

**Delovni naslov: Oddaljeno branje slovenske pripovedne proze dolgega 19. stoletja v kontekstu družbeno-kulturnih procesov**

Kandidatka: Lucija Mandić

Mentorica: doc. dr. Andrejka Žejn

Somentorica: doc. dr. Monika Deželak Trojar

Komisija: izr. prof. dr. Tomaž Erjavec, red. prof. dr. Marko Juvan, prof. dr. Katja Mihurko

Doktorska disertacija razvije novo branje daljše slovenske pripovedne proze z uporabo računalniško podprtih metod za raziskovanje literature, ki so se v zadnjih desetletjih razvile v okviru digitalne humanistike. Osrednje raziskovalno vprašanje disertacije je, ali lahko uporaba računalniških metod, s katerimi lahko razširimo predmet raziskovanja s posameznih literarnih del, avtorjev in smeri na celotno produkcijo obravnavanega obdobja (v tem primeru t. i. dolgega 19. stoletja), prispeva k razumevanju procesa literarne kanonizacije na Slovenskem. Kanonizacijo, ki jo je tradicionalna literarna veda obravnavala predvsem z gledišča družbenega konteksta literarnih tekstov, same tekste pa je analizirala predvsem z metodo t. i. natančnega branja, ki tekste izolira drugega od drugega, bomo z uporabo računalniških metod obravnavali tudi na tekstualni ravni, kjer nam bo na voljo t. i. oddaljeno branje obsežnejšega korpusa tekstov.

Predmet disertacije je izvirna slovenska pripovedna proza, ki je nastala v nekaj manj kot stoletnem obdobju od njenih začetkov do konca prve svetovne vojne. Izbiro obdobja je narekoval razvojni lok slovenske pripovedne proze od prve slovenske izvirne pripovedi (1836), prek začetka in razvoja slovenskega romana in vzporednega pojava mohorjanskih povesti v drugi polovici 19. stoletja, do estetskega vrhunca v času moderne. Poleg teh literarnozgodovinskih premen pa izbrano obdobje zaznamujejo tudi pomembne družbene spremembe, kot so industrializacija, vstop žensk in delavstva v javno sfero ter razvoj slovenske nacionalne misli od njenih začetkov do razpada Avstro-Ogrske.

Za potrebe disertacije je bil sestavljen Korpus daljše slovenske pripovedne proze (KDSP), ki vsebuje vsa izvirna pripovedna prozna besedila v slovenščini, ki so izšla med letoma 1836 in 1918. Besedila v korpusu so opremljena z bibliografskimi in biografskimi metapodatki ter so dodatno kategorizirana glede na desetletje izida, dolžino, zvrst, tematiko in stopnjo kanoničnosti (z visoko stopnjo kanoničnosti so označena besedila tistih avtorjev in avtoric, ki

so vključena v šolske učbenike po letu 1980 in/ali v zbirko *Zbrana dela slovenskih pesnikov in pisateljev*).

Disertacija je razdeljena na tri vsebinske sklope. Prvi sklop začne niz metapodatkovnih analiz, katerih predmet so zvrsti in žanri v odnosu do ravni kanonizacije. Analiza metapodatkov o zvrstni in žanrski strukturi korpusa KDSP pokaže, da so iz literarnega kanona izločena tista literarna dela, ki ne ustrezajo sodobnim pojmovanjem literarnih zvrsti, nekanonična dela v korpusu pa so zvrstno in žanrsko bolj raznolika kakor kanonična. V nadaljevanju prvega sklopa uporaba metode modeliranja tem pokaže, da je kanon bolj homogen tudi na ravni notranje strukture literarnih tekstov. Modeliranje tem izpostavi Ivana Cankarja, najbolj kanoničnega avtorja celotnega korpusa, kot avtorja, čigar dela so tematsko najbolj homogena.

V drugem sklopu so s pomočjo stilometrije, uporabe konkordančnika in analize oblikoskladenjskih oznak kanonična in nekanonična dela primerjana z vidika rabe besednih vrst. Izkaže se, da je med besedami, ki izstopajo, v kanoničnih delih več predmetnopomenskih, v nekanoničnih pa več funkcijskih besed. Te rezultate dodatno osvetli analiza razmerja med pojavnicami in različnicami, ki pokaže, da imajo kanonična dela na ravni predmetnopomenskih besed manj raznoliko besedišče kakor nekanonična. V nadaljevanju sklopa je izvedena semantična analiza besedišča s pomočjo vektorskih vložitev besed. Jezikovni model pokaže na prekrivnost semantičnih polj politike in kulture predvsem v kanonični, a tudi v nekanonični literaturi, kar empirično potrди in razširi ugotovitve tradicionalne literarne vede o t. i. prešernovski strukturi slovenske književnosti obdobja, ki ga zajema KDSP.

Tretji sklop zastavlja vprašanje spola v kontekstu kanonizacije. Metapodatkovne analize razkrijejo marginalnost avtoric v slovenskem srednješolskem kanonu. Analiza imenskih entitet podobno pristopi k literarnim junakinjam, ki so tako v kanoničnem kakor v nekanoničnem delu korpusa zastopane v manjši meri kakor junaki. Dodatna analiza s pomočjo vektorskih vložitev besed osvetli še hierarhijo med spoloma, ko gre za način pripovedovanja o ženskah na eni strani in o moških na drugi.