
A vizsgálat utolsó kérdése az, hogy vajon a metrum mellett a fonológiai ismétlődésnek egy másik tényezője, a rímelés mutat-e összefüggést a szintaktikai ismétlődés mértékével. A kérdés megválaszolásához a közvetlenül egymást követő egymással rímelő sorok és a közvetlenül egymást követő egymással nem rímelő sorok átfedési együtthatóját számítottam ki sima szó-faji címkék és kiterjesztett szófaji címkék 2- és 3-gramjai alapján. A rímelő és nem rímelő sorpárokra kapott értékeket minden szerző esetében átlagoltam. Az így kapott értékeket a *Függelék* 12. táblázata, illetve a 15. és a 16. ábra mutatja be. A vizsgálatba csak az 1000 szónál rövidebb verseket vettem bele annak érdekében, hogy az átlagok számítását ne torzítsák a hosszú elbeszélő költemények.

³⁹ Uo., 18.

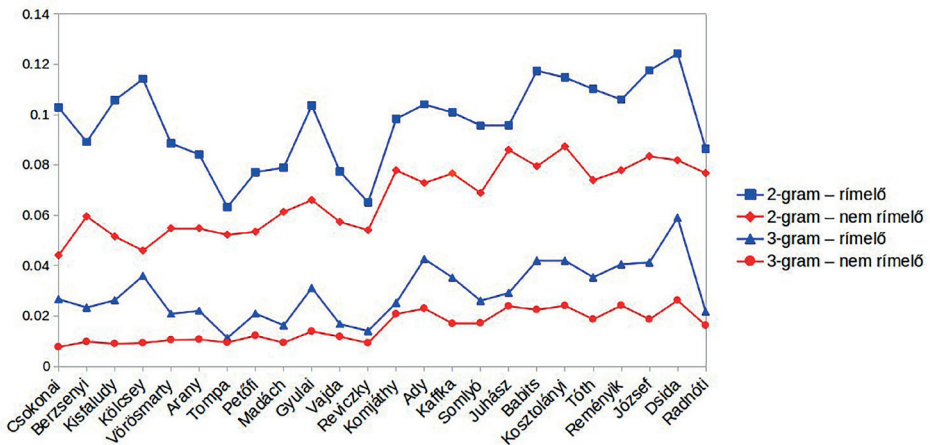
⁴⁰ SZUROMI Lajos, *A szimultán verselés* (Budapest: Akadémiai Kiadó, 1990), 58.

⁴¹ LOTZ János, „Az Uráli verselésről”, in LOTZ János, *Szonettkoszorú a nyelvről*, 237–257 (Budapest: Gondolat Kiadó, 1976), 238.

⁴² SIMON Gábor, „Az ütemhangsúlyos verselés kognitív poétikai magyarázata”, *Magyar Nyelvőr* 138, 4. sz. (2014): 460–478.



15. ábra. Egymást követő rímelő és nem rímelő sorok átfedési együtthatójának átlaga szófaji címkék 2-gramjai és 3-gramjai alapján



16. ábra. Egymást követő rímelő és nem rímelő sorok átfedési együtthatójának átlaga kiterjesztett szófaji címkék 2-gramjai és 3-gramjai alapján

Ahogy az a 12. táblázatból és az ábrákból is kiderül, a közvetlenül egymást követő rímelő sorok átfedési együtthatóinak az átlaga mindig meghaladja a közvetlenül egymást követő nem rímelő sorok átfedési együtthatóinak az átlagát. Mindez azt mutatja, hogy a rímeléssel megvalósuló fonológiai ismétlődés motiválja a szintaktikai ismétlődés megvalósulását. A vizsgálat eredménye tehát egy újabb bizonyítékát szolgáltatja annak, hogy a versekben megjelenő fonológiai és szintaktikai ismétlődés nem független egymástól.

Amennyiben a vizsgált szintaktikai ismétlődést sematikus szinten megvalósuló szemantikai ismétlődéssel azonosítjuk, akkor a rímelő sorpárok esetében a rímeléssel megjelenő fonológiai hasonlóság a sorok szintaktikai hasonlóságában megmutatkozó sematikus szemantikai hasonlóságnak az ikonikus leképeződése. Az utóbbi évtizedekben számos vizsgálat rámutatott a nyelv ikonikus természetére, azaz arra, hogy a nyelvi forma és a jelentés közötti szimbolikus kapcsolat nem teljesen önkényes, a forma, vagyis a nyelv fonológiai oldala valamilyen módon hasonlít az általa kifejezett tartalomra, azaz a nyelv szemantikai oldalára. A nyelv ikonikus természete nyilvánul meg például abban, hogy a szemantikailag közelebb álló elemek alakilag is közelebb helyezkednek el egymáshoz, hogy az információközlés szempontjából fontosabb elemek a mondatban, illetve a szövegben kitüntetett pozícióban állnak, vagy hogy a szemantikailag kevésbé komplex fogalmakat alakilag is kevésbé komplex módon fejezzük ki.⁴³ A rímelő sorpárok esetében a rímelés révén megjelenő fonológiai párhuzamosság és a szintaxison keresztül megmutatkozó szemantikai párhuzamosság tehát forma és tartalom egymást tükröző, ikonikus megfelelésre vezethető vissza. Ebben a viszonyban a rímelés egyszerre motiválja és egyszerre le is képezi a két sor szintaktikai hasonlóságában tetten érhető sematikus szemantikai hasonlóságot. Az eredmények tehát megerősítik azokat az elképzeléseket, miszerint a rím nem csupán fonológiai, hanem fonológiai és szemantikai jelenség.⁴⁴

⁴³ A nyelv ikonikus természetéről például lásd: John HAIMAN, „The Iconicity of Grammar: Isomorphism and Motivation”, *Language* 56, 3. sz. (1980): 515–540, <https://doi.org/10.2307/414448>; Claus D. PUSCH, „Ikonizität”, in *Language Typology and Language Universals: Handbücher zur Sprach- und Kommunikationswissenschaft*, Band 20.1., szerk. Martin HASPELMATH, Ekkehard KÖNIG, Wulf OESTERREICHER és Wolfgang RAIBLE, 369–384 (Berlin, New York: Walter de Gruyter, 2001); Willy van LANGENDONCK, „Iconicity”, in *The Oxford Handbook of Cognitive Linguistics*, szerk. Dirk GEERAERTS és Hubert CUYCKENS, 394–418 (Oxford: Oxford University Press, 2007); BENCZES, Rhyme over Reason...

⁴⁴ A rím jelentéskezdeményező szerepéről lásd SIMON Gábor, *Egy kognitív poétikai rímelmélet megalapozása* (Budapest: Tinta Könyvkiadó, 2014).

7. ÖSSZEGZÉS

A szófaji címkék n-gramjaira épülő vizsgálatokból kiderült, hogy a magyar kanonikus költészetnek a felfokozott fonológiai ismétlődés mellett a felfokozott szintaktikai ismétlődés is egy fontos sajátossága, hiszen a vizsgált szerzők többsége esetében a szintaktikai ismétlődés mértéke meghaladja a regények esetében kapott értéket. A vizsgálatokból azt is láthattuk, hogy a 20. század elején alkotó szerzők esetében megnő a szintaktikai ismétlődés mértéke a 19. században alkotó szerzők nagy részéhez képest. Összevetve a szintaktikai ismétlődés mértékének időbeli alakulását a jambikus és trochaikus metrumú versek metrikai szabályosságának időbeli alakulásával, amellelt érveltem, hogy a szintaktikai ismétlődés mértékének a növekedése a metrikai fellazulást, azaz a fonológiai ismétlődés mértékének a csökkenését kompenzálja. Más szavakkal a poétikai rendszer a versszerűséget azzal őrzi meg, hogy az ismétlődésnek egy adott dimenzióban történő visszaesését egy másik dimenzióban megvalósuló felerősítésével egyenlíti ki.

Emellett azt is sikerült igazolni, hogy a szintaktikai ismétlődés mértéke összefügg a versekben megjelenő fonológiai ismétlődés típusával. Láthattuk, hogy az ütemhangsúlyos metrumú versekben nagyobb mértékben jelenik meg a szintaktikai ismétlődés, mint a jambikus versekben. Ennek magyarázata az lehet, hogy szemben az időmértékes metrumok verslábaival, az ütemek sok esetben teljes szintaktikai frázisoknak feleltethetők meg, az ütemek ismétlődése így motiválhatja a szintaktikai frázisok ismétlődését.

Ugyancsak sikerült igazolni, hogy a fonológiai ismétlődés egy másik megvalósulása, a rímelés is összefüggésbe hozható a szintaktikai ismétlődéssel, hiszen a közvetlenül egymást követő rímelő sorok szintaktikailag jobban hasonlítanak egymásra, mint az egymást követő nem rímelő sorok. Ez a jelenség a rímelés és a szintaxis irányából is magyarázható. Egyrészt a rímeléssel megvalósuló fonológiai párhuzamosság feltételezhetően motiválja a szintaktikailag párhuzamos szerkesztésmódot. Másrészt, ha a szintaxisra sematikus szemantikaként tekintünk, akkor a fonológiai párhuzamosság a szintaktikai hasonlóságban testet öltő szemantikai párhuzamosságnak az ikonikus leképeződése.

A szintaktikai ismétlődés, csakúgy, mint a fonológiai ismétlődés, nem a költészet sajátossága, hanem magának a nyelvnek a sajátossága. Amellett érveltem, hogy a szintaktikai ismétlődés fokozott jelenléte a költészetben – hasonlóan a rímeléssel, ritmussal és egyéb eszközökkel megvalósított fonológiai ismétlődés felerősítéséhez – a nyelv poétikai funkcióját valósítja meg. A poétikus szöveg

esetében a fonológia, a szintaxis és minden bizonnyal a nyelv egyéb, itt nem vizsgált dimenzióiban is megvalósuló felfokozott ismétlés a nyelv egyik leg-alapvetőbb, ugyanakkor a hétköznapi nyelvhasználatban kevésbé szembetűnő mechanizmusát helyezi a figyelem előterébe.

FÜGGELÉK

7. táblázat. A regényekre és versekre kapott szintaktikai ismétlődés értékeinek mediánja szófaji címkék n-gramjai alapján

Szerző/ alkorpusz	100 token			50 token			25 token		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
Regények	0.4237	0.1472	0.0383	0.2871	0.0814	0.0205	0.1671	0.0392	0.0096
Csokonai	0.4465	0.1511	0.0383	0.3087	0.0846	0.0194	0.1841	0.0422	0.0085
Berzsenyi	0.4843	0.1831	0.0444	0.3474	0.1041	0.0232	0.2069	0.0473	0.0091
Kisfaludy	0.4447	0.1468	0.0362	0.3085	0.083	0.0178	0.1843	0.0377	0.0071
Kölcsey	0.4704	0.1829	0.0529	0.3336	0.1042	0.0232	0.2083	0.0505	0.0096
Vörösmarty	0.4381	0.1451	0.0348	0.3044	0.0805	0.0169	0.1771	0.0373	0.0054
Arany	0.4076	0.1238	0.0295	0.2763	0.0687	0.0161	0.1608	0.0335	0.0068
Tómpa	0.4057	0.1238	0.0254	0.2738	0.0654	0.013	0.1569	0.0307	0.0057
Petőfi	0.408	0.1298	0.0302	0.2755	0.071	0.0153	0.1604	0.0344	0.0063
Madách	0.4432	0.1447	0.0329	0.3067	0.0792	0.0159	0.1758	0.0371	0.006
Gyulai	0.4309	0.145	0.0403	0.3006	0.085	0.022	0.1813	0.0445	0.0101
Vajda	0.421	0.1307	0.0297	0.2845	0.0717	0.0157	0.1653	0.036	0.0075
Reviczky	0.4035	0.1192	0.0266	0.2731	0.0652	0.0126	0.1568	0.0317	0.0053
Komjáthy	0.4573	0.1779	0.0545	0.3275	0.1067	0.0298	0.197	0.0558	0.016
Ady	0.4372	0.1594	0.0488	0.3002	0.0933	0.0264	0.18	0.0487	0.0127
Kaffka	0.4123	0.146	0.0403	0.2835	0.0862	0.0222	0.1723	0.0452	0.0117
Somlyó	0.4175	0.1531	0.0416	0.2954	0.0895	0.0239	0.1819	0.0464	0.011
Juhász	0.4676	0.1871	0.0583	0.3368	0.1117	0.0305	0.208	0.0571	0.0141
Babits	0.4267	0.1539	0.0417	0.3023	0.0925	0.0261	0.1883	0.0514	0.0142
Kosztolányi	0.4748	0.1948	0.0606	0.3412	0.1165	0.033	0.2096	0.058	0.0154
Tóth	0.4563	0.1738	0.0485	0.3253	0.1015	0.0262	0.199	0.0516	0.0123
Reményik	0.4122	0.1453	0.0412	0.291	0.0879	0.0227	0.178	0.0476	0.0131
József	0.4505	0.1622	0.0425	0.315	0.0945	0.0241	0.1897	0.0454	0.01
Dsida	0.4354	0.1628	0.0479	0.3081	0.0949	0.0246	0.1875	0.0497	0.0121
Radnóti	0.4632	0.1586	0.0412	0.3342	0.0885	0.021	0.1938	0.0385	0.0069

8. táblázat. A regényekre és versekre kapott szintaktikai ismétlődés értékeinek mediánja kiterjesztett szófaji címkék n-gramjai alapján

Szerző/ alkorpusz	100 token			50 token			25 token		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
Regények	0.2340	0.0578	0.0130	0.1481	0.0324	0.0078	0.0815	0.0166	0.0044
Csokonai	0.2388	0.0509	0.01	0.1473	0.0276	0.0049	0.0822	0.0147	0.0013
Berzsenyi	0.264	0.0609	0.0109	0.1645	0.0332	0.0045	0.0931	0.0155	0.0003
Kisfaludy	0.2428	0.0569	0.0105	0.1662	0.033	0.0055	0.0906	0.0154	0.0009
Kölcsey	0.2525	0.0623	0.0113	0.1565	0.0335	0.0049	0.0857	0.0162	0

Szerző/ alkorpusz	100 token			50 token			25 token		
	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram	2-gram	3-gram	4-gram
Vörösmarty	0.24	0.0529	0.0096	0.1496	0.0281	0.0033	0.0811	0.0123	0
Arany	0.2378	0.0528	0.0107	0.1508	0.0296	0.0055	0.083	0.0158	0.0016
Tómpa	0.2333	0.0498	0.0079	0.1463	0.027	0.0039	0.0775	0.0131	0.0017
Petőfi	0.2416	0.0558	0.0103	0.1516	0.0296	0.0045	0.0833	0.0146	0.0005
Madách	0.2536	0.0563	0.0101	0.1611	0.029	0.0045	0.0853	0.0135	0.0014
Gyulai	0.2762	0.0744	0.0183	0.1879	0.0434	0.0089	0.1122	0.0246	0.0054
Vajda	0.2581	0.0591	0.0104	0.1651	0.0328	0.0059	0.0931	0.017	0.0029
Reviczky	0.2453	0.0551	0.0103	0.1572	0.0294	0.0037	0.0869	0.0157	0.0012
Komjáthy	0.2853	0.0823	0.0236	0.1897	0.0525	0.0146	0.1119	0.0291	0.0085
Ady	0.2821	0.0857	0.0233	0.1884	0.0516	0.014	0.1115	0.0287	0.0068
Kaffka	0.2604	0.0713	0.0152	0.1699	0.0393	0.0095	0.0974	0.0243	0.0055
Somlyó	0.2553	0.0686	0.015	0.1675	0.0407	0.0087	0.0998	0.0229	0.0041
Juhász	0.2949	0.0905	0.0221	0.1992	0.0519	0.0117	0.1164	0.028	0.0051
Babits	0.2665	0.0754	0.02	0.1793	0.0458	0.0136	0.1078	0.0273	0.0082
Kosztolányi	0.3202	0.102	0.0284	0.2142	0.0594	0.0145	0.1236	0.0319	0.0071
Tóth	0.292	0.0833	0.019	0.198	0.0496	0.01	0.1132	0.0251	0.0048
Reményik	0.2559	0.0735	0.0201	0.174	0.0462	0.0116	0.1044	0.0277	0.0061
József	0.2918	0.0816	0.018	0.1929	0.0441	0.0075	0.11	0.0234	0.0027
Dsida	0.2751	0.0814	0.0206	0.1869	0.0474	0.0109	0.1084	0.0263	0.0049
Radnóti	0.2918	0.0714	0.0121	0.187	0.0376	0.0053	0.1029	0.0159	0

9. táblázat. A nem szimultán jambikus és trochaikus versek száma és szabályossági pontszámainak mediánja 0,5 és 0,3 szabályossági küszöbértékkel

Szerző	Összes	Jambikus				Trochaikus			
		0,5 <		0,3 <		0,5 <		0,3 <	
		db	Medián	db	Medián	db	Medián	db	Medián
Csokonai	394	46	0.83	64	0.76	50	0.72	118	0.46
Berzsenyi	136	20	0.7	45	0.49	1	0.56	19	0.36
Kisfaludy	123	19	0.74	21	0.74	11	0.69	22	0.53
Kölcsy	150	44	0.66	58	0.64	21	0.69	23	0.68
Vörösmarty	662	64	0.76	94	0.7	38	0.78	82	0.45
Arany	417	84	0.72	96	0.72	40	0.65	63	0.56
Tómpa	491	226	0.72	239	0.72	32	0.65	38	0.63
Petőfi	839	249	0.68	292	0.66	81	0.65	102	0.64
Madách	319	97	0.72	107	0.71	45	0.74	49	0.73
Gyulai	156	33	0.66	37	0.65	27	0.64	41	0.61
Vajda	199	71	0.64	82	0.63	41	0.63	54	0.61
Reviczky	335	223	0.78	223	0.78	19	0.7	20	0.69
Komjáthy	246	169	0.77	178	0.77	9	0.7	14	0.67
Ady	1116	203	0.57	585	0.46	35	0.6	107	0.45
Kaffka	102	38	0.66	52	0.6	4	0.56	6	0.53
Somlyó	379	155	0.63	202	0.6	10	0.58	40	0.43
Juhász	1278	666	0.69	720	0.69	14	0.6	37	0.47
Babits	514	200	0.67	246	0.65	33	0.61	62	0.52
Kosztolányi	630	440	0.74	471	0.74	19	0.76	44	0.46
Tóth	451	185	0.65	274	0.6	10	0.55	44	0.41
Reményik	667	458	0.59	605	0.56	1	0.61	4	0.35
József	599	184	0.61	287	0.56	21	0.58	69	0.41
Dsida	1060	295	0.58	634	0.5	9	0.66	67	0.37
Radnóti	298	60	0.75	112	0.52	4	0.54	32	0.36

10. táblázat. A minimum 50 tokenből álló versek száma az egyes metrumkategóriákban szerzőnként, 0,5 küszöbértékkel

	Összes	Jambikus	Trochaikus	Ütemh.	TrochSzim.	JambSzim.
Csokonai	338	37	45	123	8	6
Berzsenyi	122	20	0	14	0	14
Kisfaludy	109	16	11	28	15	4
Kölcsey	129	42	19	5	13	8
Vörösmarty	424	56	33	25	55	49
Arany	315	77	35	54	84	14
Tómpa	474	222	31	38	83	73
Petőfi	738	224	79	130	143	85
Madách	245	93	44	3	54	37
Gyulai	152	33	26	22	38	10
Vajda	190	69	39	14	27	14
Reviczky	317	216	18	3	35	41
Komjáthy	220	164	7	8	9	11
Ady	1075	197	32	111	66	48
Kaffka	101	37	4	6	8	3
Somlyó	377	155	10	28	10	62
Juhász	1190	649	12	39	84	315
Babits	502	196	32	40	24	44
Kosztolányi	582	417	17	10	6	54
Tóth	407	169	9	20	22	29
Reményik	639	440	1	14	6	28
József	552	178	19	53	24	51
Dsida	987	286	8	70	30	67
Radnóti	272	59	3	9	2	12

11. táblázat. 50 szavas ablakkal szófaji címkék és kiterjesztett szófaji címkék 2-, 3-, és 4-gramjaira kapott ismétlődési értékek mediánja metrumkategóriánként

	Szófaji címkék						Kiterjesztett szófaji címkék					
	2-gram		3-gram		4-gram		2-gram		3-gram		4-gram	
	Jamb.	Ütem.	Jamb.	Ütem.	Jamb.	Ütemh.	Jamb.	Ütem.	Jamb.	Ütem.	Jamb.	Ütemh.
Szerző	0.3014	0.3203	0.0802	0.0914	0.0128	0.0219	0.1536	0.1489	0.0351	0.0276	0.0013	0.0055
Csokonai	0.3014	0.3203	0.0802	0.0914	0.0128	0.0219	0.1536	0.1489	0.0351	0.0276	0.0013	0.0055
Berzsenyi	0.3224	0.3478	0.0948	0.118	0.0227	0.0249	0.1578	0.1557	0.029	0.0378	0.005	0.0052
Kisfaludy	0.3175	0.293	0.0725	0.079	0.0162	0.0149	0.1436	0.1735	0.0273	0.0293	0.0052	0.0024
Kölcsey	0.3469	0.3373	0.1091	0.1001	0.0275	0.0238	0.1618	0.156	0.0405	0.0322	0.007	0.0064
Vörösmarty	0.2898	0.3109	0.0761	0.0888	0.0165	0.02	0.1575	0.1612	0.0336	0.0317	0.0056	0.007
Arany	0.2637	0.2848	0.0657	0.0735	0.013	0.0172	0.148	0.1525	0.028	0.0315	0.0049	0.006
Tómpa	0.2747	0.2691	0.0649	0.0644	0.0123	0.0133	0.144	0.143	0.0269	0.0249	0.0038	0.0039
Petőfi	0.2679	0.2831	0.067	0.0747	0.0138	0.0173	0.1493	0.1559	0.0274	0.0342	0.0042	0.0053
Madách	0.3089	0.3107	0.0789	0.081	0.0164	0.0167	0.1623	0.1579	0.0306	0.027	0.004	0.0046
Gyulai	0.3002	0.306	0.0831	0.087	0.0213	0.0232	0.185	0.1933	0.0417	0.0459	0.0079	0.0088
Vajda	0.2828	0.292	0.0716	0.0745	0.0143	0.0157	0.1752	0.1668	0.0341	0.0307	0.0058	0.0042
Reviczky	0.2709	0.2822	0.0633	0.0752	0.0119	0.0149	0.1552	0.1627	0.0295	0.0399	0.0039	0.0065
Komjáthy	0.321	0.3802	0.0983	0.1649	0.0264	0.08	0.1835	0.2195	0.0471	0.0629	0.0122	0.0279
Ady	0.3056	0.3008	0.0955	0.1027	0.0258	0.0323	0.1949	0.1802	0.0565	0.0554	0.0148	0.017
Kaffka	0.2816	0.328	0.0789	0.1111	0.019	0.0255	0.1617	0.1725	0.0372	0.0399	0.0091	0.0074
Somlyó	0.2938	0.3152	0.0849	0.1026	0.0222	0.0245	0.162	0.1785	0.0367	0.0452	0.0076	0.0096
Juhász	0.3402	0.3408	0.1108	0.1199	0.0291	0.037	0.1981	0.2041	0.0495	0.0546	0.0107	0.0109
Babits	0.2999	0.3096	0.0871	0.0961	0.0247	0.0235	0.1789	0.1807	0.0421	0.0514	0.011	0.015
Kosztolányi	0.341	0.3213	0.1141	0.1292	0.0316	0.046	0.2143	0.209	0.0579	0.0589	0.0141	0.0135

	Szófaji címkék						Kiterjesztett szófaji címkék					
	2-gram		3-gram		4-gram		2-gram		3-gram		4-gram	
Szerző	Jamb.	Ütem.	Jamb.	Ütem.	Jamb.	Ütemh.	Jamb.	Ütem.	Jamb.	Ütem.	Jamb.	Ütemh.
Tóth	0.334	0.3282	0.105	0.0974	0.0264	0.0263	0.2024	0.1922	0.0496	0.0527	0.009	0.0106
Reményik	0.2883	0.3097	0.0864	0.1065	0.0226	0.036	0.1737	0.1902	0.0452	0.0697	0.0113	0.0213
József	0.3025	0.3169	0.0838	0.1025	0.0208	0.0284	0.1828	0.1981	0.0407	0.0476	0.0073	0.0087
Dsida	0.3046	0.3008	0.0881	0.0867	0.0213	0.0247	0.1833	0.1711	0.0448	0.0419	0.0091	0.0111
Radnóti	0.3255	0.3117	0.0822	0.0808	0.017	0.0071	0.195	0.1874	0.0383	0.0329	0.0057	0.004

12. táblázat. A rímelő és nem rímelő sorpárok átfedési együtthatóinak átlaga az 1000 szónál rövidebb versekben

Szerző	Szófaji címkék				Kiterjesztett szófaji címkék			
	2-gram		3-gram		2-gram		3-gram	
	Rímelő	Nem rímelő	Rímelő	Nem rímelő	Rímelő	Nem rímelő	Rímelő	Nem rímelő
Csokonai	0.1871	0.1247	0.053	0.0265	0.1029	0.0441	0.0267	0.0077
Berzsenyi	0.1913	0.1606	0.0566	0.0322	0.0892	0.0596	0.0234	0.0098
Kisfaludy	0.1855	0.1435	0.0566	0.0291	0.1058	0.0516	0.0263	0.009
Kölcsey	0.2056	0.1499	0.0651	0.0306	0.1142	0.046	0.036	0.0093
Vörösmarty	0.157	0.1377	0.0405	0.0279	0.0886	0.0549	0.0209	0.0105
Arany	0.1493	0.1161	0.0385	0.024	0.0842	0.0549	0.0221	0.0107
Tómpa	0.1331	0.126	0.0267	0.025	0.0633	0.0523	0.0113	0.0095
Petőfi	0.1422	0.1153	0.0405	0.0275	0.0772	0.0535	0.021	0.0122
Madách	0.1616	0.1452	0.0363	0.0269	0.079	0.0614	0.0163	0.0094
Gyulai	0.1646	0.1253	0.0505	0.0317	0.1037	0.0661	0.0312	0.0139
Vajda	0.1387	0.1215	0.0326	0.024	0.0774	0.0574	0.0168	0.0118
Revczky	0.1268	0.1173	0.0259	0.0215	0.0652	0.0541	0.0141	0.0093
Komjáthy	0.1692	0.1566	0.0501	0.0445	0.0984	0.0779	0.0253	0.0208
Ady	0.1565	0.1379	0.0561	0.0387	0.1041	0.0729	0.0427	0.023
Kaffka	0.1789	0.1546	0.0588	0.0404	0.1009	0.0767	0.0353	0.017
Somlyó	0.1648	0.1558	0.0489	0.037	0.0956	0.0689	0.0261	0.0172
Juhász	0.1823	0.1741	0.0539	0.0463	0.0957	0.086	0.0292	0.0239
Babits	0.1883	0.1549	0.064	0.0423	0.1175	0.0795	0.0421	0.0225
Kosztolányi	0.1899	0.1732	0.065	0.0477	0.1148	0.0874	0.042	0.0241
Tóth	0.1851	0.1538	0.0558	0.0388	0.1102	0.074	0.0354	0.0187
Reményik	0.1664	0.1461	0.0611	0.0425	0.106	0.0779	0.0406	0.0242
József	0.187	0.1626	0.065	0.0397	0.1176	0.0835	0.0413	0.0187
Dsida	0.1854	0.1509	0.0788	0.0423	0.1243	0.0819	0.0591	0.0262
Radnóti	0.1554	0.1552	0.0388	0.0338	0.0865	0.0768	0.0217	0.0163

SZEMES BOTOND
Szavak a hálóban

Az irodalomtörténet hálózatos megjelenítése

ORCID: 0000-0002-0637-6776
botond.szemes@ut.ee

HELIKON

Words in the Web. Network-Based Representation of Literary History

Abstract

This paper provides a brief overview of diagrammatic techniques commonly used in (computational) literary studies to illustrate historical processes. After examining both their epistemic potential and limitations, it explores the role of networks in literary historiography, in particular how relationships between literary works, authors, and historical contexts can be visualised and analysed. Two case studies demonstrate the application of the stylometric method *Bootstrap Consensus Network* in literary history. In the first, the stylistic development of 100 Hungarian novels from 1832 to 2005 is analysed in order to identify major trends in Hungarian literary history. The second example focuses on a smaller collection to examine the stylistic diversity of a specific literary period (the early 20th century, “modernism”). Ultimately, the paper argues that network-based literary history provides a more nuanced understanding of literary evolution.

Keywords: diagram, network, literary history, Hungarian novel, stilometry

Franco Moretti *Graphs, Maps, Trees* című könyve nemcsak abban gyakorolt nagy hatást a (digitális) irodalomtudomány alakulására, hogy felhívta a figyelmet a diagramok szerepére az ismeretek előállításában, hanem abban is, hogy kijelölte a megjelenítések azon formáit, amelyek mai napig meghatározzák, miként ábrázoljuk az irodalmi szövegek közötti kapcsolatokat.¹ Ami az első felismerést illeti: bár még mindig gyakori, hogy a kartografikus-diagrammatikus eljárásokat csak mint ötletszerűen felrajzolt ábrákat, semmint valódi adatvizualizációkat értik (félre) a kutatók, egyre elterjedtebbek a diagramok episztemológiai potenciálját az érvrendszer szerves részeként hasznosító tanulmányok. A diagrammatikus megjelenítések ugyanis nem pusztán jelenségek szemléltetésére alkalmasak (érintetlenül hagyva a nem kvantifikálhatónak elgondolt lényegét), hanem az ismeretek létrehozásának nagyon is fontos ágensei.² Éppen ezért érdemes tudatosítanunk azt is, hogy az egyes megjelenítési módok milyen típusú következtetéseket tesznek lehetővé, miként rendezik el és láttatják szövegek kapcsolatát.

Az irodalomtörténet megjelenítésekor a Moretti által alkalmazott eljárások közül leggyakrabban a statisztikától (szorosabban a kvantitatív történetírástól) kölcsönzött idősoros grafikonok és – kisebb mértékben – az evolúciós elmélektől kölcsönzött leszármazási fák alkalmazása tekinthető gyakorinak. Mindkét típus szemléletesen és nagy magyarázóerővel tudja a felület síkján, térbeli viszonyok formájában bemutatni és kezelhetővé tenni a máskülönben nehezen átlátható időbeli folyamatokat. Azonban ennél többről van szó: ezek a diagramok nem ártatlan közvetítők, hanem alapvetően határozzák meg az időbeliségről alkotott elképzelésünket. Máshol mellett érveltem, hogy a viszonylag későn, a 18. század végén megjelenő „statisztikai, idősoros grafikonok hozták létre, valamint adtak formát a modernitás időszemléletének, illetve a történelem és a történetek szerkezetére vonatkozó, a mai napig érvényes elképzeléseknek.”³ A lineáris, fejlődéselvű, az időpontokat egyetlen vonal mentén összefűző, ezáltal a múlt, a jelen és a jövő kategóriáit elhatároló és megszakításmentesen összekapcsoló időszemlélet alapja az idősoros grafikon tendenciagörbéje.⁴ Napjaink-

¹ Franco MORETTI *Graphs, Maps, Trees* (London: Verso, 2003).

² Sybille KRÄMER, „Operatív képiség: A »grammatológiától« a »diagrammatológiáig«?: Gondolatok a megismerő »látásról«, ford. ROSKÓ Mira, *Helikon* 67, 2. sz. (2022): 277–308; SZEMES Botond, „A kultúra mint diagram: Előszó a *Helikon Diagrammatológia* című lapszámához”, *Helikon* 67, 2. sz. (2022): 267–276.

³ SZEMES Botond, „Az idővonal felemelkedése: Szemelvények az idő grafikus leképezésének európai kultúrtörténetéből”, *Literatura* 49, 3. sz. (2023): 269–288, <https://doi.org/10.57227/Liter.2023.3.1>.

⁴ Ezek a vonalak az egyes adatpontok vagy azok átlagának összekötéseként, esetleg (a statisztika későbbi alakulásával) az adatpontokra illesztett regressziós egyenesként állnak elő. A regressziós egyenes olyan becsült értékek alapján jön létre, amelyek y valószínűsített

ban nemcsak az ilyen grafikonok burjánzása feltűnő, hanem az is, milyen erősen strukturálja ez a leképezési mód gondolkodásunkat és nyelvhasználatunkat – ma már nem is tudunk olyan nagy kiterjedésű folyamatokról beszélni, mint a járványok terjedése, a klímaváltozás vagy az árfolyamok alakulása, anélkül hogy ne grafikonokra utalnánk (vö. „a görbe ellaposítása”, „a forint bezuhant”). Mindez a narratív ív alapján szerveződő elbeszélő művek felépítésén túl a kvantitatív irodalomtörténet-írásra is komoly hatást gyakorolt. A digitális bölcsészet bevett eljárása egy változó eredményeit és a szövegek keletkezési/megjelenési idejét egymás függvényében ábrázolni, majd a kirajzolódó tendenciákból irodalomtörténeti változásokra következtetni.

