

Kvantitatívna komparatistika v literárnovednom výskume

Marek Debnár

DEBNÁR, M.: Quantitative Comparatistics in Literary Studies
SLOVENSKÁ LITERATÚRA, vol. 68, 2021, no. 4, pp. 451-456

DOI: <https://doi.org/10.31577/slovlit.2020.68.4.6>

ORCID ID: 0000-0002-2879-6878

Key words: Richard Změlík,
Digital Humanities, quantitative
methods, distant reading, author
corpora, Jan Čep

The article focuses on distant reading approaches and the ways literary studies can use these for research of various semantic fields of authors' corpora (or subcorpora). It provides a closer look at the methodological facets of these approaches, evaluates their adequacy and the possibilities of implementation of chosen linguistic methods in literary studies. These aspects are reflected in polemics with methods and conclusions presented in Richard Změlík's *Konceptualizace barev v narativní fikci na pozadí kvantitativních modelů* ([Conceptualisations of colours in narrative fiction on the background of quantitative models] Olomouc: Univerzita Palackého, 2019, 320 pp.), in which the author employed the methodology of distant reading on the corpus of fiction written by the Czech writer Jan Čep (1902 – 1973). The paper aims at contributing to a wider discussion on quantitative methods of distant reading, their application in literary studies and the creation of authors' specialised and general reference representative literary corpora.

Kľúčové slová: Richard Změlík,
digitálne humanitné vedy,
kvantitatívne metódy, dištančné
čítanie, autorský korpus, Jan Čep

Digital humanities alebo digitálne humanitné vedy (ďalej DH) priťahujú v posledných rokoch stále viac pozornosti nielen v zahraničí, kde sú pomerne dobre etablované, čo vidieť v čoraz častejšej implementácii týchto postupov takmer vo všetkých humanitných disciplínach, ale aj v slovenskom akademickom priestore. Zatiaľ síce nemáme zavedené samostatné študijné odbory

452 DH na univerzitách ani národné DH centrum v Slovenskej akadémii vied, ako je to v ďalších európskych krajinách (napríklad v susednom Rakúsku jedno takéto centrum zamestnáva viac než sto odborníkov), ale trend implementovať do vedeckej práce aj digitálne prístupy silnie i na Slovensku. Dá sa povedať, že momentálne v tejto oblasti dobiehame zameškané, a je preto prirodzené, že naše napredovanie čiastočne kopíruje vývin, ktorý sme videli v zahraničí. Platí to aj v prípade literárnej vedy, ktorá má oproti jazykovede, kde sa začali kvantitatívne výpočtové metódy využívať už skôr, špeciálne postavenie. V porovnaní so skúmaním jazyka sa v nej totiž viac prejavuje potreba individuálneho prístupu k jedinečnému textovému materiálu a jeho originálnej interpretácii. Domnievam sa, že nejasnosť ohľadom miesta DH v literárnej vede nie je a nikdy nebola na strane dostupnosti digitálnych textových zdrojov, teda digitálnych verzií kníh, rukopisov, denníkov či korešpondencie jednotlivých autorov a podobne, ale v určení hraníc ich využitia. Viacúčelovo použiteľná digitálna zbierka literárnych textov, ktorá mimochodom na Slovensku zatiaľ absentuje (či už v reprezentatívnej podobe, ktorá by zastupovala slovenskú literatúru ako celok, alebo selektívnej, v podobe autorských korpusov), je niečo, čoho dôležitosť si s nemalou naliehavosťou uvedomujú všetci, ktorí prichádzajú do užšieho styku s výskumom literatúry. Rozličné očakávania vyvolávajú až predstavy o využití týchto dát, ktoré sa pohybujú od jednoduchého ilustračného dopĺňania klasických literárnovedných a literárnohistorických výskumov o grafy a tabuľky, až po autentické výskumy vychádzajúce z dát samotných, ktoré by umožňovali systematické uchopenie určitého literárneho fenoménu ako celku na novej metodologickej báze. Jedným z kľúčových dôvodov rozporu okolo miesta DH v literárnej vede na Slovensku (ale aj v Českej republike)¹ je nedostatok pôvodných vedeckých publikácií a metodologických opisov aplikovaných výskumov konkrétneho, digitálne spracovaného literárneho materiálu. Monografia Richarda Změlíka *Konceptualizace barev v narativní fikci na pozadí kvantitativních modelů* (Olomouc, 2019) je ojedinelým príkladom takejto publikácie, a preto si zaslúži pozornosť.

