

## Eliška Konývková: Experimentální validace metrik lexikální diverzity

*posudek oponenta bakalářské práce odevzdané na oboru Obecná lingvistika na Filozofické fakultě Univerzity Karlovy v ak. roce 2021/2022*

Bakalářská práce Elišky Konývkové se zabývá konceptem lexikální diverzity, tj. souboru metrik, které vyjadřují lexikální a lexikálně-sémantickou variabilitu textů a které jsou využívány v různých oblastech především aplikované lingvistiky. Autorka konstatuje, že tyto metriky odvozené z měřitelných vlastností textů začaly být teprve relativně nedávno konfrontovány se subjektivním vnímáním textů uživateli jazyka. Ve své práci vychází ze série studií, jejichž cílem byla právě taková konfrontace, a aplikuje tuto metodu validace na české texty.

K jednotlivým bodům hodnocení práce:

### 1. Koncepce a struktura práce

Práce má jasně a úzce vymezené téma, výzkum je zasazen do širšího odborného kontextu a přímo reaguje na nejnovější studie dotýkající se této tematiky. Výzkum je tak uspokojivě odůvodněn, byť by bylo možná zajímavé, kdyby se autorka širěji zamyslela nad tím, proč je důležité souvislost mezi metrikami a subjektivním hodnocením zkoumat.

Výzkumná otázka je jasně formulovaná a design studie je vzhledem k ní vhodně zvolen. Autorka v závěru také vyhodnocená data s cíli výzkumu jasně usouvztažňuje, ačkoli navzdory oborovým zvyklostem chybí samostatný oddíl diskuze. Práce má jasnou strukturu a jednotlivé oddíly na sebe dobře navazují, experimentální část přímo vychází z teoretické. Vzhledem ke strukturování textu je třeba ocenit poměrně svižný a úsporný styl, nezátížený příliš komplikovanými strukturami a vazbami.

Přesto mám ke struktuře práce jednu výhradu. Pro čtenáře, který se v problematice neorientuje úplně dobře, není příliš přívětivé, že je v úvodních oddílech textu výčet metrik, které jsou v práci blíže představeny až později. Bylo by vhodné k dané části (2.4.2) alespoň odkázat, aby měl čtenář jistotu, že se s obsahem zmíněných konceptů později blíže seznámí. V oddílech 2.3 a 2.4 je pak představeno několik klíčových studií, jejichž obsah je poměrně detailně shrnut, chybí však jakékoli kritické zhodnocení tohoto výzkumu. Čtenářsky poněkud nepřístupné je také autorčino implicitní očekávání předporozumění některým konceptům, které v práci zmiňuje. Ačkoli se typicky jedná o termíny či koncepty, které s vlastním tématem práce souvisí jen okrajově, bylo by dobré v poznámkách pod čarou alespoň krátce vysvětlit, jak jim čtenář má rozumět (např. termín *writing quality* (s. 9) či konstatování, že lexikální diverzita je porovnávána s diverzitou biologickou (ss. 12-13)). Podobně výčet dimenzí konstruktu lexikální diverzity (s. 14) by si zasloužil bližší komentář.

Pokud jde o strukturu práce, je třeba také konstatovat, že uvedený abstrakt vlastně není v pravém slova smyslu abstraktem, ale spíše souhrnem zadání/tématu práce.

Podněty do diskuze:

- V oddíle 2.2 autorka píše, že je “vhodné zmínit, proč je validace metrik LD důležitá” (s. 9), ve skutečnosti však pouze uvádí příklady využití těchto metrik. Nabízí se otázka, nakolik ze samotného “nasazení” metrik v praxi plyne potřeba zjišťovat, do

jaké míry korelují s “lidskými” hodnoceními. Mohla by autorka toto ještě okomentovat?

- Východiskové studie pracují s tréninkem hodnotitelů, kteří jsou mimo jiné konfrontováni s referenčními texty, pro které je na základě metrik lexikální diverzity předem uveden rating. Mohla by se autorka ještě blíže zamyslet nad vhodností takového postupu, kdy jsou ratingy hodnotitelů, pro něž autoři zjišťovali korelace s metrikami lexikální diverzity, “nakalibrované” podle textů, které byly ohodnoceny právě na základě takových metrik?

## 2. Metodologie a zpracování dat

Práce má solidní metodologické základy. Design experimentu je poměrně jasně a detailně popsán, autorka dobře zdůvodňuje kritéria pro výběr textového materiálu z dostupných korpusů. Na druhou stranu chybí důkladnější zdůvodnění nejvýznamnější odchylky od studie, která sloužila jako vzor pro výzkum, sice absence tréninku hodnotitelů a referenčního textu s předem známou mírou lexikální diverzity. Toto rozhodnutí je komentováno pouze jednou větou v závěru práce. Zpracování korpusových dat i ratingů udělených jednotlivými hodnotiteli užívá vhodné přístupy, které jsou poměrně jasně popsány, a výsledky jsou prezentovány poměrně přehledně. V tomto kontextu je škoda, že nebyla o něco detailněji popsána i metodika sběru a dalšího zpracování následných rozhovorů.

