

Posudek bakalářské práce Nikoly Pialkové Čeština čínských mluvčích v korpusu CzeSL-SGT

Bakalářská práce se zaměřuje na korpus CZESL-sgt a analýzu chyb čínských mluvčích češtiny. Ještě před samotným textem autorka vysvětluje třináct klíčových pojmů spojených s tématem práce, které převzala z publikace Barbory Štindlové. Zajímavé je, že ve slovníčku uvedený pojem jazykový výstup = output není mimo tento slovníček v práci nikde použit.

Autorka nejprve představuje korpus CZESL-sgt a jeho postavení v systému korpusů Českého národního korpusu. Jde o kapitoly svojí podstatou popisné, přesto se nemohu ubránit poznámce, že jsou poskytnuté informace velmi strohé, citace jsou někdy vybrané nevhodně (např. na str. 11 autorka cituje z příručky ČNK, citaci přejímá i včetně poděkování: *Dále patří vděčnost všech uživatelů dalším externistům z celé republiky, kteří přispívají nejen velkým množstvím materiálů, ale jsou též pomocí při koordinaci potřebných činností*). V textu je nejen velké množství překlepů a chyb, ale (a zejména) je velmi nepřesný terminologicky, autorka např. mluví o autentičnosti korpusů namísto jazykových dat (s. 12), prohledávat můžeme korpus, nikoliv lemmatizaci (*Lemmatizace u korpusu CzeSL-plain neproběhla, a proto je v ní mnohem obtížnější vyhledávat*, s. 14), titulky u tabulek jsou někdy až úsměvné (*Statistika obsazenosti studentů v CzeSL-SGT podle prvního jazyku*, s. 18), podobně i formulační neobratnosti (*Podle SERR (Společného evropského referenčního rámce) jsou v korpusu obsaženy čtyři základní skupiny (A1; A2; B1 a B2)*, s. 23; *Dalším aspektem, v němž by se úrovně mohly lišit, jsme předpokládali povolenou pomůcku*, s. 24), *publicistický spisovný jazyk*, s. 34.

Poměrně značná část práce je věnována popisu dostupných metadat. Autorka zpracovala informace o příslušné části korpusu CZESL-sgt do přehledných tabulek, v textu ale postrádám informaci o tom, jak byla metadata sbírána, zda je zprostředkoval sběrač textu, nebo sám autor, jak byla stanovena jazyková úroveň, ke které je daný text přiřazen apod. Dostupná metadata jsou pouze popsána, očekávala bych spíše jejich kritické zhodnocení. Při vyvozování závěrů bych autorce doporučila být zdrženlivější, její formulace vyznívají víceznačně (*V otázce bilingvismu bohužel značná část studentů neodpověděla. S jistotou však můžeme říci, že téměř 60 % korpusu vytvořili Číňané, kteří bilingvismus popřeli*, s. 26, *rozsah 50-99 slov je mezi čínskými studenty nejběžnější*, s. 29). Přestože je popisu metadat věnována značná pozornost, při analýze chyb se s metadaty žádným způsobem neparčuje.

Jádrem práce jsou ukázky chyb, které čínští mluvčí v češtině typicky dělají. K metodologii této části práce směřuje moje největší připomínka. Autorka uvádí: *Protože se v tomto případě jedná o kvalitativní analýzu vybraných typů chyb, které jsou typické buď přímo pro čínské studenty či pro mluvčí analytického mateřského jazyka, postup práce začal v důkladném nastudování gramatik*

čínského jazyka, a to především Gramatiky současné čínštiny (Lingea, Brno 2018). Následně jsme si určili pravděpodobné chyby, které by mohly být v korpusu obsaženy, a na základě této domněnky poté hledali chybové typy v textech korpusu (s. 32). Z textu není možné zjistit, na základě jakých kritérií byly tyto typické chyby z gramatik vybrány, zda šlo o náhodný výběr, nebo zda byl výběr něčím motivován. Metodologií výběru ukázek trpí i zbytek textu, autorka spíše dokládá jednotlivé příklady (frekvence výskytu jednotlivých typů chyb je velmi nízká), nelze tedy rozhodnout, zda jde o chyby tzv. typické pro čínské mluvčí češtiny (samozřejmě v rámci korpusu CZESL-sgt, nebo o chyby, pro které bylo vůbec možné ve zmíněném korpusu dohledat příklady).

Přes uvedené výhrady práci doporučuji k obhajobě a navrhuji hodnocení dobře.

V Praze 23. 1. 2019

Mgr. Kateřina Šormová, Ph.D.