Egy trendvonalal ellátott idősoros grafikon azonban nemcsak az időbeliségre vonatkozó elgondolásunkat befolyásolja (azaz, hogy nem tesz elgondolhatóvá nem lineáris, ciklikus, illetve töredezett folyamatokat), hanem azt is, milyen mértékben vonatkoztatunk el a konkrét megfigyelésektől az általános tendenciák azonosítása érdekében. Moretti saját korábbi munkásságát is kritizálva a túlzó általánosítás veszélyre hívja fel a figyelmet Oleg Sobchukkal közösen írt cikkében:

[a kutatók] nem látják saját adataikat? Dehogynem; de a trendvonalak megváltoztatták a *látásmódunkat*: a statisztikai absztrakciókat olyan fizikai jelenlétté alakították át, amelyek ugyanolyan valóságosak, mint maguk az adatok – sőt általában *sokkal láthatóbbak*, mint maguk az adatok. A vizualizáció az intuíciónkra apellál; ha pontok felhőjét mutatja egy vonallal a közepén, akkor csak a vonalat vesszük észre. Ez elkerülhetetlen. És így ahelyett, hogy segítettek volna elemezni a bizonyítékokat, az átlagok gyakran ahhoz járultak hozzá, hogy elfelejtsük azokat. Azért fordulunk a számszerűsítéshez, mert látni akartuk mindazokat a dokumentumokat, amelyeket a kánon túlsúlya láthatatlanná tett – és most, hogy a szemünk előtt vannak, megtaláltuk a módját, hogy ne lássuk őket!⁵

helyzetét (tehát egy függő változó mértékét) adják meg x függő változóhoz (jelen esetben az időpontokhoz) képest úgy, hogy a becslő értékek és a valódi adatpontok közti különbség a lehető legkevesebb legyen. A görbeillesztést statisztikai egyenletek vezérlik, amelyekben előre meghatározható a görbe formája is: lineáris regresszió esetén egy egyenes vonalat kapunk, más esetek megengedik a vonal görbülését.

⁵ Franco MORETTI és Oleg SOBCHUK, „Hidden in Plain Sight: Data Visualization and Digital Humanities”, *New Left Review* 118 (2019): 86–115, 94–95, <https://doi.org/10.64590/c8k>.

Ebben az idézetben jól tetten érhető a diagramok használatában mindig is jelen lévő feszültség. Míg a diagramok éppen a sematizmus és az absztrakció révén képesek érvényes összefüggéseket felismerhetővé tenni, addig jogosan merülhet fel, hogy ennek során túlságosan leegyszerűsítik és uniformizálják a vázolt jelenségeket. A nagy mintázatokban könnyen eltűnnek az egyéni pozíciók és kapcsolatok; miközben az adatpontok önmagukban való, rendezetlen bemutatása gátolhatja az történet- és/vagy elméletalkotást.

Meglehet ugyanakkor, hogy az ebből adódó kritikákat nem csupán az idő-soros grafikonok korlátozott teljesítőképessége váltja ki, hanem megszilárdult pozíciójuk is: mára máshogy se igazán tudjuk elképzelni az időbeli folyamatok ábrázolását, holott a diagrammatikus gondolkodás éppen a fennálló struktúrák folyamatos újrarendezésére irányulna.⁶ Az ilyen grafikonok begyakorolt, magától értetődő használata viszont egyetlen látásmód megszilárdulását eredményezi.

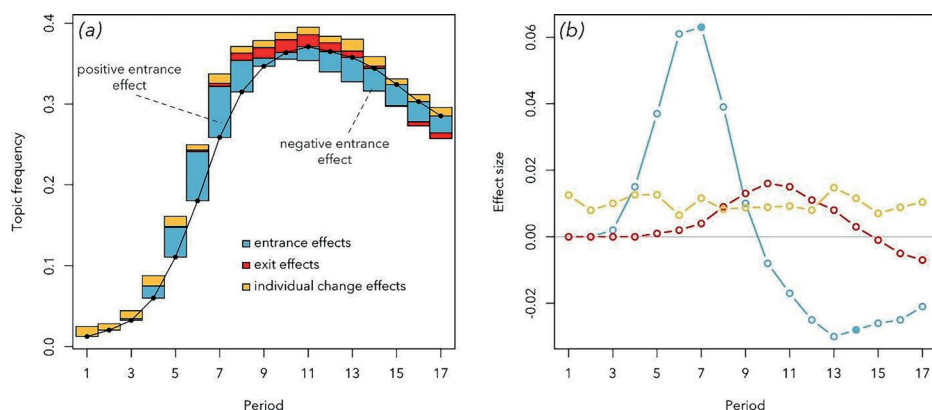
Éppen ezért kísérleteznek köztes megoldásokkal a kulturális evolúció területén a kutatók. A jelenségek mögött meghúzódó folyamatok átlátható bemutatását célozzák a komplexitás lehető legkisebb mértékű redukciójával. Ilyen például a trendvonalak felbontása különböző, egymást hol erősítő, hol egymásnak ellentartó tényezőkre – lásd például az 1. ábrát, többek között az imént idézett Sobchuk egyik újabb kutatásából.⁷ Ez esetben is a háttérben maradnak az adatpontok, a kirajzolódó mintázat azonban összetettebb bemutatást nyer, mintha egyetlen trendvonallal szemléltetnénk.

A leszármazási fákat még erősebb kritika érte az elmúlt évtizedekben – még ha sok esetben hasznos modellnek is bizonyulnak⁸ –, amennyiben nemcsak leegyszerűsítőnek, hanem egyenesen alkalmatlannak mutatkoztak a kulturális változások leírására. Sobchuk még megengedően fogalmaz, amikor azt írja, hogy a fa-szerkezet bizonyos területeken, mint például a nyelvtörténet szemléltetésére valóban hasznos lehet, csupán más jelenségek vizsgálatakor nem alkalmazható jól – például a detektívtörténet műfajának rekonstrukciójá-

⁶ Vö. Gilles DELEUZE, „Egy új kartográfus”, ford. SIPOS Balázs, *Helikon* 67, 2. sz. (2022): 324–344.

⁷ Oleg SOBCHUK és Bret BEHEIM, „Does Literature Evolve One Funeral at a Time?”, *Proceedings of the Royal Society B: Biological Sciences* 292, 2040. sz. (2025): 20242033, <https://doi.org/10.1098/rspb.2024.2033>. Lásd még Bret BEHEIM és Ryan BALDINI, „Evolutionary Decomposition and the Mechanisms of Cultural Change”, *Cliodynamics: The Journal of Quantitative History and Cultural Evolution* 3, 2. sz. (2012), <https://doi.org/10.21237/C7CLIO3212123>.

⁸ Mark COLLARD, Stephen J. SHENNAN és Jamshid J. TEHRANI, „Branching, Blending, and the Evolution of Cultural Similarities and Differences among Human Populations”, *Evolution and Human Behavior* 27, 3. sz. (2006): 169–184, <https://doi.org/10.1016/j.evolhumbehav.2005.07.003>.



1. ábra. Kísérlet a trendvonalat befolyásoló hatások megjelenítésére. Oleg Sobchuk és Bret Beheim, *Does Literature Evolve One Funeral at a Time?*

kor.⁹ Ugyanakkor a leszármazási fák nyelvtörténeti alkalmazhatósága is erősen kétségszoros.¹⁰ Farkas Zsolt szerint a diszciplína egészét siklatta ki a helytelen modell alkalmazása: „De a szakma tévedésrendszerének eredeti és legfőbb oka a fa-modell, ami a kollektív örület nemzetközi vetülete. Eszerint egy nyelvnek nem lehet több, különböző eredetű és meghatározó alkotóeleme, nincs nyelvkeveredés, a jövevényelemek lényegtelenek. Az indoeurópai fa és a finnugor fa minden egyéb nyelv komolyabb hatása nélkül fejlődik.”¹¹ A legfőbb problémát az ilyen leképezésekkel a horizontális kapcsolatok figyelmen kívül hagyása jelenti, illetve annak az esetnek a kizárása, amikor egy „utód” egyszerre több „szülőtől” származik. Ismét a túlzott sematizmus veszélye, hogy elrejtje az elemek közötti valódi viszonyrendszert. Mivel „az emberi történelemben a leszármazási útvonalak közötti átvitel talán a kulturális változás fő forrása”, a különálló ágakból felépülő „biológiai evolúció rossz

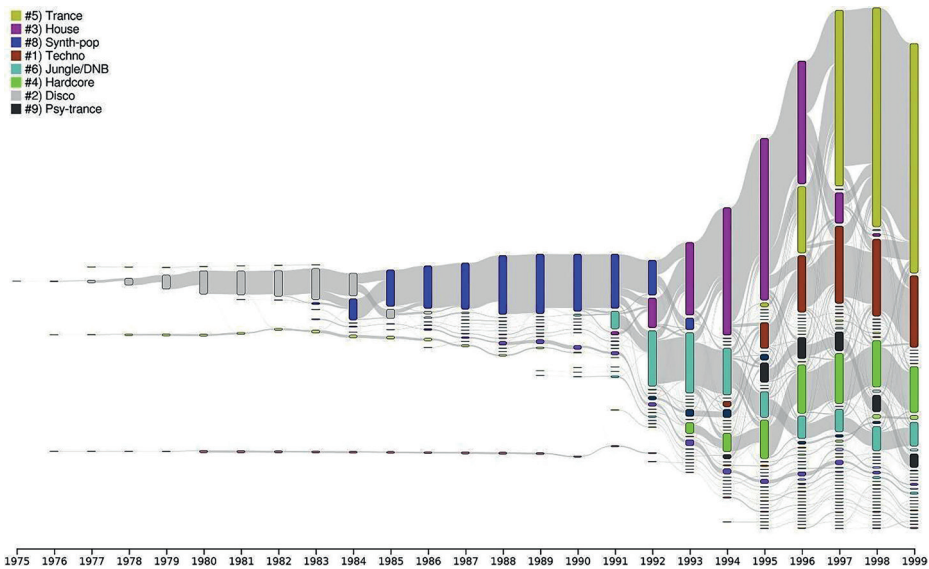
⁹ Oleg SOBCHUK, *Charting Artistic Evolution: An Essay in Theory* (Tartu: University of Tartu Press, 2018), 64, bővebben: 68. Sobchuk ezzel Morettinek a detektívtörténet műfajának kialakulását szemléltető híres ágrajzát is kritika alá vonja.

¹⁰ Stephen C. LEVINSON, Stephen C. és GRAY Russell D., „Tools from Evolutionary Biology Shed New Light on the Diversification of Languages”, *Trends in Cognitive Sciences* 16, 3. sz. (2012): 167–173, <https://doi.org/10.1016/j.tics.2012.01.007>; Remco BOUCKAERT, Philippe LEMEY, Michael DUNN, Simon J. GREENHILL, Alexander V. ALEKSEYENKO, Alexei J. DRUMMOND, Russell D. GRAY, Marc A. SUCHARD és Quentin D. ATKINSON, „Mapping the Origins and Expansion of the Indo-European Language Family”, *Science* 337, 6097. sz. (2012): 957–960, <https://doi.org/10.1126/science.1219669>. (Kiemelés tőlem – Sz. B.)

¹¹ FARKAS Zsolt és JÁNOSSY Lajos, „Digitális és drámai (interjú)”, *Litera*, hozzáférés: 2025.10.14, <https://litera.hu/magazin/interju/farkas-zsolt-digitalis-es-dramai.html>.

analógiája a kulturális változásoknak”.¹² Sőt, a mikrobiológus Eugene Koonin még azt is kétségbe vonja, hogy az életfa-moddell a sejtmaggal rendelkező állatok egy kis csoportján kívül érvényes lenne a földi élet egészére, amelyre ugyanis éppen annyira jellemző a horizontális génátvitel is.¹³

Emiatt a leszármazási fák újragondolása szintén kiemelten fontossá vált a kulturális evolúció területén. Mason Youngblood és társai például az elektronikus zenei műfajok evolúcióját a több helyről való származást is megengedve ábrázolták tanulmányukban (2. ábra).¹⁴



2. ábra. A leszármazási fák újragondolása – zenei műfajok evolúciója. Youngblood et. al., *Phylogenetic Reconstruction of the Cultural Evolution of Electronic Music via Dynamic Community Detection (1975–1999)*

¹² Stephen Jay GOULD, *Bully for brontosaurus: reflections in natural history* (New York: Norton, 1991), 64. (Kiemelés tőlem – Sz. B.)

¹³ Eugene V. KOONIN, *The Logic of Chance: The Nature and Origin of Biological Evolution* (Upper Saddle River, NJ: Pearson Education, 2012), 165. Hasonló érveket mutat be: W. Ford DOOLITTLE és Eric BAPTESTE, „Pattern Pluralism and the Tree of Life Hypothesis”, *Proceedings of the National Academy of Sciences* 104, 7. sz. (2007): 2043–2049, <https://doi.org/10.1073/pnas.0610699104>.

¹⁴ Mason YOUNGBLOOD, Karim BARAGHITH és Patrick E. SAVAGE, „Phylogenetic Reconstruction of the Cultural Evolution of Electronic Music via Dynamic Community Detection (1975–1999)”, *Evolution and Human Behavior* 42, 6. sz. (2021): 573–582, <https://doi.org/10.1016/j.evolhumbehav.2021.06.002>.

Ez az ábra már inkább hálózatos elrendezést mutat,¹⁵ mintsem egy megszokott ágrajzot, bár még őrzi annak jellegzetességeit, különösen az évekre bontott „pillanatfelvételek” és az előre meghatározott műfaji csoportok következtében. Az alábbiakban ezektől a kötöttségektől is megszabadulva, valódi hálózatokra és az irodalomtörténet hálózatos megjelenítésére hozok példát a digitális irodalomtudomány területéről. A hálózatok ugyanis megfelelő formának tűnnek, hogy az eddig vizsgált eljárások gyengeségeit orvosolni tudjuk: az elemek között bármilyen kapcsolatot megengednek (tulajdonképpen a horizontális–vertikális megkülönböztetés is érvényét veszti), és nem rejtik el az adatpontokat egyetlen tendenciagörbe mögé (ellenkezőleg: az egyes csomópontok kapcsolatát helyezik az előtérbe). Mindeközben, habár e tulajdonságait többek között annak köszönhetik, hogy nincs előre megszabott irányuk, a hálózatok mégis képesek lehetnek történeti folyamatokat szemléltetni.

Ez utóbbin magam is meglepődtem. Amikor ugyanis 100 magyar regény stílár¹⁶ hasonlóságát akartam szemléltetni egy hálózat formájában, nem vettem figyelembe a művek keletkezési idejét; a csomópontok elhelyezkedését csak a hasonlóságokon alapuló kapcsolatrendszer határozta meg. Az ábrán mégis egyértelműen kirajzolódik a magyar regényirodalom történetisége (3. ábra): a bal oldalon a 19. század közepéből származó szövegek találhatók, majd szinte időrendben haladva jutunk el a 20. század második felében írt szövegekig az ábra jobb oldalán. Az eredeti (színes, és így nem csak szebb, de az alcsoportokat is jobban elkülönítő) hálózat *Digitális irodalomtudomány* című könyvem borítóján található,¹⁷ ám a főszövegben nem szerepel, így joggal rója fel a kritikus, hogy „a kötet borítóján található hálózat, amely vizuálisan

¹⁵ Ez irányba módosítja a leszármazási fákat Tehrani is híres, *A pirooska és a farkas* történetének leszármazását vizsgáló kutatásában: Jamshid J. TEHRANI, „The Phylogeny of Little Red Riding Hood”, szerk. R. Alexander BENTLEY, *PLoS ONE* 8, 11. sz. (2013): e78871, <https://doi.org/10.1371/journal.pone.0078871>.

¹⁶ Fontos megjegyezni, hogy a vizsgálat az irodalomtörténetet stílustörténetként látta, figyelmen kívül hagyva más megközelítéseket, amelyek például a nyelvhez és annak uralhatóságához való viszony, valamint a világ megismerhetőségére vonatkozó képzetek változásai mentén tartják elmondhatónak az irodalomtörténeti korszakok egymásra következését. Itt nem az elbeszélést meghatározó ismeretelméleti előfeltevések és az irodalmi nyelv működésére adott reflexiók válnak fontossá, hanem pusztán a nyelvhasználat hasonlóságai és különbségei. Ez sok esetben ugyanakkor, ahogy látható is lesz, visszaigazolja a létező poétikátörténeti narratívákat, ám lehetővé teszi azt is, hogy ne paradigmaváltások sorozatát lássuk az irodalom történetében, hanem változások és összeköttetések – hálózatos – rendszerét.

A vizsgált korpusz összetételének leírását lásd: SZEMES Botond, „Mondathosszúság és irodalomtörténet”, *Literatura* 46, 3. sz. (2020): 337–342.

¹⁷ Elérhető a tanulmány repozitóriumban: <https://github.com/SzemesBotond/literary-networks> '100regény' mappa network-color.png fájl.

erőteljes és figyelemfelkeltő, sajnos nem kap érdemi elemzést a könyvben”.¹⁸ Az alábbi ismertetés ezt a hiányosságot is pótolni igyekszik.

Még a megjelenítés előtt fontos kérdés, hogy mire vonatkozik a „stiláris hasonlóság” megjelölés. Ennek meghatározásakor kvantitatív, a stilometriában bevett megközelítését érvényesítettem, amely nyelvi elemek eloszlására vezeti vissza a stílus fogalmát – eltérően más, kvalitatív szempontoktól.¹⁹ Egy-szerűen összeszámoltam a leggyakoribb szavak relatív gyakoriságát a regényekben, hogy milyen eséllyel fordulnak elő egy szövegben. Két regény stílusa ebben a keretben akkor hasonlít egymásra, ha hasonló mértékben fordulnak elő bennük a leggyakoribb szavak (amelyek elsősorban olyan tartalmas jelentés nélküli, ún. funkciósavak, mint a névelők vagy a kötőszavak).²⁰ Ez a szerzőazonosítást célzó kutatások alapvető eljárása.²¹ Ilyen célú alkalmazhatóságát a 3. ábra is jól szemlélteti: a történetiség mellett a szövegek helyzetére a szerzőség van a legnagyobb hatással, vagyis az azonos szerzőségű művek kerültek a legközelebb, legszorosabb kapcsolatba egymással.

Ezt követően a hálózatot a Maciej Eder által kifejlesztett *Bootstrap Consensus Network* módszerével hoztam létre.²² Ez azon a felismerésen nyugszik, hogy a stilometriai kutatásokban sok múlik azon, hogy milyen beállítások mentén számítjuk ki szövegek hasonlóságát (Milyen elemek gyakoriságát vizsgáljuk? Hány ilyen elem alapján végezzük a számítást? A szövegek egészét

¹⁸ MAKKAJ T. Csilla, „Az adatok vonzásában”, *Szépirodalmi Figyelő* 23, 4. sz. (2024): 116–120, 118.

¹⁹ Részletesebben lásd J. Berenike HERRMANN, „Revisiting Style a Key Concept in Literary Studies”, *Journal of Literary Theory* 9, 1. sz. (2015): 25–52, <https://doi.org/10.1515/jlt-2015-0003>; SZEMES Botond, *Digitális irodalomtudomány* (Budapest: Kijárat Kiadó, 2024), 50–70.

²⁰ Felvethető ugyanakkor, hogy a helyesírás változása befolyásolja az eredményeket. A korpusz leggyakoribb szavait vizsgálva látható a 19. században jellemző 'c' hang „cz”-ként való átírása; esetenként rövid magánhangzó szerepeltetése hosszú helyett („ugy”, „igy”, „mig”); a 'j' hang kettőzése („olly”) vagy a vonatkozó névmások különírása: „a ki”, „a mi”, „a mely”. Ez utóbbi növeli egy szövegben az „a” névelők és a „ki”, „mi” névmások számát, míg csökkentik az „aki”, „ami”, „amely” alakokat. Ezért ezeket a legtöbbször előforduló helyesírási különbségeket egységesítve is elvégeztem a méréseket és a hálózatos megjelenítést. A tanulmány repozitóriumában (<https://github.com/SzemesBotond/literary-network>) közzétett ábrán látható, hogy a változtatások nem befolyásolták érdemben az eredményeket; a stílus időbelisége nem vezethető vissza kizárólag a helyesírás különbségeire. Ennek kapcsán fontos megjegyezni, hogy a korpusz 1000 leggyakoribb szavában így is találhatók régies kifejezések, amelyek hatással vannak a stiláris hasonlóság mértékére: *állott, kérdé, mondá, ama, folytató, lőn, minő*.

²¹ Részletesebben lásd SZEMES Botond, „Kvantitatív stilisztika”.

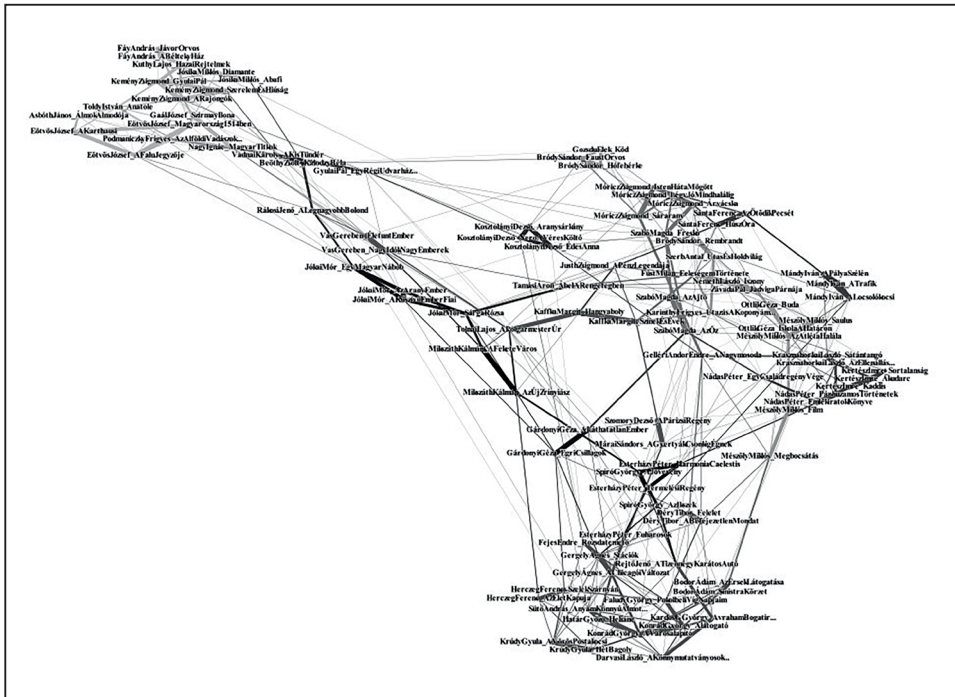
²² Maciej EDER, „Visualization in stylometry: cluster analysis using networks”, *Digital Scholarship in the Humanities* 32, 1. sz. (2017): 50–64, <https://doi.org/10.1093/lhc/fqv061>.

vagy egyenlő hosszúságú részeit hasonlítjuk össze? Milyen matematikai eljárást alkalmazunk? stb.). A különböző beállítások ugyanis eltérő eredményekre vezethetnek. Maciej Eder módszere ezt az esetlegességet azáltal küszöböli ki, hogy több beállítás eredményét is figyelembe veszi a vizualizációk elkészítésekor. Minden beállítást követően rögzíti, hogy egy szöveg melyik másik háromhoz hasonlít a leginkább a mért értékek szerint (ezt a hármat is rangsorolva adja meg). Minél inkább megegyeznek a vizsgált nyelvi elemek eloszlásai két szövegben, annál kisebb közöttük a matematikai értelemben vett távolság, vagyis annál nagyobb a stiláris hasonlóság mértéke. Az összes beállítás után a szövegek közötti viszony az egyes eredmények összegzéseként jön létre: minél többször került közel egy szöveg egy másikhoz az első, második, illetve harmadik helyen, annál magasabb értéket mutat a hasonlóságot jelölő pontszám – amely ugyanakkor képes jelölni a csak eseti, lazább kapcsolatokat is. Ezekből az eredményekből jönnek létre végül a hálózatok, amelyekben a szövegek szerepelnek csomópontként, a köztük lévő élek (valamint azok közelsége és vastagsága) pedig a kiszámított posztszámok alapján létesülnek. Minél vastagabb egy kapcsolat, annál erősebb a stiláris hasonlóság két szöveg között; ugyanezt jelzi a közelség is, ám itt számításba kell venni, hogy a szövegek összességének viszonyrendszere határozza meg két csomópont elhelyezkedését és közelségét, nem csupán hasonlóságuk mértéke. Így megeshet, hogy távolabb/közelebb kerülnek egymáshoz vastagabb/vékonyabb élekkel összekötött csomópontok.

Ezt az eljárást alkalmazva rögzítettem tehát szövegek hasonlóságát (az első három helyen) a 100, 200, 300... 1000 leggyakoribb szó alapján, majd az eredményeket egy hálózatként összegeztem a Gephi szoftver segítségével. A jobb elrendezés érdekében a *ForceAtlas2* algoritmust használtam, amely a szorosan kapcsolódó részeket közelebb, a nem vagy lazán kapcsolódókat pedig távolabb helyezi el az ábrán. a fentebb leírt módon²³ Tulajdonképpen ez az algoritmus eredményezi a történetiség kirajzolódását. Végül az egymással szorosabb összeköttetésben álló csomópontok csoportjait is azonosítottam a hálózatban a

²³ A *ForceAtlas2* egy gráfelrendező algoritmus, amely a csomópontokat úgy mozgatja, mintha fizikai erők hatnának rájuk: a kapcsolódó csomópontok vonzzák, a nem kapcsolódók taszítják egymást. Az elrendezés iteratív módon jön létre, így a hálózat fő formájának kialakulását követően is alakulhat. Némileg – de nem számottevő mértékben – különböző elrendezés jön létre attól függően, hogy mennyi ideig hagyjuk dolgozni az algoritmust. Egy idő után azonban beáll egy egyensúly, amely már csak elenyésző mértékben változik. Az ábra létrehozásakor nem vártam meg ezt az egyensúlyi állapotot, csupán a hálózat formájának kialakulásáig (1-2 másodperc) futtattam az algoritmust.

Modularity függvény segítségével (resolution = 1) – ezt az eredeti ábrán külön színnel jelöltem, ami a fekete-fehér változatban kevésbé érzékelhető.

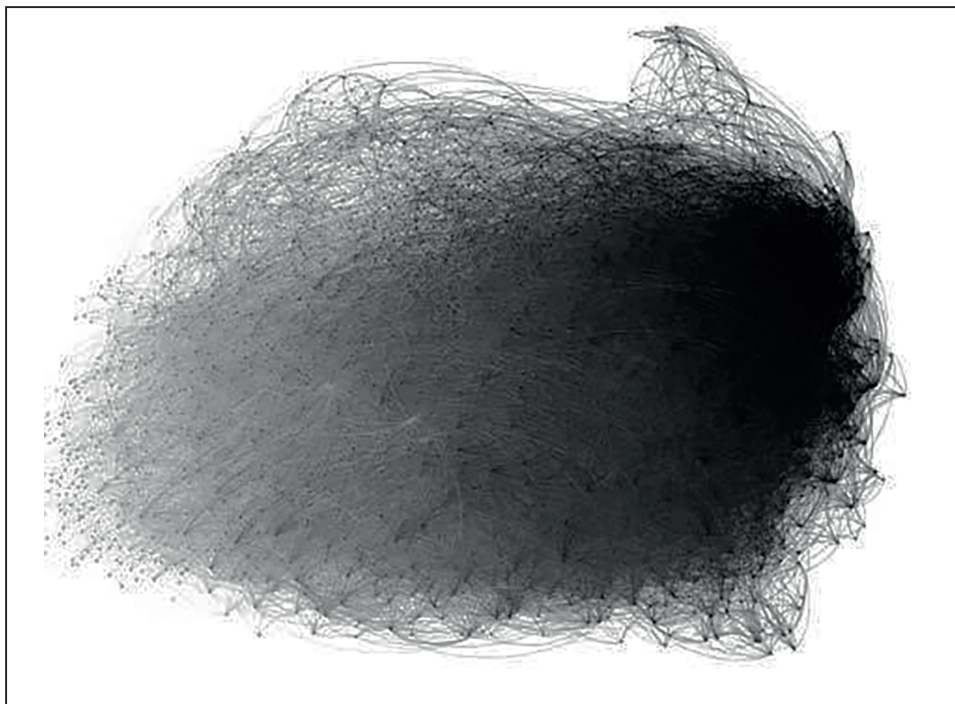


3. ábra. 100 magyar regény stílusár hasonlóságát bemutató hálózat (saját szerkesztés)

Ismétlem: a hálózatban a történetiség a beállításoktól függetlenül jelenik meg; egyszerűen az egy időben keletkezett szövegekben jobban hasonlít a szavak gyakorisági mintázata, mint az időben távol levőkben. Meggyőző bizonyítéka ez annak, hogy a stílus nemcsak egyénileg és szociokulturálisan meghatározott,²⁴ hanem korszakokhoz is kötődik. Ugyanilyen eredményre jutott Matthew Jockers *Macroanalysis* című könyvében, amikor több mint 3300 angol nyelvű regény kapcsolatát ábrázolta hálózatosan (4. ábra). Ekkor az imént leírtához hasonló módszertant alkalmazott, ám nem a figyelembe vett szavak számát változtatta az egyes beállításokban, hanem azok körét: a 42

²⁴ TOLCSVAI NAGY Gábor, „A stílus szociokulturális tényezőinek kognitív nyelvészeti megalapozása”, in *A stílus szociokulturális tényezői*, szerk. TOLCSVAI NAGY Gábor és TÁTRAI Szilárd, 19–49 (Budapest: Eötvös Loránd Tudományegyetem, 2012), különösen: 26–28.

leggyakoribb tokenre (szóra és írásjelre) vonatkozó eloszlásokat tematikus szócsoportok gyakoriságával egészítette ki. A csomópontokat a szövegek megjelenési éve alapján a szürke különböző árnyalataival, az éleket pedig a csomópontoknak megfelelő színnel jelölte:



4. ábra. 3346 angol nyelvű regény hálózatos elrendezése stiláris és tematikus jegyek alapján a *ForceAtlas2* algoritmus használatával – Matthew Jockers, *Macroanalysis*

A csomópontok évszámok szerinti árnyékolása a stilisztikai-tematikai adatok egyértelmű időbeli jellegét mutatja. A grafikon világosabb, nyugati részénél kezdődően haladunk kelet felé az időben. Fontos szem előtt tartanunk, hogy a regények nem a megjelenési dátumuk alapján kerültek csoportosítva a hálózatban; valójában a dátumoknak semmilyen szerepe nincs annak meghatározásában, hogy a könyvek milyen közel vannak egymáshoz, vagy hogyan helyezkednek el a vizualizációban. A könyveket a kiszámított stilisztikai és tematikai távolságuk hasonlósága alapján húzzuk össze (és toljuk szét). Az, hogy kronologikus sorrendben állnak össze, csupán véletlen egybeesés – de egyben rendkívül figyelemre méltó is. Az időrendi elrendezés

megmutatja, hogy az idő múlásával valóban bekövetkeznek tematikai és stilisztikai változások.²⁵

A magyar regényeket bemutató hálózat az időbeliség pusztá regisztrálásán túl részletesebb elemzést is lehetővé tesz. Voltaképp a különböző léptékek egyidejű bemutatásában rejlik jelentősége: a nagyobb tendenciák és a helyi kapcsolatok is leolvashatók róla. Feltűnő például a 19. század közepéről származó szövegek egységessége, amelyet Eötvös József, Kemény Zsigmond vagy Fáy András regényeivel jellemezhetünk. Jól látszik az is, hogy a 20. századi prózapoétikákhoz Vas Gereben és mindenekelőtt Jókai Mór szövegei biztosítanak átjárást. Nem véletlen, hogy ez a két szerző reagál a leggyorsabban és leghatározottabban a szövegek megjelenésének megváltozott körülményeire, a folyóiratokon alapuló médiarendszer kiépülésére. Közelebbről szemügyre véve az is kijelölhető, hogy mely hagyomány felé nyitottak elsősorban ezek a szövegek. A Vas–Jókai lánc a hálózatban leginkább a Mikszáth-féle próza irányába mutat (ide tartozik még Tolnai Lajos 1885-ös regénye is, *A polgármester úr*), némileg elkülönülve a Bródy Sándor és Móricz Zsigmond nevével jellemezhető realista iránytól és az önálló egységet képező Kosztolányi-prózától. Szembetűnő továbbá, hogy Jókai vizsgált négy regénye is időrendbe rendeződik: az 1854-es *Egy magyar nábobtól*, az 1869-es *Kőszívű ember fiai* és az 1872-es *Az arany ember* kettősen át az 1892-es *Sárga rózsáig*. A korábbi stilometriai mérésekkel összhangban²⁶ azt mondhatjuk, hogy az időben kiterjedt Jókai-próza leköveti a korabeli magyar irodalom változásait, illetve e változások egyik előmozdítójaként tűnik fel. Mindeközben fontos azt is megjegyezni, hogy az ábrán látható hasonlóságok csak az elemzett korpuszon belül érvényesek: egyes Vas-, Jókai- és Mikszáth-regények közelsége nem jelenti azt, hogy bármely korpuszra vonatkozva is szükségszerűen a legközelebb kerülnének egymáshoz. Ezért az ebből adódó esetlegességeket is figyelembe kell venni, hogy elkerüljük az ábra túlértelmezését.

