



Ústav Českého národního korpusu

**Posudek na habilitační spis Richarda Změlíka  
Konceptualizace barev v narativní fikci na pozadí kvantitativních modelů**

Předložená habilitační práce Richarda Změlíka „Konceptualizace barev v narativní fikci na pozadí kvantitativních modelů“ vyšla jako monografie v olomouckém univerzitním nakladatelství v roce 2019 a navazuje na autorovy dlouhodobé badatelské zájmy v několika směrech. Jednak je to soustavná pozornost, kterou R. Změlík věnuje prózám Jana Čepa (viz jeho předchozí monografie z roku 2015 a zejména korpus sestavený z Čepových textů), a jednak je to metodologická pozice, v níž úspěšně kombinuje metody literární vědy a korpusové a kvantitativní lingvistiky. V tomto smyslu je předložená práce završením několikaletých badatelských snah, což se odráží i v tom, že dílčí části práce byly publikovány jako samostatné studie.

Předložená práce materiálově vychází z výše jmenovaných korpusových zdrojů (tedy z textů Jana Čepa), zaměřuje se nicméně specificky na oblast barev a jejich využití v literárním díle. Monografie postupuje od obecných výkladů o konceptualizaci barev (od materiálně-fyzikálních, přes kognitivně lingvistické až k literárním), do nichž zasazuje svoji vlastní interpretační pozici. Ta, jak autor na mnoha místech zdůrazňuje, sice vychází z korpusových (tedy empirických a kvantifikovaných) zjištění o textu či korpusu, nicméně tyto výsledky pro něj nejsou konečným cílem, ale pouze ukazatelem toho, jakým směrem se má interpretace ubírat. Je třeba zdůraznit, že autor si dobře uvědomuje, že kvantitativní výsledky/modely (v nejjednodušším pojetí si pod nimi můžeme představit frekvenci slov vyjadřující určitou barvu v textu) jsou ve své podstatě dekontextualizované (počítáme frekvenci *typu*, nikoli *tokenu*) a až jejich zpětným zapojením do vztahu s okolními jednotkami se dobíráme nějakého významu nebo smyslu. R. Změlík tuto kontextualizaci činí na několika úrovních: jednak na úrovni úzce textové, kdy probírá typické kolokace a jejich implikace pro význam barev, jednak na úrovni textu, kdy zapojuje širší souvislosti daného literárního díla, tak i na úrovni registru či jazyka (tedy ve vztahu ke korpusu české prózy jiných autorů, než je J. Čep, či ve vztahu k reprezentativnímu korpusu psaného jazyka SYN2015). Právě tento mnohohrstevnatý



## FILOZOFICKÁ FAKULTA Univerzita Karlova

Ústav Českého národního korpusu

postup, který bere v potaz různě široký kontext, je na předložené práci cenný a otevírá nové perspektivy korpusově založené literární interpretace.

Necítím se být kompetentní komentovat literárněvědnou stránku celé práce, omezím proto své poznámky a otázky na témata, která se týkají lingvistické části práce. Když autor mluví o jazyku uměleckého díla jako o idiolektu (např. s. 58), nabízí se otázka, do jaké míry je taková charakteristika přiléhavá. Specifikum jakéhokoli textu se odehrává na několika rovinách, z lingvistického pohledu bychom asi nejspíš akcentovali dvě: rovinu tematickou (obsahovou) a rovinu registrovou, jejíž součástí by byl i onen idiolekt, jako souhrn charakteristik do textu vnesených osobou autora. Idiolektové charakteristiky jsou ovšem v kompetenci s prostředky, které text utvářejí z důvodů jak žánrových, tak funkčních (registr, styl). Toto rozlišení (téma vs. registr a v rámci registru pak idiolekt vs. žánr/registr v užším smyslu) jsem v práci postrádal, přitom by pro další promýšlení metod mohlo být určitým vodítkem. Pro obsahovou analýzu se zejm. v diskurzních studiích s úspěchem využívá metoda identifikace klíčových slov, která by rovněž mohla být inspirací a přínosem v dalším korpusovém studiu literárního textu.