Je třeba ocenit, že autorka na základě výsledků hodnocení i následných rozhovorů zajímavým způsobem reaguje na předchozí výzkum a přichází s některými podnětnými závěry.

Níže uvádím několik doplňujících otázek a doporučení:

- Jaká byla forma následných rozhovorů (ústní/písemné)? Jakým způsobem byly odpovědi zpracovány, resp. na základě jakého klíče autorka vybírala tematické okruhy a konkrétní odpovědi prezentované na s. 37-39?
- Uvítal bych, kdyby byl v příloze uveden soupis použitých textů s naměřenými hodnotami metrik a udělenými ratingy.
- Podobně v textu chybí přesné znění instrukcí pro hodnotitele, citovány jsou pouze neupravené instrukce ze studie Jarvis 2017.
- Z popisu není zcela jasné, jak přesně byl sestaven korpus hodnocených textů. Platí, že každý text měl unikátního autora? Co přesně znamená sloupec “Identifikace” v přehledové tabulce textů (s. 18)?
- Na s. 31 autorka píše, že “[n]aměřené hodnoty [korelace] se pohybují u všech hodnotitelů kolem nuly, ať už do plusu nebo do minusu, což je přesně to, co bychom očekávali”. Mohla by toto autorka rozvést, proč bychom takový výsledek zde měli očekávat?
- Jakým způsobem byli vybíráni hodnotitelé? Byl před jejich účastí ve výzkumu získán informovaný souhlas?
- Hraje v oblasti lexikální diverzity a jejího případného vnímání nějakou roli typologický profil jazyka?
- V oddíle 4.5.1 soudím, že jde spíše o obecný popis pozorovaných tendencí než o výklad o žánru “popis místa”.
- Vzhledem k velikosti vzorku bych osobně považoval za vhodnější použít namísto

houslových grafů jitter plot nebo beeswarm plot.

- Na s. 19 autorka píše, že sledovala i známky udělené jednotlivým textům vyučujícími, protože “lexikální diverzita se do celkového hodnocení, známky, také promítá”. Následně však dochází k závěru, že tomu tak nejspíše není. Bylo by tedy lepší na tomto místě volit poněkud jinou formulaci.

### 3. Práce s odbornou literaturou

Autorka prokazuje solidní orientaci v tématu, uvedená literatura je vhodně volena. Je třeba ocenit, že autorka v seznamu literatury uvádí i softwarové nástroje, s nimiž pracovala při vytěžování dat i jejich dalším zpracování. Bylo by nicméně vhodné i u opakovaně citovaných východiskových studií uvádět odkaz i s vročením (viz např. s. 24 “[p]ři posuzování shody ve svém experimentu měřili Kyle et al. shodu pouze mezi dvěma respondenty...”).

### 4. Formální úroveň práce

Práce má formální náležitosti bakalářské práce, je zvolen vhodný citační aparát, který je užíván systematicky napříč prací. Formální úroveň je ovšem jedním ze slabých míst celé práce a bohužel poměrně sráží celkový čtenářský dojem.

Největším formálním problémem je absence jakéhokoli popisného aparátu u prezentovaných tabulek a detailnějšího popisu u grafů. Především u grafů korelace 1, 8, 10 a 11 by méně informovaný čtenář mohl mít problém cokoli vyčíst. Napříč prací se pak opakují pravopisné chyby, nejvíce v přičestích (např. s. 18 “[d]alší údaje ke vzniku samotných textů nebyli v korpusu k dispozici”). Velmi rušivě také působí typografická nejednotnost. Některé odstavce jsou zarovnané do bloku, jiné doleva. Mezi některými odstavci jsou mezery, jinde ne, obdobně jsou některé (pod)kapitoly odděleny poměrně velkými prázdnými místy.

## 5. Celkové hodnocení

Bakalářská práce Elišky Konývkové splňuje jak formální, tak obsahové požadavky na závěrečné práce tohoto stupně. Přes uvedené výtky je třeba konstatovat, že se jedná o solidní práci, která je zajímavým příspěvkem do této oblasti lingvistiky.

Práci doporučuju k obhajobě a navrhuju s přihlédnutím k průběhu obhajoby hodnotit *velmi dobře -- výborně*.

Ve Včelné pod Boubínem 2. 2. 2022

Michal Láznička