

Láska v číslach: Sládkovičova Marína pohľadom kvantitatívnej lingvistiky

Miroslav Zumrík

ZUMRÍK, M.: *Love in Numbers: Sládkovič's Marína from the viewpoint of quantitative linguistics*

SLOVENSKÁ LITERATÚRA, vol. 69, 2022, no. 1, pp. 67-78

DOI: <https://doi.org/10.31577/slovlit.2022.69.1.6>

ORCID ID: <https://orcid.org/0000-0001-5749-082X>

Key words: quantitative linguistics,
text analysis, methods of literary
research, interdisciplinary research

This review essay discusses the monograph *Love in Numbers: Sládkovič's Marína from the viewpoint of quantitative linguistics* authored by Michal Místecký, Natália Kolenčíková, Martin Navrátil and Gabriel Altmann (Bratislava: Veda, 2020). The book is a significant contribution to quantitative research of literary texts, a field currently viewed to a large extent as part of Digital Literary Studies. Using mathematical tools in the research of literary genres is not incoherent with traditional literary studies, on the contrary – it enables to create a fruitful dialogue between these approaches. This is also testified by the quantitative analysis of the key poem of Slovak Romanticism, Andrej Sládkovič's (1820 – 1872) *Marína* (1846). The essay discusses the results of the collective efforts employing quantitative methods to the research of Sládkovič's work. It concludes that the authors of *Love in Numbers* were able to put statistical methods and recent advancements in the field into a dialogue with philosophy of technology and that their publication benefited from the fruitful interdisciplinary character of their work.

Kľúčové slová: kvantitatívna
lingvistika, analýza textov,
metódy literárnovednej analýzy,
interdisciplinárny výskum

Názov kolektívnej monografie Michala Místeckého, Natálie Kolenčíkovej, Martina Navrátila a Gabriely Altmanna *Láska v číslech: Sládkovičova Marina pohľadom kvantitatívnej lingvistiky* (Bratislava: Veda, 2020) možno v niekom vyvolá otázku: Dá sa vlastne láska vtiesnať do čísel, možno ju spočítať, bezo zvyšku vyjadriť matematickými vzorcami? Nie je láska výsostným prejavom subjektivity, jedinečného a nevysvetliteľného zážitku? A nesmeruje preto názvom naznačený výskumný zámer – analyzovať Sládkovičovo dielo nástrojmi kvantitatívnej lingvistiky – priamo proti podstate lásky? Predkladá sa tu po vzore Schopenhauerovej metafyziky lásky jej matematika? A nejde takýto zámer aj proti výkladovým postupom literárnovedného myslenia a filológie? Očakávaniam v tomto duchu zdanlivo napovedá grafika obálky s portrétom Maríny zo znakov počítačového kódu. Názov monografie si však podľa mňa zámerne zachováva otvorenosť, súvisí s metódou abstrakcie a princípom *ceteris paribus* (všetko ostatné je rovnaké), ktoré patria k základným postupom vo vede, a tento názov možno využiť i ako nástroj pri interpretácii zámeru autorského kolektívu: ukázať na schodný, či skôr vhodný prienik kvantitatívnych metód, lingvistiky a tradičnej literárnej vedy a filológie.

Istú indíciu pri interpretácii poskytuje už edičná úprava knihy, konkrétne fakt, že monografia vychádza dvojjazyčne, v slovenčine a angličtine. Toto rozhodnutie možno čítať ako ambíciu osloviť čo najširšiu odbornú i menej odbornú čitateľskú obec na Slovensku a v zahraničí. Takému komunikačnému zámeru by zodpovedal aj čitateľsky ústretový rozsah monografie, čosi vyše sto strán na jednu jazykovú mutáciu, pričom podľa vlastných slov sa autori pre dvojjazyčnosť knihy rozhodli so zámerom vyvážiť jej vedecký a spomienkový charakter (strana 11). Ďalšiu interpretačnú indíciu vnímam vo fakte, že cez prizmu čísel, kvantitatívnej lingvistiky nazierajú autori na jeden z vrcholov slovenskej romantickej literatúry, ktorým Sládkovičova *Marína* nepochybne je. V úvode formulujú svoj zámer tak, že cítili potrebu overiť životnosť klasického diela stále rozvinutejšími analytickými nástrojmi (strana 10), no keďže z ich úvodných slov explicitne vyplýva aj úcta a obdiv k básnickej skladbe, možno azda zároveň povedať, že tiež overujú nosnosť spomenutých matematických a počítačových nástrojov kvantitatívnej lingvistiky na vybranom materiáli. Poslednou indíciou, ktorú by som chcel spomenúť a vo svojom úvode ju komentujú aj autori (strana 12), je interdisciplinárne zloženie kolektívu, ktorý tvorí lexikografka, literárny vedec a dvaja kvantitatívni lingvisti rôznej generačnej príslušnosti. Tento fakt vnímam ako overovanie a súčasne demonštráciu i nosnosti interdisciplinárneho prístupu, objavovanie vzájomných tematických synergií rôznych oblastí a metód.

Na základe uvedených predbežných pozorovaní sa ukazujú ako pravdepodobné, že názov monografie metaforicky vyjadruje nielen zámer skúmať konkrétne dielo matematickými metódami, ale aj presvedčenie, že komunikácia medzi rôzne navzájom vzdialenými oblasťami poznania, skúmania, myslenia a tvorby by sa mohla uskutočňovať v duchu dialógu, čo platí jednak pre celé skupiny vied, ako sú prírodné a humanitné vedy, ako aj pre konkrétne vedy, napríklad kvantitatívnu a tradičnú lingvistiku: „Základným cieľom kvantitatívni lingvistiky je poznání přirozeného jazyka z nejrůznějších hledisek. Užití matematických metod tedy v žádném případě neznamená odvrácení pozornosti od jazyka jiným směrem, rozdíl spočívá v cestě, kterou se k poznatkům badatelé dostávají“ (Kubát 2016: 14). Analogicky, podľa môjho názoru, ani interdisciplinárny prienik

kvantitatívnej lingvistiky a tradičných literárnych vied či filológie sa pri analýze *Maríny* neodvracia od literárneho diela.