Při analýze lexikálního významu výrazů vyjadřujících barvy se R. Změlík přirozeně musí vypořádávat s tím, že vedle nociónální složky je sémantika, zejm. v rámci literárních děl, utvářena systémem asociací („jestliže v rámci přirozeného světa je konceptuální význam hodnotou základní, ke které se připojují konotované významy, ve fikčním světě je základní hodnotou textově konotovaný lexikální význam barvy, jenž má fikční kvalitu“, s. 102). V této souvislosti se nabízí využití kolokací, které jsou schopny zprostředkovat informaci o bezprostředním kontextu, což autor v materiálové části práce poctivě činí. Vzdálenější asociace ovšem R. Změlík empiricky nepodchycuje, což je pochopitelné, protože korpusové nástroje pro takovou analýzu nejsou k dispozici v takové míře jako kolokační míry, a supluje tento nedostatek introspekci (kontext fikčního světa) a důkladnou znalostí analyzovaných textů (close reading). Analýza výrazů pro barvy je tedy ve výsledku úplná, odchyluje se tím nicméně od metodologického cíle, jímž je propojení empirických metod korpusové analýzy textu a literární interpretace.



## FILOZOFICKÁ FAKULTA Univerzita Karlova

### Ústav Českého národního korpusu

V oblasti materiálové analýzy lze k předložené práci připojit několik připomínek, které by mohly přispět k větší robustnosti interpretace, která je na výsledcích založena. Obecný kvantitativní model pro češtinu a angličtinu (prezentovaný na s. 119) je založen na srovnání dvou referenčních a reprezentativních korpusů: SYN2015 pro češtinu a BNC pro angličtinu. Interpretace výsledků na nich založených by měla na začátek vytknout, že tyto korpusy nejsou zcela srovnatelné co do svého složení (např. BNC obsahuje podstatný komponent mluveného jazyka, zatímco SYN2015 je čistě psaný; BNC je striktně synchronní a nepřekladačový korpus britské angličtiny, zatímco SYN2015 má větší diachronní hloubku a podstatný komponent překladových textů; proporce jednotlivých žánrů nejsou zcela srovnatelné). Je nicméně pravda, že jsou tyto korpusy představují nejlepší aproximace pro obecná srovnání, které máme.

Výsledný graf frekvencí barev v češtině a angličtině (Graf 1) pak prezentuje relativní frekvence, což je jistě významně lepší, než pokud by obsahoval frekvence absolutní, nicméně komplexní a vyčerpávající kvantifikaci běžnosti jednotlivých barev asi není, protože zanedbává disperzi – zde se nabízí třeba použití míry ARF. Druhou připomínkou k tomuto grafu (a obecně k prezentování frekvencí v práci), je absence konfidenčních intervalů pro jednotlivé hodnoty – výsledkem je určitá absolutizace konkrétních frekvencí pro jednotlivé barvy, přičemž realita v jazyce je mnohem méně ostrá; v jiných korpusech, které by byly vytvořené stejnou metodologií ovšem z jiných textů by výsledky byly drobně jiné. Tuto přirozenou a nevyhnutelnou variabilitu zprostředkovávají právě konfidenční intervaly (viz např. nástroj Calc, <https://korpus.cz/calc/>), jejichž použití může varovat před některými přímočarými závěry (např. pozice barev 6–8 v Grafu 1 vykazují malé rozdíly, je tedy možné, že jejich konfidenční intervaly se významně překrývají, a tudíž nelze na jejich pořadí dané přesnou hodnotou relativní frekvence zcela spoléhat). Absenci konfidenčních intervalů lze vyčíst i dalším výsledkům, nicméně je třeba na druhou stranu uvést, že autor se při interpretaci zaměřuje především na frekvenčně nejvýznamnější barvy a nejvýraznější rozdíly mezi nimi, které by analýza s pomocí konfidenčních intervalů pravděpodobně stejně neovlivnila.



