

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 ŠELA, *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 metrumcímeknek), 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 megfele-

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>.

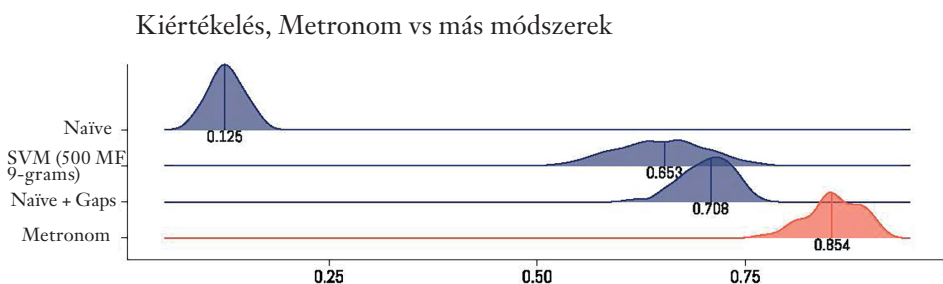
¹⁸ Klemens BOBENHAUSEN és Benjamin HAMMERICH, „Métrique littéraire, métrique linguistique et métrique algorithmique de l’allemand mises en jeu dans le programme Metricalizer 2:”, *Langages* 199, 3. sz. (2015): 67–88, <https://doi.org/10.3917/lang.199.0067>.

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.



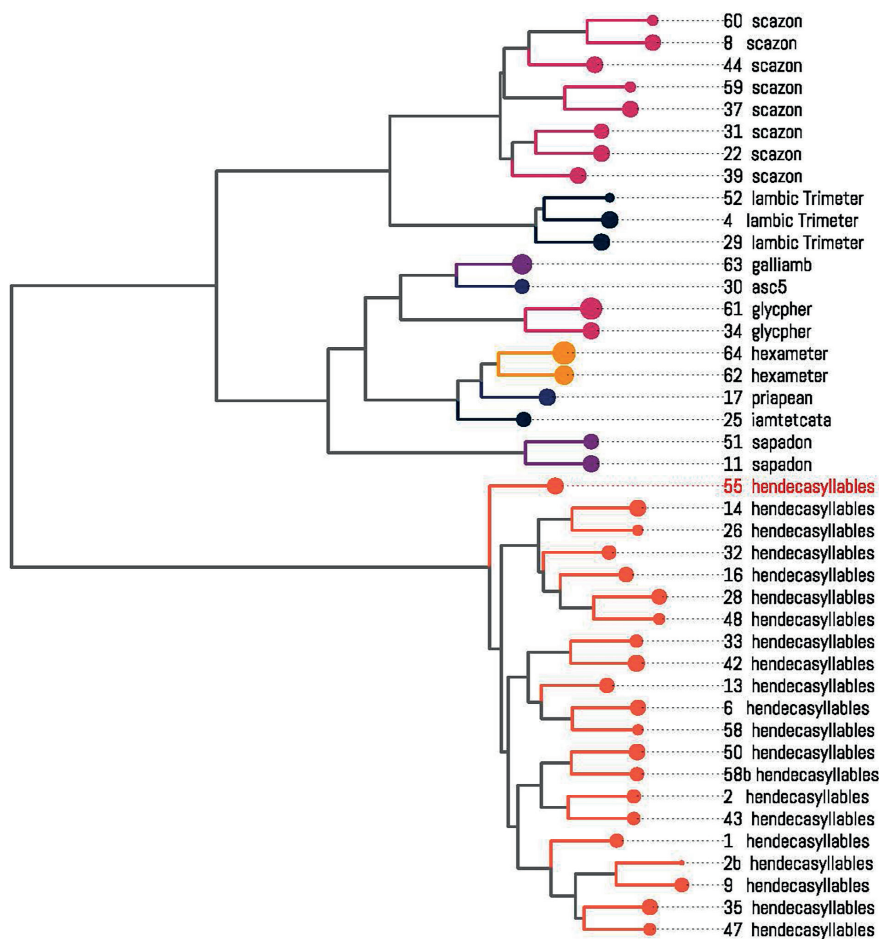
1. ábra. A Metronom osztályozási teljesítménye az SVM-mel és a másik két illesztésen alapuló módszerrel szemben. Pontossági értékek simított eloszlása 50 véletlenszerű mintavételezés után, medián pontszámokkal.

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*mel, 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 tanügyi 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 mindegyik 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-

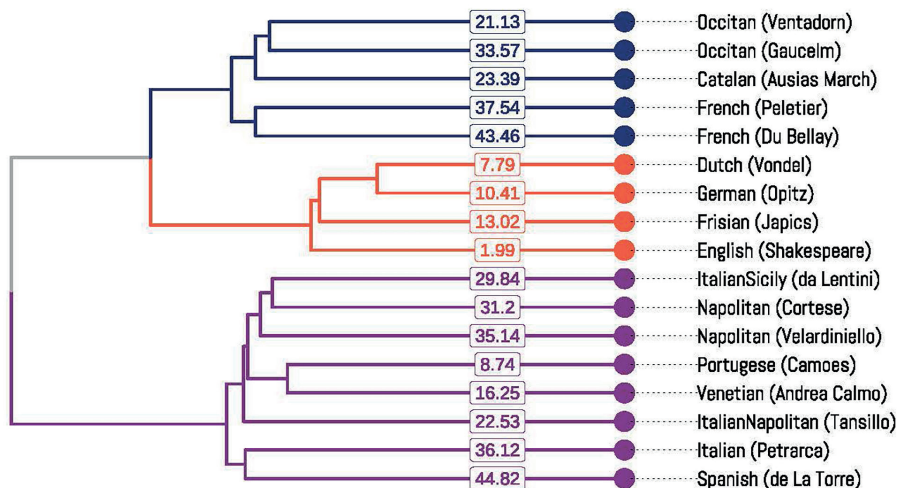
²² Mirella DE SISTO, „The interaction between phonology and metre: Approaches to Romance and West Germanic Renaissance metre” (Disszertáció, Radboud University, 2020).