Knihu tvorí okrem stručného úvodu a záveru šesť hlavných kapitol, zoznam literatúry a register, všetko v oboch jazykoch. Hlavné kapitoly najprv načrtávajú Sládkovičovu skladbu ako predmet literárnovedného výskumu, potom sa postupne venujú jednotlivým jazykovým rovinám a vlastnostiam básnického textu, na ktoré je možné zamerať kvantitatívny výskum: textovej vlastnosti označovanej ako slovné bohatstvo, rýmu, slovným druhom rýmových slov, syntaktickým vlastnostiam básne a napokon kvantitatívnym ukazovateľom súborne označovaným ako aktivita. Tu vidieť, že kvantitatívne metódy možno uplatniť prakticky na všetky jazykové roviny od fonetickej cez morfológickú, lexikálnu, sémantickú až po syntaktickú. Druhá a šiesta kapitola sú okrem toho príkladom postupu, keď sa intuitívne používané výrazy (slovné bohatstvo, aktivita) preberajú z bežnej reči, definujú sa a dodáva sa im presný význam. Takýto postup sa napríklad vo filozofii považuje za jednu z možných podôb filozofovania a označuje sa ako konceptuálna analýza (Zouhar 2016), teda analýza významu používaných výrazov, spresňovanie týchto významov sa zasa pomenúva ako explikácia.

Intuitívny pôvod a bežné používanie pôvodných výrazov slovné bohatstvo a aktivita sú známe: ak ostaneme na rovine čitateľskej fenomenológie, pravdepodobne vnímame, že je istý text bohatý na rôznorodé slová, čo vyjadrujeme napríklad tak, že má daný autor bohatý, bohatší či naopak chudobný alebo chudobnejší slovník než iný autor. Ak siahneme po klasikoch modernej literatúry, počet rôznych slov v základnom, slovníkovom tvare (takzvaných lem) v modernistických dielach ako *Ulysses* od Jamesa Joycea bude zrejme vysoký a vyšší než počet lem u rovnako modernistického Franza Kafku, a to aj vzhľadom na počet neobvyklých výrazov či idiolektických výrazov, ktoré Joyce používal. Samotná lexikálna bohatosť prirodzene ešte nič nehovorí o literárnej presvedčivosti a hodnote porovnávaných diel. Podobne to platí pre vlastnosť označovanú ako aktivita: aj tu intuitívne vnímame rozličnú prítomnosť dynamických príznakov v zmysle „akčnosti“, intenzity deja, ktoré sa v texte vyjadrujú prostredníctvom sloviac či prídavných mien. Na základe takejto čitateľskej, prirodzenej skúsenosti s textom nie je možné stanovovať a overovať tézy založené na kvantifikovateľných, intersubjektívne platných postupoch a overiteľných veličinách.

Na rozdiel od intuitívneho chápania a používania výrazov ako slovné bohatstvo či aktivita fungujú výrazy vo vede – v tomto prípade v kvantitatívnej lingvistike – ako termíny, čo znamená, že im je priradená definícia a ako metódy sa realizujú pomerne presne stanovenými matematickými nástrojmi. Okrem slovného bohatstva a aktivity sa v prácach korpusových a kvantitatívnych lingvistov definujú a skúmajú aj mnohé ďalšie vlastnosti vyjadrované i pomocou niekoľkých ukazovateľov súčasne. Takými vlastnosťami môžu byť prekladovosť textov (Chlumská 2017), teda fakt, že text potenciálne preberá isté črty východiskového jazyka; ich termínovosť (Kováříková 2017), to jest škálovitá prítomnosť viac-menej presne definovaných odborných výrazov s presne definovaným významom; reflexívnosť, respektíve „savosť“ (Garabík – Karčová – Žáková 2018), čiže textová prítomnosť sloviac so zvratným komponentom sa; alebo formulaickosť (Trklja – Grabowski 2021), čo znamená prítomnosť frekventovaných slovných sekvencií v bežnej reči i špeciálnych diskurzoch, napríklad v akademickej, vo farmaceutickej

70 alebo v právnej oblasti. Konkrétne matematické termíny patriace k oblasti slovného bohatstva a aktivity, ako vyplýva už zo zbežného oboznámenia sa s obsahom podkapitol v kolektívnej monografii, sa už na prvý pohľad vzdďalujú intuitívnosti a bežnej reči. Je teda potrebné vytvoriť spoločnú platformu kvantitatívnych a tradičných prístupov k textu, aby bol ich vzájomný dialóg produktívny.

Prvá kapitola práce predstavuje práve také východisko zasadené do literárnej vedy či klasickej literárnej filológie. *Marína* sa tu vníma ako svedectvo a zároveň výsledok básnikovho „zápasu o celistvosť“, čo možno chápať tak, že dielo sa dotýka metafyzických tém, predovšetkým času, večnosti, pomínutelnosti a trvania. V diele sa nastoľuje otázka, či a ako sa možno pomínutelnosti a zániku zmyslu brániť, ako docieľiť syntézu času a večnosti, a teda ako prostredníctvom umenia dosiahnuť osobnú i výrazovú celistvosť. Výsledkom v prípade *Maríny* nie je len už spomenuté vrcholné dielo slovenského literárneho romantizmu, ale aj istá pridaná hodnota: básnikovo objavovanie nových možností vyjadrenia. Autori vychádzajú z toho, že Sládkovič nachádza v básnení intímny priestor, ktorý na rozdiel od vonkajšieho okolia nepodlieha pomínutelnosti, a tak sa v podobe tvarovania diela azda môže stať zdrojom zmyslu a skrytej trvácnosti. Pomínutelnosť fyzického sveta možno napokon zasadiť aj do súvislosti s nástupom digitálnej reality v podobe médií, sietí a technológií, založených práve na číslach a znakoch, rovnakých, aké nachádzame v anonymnom kóde na obálke monografie. Môže to znamenať ďalšie spracovanie metafyzickej otázky vzťahu medzi pomínutelným a nepomínutelným, časom a večnosťou, zmyslom a absenciou zmyslu? Podľa mňa áno. V tejto súvislosti by som chcel preto odcitovať scénu z románu *Tridsaťtri* od súčasnej nórskej autorky mladšej generácie Kjersti Annesdatter Skomsvoldovej. Rozprávačka v texte uvažuje o svojich dvoch životných partneroch, jednom žijúcom a jednom už nežijúcom, prostredníctvom nasledujúceho obrazu: „*Keď som sa s ním na jednej konferencii tu v Írsku zoznámila, v myslí som si ho chcela premeniť na binárny kód. Ferdinand zomrel iba nedávno, a hoci som sa tej straty dlho bála, udiala sa rýchlo, Ferdinand jednoducho prestal dýchať, akoby sa vzniesol albatros, a odvtedy som sa pokúšala predstavovať si Samuela ako jednotky a nuly, ale nedarilo sa mi to. Nechcela som, aby bol príliš skutočný, pretože som sa bála, že stratím aj jeho*“ (Skomsvold 2020: 11). Digitálny, binárny kód ako symbol novej doby (a s ňou spojených nových, aj kvantitatívnych analytických metód) sa tak podľa mojej mienky načrtáva ako nová podoba večnosti, večného kolobehu informácií o našom živote, našich digitálnych odtlačkov v tom, čo iný nórsky autor, Thure Erik Lund, označuje v románe *Smutný a hlúpy príbeh* metaforou „*digitálne masívy*“ (Lund 2011: 200).

