

[My Submissions](#)[NWS2021](#)[Conference](#)[News](#)[EasyChair](#)

NWS2021 Submission 43

Submission information updates are disabled.

For all questions related to processing your submission you should contact the conference organizers.

[Click here to see information about this conference.](#)

Submission 43

Title:	A magyar regény korpusza és a "távoli olvasás"
Author keywords:	stilometria regénykorpusz TEI XML distant reading távoli olvasás
Abstract:	Az ELTE Digitális Bölcsészeti Központban készülő regénykorpusz újszerű elérési lehetőséget kínál a „távoli” vagy számítógépes olvasás számára úgy, hogy a szabadon hozzáférhető magyar nyelvű regények annotált és kereshető adatbázisát szolgáltatja. A projekt szervesen illeszkedik az ELTeC (European Literary Text Collection https://www.distant-reading.net/eltec/) nemzetközi gyűjteményébe azáltal, hogy a korpusz magját alkotó első 100 regény ezen gyűjtemény magyar nyelvű alkorpuszát képezi. A szövegek feldolgozásának és tárolásának formátuma a TEI XML, amely lehetővé teszi a strukturált szövegfeldolgozást és megosztást az interneten. Az előadás részletesen beszámol a regénykorpusz kvantitatív elemzésének lehetőségeiről is konkrét esettanulmányokon keresztül. Ennek során azt kívánjuk bemutatni, hogy a statisztikai és nyelvészeti megközelítések együttes alkalmazásával milyen irodalomtörténeti belátások nyerhetők, illetve hogy a keresőfunkciók használata milyen szerepet tölthet be a szövegek értelmezésének folyamatában. A regénykorpusznak mint nagy mennyiségű adatnak az egyes szempontok mentén történő kvantitatív vizsgálata és az erre épülő vizualizáció ugyanis az egyes művek, vagy épp történeti korszakok olyan jellemzőit képes láthatóvá tenni, amelyek az olvasás tapasztalatában vagy reflektálatlanok maradnak, vagy a mérések hiányában nehezen igazolhatók. Az előadásban egyéb stilometriai jellemzők mellett a mondathosszúság, valamint az összetett mondatok fajtáinak relatív gyakoriságában megfigyelhető változásokat ismertetjük. A kirajzolódó tendenciák segítségével a magyar regényirodalom egy újszerű stílustörténete vázolható fel, míg az egyes szövegeknek a korpusz többi eleméhez viszonyított sajátosságai az adott szerzők és regényeik poétikájának leírásában tehetők jelentéssé.
Submitted:	Feb 03, 13:10 GMT
Last update:	Feb 03, 13:10 GMT
Angol nyelvű absztrakt	Distant Reading of the Corpus of Hungarian Novel The corpus of the Hungarian novel created at the ELTE Center of Digital Humanities offers new methodologies for the philological research and “distant reading” approaches by providing a digitized, annotated and searchable database of freely accessible novels from the Hungarian literary history. The database fits organically into the international collection of the ELTeC COST Action Project (European Literary Text Collection https://www.distant-reading.net/eltec/), since the first 100 novels of the database are part of the Hungarian sub-corpus of that collection. The TEI XML

format of the texts allows structured text analysis and the possibility of an online publication. Beside the description of the corpus and the aspects of the selection, the presentation also reports in detail on the possibilities of the quantitative analysis of the novels. In doing so, we want to present what kind of knowledge of the Hungarian literary history can be produced by applying statistical and linguistic approaches, and what role these methods can play in the process of the interpretation of the texts. The visualizations of the corpus handled as Big Data can make visible such features that are either difficult to verify in the absence of the statistical measurements or even remain unnoticed in the reading experience. In the presentation, we show the changes that can be observed in the average sentence length of the novels, as well as in the relative frequency of the different types of complex sentences. Through the visualized tendencies, a new history of the style of the Hungarian prose can be outlined, while the peculiarities of some texts in relation to others can lead to the description of the poetics of the given authors and their novels.

Telefonszám +36204200096

Authors						
first name	last name	email	country	affiliation	Web page	corresponding?
Tímea Borbála	Bajzát	bajzat.timi9696@gmail.com	Hungary	ELTE BTK TI Digitális Bölcsészet Tanszék		✓
Botond	Szemes	boboszemes@gmail.com	Hungary	ELTE BTK Irodalomtudományi Doktori Iskola		✓
Eszter	Szlávich	szlavicheszter@gmail.com	Hungary	ELTE BTK TI Digitális Bölcsészet Tanszék		✓