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”.

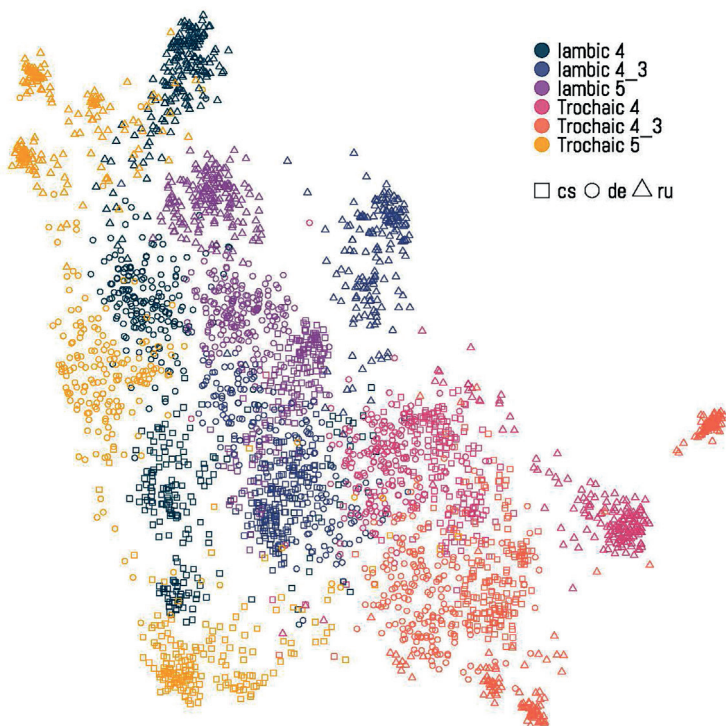
²⁹ GASPAROV, *A History of European Versification*, 207–209; Evgeny KAZARTSEV, „Iambic Verse in Different Literary Traditions”, *Glottottheory* 13, 1. sz. (2022): 3–19, <https://doi.org/10.1515/glot-2021-2004>.

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.

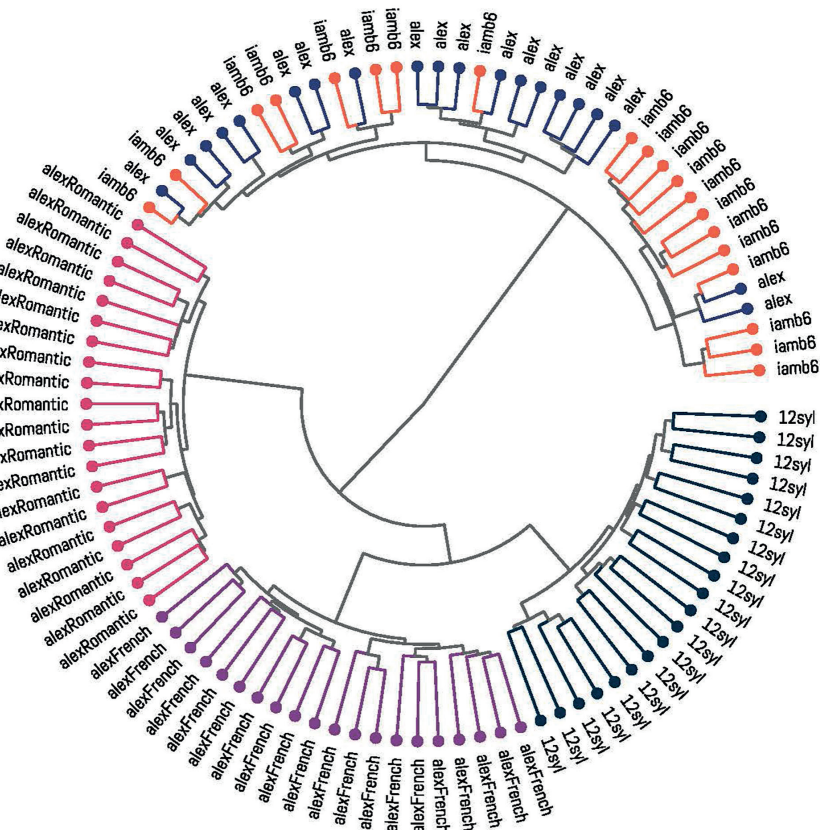
A generált versek hossza Poisson-eloszlást követ ($\lambda = 14$), így a mesterséges tesztünkben a versek átlagosan 14 sorosak, azaz szonettnyi hosszúak.

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