Kým prvá kapitola zasadzovala *Marínu* do kontextu literárnovedného uvažovania, druhá časť predstavuje východiská iného prístupu, východiskové matematické nástroje a veličiny na skúmanie textu *Maríny* z pohľadu kvantitatívnej lingvistiky. Dochádza tak ku skutočnému prieniku diskurzov: literárnovedného, založeného či už na štrukturalistických, alebo hermeneutických premisách a metodike, a kvantitatívneho, ktorý pristupuje k textu ako k objektívnej realite, ktorú je možné pomocou matematických metód skúmať z hľadiska jeho vnútornej štruktúry, prítomnosti a početnosti zložiek a ich vzájomných vzťahov. Metaforické „rozobratie“ diela na jeho štruktúrne zložky a kvantitatívne charakteristiky však neznamená koniec analýzy či rezignáciu na následnú kvantitatívnu

analýzu, ale je práve východiskovým bodom jej ďalšieho literárnovedného, napríklad komparatívneho výskumu. Z prístupov k textu, ktoré sa občas vnímajú ako konkurenčné (tradičný verzus moderný, kvalitatívny verzus kvantitatívny), sa tak stávajú momenty a súčasti celostného výskumného cyklu. Spomenutá predstava „rozloženia“ diela pritom predstavovala a sčasti naďalej predstavuje azda hlavný metodologický i myšlienkový kameň úrazu v očiach „klasickej“ literárnej vedy a filológie, ktorej námietka by sa dala pravdepodobne sformulovať tým spôsobom, že k sémantickému gestu diela neprenikneme jeho rozložením na súčiastky, štatistickým vyhodnocovaním najčastejších (nie nevyhnutne najvýznamnejších) slov ani skúmaním surového materiálu (v tomto prípade textu), prostredníctvom ktorého je dielo realizované, rovnako ako neprenikneme k podstate hudobnej skladby či hudby všeobecne tak, že budeme analyzovať mechanické vlastnosti alebo chemické zloženie napríklad dreva gitary. Tu by som však opäť rád zdôraznil, že ide o komplementárne prístupy k predmetu štúdia, pričom zástupcovia obidvoch výskumných línií si uvedomujú, že celok je – parafrázujúc podnetnú prácu Miroslava Svíteka o systémovom pohľade na proces ľudského poznania – čímsi viac než sumou častí (Svítek 2013).

Druhá kapitola z obsahového hľadiska predkladá východisko kvantitatívnej analýzy textu *Maríny*, ktorý tvorí 291 strof, najprv so zameraním na vlastnosť označovanú ako slovné bohatstvo. Na určenie slovného bohatstva sú potrebné východiskové výpočty kvantitatívnych charakteristík: ide o výpočet takzvaného h-bodu, výpočet indexu opakovania a výpočet veličiny (textová) entropia. H-bod, ktorý do textovej analýzy zaviedol rumunský fyzik a lingvista Ioan-Iovitz Popescu, sa napríklad určuje alebo vypočítava ako bod, v ktorom sa frekvencia nejakého výrazu (počet jeho výskytov v texte) rovná jeho rangu (jeho poradiu podľa početnosti výrazov v texte) (Popescu 2007). Význam určenia tohto bodu spočíva v tom, že viac-menej delí výrazy v texte na synsémantiká (funkčné slová bez samostatného významu) a autosémantiká (slová so samostatným významom). Toto vypočítané rozhranie možno použiť pri interpretácii výskytu výrazov, ktoré sa nachádzajú pod i nad bodom. Kým synsémantiká sa objavujú najmä nad bodom h (majú veľkú frekvenciu), autosémantiká spravidla tvoria slovnú zásobu pod bodom h (s nižšou frekvenciou). Z hľadiska literárnovedných kvantitatívnych analýz je pri tomto bode podstatné východisko, že ak sa nejaký autosémantický výraz (napríklad podstatné či prídavné meno) objaví medzi synsémantickými výrazmi nad bodom h, je možné pririeknuť mu funkciu pravdepodobne príznakového výrazu, napríklad literárneho motívu. Takýto postup rozdeľovania slovnej zásoby bol už využitý napríklad pri skúmaní a porovnávaní jednotlivých autorských žánrov u Karla Čapka, kde sa takisto pomocou bodu h identifikujú možné Čapkove motívy v jednotlivých útvaroch (Kubát 2016). H-bod sa tiež využíva ako nástroj na vymedzenie predmetu viacrozmerých lingvistických analýz: v práci Václava Cvrčka a kolektívu sa použil ako prah, nad ktorým sa v jednotlivých porovnávaných českých textoch porovnávala početnosť zaužívaných slovných sekvencií, teda vyššie spomenutá formulaickosť (Cvrček – Lauberová – Lukeš – Poukarová – Řehořková – Zasina 2020). V tomto zmysle slúži h-bod ako matematický ukazovateľ istého sémantického rozhrania, vnútornej štruktúry skúmaného textu, ktorý nie je stanovený úplne arbitrárne (napríklad stovky najčastejších slov), ale vyplýva z povahy skúmaného textu. Autori tiež definujú spomenutý index