A hálózat másik végén a „prózafordulat” szerzői sorakoznak a 20. század végéről: Kertész Imre, Krasznahorkai László és Nádas Péter munkái találhatók itt. Ez alapján Kertész Imre prózapoétikája szorosabban tűnik integrálhatónak a korszak más jelentős alkotásai közé, mint ahogyan a szakirodalmi konszenzus tartja.²⁷ Jobban elkülönülnek e szövegektől Esterházy Péter és

²⁵ Matthew JOCKERS, *Macroanalysis: digital methods and literary history* (Illinois: University of Illinois Press, 2013), 164, <https://doi.org/10.5406/illinois/9780252037528.001.0001>.

²⁶ SZEMES, *Digitális irodalomtudomány*, 117–137.

²⁷ Egy Kelemen Pál és Györffy Miklós által szerkesztett, német nyelvű tanulmánykötet tanulságait (*Kertész und die Seinigen: Lektüren zum Werk von Imre Kertész*, 2009) így összegzik a szerkesztők és Szirák Péter, Kertész monográfiája: „Kertész munkássága úgymond párhu-

Spiró György regényei (a *Termelési-regény*, a *Harmonia caelestis*, a *Fuharosok*, illetve *Az ikszek* és a *Jövevény*), amelyek szorosabb kapcsolatot ápolnak a magyar irodalom mikszáthi hagyományával, és stílusban kevésbé radikális poétikát érvényesítenek – bár például a (maga is kétarcú) *Harmonia* erősen hasonlít az elemzett Nádas-regényekre. Ám a fő átjárást az ábra jobb oldalán csoportosuló szerzőkhöz elsődlegesen Mészöly Miklós művei (különösen a *Film*) biztosítják, megerősítve a szerző irodalomtörténeti helyét a prózafordulat előzményeként. Mészöly összekötő szerepe ugyanakkor ennél is kiterjedtebb. Vizsgált regényei szoros kapcsolatot tartanak az ábra különböző alcsoportjaival, nem csak a prózafordulat szerzőivel. Munkássága egyaránt kapcsolódik, főként *Az atléta halála* és a *Saulus* révén, a Móricz–Bródy irányból továbbfejlődő realista törekvésekhez és a „szabályos” (vagyis a prototipikus mondat sémájára épülő) mondatszerkezeteket és/vagy a realizmus valóságigényét fellazító irányzatokhoz, elsősorban a *Megbocsátás* révén. Az előbbi láncolatba tartozik többek között Szerb Antal *Utas és boldvilág* és Füst Milán *Feleségem története* című regénye, valamint Sánta Ferenc, Németh László, Ottlik Géza és Mátyás Iván prózája, míg a másodikba Krúdy Gyula és Herczeg Ferenc vagy Darvasi László és Bodor Ádám művei sorolhatók.

A hálózat tanúsága szerint a stílusis alcsoportokra való elkülönülés a 20. század első évtizedeiben erősödik fel. Korábban a stílus tekintetében jóval egységesebbnek mutatkozik a magyar irodalom, legalábbis a 100 regény és a leggyakoribb szavak eloszlása felől nézve. A 20. századi elkülönülés ugyanakkor radikális; ezt mutatja a hálózat jelentős „észak–déli” kiterjedése. Ebből az időszakból származnak e tengelyen egymástól legtávolabb kerülő szerzők (Gozsdu–Bródy vs. Krúdy). Ezt a jelenséget a stílusirányzatok térnyerésével szokás leírni az irodalomtörténet-írásban. Az ábra meggyőző bizonyíték lehet amellett, hogy a századfordulót és az azt követő éveket valóban a stílusis „burjánzás” és a kifejezésmódok sokszínűsége jellemzi. Jól elkülöníthetők például a realista törekvések felül, valamint a különféle értelemben felfogott impresszionista szerzők középen (például Kaffka Margit és Szomory Dezső) és alul (például Krúdy).

A nagyfokú elkülönöződés és összeköttetés, valamint az „impresszionizmus” megjelölés ilyen többértelműsége a stílusirányzatok hagyományos meg-

zamosan halad a prózafordulat történetével, ami folyamatos reflexió tárgya a kötetben, hiszen bizonyos értelemben alternatívát kínál a prózanyelv megújítására, és kísérletei nem írhatók le teljes mértékben a prózafordulat kanonizált szövegein kidolgozott értelmezési készséggel.” GYÖRFFY Miklós, KELEMEN Pál, PALKÓ Gábor és SZIRÁK Péter *A „prózafordulat” mint korszakalakzat és elemzési szempont*, OTKA Munkabeszámoló, hozzáférés: 2025.06.10, https://real.mtak.hu/1712/1/47164_ZJ1.pdf.

közelítésének nehézségeire is rávilágít. A korábbi leírások ugyanis egy szigorú kategorizációt kívántak érvényesíteni, amely egymástól jól elhatárolható irányzatokat és ezek alá egyértelműen besorolható irodalmi alkotásokat feltételez. Ez a gondolkodás gyakori ellentmondásokhoz (hiszen az élesen elkülöníteni kívánt kategóriák jellemzői sokszor megegyeztek egymással) és természetlen vitákhoz vezetett szövegek stílusának besorolását illetően. Ehelyett érdemes egy lazább, családi hasonlóságon alapuló, prototípus-elvű²⁸ kategorizációt előnyben részesítenünk. Ezt az elképzelést erősíthetik a stiláris kapcsolatokat megjelenítő hálózatok. Egy ilyen hálózatban is fellelhetők jobban elkülönülő részek, amelyek az elemek szorosabb összetartozását mutatják más elemekhez képest, és amely részeket azonosíthatjuk az egyes irányzatokkal. Ugyanakkor ez az elképzelés lehetővé teszi annak belátását is, hogy az irányzatoknak vannak központi, prototipikus alkotásai, és vannak ezekhez lazábban kapcsolódó, periférikus művek, amelyek nem minden jellemzőjükkel kötődnek az adott irányzathoz. Beláthatóvá válhat továbbá az is, hogy egy szöveg nem feltétlenül sorolható problémátlanul egyik kategória alá sem, és hogy az irányzatokhoz társított sajátosságok ritkán fordulnak elő egyszerre egy alkotásban, ám egyenként akár különböző stílusú művekben is megjelenhetnek.

A több mint 170 év irodalomtörténetét bemutató 3. ábra ezt a hálózatos szemléletet juttatja érvényre, ám kiterjedtsége miatt inkább csak felveti egyetlen időszak szorosabb vizsgálatának lehetőségét. Ebből kiindulva, végezetül, egy olyan hálózatot hoztam létre (5. ábra), amely nem az irodalomtörténet *folyamatát* ragadja meg, hanem a nagy hálózat egy részére ráközelítve egy irodalomtörténeti *korszak* jobb megértéséhez vezethet el. Célom a 20. század elején végbement stiláris újítások, vagyis a stílusirányzatok kérdésének újragondolása a hálózatos elrendezésen keresztül.

Az ábra létrehozásához ismét a fenti módszertant alkalmaztam, ugyanakkor ez esetben nem egyszerűen a leggyakoribb szavak eloszlását számítottam ki. Ennek oka, hogy az irányzatok sajátosságait hagyományosan szófajok alapján határozták meg;²⁹ így én is inkább szófajok eloszlásából indultam ki. Ennek érdekében a szövegek szavaihoz szófaji címkéket rendeltem az e-magyar

²⁸ Vö. TOLCSVAI NAGY Gábor, „A nyelvi kategorizáció kognitív nyelvészeti keretben”, *Az Eszterházy Károly Főiskola tudományos közleményei: Tanulmányok a magyar nyelvről* 32 (2005): 5–20.

²⁹ Lásd például: HERCZEG Gyula, *A modern magyar próza stílusformái* (Budapest: Tanácskönyvkiadó, 1979), SZABÓ Zoltán, *A magyar szépirodalmi stílus történetének fő irányai* (Budapest: Corvina Kiadó, 1998).

nyelvi elemző segítségével.³⁰ Az eredeti szöveg szavait az alapszófajok esetében ezekre a címkékre cseréltem (főnév – NOUN, melléknév – ADJ, ige – VERB, határozószó – ADV, tulajdonnév – PROPN).³¹ Azért csak az alapszófajokat vettem figyelembe, mert ezek gyakorisága kerül az előtérbe a korábbi meghatározásokban is, valamint így megmaradnak a műfajok, korszakok és szerzők stiláris „ujjlenyomatát” leginkább hordozó funkciószavak is a szövegekben.³² Az előfeldolgozás eredményét (és esetlegességét a tulajdonnevek azonosításában) az alábbi példa – Kosztolányi *Édes Annájának* nyitánya – szemlélteti:

[1] Kun Béla repülőgépen menekült az országból. / Délután – úgy öt óra felé – a Hungária-szállóban székelő szovjetház körül fölrebbent egy repülőgép, átrepült a Dunán, a Várhegyen, s merész kanyarodással a Vérmező felé tartott.

[2] PROPN PROPN NOUN VERB az NOUN . / ADV - úgy NOUN felé - a PROPN ADJ NOUN körül VERB egy NOUN , VERB a PROPN , a NOUN , s ADJ NOUN a NOUN felé VERB .

A korábbi munkák tanúsága szerint nemcsak szófajok pusztá gyakorisága, hanem az összekapcsolódásukból létrejött grammatikai szerkezetek (például jelzős főnevek) vagy retorikai alakzatok (például halmozás) képesek leginkább leírni az irányzatok sajátosságait. Ezért az átalakított prózaszövegekben három egymást követő egységek eloszlását vizsgáltam, és nem egyszerűen tokenek gyakoriságát. Ekkor az írásjeleket nem vettem figyelembe, hiszen például névszói halmozás esetében az elemek egymásra következését vesszők tagolják, így nem lennének azonosíthatók csupán három token segítségével. A [2]-es mondatban olyan hármas egységek találhatók, mint például „úgy

³⁰ VÁRADI Tamás, SIMON Eszter, SASS Bálint, MITTELHOLCZ Iván, NOVÁK Attila és INDIG Balázs, „E-magyar – A Digital Language Processing System”, in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (Miyazaki: ELRA, 2018), 1307–1312.

³¹ Ekkor szófajok szomszédsági viszonyából indulok ki, habár mindezt kitágíthatónak gondolom a mondattani felépítés vizsgálatára is. Ugyanis, ahogy Horváth Péter „Szintaktikai ismétlődés a magyar kanonikus költészetben” című tanulmányban is írja jelen számban, „a szófajok közötti szomszédsági viszonyok ismétlődése sok esetben mondattani értelemben vett szintaktikai viszonyok, azaz függőségi viszonyok ismétlődésére utal”.

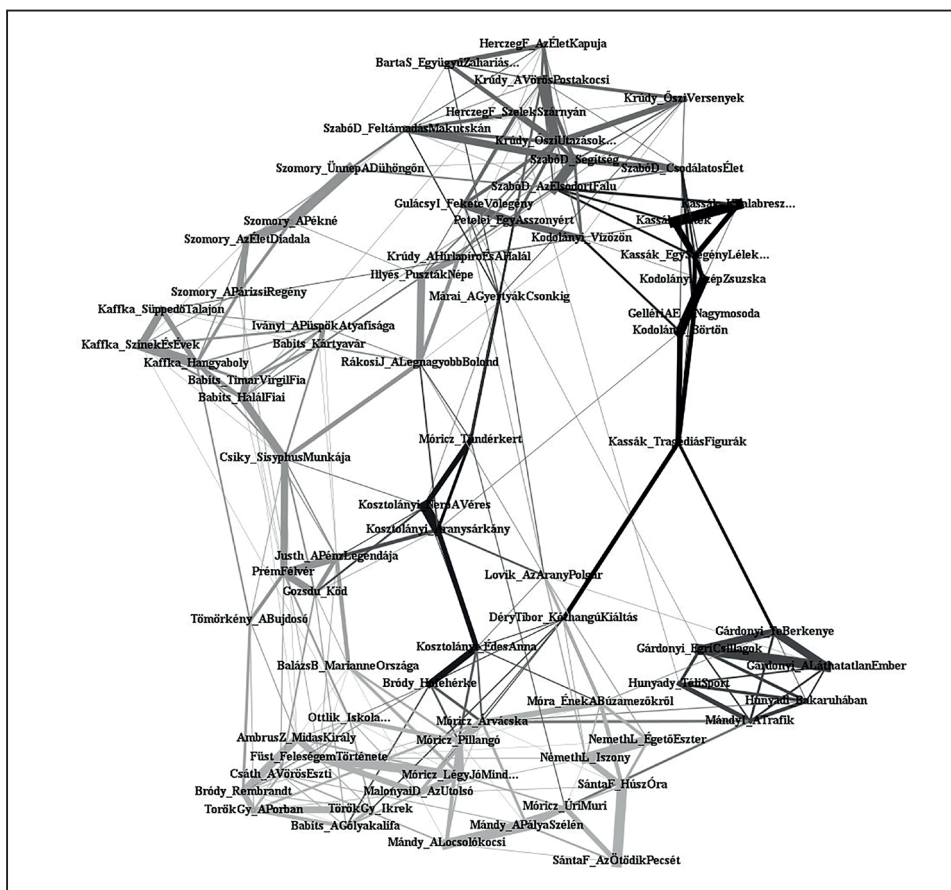
³² Ehhez lásd Joanna BYSZUK és SZEMES Botond, „A krakkói Computational Stylistics Group bemutatkozása: Előszó a Digitális Bölcsészet folyóirat tematikus lapszámához”, *Digitális Bölcsészet* 5 (2021): 3–19, 4, 7, <https://doi.org/10.31400/dh-hun.2021.5.3147>.

NOUN felé” , „VERB a PROPN” vagy „egy NOUN VERB” – ez utóbbi a főnévi és igei címke között lévő vessző elhagyásával jön létre. Ezt követően a *Bootsrap Consensus Network* módszerét alkalmazva rögzítettem szövegek hasonlóságát (az első három helyen) a 20, 40, 60, 80 és 100 leggyakoribb hármas szerkezet alapján, majd az eredményeket egy hálózatként összegeztem a Gephi szoftver és a *ForceAtlas2* algoritmus segítségével.

Annak érdekében, hogy részletesebb képet kapjunk a korszakról, kitágítottam a 100 regényt tartalmazó lista vonatkozó szövegeinek körét. Összesen 74 prózai szöveget gyűjtöttem egybe, elsősorban regényeket, de legalább 2000 szó hosszúságú elbeszéléseket (hogy a gyakorisági mérések statisztikailag hiteles eredményre vezessenek). A gyűjtemény magját olyan szövegek alkotják a 20. század első évtizedeiből, amelyeket gyakran sorolnak³³ valamely stílusirányzat alá (Kaffka Margit, Kassák Lajos, Krúdy Gyula, Szabó Dezső, Szomory Dezső fejenként 3–4 szövege). Ezeket egészítettem ki olyan művekkel, amelyek csak helyenként merülnek fel egy-egy irányzat képviselőiként (Babits Mihály négy szövege, Balázs Béla *Marianne országa*, Barta Sándor *Együgyű Zahariás*, a *Megváltó*, Bródy Sándor *Rembrandt*, Déry Tibor *A kétbangú kiáltás*, Gárdonyi Géza *Té Berkenye*, Gelléri Andor Endre *A nagymosoda*, Gulácsy Irén *Fekete vőlegények*, Herczeg Ferenc *Az élet kapuja* és *Szelek szárnyán*, Hunyady Sándor *Bakarubában* és *Téli sport*, Kodolányi János három szövege, Márai Sándor *A gyertyák csonkig égnek*, Malonyai Dezső *Az utolsó*, Tömörkény István *A bujdosó*), illetve amelyek az irányzatok előzményeiként említhetők (Ambrus Zoltán *Midas király*, Bródy Sándor *Hófehérke*, Csáth Géza *A vörös Eszti*, Csiky Gergely *Sisyphus munkája*, Gozsdu Elek *Köd*, Iványi Ödön *A püspök atyafisága*, Justh Zsigmond *A pénz legendája*, Petelei István *Egy asszonyért*, Prém József *Félvér*, Rákosi Jenő *A legnagyobb bolond*). Szintén felvettem a szövegek közé a korszak olyan alkotásait, amelyeket korábban nem hoztak összefüggésbe semelyik irányzattal, vagy amelyeket kifejezetten a realista próza paradigmájába soroltak (Füst Milán *Feleségem története*, Gárdonyi Géza *A láthatatlan ember* és *Egri csillagok*, Illyés Gyula *Puszták népe*, Kosztolányi Dezső három regénye, Móra Ferenc *Ének a búzamezőkről*, Móricz Zsigmond öt szövege, Török Gyula *A porban* és *Ikrek*). Végezetül, hogy az időbeliség mérésekre gyakorolt hatása is jobban vizsgálható legyen (azaz mennyire befolyásolja a mért stílárís hasonlóságot az időbeli közelség), felvettem a második világháborút követően megjelent regényeket is a korpuszba, amelyeket leginkább a visszatérő vagy megerősödő realista paradigma része-

³³ A korpusz összeállításakor elsősorban a szócikkhez megadott bibliográfiában szereplő stílustörténeteket vettem alapul.

ként tárgyalnak az említett stílustörténetek.³⁴ (Némileg önkényesen, a száz magyar regényt tartalmazó gyűjteményből szemezgettem; az így összeállított lista tartozik Mátyás Iván három szövege, Németh László *Égető Eszter* és *Iszony*, Ottlik Géza *Iskola a határon*, Sánta Ferenc *Az ötödik pecsét* és *Húsz óra*).



5. ábra. Stílusirányzatok a modernségben

³⁴ Elsősorban HERCZEG Gyula, *A modern magyar próza stílusformái* (Budapest: Tankönyvkiadó, 1979).

Az ábra jól mutatja a szövegek közötti sokrétű kapcsolatot és az abból létrejövő komplex viszonyrendszert, valamint alcsoportok – nem éles – elkülöníthetőségét. Az alsó csoport mintegy stilisztikai kiindulópontként a realista próza címkével jellemezhető. Ide tartozik többek között Móricz négy szövege, Török Gyula regényei, Móra Ferenc *Ének a búzamezőkről*, Babits *A gólyakalifa*, Füst *Feleségem története*; valamint kivétel nélkül a második világháború után született írások. Ez azt is jelenti, hogy a hálózatban valóban az irányzatokba sorolható írásmódok jelentik a különbséget, és az időbeliség nem befolyásolja számottevően a csomópontok elhelyezkedését. A realizmus egy alcsoportját képezik továbbá Gárdonyi és Hunyady szövegei és Mándy *Trafikja*, amelyekben feltehetőleg az egyenes beszéd nagy száma eredményez stiláris különbséget – bár ez a hipotézis egyelőre nem igazolt. Minél távolabb találhatók szövegek a realista csoportoktól az ábrán, annál inkább sajátos, a korabeli irodalmi nyelvhez képest normasértó vagy újító prózapoétika jellemző rájuk. Az ábra jobb oldalán egy kisszámú csoportot olyan szövegek alkotják, amelyeket az expresszionista próza kategóriájába szokás sorolni: Kassák és Kodolányi munkái találhatók itt, valamint Gellértől *A nagymosoda*, megerősítve az elgondolást, miszerint prózájában expresszionista jegyek fedezhetők fel.³⁵ Érdemes kiemelni Déry Tibor *A kéthangú kiáltás* című, szintén gyakran expresszionistának címkézett művét, amely a szófaji hármassok eloszlása alapján inkább a realista csoporthoz áll közel, ám kapcsolatot biztosít az expresszionista szerzők felé. Az ábra bal oldalán elkülöníthető csoportot impresszionista jelzővel illethetjük, ide tartoznak Kaffka Margit és Szomory Dezső művei, valamint Babits Mihály három regénye is (míg *A gólyakalifa* inkább realista jegyeket mutat). E poétikák felé a századforduló szerzői mint előzmények vezetnek át: közéjük sorolható Csiky Gergely, Gozdsu Elek, Iványi Ödön, Justh Zsigmond, Prém József és Rákosi Jenő. Szintén hozzájuk kapcsolódik Tömörkény *A bujdosó* című novellájával, akit Herczeg Gyula a kollektív nézőpontot érvényesítő, élőbeszédszerű-népies realizmushoz sorol, ám gyakran tartják impresszionista szerzőnek is³⁶ – ez utóbbi kapcsán Krúdyval való rokonságát hangsúlyozva,³⁷ amely hasonló-

³⁵ BORI Imre monográfiájában *Az expresszionizmus vonzásában* című fejezet tárgyalja munkásságát: *A szecessziótól a dadáig*, Symposium könyvek 20 (Újvidék: Forum Kiadó, 1969), 219–220. Bori szerint a magyar expresszionizmus fő képviselője Kassák mellett Barta Sándor és Szabó Dezső, akik az ábrán viszont a prototipikus expresszionizmus és impresszionizmus közötti kategóriához tartoznak.

³⁶ Például BARÁNSZKY Jób László, „Tömörkény táji impresszionizmusa”, *Irodalomtörténet* 57, 4. sz. (1969): 703–73.

³⁷ LENGYEL András, *Tömörkény-tanulmányok* (Szeged: Tiszatáj könyvek, 2019), 173, 99.

ság azonban „nem túl meggyőző”³⁸ a korszak más kutatói számára. Az ábra alapján a Krúdyval való stílári párhuzam kevésbé tűnik megalapozottnak; ahogyan egyértelműen impresszionistának sem mutatkozik a vizsgált novella, inkább a „protoimpresszionista” művekkel hozható kapcsolatba.

Tanulságos felidézni, hogy a száz regényt ábrázoló hálózatban (3. ábra) Kosztolányi prózája egy viszonylag független „szigetet” alkotott a magyar próza két nagy fejlődési iránya között. Így van ez most is, amennyiben regényei saját, köztes kategóriát hoznak létre (Bródy Sándor *Hófehérke*, Móricz Zsigmond *Tündérvár* és Márai *A gyertyák csonkig égnek* című szövegeivel kiegészülve, ám közöttük is jelentős a távolság, így nem igazán beszélhetünk a csoport egységességéről). Kosztolányi prózáját más csoportoktól leginkább a stílári visszafogottság különíti el: az impresszionizmustól a mondatrészek halmozásának hiánya, a realizmustól (például a hozzá közel álló Móricztól)³⁹ a szabad függő beszéd kis száma és rövidebb, jól szerkesztett mondatok alkalmazása. Vagyis míg más szerzők az érzéki benyomások nagy számával (impresszionizmus) vagy a szereplői szövegek beépítésével (realizmus) bővítik a mondatstruktúrákat, addig Kosztolányi tartózkodik az ilyen megoldásoktól. A hálózat ugyanakkor nemcsak a különbségeket, hanem a hasonlóságokat is megmutatja, Kosztolányinak például Móricz mellett a századelő prózájával való rokonságát (Gozsdu, Justh, Csiky, Rákosi).

A hálózatos gondolkodás szempontjából az ábra legtanulságosabb csoportja a prototipikus impresszionizmus és expresszionizmus között elhelyezkedő szövegegyüttes Szabó Dezső, Krúdy Gyula és Herczeg Ferenc szövegeivel (ide tartozik még Gulácsy Irén és Petelei István egy-egy regénye is). Ez egy olyan átmeneti kategória, amely sok tekintetben hasonlít a másik két irányzatra, azoktól élesen nem elkülöníthető, mégsem tartoznak minden szempontból hozzájuk. A hagyományos stilisztika kategorizációja nem képes ilyen árnyalt megkülönböztetésekre (Krúdyt inkább az impresszionizmus, Szabó Dezsőt pedig az expresszionizmus⁴⁰ címkéi alá szokás sorolni). Külön érdekes Herczeg Ferenc szereplése itt, akit csak elvétve neveznek impresszionistának,⁴¹ inkább a nemzeti-konzervatív irodalom képviselőjének szokás tekinteni. A mérések azt mutatják, hogy ilyen ideológiai különbségek ellenére, stílárisan Herczeg prózája is része az esztétizáló modernség újító irodalmának.

³⁸ RADNAI Dániel Szabolcs, „Egy életmű számbavétele”, *Jelenkor* 63, 6. sz. (2020): 699.

³⁹ HERCZEG Gyula, *Móricz Zsigmond stílusa* (Budapest: Tankönyvkiadó, 1982).

⁴⁰ Szabó Dezső esetében már Herczeg Gyula is jelzi, hogy nem tekinthető valójában expresszionistának, és bizonyos szempontokból inkább Krúdy impresszionizmusához állhat közel. HERCZEG, *Móricz Zsigmond stílusa*, 139–140.

⁴¹ VÖ. KEMÉNY Gábor, „A Nyugat jelentősége a modern magyar szépirodalmi stílus kiteljesedésében”, *Magyar Nyelvőr* 113, 3. sz. (2008): 279–302, 288.

Röviden érdemes még kitérni arra is, hogy a stilometriai módszer a szövegek és csoportok jellemző nyelvi szerkezeteit is azonosítani képes. Ekkor megint csak egy prototípus-elvű kategorizáció jut érvényre: az egyes jellemzők önmagukban nem különítik el a kategóriákat egymástól, csupán azok együttes figyelembevételével, és ekkor sem éles határokkal rendelkező kategóriákat eredményezve. Vagyis a sajátosságok nem csak egy csoport szövegeire jellemzők, ahogyan egy csoportba tartozó szövegekben sem egyenlő mértékben azonosíthatók. Jó példa erre a melléknév, illetve jelzőhalmazt indikáló „ADJ ADJ ADJ” szófaji hármas, amely minden, a realista hagyománytól eltérő kategóriában gyakran előfordul. A módszer egyaránt azonosítani tudja a vesszővel ellátott felsorolásokat (az írásjelek figyelmen kívül hagyása miatt) és a vessző nélküli melléknévhalmozásokat, beleértve azokat az eseteket is, ahol az egyik melléknév egy másik bővítménye. Példaként álljon itt egy-egy mondat az expresszionistának, az impresszionistának és a köztes stílusúnak címkézett csoportok központi képviselőitől: „a nagy, fehérre meszelt négyszögletű [ADJ, ADJ ADJ ADJ – a „fehérre” a „meszelt” melléknévi bővítménye] templomban szürkén és didergetőn borong a február délelőtt, csupán a szószerk vörös bársonya s az úrasztal terítője izzik meleg-pirosan és tompa-aransárgán” (Kodolányi: *Szép Zsuzska*); „A szürke falalánk mögött messzi, mozdulatlan, mélylő kék nyári [ADJ, ADJ, ADJ ADJ ADJ – a „mélylő” a „kék” bővítménye] ég és nagy, telitarka, álomszínes [ADJ, ADJ, ADJ] világ odakinn a napon!” (Kafka Margit: *Süppedő talajon*); „A másik nagy szakállú, tántorgó, ezerkarú” [ADJ ADJ ADJ, ADJ, ADJ – a „másik” melléknévi értelmezése vitatható, az e-magyar mégis ezt a címkét rendelte hozzá, továbbá alanya is a mondatnak, így nem része az állítmányi melléknévhalmozásnak; a példa azonban enélkül is szemléletes]. (Szabó Dezső: *Az elsodort falu*).

A mérések alapján a nominális stílus önmagában nem megkülönböztető jelle a prototipikus impresszionista szövegeknek. Sajátosságuk inkább az egyes névszói szerkezetek – például (jelzővel ellátott) birtokos szerkezetek – nagy számában jelölhető ki, valamint abban, hogy inkább szerepeltetnek kategória-megnevezéssel, azaz köznévvél elérhetővé tett szereplőket, amelyek a típusjelentésen keresztül utalnak egy egyedi példányra (például: a fiú), szemben többek között Krúdy prózájával, ahol jelentősen gyakoribb az egyedi példányokat megnevező tulajdonnevek használata.⁴² Az impresszionista csoportra jellemző továbbá a megszólaló nézőpontját hangsúlyozó, az élőbeszédhez közelítő, gyakran parti-

⁴² Ez a megfigyelés fontos kiegészítője lehet Kemény Gáborénak, aki Krúdy prózájában a „dezanropomorfizáló” színekdoché gyakoriságára hívja fel a figyelmet, vagyis amikor egy – főként női – szereplőt egy testrész vagy ruhadarab helyettesít a leírásban. KEMÉNY Gábor, *Krúdy*

kulákra (az e-magyar elemzésében határozószavakra) épülő szerkezetek alkalmazása – főként párbeszédekben, de a narrációban is, például: „már rég nem [ADV ADV ADV]”, „most már igazán [ADV ADV ADV]”, „tán nem is [ADV ADV ADV]”, „ő is csak ilyen [ő ADV ADV ADJ]” stb. –, illetve a segédigés szerkezetek alacsony száma. Ez utóbbi inkább az expresszionista szövegek esetében gyakori. Vagyis ekkor sem egyszerűen az igék nagy számáról beszélhetünk, hiszen egyrészt a főnévi és melléknévi mondatrészek halmozása is jellemző ezekre a művekre, másrészt inkább az igei állítmányok módosítása, pontosítása jelölhető ki gyakori elemként. Ide tartoznak szerkezetek a feltételes módot kifejező *volna* segédigével („mintha valami mély, fekete veremre csukta volna rá”), főnévi igenévvel („énekelni kezdett”) vagy módhatározóval („őszintén gyötrődve elsírja [magát]”, „vérmes-dühösen hirtelen felrántotta”), valamint passzív szerkezetek is („[arca] fájdalmasan el volt torzulva”). Az ilyen típusú szerkezetek nagyrészt hiányoznak a köztes kategória szövegeiből, amelyekben azonban a tipikus impresszionistáknál gyakoribb a főnévi és igei elemet is tartalmazó szófaji hármasok (amelyek tipikusan egy szereplőt és cselekvést egymáshoz közel, más bővítmények nélkül fejeznek ki). E csoportra a tulajdonnevek már említett szerepeltetésén túl jellemző még különböző alanyú és állítmányú tagmondatoknak az és kötőszóval való összekapcsolása („rettenetes mélységek ásítottak és összezúzott boltívek hosszú sora kapaszkodik...”) – szemben egyetlen figura részletes kidolgozásával.

A hálózatos gondolkodásra és a prototípus elvű kategorizációra azonban még inkább igaz, hogy a lényeg – általánosító megállapításokon túl – a részletekben rejlik. Ezek felderítését az ábrák aprólékos tanulmányozása mellett a mérési eredmények vizsgálata segítheti, amelyek elérhetők a tanulmány repozitóriumban.⁴³ Persze ezek sem önmagukban érdekesek, hanem csak a szövegek olvasásának és értelmezésének függvényeként; amelyekhez viszont értékes szempontokat kínálhatnak. Ekkor válhat a stilometria a stíluskutatás, a hálózatok pedig az irodalomtörténet-írás – egyszerre módszertani és elméleti – megújulásának ösztönzőjévé.

Lektorálta: Horváth Péter

képalkotása, Nyelvtudományi értekezések 86 (Budapest: Akadémiai Kiadó, 1974), 26–28. Úgy tűnik emellett a tulajdonnevek (szereplők, helyszínek) gyakori felidézése is hozzátartozik e próza stílusához és hangulati hatásához.

⁴³ <https://github.com/SzemesBotond/literary-network>.

PÉTER RÓBERT

Rejtett hatások feltárása az eszmetörténetben a *Reception Reader* kutatási eszközzel¹

ORCID: 0000-0002-7972-4751

robert.peter@ieas-szeged.hu

HELIKON

Uncovering Hidden Influences: The *Reception Reader* as a Tool for Intellectual Historians

Abstract

This article critically examines the *Reception Reader*, a digital tool developed by the Helsinki Computational History Group to explore text reuse in early modern print culture. Positioned at the intersection of intellectual history and digital humanities, the tool leverages large-scale corpora such as EEBO-TCP and ECCO to trace borrowing, reprinting, and repurposing of texts between 1475 and 1800. Using a modified BLAST algorithm to address noisy OCR data, the *Reception Reader* enables scholars to combine distant and close reading in the study of reception, influence, and intertextuality. Three case studies – on Bernard Mandeville, David Hume, and eighteenth-century utopian literature – demonstrate its capacity to reassess authorship attribution, map the circulation of ideas, and reveal hidden intellectual influences. While the tool's visualisations and databases present methodological limitations, its integration of computational analysis with traditional scholarship underscores its value for intellectual historians. Looking ahead, the COMHIS group plans to integrate large language models to detect semantic similarity and to develop an edition comparison tool that will help scholars chart changes across multiple editions and works of the same author. By democratising access to vast digitised archives and advancing text reuse detection, the *Reception Reader* exemplifies how interdisciplinary digital methods can transform the study of intellectual history.