Hneď na začiatku je potrebné povedať, že ide najmä o metodologickú prácu. V názve deklarovaná konceptualizácia farieb v dielach českého spisovateľa Jana Čepa (1902 – 1973) na pozadí kvantitatívnych modelov je spojená so široko koncipovaným metodologickým a metateoretickým základom. Cieľom presunu ťažiska do oblasti teórie a metodológie digitálnych literárnych štúdií (*Digital Literary Studies*) je objasnenie významu a miesta dištančného čítania (*distant reading*) v procese komplexnej literárnej interpretácie. Autor monografie vychádza z toho, že si uvedomuje ambivalentnosť, ktorú tento pojem vyvoláva. Dištančné čítanie je v súčasnosti skôr zastrešujúci pojem, ktorý viac než jednu konkrétnu metódu postihuje celú škálu kvantitatívnych prístupov literárnovedných a literárnohistorických skúmaní. Nakoniec aj sám autor tohto pojmu, Franco Moretti, ho „definuje“ pomerne vágne, ako protipól k metóde podrobného čítania (*close reading*), ktoré vychádza len z textu samotného (Moretti 2014: 7). Žiada sa

1 Ako uvádza Richard Změlík v štúdiu o digitálnej literárnej kartografii v súvislosti s vymedzením predmetu výskumov Digital Humanities v literárnej vede (Digital Literary Studies) a počítačnej literárnej vedy (Computational Literary Studies), „v českém literárněvědném prostředí vzniklo jen pár takto orientovaných prací“ (Změlík 2020: 524).

doplniť, že to bol najmä F. Moretti, kto mal leví podiel na ambivalentnom vnímaní daného pojmu, pretože nie vždy zastával umiernené stanovisko vo vzťahu ku klasickej interpretácii a k podrobnému čítaniu. Svoju úlohu na formovaní diskurzu samozrejme zohráva aj to, že u pôvodcov istého typu diskurzivity sú výskumníci naučení opakovane čítať a interpretovať práve ich najradikálnejšie myšlienky, čo platí aj v Morettiho prípade.

R. Změlík nezačína svoju monografiu úvahami o podrobnom a dištančnom čítaní, túto tému necháva naplno vyznieť až v záverečnej kapitole knihy. S ohľadom na vlastné situovanie opiera svoje počiatkové úvahy a metodológiu o už etablovanú oblasť kognitívnej a korpusovej lingvistiky. V princípe však možno povedať, že ak by sa aj chcel v tejto časti nejako oprieť o Morettiho, nebolo by to veľmi možné, keďže technické detaily svojej práce si F. Moretti v zásade nechával pre seba. Monografia preto začína priblížením synchronných a diachronných lingvistických východísk ku konštituovaniu pojmov jednotlivých farieb, respektíve farebného spektra, pričom táto rozsiahla reflexia ďalej slúži na postupný prechod od opisu recepcie farieb k ich konceptualizácii a sémantike. Úvodná časť (druhá kapitola) je však aj ilustráciou hľadania miestami až empirickej exaktnosti, určitých výrazov (v danom prípade farieb, no rovnako by mohlo ísť o iné motivované výrazy) v oblasti lingvistiky, čím sa načrtáva paralela s ďalšími kapitolami knihy, ktoré sa zameriavajú na hľadanie exaktnosti a jej miery vo sfére fiktívnych svetov spadajúcich do oblasti literárnej vedy. Toto usúvzťažnenie tematiky farieb s hľadaním exaktnosti, ktorú R. Změlík ilustruje na príklade lingvistiky, sa ukazuje ako mimoriadne prínosné v nasledujúcich častiach, kde prechádza od lingvistiky k literárnej vede.

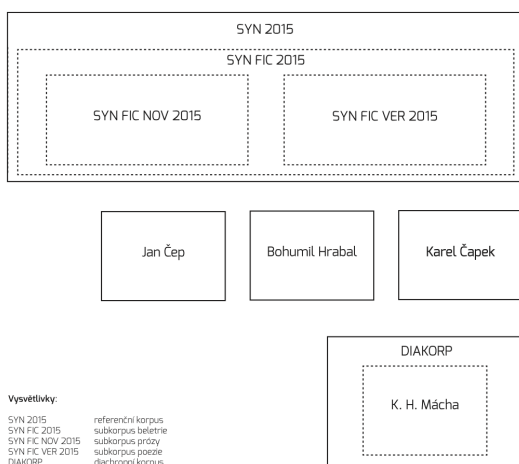
Načrtnutú teoretickú bázu autor ďalej rozvíja a usúvzťažňuje v tretej a vo štvrtej kapitole, a to na jednej strane s teóriou dištančného čítania, o ktorú sa opiera pri popise výstavby a analýzy kvantitatívnych modelov, pričom práve v tomto kontexte odkazuje na viaceré tvrdenia F. Morettiho, a na druhej strane so svojou vlastnou prácou s korpusmi. V týchto pasážach načrtáva prvé kontúry vzťahov medzi kvantitatívnou analýzou a interpretáciou:

„Základní smysl kvantitativní analýzy spočívá především v tom, že systematicky modeluje nejen makrostrukturní vlastnosti jevu, ale umožňuje mnohačetné komparace, které jsou závislé na badatelské strategii. Kvantitativní analýza sice sama o sobě není interpretací v literárněvědném smyslu, avšak literárněvědné zkoumání je schopna účelně metodologicky fundovat a determinovat. Nejedná se tedy o samoučelné doplňování literárněvědného výkladu o grafy, schémata či tabulky, ale především o úsilí po objektivizaci, systematizaci a distribuci makrostrukturních kvalit reprezentovaných kvantitativními ukazateli“ (Změlík 2019: 113-114).