72 opakovania (slov) a (textovú) entropiu, pričom opäť možno povedať, že tieto a ďalšie termíny či výrazy, ktoré v nejakom význame alebo iba vágne registrujeme vo vede (fyzike, informatike a iných), alebo aj v rámci bežného používania reči, získavajú z pohľadu kvantitatívnej lingvistiky presne určený význam. Navyše, nadväzujú na tradíciu matematizácie lingvistiky podľa Leonarda Bloomfielda (Dolník 2009: 12), skúmaný predmet sa charakterizuje prostredníctvom mier, čo znamená, že kvantifikovateľné vlastnosti textov sa určujú na škálach alebo stupniciach. Inými slovami, je pre ne príznačný do istej miery skôr kontinuálny prechod a výskyt „štýlových zmesí“ a „prechodných oblastí“ (Slančová 2009, citované podľa Imrichová 2014: 208) než jednoznačných binárnych (pojmových) opozícií. V kapitole sa potom uvádzajú ďalšie matematické nástroje a veličiny slúžiace na zistenie takzvanej frekvenčnej štruktúry textu (Kubát – Matlach – Čech 2014: 10-41), využiteľné na kvantifikáciu slovného bohatstva (dĺžka oblúka a lambda, Ordovo kritérium, Giniho koeficient), a v závere kapitoly sa prostredníctvom tabuliek prezentuje výpočet týchto vybraných ukazovateľov pre jednotlivé strofy *Maríny*. Z výpočtov je okrem samotných údajov zrejmé, že slovné bohatstvo sa takýmto spôsobom kvantifikuje prostredníctvom viacerých metrík, čo konvenuje so základnou intuíciou, že jazyk ako mimoriadne komplexný fenomén nemožno dostatočne opísať využitím jediného uhla pohľadu, preto sa naň uplatňujú viaceré, kombinované či komplementárne metriky a indexy. Takisto to znamená, že javy skúmané vedou sú azda najčastejšie viacrozmerné a na ich analýzu je preto s odkazom na prácu *Statistická analýza vícerozmerných dát* vhodné uplatňovať metódy viacrozmernej analýzy dát, keďže „príroda je vícerozmerná“ (Meloun – Militký – Hill 2017: 17). Takáto viacrozmerná (a konkrétne faktorová) analýza sa presvedčivo uplatnila v aktuálnej kolektívnej práci o registroch súčasnej češtiny (Cvrček – Lauberová – Lukeš – Poukarová – Řehořková – Zasina 2020). Čo sa týka záverov o miere slovného bohatstva v Sládkovičovej *Maríne*, autori konštatujú jej relatívne lexikálne bohatstvo a potrebu komparatívneho kvantitatívno-literárneho výskumu: „dôležitou úlohou do budúcnosti je jej komparácia s ďalšími slovenskými literárnymi dielami“ (strana 102).

Tretia kapitola nadväzuje na takpovediac tradičné využitie kvantitatívnych a počítačových metód, a síce ich aplikáciu v skúmaní verša. Skúmanie merateľných vlastností verša, eufónie alebo rýmu ako príznaku a štruktúry a zároveň výrazového prostriedku v poézii sa napríklad v českom prostredí v nadväznosti na výskumy Miroslava Červenku a Květy Sgallovej v posledných rokoch rozvíja pomerne intenzívne a medzičasom sa pretavilo do práce o určovaní autorstva básní (Plecháč 2021) či do kolektívnej monografie o korpusovej verzológii (Kolár – Plecháč 2017). Autori práce o *Maríne* sa v tretej kapitole postupne venujú dĺžke rýmového slova, rýmovým slovám v strofe, dĺžkovým úsekom, rozdielom medzi susediacimi slovami, pričom opäť uplatňujú aj špecifické kvantitatívno-lingvistické termíny, konkrétne Köhlerove dĺžkové motívy. Tieto nástroje im umožňujú v básnickom texte, ktorý čitateľ a interpret vnímajú zážitkovo-intelektuálne, objavovať isté formálne štruktúry, pravidelnosti či nepravidelnosti ukryté prirodzenému čítaniu a výkladu, ktoré sa môžu uplatniť pri následnej kvalitatívnej analýze a interpretačnom výkone. Autori už vyššie uvádzanej práce o viacrozmerných štatistických metódach hovoria v tejto súvislosti o hľadaní štruktúry v znakoch a dátach (Meloun – Militký – Hill 2017: 55), ktoré možno efektívne odhaliť práve

vo viacrozmerných dátach. Dalo by sa teda povedať, že takýmto spôsobom sa okrem čítanej knihy, respektíve básnickej skladby, odhaľuje prostriedkami kvantitatívnej lingvistiky ďalšia štruktúrna úroveň – javy ukryté prirodzenému oku, akási kniha v knihe, ktorá mohla v autorovom texte vzniknúť aj nezamýšľane, nielen na základe vedomého výberu výrazových prostriedkov. Znova treba zdôrazniť, že odhalením prípadnej takejto skrytej štruktúry sa interpretačný výkon v uvedenom prístupe nevyčerpáva, iba sa mu poskytuje pomerne exaktne získané východisko. V tretej kapitole tak autori pomocou kvantitatívno-lingvistických nástrojov dokázali „identifikovať vplyv jazykovotypologického charakteru slovenčiny či programového zakomponovania prvkov ľudovej slovesnosti do literárnej tvorby obdobia romantizmu“ (strana 102).