Keywords: text-reuse, evolution of ideas, intellectual history, Reception Reader, early modernity

¹ Ez a cikk az alábbi, szabadon felhasználható mű alapján készült, de több tekintetben eltér az eredeti tanulmánytól: PÉTER Róbert, „Uncovering Hidden Influences: The Reception Reader as a Tool for Intellectual Historians (Reception Reader, Helsinki Computational History Group, <https://receptionreader.com/>)”, *Global Intellectual History*, 1–13, hozzáférés: 2025.11.02, <https://doi.org/10.1080/23801883.2025.2474486>. Eredetileg megjelent az Informa UK Limited (Taylor & Francis Group) kiadásában.

1. HÍD AZ ESZMETÖRTÉNET ÉS A DIGITÁLIS BÖLCSESZET KÖZÖTT

A nyílt hozzáférésű, illetve előfizetéshez kötött digitális archívumokban elérhető, hatalmas és egyre bővülő adatmennyiség (*big data*) olyan kereshető korpuszokat biztosít a kutatók számára, amelyek gyakran ritka és sérülékeny könyveket, kéziratokat, újságokat, leveleket és efemerákat is tartalmaznak. A digitális bölcsészet megjelenésének köszönhetően az irodalomtudományban az elmúlt húsz év során viszonylag elterjedtté vált az adatalapú módszerek alkalmazása, amelyek új felismerések, mintázatok és összefüggések feltárásában segítik a kutatókat. Ugyanakkor, ahogyan Mark J. Hill 2016-ban megjegyezte a *Global Intellectual History* folyóiratban, „a digitális bölcsészetről bár sok szó esik, mégis alig feltárt terület az eszmetörténet szempontjából”. Tanulmányának záró gondolataiban arra biztatta a digitális bölcsészeket és az eszmetörténészeket, hogy működjenek együtt, és hozzanak létre olyan új eszközöket és módszereket, mint például az e tanulmányban bemutatott *Reception Reader*;² amelyek gazdagítják és elmélyítik kutatások minőségét, és a gondolatok terjedésének nyomon követését nagy idő- és térbeli távolságokban is lehetővé teszik.³ Hill javaslata óta jelentős digitális bölcsészeti projektek indultak el, fejlődtek tovább és értek el komoly eredményeket a történettudomány területén, ezek közül néhányat röviden ismertetek az alábbiakban.

A 2007-ben indított és azóta bővített *French Book Trade in Enlightenment Europe* (FBTEE) projekt a régi rend (*ancien régime*) korszakának könyvkereskedelmét térképezi fel, bibliometrikus módszereket alkalmazva az eszmék terjedésének elemzésére.⁴ A projekt korábban ismeretlen összefüggésekre világított rá: például olyan vallásos szövegek – mint a *L'Ange conducteur* – széles körű elterjedésére, amelyek megkérdőjelezik a felvilágosodás szekularizációjáról alkotott narratívákat.⁵ A *Mapping the Republic of Letters* projekt vizualizációs technikákat alkalmaz a felvilágosodáskori gondolkodók levelezési hálózatainak azonosítására.⁶ Többek között az *Electronic Enlightenment* adatbázis

² ROSSON, „Reception Reader”, *Journal of Open Humanities Data* 9, 1. sz. (2023), <https://doi.org/10.5334/johd.101>, <https://receptionreader.com/>.

³ MARK J. HILL és SIMON HENGCHEN, „Invisible Interpretations”, *Digital Scholarship in the Humanities* 34, 4. sz. (2019): 130–50.

⁴ *The French Book Trade in Enlightenment Europe*, hozzáférés: 2025.02.06, <http://fbtee.uws.edu.au/main/>.

⁵ SIMON BURROWS, „The FBTEE Revolution,” in *Digitizing Enlightenment: Digital Humanities and the Transformation of Eighteenth-Century Studies*, szerk. SIMON BURROWS és GLENN ROSE, 94 (Liverpool: Liverpool University Press, 2020).

⁶ *Mapping the Republic of Letters*, hozzáférés: 2025.02.06, <http://republicofletters.stanford.edu/>.

metaadatait felhasználva elemzi a levelezéseket, kapcsolatokat teremtve olyan kulcsfigurák között, mint Locke, D’Alembert és Franklin, hogy ezek révén szélesebb intellektuális közösségeket tárjon fel. Tér- és időbeli elemző eszközök új generációját alkalmazva a kutatás jelentős felismerésekre jutott – például Voltaire korlátozott brit kapcsolatait illetően –, és megkérdőjelezett általános, a felvilágosodás kapcsolatrendszereire vonatkozó feltételezéseket.⁷ A *Reassembling the Republic of Letters (1500–1800)* projekt szintén digitális ke-retrendszerben tette kutathatóvá a kora újkori levelezési hálózatokat.⁸ Ezek a kezdeményezések összekapcsolt nyílt adattechnológiák (*linked open data*) segítségével interoperábilis metaadat-forrásokat hoztak létre, ezáltal lehetővé téve összekapcsolásukat és a transznacionális intellektuális kapcsolatok átfogó vizsgálatát.

A *Measuring Enlightenment: Disseminating Ideas, Authors, and Texts in Europe* (MEDIATE) projekt magánkönyvtárak aukciós katalógusait digitalizálja annak érdekében, hogy a könyvek és – a felvilágosodás fogalmához hagyományosan társított – eszmék áramlását kövesse nyomon 1665 és 1830 között bibliometrikus módszerekkel és az adatmodellezés eszköztárát mozgósítva.⁹ A *Comédie-Française Registers Project* (CFRP) a Comédie-Française színházi társaság 1680 és 1793 közötti előadásainak és pénzügyi nyilvántartásainak digitalizálását végzi.¹⁰ Az adatok vizualizációját és a metaadatok modellezését ötvözve a projekt a jegyeladási adatokat és a repertoár alakulását vizsgálja.¹¹ A tárgyaltak közül a legújabb, *Modelling Enlightenment: reassembling inter-textual networks through data-driven research* (ModERN) projekt célja pedig egy átfogó francia szövegkorpusz létrehozása, valamint új kutatási módszerek

⁷ Dan EDELSTEIN, „Mapping the Republic of Letters”, in BURROWS és ROSE, szerk., *Digitizing Enlightenment: Digital Humanities and the Transformation of Eighteenth-Century Studies*, 73–88.

⁸ TUOMINEN, „Reassembling the Republic of Letters”, in *Proceedings of the Digital Humanities in the Nordic Countries 3rd Conference (DHN 2018)*, szerk. Eetu MÄKELÄ, Mikko TOLO-NEN, Jouni TUOMINEN (CEUR Workshop Proceedings, 2018), 76–88, <https://doi.org/10.5617/dhnpub.11011>.

⁹ Measuring Enlightenment: Disseminating Ideas, Authors, and Texts in Europe (MEDIATE), hozzáférés: 2025.02.06, <https://mediate18.nl/>; Alicia C. MONTROYA, „Enlightenment? What Enlightenment? Reflections on Half a Million Books,” *Eighteenth-Century Studies* 54, 4. sz. (2021): 909–934, <https://doi.org/10.1353/ecs.2021.0097>.

¹⁰ Jeffrey S. RAVEL, „The Comédie-Française Registers Project”, in *Digitizing Enlightenment: Digital Humanities and the Transformation of Eighteenth-Century Studies*: 133–150.

¹¹ Comédie-Française Registers Project, hozzáférés: 2025.02.06, <https://cfregisters.org/>.

kidolgozása, hogy mélyebb betekintést nyerhessünk a kora újkori francia nyomtatott kultúrába és annak intertextuális dinamikájába.¹²

Az említett projektekben alkalmazott adatalapú eszközök és algoritmusok jelentősen előmozdították az eszmetörténeti kutatásokat, elősegítve például a fogalmak megjelenésének és átalakulásának elemzését, a hosszú távú ideológiai változások, továbbá az intellektuális hálózatok rekonstrukcióját. Az eszközök ugyanakkor nemcsak a fogalmi ideológiai változások időbeli modellezésében segítik a kutatókat, hanem abban is, hogy megadják a fogalmak és gondolatok jellemző kontextusait (például műfajok szerint), ezzel megkönnyítve a hagyományos, szoros olvasásra épülő szövegelemzést is. A mesterséges intelligencia legújabb fejlesztései ugyancsak gazdagították az eszmetörténeti kutatásokat. Például az MI-alapú eszközök segítik a többnyelvű korpuszok gyorsabb és hatékonyabb elemzését, ezáltal a kutatók az eredeti nyelveken tanulmányozhatják a szövegeket, miközben azonosíthatják a kulturális hatásokat, és transznacionális összehasonlításokat végezhetnek.

2. INTERDISZCIPLINÁRIS EGYÜTTMŰKÖDÉS: A *RECEPTION READER* MÖGÖTT ÁLLÓ CSAPAT

Az elmúlt évtizedben a Helsinki Computational History Group (COMHIS), amelynek Hill is tagja volt, jelentős mértékben hozzájárult a bölcsészettudományi kutatások digitális átalakulásához, különösen az eszmetörténet területén.¹³ A Mikko Tolonen által vezetett kutatócsoport az eszmetörténetet a modern adattudomány, a gépi tanulás módszereit pedig a történettudomány és a nyelvészet szemléletével ötvözi. Számukra a „»számítógépes történelem« egy kevert módszertant jelenti, amelyben a nagy adatok elemzését az eszmetörténet és könyvtörténet szakértői tudásával kombinálják”.¹⁴ A COMHIS csoport számos jelentős nemzeti és nemzetközi kutatási projektben vett részt ezen a területen, amelyek alapvető és gyakran hivatkozott publikációkhoz vezettek. A jelen tanulmányban tárgyalt *Reception Reader* – egy hiánypótló, webes alkal-

¹² Dario Maria NICOLSI és Glenn ROE, „Modelling Enlightenment: Reassembling Intertextual Networks through Data-Driven Research (ModERN)”, *Digital Enlightenment Studies* 2, 1. sz. (2024): 62–84, <https://doi.org/10.0.238.219/des.22>.

¹³ Helsinki Computational History Group, hozzáférés: 2025.02.06, <https://researchportal.helsinki.fi/en/projects/helsinki-computational-history-group/publications>.

¹⁴ Helsinki Computational History Group (COMHIS), hozzáférés: 2025.02.06, <https://www.helsinki.fi/en/researchgroups/computational-history>.

mazás – szintén e csapat munkája, és 2023-ban vált elérhetővé a nagyközönség számára.

A *Reception Reader* olyan tudományos eszköz, amely bölcsészettudományi kutatóknak nyújt segítséget szövegek időbeli újrafelhasználásának (*text reuse*) vizsgálatában. Az *Early English Books Online-Text Creation Partnership* (EEBO-TCP)¹⁵ és az *Eighteenth Century Collections Online* (ECCO)¹⁶ adatbázisok – amelyek együttesen jelentős mennyiségű, 1475 és 1800 között megjelent szöveget tartalmaznak – gazdag forrásanyagát használva választhatják ki a felhasználók azt a művet, amelynek recepcióját vizsgálni kívánják. Az eszköz megmutatja szövegek újrafelhasználásait, a találatokat kronológiai előfordulások szerint rendszerezi, és dokumentumrészek alapján csoportosítja. Tervezése során elsődleges szempont volt a távoli és a szoros olvasás módszertanainak együttes alkalmazhatósága, ami művek recepciójának és hatásának árnyalt vizsgálatát biztosítja időbeli távlatban.

3. A SZÖVEG-ÚJRAFELHASZNÁLÁS FELTÉRKÉPEZÉSE AZ ESZMETÖRTÉNETBEN ÉS A RECEPCIÓKUTATÁSBAN

A szöveg-újrafelhasználás tanulmányozása a bölcsészettudományok egyik központi eljárása, amely betekintést nyújt abba, miként terjednek és alakulnak az eszmék és szövegek az idők során. Az ismétlődő vagy közös szövegrészek azonosításával a kutatók nyomon követhetik az idézetek, kölcsönzések, újrayomatások, parafrázisok vagy utalások mintázatait. Mindezek hozzájárulnak a hatástörténeti összefüggések és azok tágabb intellektuális kontextusainak mélyebb megértéséhez.

A kora újkori nyomtatott kultúrában a szövegek újrakiadása megszokott gyakorlat volt: az újságok előszeretettel közöltek teljes bekezdéseket forrás-megjelölés nélkül, a kiadók pedig frissített vagy kibővített kiadásokkal próbálták új vásárlók figyelmét felkelteni, míg az antológiák különféle prózai és lírai műveket gyűjtöttek egybe szélesebb olvasóközönség számára. Az olvasók maguk is részt vettek ezekben a gyakorlatokban – például közhelygyűjteményeket (*commonplace books*) vezettek, amelyekben témák szerint gyűjtötték az idé-

¹⁵ Early English Books Online – Text Creation Partnership (EEBO-TCP), hozzáférés: 2025.02.06, <https://textcreationpartnership.org/tcp-texts/eebo-tcp-early-english-books-online/>.

¹⁶ Eighteenth Century Collections Online – Text Creation Partnership (ECCO-TCP), hozzáférés: 2025.02.06, <https://textcreationpartnership.org/tcp-texts/ecco-tcp-eighteenth-century-collections-online/>.

zeteket, hogy azokat később újra felhasználják. Az efféle átvételek egy olyan kultúra normáit tükrözték, amely nagyra értékelte a klasszikus és kortárs források ismeretét. Ugyanakkor olyan viták, mint a *Régiek és Modernek vitája* (*Quarrel of the Ancients and Moderns*), arra utalnak, hogy a gyakran idézett klasszikus szerzők tekintélyéhez való hozzáállás változóban volt: noha műveiket gyakran utánozták, a modern szerzők bírálták is ezeket. A kora újkorban a plágium fogalma inkább az engedély nélküli újradíjazásra vonatkozott, semmint az ötletek eltulajdonítására. Ebben a kontextusban az eredetiség jól megférhetett a kiterjedt szövegátemelésekkel.

A szöveg-újrafelhasználás szorosan kapcsolódik a recepciótörténethez is, hiszen a későbbi művekben ismétlődő szövegrészekből következtethetünk arra, hogy miként értelmezték és hasznosították újra a korábbi írásokat. Mindazonáltal fontos megjegyezni, hogy ugyanannak a részletnek az idézése nem feltétlenül jelent közvetlen hatást vagy befolyást: a szövegek gyakran merítnek egy közös, ismert idézetekből vagy motívumokból álló repertoárból is. Ahogyan a *Reception Reader* alkotói is hangsúlyozzák: „a szöveg-újrafelhasználás nem szükségszerűen jelent befolyást, mégis értékes információkat nyújt a recepció nagyléptékű vizsgálatához.”¹⁷ Valóban, egy adott szövegrész – például egy bibliai idézet – pusztán ismétlése nem feltétlenül utal közvetlen hatásra, inkább egy hasonló kulturális háttérrel jelölhet, azonban tanulságos lehet az is, ha különböző szerzők eltérően értelmezhetik vagy más célokra használhatják fel ugyanazokat a sorokat.

Neves kutatók rávilágítottak arra, hogy a szöveg-újrafelhasználás mögött összetett dinamikák rejlenek. Harold Bloom például azt vizsgálta, miként birkóznak meg a későbbi szerzők elődeikkel annak érdekében, hogy saját hangjukat érvényesítsék. Julia Kristeva intertextualitás-fogalma (amely Mihail Bahtyin munkásságára épül) azt hangsúlyozza, hogy a jelentés mindig is szövegek közötti kölcsönhatás révén jön létre. Gérard Genette transztextualitás-elmélete tovább finomítja ezeket a kapcsolattípusokat: az intertextualitás megjelenhet például idézetek, plágium és allúziók formájában is.¹⁸ Ezek az elméleti keretek megvilágítják, miért foglal el kiemelkedő helyet az irodalomtudományi kutatásokban a szöveg-újrafelhasználás, amely az intertextualitás egy specifikus formájaként tekinthető. Ezekből kiindulva egy kutatás irányulhat arra, hogyan és miért maradtak fenn bizonyos eszmék; hogyan viszonyultak a szerzők a meglévő művekhez; illetve miként alakították a kulturális

¹⁷ ROSSON, „Reception Reader”.

¹⁸ Glenn ROE, „Text Reuse as Cultural Practice,” *Digital Enlightenment Studies* 2, 1. sz. (2024): 1–30, <https://doi.org/10.61147/des.23>.

normák ezeket az adaptációs és kölcsönzési folyamatokat. Megfigyelhető továbbá az is, hogyan emeltek át szövegeket egyik korszakból a másikba, néha megerősítve egy korábbi szerző hírnevét, máskor pedig újraértelmezve az eredeti anyagot új célok érdekében. Annak vizsgálata, hogy az írók hogyan választottak ki, módosítottak vagy hasznosítottak újra bizonyos szövegrészeket, betekintést nyújthat a retorikai stratégiákba és az olvasói elvárásokba is. Ily módon a szöveg-újrafelhasználás feltárja az irodalom és a kulturális termelés közötti kapcsolatokat, és szélesebb történeti összefüggéseket vázolhat fel anélkül, hogy szükségszerűen közvetlen hatást vagy lineáris eszmei láncolatot feltételezne.

Ahogy a *Reception Reader* fejlesztői is hangsúlyozzák, „a kontextus és a recepció vizsgálata alapvető fontosságú a kora újkori intellektuális, irodalmi és kulturális történelem megértéséhez.”¹⁹ A digitális fordulat előtt a szövegátvétel azonosítása kézi módszerekkel történt, aminek következtében például a kritikai kiadásokon dolgozó kutatók kisebb léptékben tudták feltárni a szövegszintű kapcsolatokat. Az elmúlt húsz év során azonban számos számítógépes eszköz és módszer született e célra, például a *passim*, a *TRACER* és a *TextPair*, amelyek már nagy számú szövegre vonatkozóan hoznak megbízható eredményeket.²⁰ Ezeket a lehetőségeket hasznosító projekt például az irodalmi szövegekre összpontosító *Versioning Machine*, a latin költészetet vizsgáló *Tesserae*, valamint a Vergilius *Aeneis*-ének tudományos fogadtatását tanulmányozó kutatások.²¹ Számos projekt, mint például az *impresso Media Monitoring of the Past*, az *Oceanic Exchanges* vagy a *Viral Texts*, egyebek mellett a történeti újságokban található szövegátvételek felismerését is lehetővé teszi, így a hírek áramlását is azonosítani képes kiterjedt idő- és térbeli távolságokban.²²

¹⁹ ROSSON, „Reception Reader”.

²⁰ *Passim*, hozzáférés: 2025.02.06, <https://github.com/dasmiq/passim/>; Matteo ROMANELLO és Simon HENGCHEN, „Detecting Text Reuse with Passim”, *Programming Historian*, 10. sz. (2021), <https://doi.org/10.46430/phen0092>; Pairwise Alignment for Intertextual Relations, hozzáférés: 2025.02.06, <https://github.com/ARTFL-Project/text-pair/>. *TRACER*, hozzáférés: 2025.02.06, <https://www.etrapp.eu/research/tracer/>; BÜCHLER és mások, „Towards a Historical Text Re-Use Detection”, in *Text Mining*, szerk. Chris BIEMANN, Alexander MEHLER (Berlin: Springer, 2014): 221–238, 238, https://doi.org/10.1007/978-3-319-12655-5_11.

²¹ SCHREIBMAN, „The Versioning Machine”, 101–107; LEE, „A Computational Model of Text Reuse in Ancient Literary Texts”, 472–79. Neil COFFEE, „The Tesserae Project: Intertextual Analysis of Latin Poetry” *Literary and Linguistic Computing* 28, 2. sz. (2013): 221–228, <https://doi.org/10.1093/llic/fqs033>.

²² DÜRING és mások., „Impresso Text Reuse at Scale,” *Frontiers in Big Data* 6 (2023), <https://doi.org/10.3389/fdata.2023.1249469>; CORDELL, „Oceanic Exchanges,” *Oceanic Exchanges: Tracing Global Information Networks In Historical Newspaper Repositories, 1840–1914*; David

Számos szöveg-újraterjesztéssel foglalkozó projekt, köztük a *Reception Reader* is, kifejezetten a 18. századra fókuszál. Ilyen az *Intertextual Hub*, egy innovatív digitális bölcsészeti kutatási környezet is – az ARTFL projekt²³ részeként –, amely a dokumentumokat szélesebb intertextuális kapcsolathálózatokban helyezik el, legyen szó akár közvetlen vagy közvetett átvételekről, más művekkel vagy szövegrészekkel közös témákról, illetve lexikai átfedésekről.²⁴ Ez segíti a felhasználókat abban, hogy ne csak az egyes szövegeket, hanem a közös eszmék, témák és szövegrészek révén összekapcsolódó szöveg-halmazokat is feltárják. Például az ARTFL Encyclopédie projektjének keretében digitalizálták Diderot és D’Alembert *Enciklopédiáját*, amelyet a felvilágosodás kori intellektuális hálózatok viszonyában tettek kutathatóvá: szekvenciaillesztési algoritmusokat²⁵ alkalmazva rejtett, korábban nem ismert idézési mintázatokra mutatnak rá. Kimutatták többek között, hogy például 13 szócikk az *Enciklopédiából* jelentős részben Émilie Du Châtelet *Institutions de physique* című művéből „kölszönzött” részeket, a *Contradiction* szócikk esetében a szöveg 95%-át.²⁶

4. A RECEPTION READER ALAPJÁUL SZOLGÁLÓ ADATBÁZISOK ÉRTÉKELÉSE: KORLÁTOK, KIHÍVÁSOK ÉS MEGOLDÁSOK

Mielőtt bemutatnánk, hogyan működik a *Reception Reader*, illetve összehasonlítani más hasonló szoftverekkel, fontos röviden áttekinteni azokat a digitális gyűjteményeket, amelyekre ezek az eszközök épülnek. Az első az *Eighteenth-Century Collections Online* (ECCO), amely az angolszász világban a 18. században nyomtatott anyagok egyik legnagyobb digitális gyűjteménye. A *Reception Reader* fejlesztői arról tájékoztattak, hogy az ezt üzemeltető Gale cég 2016-ban a gyűjtemény első két verzióját, összesen 207 613 művet bocsá-

A. SMITH, Ryan CORDELL, és Elizabeth Maddock DILLON, „Infectious Texts”, in *2013 IEEE International Conference on Big Data*: 86–94, <https://doi.org/10.1109/BigData.2013.6691675>.

²³ The Project for American and French Research on the Treasury of the French Language (ARTFL), hozzáférés: 2025.02.06, <https://artfl-project.uchicago.edu/>.

²⁴ Clovis GLADSTONE, „The Intertextual Hub”, *Digital Enlightenment Studies* 1, 1. sz. (2023), hozzáférés: 2025.02.06, <https://intertextual-hub.uchicago.edu/>.

²⁵ A módszer lényege, hogy a nem pontos egyezéseket is képes azonosítani; a részletekhez lásd például a lapszám *Metronom: versmértékek változatainak azonosítása a lokális szekvenciaillesztés módszerével* című tanulmányát – a szerk.

²⁶ ROE, „Text Reuse as Cultural Practice”.

tott a kutatócsoport rendelkezésére.²⁷ 2020-ban a Gale hozzáfogott további mintegy 90 000 könyv digitalizálásához, amelyek a jövőben az ECCO III részeként kerülnek kiadásra.²⁸ A *Reception Reader* másik jelentős forrása az *Early English Books Online-Text Creation Partnership* (EEBO TCP) 60 327 szövege. Az EEBO gyűjteményben található művek képfájllai azonban gyenge minőségűek, és tudomásom szerint még nem alakították át őket géppel olvasható szöveggé.

Összességében tehát a *Reception Reader* mintegy 56%-át fedi le az *English Short Title Catalogue*nak (ESTC), amely körülbelül 480 000 rekordot tartalmaz olyan nyomtatványokról, amelyeket 1473 és 1800 között adtak ki a Brit-szigeteken, a brit gyarmati fennhatóság alatt álló tengerentúli területeken, valamint az Egyesült Államokban.²⁹ Valójában azonban ennél jóval nagyobb a lefedettség, ha figyelembe vesszük, hogy az ECCO projekt célja az összes egyedi mű és nem feltétlenül azok minden újrakiadásának feldolgozása volt.³⁰ Fontos hangsúlyozni azt is, hogy az ESTC-ben rögzített művek közül több ezernyi – különböző okok miatt – még nem került digitalizálásra, így a COMHIS csapat ezeket nem is tudta felhasználni. Emellett maga az ESTC – mint minden katalógus – sem teljes. Például jelen sorok szerzője 63 olyan művet azonosított szabadkőműves archívumokban, amelyek nem szerepelnek az ESTC-ben; ezek közül néhány a *British Freemasonry, 1717–1813* című, ötkötetes forráskiadásban jelent meg.³¹ Mivel az ECCO TCP gyűjtemény mindössze 3100 címet tartalmaz, a *Reception Reader* fejlesztői kénytelenek voltak az OCR hibákkal terhelt ECCO verzióra is támaszkodni, ami nemcsak az adatbázis kereshetőségére, hanem a szövegelemzés minőségére is hatással van.

A COMHIS kutatócsoport emiatt igyekezett meghatározni az OCR-hibák arányát, és felmérte ezek hatását az adatalapú szövegelemzésekre. Kimutat-

²⁷ Hálás vagyok Mikko Tolonennek és Ville P. Vaara-nak a *Reception Reader* eszközzel kapcsolatos konkrét kérdéseimre adott válaszáért, amelyek jelentősen gazdagították ezt az áttekintő irodalmat.

²⁸ Stephen H. GREGG, *Old Books and Digital Publishing: Eighteenth-Century Collections Online*. Elements in Publishing and Book Culture (Cambridge: Cambridge University Press, 2021), <https://doi.org/10.1017/9781108767415>.

²⁹ English Short Title Catalogue (ESTC), hozzáférés: 2025.02.06, <https://web.archive.org/web/20230812200936/>; ESTC, hozzáférés: 2025.02.06, <https://www.bl.uk/projects/english-short-title-catalogue>.

³⁰ Mikko TOLONEN, Eetu MÄKELÄ és Leo LAHTI, „The Anatomy of Eighteenth Century Collections Online (ECCO),” *Eighteenth-Century Studies* 56, 1. sz. (2022): 95–123, <https://doi.org/10.1353/ecs.2022.0060>.

³¹ Róbert PÉTER, Cécile RÉVAUGER és Jan SNOEK, *British Freemasonry, 1717–1813* (London: Routledge, 2016), <https://doi.org/10.4324/9781315639901>.

ták, hogy az olyan hibák következtében, mint például a hosszú „s” betű vagy a ligatúrák helytelen felismerése, a teljes korpusz tokenjeinek 26,9%-a torzulhatott az OCR-folyamat során.³² Ez azt jelenti, hogy a hagyományos kulcsszavas keresések a szavak legalább 26,9%-át nem képesek megtalálni. Csak ebből a nézőpontból érthetjük meg igazán, hogy a szoftverfejlesztőknek milyen jelentős módszertani kihívásokkal kellett szembenézniük, amikor a rendkívül „zajos” adatbázisokra – mint például az ECCO – épülő szövegátvételt elemző programokat készítették.

Ez a kihívás ösztönözte Clovis Gladstone és Charles Cooney kutatókat, amikor a *Reception Reader*höz hasonlóan a szöveg-újrafelhasználást tanulmányozták az ECCO és az EEBO TCP adatbázisokban, valamint egy klasszikus latin szövegekből álló korpuszban.³³ Céljuk az intellektuális hatások és a szövegátvétel feltérképezése volt ezekben a gyűjteményekben. Ehhez kifejlesztettek egy szekvenciaillesztési algoritmust, a *PhiloLine*-t, amely képes azonosítani a szövegátvétel méreteket és idézeteket nagy méretű, „rendezetlen” korpuszokban is, hiszen kisebb és bizonyos fajtájú (például különböző karakterek ugyanabban a szóban) eltérések ellenére is felismerhetővé teszi az egyezéseket. A megbízhatóság növelése érdekében figyelmen kívül hagyták ugyanazon szerzőhöz tartozó műveket, kiszűrték az újrakiadásokat (a korpusz 43%-át), egységesítették a helyesírást, és szótrigramokká alakították a szövegeket a hasonlóságok elemzése érdekében. Így több mint 60 millió közös szövegrészletet sikerült azonosítaniuk. Az eszköz lehetővé teszi továbbá a felhasználók számára, hogy metaadat-szűrők (úgy, mint szerző, cím, dátum, egyezés mérete), idősoros és kontextusalapú elemzések révén kutassák az újrafelhasznált szövegeket. A projekt többek között betekintést nyújt abba, milyen volt az egyes szerzők, mint például Lucretius recepciója a kora újkori Angliában, illetve rávilágít a Biblia széles körű használatára is (a szövegegyezések több mint 58,5%-a; a felhasználóknak ugyanakkor lehetőségük van kizárni a bibliai forrásokat a keresésekből, ha más típusú összefüggésekre kíváncsiak). Számos olyan átvett szövegrészlet esetében nem lehetett megfelelő forrásazonosítást adni, amelyek az ECCO adatbázis állományát megelőző időszakból (1700 előtt) származtak. Ennek következtében Shakespeare, Milton, Locke és számos más szerző szövegrészletei sem voltak pontosan azo-

³² Mark J. HILL és Simon HENGCHEN, „Quantifying the Impact of Dirty OCR on Historical Text Analysis”, *Digital Scholarship in the Humanities* 34, 4. sz. (2019): 825–843, <https://doi.org/10.1093/lilc/fqz024>. A ligatúra több betűjel összekapcsolása és összesűritése úgy, hogy egyetlen betűtestben jelenjen meg.

³³ Clovis GLADSTONE és Charlse COONEY, „Opening New Paths for Scholarship”, in *Digitizing Enlightenment: Digital Humanities and the Transformation of Eighteenth-Century Studies*: 353–384. Lásd még: <https://commonplacecultures.uchicago.edu/>, hozzáférés: 2025.02.06.

nosíthatók. E szerzők műveinek 18. századi kiadásai és kommentárjai szolgáltak az újr felhasznált szövegrészek legkorábbi forrásaként.

A következőkben azt kívánom bemutatni, miként oldotta meg a COMHIS csapat ugyanezeket a kihívásokat a *Reception Reader* alkalmazás keretein belül. Mindenekelőtt eltérő módszert alkalmaztak a szövegátfedések azonosítására. A *Basic Local Alignment Search Tool* (BLAST) helyi szekvenciaillesztő algoritmus módosított változatát használták, amelyet eredetileg biomolekuláris szekvenciák – például DNS-láncok – összehasonlítására fejlesztettek ki. Ezzel nagyrészt sikerült kiküszöbölni a gyenge OCR minőségből fakadó hibákat, ugyanis a kisebb eltérések ellenére is felismerhetővé váltak a hasonló szekvenciák (például, ha ugyanabban a sorban találhatók különbözőképp átírt szavak). A Turku Egyetem TurkuNLP kutatócsoportja fejlesztette ezt az algoritmust, amelyet korábban már alkalmaztak más digitális bölcsészeti projektekben is, például finn újságcikkek szöveg-újr felhasználásának felismerésére.³⁴ A *Reception Reader* összevetve a *PhiloLine*-alapú eszközzel úgy tűnik, hogy az előbbi több szövegátfedést képes azonosítani. A két eszköz pontos összehasonlítása azonban jelenleg több okból sem lehetséges. Egyrészt a szövegegyezések töredezettsége mindkét algoritmus számára kihívást jelent. Ez a töredezettség több tényezőtől fakad, többek között a gyenge OCR-minőségből, valamint abból, hogy a hosszú fejléc- és lábléc-szakaszok megszakítják a szövegfolyamot, még akkor is, ha a fő szövegtörzs folyamatosnak tűnik. Másrészt ennél is fontosabb, hogy – ellentétben sok természetesnyelv-feldolgozási (NLP) feladattal – ebben az esetben nem létezik arany standard egyértelmű mérőszámokkal, amely alapján meg lehetne mérni és összehasonlítani a szöveg-újr hasznosítást azonosító algoritmusok teljesítményét és sikerességét, illetve az általuk detektált szövegátfedések mértékét.