Pri práci s autorskými korpusmi sa naplno prejavuje autorova erudícia v oblasti korpusovej lingvistiky, ktorú mimochodom prejavil už vo svojej o niečo staršej publikácii *Kvantitativně-korpusová analýza a literární věda: model a realizace autorského korpusu a slovníku Jana Čepa v kontextu zahraniční a české autorské lexikografie* (Olomouc, 2015). Je zřejmé, že najmä v oblasti literárnych skúmaní má mimoriadne postavenie autorský korpus, ktorý umožňuje skúmať lexikálne kontexty jednotlivých výrazov. Samozrejme, ako literárnemu vedcovi

454 je R. Změlíkovi zřejmé, že sa nemôže uspokojiť len s kvantifikáciou a jej modelmi, naopak, cieľom korpusového spracovania literárnych textov je vytvoriť (základnú či rozšírenejšiu) bázu kvantitatívnych dát a modelov, ktoré by bolo možné využiť v ďalšom literárnovednom výskume a interpretácii textu. Zjednodušene povedané, kvantitatívny model upozorňuje na isté symptomatické situácie v oblasti korpusu, ktoré vysvetlí literárnovedná interpretácia.

Pre zachovanie objektivity s cieľom demaskovať základný rámec konceptualizácie farieb druhého stupňa² v prózach J. Čepa však nestačí zostaviť či vybudovať len jeden autorský korpus (tu konkrétne Čepov). V tomto procese sa ako nesmierne dôležitá ukazuje komparácia medzi korpusmi a subkorpusmi, smerujúca k hlbšej explikácii problematiky. Stratégiu jednotlivých komparácií medzi viacerými korpusmi a subkorpusmi zachytáva nasledujúca schéma (Změlík 2019: 115):



Obr. 20: Schéma zobrazuje oblasti kvantitatívneho modelovania jevů. Korpus synchronních textů SYN 2015 zahrnuje subkorpus beletrie (SYN FIC 2015) a tento subkorpus prózy (SYN FIC NOV 2015) a poezie (SYN FIC VER 2015). Korpusy próz Jana Čepa, Bohumila Hrabala a Karla Čapka jsou samostatnými autorskými korpusy. Diakorp je diachronním korpusem, jehož součástí jsou i prózy Karla Hynka Máchy. Analýzu barev na úrovni celého Diakorpu neprovádíme, neboť s výjimkou Máchových prozaických textů nejsou ostatní texty lematizovány.

Kvantitatívno-korpusová analýza vychádzajúca z tejto schémy postupu je od vytvorenia a vzájomného porovnávania modelov frekvenčnej stratifikácie lexém základných farieb na úrovni jednotlivých korpusov a subkorpusov cez kontextovú analýzu lexém pomocou vybraných asociačných mier a výpisu konkordancií až k vytvoreniu takzvaných tematicko-motivických modelov. Dôležitým aspektom predkladanej analýzy je komparácia autorských subkorpusov J. Čepa, Bohumila Hrabala a Karla Čapka – práve na nej sa totiž ukazuje, že čím je diskurz špecifickejší, tým viac sa jeho model odchyľuje od všeobecnejšieho referenčného rámca (korpusy SYN 2015). V tejto súvislosti sa žiada dodať, že v budúcnosti literárnovedných skúmaníach bude využívanie autorských korpusov štandardom, ktorý bádatelom umožní identifikovať dosiaľ skryté vzťahy nie iba medzi jednotlivými

2 Pod „konceptualizáciou druhého stupňa“ (kapitola 3.2) má autor na mysli evokáciu fikčného, a nie obraz reálneho sveta.

autorskými korpusmi navzájom, ale aj na vyššej úrovni medzi rôznymi národnými literatúrami. Možno práve konceptualizácia farieb druhého stupňa je jednou z ciest, ako identifikovať skryté vzťahy medzi rôznymi literatúrami.