Štvrtá kapitola sa venuje rýmovým slovám z hľadiska ich potenciálne príznačnej slovnodruhovej príslušnosti. Aj v prípade slovných druhov a ich špecifického autorského používania sme si na základe čitateľskej introspekcie schopní všimnúť idiosynkrázie alebo afinity pri používaní slovných druhov u daného autora, prípadne historického literárneho obdobia. K tomuto však znova pristupuje kvantitatívny prístup s jeho ponukou oprieť svoje pozorovania a výklad skúmaných diel o exaktné, merateľné a intersubjektívne replikovateľné prvky. Okrem všeobecných východísk slovnodruhovej analýzy rýmových slov a problematiky vývoja autori skúmajú básnickú skladbu z hľadiska takzvanej tendencie uhlopriečky v texte, z hľadiska symetrie slovných druhov v rýme, a zaoberajú sa tiež poradím jednotlivých riadkov a stĺpcov slovných druhov. Zvláštnu pozornosť venujú kvantifikácii prítomnosti a charakteristike prísloviak a predložiek v *Maríne*. Zameranie na tieto slovné druhy azda prekvapí, pretože ich pri individuálnom čítaní a vykladaní poézie môžeme ľahko prehliadnuť, z kvantitatívneho hľadiska sú však pomerne ľahko spracovateľné a zákonitosti ich distribúcie a autorského používania sa môžu ukázať ako zdroj možnej interpretácie a špecifickosti daného autora. V tejto súvislosti by som chcel spomenúť poľského básnika z medzivojnového obdobia Bolesława Leśmiana (Leśmian 2000), ktorého literárnohistoricky ťažko zaraditeľné básne čerpajú časť svojej sugestívnosti práve z inovatívneho zaobchádzania s narúšaním prirodzených pádových väzieb v poľštine, čo vyvoláva ozvlášťujúci efekt a vyjadruje básnikovo videnie paradoxnej metafyziky sveta. Z hľadiska slovnodruhovej príslušnosti slov tvoriacich rým v *Maríne* autori konštatujú „signifikantnú tendenciu k rýmovaniu slov rovnakých slovnodruhových tried“ a špeciálne vo vzťahu k príslovkám a predložkám dospievajú k záveru, že „istým matematickým zákonitostiam zodpovedá aj distribúcia prísloviak a predložiek v básni, no ich modelovanie komplikuje sémantika, či skôr polysémia“ (strana 103). Vidieť tu, že ani aplikácia matematických metód neprináša jednoznačné, sémanticky uzatvorené zistenia a odpovede na kvalitatívne postavené otázky, ale práve umožňuje pripraviť pôdu pre pozorovanie jemných kvalitatívnych rozlíšení.

V piatej kapitole sa od jednotlivých slov a slovných druhov postupuje k vyššej, syntaktickej úrovni: skúmajú sa syntaktické kvantitatívne vlastnosti viet v *Maríne*, ako je počet viet v strofe, alebo vlastnosti označované ako centralizovateľnosť a lateralita vety. Znova ide o jazykové štruktúry a javy, ktoré si prinajmenšom v prípade počtu viet dokáže čitateľ vedome všimnúť alebo podvedome registrovať, pre formuláciu verifikovateľných či falzifikovateľných tvrdení o charakteristickom tvarovaní viet u vybraného autora je však potrebné disponovať

74 kvantifikovateľnými veličinami a mierami takýchto textových vlastností. Centralizovanosť a lateralita vety nemusia byť výsledkom úplne vedomého stvárnovania zo strany Sládkoviča, no umožňujú objektívne opisovať jednak zákonitosti jeho zaobchádzania s jazykom, jednak jeho autorský rukopis zo syntaktického hľadiska a na objektívnom základe ho porovnať so štýlom iných autorov. Z aspektu syntaktických príznakov *Maríny* analýza podľa autorov naznačuje prítomnosť troch celkov v básni: „úvodné časti sa vyznačujú paralelnými kompozíciami a presne štruktúrovanými vetami charakteristickými pre neoklasicizmus“, v strednej časti sa uplatňuje „emocionalita a expresivita básnického výrazu“ a v záverečnej časti autori nachádzajú „množstvo eliptických štruktúr“ (strana 103).

Posledná, šiesta kapitola sa zaoberá vlastnosťou, ktorá je zdanlivo iba prevzatá z bežného používania v reči: takzvanou aktivitou. Tá sa chápe ako „súbor kvantitatívnych ukazovateľov, ktoré sú založené na frekvenciách prídavných mien a sloviess v texte“ (strana 83). Autori v tejto súvislosti prezentujú takzvaný Busemannov kvocient, štatistickú signifikantnosť dát z *Maríny* podrobujú overovaniu v podobe chí-kvadrátového testu, pokračujú skúmaním vývojových trendov tohto javu a v závere vysvetľujú adjektívno-verbálne reťazce v jednotlivých strofách básne, a ponúkajú niekoľko interpretácií tohto štruktúrneho javu v básni (strana 83). Pripomínajú dôležitosť uplatňovania „viacerých metodických prístupov“ (strana 103), teda uplatnenie viacerých vedeckých nástrojov a vlastne i optík, čo znamená, že perspektívu zameranú na báseň ako celok treba dopĺňať perspektívou zameranou na vyššie spomenuté trojité delenie *Maríny*. Na základe takejto kombinovanej či komplementárnej perspektívy autori považujú distribúciu prídavných mien a sloviess v celku básne za rovnomernú, hoci ich distribúcia v jednotlivých častiach vykazuje „isté tendencie k prevahe sloviess v úvodných častiach básne a prevahe prídavných mien v závere“ (strana 103), a na základe tohto zistenia možno „prvú časť *Maríny* vnímať skôr ako príbehovú, druhú zase ako reflexívnu“ (strana 103). Tu zároveň vidieť, ako sa kvantitatívno-lingvistický prístup usiluje vyjadriť vlastnosti textu pomocou škály a súčasne objavuje skryté textové a potenciálne významové štruktúry v diele.