5. A RECEPTION READER HASZNÁLATA: FUNKCIÓK ÉS LEHETŐSÉGEK

Ebben a fejezetben azt mutatom be, hogyan működik a *Reception Reader* a gyakorlatban.³⁵ A rendszer egy metaadat-alapú keresőfelülettel indul, ahol a felhasználók szerzőnév és könyvcím (egy vagy akár több tag szava) alapján, illetve ezek kombinálva adhatnak meg lekérdezéseket. A keresőmotor támogat-

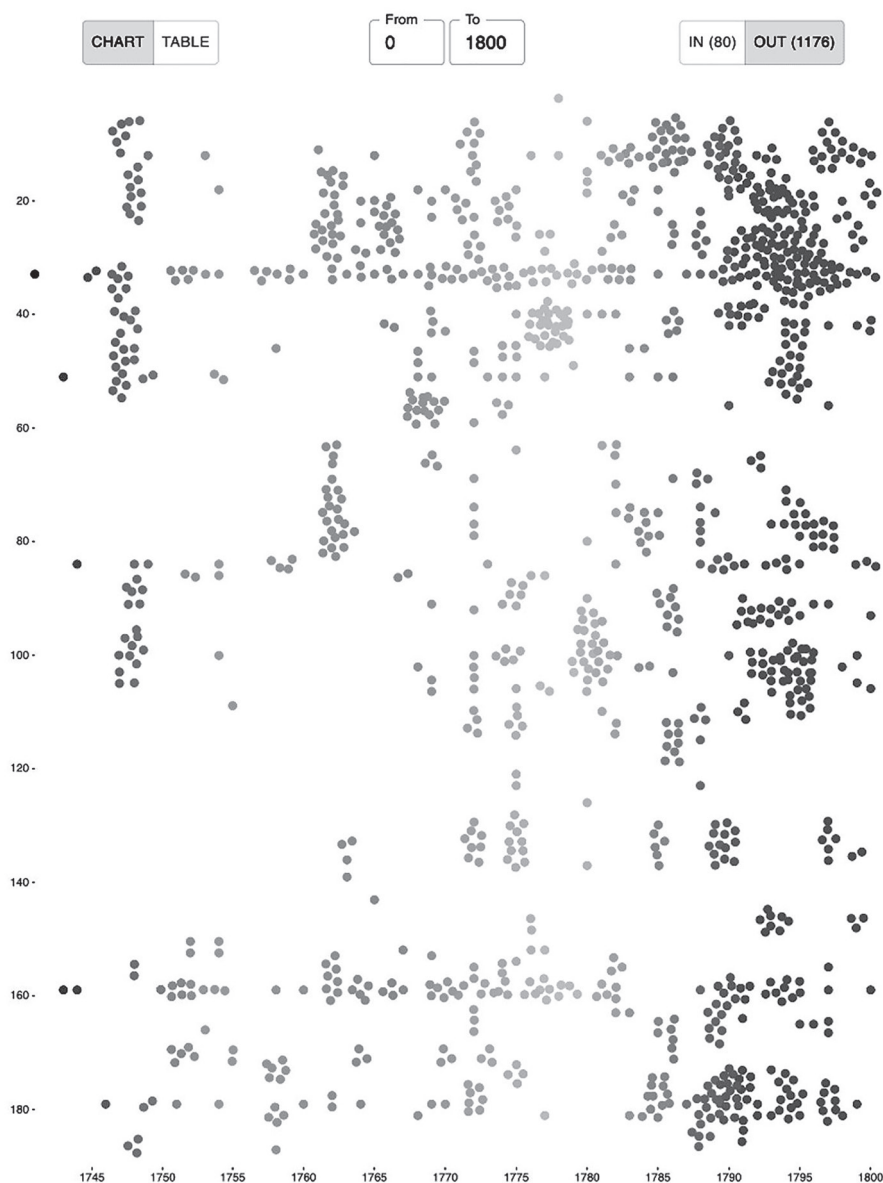
³⁴ Aleksis VESANTO, Asko NIVALA, Heli RANTALA, Tapio SALAKOSKI, Hannu SALMI és Filip GINTER, „Applying BLAST to Text Reuse Detection in Finnish Newspapers and Journals 1771–1910”, in *Proceedings of the NoDaLiDa 2017 Workshop on Processing Historical Language*, szerk. Gerlof BOUMA és Yvonne ADESAM (Gothenburg, 2017), 54–58.

³⁵ ROSSON, „Reception Reader”.

ja az előtag-alapú (*prefix*) és az úgy nevezett „hibatűrő” (*fuzzy*) keresést is, így a csak megközelítőleg azonos esetek is megjeleníthetők. A keresés indítása után a rendszer az első száz találatot mutatja, amelyek relevancia szerint, a találati feltételek alapján rangsorolva jelennek meg. Minden találatnál feltűntetésre kerül a dokumentumazonosító, a kiadás éve, a szerző neve és a cím rövidített alakja. Az oszlopokban található adatok tetszés szerint, az oszlopok fejlécéreire kattintva, növekvő vagy csökkenő sorrendbe rendezhetők. A dokumentumazonosítóra kattintva a felhasználó kiválaszthatja azt a művet, amelyet az elsődleges elemzéshez szeretne használni.

Miután kiválasztottuk az elsődleges dokumentumot, megjelenik a fő kezelőfelület, amely bal oldalon egy „méhraj” diagramot (*bee swarm chart*), jobb oldalon pedig egy oldalkép-megjelenítőt tartalmaz. A méhraj diagram minden egyes pontja egy szöveg-újrafelhasználási esetet jelöl. Alapértelmezés szerint az ábra a „kimenő” kapcsolatokat mutatja, vagyis azokat az újrafelhasznált szövegrészeket, amelyek a vizsgált dokumentummal egy időben vagy annál később megjelent művekben fordulnak elő. Amennyiben a felhasználó arra kíváncsi, hogy korábbi művek szövegei hol bukkannak fel az elsődleges műben, akkor átválthat a „bejövő” nézetre. Az x tengely a kapcsolódó művek kiadási évét ábrázolja, míg az y tengely az azonosított részletnek az elsődleges dokumentumon belüli helyzetét mutatja (legtöbbször oldalszám alapján). Például ha egy 100 oldalas mű 40. oldaláról idéztek később, akkor az y tengelyen a 40-es értéknél jelennek meg az idéző művek adatpontjai. Az egyes pontok színe a kiadások közötti időbeli távolságot tükrözi: a „kimenő” nézetben a kék szín a vizsgált dokumentumhoz közeli kiadási dátumokat, míg a piros pontok fél évszázaddal vagy még később megjelent műveket jelölnek – lásd az 1. ábrán, ahol a szürke minél világosabb árnyalata felel meg a minél közelebbi időbeli távolságnak.

A felhasználó megadhat konkrét évszámtartományokat, így csak az adott időszakon belül megjelent könyvek pontjai maradnak láthatók. A diagram egy szimuláció segítségével kissé eltávolítja egymástól a pontokat, hogy ne fedjék le egymást – ezáltal kirajzolódnak a szövegátvétel időbeli mintázatai. Azok a szövegrészek, amelyeket az adott időszak során többször is idéztek, vízszintesen kirajzolódó pontsorokként jelennek meg egy meghatározott oldaltartományban (y-érték). A függőleges sávok segítségével pedig azonosítható a szövegrészlet különböző időszakokban előforduló újrafelhasználása. Egy oldalirányban kiszélesedő „kúpszerű” alakzat fokozatosan növekvő érdeklődést jelezhet a vizsgált szöveg egy adott szakasza iránt az idő előrehaladtával. Ha a felhasználó egy adatpontra viszi a kurzort, megjelenik egy információs buborék, amely



1. ábra: A *Reception Reader* megjelenítője. Az x tengely a kiadási évek szerint rendezi a szövegeket, míg az y tengely az elsődleges dokumentum oldalszámai alapján az azonosított részleteket. Minden adatpont egy dokumentumot jelöl, amelyben szövegátvétel azonosítható.

tartalmazza a kiadás évét, a szerző nevét, a címet és az elsődleges dokumentum oldalszámát. Ha pedig rákattint egy pontra, a jobb oldali nézetben – amennyiben elérhető – megjelennek a szkennelt oldalképek is, amelyeken kiemelve látható az újrafelhasznált szövegrész mind az elsődleges, mind a kapcsolódó dokumentumban. Ez lehetővé teszi annak vizsgálatát, hogy az idézett szakaszt hogyan illesztették be vagy kommentálták későbbi művekben.

A méhrajdiagram alternatívájaként választható táblázatos nézet is, amely soronként tünteti fel ugyanezeket a kapcsolatokat: minden sor megfelel egy diagramon szereplő pontnak. A sorok tartalmazzák az újrafelhasznált szövegrészek adatait (a kiadás évét, szerző nevét és a mű címét), valamint az eredeti dokumentum idézett oldalszámát. Itt szintén rendezhetők a sorok az oszlopfejlécekre kattintva, illetve kereshetők a böngésző beépített szövegkeresőjével. A táblázat egyik sorát kiválasztva ugyanaz a nézet jelenik meg jobb oldalon, mint a diagram nézet esetében: a kiemelt, újrafelhasznált szövegrészek oldalképei. Összességében ezek a funkciók ötvözik a szöveg-újrafelhasználási mintázatok közvetlen vizuális elemzését a beszkennt forrásanyagok részletes megtekintésének lehetőségével. Így a történések átfogó képet kapnak arról, hogy egy művet miként idézték, dolgozták át vagy bírálták különböző időszakokban. A *Reception Reader* visszajelzési lehetőséget is kínál a felhasználóknak, amely hozzájárul az eszköz további fejlesztéséhez és finomhangolásához.

6. A *RECEPTION READER* KORLÁTAI

A *Reception Reader* innovatív vizualizációs megoldása különösen hasznos, mivel kiemeli az ismétlődő szövegátvételeket, és jól szemlélteti, hogyan idéztek vagy alakítottak át részeket az idők során, ami végső soron a gondolatok és eszmék evolúciójába adhat betekintést. Mindezt pedig a hagyományos lineáris narratívákhoz képest összetettebben képes láttatni: az összefüggések időbeli és tematikus hálózata ugyanis bonyolult kapcsolatrendszerekre világít rá, kiemelve a hivatkozások sűrűségét és időrendi eloszlását. Ugyanakkor fontos figyelembe venni, hogy a digitális bölcsészeti eszközök vizualizációi sohasem objektívek vagy semlegesek. Többek között a szoftverfejlesztés során hozott *design* döntések – például a színskála, a pontméret, az adatelrendezés vagy a szűrés alapbeállításai – óhatatlanul befolyásolják az értelmezést, és elhomályosíthatnak más-különb fontos jellemzőket. Például a felhasználóknak tudniuk kell, hogy a *Reception Reader* alapértelmezetten a legjobb egyezési pontszám alapján rende-

zett első száz találatot jeleníti meg. A méhrajdiagram sűrűsége vagy a színátmenetek hatására a kutatók esetleg túlhangsúlyozhatják egyes szövegrészek jelentőségét, vagy éppen figyelmen kívül hagyhatják az adatokban lévő következetlenségeket és hiányokat. A méhrajdiagram sűrűsége és dokumentumcsoportosításai azt hivatottak jelezni, hogy „mely könyvrészeket tárgyalták a legtöbbet, és melyeket nem”, valamint hogy időbeli trendeket világosítsanak meg. A diagram célja a recepcióban megjelenő újrachasznosítás „volumenének” vizualizálása. Ugyanakkor ez a „valódi volumen” torzulhat, mivel járulékos elemek – mint például kiadói impresszumok vagy reklámok – is gyakran újrachasznált szövegekként jelennek meg, és előfordulhat, hogy egyetlen dokumentum a forrásszöveg több részét is kommentálja, melyet később újra kiadnak; ez felnagyítja az újrachasználás mértékét és torzíttja a vizuális reprezentációt. A közeljövőben a *Reception Reader* alkotói terveik szerint összevonják a szomszédos újrachasználási eseteket, kiszűrik a zavaró járulékos szövegelemeket, és finomítják a diagramkészítési mechanizmust, hogy az adatok alapján pontosabb vizuális narratívát nyújtsanak.

A *Reception Reader* jelenlegi verziója dokumentumszintű áttekintést nyújt a szövegátfedésekről, ezért a kutatók nem tudják például egy könyvfejezet utóéletét külön vizsgálni. A szöveghiányokat illetően fontos megemlíteni, hogy az eszköz szándékosan kizárja ugyanazon szerző más műveit, valamint a teljes egészében újranyomott szövegeket. Emellett nem célja a programnak az allúziók vagy szemantikai hasonlóságok azonosítása sem. Mindezek miatt elengedhetetlen az ábrakon alapuló felismeréseket kiegyensúlyozni a szoros olvasás eredményeivel, amelyhez maga az eszköz is támogatást nyújt.

A BLAST-alapú módszer lenyűgözően átfogó, mivel szinte minden releváns szövegátfedést képes azonosítani. A készítőik azonban hangsúlyozzák, hogy a jelentős OCR-zaj miatt az eredmények nem hibátlanok. A *Reception Reader* nem feltétlenül azonosítja a két jelenleg használt digitális gyűjtemény összes újrachasznált szövegrészletét; nem is beszélve arról, hogy a gyűjtemények lefedettsége is hiányos az *English Short Title Catalogue*-hoz viszonyítva. A lábjegyzetek gyakran kimaradnak, és a sokszor ismétlődő szövegrészek pedig felnagyíthatják az újrachasználási esetek számát. Egyes hosszabb átvett szövegrészek töredezett formában jelennek meg, mivel lábjegyzetek, fejlécek, valamint az OCR hibái szét darabolják őket. Ráadásul a páronkénti hasonlóságvizsgálat nem veszi figyelembe a szövegkölcsonzés történeti irányát, és számos „duplikált” eredményt hoz létre, amikor egy szövegrész sok dokumentumban is újrachasználásra kerül. Noha ezek a tényezők megnehezítik az elemzést az adatok mennyiségének és töredezettségének növekedésével,

együttal rámutatnak arra is, mennyire fontos a kritikus értelmezés: az eszmétörténészeknek a detektált példákat továbbra is manuálisan kell ellenőrizniük és kontextusba helyezniük, nem támaszkodhatnak kizárólag a módszer automatikus kimenetére.

7. HÁROM ESETTANULMÁNY

Ebben a fejezetben három esettanulmányt mutatok be, amelyek szemléltetik a *Reception Reader* használatát, és kiemelik a legfontosabb kutatási eredményeket. Az első tanulmány címe *A Comparative Text Similarity Analysis of the Works of Bernard Mandeville*, amely Bernard Mandeville műveinek és recepciójának feltárására törekszik nagyszabású szöveg hasonlósági vizsgálat segítségével.³⁶ Az elemzés empirikus alapokon járul hozzá az eszmétörténeti diskurzushoz azáltal, hogy újraértékeli Mandeville hatását a 18. századi brit és európai gondolkodásra. Pontosabban Mandeville szövegátemelési szokásait, hatásmintázatait és a szerzőség kérdéseit vizsgálja, különös tekintettel az *A Modest Defence of Publick Stews* című művére. Ez utóbbi kapcsán a kutatás felismerése, hogy Mandeville előszeretettel kölcsönöz saját korábbi műveiből is, miközben a *Publick Stews* esetében ilyen átvételek nem találhatók, ezáltal – a negatív bizonyíték elvét használva – megkérdőjelezhető a szöveg Mandeville-nek tulajdonított szerzősége. Ez a megközelítés árnyaltabb szerzőségvizsgálatot eredményezett a pusztán stilisztikai elemzésre építő hagyományos módszereknél.

A módszer továbbá feltárta Mandeville-recepció mintázatait is: írásai nemcsak politikai és filozófiai művekre hatottak, hanem irodalmi és vallási szövegekben is visszhangoztak. Az intertextuális hálózatok segítségével a tanulmány bemutatja, hogyan terjedtek el és alakultak át Mandeville gondolatai az idők során. Mandeville íásaiból 2557 mű (3989 kiadásban) vett át szövegrészeket: például Frederick Eden (akinek társadalmi kérdésekről, például a szegénységről szóló írásai természetes módon kapcsolódnak Mandeville-hez), a régiségtudós John Smith (1747–1807), valamint a whig történész Catharine Macaulay. Az elemzés meggyőzően bizonyítja, hogy a szövegújrafelhasználás-felismerés kulcsfontosságú az eszmétörténet számára, mivel skálázható módszereket kínál az eszmék terjedésének nyomon követésére, a hatás mérésére és a kritikai kiadások hatékonyabb elkészítésére.

³⁶ Yann Ciarán RYAN, Ananth MAHADEVAN, és Mikko TOLONEN, „A Comparative Text Similarity Analysis of the Works of Bernard Mandeville,” *Digital Enlightenment Studies* 1, 1. sz. (2023): 28–58, <https://doi.org/10.61147/des.6>.

A második esettanulmány Mark G. Spencer és Mikko Tolonen kutatása, amely David Hume *Essays* című művének 18. századi brit fogadtatását tárgyalja.³⁷ Megközelítésük ötvözi a hagyományos eszmetörténeti módszereket az automatikus szövegújrafelhasználás-elemzéssel. A kutatás számos, korábban figyelmen kívül hagyott esetet azonosított, amikor Hume esszéit átvették, parafrázták vagy újranyomtatták, ami befolyásolta a közéleti diskurzust a kereskedelem, kormányzás és szabadság kérdéseiről. A szerzők bizonyítják, hogy Hume művei sokkal szélesebb körben terjedtek, és jóval nagyobb hatást gyakoroltak a korabeli értelmiségi kultúrára, mint azt korábban gondolták. Esszéit gyakran név nélkül terjesztették, és váratlan helyeken – például antológiákban, fejezetmottókban vagy folyóiratokban – is felbukkantak. Ezek a szövegátvételek, amelyeket szerkesztők gyakran új kontextusba helyeztek, alakították a kortárs viták kereteit. A számítógépes és a kézi elemzés ötvözésével a tanulmány ezúttal is új utakat nyit a kora újkori szövegek recepciójának és átalakulásának nyomon követésére, és árnyaltabb képet ad az eszmék nyilvános térben való terjedéséről és evolúciójáról.

A harmadik esettanulmány Kira Sophie Hinderks mesterszakos diplomamunkája, amely a *Reception Reader* használta a 18. századi utópisztikus művek kapcsolatainak elemzésére.³⁸ Kutatásának célja, hogy e szövegátvételi gyakorlatokat összekapcsolja a tágabb nyomtatott kultúrával, az eszmetörténettel és a társadalmi-politikai diskurzussal. A tanulmány azt vizsgálja, miként használták az utópisztikus szerzők a szövegátvételt az értelmiségi hagyományok megerősítésére vagy megkérdőjelezésére. Például Thomas Northmore 1795-ben megjelent *Memoirs of Planetes* című művében szándékosan idézett olyan klasszikus szerzőket, mint Cicero és Tacitus, akiket hagyományosan Nagy-Britannia kevert államformájának és a parlament legfőbb hatalmának igazolására alkalmaztak. Northmore azonban e bevett érvek kiforgatására törekedett: ugyanazon szövegeket arra használta, hogy bírálja Nagy-Britannia politikai berendezkedését, és a korabeli radikális eszmének számító népszuverenitás mellett érveljen. A *Reception Reader* kulcsszerepet játszott a kutatásban, bár korlátai is megmutatkoztak: például az OCR-zaj, illetve a kiadói impreszumok és hirdetések ismétlődéséből fakadó álpozitív találatok rontották a pontosságot. Az első két esettanulmányhoz hasonlóan, ezeket a hiányosságokat kézi ellenőrzéssel és műfajalapú kategorizálással sikerült enyhíteni és így

³⁷ Mark G. SPENCER és Mikko TOLONEN, „The Reception of Hume’s Essays in Eighteenth-Century Britain”, in *Hume’s Essays*, szerk. Max SKJÖNSBERG és Felix WALDMANN (Cambridge: Cambridge University Press, 2025): 15–35, <https://doi.org/10.1017/9781009047227.003>.

³⁸ Kira Sophie HINDERKS, *Cromwell on the Moon; Or, Printing, Popularity, Persuasion*, MA tézis, Uppsala University, 2023.

biztosítani az elemzések megbízhatóságát. Hinderks a hagyományos szoros olvasás módszerét alkalmazta az újrafelhasznált szövegrészek tágabb történeti kontextusban való értelmezésére. Kutatása feltárta, hogy az utópisztikus művekben a szöveg-újrafelhasználás nem véletlenszerű, hanem gyakran tudatos stratégia volt az intellektuális hagyományokhoz való kapcsolódás vagy azok kritikája céljából. Az utópisztikus szerzők az átvett szövegeket a tekintély jelzésére, a kortárs politikai vitákba való bekapcsolódásra és társadalom-kritikai nézőpontjaik kiemelésére használták. Hinderks szerint „a szövegújrafelhasználás-felismerő eszközök – mint a *Reception Reader* – legnagyobb erénye abban rejlik, hogy felszínre és előtérbe hozzák ezeket a rejtett szövegeket, így lehetővé teszik a történészek számára, hogy szakterületi tudásukkal értelmezzék a feltárt összefüggéseket.”³⁹

8. ÖSSZEGZÉS ÉS JÖVŐBELI KILÁTÁSOK

Az elmúlt két évtized tömeges szövegdigitalizálási hulláma hatalmas digitális archívumokat hozott létre, amelyek teljes körű feldolgozása új, adatalapú módszereket igényel. Ezek közé tartozik a szövegújrafelhasználás-felismerés is, amely kulcsfontosságú eszköz a recepció és a szövegek utóéletének feltárására: segítségével azonosíthatók az idézetek, átvételek, parafrázisok és szerkesztői módosítások mintázatai. Amint azt a fentiekben bemutattuk, a *Reception Reader* különösen hasznos lehet ebből a szempontból, mivel lehetővé teszi az eszmetörténészek számára a közeli és távoli olvasás együttes alkalmazását. A kutatók nemcsak a visszatérő szövegátvételek időbeli eloszlását vizsgálhatják, hanem azok kontextusát és más művekben való újrakontextualizálását is. Ez a képesség jelentősen gazdagítja az intellektuális hálózatok, az ideológiai átalakulások és az eszmék fejlődésének vizsgálatát. Emellett új lehetőségeket nyit a 17–18. századi hivatkozási stratégiák és a szövegátadási módok feltérképezésére, összekapcsolva kevésbé ismert műveket a kor meghatározó társadalompolitikai vitáival és kulturális gyakorlataival.

A *Reception Reader* módszertani és vizuális kialakítása jelentős technológiai újítást képvisel: a BLAST szekvenciaillesztő algoritmusra épülő robusztus szövegújrafelhasználási eljárás segítségével képes leküzdeni a nagyméretű és „zajos” korpuszokban – például az ECCO-ban és az EEBO-ban – előforduló OCR-hibák okozta kihívásokat. A jövőbeli fejlesztések célja a rendszer korpuszának bővítése – például a British Library *Burney* és *Nichols* gyűjteményei-

³⁹ Uo., 36.

nek történeti újságanyagaival –, valamint az ECCO és a brit digitális sajtóarchívumok közötti szövegátfedések azonosítása.⁴⁰ Ez az integráció tovább növeli az eszköz jelentőségét a történeti, irodalmi és kulturális kutatások területén. A *Reception Reader* egyedülálló digitális ökoszisztémája példaértékű interdiszciplináris együttműködés eredménye adattudósok, történészek és nyelvészek között. Ráadásul a platform szabadon elérhető felületének köszönhetően demokratizálja a tudományos kutatáshoz való hozzáférését. A *Reception Reader* segít megkerülni az alapjául szolgáló adatbázisok fizetős falait, így használata az anyagi erőforrásokban szűkölködő intézmények kutatói számára is új lehetőségeket kínál.

A *Reception Reader* jövőbeli fejlesztési irányai kapcsán a COMHIS-csoport nagy nyelvi modellek integrálását tervezi, hogy a szemantikai hasonlóságot is felismerje, így a szó szerinti szövegátvételeken túl a parafrázisok és allúziók azonosítása is lehetővé váljon. Ez az innováció árnyaltabb, „eszmeközpontú” elemzést kínál, amely átformálhatja a meglévő értelmezéseket, és eddig ismeretlen intellektuális hatásokat tárhat fel. Emellett a COMHIS-csoport egy kiadás-összehasonlító eszköz fejlesztésén is dolgozik, amely segít felmérni ugyanazon mű több kiadásán végbement változásokat – ez különösen nagy segítséget nyújt a kritikai kiadások készítőinek. E fejlesztéseket a gyakorlati munka során tesztelik: a kutatócsoport készíti többek között David Hume *The History of England* című munkájának kritikai kiadását az Oxford University Press gondozásában. E fejlesztések rávilágítanak a számítógépes szöveg-elemzés átformáló erejére: új lehetőségeket kínálnak az eszmetörténészek számára az eszmék, szövegek és hagyományok összekapcsolt evolúciójának nyomon követésére, olyan léptékben, amelyre korábban nem volt példa.

⁴⁰ MAHADEVAN, „Optimizing a Data Science System for Text Reuse Analysis,” *arXiv*, 2024.

IN MEMORIAM

Bezeczky Gábor
(1955–2025)

Ha Bezeczky Gábor tudományos pályájának összefoglalására vállalkoznék, minden bizonnyal – és talán nem alaptalanul, hiszen ez állt hozzá a legközelebb – Krúdy-kutatásaival kezdeném a felsorolást. Kulcsszerepe volt abban, hogy a korábbi, főként a korai írásokat mellőző Krúdy-kiadások mellett elindulhasson egy új, átfogó összkiadás-sorozat. De nem csupán a szövegek közreadása köthető hozzá, hanem azok értelmezése és ezzel együtt a Krúdy-œuvre átértékelése is. Életművének áttekintésekor nem hagynám ki azt sem, hogy tematikus lapszámot állított össze a *Helikon* számára *A jelentésteremtő metafora* címmel 1990-ben. Ez a szám és további metafora-kutatásai is jól mutatják, milyen kérdések izgatták Gábort: a téma az irodalom és nyelvészet határterületeire, a nyelv- és jelentéstudomány felé vezette, eredményeit pedig irodalmi szövegek olvasásakor is kamatoztatta.

Még hosszasan folytathatnám a sort, hiszen Gábor szerteágazó érdeklődése számos, többek között sajtótörténeti vagy narrációelméleti fókuszú írásban öltött testet. Mégsem kísérel meg tudományos életművének átfogó összegzését, indoklásul pedig éppen az ő szavait veszem kölcsön. A 2021-ben elhunyt Kálmán C. Györgyről ezt írta a *Literatura* emlékszámban: „Kálmán C. György nekem nem tudós, nem tanár, nem megértendő probléma volt, hanem barát. Személyes ügy.”¹ Nem tudnám pontosabban megfogalmazni, mit jelentett nekem Gábor.

Ennek ellenére gyakran diskuráltunk szakmai ügyekről, meséltünk egymásnak aktuális kutatásunkról. Mindig lenyűgözött széles tájékozottsága és olvasottsága; nem volt olyan téma, amihez ne tudott volna legalább két-három könyvet ajánlani a kimeríthetetlen könyvtárából. Olyan kutatónak ismertem, akit soha nem a kimenet, a felmutatható eredmény érdekelt, hanem maga a folyamat. Nem riadt vissza az aprólékos adatgyűjtéstől; heteket, hónapokat szánt olyan projektekre is, amelyek végül az asztalfiókban maradtak. A kényszerfeladatok nemhogy nem hozták lázba, hanem egyenesen bosszantották,

¹ BEZECZKY Gábor, „Nehezemre esik”, *Literatura* 48, 1. sz. (2022): 14–16, 14.

míg a saját, önmaga által kijelölt témákkal kitörő örömmel foglalkozott. Szenvedélyesen kutatta például a titokzatos, ismeretlen nyelven írt, máig megfejtetlen Rohonci kódexet, és szívesen osztotta meg kollégáival a századfordulós írókat, hírlapírókat, szerkesztőket feltérképező adatbázisát.

De ugyanezzel a lelkesedéssel vetette bele magát minden másba is, ami érdekelte; élete végén ugyan már kevésbé foglalkozott irodalommal, de szívesen olvasott a pénz történetéről, és szinte szünet nélkül zenét hallgatott. Barátságunkat alapjaiban határozta meg a zene, legyen szó Mozarttról, Telemannról, a Beatlesről vagy a King Crimsonról, akikhez képest – ahogy ő mondta – még a Led Zeppelin is túl hagyományosnak hatott a hetvenes években, amikor töretlen rajongása elkezdődött. A Hammond-orgonáról ugyan megosztott a véleményünk, de ez semmit nem vont le beszélgetéseink értékéből. Csípős, sziporkázó humora az utolsóig vele maradt, ahogyan nyelvi eleganciája is. Sosem volt pongyola; még egy hétköznapi, baráti levélváltás során is gondosan és választékosan fogalmazott, és csak a legindokoltabb esetben használt felkiáltójelet. Gyakran és szívesen mesélt az életéről, gyerekkori emlékeiről, utazásairól, nyelvtanári élményeiről, és mindig meg tudott lepni egy-egy meghökkentő, váratlan életrajzi ténnyel. Élete utolsó heteiben tudtam meg például, hogy a kilencvenes években lefordított egy Tarzan-kötetet.

Gábor korábban egy másik titkot is megosztott velem – bár ez inkább afféle műhelytitok volt, semmint hétpecsétes misztikum. Ráadásul nem is szakmai, hanem gasztronómiai természetű: elárulta a tökéletes palacsinta receptjét. Ugyanis Gábor – ezt sokan tudják róla – valóban a legjobb palacsintát sütötte, és gyakran hívta össze barátait, kollégáit, hogy közösen fogyasszák el a legendás ételt. A titkos összetevő valójában nem különösebben rendkívüli, mégis tagadhatatlan, hogy a palacsinta íze és állaga is jobb tőle. Amikor én készítettem el az édességet immár a leleplezett hozzávalóval, és Gábor érdeklődött, hogy sikerült, be kellett vallanom: nem rossz, de valahogy mégsem az igazi. Nem sokkal később meglepett egy profi palacsintasütővel, mondván, ha ebben sütöm, hibátlan lesz. A végeredmény valóban még jobb lett, de – bár neki nem vallottam be – valami még mindig hiányzott. Azóta rájöttem, mi az az igazi és most már pótolhatatlan összetevő: hogy nem ő sütötte.

Major Ágnes

KÖNYVEK

TARNAI CSILLAG

tarnai.csillag@gmail.com

ORCID: 0000-0002-6019-7939

Robert BOYD és Peter J. RICHERRSON. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press, 1988; Robert BOYD és Peter J. RICHERRSON. *Not By Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press, 2004.

„Kultúra” alatt különböző tudományterületek képviselői mást és mást értenek, így olyan nehezen definiálható fogalomként működik, amely ugyanakkor alapvetően határozza meg kérdésfeltevéseinket. A bölcsészet- és társadalomtudományoknak egyik legfontosabb tárgyaként jelenik meg, gondoljunk akár a kritikai kultúrákutatók (*cultural studies*) vagy a kulturális antropológia gyakran nagyon is eltérő törekvéseire. A neveléstudományi, pedagógiai és pszichológiai elméletek, valamint a természettudományok szintén létrehozták saját kultúrafogalmukat, ennek megfelelően az emberi fejlődés, nevelődés, biológia területén különböző meggyőződések léteznek a kultúra mibenlétéről. A nézőpontok sokszínűsége mellett érdemes megjegyezni, hogy a kultúra történetiségében is változó fogalom; magyar viszonylatban többek között Márkus György összegző igényű munkája, a *Kultúra, tudomány, társadalom: A kultúra modern eszméje* tárgyalja a kultúra jelentésváltozásait a 18. század második felétől.

Mindezeknek a meggyőződéseknek az összefűzésére, a darwini evolúcióelmélet mintájára egy általános kulturális modell kidolgozására vállalkozott 1988-as *Kultúra és az evolúciós folyamat* (*Culture and the Evolutionary Process*) című könyvében Robert Boyd antropológus és Peter J. Richerson környezettudós. A különböző kultúrafogalmak útvesztőjében való eltévedést megelőzendő Boyd és Richerson saját kultúrafogalmukat már az első oldalakon világosan lefektetik: „[k]ultúra alatt

a tudás, az értékek és más, a viselkedést befolyásoló tényezők nemzedékről nemzedékre történő átadását értjük, tanítás és utánzás útján” (2). Vagyis ezen a területen is az egyedek közötti átöröklítés a folyamat lényege, akárcsak a biológiai evolúció során. A két folyamat tehát egymás mintájául szolgálhat – sőt a szerzőpáros szerint nem is elválaszthatók egymástól. Érvelésük szerint ugyanis az egyedek külső, megfigyelhető tulajdonságai és/vagy viselkedési mintázatai (fenotípus) három feltevel – a gének, a környezet és a kultúra – folyamatos interakciójának az eredménye. A kultúra fogalmát nem szűkítik le kizárólag az emberekre, ám mivel a kulturális folyamatok az emberi evolúcióban játszanak kitüntetett szerepet, ezért elsősorban arra összpontosítanak könyvükben.

Boyd és Richerson elsősorban a szociobiológia eredményeire támaszkodva dolgozzák ki elméleteiket, de vitatkoznak is velük. A szociobiológusok munkájában konszenzusnak tekinthető, hogy mivel a kultúrára való képesség maga is az evolúció terméke, a kultúra legtöbbször maximalizálja az egyedek evolúciós sikerességét és rátermettségét (*inclusive fitness*). Boyd és Richerson ugyanakkor rámutatnak arra, hogy a kulturális evolúció olyan eredményekhez is vezethet, amelyek a biológiai evolúció szempontból nem segítik az adaptációt. Többek között ennek a felismerésnek, és annak köszönhetően, hogy a szerzőpáros a társadalomtudományok eredményeit is felhasználja az elmélet kidolgozásához, a könyv arra vállalkozik, hogy árnyaltabban tárja fel az összefüggéseket a kultúra mint az átadásra alkalmas rendszer és a genetikai evolúció mint neodarwinista folyamat között. Ennek eszköze a Boyd és Richerson által „kettős öröklélméletnek” (*dual inheritance theory*) nevezett modell kidolgozása: a genotípusok és fenotípusok, a gének és a viselkedési mintázatok – tágan értve: biológia és kultúra – egymást dinamikusan alakítva változnak.