Piata kapitola, zameraná na samotnú fikčnú sémantiku farieb v diele J. Čepa, je vyvrcholením všetkých predchádzajúcich krokov. Autor v nej predstavuje takzvanú riadenú literárnu interpretáciu, čo znamená, že „naši výchozí interpretační mapou jsou kvantitatívni modely [...] tyto modely nemohou být pouze pasivně opisovány, ale musí být konfrontovány jednak se znalostmi vyvozenými ze zdrojových textů a jednak s vlastní literárněvědnou metodou, která v našem případě vychází ze strukturálně-sémiotických základů. Z kritické a především funkční kombinace těchto dvou hlavních metodologických aspektů se potom odvíjí i následující interpretace“ (Změlík 2019: 204-205). Změlíkove interpretácie farieb v diele J. Čepa sú presvedčivé, exaktné, analytické a tým, že súčasťou monografie sú tiež dáta extrahované z kvantitatívno-korpusovej analýzy, aj verifikovateľné. Vyčerpávajúci opis jednotlivých farieb palety Čepovho diela by na tomto mieste nemal veľký význam, v podstate to ani nie je žiaduce; každý takýto pokus by bol len sploštením a redukciou významu. Dôležitý je fakt, že táto kapitola je jednoznačným potvrdením prínosu postupov DH pre literárnovednú interpretáciu.

Knihá R. Změlíka úspešne ilustruje nielen prínos kvantitatívnych metód v literárnej vede, ale načrtáva aj cestu k porozumeniu ich miesta v súčasnom diskurze DH. Robí to neľahkým spôsobom, pretože sa nevyhýba páľčivým teoretickým a ani technickým otázkam a problémom. Z formálneho hľadiska je monografia spracovaná veľmi dobre, grafy, tabuľky, štatistiky a náčrty funkčne dopĺňajú a sprehľadňujú nastolené koncepcie a výsledky, ktoré je možné dohľadať a ďalej ich interpretovať. Za mierne problematické považujem len to, že autor neprekladal citáty zo zahraničnej odbornej literatúry (ide najmä o teoretické, úvodné časti práce). V angličtine to vzhľadom na jej rozšírené používanie v zásade neprekáža, ale v origináli zostávajú aj citáty v iných jazykoch, napríklad v poľštine, čo môže byť pre niekoho problém.

Monografia *Konceptualizace barev v narativní fikci na pozadí kvantitatívnych modelů* je vzácnym príkladom komplexnej práce DH vedca na poli literárnej vedy. V zahraničí sú takéto výstupy zvyčajne výsledkom viacčlenného riešiteľského tímu, v našich podmienkach je často nevyhnutné, aby si vedec sám zaobstaral textové zdroje, upravoval dostupné skripty, vypracoval metodológiu a podobne, takže okrem znalostí z vlastného vedeckého odboru sa vyžadujú aj schopnosti zo sféry informačných technológií, ktoré sú, žiaľ, väčšine humanitných vedcov cudzie. Nie je to jednoduchá práca a len málokedy je náležite pochopená a docenená. DH vedci to vedia, no aj napriek tomu (alebo práve preto) pokračujú vo svojich výskumoch ďalej, na čo je potrebná nemalá dávka nadšenia. Keby som mal potenciálnym čitateľkám a čitateľom tejto monografie – najmä tým, ktorí nikdy neprišli do styku s digitálnymi humanitnými vedami – odporučiť nejaký prístup, bola by to rada, aby skúsili danú knihu čítať práve s takýmto nadšením z nových možností, ktoré prinášajú zmeny perspektívy.

456 Literatúra

roč. 68, 2021, č. 4

- MORETTI, Franco, 2014. *Grafy, mapy, stromy. Abstraktní modely literární historie*. Přeložila Olga Čaplyginová. Praha: Karolinum. ISBN 978-80-246-2609-3.
- ZMĚLÍK, Richard, 2015. *Kvantitativně-korpusová analýza a literární věda: model a realizace autorského korpusu a slovníku Jana Čepa v kontextu zahraniční a české autorské lexikografie*. Olomouc: Univerzita Palackého. ISBN 978-80-244-4766-7.
- ZMĚLÍK, Richard, 2019. *Konceptualizace barev v narativní fikci na pozadí kvantitativních modelů*. Olomouc: Filozofická fakulta. ISBN 978-80-244-5479-5.
- ZMĚLÍK, Richard, 2020. Digitální literární kartografie a kvantifikace segmentů literárně narativního textu – metodologie, nástroje a interpretační strategie (popis jednoho projektu). *Slovenská literatúra*, roč. 67, č. 6, s. 523-551. ISSN 0037-6973.

Elektronické zdroje

- MORETTI, Franco, 2017. Patterns and Interpretation. Pamphlet 15. *Pamphlets of Stanford Literary LAB*. Dostupné online: <https://litlab.stanford.edu/LiteraryLabPamphlet15.pdf>

Mgr. Marek Debnár, PhD.
Centrum digitálnych humanitných vied
Filozofická fakulta UKF v Nitre
Hodžova 1
949 74 Nitra
Slovenská republika
E-mail: mdeb nar@ukf.sk