V čom teda na základe reprodukovateľného a komentovaného obsahu publikácie vidím jej hlavné prínosy? Predovšetkým v tom, že živo nadväzuje na nie úplne všeobecne známe, no mimoriadne záslužné analýzy, nástroje a všeobecne metodológiu kvantitatívneho skúmania jazyka a jazykových javov, tak ako tieto výskumy dlhé roky viedol, individuálne i kolektívne rozpracoval a inšpiroval jeden z najvýznamnejších kvantitatívnych lingvistov, profesor Gabriel Altmann (1931 – 2020). V tejto spojitosti by som chcel spomenúť na prvý pohľad (vzhľadom na množstvo matematických zápisov) ťažšie prístupnú slovenskú kolektívnu monografiu *Úvod do analýzy textov* (Wimmer – Altmann – Hřebíček – Ondrejovič – Wimmerová 2003), kde sa kolektívu, ktorého členom bol aj G. Altmann, syntetickým spôsobom podarilo prezentovať celú škálu rôznych kvantitatívnych metód, aplikovateľných na jednotlivé jazykové úrovne a textové žánre. V kolektívnej práci G. Wimmera a ostatných sa tak ponúka celá konceptuálna batéria veličín, indexov, postupov a metrick na kvantifikáciu javov na viacerých úrovniach: na úrovni fóniky (kde možno skúmať eufóniu, asonanciu, aliteráciu, agregáciu, rým); na úrovni slova (kde sa dá skúmať dĺžka slova, rozdelenie slov v texte, rangové rozdelenie, frekvenčné rozdelenie, slovné bohatstvo textu, frekvenčný aspekt, bodové

charakterizácie, priebehové charakterizácie); na úrovni vety (tu možno skúmať dĺžku vety cez prizmu takzvaného Menzerathovho-Altmanovho, prípadne mocninového zákona). Okrem týchto rovín kvantitatívne metódy umožňujú skúmanie textov aj v podobe takzvanej obsahovej analýzy (ktorá sa zameriava na meranie akčného kvocientu textu, odhaľovanie asociácií, stanovovanie takzvanej čitateľnosti textu, čo je opäť termín iba zdanlivo prevzatý z bežnej reči). Alebo sa dá pri výskume textu pohybovať na rovine denotačnej analýzy (kde sa operuje s pojmi ako koincencia, súvislosť, stupeň vrcholu, vzdialenosti v grafe). Okrem týchto matematicky definovaných vlastností možno postúpiť takisto k ďalším, čitateľsky menej vnímaným alebo vôbec nevnímaným jazykovým rovinám, na úroveň takzvaných nadvetných a lineárnych štruktúr. V prípade nadvetných štruktúr sa uvádzajú pojmy ako slovné asociácie a hreby, text je vnímaný ako dynamický systém obsahujúci jazykové štruktúry a fraktály. Jazykové úrovne sa zasa skúmajú ako istý rad, v rámci ktorého existuje fázový prechod k sémantickej úrovni textu a uplatňuje sa princíp emergencie jazykových javov. Čo sa týka lineárnych štruktúr, tieto sa skúmajú prostredníctvom ukazovateľov ako dĺžka vety, Ljapunov koeficient či Hurstov indikátor. Lineárne štruktúry sa pozorujú vo vzťahu k radu slovných frekvencií, k intervalom slov tvoriacich základ hrebu, z hľadiska rôznych modelov zhlukov a intervalov, ako aj z hľadiska takzvanej koreferencie a pozičnej analýzy. Tieto miery a postupy možno potom po zohľadnení konkrétnych dát a špecifik výskumnej otázky využiť na porovnávanie a prípadnú vizualizáciu textov: to sa deje pomocou informačnej štatistiky a rozdielneho rozdelenia kriviek údajov v prípade jednotlivých textov či skupín textov. Z uvedených rovín a metrik je po zohľadnení limit a vhodnosti použitia jednotlivých testov a štatistik možné zostaviť čosi ako kvantitatívny profil skúmaného diela alebo textov – tak je to i v monografii venovanej *Marine*. Podobnú množinu charakteristík však uplatňuje aj Miroslav Kubát v kvantitatívnej, respektíve štylometrickej analýze žánrov u K. Čapka (Kubát 2016), kde okrem slovného bohatstva ide o takzvanú tematickú koncentráciu, jej sekundárnu i proporciálnu verziu, vzdialenosti slovies, priemernú dĺžku jazykovej jednotky (tokenu), podiel aktivity a deskriptivity či viacznakové alebo viacslovné útvary a sekvencie (n-gramy), alebo ide o analýzu najfrekvencovanejších slov.

Zaujímavým a možno až literárnym motívom je podľa mňa fakt, že G. Altmann po sebe zanechal celý rad (kolektívnych) prác zameraných takpovediac na virtuálne, budúce ciele, v ktorých vytýčil i zaujímavé budúce otázky, akými by sa kvantitatívna lingvistika pravdepodobne aj v interdisciplinárnom dialógu s inými oblasťami poznania mohla zaoberať (Altmann – Strauss – Fan 2008, Altman – Köhler 2009, 2014, Altmann – Čech 2011, Altmann 2015, Altmann – Kelih 2018). Tento fakt vnímam ako potvrdenie istej devízy vedeckého poznávania a metodiky, ktorou je možnosť replikácie experimentov a záverov výskumu do istej miery bez ohľadu na prístup individuálneho vedca. Skúmaný materiál v jeho fenomenálnych prejavoch a systematických súvislostiach sa nevyčerpáva aplikovaním merania, skúma sa však overiteľnou, objektívnou a opakovateľnou metódou, ktorá umožňuje navrhnúť, uskutočniť a zhodnotiť experiment v prostredí vedeckého diskurzu (Wimmer – Altmann – Hřebíček – Ondrejovič – Wimmerová 2003: 13-18). To znamená akýsi kontrast s romantickou koncepciou individuálneho tvorcu-génia (alebo i vykladača), ktorého vystúpenie a dielo sú nenapodobiteľné, nevypočítateľné

76 a neopakovateľné: „Kvantitatívni lingvisté dochádzajú k záverom pouze na základe experimentu, čo umožňuje intersubjektívny pohľad na zkoumanou problematiku. Základní výhodou takového přístupu je to, že kdokoliv daný experiment zopakuje, získá stejné výsledky. Pokud bychom však zadali kvalitativní analýzu jednoho textu několika lidem, jen stěží budou výsledky totožné. Kvantitatívni přístup nám umožňuje ověřovat naše intuitivní a do značné míry subjektivní předpoklady“ (Kubát 2016: 14-15). Uvedená charakteristika vedeckej metodiky zároveň pripomína, že vedecký diskurz pracuje s objektívnymi entitami, nie individuálnymi vecami, a teda že vedecké poznanie je, slovami Františka Mika, zamerané na to, čo je opakovateľné (Miko 1970: 40).