Boyd és Richerson modelljük aprólékos kifejtését követően különböző erőket azonosítanak és

modelleznek, amelyek hatnak a kulturális öröklődés tendenciára. Ezekkel (irányított variáció, közvetlen torzított átvitel, természetes szelekció, frekvenciafüggő torzított átvitel és közvetett torzított átvitel) a könyv különböző fejezetekben foglalkozik. Az irányított variáció (*guided variation*) során az egyének nem véletlenszerű, hanem célirányos vagy problémamegoldó tanulás révén hoznak létre új kulturális viselkedésformákat vagy módosítanak meglévőket. Ez a folyamat elsősorban egyéni tanuláson alapul, és gyakran az adott környezeti feltételekhez való adaptív alkalmazkodást szolgálja. Mivel nem pusztán utánzásról van szó, az irányított variáció lehetőséget ad az újításra és a kulturális sokféleség megjelenésére. A természetes szelekció a kulturális evolúcióban azt jelenti, hogy bizonyos kulturális viselkedésformák vagy eszmék nagyobb valószínűséggel maradnak fenn és terjednek el, mert növelik az egyének túlélési vagy szaporodási esélyeit, esetleg javítják gazdasági vagy társadalmi pozíciójukat. Más szóval: a kulturális „változatok” versenyeznek egymással, és az adaptívabb, hasznosabb – gyakran kognitív vagy funkcionális előnnyel bíró – formák idővel elterjednek, míg a kevésbé hatékonyak kikopnak a populációból. A közvetlen torzított átvitel (*direct bias*) olyan társas tanulási mechanizmus, amely során az egyének nem véletlenszerűen, hanem a tartalmuk alapján preferálva választanak kulturális változatokat. Egy adott viselkedés vagy ötlet nagyobb eséllyel kerül átvételre, ha azt az egyén hasznosnak, vonzóknak, érdekesnek vagy könnyen megérthetőnek találja. Itt tehát a választás oka maga a kulturális információ tartalma, nem pedig annak forrása vagy elterjedtsége. A frekvenciafüggő torzított átvitel (*frequency-dependent bias*) esetében a választás alapja, hogy egy kulturális változat mennyire gyakori vagy ritka. A legelőnyösebb változat legtöbbször a leelterjedtebb, így a leggyakoribb változat elsajátítása gyakran a „helyes” kulturális változat elsajátításának legkönnyebb módja. A közvetett torzított átvitel (*indirect bias*) során az egyének nem közvetlenül a kulturális információ tartalmát értékelik, hanem a tudás forrását veszik alapul, amikor eldöntik, mit tanuljanak el. Az emberek ilyenkor olyan modelleket választanak tanulásra – például sikeres, tekintélyes, népszerű vagy megbízható személyeket –, akiknek a viselkedése követésre érdemesnek tűnik. Ez a mechanizmus lehetővé teszi, hogy gyorsan alkalmazkodjunk egy

társadalmi környezethez, de nem garantálja, hogy a tanult viselkedés adaptív lesz.

Eric A. Smith, a kötet egyik recenzense kiemeli, hogy a szerzők kivételes éleslátással határolják el – ahol az elhatárolás egyben kölcsönhatásuk feltárásának is az alapja – az egyéni tanulás adaptív okait és következményeit a kulturális átadás útján megvalósuló társas tanulás okaitól és következményeitől, azaz a folyamatok mikro- és makroszintjét. Ez különbözteti meg a szerzőket azoktól az általuk „holisztikusabbnak” nevezett paradigmákat követőktől (például a szociológiai és antropológiai funkcionalisták vagy a marxisták), akik azt állítják, hogy a társadalom formálja az egyéni viselkedést. Boyd és Richerson modelljeiben ezzel szemben „a két szint kölcsönösen kapcsolódik egymáshoz; a nagyléptékű folyamatok hatással vannak a kisléptékű jelenségekre, és fordítva” (23). Ezzel a megközelítéssel képesek arra, hogy az egyén társadalomra gyakorolt hatását is érzékenyen közelítsék meg, miközben a nagyobb rendszer hatásával is számolnak.

Smith emellett a matematikai modelleket is dicséri, amelyek aprólékosan modellezik a gének és a kultúra közötti interakció számos különböző formáját. A könyv ugyanis rengeteg matematikai képletet tartalmaz. Ennek szükségességét azzal indokolják a szerzők, hogy empirikus vizsgálódások hiányában munkájuk „a tudomány feladásával és a spekulatív filozófiába való visszahúzóódással járna” (28). Tisztában vannak ugyanakkor azzal is, hogy könyvüket nem kizárólag természettudósok fogják majd forgatni, így alapos és részletes leírásokat mellékeltek a képletekhez. A könyv egyértelműen egyik legnagyobb erénye ennek köszönhetően az a közérthető nyelvezet, amelynek segítségével a szerzők bonyolult elméleti belátásokat összegeznek, és tesznek könnyen hozzáférhetővé mind a természet-, mint a bölcseszeti- és társadalomtudományok felől közelítő olvasók számára.

Minde mellett, hogy elismeri a vállalt tudományos feladat relevanciáját és a befektetett munkát, Smith két ellenvetést fogalmaz meg a kidolgozott elméleti gondolatmenettel kapcsolatban. Mindezekelőtt hangsúlyozza, hogy bár érti a szerzőpáros kritikáját a szociobiológia korábbi irányzataival szemben, nincs teljesen meggyőződve arról, hogy a kettős öröklélmélet modellje és a kulturális szelekció valóban annyira jelentősek, hogy a genetikailag öröklődő mintákhoz hasonlóan szignifi-

káns hatást gyakoroljanak az evolúciós folyamatokra. Smith ezzel összefüggésben úgy gondolja, hogy a szerzők nem támasztják alá elégséges, a szövegben szerepeltetett empirikus bizonyítékkal a változt jelenségeket.

Smith ellenvetései döntően a bemutatott elmélet empirikus bizonyíthatóságát kérdőjelezi meg, ugyanakkor recenzióban egy, a kutatás kérdésfeltevésének ideológiáját érintő aggályomnak szeretnék hangot adni. Ez nem példa nélküli a kötet recepciójában: korábban is többen rámutattak, hogy a szerzők kultúrafelfogása nem elég „aktivista”, és nem számol azzal, hogy a kultúrát az egyének nemcsak elfogadják vagy elutasítják, akár módosítják, hanem a kultúrának sokan aktívan ellenállnak, ráerőltetik azt másokra vagy manipulálják saját egyéni haszonszerzésük reményében.

A szerzőpáros már 1988-as könyvében is jelezte, hogy tervezik elmélyíteni kutatásaikat, empirikus vizsgálataikat, és ez valóban így történt: a következő csaknem húsz év munkáját összegezve *Nem csak a géneken keresztül: Így alakította át a kultúra az emberi evolúciót (Not By Genes Alone: How Culture Transformed Human Evolution)* címmel jelent meg ismét közös könyvük 2006-ban. Korábbi munkájuktól eltérően a szerzők expliciten nem mutatnak be, csupán mellékelnek matematikai modelleket és képleteket. Az elméletalkotásról – „Bolondok azok a matematikai modellezők, akik összekeverik absztrakcióikat a valósággal” (255) – az empiriára kerül a hangsúly. Ebben az írásukban a szerzők rövid, tételmondatyszerű címmel emelik ki a kultúra jellegéről alkotott elképzeléseiket, és különböző tudományterületekről származó példákkal, jelenségekkel (néprajz, régészet, pszichológia stb.) támasztják alá érveiket.

Érvelésük középpontjában továbbra is az a tételmondat áll, hogy a kultúra a Darwin által felrajzolt evolúciós folyamatnak megfelelő módon alakult ki és változik, ennek megfelelően pedig a kultúra értelmezése létfontosságúvá válik az emberi viselkedés elemzésekor. Augustín Fuentes recenziójában a szerzőpáros öt legfontosabb állítását összegzi a néhány évvel korábban megjelent *A kultúrák eredete és evolúciója (The Origin and Evolution of Cultures)* című tanulmánykötetük alapján: 1. a kultúra olyan információ, amelyet emberek egymástól sajátítanak el utánzás, tanítás, vagy más szociális tanulás útján, 2. a kultúra változását a darwini evolúciós folyamathoz hasonlóan kell el-

képzelnünk, 3. a kultúra az ember biológiájának része, 4. a kultúra teszi az emberi evolúciót különlegessé más élőlények evolúciójához képest, 5. a gének és a kultúra együtt változnak, fejlődnek.

Az első könyvhöz hasonlóan ennek az írásnak is az egyik legfontosabb eredménye annak felmutatása, hogy a kultúra, a környezet és a biológia hogyan fonódnak szorosan össze. A harmadik fejezetben részletesen tárgyalt gén–kultúra koevolúció kapcsán a kulturális evolúciót alakító szelekciós erők működését gondolják újra. A genetika kapcsán a mutáció, a természetes szelekció és a sodródás merülnek fel, míg a kultúra esetében a „véletlen erők” (kulturális mutáció és kulturális sodródás), a „döntéshozó vagy tanulási torzítások” és a már tárgyalt „kulturális természetes szelekció” érvényesítik hatásukat. Ezek nagyrésze már a korábbi könyvben is kifejtésre került. Kulturális mutációról véletlenszerű, egyéni szintű folyamatokból eredő hatások esetén beszélhetünk, például akkor, ha egy egyén tévesen jegyez meg egy kulturális elemet. A „véletlen erők” másik fajtája a kulturális sodródás (*drift*), amikor statisztikai anomáliák gyakorolnak kiterjedt hatást kis populációkra. Például olyan társadalmakban, ahol bizonyos tudás kizárólag meghatározott személyekhez kötődik, előfordulhat, hogy egy generációban minden ilyen tudáshordozó fiatalon meghal, vagy nem adja tovább tudását – így az adott kulturális elem teljesen eltűnhet. A döntéshozó erők közé tartoznak azok a társas tanulási folyamatok, amelyek során az egyének nem véletlenszerűen, hanem bizonyos szabályszerűségek szerint vesznek át kulturális elemeket: például az irányított variáció (újítás, célzott tanulás), vagy a különféle torzított átvittelek (például a gyakori, tartalmilag vonzó vagy presztízshez kötődő változatok előnyben részesítése).

Azzal, hogy ebben a könyvben már nem szerepelnek bonyolult, a bölcsészettudományok területén kutatók számára nehezen értelmezhető képletek, a szöveg érvelése és felépítése is jóval közérthetőbb. Ez a beszédmód is hozzájárul ahhoz, hogy – megértve a felvázolt modelleket – Fuentes megállapításával egyetértve képesek legyünk arra, hogy a tisztán evolucionista pszichológiával és a hiperreflexív szociokulturális antropológiával szemben ezen az új irányon keresztül kerüljünk közelebb ahhoz, hogy megállapítsuk, pontosan mit jelent embernek lenni.

A továbbfejlesztett elmélet kapcsán több kritikát is megfogalmaztak. Fuentes például azzal vádolja a szerzőket, hogy túlzottan leegyszerűsítik a darwini modellt, amikor a kulturális változásokat szinte kizárólag a természetes szelekcióval magyarázzák, és nem térnek ki kellő mélységgel az alternatív, komplexebb kulturális mechanizmusokra. Véleménye szerint a kulturális szelekció ilyen értelmezése túlságosan determinisztikus, és nem elég árnyalt ahhoz, hogy megragadja a kulturális dinamika valódi sokféleségét. A recenzió hangsúlyozza, hogy a szerzőpáros érvelése nem elég meggyőző addig, amíg nem tájrák fel részletesebben a kultúra és más evolúciós erők (például szimbolikus gondolkodás, társadalmi struktúrák) közötti bonyolult kölcsönhatásokat.

Egy másik recenzens, Hourdequin szintén kritikai észrevétellel él, amikor egy konkrét példán keresztül kérdőjelezi meg a szerzők fogalmi keretét. A könyv egyik érve szerint a kulturális természetes szelekció akkor is működik, amikor egy korábban utáncs, sikeres személy – például egy kábítószerfüggő, aki börtönbe kerül – kikerül a társadalmi modellkínálatból. Ezzel a társadalom gyakorlatilag „negatív szelekciót” hajt végre, hiszen csökken annak az esélye, hogy a kábítószer-függőség mint kulturálisan másolható viselkedésforma, fennmaradjon. Hourdequin kritikájának lényege, hogy a szerzők példája – miszerint egy negatív viselkedésformát a társadalom kiszorít – nem illeszkedik világosan a természetes szelekció elvébe, ha nincs megfelelően definiálva, mi alapján „szelektálódik” egy kulturális minta. A fogalmi keret hiányosságai miatt az olvasó számára elmosódik a határ a kulturális és biológiai fitness között, így a szerzők érvelése nem elég meggyőző vagy következetes.

Mindazonáltal Boyd és Richerson munkája hiánypótlónak tekinthető abban az értelemben, ahogyan hatalmas elméleti apparátus és szakirodalom mozgatószárával letisztult, közérthető nyelven fogalmaznak meg a humán- és természettudományokat egyaránt érintő elméletet a kultúra jellegéről. Az elmélet egy „harmadik útnak” tekinthető a kultúraellenes, evolúcióban gondolkodó pszichológusok és evolúcióellenes társadalomtudósok álláspontja között. A tudományterületek közelítése pedig olyan termékeny eredményekkel járhat, amelyek kölcsönösen befolyásolják és alakítják egymás kultúrafogalmát.

SÁNDOR LÁSZLÓ
sandorl195@gmail.com
ORCID: 0000-0003-3287-6165

Alberto ACERBI. *Cultural Evolution in the Digital Age*. Oxford: Oxford University Press, 2020.

Alberto Acerbi olasz teoretikus *Cultural Evolution in the Digital Age* című munkája a kognitív antropológia, a pszichológia és a társadalomtudomány határterületén tárgyalja a kulturális gyakorlatok és szimbolikus (írott és multimediális) tartalmak fennmaradását és terjedését az általános digitalitás korában. Újszerű szempontokat szolgáltat nemcsak az irodalomelmélet kánonképződést tárgyaló munkáihoz, de a filozófiai megalapozottságú médiaelméletekhez képest is: a tartalmak vagy éppen a közvetítő médiumok sajátosságainak vizsgálata helyett azokra a pszichés folyamatokra koncentrál, amelyek mindig is meghatározták a kultúra továbbörökítésének folyamatait. Módszertana is újszerű: kvalitatív elemzések helyett kvantitatív kutatási eredmények segítségével igyekszik bemutatni, hogy miért terjedhetnek jobban bizonyos tartalmak (például a negatív hangvételű hírek) az online térben, és hogy igazolhatóak-e az olyan általános vélekedések, mint hogy a közösségi média az ideológiai polarizációt vagy az álhírek térnyerését segítené elő.

Acerbi a kulturális evolúció mozgatórugóit, azaz a kulturális tartalmak és gyakorlatok terjedésére, öröklődésére és változására vonatkozó ökölszabályokat az egyén pszichés működésének szabályszerűségeiből vezeti le. Induktív módon jár el; teoretikus kiindulópontja, hogy a kultúra fő jellemzői az egyének szintjén megvalósuló interakciók statisztikai aggregátuma (23). Állítása szerint az ember társas lény, akinek alapvonása a másolásra való hajlam, a másolás azonban nem véletlenszerűen történik, hanem olyan kognitív imperatívuszok szerint, mint hogy „másold a többséget”, „másold a sikeres és elismert embert” és „másold azt, aki hasonlít hozzád”. A kulturális átöröklődés tehát nem átgondolt és pragmatikusan megalapozott szelekcióként írható le, hiszen ennél fontosabb összetevő a hitbéli ráhatározás. Ahogy a múlt embere számára sem volt pontosan érthető, hogy miért hasznos megfőzni a húst, úgy általában mi sem vagyunk vele tisztában, miért lehet érdemes okoseszközünk operációs rendszerét

frissíteni, vagy hogy biztonságos-e a felhőalapú adattárolás – mégis csináljuk.

Habár a könyvben ismertetett kutatások az említett másolási szabályszerűségek sokaságát sorolják fel, sőt már-már elrettentő módon hangsúlyozzák érvényességüket, Acerbi ezeknek inkább a kontextuális korlátait emeli ki. Példaértékű esete ennek az Asch-kísérlet ismertetése (34–38), amely az 1950-es években azt látszott bizonyítani, hogy az emberek – az esetek mintegy harmadában – hajlamosak elfogadni és képviselni a többségi véleményt, még akkor is, amikor az egyértelműen téves. Az ezzel kapcsolatos konformitás-elv (másként: alkalmazkodási torzítás) általános érvényét azonban Acerbi további kísérletek tanulságai alapján kétségbe vonja, amellet érvelve, hogy az egyén akkor hajlandó vakon követni a többséget, amikor bizonytalan, vagy ennek nincs igazi téje (mint például a kísérlet során). Hiszékenyek vagyunk tehát, állítja ő, de nem olyan mértékben, ahogy azt általában gondoljuk, és ahogy azt egyes kutatások sugallják.

A kulturális öröklődés (azaz a kulturális tartalmak másolása) tehát nem kritika nélkül megy végbe, egy-egy tartalom terjedése sokszor lehet öngerjesztő folyamat „a gazdag még gazdagabb lesz” elv mintájára. Ennek példái nemcsak a keresztnevek vagy a kutyafajták, amelyek esetében korszakonként és régióként néhány nagyon elterjedt és rengeteg kis példányszámú egyed jellemző (sok *Anna* és kevés *Adélia*), de igaz ez az internetes tartalmakra is: a kevés nagy elérésű Instagram-influenszer, YouTube-csatorna és hírportál mellett számtalan alig ismert is létezik. Az ilyen, nem egyenletes eloszlások esetében a jelenlét vagy láthatóság meghatározó tényező, nem véletlen, hogy bevezető a digitális kor egyik legfontosabb tényezőjének az információk hozzáférhetőségét (*availability*) és a tömegek elérésének lehetőségét (*reach*) gondolja (xviii). Még akkor is, amikor más terjedési mintázatokat, szabályszerűségeket ismertet és kritizál, mindig visszatér ennek hangsúlyozásához, és kiemeli más hatásoktól való szétszálazhatatlanságát – vagyis, hogy valaminek mindenekelőtt elérhetőnek kell lennie ahhoz, hogy széles körben elterjedjen.

Az egyéb tényezők egyike a presztízshatás, azaz az elismert emberek követésének jelensége. Az influenszerek korában azt gondolhatnánk, hogy az elismertség önmagában jelentős szellemi

és piaci befolyást biztosít, amennyiben a hírnév elősegítheti a szellemi vagy akár a fizikai termékek terjedését. Acerbi szerint azonban a képlet ennél bonyolultabb. Igaz ugyan, hogy a közösségi médiában a nagyobb követőszám nagyobb elérést jelent, de a hírnév önmagában nem garanciája az eladhatóságnak – sem kulturális, sem piaci értelemben. Azok a kutatások, amelyek a hírességek reklámarcként való felhasználását vizsgálták, akár a hagyományos, akár a digitális médiában, nem igazolták vissza ennek egyértelmű pozitív hatását (59–62). Állításuk szerint a hírvérésnek ez a formája leginkább akkor működik, amikor az adott híresség a reklámhoz kapcsolódó terület szakértőjének tekinthető (például Michael Jordan a sportcipők esetében), ekkor azonban már nem pusztán a presztízskövető tendencia érvényesül (61). Ha tehát el szeretnénk adni valamit, a presztízshatás kihasználása helyett érdemesebb az elérés mértékét maximalizálni, azaz drága Michael Jordan-es reklámok helyett inkább olcsóbb, de nagyobb elérésű közösségimédia-hirdetésekre költeni a költségvetésünkéből. Az influenszerek is akkor hasznos közvetítői a tartalmaknak, ha a körük szerveződő közösségek megbíznak bennük, hozzáértőnek tekintik őket, illetve a követőtáboruknak még van lokális, személyes jellege – számszerűen 10 000 és 100 000 közötti optimális követőszámmal, tudniillik a követők száma és az aktív bevonódás aránya fordítottan arányos (66–67).

A (hír)név és a tartalom viszonyának másik fontos példája a „hamis idézetek” esete. Ekkor a közösségi médiában hangzatos idézetek mellett olyan hírességek nevei szerepelnek, akik azokat soha nem mondták („A tájékozódás okos emberre vall, de bölcsesség az, ha nem hiszünk el mindent, amit a Facebookon olvasunk.” – Albert Einstein). Ez a rész ismerteteti az ezt vizsgáló kutatást, amelynek során véletlenszerűen rendelték híres és kevésbé híres emberek nevéhez idézeteket, és megkérték a kísérleti alanyokat, hogy válasszák ki a számukra tetszetősebb mondatokat. Ezzel párhuzamosan ugyanezt a kísérletet elvégezték nevek nélkül is. Az eredmények azt mutatják, hogy a résztvevők az első esetben is a nevektől függetlenül, pusztán a szöveg alapján ítélték meg az idézeteket. Acerbi szerint ez azt jelzi, hogy a helytelen forrással megjelölt idézetek éppen nem a hozzájuk rendelt nevek miatt terjednek: ha egy idézet megkapó, azt sokan másolják, és egy idő után annak

járolékos eleme, a forrásmegjelölés vesz el vagy módosul, éppen járulékos volta miatt. Ebben az esetben tehát a tartalomnak nem annyira a forrása vagy kulturális kontextusa lényeges, hanem az érzelmi hatásmechanizmusa (57–59). A presztízs-követés tendenciájával tehát a könyv úgy jár el, mint sok más kognitív torzítással vagy evolúciós ökölszabállyal kapcsolatban: először ismerteti ezeket, bemutatja a mellettük szóló érveket és kutatásokat, majd kérdőre vonja azok totális érvényességét.

Arról, hogy a tartalomkövetést milyen érzelmi hatások segítik elő, és ezt milyen módszerekkel lehet kutatni, a könyv egy másik, az állhírek terjedését tárgyaló fejezetében olvashatunk részletesebben (117–155). Itt a *sorozatos reprodukció* (*serial reproduction*) vagy más néven *átviteli láncok* (*transmission chains*) kísérleti módszertanát és annak eredményeit helyezi előtérbe. A sokat hivatkozott módszer valójában a gyermekkorunkból ismert „sugdosós játék” laboratóriumi változata, amely során a résztvevők történeteket mesélnek tovább egymásnak, majd a kutatók a lánc végére létrejövő történeteket összehasonlítják az eredeti változattal. Ezzel a módszerrel kimutatható, hogy az emberek körülbelül kétszer nagyobb valószínűséggel jegyzi meg és adják tovább a negatív töltetű üzeneteket, mint a pozitívakat (ennek a kognitív torzításnak a neve *negativity bias*). Acerbi szerint ennek evolúciós okai lehetnek, hiszen mindig is fontosabb volt a fenyegetések, a rossz hírek továbbadása, mint az örömhíreké. Ha azt halljuk, hogy „a szikla alatt mérgeskígyók vannak”, azt nemcsak kisebb valószínűséggel ellenőrizzük (hiszen ez veszélyes lenne), mint hogy „a szikla alatt gyémántok vannak”, de evolúciós érdekünk terjeszteni is a híret (142). Az említett módszerrel kimutatható például, hogy a fenyegető és az undorkeltő tartalmi elemek nagyobb hatással bírnak az emberekre, azaz hajlamosabbak azokra jobban figyelni, azokat megjegyezni és továbbítani. Ezt a hatást bizonyítják nemcsak az évszázadokat átölő szövegbányászati -vizsgálatok (139), de a közösségi médiára vonatkozó kutatások is. Az állhírek terjedésében sem maga az információs tartalom, a hamis tények a meghatározóak, hanem az érzelmi hatáskeltő elemek.

Az állhírekkel kapcsolatos előítéletek mellett a könyv hosszan elemzi a véleménybuborékok kialakulását és hatásának jelenségét is (96–116). Állítása szerint a közvélekedéssel ellentétben a közös-

ségi háló éppen nem a polarizációt segíti elő, hanem inkább az ellen hat. Ennek oka, hogy a közösségi platformokon – bár tagjai gyakran megerősítik egymás vélelmeit, kialakulnak véleményközösségek, és hajlamosak vagyunk a saját véleményünket támogató tartalmakat terjeszteni (*myside bias*) – nemcsak a saját véleményünkkel és nézeteinkkel, de az azoknak ellentmondó tartalmakkal is találkozunk, amelyek hatnak ránk – sokkal inkább, mint az offline térben. Acerbi bemutat egy kutatást, amely során három országban mérték az emberek elfogultságának mértékét választások előtt és után. Eszerint a politikai tartalmaknak való kitettség inkább csökkentette a polarizációt (113–114). Azt a hipotézist, hogy a közösségi média inkább a polarizáció ellen hat, megerősíti továbbá az a tény, hogy az idősebbek (akik kevésbé vannak jelen az online térben) szélsőségesebb nézeteket vallanak, mint a fiatalok (111), valamint hogy más kutatások a polarizáció tendenciaszerű növekedését már az 1970-es évektől detektálni vélik, és így annak inkább szocioökonómiai okait feltételezik (114–115). Összességében tehát érvényesülhetnek a digitális világban a saját véleményeket megerősítő kognitív torzítások (*self-similarity bias, myside bias*), de a különböző véleményeknek való kitettség mégis inkább a polarizáció ellen hat.

Utóbbi hipotézissel szemben a felkészült olvasó felvetheti, hogy a közösségi média hírfolyamai egyre inkább személyre szabottak, így a sajátunknak ellentmondó vélemények egyre kevésbé jelennek meg számunkra ezeken a felületeken. Ennek az ellenérvnek tudatában van Acerbi is, bár – a korábban említett folyamatokhoz és kognitív meghatározottságokhoz képest – viszonylag röviden és felszínesen foglalkozik csak vele. Az online algoritmusoknak három jellemzőjét írja le és teszi rögtön zárójelbe (204–211). Először is elfogadja, hogy az algoritmusokra jellemző a részrehajlás (példaképp az Amazon munkaerő-toborzó algoritmusát említi, amely a belétáplált adatok alapján a női felvételizőket diszkriminálta), de kiemeli, hogy az emberek sem mentesek ettől, és az ilyen típusú „hibák” korrigálhatók. Ezen túl felveti problémaként, hogy az algoritmusok működése nem belátható olykor a programozók számára sem, de szerinte ez sem egyedi jelenség, hiszen a kulturális folyamatok többsége szintén nem az. Harmadikként pedig a személyre szabott működést mint *algoritmikus torzítást* veti fel, de állítása

szerint az elvégzett kutatások eddig nem erősítették meg, hogy ez számottevő tényező lenne a kulturális átadás során.

A könyv egyik legnagyobb gyengéje éppen az ehhez (az algoritmikus működéshez) hasonló sajátosságok figyelmen kívül hagyása, vagy legalábbis alulértékelése. Ez a típusú korlátoltság persze magából a nézőpontból, a jelenségek tárgyalásának elméleti keretéből következik, de a médiaelméletek követői számára, akik éppen a médiumok jelentésalkotó szerepére kívánják helyezni a hangsúlyt, reflektálatlanságnak tűnhet. A szerző által tárgyalt kutatások jelentős része nem a digitális és digitalitás előtti kulturális öröklődési mintázatokat ütközteti, inkább egy, e differenciát nélkülöző nézőpontból feltételezi azok folytonosságát. További problémaként merülhet fel az online tér kulturális folyamatainak változékonysága is, amelyre Acerbi maga is többször utal, ahogy magának a kulturális evolúciókutatás diszciplínájának a gyors fejlődésére is (xiii, xvii). E változékonyság miatt azonban felmerülhet a gyanú, hogy a bemutatott, konkrét jelenségekre, digitális folyamatokra vonatkozó (általában 8–10 éves) kutatások eredményei talán már nem is egészen relevánsak – például ami a különböző közösségi platformok perszonalizációs algoritmusait illeti (hiszen emlékezhetünk, hogy tíz éve még egészen másként nézett ki a Facebook-hírfolyamunk). Hogy a két fenti szempont, tehát a digitalitás specifikus működésének feltárása és a digitális platformok változékonysága között feszülő ellentmondás mennyiben feloldható, szintén nem merül fel a könyvben. Mondhatjuk persze, hogy ez nem is számonkérhető Acerbi munkáján, hiszen nem erről kíván beszélni. Azt azonban talán joggal róhatjuk fel a szerzőnek, hogy a cím többet ígér annál, mint amit nyújt: az ugyanis a digitális korba ígér betekintést a kulturális evolúció szempontjából, de a digitalitásnak valójában csak egy szegmensét, a közösségi média működését tárgyalja. Részletesebben a Twitterről (mai nevén X-ről) és a Facebookról van benne szó, de a kép- vagy videómegosztó platformokat legfeljebb érintőlegesen, említés szinten tárgyalja, ahogy a blogok, hírportálok, videójátékok egyéb alkalmazások sem kerülnek említésre, ezek kulturális hatásait és evolúciós működését nem elemzi.

Összességében azonban, minden hiányossága ellenére a *Cultural Evolution in the Digital Age* olvasásra érdemes. Nemcsak azért, mert összetetten, a

megszokott bölcsészettudományi beszédmódokhoz képest újszerű módon tárgyalja az információk és kulturális kódok terjedésének kortárs folyamatait, de azért is, mert kritikai elemzései sok közhelyeszerű vélekedést megkérdőjeleznek. Habár végső következtetéseket ritkán von le, a különböző nézőpontok ütköztetése és a statisztikai bizonyítékok árnyaltabbá tehetik a kulturális tartalmak terjedésének és átöröklődésének diskurzusát. Még ha a cím többet ígér is, mint amennyit ez a bevezető jellegű mű kínál, a mediális megelőzöttség dogmatikus tételét („a médiumok határozzák meg tevékenységeinket”) figyelmen kívül hagyva hasznos szempontokkal gazdagíthatja a digitális tartalmak terjedéséről való gondolkodásunkat.

HELIKON

TÓTH BARBARA

toth.barbara0517@gmail.com

ORCID: 0009-0009-1799-7955

Ronald INGLEHART. *Cultural Evolution – People's Motivations are changing, and Reshaping the World*. Cambridge: Cambridge University Press, 2018.

Tudományos igénnyel, ám közérthetően megírt szociológiai mű, amely a kulturális evolúciót mint az emberek értékválasztásai mögött meghúzódó motivációkat vizsgálja empirikus kutatásokból kiindulva. A szerző Ronald F. Inglehart, amerikai politológus, közreműködője többek között az *EuroBarometer* közvéleménykutatásainak és alapító elnöke a *World Values Survey Association*nek, amely vezetésével 105 ország lakosságának (a világ népességének több mint 90 százalékának) értékeit és hitrendszerét vizsgálta 1981 és 2014 között. Ezekkel a statisztikai adatokkal, diagramokkal illesztve kutatásokra hivatkozva érvel *Cultural Evolution* című, pályáját összegző könyvében hipotézise mellett, miszerint az emberek együttélésének módjait a gazdasági körülmények határozzák meg – és hogy a modernizáció nagyjából kiszámítható módon zajlott és zajlik a mai napig. Konkrét meglátásai ugyanakkor sok újdonságot nem hoznak (a könyv 2018-as megjelenésű); a nyugati típusú demokráciák kiépülésének ünneplését adják a populizmus térnyerésének kritikája mellett. Habár felszólal a társadalmi egyenlőtlenségek és a ki-

zsákmányolás különféle módozatai ellen, könyve megmarad a centrumországok neoliberális diskurzusán belül – miközben az adatelemzés és a szociológiai-gazdaságtani megközelítést kiegészítő kulturális evolúció elmélete a már sokat hallott érvrendszer árnyaltabb kibontását teszi lehetővé.

Már a bevezetőben leszögezi, hogy a korábbi modernizációs elméletek azért vallottak kudarcot, mert az emberi viselkedést összetett tényezők széles skálája befolyásolja, így irreális minden olyan állítás, amely pontos, determinisztikus előrejelzéseket akar tenni. Maga is szembesül ezzel a problémával, és ezért választja a kulturális evolúció elméleti hátterét saját eredményeinek értelmezésekor: ez ugyanis erősen interdiszciplináris megközelítés, az antropológia, a régészet, a biológia, a közgazdaságtan, a történelem, a nyelvészet, a pszichológia és a szociológia határterületén. A kulturális evolúció azt járja körül, hogy az egyéni értékítéletet és döntésmechanismusokat milyen közösségi és kognitív tényezők befolyásolják; illetve megfordítva: ezek az egyéni döntések miként eredményeznek kulturális folyamatokat. Inglehart látszólag szintén ezekre támaszkodik, ám döntő faktorként az ezek mögött meghúzódó gazdasági folyamatokat tartja, érdekes egyvelegeként a gazdasági-politikai és kognitív-társadalmi megközelítéseknek. Ennek eredménye, hogy elméletileg nem teljesen megalapozott munkája: a kulturális evolúció inkább mint metafora és gondolkodásmód jelenik meg esetében, semmint kidolgozott és következetesen érvényesített keretrendszer. A könyv ugyanakkor nem a tudományos elit számára készült, megértése nem igényel előtudást, s habár sok grafikont tartalmaz, a nagy mennyiségű adaton alapuló összefüggéseket képesek egyszerű, szemléletes mintázatokban összefoglalni.

