

Digitális eszközökkel támogatott verstani kutatások

Gortva Tamás¹, Kelemen Péter², Kelemen András³

¹tgortva@gmail.hu

SZTE BTK

²peter.balazs.kelemen@gmail.com

ELTE IK

³kelemen.andras.felix@szte.hu

SZTE JGYPK

Absztrakt. Az 1960-as évektől ismét a nyelv- és irodalomtudomány homlokterébe kerülő műfajkutatások hajtómotorja a kommunikáció matematikai leírásának megvalósulása [1], és az ezzel szinte egy időben zajló *pragmatikai fordulat* [2] voltak. Továbbá a digitális bölcsészet egyik alapvető elvárása a *forrásanyag feldolgozásához szükséges folyamatok automatizálhatósága, ami minőségében és látásmódját tekintve újszerűen kutatható adatbázist eredményez* [3]. Jelen projekt e kutatásokhoz kíván hozzájárulni a verstani elemzés és a stilometria számítógéppel automatizálható eszköztárának fejlesztése révén. Ezek a szoftverek végső soron az egyes szövegek objektív összehasonlíthatóságának egyik alapfeltételét teremtik meg, hosszú távon pedig hozzájárulhatnak a korpuszban lévő szövegek hálózatainak feltárásához, tehát a stilisztikai-retorikai (részben műfaji) sajátosságok leírásához.

Kulcsszavak: stilometria, verstan, digitális bölcsészet, szövegfeldolgozás, algoritmus

1. Bevezetés

Az 1960-as évektől ismét a nyelv- és irodalomtudomány homlokterébe kerülő műfajkutatások hajtómotorja a kommunikáció matematikai leírásának megvalósulása (Shannon – Weaver 1964), és az ezzel szinte egy időben zajló *pragmatikai fordulat* (Tátrai 2011: 279) voltak. Ugyancsak erre az időre tehető a *hálózatokutatás* megújulása is, melynek ma egyik – sokat vitatott, a bölcsészettudomány terén alkalmazhatónak látszó – elméleti kerete a skálafüggetlen hálózatok ábrázolása (Barabási 2002, Csermely 2005), mely az informatikai technológia bevonása nélkül elképzelhetetlen. Továbbá a digitális bölcsészet egyik alapvető elvárása a *forrásanyag feldolgozásához szükséges folyamatok automatizálhatósága, ami minőségében és látásmódját tekintve újszerűen kutatható adatbázist eredményez* (Sennyey 2018: 115–116).

Jelen projekt e kutatásokhoz kíván hozzájárulni a verstani elemzés és a stilometria számítógéppel automatizálható eszköztárának fejlesztése révén. A lírai nyelvhasználat felszíni struktúrája (parole; nyelvhasználat) és mélyszerkezete (langue; pragmatikai és szemantikai meghatározottsága) között mintegy átmenetet képez a verstan, amely egyszerre kíván megfelelni a nyelv diktálta szabályszerűségeknek, és a nyelvhasználó elvárásainak, kommunikációs szándékának. E jellemzők adatolható feltárása végső soron az egyes szövegek objektív összehasonlíthatóságának egyik alapfeltételét teremtik meg, hosszú távon pedig hozzájárulhatnak a korpuszban lévő szövegek hálózatainak feltárásához, tehát a stilisztikai-retorikai (részben műfaji) sajátosságok leírásához. Kutatásunk szövegtörzsét a gimnáziumi tananyag versanyaga szolgáltatja, hiszen e szövegek magas kanonizáltságuk és közismertségük által nagyban hozzájárulnak szövegfogalmunk és stilisztikai-retorikai érzékünk kialakulásához.

2. Időmértékes verselés

Az időmértékes verselés nem más, mint a hosszú és rövid szótagok – jobban mondva morák – szabályos ismétlődésén alapuló versrendszer. Egy szótagot rövidnek nevezünk akkor, ha benne a magánhangzó rövid és utána maximum egy mássalhangzó áll. Verstani jelölése: **U**. Egy szótag hosszú, ha a benne található magánhangzó hosszú, és/vagy rövid, bár utána több mint egy mássalhangzó áll. Verstani jele: **—**. (A sorvégi szótagok – annak függvényében, hogy a következő sor elején mássalhangzó vagy magánhangzó áll – lehetnek rövidek vagy hosszúak egyaránt.)

2.1. Felépítése

Az időmértékes verselés alapegysége a **mora**. Egy morának egy rövid szótag felel meg. A hosszú szótag két mora. A verslábak morákból épülnek fel. A **versláb**akat annak alapján különböztetjük meg, hogy hány morásak. A leggyakoribb verslábakat az 1. táblázat tartalmazza.