Objavovanie skrytých štruktúr v skúmanom texte by som napokon chcel dať do súvislosti s podstatou, respektíve takzvanou „bytnosťou“ vedy, ako ju vysvetľuje Martin Heidegger v prednáške *Věk obrazu světa*, kde hovorí: „Věda se stává výzkumem, když při rozvrhování a zajišťování svého předmětného okrsku postupuje s [...] přísností [...]. Má-li být rozvržený okrsek okrskem předmětů, pak je třeba nechat jej vystoupit v celé jeho vrstevnatosti a spletnosti“ (Heidegger 2013: 13). Matematické nástroje a metódy, aké uplatňuje autorský tím pri analýze Sládkovičovej *Maríny*, potom možno považovať za práve také postupy, ako s vedeckou prísnosťou nechať vyvstať text diela (ako predmet, respektíve „okrsok predmetov“) v celej jeho vrstevnatosti. Tá je, ako vidieť napríklad pri skúmaní nadvetných a lineárnych štruktúr, prirodzenému čítaniu prirodzene skrytá. Znamená to, že ak ich nemožno odhalíť „voľným okom“ a zrejme si ich nevedomoval autor, rovnako ako ich nevníma zväčša ani čitateľ, nemá zmysel sa takýmito štruktúrami a vrstvami diela zaoberať? Myslím si, že má, a to najmä v situácii, keď sa v literárnovednom diskurze uplatňuje na jednej strane až nehatená diverzifikácia prístupov, postojov a interpretácií, ktoré sa niekedy ako keby navzájom míňali, uzatvárali do mikropriestoru vlastnej perspektívy, a na druhej strane nemožno prehliadnúť prítomnosť stmelujúcich, zastrešujúcich prístupov k interpretácii literárneho diela, najmä z hľadiska jeho diskurzívneho a ideologického kontextu. Kvantitatívny, empirický prístup k textu literárneho diela podľa mňa poskytuje vhodné východisko akejkoľvek, právom rôznorodej a inovatívnej diskusii o možných významových vrstvách v diele. Odhalenie spomenutých ukrytých vrstiev textu teda literárnovednú analýzu diela neukončuje, ale naopak otvára a potenciálne rozširuje možnosti interpretácie diel, autorov, štýlov, žánrov a historických literárnych období. Matematické, počítačové, korpusové či kvantitatívne metódy, postupy a zdroje, ktoré sú napríklad v prípade korpusovej lingvistiky výrazom jej „nekompromisného empirizmu“ (Jarošová 2018: 298), presvedčivým spôsobom napomáhajú odkryť ďalšie významotvorné vrstvy či „záhyby“ skutočnosti (Deleuze 2015), ktorá sa ukazuje ako pôsobivá, hoci potenciálne ochromujúca „spleť“ interpretačných podnetov (Nitsche 2016). V literárnovednom diskurze sa medzičasom vytvorila a stále sa rozširuje oblasť takzvaných digitálnych literárnych štúdií (*Digital Literary Studies*), ktorá nielen početnými analytickými výstupmi, ale aj metodologickou sebareflexiou preukazuje, že ide o schodnú cestu v uvažovaní o literatúre (Siemens – Schreibman 2008). Takto sa napríklad reflektujú otázky, akým spôsobom je možné konceptuálne uchopiť a prakticky realizovať kombinovanie povedzme klasických metód v humanitných vedách s počítačovo podporovanou analýzou textu (Kuhn 2019) alebo ako sa prostriedkami matematiky,

konkrétne prostredníctvom takzvaného vektorového priestoru, skúma literárnosť románov ako ďalšia z kvantifikovateľných vlastností literárneho textu (van Cranenburgh – van Dalen–Oskam – van Zundert 2019). Preto treba uvítať, že k tejto ceste sa v podobe viacrozmernej kvantitatívnej analýzy Sládkovičovej *Maríny* v monografickej podobe pripája aj česko-slovenská literárna veda a jazykoveda.

Táto práca bola podporená Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-20-0414. Projekt *Digitálna zbierka slovenskej prózy*. Zodpovedný riešiteľ: Mgr. Marek Debnár, PhD. Doba riešenia: 2021 – 2025.

Pramene

- LEŠMIAN, Boleslaw, 2000. *Lúka*. Vybral a preložil Juraj Andričík. Spolupráca pri preklade básní Pavol Winczer. Bratislava: Studňa. ISBN 80-967534-8-7.
- SKOMSVOLD, Kersti Annesdatter, 2020. *Tridsaťtri*. Preložil Miroslav Zumřík. Bratislava: Premedia. ISBN 9788081596599.