Segít megérteni laikusok számára, hogy az emberek értékrendjének és céljainak változása hogyan változtatja meg közös világunkat. A modernizációs elmélet új, empirikusan tesztelt változatával magyarázza például a második világháborút követő társadalmi változásokat a nyugati világban. Bemutatja, hogy a társadalmi és gazdasági biztonság eredményezte az ún. posztmaterális értékek térnyerését, például a környezetvédő pártok előretörését vagy a nemek közötti egyenlőség növekedését. Ennek megfelelően a gazdasági bizonytalanság és annak pszichés következményei okozzák szerinte az olyan politikai jelenségeket, mint

Trump elnöksége, a Brexit vagy a francia Nemzeti Front előtörése, amelyeket a materiális értékek visszatéréseként azonosít.

A könyv bevezetőjében olvashatjuk, hogy a nemek közötti egyenlőséggel és a homoszexualitással kapcsolatos tömeges hozzáállás kétlépcsős folyamat során változott a nyugati társadalomban. Az első szakasz a melegekkel szembeni nagyobb tolerancia és a nemek közötti egyenlőség nagyobb támogatása felé történő fokozatos, lassú elmozdulás volt, amely az idősebb generációk leváltásával ment végbe. Végül ez elérte azt a küszöböt, amikor a magas jövedelmű társadalmakban az új nézetek normává váltak. Ekkor jön képbe a kulturális evolúcióban jól ismert alkalmazkodási torzítás (*conformity bias*). Ennek lényege, hogy az egyén döntése során a többségi véleményre hagyatkozik, mivel az kipróbálnak és sikeresnek mutatkozik – így csökkentve a döntéssel járó kognitív erőfeszítéseket. Amikor a homoszexuálisok elfogadottsága elérte a küszöbértéket, a korábbi ellenzők is elkezdtek támogatón fellépni, ami sokkal gyorsabb változást eredményezett, mint a népességcseré által előidézett folyamatok. 2015-re az Egyesült Államok Legfelsőbb Bíróságának többsége támogatta az azonos neműek házasságát: még az idős bírák is a történelem „jó” oldalán akartak állni. Vagyis e történet szerint gazdasági folyamatok fokozatosan és lassan eredményeznek mentálisbeli változásokat, amelyek a társadalmi dinamikáknak köszönhetően egy idő után felgyorsulnak és paradigmaváltónak bizonyulnak.

Inglehart szerint a kultúra a társadalom túlélési stratégiáját alkotó, tanult viselkedési formák összessége. Megállapítja ugyanakkor, hogy a második világháború után a nyugati társadalmakban a túlélés mindennapi kérdései háttérbe szorultak a gazdasági jólét következtében, s már nem a fizikai biztonság nemzetenként eltérő különbsége határozza meg az emberek világnézetét. Ennek eredményeként a kultúra funkciója is módosult, és már nem csak mint a kollektív adaptációt szolgáló gyakorlatok összességeként, hanem egyéni szempontokat is mind erősebben figyelembe véve kezdett működni. Mivel a közösségi funkció mindig szabályok és tiltások rendszerét is jelenti, ennek gyengülésével erősödnek fel a liberalizációt jelentő posztmaterális értékek: az egyéni szabadságjogok, az individualizmus és az önkifejezés szempontjai. A hagyományos, kollektivisták működés

inkább a periféria országaira jellemző Inglehart szerint, ahol továbbra is előtérben áll a túlélés és az adaptáció kérdése. Vagyis a társadalmi kérdések ezúttal is elsősorban gazdaságilag determináltak, az értékpreferenciákban bekövetkezett változások pedig a kulturális evolúció által leírt hatások szerint alakulnak. Ugyanakkor, amint a gazdasági egyensúly megbillen, a nyugati társadalmak értékpreferenciája is változik – ismét a szabályozások és tiltások, valamint a társadalmi rendet ígérő politikák térnyerése lesz megfigyelhető.

A nagyvívú folyamatok vizsgálatakor a könyv figyelmeztet, hogy az intézmények és az államok felett mindig egy szűk, jómódú elit dönt, amelynek értékértékeit nehezen követi le a többségi társadalom. Habár sokan elfogadnak egy-egy politikai párt vagy véleményvezérek által képviselt nézetrendszert, életmódjuk és jövedelmük nem engedi meg, hogy azokat megvitathassák és valóban magukévá tegyék. E diagnózis szerint az alacsony jövedelmű és kevesebb iskolával rendelkezők sok esetben saját érdekeik ellen szavaznak, és világnézetüket valójában jól marketingelt politikai programok vezérlik. Ez is az alkalmazkodási torzítás egy fajtája, ám ekkor a többség véleményét tudatosan befolyásolja egy szűkebb elit.

Az individualista társadalmakban a nemi egyenlőségnek és az azonos neműek házasságának normalizációja évtizedek munkájába került, vagyis nagymintás kutatásai szerint az 1968-as szexuális forradalom és a feminizmus második hullámának eredményei a 2000-es évekre váltak láthatóvá és domináló alapeszmékké, ami a politikai döntéshozók generációs leváltásával és az egzisztenciális biztonság mindennapiságával valósulhatott meg. Habár manapság mindezt már a törvénykezés is támogatja a gazdag országokban, a szegényebb területeken változatlanul inkább a termékenység elősegítő normák és a vallásos szemlélet dominanciája számottevő, vagyis az abortusz és a válás megnehezítése vagy tiltása, a heteroszexualitástól eltérő, nem reprodukcióval járó orientáció stigmatizálása, valamint a tradicionális nemi szerepek fenntartása. A növekvő biztonságérzet és az egyéni jogok hangsúlyozása általában erodálja a tekintélyelvű rendszereket, amelyek viszont válság idején képesek sikeresen fellépni a hatalomért vívott harcban.

Inglehart a 2008 utáni nyugati kultúrát – szintén nem túl meglepő módon – a bizonytalanság, és

az ebből következő társadalmi regresszió korszakaként írja le. Ezeknek az egyik tünete és okozója a mesterséges intelligencia térnyerése, ami a fejlődésnek egy olyan szakaszát hozza el, amelyben gyakorlatilag bárki munkája helyettesíthető számítógépes programokkal. Prognózisa szerint a gépekkel kifejlesztett csúcstermékek monopóliumot fognak élvezni, az automatizáció pedig tömegeket szorít ki a munkaerőpiacról, miközben a vezetőkhez jut a cég nyereségének egésze. A nyomtatott újságírás szakmája eltűnőben van, s egyre nehezebb megkülönböztetni az állhíreket a valódiaktól, gyengítve ezzel a demokrácia egyik alapvető biztosítékát. Hasonló tendenciát lát abban, hogy az egyetemek vezetése a vállalatok vezérigazgatóihoz hasonlóan megszüntetik a biztos állásokat, és helyettük adjunktusokat és rész munkaidős oktatókat alkalmaznak, akiknek alacsony a fizetésük és bizonytalan a munkájuk.

Mindemögött pedig a modern társadalmak legsúlyosabb konfliktusa húzódik meg az erőforrások többsége felett uralkodó 1 százalék és az őket kiszolgáló 99 között. Ezt a helyzetet tovább fokozza az általános automatizáció, ezért a mesterséges intelligenciával való megbirkózásra irányuló sikeres stratégiák kidolgozása Inglehart szerint kulcsfontosságú feladat a jövőre nézve. Történeti áttekintését onnan indítja, hogy a nagy gazdasági világválságra adott reakcióként sok országban fasiszizmus vagy kommunizmus képeiben egy centralizált gazdaság jött létre. Ugyanakkor a huszadik század második felében a világ azt tanulta meg, hogy az államilag irányított gazdaságok nem működnek jól. A huszonegyedik században viszont azt vagyunk kénytelenek megtanulni, hogy a mesterséges intelligencia társadalmának hiperkapitalista, „a győztes mindent visz” logikáját csak kormányzati beavatkozással lehet ellensúlyozni. A piac és az állam közötti helyes egyensúly kialakítása szintén kísérletezést és éleslátó innovációt igényel, de amikor a túlélés a tét, az emberek általában felnőnek a feladathoz – bizakodik Inglehart.

Az amerikai politológus könyve valóban több tudományág ismeretanyagának összehangolásával íródott, azonban a fejezetek ismétlései didaktikusá teszi az olvasmányt. Ráadásul látéletei nagyon is ismerősek lehetnek, tulajdonképpen a mainstream neoliberális társadalomelméleteket mondja újra, még ha esetenként annak kritikáját is kívánja adni. Ami felveti a lehetőségét, hogy Inglehart

gondolkodására is az alkalmazkodási torzítás gyakorol erős hatást (azaz a többségi véleménnyel tart). Megtévesztő lehet továbbá a könyv címe, hiszen hiányzik az időbeli kontextualizálás: megfontolandó lenne alcímként hozzátenni, hogy minden eredmény és maga a teória a második világháború óta eltelt időszakról ad számot, nem átfogó történelmi láttelest. Ez esetben is a domináns, nyugati országok nézőpontjából beszél, és egy neoliberalis demokráciafelfogást érvényesítve láttatja centrum-periféria viszonyát. Habár ő is hangsúlyozza a szegényebb rétegek és elnyomott társadalmi osztályok érdekérvényesítési nehézségeit, érdemes volna társadalmi osztályok és iskolázottság szerint külön is elemeznie a bemutatott adatokat. Az adatelemzések és értelmezések általánosításától eltekintve azonban izgalmas olvasmány, amely gondolkodásra és cselekvésre is buzdít. Célja, hogy a 99%-ba tartozók felismerjék manipulálhatóságuk okát, és a passzivitásból kitörve megpróbálják saját javukra fordítani a gazdasági játékszabályokat.

— HELIKON —

SZEMES BOTOND
 botond.szemes@ut.ee
 ORCID: 0000-0002-0637-6776

Olivier MORIN. *How Traditions Live and Die*.
 Oxford: Oxford University Press, 2015.

Olivier Morin a kvantitatív kutatások minőségi művelője. Ami alatt nem csak azt értem, hogy milyen példaértékű körültekintéssel hozza létre (gyakran statisztikai) modelljeit, és milyen óvatossággal, több nézőpontot is figyelembe véve értelmezi azok eredményeit. Hanem azt is, hogy munkájának meghatározó felismerése, hogy kultúrtörténeti és antropológiai vizsgálatok sikeréhez elengedhetetlen a mennyiségi szemlélet érvényesítése a módszertanon túl a jelenségek magyarázatakor is – a (makro)folyamatok egyik legfontosabb összetevője ugyanis a gyakoriság. Ugyanez igaz megfordítva is: kvantitatív kérdései (jelen esetben: mennyi ideig él egy hagyomány? hogyan lehetséges hosszantartó kulturális stabilitás?) és kvantitatív módszerei (milyen folyamatokra következethetünk a rendelkezésre álló adatokból, modellekből és kísérleti eredményekből?) esetében *elméletté* állnak össze. Az adatokból nem

von le hirtelen következtetéseket, nem vezet le belőlük csábítóan egyszerű összefüggéseket; éppen a kulturális folyamatok lassú, összetett és sok esetben véletlenszerű változását szemlélteti velük. Mindez a hagyományos evolucionista, fejlődéselvű felfogással szemben nem egyenesvonalú és nem szükségszerűségekből építkező narratívákat eredményez, sőt sokkal inkább az ilyen elgondolások buktatóira mutat rá.

Igaz ez *How traditions live and die* című művére is, amelyben Morin azt vázolja, hogy az emberi kultúra nem a fajra jellemző különleges kognitív képességek következménye, azaz nem kitüntetett-ségünkre vezethető vissza, hanem társas viszonyaink melléktermékeként értelmezhető. Elgondolása jelen esetben, szemben más munkáival, nem annyira módszertanában kvanitativ (a könyv kis részét teszik ki statisztikai modellek, nagyobb részét az elméletalkotás és korábbi munkák áttekintése), inkább az alapvető összefüggéseket tekintve az: meggyőzően érvel amellett, hogy a kultúra kialakulása elsősorban nem minőségi szempontokkal, hanem az egyre kiterjedtebb közösségekben az interakciók megnövekedett számával magyarázható.

Fő kérdés tehát, hogy miként lehetségesek hosszú időn át fennálló hagyományok az emberi közösségekben, amivel tulajdonképpen arra kérdez rá, hogy miként lehetséges emberi kultúra. E megközelítés szerint ugyanis a kultúra nem más, mint gondolatok és gyakorlatok térben és időben kiterjedt rendszere, amely kiterjedtséget az egyénekből felépülő átadási láncok biztosítják. Kiindulópontja, hogy a kulturális átadás (*cultural transmission*) során hagyományosan két problémát azonosítanak a kutatók. Az egyik az Elfáradási kérdés (*Wear-and-Tear Problem*), amely arra a jelenségre utal, hogy az átadási láncokban már néhány lépést követően is torzulhat az információ (gondoljunk gyermekkorunk „sugdosós” játékára). A másik nehézség a Leállási kérdés (*Flop Problem*), amely értelmében még a hűségesen átadott információk láncolata is gyakran megszakad, mert nem tartjuk őket elég fontosnak vagy érdekesnek ahhoz, hogy megosszuk másokkal. E két nehézség a hosszantartó hagyományok kiépülését gátolja. Morin érvelése szerint a kulturális evolúció kutatói korábban szinte kizárólagosan az Elfáradási kérdést helyezték a középpontba, abból az előfeltevésből kiindulva, hogy ha torzításmentes, hűséges egy átadási lánc, akkor és csakis akkor lehetnek

működőképesek az egyes hagyományok. Ez azt a további elgondolást vonja maga után, miszerint a kultúra alapja az imitáció.

Az „imitáció-elmélet” képviselői szerint a kultúránk forrása az a kitüntetett emberi képesség, mely szerint pontos és hűséges utánzásra vagyunk képesek, amelynek segítségével hosszan tartó információs láncokat tudunk létrehozni. Ugyanakkor Morin rámutat e megközelítés vakfoltjaira. Egyfelől a kulturális átadás sok olyan elemet tartalmazhat, amelyet egyáltalán nem kell utánoznunk (például egy útvonalat különböző módokon is megjelölhetünk, de nem a megjelölés módját, hanem magát az útvonalat akarjuk továbbítani és megőrizni). Ha a sikeres hagyományok kizárólag a helyes/igaz imitáción alapulnának, akkor a kulturális közvetítés egy meglehetősen költséges és sekélyes művelet volna. Ezzel szemben kontextusfüggő szelekció jellemez minden elsajátítást, amely során az adott jelenség szempontjából nem fontos részleteket figyelmen kívül hagyhatjuk. Másrészt egy hagyomány kialakulását sok esetben nem imitáció vagy tanulás eredményezi, hanem egyszerű reprodukcióként jön létre. Morin példája az angol nyelvben a segédigék formai elkülönülése a többi igétől: az óangolban ez az elkülönülés még nem létezett, vagyis a segédigéket is ragozták mód/idő/személy szerint; ám a két csoport közötti kiejtésbeli, hallható különbségek generációk során annyira felerősödtek, hogy a segédigéket az igéktől eltérő, saját grammatikai szabályok szerint működő nyelvi csoportként kezdték kezelni. Ekkor nem tudatos műveletek és hűséges imitációk álltak a folyamat hátterében, hanem az örökölt különbségek újratermelése és felerősítése a nagyszámú átadás következtében. Harmadrészt felhívja a figyelmet arra is, hogy az újítás és az átadás nem ellentétes fogalmak, ahogy azt a hűséges, torzításmentes imitáció elmélete sugallja. Bármely gyakorlat vagy gondolatsor elsajátítása az újraalkotás mozzanatát is magában hordozza, miközben ezek a módosítások nem gátolják (ellenkezőleg: elősegítik!) stabil hagyományok kiépülését.

Az imitáció-elmélet másik összetevője a kényszeres, kötelező utánzás elképzelése. Eszerint az imitáció alkalmával két fő hatásnak engedelmessé válnak: mindenekelőtt a népszerűség és a tekintély befolyása alatt állunk. Vagyis azokat a jelenségeket vagyunk hajlamosak utánozni, amelyeket a többség is utánoz (népszerűség), vagy amelyet si-

keres emberek közvetítenek (tekintély). Morin nem veti el ezt, a kulturális evolúció alapvetésének számító elképzelést, amennyiben bizonyosan ezek az összetevők is szerepet játszanak döntéseinkben, ám túl leegyszerűsítőnek gondolja. Egyrészt hangsúlyozza, hogy akkor hagyatkozunk ilyen előzetes és külső megfontolásokra, amikor nincs nagy tétje a döntésnek egyéni boldogulásunkra nézve. Amikor a tételek nagyok, más szempontokat is hajlamosak vagyunk mérlegelni, és kevésbé engedelmessé válnak a „vak” imitáció kényszerének (ennek kapcsán több korábbi kutatás érvényességét is megkérdőjelezi a negatív tömeghatás kiterjedtségéről, amikor az egyén saját érdekei ellenére cselekszik a környezetének nyomására). Ennek érdekében vezeti be a „rugalmas imitáció” fogalmát, hiszen kontextusfüggő az is, hogy éppen milyen ráhatásoknak engedelmessé válnak a kulturális átvétel során. Másrészt ugyan elismeri, hogy ezek a hatások rövid távon módosíthatják egy hagyomány sikerét, ám jelzi azt is, hogy hosszú távon nem játszanak számottevő szerepet annak stabilitásában. Szemléletes példája a kísérlet, amelyben a résztvevőknek zeneszámokat kellett pontszámokkal értékelniük; az egyik esetben anélkül, hogy látnák mások értékelését, a másikban az eddig adott pontszámok ismeretében. A legsikeresebb zeneszámok csoportja mindkét esetben elkülönült a legkevésbé sikeresektől, azaz, úgy tűnik, valamely belső tulajdonságukra, vonzerőre visszavezethető, milyen pontszámot kapnak – miközben a siker *mértékét* erősen befolyásolta, ha a résztvevők ismerték a korábbi értékeléseket. Így tehát azt, hogy mi lesz igazán *big hit*, erősen befolyásolja a népszerűsége alapuló megerősítési hatás (*comformity bias*), de nem azt, hogy valami sikeres-e egyáltalán. Vagyis a népszerűség és a tekintély megerősíthet folyamatokat, de csak már korábban, ettől függetlenül stabilizált hagyományok esetében.

Morin mindezek alapján szakít a korábbi megközelítésekkel, és amellett érvel, hogy az Elköpási kérdés helyett a Leállási kérdés megoldásában keresendő a hosszú és stabil hagyományok, és az azokból felépülő kultúra magyarázata. Azaz első sorban kellő motivációval kell rendelkezniük a résztvevőknek a gyakori utánzásra, és nem hűségnek lenniük forrásukhoz. A legnagyobb probléma az Elköpási kérdés megoldani hivatott imitáció-elmélettel, hogy összekeveri a nagy léptékű terjedést (*diffusion*) az emberközi, kis léptékű át-

adással, végeredményben a kultúrát a pszichológiával. Azt feltételezi, hogy két egyén közötti torzításmentes továbbításon múlik minden, miközben a hagyományok kialakulása nem a továbbítás minőségétől, hanem azok *számától* függ. Egyszerűen elég sok átadás kell ahhoz, hogy *robosztus és redundáns* láncolatok jölessenek létre. Egy ilyen láncolatban a párhuzamosan és változatokban terjedő információ mintegy magát szabályozza, amely lehetővé teszi, hogy az egyes megvalósulásokat *összehasonlítsák és kijavítsák* a résztvevők, ha azt fontosnak tartják. Vagyis ha egyszerre sok résztvevő, kellően gyakran ismétli meg az átadott információt, akkor az esetleges hűtlenségek ellenére is stabil hagyományok jöhetnek létre. Így tehát a Leállási kérdés megoldásával (ha valamely motiváció, például az adott információ vonzereje folyamatos átadásra sarkallja az egyéneket) egy csapásra az Elkopási kérdés problémája is megoldódik, amennyiben már nem gátja a hosszú távú megőrzésnek, ha egy átadás nem tökéletes. Mindez azt is jelenti, hogy az Elkopási kérdéskör elgondolt lineáris átadási láncok (A → B → C) csak laboratóriumi kísérletek sajátjai, a valóságban ehelyett a hagyományok hálózatosan terjednek, amely hálózatokra 1) az ismétlés jellemző (A többször átadhatja ugyanazt az információt B-nek); 2) redundánsak működnek, azaz egyszerre több helyről is érkezhethet ugyanaz az információ; illetve ugyanígy 3) szerzte is ágazhat (egy forrásból több helyre).

Vannak olyan újítások, amelyek hirtelen érnek el sok embert; ezek általában rövid távú sikert jelentenek, de kismértékű stabilitás jellemző rájuk. Ezzel szemben a kiterjedt, de kis léptékű kapcsolatokból álló hálózatok védettek a divatjelenségekkel szemben. A kiterjedtség egyik legfontosabb előfeltétele pedig nem a hűség imitáció, hanem – az átadott tartalom vonzereje mellett, amelyet Morin szintén részletesen tárgyal – az információ *elérhetősége*. Ez az egyének magas élettartalmán, az információátvitel megbízható módjain és intézményein túl mindennekelőtt sűrű társas kapcsolatokat követel meg. Ugyanis azok a láncok tudnak nagy tér- és időbeli távolságot felölelni, amelyek sok, apró kapcsolatból jönnek létre, és amelyek így könnyen elérhetők lokális szinten. A nehezen hozzáférhető információk, legyenek bármilyen vonzó is, kevés eséllyel őrződnek meg generációkon keresztül. Így tehát nem szükséges, hogy egyetlen átadás nagy távolságot öleljen fel, hiszen a stabili-

tást éppen az szavatolja, hogy sok egyénen halad át a gyakran megismételt, lokálisan jól elérhető információ.

Mindez a kulturális evolúcióban sokat hangsúlyozott vertikális átadás (szülőtől gyerekeig, generációk között) mellett a horizontális (generáción belüli) kapcsolatok fontosságára hívja fel a figyelmet – ismét a hálózatok metaforáját előtérbe helyezve az egyenes vonalú láncolatokkal szemben. Morin egyik meggyőző példája ilyen hálózatok működésére a gyerekek kultúrája, amelyben gyakran váltják egymást a generációk és cserélődik ki a teljes népesség (5-6 év után elhagyják ezeket a közösségeket, és az ott felvett szokásokat – például a játékokat, mondókákat – már nem gyakorolják tovább). Mégis kifejezetten hosszú hagyományok tudnak kialakulni ilyen közösségekben is: akár négy-öt évszázadon keresztül fennmaradnak rímek, dalok és mondókák anélkül, hogy felnőttek beleavatkoznának ebbe a folyamatba. Morin részletesen elemzi ezt a jelenséget (a könyv *Függelékében* statisztikai módon is bizonyítja állításait), amelyet hosszú, kiterjedt, de lokális szinten működő átadás-rendszerként jellemez, és amelyet minden sikeres hagyományterjedés modelljének tekint. A gyerekközösségekben egyszerűen a nagyszámú helyi, kvázi-horizontális kapcsolatok sűrűsége teszi lehetővé, hogy a generációk gyors váltakozása ellenére is megőrződjenek bizonyos formák akár extrém hosszú ideig.

Miközben a mondókák vizsgálata azért is szemléletes, mert rávilágít, hogy egy ilyen közegben sem *minden* újítás lesz sikeres. A dalok, rímek, szójátékok java része eltűnik pár generáció alatt, és csak egy-kettő képes több száz éven át életben maradni. Ez ismét a mechanizmusok kvantitatív alapjára mutat rá, amennyiben nagy számú hagyománykezdeménynek kell egyszerre jelen lennie egy közösségben ahhoz, hogy közülük kevés kiválasztódjon és stabil hagyománnyá alakuljon. Kis számú újítás esetében nagyon kevés az esélye ilyen hosszantartó láncok kiépülésének.

Ezen a felismerésen keresztül jut el Morin legáltalánosabb állításaihoz. Korábban antropológusok, filozófusok és kultúrakutatók széles körében egyetértés övezte az elképzelést, miszerint az emberi fajt egy külön, a kultúra létrehozására, megőrzésére és továbbítására kifejlett képesség jellemzi. Ez az ember mint „kulturális állat” képzete. Morin némi ironiával jegyzi meg, hogy valóban jó

érzés ebben a naiv humanizmusban hinni, ám ha szem előtt tartjuk az eddig elmondottakat, lelepleződik az ilyen elméletek vágyvezéreltsége. Ezekből ugyanis az következik, hogy egyszerűen általános kommunikációs képességünk és gyakorlataink *mellékterméke* a kultúra. Nem biológiai oka van annak, hogy kulturális lények lettünk, sőt közösségeink sem lettek okosabbak vagy jobban utánpótlás az idők során. Ezzel szemben csupán kellően nagy számú hagyománykezdeményre volt szükség ahhoz, hogy közülük néhány extrém hosszú ideig fennmaradjon, vagy másképp: több újításra, mint amennyi kezdemény egységnyi idő alatt elhalt. Ez persze kommunikáción alapuló, békés és sűrű kapcsolatkból álló közösségeket feltételez, ám önmagában a kooperáció képessége nem vezet még kultúrához (ahogy nem is kizárólag emberi tulajdonság). Ehhez az újítások – nem tudatos – felhalmozódására és azok gyakori megismétlésére is szükség van. Ennek kedveznek az említett társas viszonyok. Fontos látni azonban, hogy nem azért jöttek létre ilyen közösségek, hogy stabil hagyományokat tartsanak fent. Ellenkezőleg, más célokat szolgáló közösségi működések és kommunikációs rendszerek előre nem kódolt következménye az emberi kultúra kialakulása. Morin provokatív megjegyzése szerint könnyen el tudunk képzelni kognitív, érzelmi és társas szempontból normális életet élő embereket – kultúra nélkül is.

A kultúra kialakulása tehát egy lassú, felhalmozáson alapuló folyamatot jelent, nem pedig pontszerűen kijelölhető minőségbeli változást. És ennyiben inkább mérték kérdése: minél több hagyománykezdemény jellemez egy közösséget, annál nagyobb az esély ezek fennmaradására, valamint minél több hosszútávú és nagy hatókörű, stabil hagyomány alakul ki, annál inkább válik az ember kulturális állattá. Innen nézve az emberiség történetét egy folyamatosan növekvő, de lassú kulturizációként is leírhatjuk.

A lebilincselő gondolatmenetet Morin maga is lassan, apró lépések sorozataként építi fel. (Ezért sem helyettesítheti a könyv elolvasását a mostani hevenyészett áttekintés). Könyve minden egyes érvet, előfeltevést részletesen megvizsgál, elemzését kiterjedt szakirodalmi tájékozódás kíséri. Alapossága ugyanakkor szórakoztató, esszéisztikus nyelvvél párosul, olyan szöveget eredményezve, amely stílusában is mintát adhat a kultúrákutatás

művelőinek. A világos, precíz érvrendszer pedig nagyon is radikális gondolatokat fogalmaz meg, amely során nem csak a kulturális evolúció és a kultúranropológia alapvetéseit kérdőjelezi meg, hanem a kultúra egészére vonatkozó belátásainkat is alapjaiban írja újra.

HELIKON

SCHÄFFER ANETT

anett.schaffer@uni-miskolc.hu

ORCID: 0009-0003-6331-9019

ADAM HAMMOND, szerk. *The Cambridge Companion to Literature in a Digital Age*. Cambridge: Cambridge University Press, 2024, <https://doi.org/10.1017/9781009349567>.

„A közelmúltig a Cambridge kézikönyv és a krimi kifejezés egymást kizáró fogalmaknak tűntek volna” – így kezdődik a 2003-as *The Cambridge Companion to Crime Fiction* (Cambridge kézikönyv a krimiről). Azóta a krimi és általában a populáris műfajok vizsgálata már nem ilyen meglepő, sőt meglehetősen elterjedt az irodalomtudományban, ennek egyik jele pedig maga a *The Cambridge Companion to Crime Fiction*. Ugyanígy a digitális irodalomtudomány (DLS – *Digital Literary Studies*) és a digitális közegben létrehozott, illetve fogyasztott irodalom intézményesülésének egyik lépése lehet a *The Cambridge Companion to Literature in a Digital Age* (Cambridge kézikönyv az irodalomról egy digitális korban) kiadása. Magyar közegben is megfigyelhetők olyan kezdeményezések, amelyek a digitális irodalomtudomány és a digitális bölcsészet egyre szélesebb körű elfogadottságát bizonyítják: ide sorolható a Digitális Örökség Nemzeti Laboratórium, az ELTE BTK Régi Magyar Irodalom Tanszéke által 2018 óta kiadott *Digitális Bölcsészet* folyóirat, az egyetemi oktatásban az ELTE Digitális Bölcsészet Tanszéke; de említhetők olyan tudományos munkák, mint Szemes Botond *Digitális irodalomtudomány* című kismonográfiája, a *Helikon Irodalom- és Kultúratudományi Szemle* jelen, illetve a 2020-as *Számitógépes irodalomtudomány* című száma vagy a *Literatura* 2023-as *Távolság olvasás* nevet viselő blokkja. A bizonyosan további elemekkel kiegészíthető lista hossza már önmagában mutatja, hogy bár új tudományterületről van szó, amelynek alakulásáról

még bőven merülhetnek fel kérdések, egyre nagyobb figyelem irányul rá.

A kutatási terület ambivalens helyzetét egy anekdotával írja le a kötet bevezetőjében Adam Hammond, a kötet szerkesztője: 2013-ban egy konferencián Hammond digitális irodalomtudományi poszterprezentációjára senki sem volt kíváncsi, sőt olyannyira nem figyeltek rá, hogy egy félíg megevert sajtodarabot tett laptopjára egy másik résztvevő. Úgy véli, hogy ez jól mutatja, hogy „az irodalomtudomány fősodrának nehéz volt elfogadnia vagy befogadnia a digitális irodalomtudomány eredményeit” (1). Meglepő, de azóta sem lát ebben pozitív változást, sőt úgy véli, hogy a közönyt „nagyreszt szilárd és sajátos megvetés váltotta fel” (2).

Hammond a bevezetőben összefoglalja a tudományterület fogadtatásának rögs történetét, és kiemeli, hogy jelen kötet célja a „hagyományos” és digitális irodalomtudomány közötti kapcsolat megteremtése, illetve a digitalitásnak általában az irodalomra és az irodalomtudományra gyakorolt hatásának feltérképezése: „Ahelyett, hogy úgy beszélne a digitális irodalomtudományról, mintha a »hagyományos« irodalomtudományon kívül esne, ez a könyv azt a széleskörű hatást tárja fel, amit a digitális technológia gyakorol az irodalom létrehozásának, recepciójának és elemzésének minden vetületére” (6). Hammond úgy véli, hogy három nagy médiumváltás volt eddig az irodalom történetében, és az utolsót épp manapság éljük át (oralitás – írás, írás – nyomtatás, nyomtatott – digitális). A kötet „fő kérdése, hogy mi forog kockán az irodalomtudomány számára az utolsó átmenet során” (6).

Az első fejezetben az irodalmi adat fogalmával és az arról való gondolkodás változásaival foglalkozik Yohei Igarashi. A tanulmány mindezt az 1800-as évek végétől vizsgálja, elsősorban az Egyesült Államokra fókuszálva, továbbá részletesen ír az 1906-ban létrehozott The Concordance Society (*Konkordancia Szövetség*) működéséről, azon belül is Lane Cooper *A Concordance to the Poems of William Wordsworth* (*Konkordancia William Wordsworth verseiből*) című kötetéről. A fejezet ezáltal jól példázza, hogy az adatalapú gondolkodás egyáltalán nem a számítógépek és az internet elterjedésével jelent meg az irodalomtudományban, és hogy az új módszertanok is egy tágabb kultúrtörténet részeként értelmezhetők.