## FILOZOFICKÁ FAKULTA Univerzita Karlova

### Ústav Českého národního korpusu

Za diskusi by nicméně stál vliv velikosti korpusů na jednotlivé výsledky (pro představu, největší korpus používaný v práci má 100 mil. tokenů, nejmenší má desítky tisíc slov). Např. frekvenční profily základových barev v Grafu 7 ukazují na dvě skupiny korpusů, které nápadně korelují s jejich velikostí. Je možné, že rozdělení, které zde sledujeme, je „pouze“ artefaktem velikosti korpusu? Stejně tak by bylo dobré prověřit vliv velikosti korpusu na hodnotu D (poměr nejfrekventovanější chromatické barvy na součtu obou achromatických barev). Pokud zde takový vliv je, funguje jako rušivý faktor a bylo by dobré ho buď eliminovat nebo zohlednit v interpretaci.

Postup, který R. Změlík volí při popisu výsledků, sleduje směr od obecnějších korpusů ke konkrétnějším až na úroveň jednotlivých textů J. Čepa. V této souvislosti se odhaluje vliv agregace, která může způsobit, že některé trendy viditelné na jedné úrovni (např. na úrovni jednotlivých děl), se na jiné úrovni (třeba v celku korpusu všech děl J. Čepa) vzájemně vyruší a ztrácejí či naopak násobí. V této souvislosti nelze nezmínit to, že většina analyzovaných textů jsou sbírky povídek, tedy rovněž svého druhu agregáty, a v rámci povídek bychom mohli identifikovat menší narativní úseky, na něž by bylo možné texty dále rozčlenit atp. Problém agregace tedy nelze beze zbytku eliminovat, o to cennější je, že si je toho autor vědom a výsledky z jednotlivých stupňů agregace porovnává.

Mezi drobné nedostatky práce bych zařadil to, že při prezentaci údajů jsou některé grafy informačně přetíženy – za extrémní v tomto ohledu považuju graf 4, který prezentuje jednotlivé barvy v různých korpusech spolu s jejich disperzí. Za drobný nedostatek lze rovněž považovat to, že nejsou podrobněji specifikovány parametry hledání kolokací (min. frekvence v korpusu a v kontextovém okně). Jinak je ale práce po formální stránce perfektní a velmi přehledná (doplněná přílohami a rejstříkem).

V závěru se R. Změlík vrací k obecnému tématu monografie, kterým je metodologické obohacení literárněvědné interpretace o empirické metody korpusové lingvistiky. Pod pojmy *close* a *distant reading* zde rozebírá možnosti obou přístupů a konstatuje nejen velký funkční překryv mezi nimi, ale rovněž rozsáhlý potenciál vzájemného obohacení. Přínosem korpusových metod je přitom obecně menší subjektivita a možnost zpracovávat větší objem



## FILOZOFICKÁ FAKULTA Univerzita Karlova

Ústav Českého národního korpusu

dat. Stále nicméně platí – a práce to na několika místech přesvědčivě dokládá – že všechny výsledky korpusové sondy jsou pouze ukazatele, které musí uchopit kompetentní interpret, aby jim dal smysl.

Přes některé dílčí výhrady a polemické komentáře uvedené výše není pochyb, že předložená práce představuje významný počín na poli propojování korpusové (lingvistické) expertízy a literární interpretace. Aniž bych chtěl jakkoli zanedbávat přínos práce pro poznání tvorby Jana Čepa, osobně si nejvíc cením právě jejího metodologického přínosu, který má potenciál inspirovat další badatele z obou disciplín k překračování hranic a vytváření dalších obdobně pronikavých syntéz. Z těchto důvodů a z důvodů uvedených výše v textu považuju práci R. Změlíka za plně vyhovující standardům habilitačního řízení, a doporučuji ji tak k projednání před vědeckou radou.

V Praze, 9. ledna 2024