Literatúra

- ALTMANN, Gabriel – STRAUSS, Udo – FAN, Fengxiang, 2008. *Problems in Quantitative Linguistics 1*. Lüdenscheid: RAM. ISBN 978-3-9802659-4-2.
- ALTMANN, Gabriel – KÖHLER, Reinhard, 2009. *Problems in Quantitative Linguistics 2*. Lüdenscheid: RAM. ISBN 978-3-9802659-5-9.
- ALTMANN, Gabriel – ČECH, Radek, 2011. *Problems in Quantitative Linguistics 3*. Lüdenscheid: RAM. ISBN 978-3-942303-08-8.
- ALTMANN, Gabriel – KÖHLER, Reinhard, 2014. *Problems in Quantitative Linguistics 4*. Lüdenscheid: RAM. ISBN 978-3-942303-22-4.
- ALTMANN, Gabriel, 2015. *Problems in Quantitative Linguistics 5*. Lüdenscheid: RAM. ISBN 978-3-942303-33-0.
- ALTMANN, Gabriel – KELIH, Emmerich, 2018. *Problems in Quantitative Linguistics 6*. Lüdenscheid: RAM. ISBN: 978-3-942303-57-6.
- van CRANENBURGH, Andreas – van DALEN-OSKAM, Karina – van ZUNDEBT, Joris, 2019. Vector space explorations of literary language. *Language Resources and Evaluation*, roč. 53, č. 4, s. 625-650. ISSN: 1574-020X.
- CVRČEK, Václav – LAUBEOVÁ, Zuzana – LUKEŠ, David – POUKAROVÁ, Petra – ŘEHOŘKOVÁ, Anna – ZASINA, Adrian Jan, 2020. *Registry v češtině. Studie z korpusové lingvistiky 27*. Praha: Nakladatelství Lidové noviny. ISBN 978-80-7422-754-7.
- DELEUZE, Gilles, 2015. *Záhyb*. Praha: Hermann & synové. ISBN 9788087054383.
- DOLNÍK, Juraj, 2009. *Všeobecná jazykoveda (Opis a vysvetľovanie jazyka)*. Bratislava: Veda. ISBN 978-80-224-1078-6.
- GARABÍK, Radovan – KARČOVÁ, Agáta – ŽÁKOVÁ, Adriána, 2018. Kvantitatívna analýza polyfunkčného komponentu na báze Slovenského národného korpusu. *Slovenská reč*, roč. 83, č. 1, s. 23-34. ISSN 0037-6981.
- HEIDEGGER, Martin, 2013. *Věk obrazu světa*. Praha: Oikoymenh. ISBN 978-80-7298-490-9.
- CHLUMSKÁ, Lucie, 2017. *Překladová čeština a její charakteristiky. Studie z korpusové lingvistiky 26*. Praha: Nakladatelství Lidové noviny. ISBN 978-80-7422-621-2.
- IMRICHOVÁ, Mária, 2014. Jazyk právnych textov so zreteľom na špecifiká právnej lexiky. In KESELOVÁ, Jana – IMRICHOVÁ, Mária – OLOŠTIAK, Martin, ed. *Registre jazyka a jazykovedy (I). Na počesť Daniely Slančovej*. Prešov: Filozofická fakulta Prešovskej univerzity, s. 207-212. ISBN 978-80-555-1111-5, 978-80-555-1161-0 (CD-ROM).
- JAROŠOVÁ, Alexandra, 2018. Pohľad na pomenovanie cez prizmu teoretických rámcov a slovníkového hesla. *Jazykovedný časopis*, roč. 69, č. 3, s. 277-301. ISSN 0021-5597.
- KOVÁŘÍKOVÁ, Dominika, 2017. *Kvantitatívni charakteristiky termínů*. Praha: Nakladatelství Lidové noviny. ISBN 978-80-7422-561-1.
- KUBÁT, Miroslav – MATLACH, Vladimír – ČECH, Radek, 2014. *QUITA – Quantitative Index Text Analyzer*. Lüdenscheid: RAM. ISBN 978-3-942303-28-6.
- KUBÁT, Miroslav, 2016. *Kvantitatívni analýza žánrů*. Ostrava: Ostravská univerzita. ISBN 9788074648687.

- KUHN, Jonas, 2019. Computational text analysis within the Humanities: How to combine working practices from the contributing fields? *Language Resources & Evaluation*, roč. 53, s. 565-602. ISSN 1574-0218.
- LUND, Thure Erik, 2011. *En trist og dum historie*. Oslo: Aschehoug. ISBN 9788203350825.
- MELOUN, Milan – MILITKÝ, Jiří – HILL, Martin, 2017. *Statistická analýza vícerozměrných dat v příkladech*. Praha: Academia. ISBN 9788024636184.
- MIKO, František, 1970. *Text a štýl. K problematike literárnej komunikácie*. Bratislava: Smena.
- NITSCHKE, Martin, 2016. *Metodická přednost spleti. Transzitivně-topologický model fenomenologie*. Praha: Togga. ISBN 9788074761126.
- KOLÁR, Robert – PLECHÁČ, Petr, 2021. *Kapitoly z korpusové versologie*. Praha: Akropolis. ISBN 9788074701498.
- PLECHÁČ, Petr, 2021. *Versification and Authorship Attribution*. Praha: Karolinum. ISBN 9788024648712.
- POPESCU, Ioan-Iovitz et al., 2007. Text ranking by the weight of highly frequent words. In GRZYBEK, Peter – KÖHLER, Reinhard, ed. *Exact Methods in the Study of Language and Text*. Berlin – New York: Mouton de Gruyter, s. 555-566. ISBN 9783110193541.
- SLANČOVÁ, Daniela, 1999. *Reč autority a lásky: reč učitelky materské školy orientovaná na dieťa. Opis registra*. Prešov: Filozofická fakulta Prešovskej univerzity. ISBN 8088722306.
- SVÍTEK, Miroslav, 2013. *Víc než součet částí: Systémový pohled na proces lidského poznání*. Praha: Academia. ISBN 9788020022868.
- WIMMER, Gejza – ALTMANN, Gabriel – HŘEBÍČEK, Luděk – ONDREJOVIČ, Slavomír – WIMMEROVÁ, Soňa, 2003. *Úvod do analýzy textov*. Bratislava: Veda. ISBN: 80-224-0756-9.
- ZOUHAR, Marián, 2016. Konceptuálna analýza v analytickej filozofii. *Filozofia*, roč. 71, č. 5, s. 410-424. ISSN 0046-385X.

Elektronické zdroje

- SIEMENS, Ray – SCHREIBMAN, Susan, ed., 2008. *A Companion to Digital Literary Studies*. Oxford: Blackwell. Dostupné online: <http://www.digitalhumanities.org/companionDLS/>
- TRKLJA, Aleksandar – GRABOWSKI, Łukasz, ed., 2021. *Formulaic language: Theories and methods (Phraseology and Multiword Expressions 5)*. Berlin: Language Science Press. Dostupné online: <http://langsci-press.org/catalog/book/304>

Mgr. Miroslav Zumrik, PhD.
Jazykovedný ústav
Ludovíta Štúra SAV, v. v. i.
Panská 26
811 01 Bratislava
Slovenská republika
E-mail: miroslav.zumrik@gmail.com