Ted Underwood a kvantitatív elemzés irodalomtörténetírásban való használatáról ír. Úgy véli, hogy a kvantitatív megközelítés ellentéte olyan történeti narratívák létrehozása, melyek „különálló történeteket helyeznek egyszerre kronológiai és implicit módon hierarchikus sorrendbe. A hierarchia alján egyes történelmi figurák és események állnak [...] Magasabb szinten ezeket az történeteket mintákba rendezik, amelyek önmagukban szintén események egy nagyobb történetben” (34). Nem titkolja, hogy a történeti narratíváknak is vannak előnyei (például ok-okozati összefüggések azonosítása), ugyanakkor a kvantitatív módszer erősségének tartja, hogy kiválóan tudja folyamatként ábrázolni az irodalom alakulását, ez azonban nem jelenti azt, hogy a hirtelen, nagy változásokat ne lenne képes megjeleníteni. Példaként saját kutatását említi: ebben egy olyan statisztikai modellt hozott létre, amely nyelvhasználat alapján különíti el a krimiket más műfajoktól, és képes meghatározni a műfaj első megjelenését is. Underwood talált egy érdekes eltérést a kvantitatív vizsgálatok eredményei és a korábbi irodalomtörténetek között: „a korszakok és mozgalmak ismerős körvonalai nem láthatók azokon a grafikonokon, amelyeket a kvantitatív irodalomtudománnyal foglalkozó kutatók eddig létrehoztak. [...] Vagy a kvantitatív módszert használó tudósok nem tudták bemérni az irodalmi változás legfontosabb aspektusait, vagy a korszakkonceptciók nem annyira nélkülözhetetlenek, mint ahogy azt az eddigi irodalomtörténetek sugallták. Vagy, és talán ez a legvalószínűbb, a két állítás egyszerre igaz” (38–39). A kvantitatív módszerektől nem a korszakolás eltörlését várja – hiszen azt abban az esetben hasznosnak tartja, ha történetként gondolunk az irodalomtörténetre –, hanem „egy új módszert a változás ábrázolására” (42–43). A hagyományos korszakolás egyeduralmától való elmozdulás már a digitális vizsgálatokat nem tartalmazó irodalomtörténetekben is megjelent. Magyar közegben jó példa erre a Bényei Tamás és Kállay Géza főszerkesztésében az utóbbi években megjelent *Az angol irodalom története* sorozat 6–7. kötete (Kijárat Kiadó, 2024). Az előszóban is olvasható, hogy angol-szász példákat követ a kötet felépítése: „Az angol nyelvű irodalomtörténet-írásban a huszadik századi fejlemények tárgyalására egyre inkább az évtized-alapú logika jellemző, nem pedig a korszakok vagy irányzatok szerinti tárgyalás. Vagyis az

újabb irodalomtörténeti vállalkozások inkább egy-egy időszak keresztmetszetének megrajzolására, a párhuzamos és egymással kölcsönhatásban lévő jelenségek, törekvések feltérképezésére törekednek, mintsem életművek portrészerről ismertetésére vagy a »nagy művek« kiemelésére.” (11-12)

Mark Algee-Hewitt szintén egy, az irodalomtörténethez szorosan kötődő fogalmat vizsgál, a kánonét. Tanulmánya elején kiemeli, hogy bár a digitális bölcsészet megjelenésekor még azt várták, hogy eltörli a kánon fogalmát, az a mai napig fennmaradt, és a digitális bölcsészettudományi vizsgálatok sem tudtak teljesen elszakadni tőle. Ezután a kánon és a kánonról való gondolkodás alakulására tér ki, majd a kánon meghatározására irányuló számítógépes elemzési módszereket használó kutatásokról és a meglévő kánon ilyen kutatásokban való felhasználásáról ír, miközben a kánon hiányosságait is figyelembe veszi. A kánonalkotó művek számítógépes modellel való meghatározásának két lehetőségét különíti el: az egyik a kánonba tartozó művek közös formai jellemzőinek meghatározására irányul, ebben az esetben abból a feltételezésből indulunk ki, hogy bizonyos formai jellemzők miatt válnak egyes művek a kánon részeivé; „ha azonban azt feltételezzük, hogy a kánon létezése a folyton változó társadalmi és kulturális jelenségekre vezethető vissza, akkor arra használhatjuk a számítógépes elemzés módszereit, hogy ezek történetének specifikus vonásait vizsgáljuk meg” (54). Véleményem szerint az első feltételezés roppant utópisztikus és veszélyeket is hordoz magában, például azt, hogy ez a hozzáállás az irodalom lényegét a meghatározott vonásoknak való megfelelésre szűkítené le, ami óhatatlanul hatással lenne arra is, hogy milyen műveket részesítenek előnyben a könyvpiaci szereplők. Bár vannak ilyen kezdeményezések, Algee-Hewitt szerint eddig még nem tudták formai vonások alapján elkülöníteni a kanonikus és nem kanonikus műveket, amiből azt szűri le, hogy valószínűsíthetően „a kánon létrejöttében szerepet játszó társadalmi tényezők elég sokfélék ahhoz, hogy ne legyen a formai jellemzőknek olyan csoportja, amely minden kanonikus műre érvényes korszakokon, térségeken vagy műfajokon át” (59).

Számos fejezet olyan jelenségeket tárgyal, amelyek nem tartoznak az irodalomtudomány klasszikus tárgykörébe. Az első ezek közül a negyedik fejezet, melyben az irodalmi művek felol-

vasása során használt hangot, pontosabban az úgynevezett „költői hang” (*Poet Voice*) mibenlétét elemzi Marit J. MacArthur és Lee M. Miller. Ez egy olyan kérdés, hogy a digitális eszközök nélkül nehezen lenne pontosan vizsgálható, illetve a szerzők által is megemlített hangoskönyvek elterjedése miatt egyre nagyobb jelentőségre tesz szert.

Katherine Bode a digitális archívumokkal foglalkozik tanulmányában meggyőzően ír arról, hogy milyen téves elképzelések terjedtek el velük kapcsolatban, valamint hogy milyen korlátaik vannak és miképpen kutathatóak, illetve arról, hogy miért nem kezelhetők a hagyományos könyvgyűjteményektől teljesen különállóként: „a digitális archívumokban található dokumentumok és azok elrendezése nagyrészt az analóg archívumokból származik és azon világnézetekből, irodalmi értékmeghatározásokból, amelyet az ilyen infrastruktúrák megtestesítenek” (106).

Claire Battershill, Anna Mukamal és Helen Southworth a tudományos igénnyel készült digitális kiadások kérdéseit taglalják. Példákkal illusztrálják, hogy „miképp lehetnek a digitális kiadások inkluzívabbak, kiterjedtebbek, egyszerűen többek, mint egy nyomtatott kiadás” (kiemelés az eredetiben, 113). A digitális kiadások felépítése, jellemzői és a bennük rejlő új lehetőségek mellett az ezekhez tartozó olvasási, illetve felhasználási stratégiákat is tárgyalják. Ilyen például, hogy az olvasók másokkal megosztott széljegyzeteket készíthetnek vagy hogy a digitális kiadások a számítógépes szövegelemzésre is különösen alkalmasak.

Dennis Yi Tenen a materialitás kérdését járja körül. A fejezet egy anekdotából indul ki: mikor Tenen Lahorban járt Pakisztánban, beszerzett egy „kalóz” változatot Bertrand Russell *A hatalom. A társadalom újszerű elemzése* című könyvének indiai kiadásából. Annak végiggondolása során, hogy miképp jutott ez a lemásolt könyv Lahor egy kis boltjába, a könyvvásárlással, könyveladással kapcsolatos szabályozások áttekintése mellett arra is rámutat, hogy egy-egy irodalmi mű milyen szorosan kötődik a fizikai mivoltához és környezetéhez azáltal, hogy kinek és miképpen van hozzáférése az adott szöveghez. Ezután arra tér rá, hogy a könyv kezelése és birtoklása miképp változik meg a digitális kiadások esetében, illetve milyen igazságtalanságokat hoz létre vagy termel újra az ilyen szövegekhez való hozzáférhetőség.

Tully Barnett az irodalmi piac változásait elemzi, nagy figyelmet fordítva az Amazon vállalat szerepére. Először Robert Darnton híres ábráját, amely az irodalmi piac szereplői közötti kommunikációs útvonalatokat mutatja be, és az ennek megújítását célzó kutatásokat foglalja össze, majd az e-book és a fanfiction irodalmi piacot átalakító hatásait vizsgálja, két példával illusztrálva az olvasók és írók közti út lerövidülését. Az egyik Adam Croft munkássága, aki e-könyvként kezdte el publikálni a krimijeit az Amazon self-publishing programján keresztül, és ért el viszonylag nagy olvasóközönséget és sikert; a másik pedig E. L. James, aki online publikáló fanfiction szerzőből lett hagyományos kiadóval le szerződött író.

A következő négy fejezet olyan jelenségeket vizsgál, amelyek kifejezetten a digitális világban jöttek létre vagy kerültek előtérbe. Anna Wilson a fanfiction jelenséget járja körül. A tanulmány első részében bemutatja a fanfiction és az azzal kapcsolatos szabályozások történetét. Majd a műfajt a közösségi olvasás (*social reading*) fogalmán keresztül vizsgálja. Szociális olvasásnak adnak teret például azok a weboldalak, amelyekben „a felhasználók könyvértékeléseket tesznek közzé, különböző olvasási kihívásokban vesznek részt és kommunikálnak a szerzőkkel” (166–167). Ilyen például a Goodreads (vagy a magyar moly.hu), de azok az e-könyvek és digitális kiadások is, „amelyeket közösségek együtt szeljegyzetelnek, emelnek ki [a szövegkiemelő digitális változatával] és állítanak össze” (167). Wilson a fanfiction megosztására specializálódott weboldalakat (például Fanfiction.net, Wattpad, AO3), melyek lehetővé teszik a fanfiction szerzők és olvasók közötti sajátosan direkt kommunikációt, is ehhez kapcsolódóan tárgyalja. Végül a peritextusokkal foglalkozik, mely fogalom „digitális kontextusban gyakran a metaadat szinonimájának tekinthető” (170): itt az AO3 weboldal felépítése kapcsán arra fókuszál, hogy egyes szövegekről megadott bibliográfiai és tartalomra vonatkozó adatok hogyan és milyen sorrendben jelennek meg a felhasználók számára, irányítva ezáltal a befogadás folyamatát.

Emily Short a szöveges kalandjátékok, szó szerinti fordításban interaktív irodalom (*interactive fiction*) kérdéseivel foglalkozik. „A fejezet első fele olyan szövegekkel kapcsolatos általános elképzeléseket vizsgál, amelyek gyakran nem igazak az interaktív narratívákra” (177), tulajdonképpen a ha-

gyományos irodalmi szövegekkel (*noninteractive literature*) hasonlítja őket össze, majd a szöveges kalandjátékokkal kapcsolatos terminológiát és ehhez kapcsolódóan legfontosabb sajátosságaikat tárgyalja. Short érzékletes példákkal szemlélteti a jellegzetességeket azon olvasók számára is érthetően, akik nem vagy csak kevésbé ismerik a szöveges kalandjátékok világát.

Nick Montfort és Judy Heflin a számítógépes programokkal generált irodalmat vizsgálják, egészen az 1950-es évek elejétől, megjelenésétől kezdve. Írnak a számítógéppel generált irodalom történetéről, különböző típusairól, a „gépi hangról” és arról, hogy milyen következtetéseket vonhatunk le belőle, illetve a számítógéppel generált irodalmi művek vizsgálatának módszereiről. A szerzők a fejezet elején megadják, hogy „[a]z irodalom vagy irodalmi művészet általános meghatározását használ[ják], amibe beletartozik a költészet és a próza, de nem csak ezekre korlátozódik” (194), arra a mindenki számára elérhető mesterséges intelligencia programok megjelenésével egyre aktuálisabbá váló kérdésre azonban nem térnek ki, hogy mennyiben és milyen feltételekkel tekinthetők egyáltalán a számítógépes programokkal generált szövegek irodalomnak.

Timothy Welsh „az irodalom és a videójátékok kapcsolatát” (212) bontja ki. A fejezet kitér a videójáték-kutatás történetére, a videójáték-kutatás és az irodalomtudomány kapcsolatára, miközben a videójáték-szektor változásaira is figyelmet fordít. A tanulmány a videójátékok *Aranypolgára* utáni keresésből indul ki, mely kivételes videójáték azt „demonstrálná, hogy a médium készen áll arra, hogy komolyan vegyük, nem csupán a kereskedelmi lehetőségek miatt, hanem a kultúra részeként” (212). Welsh azonban arra jut, hogy ez a kérdés már elvesztette a jelentőségét, hiszen „a művészek és az írók a huszadik században [...] megmutatták, hogy a művészet nem teljesen a tárgyról szól, hanem arról, hogy mit csinálunk vele, hogy reagálunk rá, hogyan lépünk vele kapcsolatba”. (232)

Inge van de Ven a nyomtatott könyv szerepére kérdez rá a digitális korban. Először „a regény halálát” (234) hirdető, újra és újra feltűnő elképzeléseket tárgyalja, melyek eddig nem bizonyultak igaznak. Ezt követően arra tér ki, hogy miként hívták fel az irodalmárok figyelmét az „elektronikus szövegek, online platformok és a *print on demand* (igény szerint nyomtatott) kiadások megjele-

nése” (239) a nyomtatott könyv materialitására. Ezután a nyomtatott könyveket alanyagként felhasználó művészeti alkotásokról, illetve Jessica Pressman fogalmát használva a „bookish” könyvekről – amelyek „a materiális dimenzióikat és a papír anyagiságát olyan módon hasznosítják, ami nem a narratíván kívül áll, hanem része a jelentésnek” (244) – és a kortárs irodalmi művek méretének jelentőségéről ír.

Garrett Stewart a szöveg és a kép kapcsolatát vizsgálja az irodalmi művekben, „a medialitás reflexív megtestesülés[ét] az irodalmi narratívában” (kiemelés az eredetiben, 251). Első példája a Henry James *What Maisy Knew* című művében szereplő „laterna magica párhuzam” (250), de emellett George Eliot és Marcel Proust műveire is kitér, majd a tanulmány második felében Richard Powers technológiaközpontú regényeivel foglalkozik.

A kötet utolsó fejezetében Gabriel Hankins elméleti nézőpontból közelít a kritikai elméletek (*critical theory* vagy Hankins szóhasználatában *critique*) és a digitális irodalomtudomány kapcsolatához. Ehhez áttekinti a kritikai elméleti nézőpont történetét, illetve a vele kapcsolatos vitákat, mielőtt rátérne, hogy miképp jelenik meg a digitális irodalomtudományban. Tanulmánya elején kijelenti „egy olyan önreflexív digitális irodalomelmélet és irodalomértelmezési gyakorlat mellett érvelek, amely a kritikai elméletek hosszú intellektuális és politikai történelmét komolyan veszi, nem úgy kezeli azt, mint különálló, elemzésre váró diskurzusok sorozatát, hanem mint a gondolkodás és cselekvés domináns módjainak, amelybe a saját bevett kritikai beszédmódjaink is beletartoznak, folyamatos megkérdőjelezését és kiegészítését”. (273)

A kötet fejezetei jól példázzák, hogy a digitális világ milyen jelentős mértékben határozza meg a kortárs irodalom és irodalomtudomány szinte minden szegletét, a digitális irodalomtudomány új kutatási módszerei mellett a meglévő irodalmi művek kezelésére és az újak kiadására is hatással van, emellett pedig olyan új jelenségek tűnnek fel, melyek lazábban vagy szorosabban kötődnek az irodalomhoz, illetve az irodalomtudományos fogalmakhoz. Az egyes tanulmányok azonban nem messiásként tárgyalják a digitális világot vagy a digitális irodalomtudományt, hanem a lehetőségek mellett a korlátokat is számba veszik. A jelen

lapszám bevezetőjében olvasható, hogy „a kultúra nem más, mint társas kapcsolatok során átadott információk és ezek időbeli alakulása.” Azt, hogy mit fogunk átadni a jövő generációinak az irodalomból és az irodalomról, biztosan meg fogja határozni, hogyan tesszük közzé a szövegeket – erről olvashatunk a *Kiadások* (*Editions*), *Materialitás* (*Materiality*), *Az irodalmi piac* (*The Literary Marketplace*) és *A nyomtatott könyv a digitális korban* (*The Printed Book in the Digital Age*) című fejezetekben – és hogy mi és mennyire érhető el könnyedén a nagyközönség számára a múlt irodalmából –, ami *Az archívum* (*The Archive*) és *A kánon* (*The Canon*) című fejezetekben tárgyalt kérdésekhez kapcsolható. Az pedig, hogy pontosan mit fogunk átadni a „magas irodalom” mellett még jelentősebb kérdés, mely összekapcsolódik azzal a dilemmával, hogy mi és miképp képezheti tudományos kutatás tárgyát – ezt járják körül a fanfictionról és a videójátékokról szóló fejezetek.

 HELIKON

TARTALOM

Kulturális evolúció: Nyelv és irodalom

TANULMÁNYOK

SZEMES BOTOND:

Kultúra mint információátadás. Előszó a *Helikon Kulturális evolúció* című lapszámához 613

INBAL ARNON, SIMON KIRBY:

A kulturális evolúció hozza létre a nyelv statisztikai szerkezetét
(Fordította: *László István*) 625

OLIVER MORIN:

Az írás fokozatos evolúciója (Fordította: *László István*) 652

BEN NAGY, ARTJOMS ŠELA, MIRELLA DE SISTO, PETR PLECHÁČ:

Metronom. Versmértékek változatainak azonosítása a lokális
szekvenciaillesztés módszerével (Fordította: *Szemes Botond*) 674

HORVÁTH PÉTER:

Szintaktikai ismétlődés a magyar kanonikus költészetben 695

SZEMES BOTOND:

Szavak a hálóban. Az irodalomtörténet hálózatos megjelenítése 734

PÉTER RÓBERT:

Rejtett hatások feltárása az esztétörténetben a Reception
Reader kutatási eszközzel 755

IN MEMORIAM

MAJOR ÁGNES

Bezeczky Gábor (1955–2025) 774

KÖNYVEK

- ROBERT BOYD ÉS PETER J. RICHERSON
Culture and the Evolutionary Process;
PETER J. RICHERSON, ROBERT BOYD Not By Genes Alone:
How Culture Transformed Human Evolution
/ TARNAI CSILLAG 776
- ALBERTO ACERBI
Cultural Evolution in the Digital Age / SÁNDOR LÁSZLÓ 779
- RONALD INGLEHARDT
Cultural Evolution. People's Motivations are Changing,
and Reshaping the World / TÓTH BARBARA 782
- OLIVIER MORIN
How Traditions Live and Die / SZEMES BOTOND 785
- ADAM HAMMOND, SZERK
The Cambridge Companion to Literature in a Digital Age
/ SCHÄFFER ANETT 788

CONTENTS

Cultural Evolution: Language and Literature

BOTOND SZEMES:

- Culture as Information Transmission. Preface to the *Helikon*
issue titled Cultural Evolution 613

INBAL ARNON, SIMON KIRBY:

- Cultural Evolution Creates the Statistical Structure
of Language (Translated by *István László*) 625

OLIVER MORIN:

- The Gradual Evolution of Writing
(Translated by *István László*) 652

BEN NAGY, ARTJOMS ŠELA, MIRELLA DE SISTO, PETR PLECHÁČ:

- Metronome. Identifying Variants of Metrical Patterns
Using Local Sequence Alignment (Translated by *Botond Szemes*) 674

PÉTER HORVÁTH:

- Syntactic Repetition in Canonical Hungarian Poetry 698

BOTOND SZEMES:

- Words in the Web. A Network-Based Representation
of Literary History 734

RÓBERT PÉTER:

- Revealing Hidden Effects. The Reception Reader as a Tool
of Intellectual History 755

IN MEMORIAM 774

BOOKS 776

INHALT

Kulturellen Evolution: Sprache und Literatur

STUDIEN

- BOTOND SZEMES:
Kultur als Informationsübertragung. Vorwort zum
Helikon-Heft mit dem Titel Kulturelle Evolution 609
- INBAL ARNON, SIMON KIRBY:
Die kulturelle Evolution erzeugt die statistische Struktur
der Sprache (Übersetzt von *István László*) 625
- OLIVER MORIN:
Die allmähliche Evolution des Schreibens (Übersetzt von
István László) 652
- BEN NAGY, ARTJOMS ŠELA, MIRELLA DE SISTO, PETR PLECHÁČ:
Metronom. Identifizierung von Varianten metrischer Muster
mittels lokaler Sequenzanpassung
(Übersetzt von *Botond Szemes*) 674
- PÉTER HORVÁTH:
Syntaktische Wiederholung in der kanonischen ungarischen
Dichtung 695
- BOTOND SZEMES:
Wörter im Netz. Eine netzwerkbasierte Darstellung
der Literaturgeschichte 734
- RÓBERT PÉTER:
Verborgene Effekte aufdecken. Der Reception Reader
als Werkzeug der Ideengeschichte 755

IN MEMORIAM 774

BÜCHER 776

HELIKON IRODALOM- KULTÚRATUDOMÁNYI SZEMLE

1955–1962

Vegyes tartalmú számok

1963

1. sz. A komplex összehasonlító kutatások elvi kérdései
2. sz. Nemzetközi Összehasonlító Konferencia (Budapest, 1962)
3. sz. Amerikai prózairodalom
4. sz. Viták a realizmusról

1964

1. sz. Az összehasonlító irodalomtudomány nemzetközi szemléje
- 2 – 3. sz. A kelet-európai avantgárd
4. sz. Shakespeare-évforduló (vegyes szám)

1965

1. sz. Mai világirodalmi mozgalmak és irányok
2. sz. A szocialista realizmus kérdéseiről
3. sz. Nacionalizmus és kozmopolitizmus; eredetiség – utánzás – hatás fogalmai
(Az AILC IV. Kongresszusa – Fribourg, 1964. – előadásaiból)
4. sz. A kelet-európai összehasonlító irodalomtörténet kérdései

1966

- 1 – 2. sz. Irányzatok és csoportok az 1920–30-as évek szovjet irodalmában
3. sz. Eszmék és művek a modern polgári irodalomban
4. sz. Irodalom és szociológia

1967

1. sz. Irodalom és folklór
2. sz. Pártosság, elkötelezettség, elkötelezetlenség
- 3 – 4. sz. A szovjet irodalomtudomány legújabb eredményeiből

1968

1. sz. A strukturalizmusról
2. sz. Az irodalmi irányzatok mint nemzetközi jelenségek
(Az AILC V. Kongresszusa – Belgrád, 1967. – anyagából)
- 3 – 4. sz. Az irodalom és a társzművészetek

1969

1. sz. Kelet-európai irodalmak a századfordulón
2. sz. Művészet – tömegkultúra – irodalom
- 3 – 4. sz. A számítógépek és a humán tudományok (vegyes szám)

1970

1. sz. A Fekete-Afrika irodalmáról
2. sz. Irodalom és összehasonlító módszer (vegyes szám)
- 3 – 4. sz. Modern stilsztika

1971

1. sz. Irodalom és társadalom (AILC VI. Kongresszus. Bordeaux, 1970)
2. sz. Irodalomelméleti viták Franciaországban
- 3 – 4. sz. A közép-európai humanizmus kérdései (Sopron, 1971)

1972

1. sz. Science fiction (a műfaj esztétikai és poétikai kérdései)

2. sz. Klasszikusaink és Európa
- 3–4. sz. A szocialista országok irodalmának másfél évtizede
- 1973
1. sz. Műelemzés és műfajelmélet (vegyes szám)
- 2–3. sz. Irodalomtudomány és szemiotika
4. sz. A XVIII. század és a felvilágosodás irodalma
- 1974
1. sz. Az AILC kanadai kongresszusa (Montreal, 1973. augusztus 13–19.)
2. sz. Az elsüllyedt kultúrák irodalma
- 3–4. sz. Modern poétika
- 1975
1. sz. Irodalom, világirodalom, nemzeti irodalom
2. sz. Az újabb Délkelet-Európa-kutatások
- 3–4. sz. Az európai romantika
- 1976
1. sz. Szubkultúra és underground
- 2–3. sz. Irodalom és irodalomtörténet Ausztriában
4. sz. Tudomány-e az irodalomtudomány?
- 1977
1. sz. A retorika újjászületése
2. sz. Az AILC VIII., budapesti kongresszusa (1976. augusztus 12–17.)
Különböző kultúrákban eredő irodalmak kapcsolatai a XX. században
3. sz. Az AILC VIII., budapesti kongresszusa.
Összehasonlító irodalomtudomány és irodalomelmélet
4. sz. A budai Egyetemi Nyomda szerepe a kelet-európai népek társadalmi,
kulturális és politikai fejlődésében (1777–1848)
Budapest, 1977. szeptember 5–8.
- 1978
- 1–2. sz. Kutatási irányok a 20-as évek szovjet irodalomtudományában
3. sz. Érték és társadalom. Transzgresszív elemzések
4. sz. Új magyar világirodalom-történet
- 1979
- 1–2. sz. Az ázsiai népek irodalma
3. sz. A jugoszláv népek irodalma
4. sz. Az egyéni és a kollektív a nyelvben és az irodalomban
(FILLM XIV. Kongresszus, Aix-en-Provence, 1978)
- 1980
- 1–2. sz. Recepciókutatás és befogadásesztétika
- 3–4. sz. Az orosz szimbolizmus
- 1981
1. sz. Az irodalom klasszikus modelljei – Az irodalom és a társművészetek –
A regény fejlődése
- 2–3. sz. Régi és új hermeneutika
4. sz. Irodalom és felvilágosodás
- 1982
1. sz. A Vormärz-irodalom és néhány magyar vonatkozása
- 2–3. sz. Új kutatási irányok a szovjet irodalomtudományban
4. sz. Művelődéstörténet és Kelet-Európa

1983

1. sz. Az AILC X. Kongresszusa
2. sz. Irodalomelmélet és beszédaktus-elmélet
- 3 – 4. sz. Irányzatok a mai francia irodalomtudományban

1984

1. sz. Polémiák a francia forradalom előtt
- 2 – 4. sz. Svájc népeinek irodalma – svájci irodalom?

1985

1. sz. A FILLM budapesti kongresszusa – A polonisztika Magyarországon
- 2 – 4. sz. Az olasz irodalomtudomány napjainkban

1986

- 1 – 2. sz. A műfordítás távlatai
- 3 – 4. sz. Szájhagyományok és irodalom a mai Afrikában

1987

- 1 – 3. sz. A posztmodern amerikai irodalom
4. sz. Hlebnyikov és az orosz avantgárd

1988

- 1 – 2. sz. Kanadai irodalmak
- 3 – 4. sz. A stilisztika útjai és lehetőségei

1989

1. sz. Az empirikus irodalomtudomány elmélete
2. sz. A felvilágosodás és nemzeti fejlődés
(A budapesti Nemzetközi Felvilágosodás Kongresszus anyagából)
- 3 – 4. sz. A szövegkiadás új elmélete és gyakorlata: a szövegek keletkezés-kritikája

1990

1. sz. A mai nemzetközi folklorisztika
- 2 – 3. sz. Irodalom és pszichoanalízis
4. sz. A jelentésteremtő metafora

1991

- 1 – 2. sz. A biedermeier kora – nálunk és Európában
- 3 – 4. sz. Hagyomány és modernizáció a mai kínai kultúrában

1992

1. sz. A frankofon irodalmak sajátosságai
2. sz. Profizmus az irodalomtudományban
- 3 – 4. sz. A Név hatalma

1993

1. sz. A konstruktivista irodalomtudomány
- 2 – 3. sz. Elsikkasztott orosz irodalom
4. sz. A mai lengyel irodalomtudomány

1994

- 1 – 2. sz. Az amerikai dekonstrukció
3. sz. A kortárs olasz irodalom
4. sz. Feminista nézőpont az irodalomtudományban

1995

- 1 – 2. sz. Posztszemiotika. A szubjektum-elméletek és a mai irodalomtudomány
3. sz. A stílus diszkurzív elmélete
4. sz. Rendszerelvű irodalomtudomány

1996

- 1 – 2. sz. Intertextualitás
- 3. sz. Újraegyesült Németország – egységes német irodalom?
- 4. sz. A posztkoloniális művelődésemélet

1997

- 1 – 2. sz. A félmúlt klasszikusai
- 3. sz. Hermeneutika az orosz századelőn
- 4. sz. A lehetséges világok poétikája

1998

- 1 – 2. sz. Az újhistorizmus
- 3. sz. Kánonok a kis népek irodalmában
- 4. sz. Textológia vagy textológiák?

1999

- 1 – 2. sz. A szó poétikája
- 3. sz. Latin-amerikai irodalomelmélet
- 4. sz. Kulturális antropológia és irodalomtudomány

2000

- 1 – 2. sz. A romantika tétjei
- 3. sz. A korszakok alakzatai
- 4. sz. (Új) filológia

2001

- 1. sz. Változatok a dialógusra
- 2 – 3. sz. Dante a XX. században
- 4. sz. Az interpretáció érvényessége

2002

- 1 – 2. sz. *A Poétika* újraolvasása
- 3. sz. Autobiográfia-kutatás
- 4. sz. A multikulturalizmus esztétikája

2003

- 1 – 2. sz. A minimalizmus
- 3. sz. Mikrotörténetírás
- 4. sz. Kísérleti irodalom

2004

- 1 – 2. sz. Petrarca: hermeneutika és írói személyiség
- 3. sz. A hipertext
- 4. sz. Wittgenstein poétikája

2005

- 1 – 2. sz. A kritikai kultúrakutatás
- 3. sz. Régi az újban
- 4. sz. Vico körei

2006

- 1 – 2. sz. Kritikai szubjektivizmus
- 3. sz. Frege aktualitása
- 4. sz. Relevancia

2007

- 1 – 2. sz. Alteritás, poétika, filozófia
- 3. sz. Ökokritika
- 4. sz. Etikai kritika

2008

- 1. sz. A második olvasat
- 2 – 3. sz. A közvetítés poétikája
- 4. sz. Az autonómia új esélyei

2009

- 1 – 2. sz. Eszmetörténet és irodalomtudomány
- 3. sz. Szimbólum- és allegóriaelméletek
- 4. sz. A jövőbelátás poétikái

2010

- 1 – 2. sz. Térpoétika
- 3. sz. A félre-értelmezett futurizmus
- 4. sz. A szerző poétikája

2011

- 1 – 2. sz. Testírás
- 3. sz. Meghekkelt valóságok
- 4. sz. Új gazdasági kritika

2012

- 1 – 2. sz. A Helikon repertóriuma 1955–2011
- 3 – 4. sz. Boccaccio 700

2013

- 1. sz. Narratív design
- 2. sz. Kognitív irodalomtudomány
- 3. sz. Corpus alienum
- 4. sz. Esztétika és politika

2014

- 1. sz. Cseh dekadencia
- 2. sz. Funkció történet és kontextuális-kulturális narratológia
- 3. sz. Az archívumok elméletei
- 4. sz. Komparatistikai kutatások az ezredfordulón

2015

- 1. sz. A kreatív írás-oktatás és a kortárs amerikai próza
- 2. sz. Transznacionális perspektívák az irodalomtudományban
- 3. sz. Horatius *Ars poetica*-ja
- 4. sz. Az ólomévek kultúrája és utóélete – A terrorizmus az olasz művészetekben és irodalomban

2016

- 1. sz. Az addikció kulturális és kritikai elméletei
- 2. sz. Biblioterápia, irodalomterápia
- 3. sz. Műfaj és komparatistika
- 4. sz. Szabadkőművesség. Új irányok a 18. századi európai maszonéria kutatásában

2017

- 1. sz. A százéves dada
- 2. sz. Hálózati elmélet és irodalomtudomány
- 3. sz. Sokarcú modernség és irodalomtörténet-írás
- 4. sz. Fénykép és irodalom

2018

- 1. sz. Tzvetan Todorov, a közvetítő
- 2. sz. Nem természetes narratológia

3. sz. Próza a posztmodern (próza) után

4. sz. Poszthumanizmus

2019

1. sz. Dehumanizáció: az elkövető alakja

2. sz. Térkép és irodalom

3. sz. Közép-európai komparatiztika

4. sz. Gérard Genette, a szerszámkészítő

2020

1. sz. Számítógépes irodalomtudomány

2. sz. Posztmodern gótika

3. sz. Posztszekularizáció

4. sz. Irodalmak Afrikában

2021

1. sz. Genetikus kritika

2. sz. Realizmusok

3. sz. Költői igazságosság

4. sz. Irányok a gyerekirodalom-kutatásban

2022

1. sz. Kultúrorvostan/Orvosbölcsezet

2. sz. Diagrammatológia

3. sz. Transzkulturális emlékezetkutatás

4. sz. Francia színházfilozófia. Mimészis

2023

1. sz. Krimi

2. sz. „Inklinációk”. Kortárs női filozófiák

3. sz. Narratív fordulat

4. sz. Ökodiskurzusok

2024

1. sz. Irodalmi adaptáció

2. sz. Etnofuturizmusok

3. sz. Az angol romantika a kortárs elméletekben

4. sz. Új elméletek kereszteződésében: A romantika francia aspektusai

2025

1. sz. Emlékezetkutatás Közép- és Kelet-Európában

2. sz. A neoavantgárd (kelet-közép-európai tükröben)

3. sz. Design és irodalom

4. sz. Kulturális evolúció: Nyelv és irodalom

HU ISSN 0017-999X



ELTE | HTK
TÖRTÉNETTUDOMÁNYKUTATÓINTÉZET



Kiadja az ELTE Humán Tudományok Kutatóközpontja
Irodalomtudományi Kutatóintézet,
Felelős kiadó: Balogh Balázs főigazgató, Kecskeméti Gábor igazgató
Nyomdai előkészítés:
ELTE Humán Tudományok Kutatóközpontja Történettudományi Kutatóintézet
Tudományos Információs Osztály
Vezető: Kovács Éva
Tördelőszerkesztő: Hudecz Andrea
A fedél és a tipográfia Baranyai András munkája
Nyomdai munkák: Prime Rate Zrt., Budapest
Felelős vezető: Dr. Tömcsányi Péter