Név	Versláb
Jambus	U —
Trocheus	— U
Spondeus	— —
Daktilus	— U U
Anapesztus	U U —

1. táblázat: Verslábak

A verslábak szabályos ismétlődése (Szepes – Szerdahelyi (1981) nyomán) olyan állandósult metrikai képletté szilárdul, amely – az állandósult szókapcsolatokhoz hasonló módon – a legkülönbélebb sorokba, periódusokba épül be, s így mintegy a versek „metrikai közhelyévé” válik, tehát a versláb nál nagyobb, a periódusnál és a versornál kisebb metrikai egységgé válik. Ezt a kettő vagy több versláb alkotta állandósult kapcsolatot nevezzük **kolón**nak, amely önállóan is létrehozhat ugyan verssort (legalábbis sortöredéket), azonban két vagy több kolón, illetve sor állandósult kapcsolata olyan igen gyakori, széles körben használt metrikai képletet szilárdíthat meg egy adott versrendszerben, amely rendre ugyanabban a formában jelentkező **periódus**ként határozható meg, és amely akár verssorokban, akár azok egy szakaszában vagy azokon túlnyúlóan is megjelenhet. Felületesen szemlélve a **verssor** az az egység, amelyet a költő egy sorba ír, azonban a szövegek mélyebb struktúráját szemlélve könnyen belátható, hogy ez szükséges, de nem elégséges definíció. Minthogy a verssorok több esetben a prózai sorokhoz hasonlóan egymásba hajlanak (enjambement) vagy – mint a XVIII. század előtt sok esetben – folyószöveggé lejegyzett ritmikai egységek, így a verssor valójában nem ritmikai egység, a látási forma kialakítása révén csupán hozzájárul a ritmusérzet kialakulásához. Ugyancsak az alkotó egyéni leleménye és a versritmus által diktált elvárás közé helyezhető a **versszak** vagy **strófa** egysége, amely a vers olyan két vagy több sorból álló, önmagában zárt metrikai egysége, amely a szöveg nagyobb részében egészében szabályszerűen ismétlődik.

2.2. Sorok, strófák, versformák

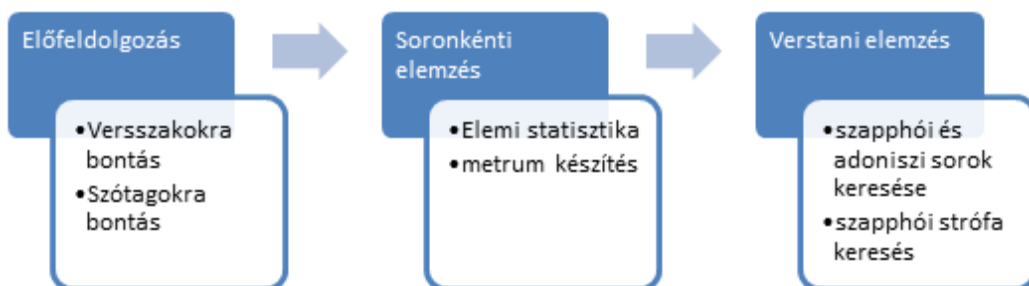
A *daktilikus* hexameter 23-24 moraértékű, hat verslábból álló, antik eredetű sorfajta, melynek első négy lába lehet daktilus vagy spondeus, az ötödik kötelezően daktilus, a hatodik trocheus vagy (a következő

sor kezdőbetűjétől is függően) spondeus. Ez a sor valójában felfogható két kolón állandósult kapcsolataként, melynek legjellemzőbb periódusa az utolsó két lábat alkotó daktilus és a hosszú vagy közömbös szótagot tartalmazó spondeus vagy trocheus: — U U / — U . Ez a kolón visszaköszön számos sorban, és önmagában **adoniszi sorként** ismeretes, többek között a szapphói strófa zárószakaként is azonosítható.

A szapphói strófa (hagyományos felfogás szerint) három tizenegy szótagos, jellemzően öt verslábból fölépülő, ún. **szapphói sorból** és egy adoniszi sorból áll. További megkötés a szapphói sor kapcsán, hogy az öt verslábból a középső (harmadik) mindig daktilus, a többi pedig jellemzően trocheus, mely(ek) spondeussal felcserélhetők.

3. Algoritmus

A következőkben az időmértékes verselemzés egy általunk létrehozott számítógépes algoritmusát és annak implementációját mutatjuk be. Az elemző bemenetként folyó szöveget vár. Ez lehet maga a teljes vers vagy annak részlete. Amennyiben a versszakokat az elemzőben is szeparáltan szeretnénk látni, akkor a bemenetben a versszakokat üres sorral kell egymástól elválasztani. Az elemző jelenlegi verziójában a következő műveleteket végzi (1. ábra):



1. ábra: Az elemző algoritmus folyamatábrája.

Részletezve:

1. az egyes verssorokat a magánhangzók mentén szótagokra bontja. A bontással egy időben kategorizálja a szótagot (hosszú, rövid, magas, mély hangrendű).
2. Minden egyes verssornál megszámolja rövid és a hosszú szótagok számát.
3. Meghatározza az adott verssor milyen verslábakból áll (metrum).
4. A metrum ismeretében eldönti, hogy a verssor a szapphói vagy adoniszi, vagy valami más.
5. Amennyiben a versszak 4 sorból áll, akkor a sorok metrumának ismeretében eldönti, hogy a versszak szapphói strófa-e.

3.1. Implementáció

Az elemző az internetes használhatóság miatt Java szervletként lett implementálva. A 2. ábrán a rezponzív felhasználói felület látható. A „Versválasztás” menüpont szolgál a versek feltöltésére a lokális gépről a szövegdobozba. Az „Elemi statisztika” menüponttal lehet az elemzést elindítani.



Vers

```
Partra szállottam, levonom vitorlám,  
A szelek mérgét nemesen kiálltam.  
Sok Charybdís közt, sokezer veszélyben  
Izzada orcám.
```

Statisztika

1. versszak

Hosszú szótagok: 7 Rövid szótagok: 4

P'artr'a sz'áll'ott'am, l'ev'on'om v'it'orl'ám,

- U - - - U U - U - - szapphai sor

2. ábra: Az elemző felhasználói felülete.

3.1.1. Adatábrázolás

A szoftver a verset versszakonként egy listában tárolja. A versszakokon belül listában tároljuk a verssorokat. Mivel soronkénti elemzést végzünk, minden egyes verssorhoz tároljuk a sor szótag szerinti felbontását egy szótag listában, valamint egy sztringben a metrumát.

3.1.2. Publikus metódusok

Az alábbiakban bemutatjuk az egyes osztályok publikus metódusait

VERS	
Metódus	Funkció
<code>void felbont()</code>	Felbontja a verset az üres sorok mentén versszakokra
<code>void elemez(PrintWriter out)</code>	Elvégzi az elemzést és az eredményt megjeleníti a weboldalon

VERSSZAK	
Metódus	Funkció
<code>void verszakToSzotag()</code>	Soronként szótagokra bontja a versszakot.
<code>void sorToSzotag(String txt)</code>	Magánhangzók mentén szótagokra bontja a verssort.
<code>String sorMetrum(String txt)</code>	Meghatározza egy verssor metrumát
<code>void Metrum()</code>	Soronként meghatározza a versszak metrumát.
<code>boolean isAdonisziSor(String sor)</code>	Eldönti egy sorról, hogy adoniszi-e
<code>boolean isSzapphoiSor(String sor)</code>	Eldönti, egy sorról, hogy szapphói-e
<code>boolean isSzappoiStrofa()</code>	Eldönti, hogy a verssor szapphói strofa-e
<code>int hosszuSzotag(String sor)</code>	Megadja, hogy hány hosszú szótag van a verssorban
<code>int rovidSzotag(String sor)</code>	Megadja, hogy hány rövid szótag van a verssorban

Következtetések

A természetes szövegek számítógépes felismerése a neurális hálókön alapuló gépi tanulás alkalmazásával az elmúlt pár évben rendkívül látványos fejlődésen ment keresztül. Ez a fejlődés megteremtette az irodalomtudományban és a nyelvészetben történő alkalmazás lehetőségét is. Egy szövegfelismerő rendszer gépi tanítása lényegében pozitív és negatív példákon hosszú iterációkkal történik. Ehhez nagyméretű pozitív és negatív korpuszra és vele együtt jelentős számítógépes kapacitásra van szükség. Meglátásunk szerint a verstanban ezt a hosszú és drága tanítási folyamatot algoritmikus eszközökkel drasztikusan le lehet rövidíteni. Ez egyben a felhasználó részéről sokkal kisebb számítástechnikai erőforrást igényel.

Irodalom

1. SHANNON, Claude E. – WEAVER, Warren (1964): *Theory of Communication*. The University of Illinois Press, Urbana
2. TÁTRAI Szilárd (2011): *Pragmatika*. In: BALÁZS Géza (szerk.) (2011) *Nyelvészetről mindenkinek, 77 nyelvészeti összefoglaló*. Inter, Bp.
3. SENNYEI, Pongrácz (2018): *Viták és vizsgák a digitális bölcsészetről*. In: *Digitális Bölcsészet 1* (szeptember), 111-120.
<https://doi.org/10.31400/dh-hun.2018.1.228>.
4. SZEPESE Erika – SZERDAHELYI István (1981): *Verstan*. Gondolat, Bp.