

Ústav romanistiky Filozofické fakulty Jihočeské univerzity v Českých Budějovicích  
Oponentský posudek diplomové práce

Název práce: Překlad vybraných lexikálních jednotek nástroji podporujícími strojový překlad – zaměřeno na frazeologismy obsahující názvy zvířat.

Autorka práce: Bc. Simona Venkrbcová

Vedoucí práce: PhDr. Jana Pešková, Ph.D.

Předkládaná diplomová práce se zabývá překladem frazeologismů obsahujících zooapelativa nástroji podporujícími strojový překlad. Jejím cílem je prostřednictvím kontrastivní analýzy zhodnotit úspěšnost překladu vybraných frazeologismů ze španělského jazyka do českého (s. 4). Téma práce koresponduje s narůstajícím trendem využití strojového překladu v běžném životě a zároveň adekvátně poukazuje na úskalí, která s sebou přináší idiomatická podstata a nekompozicionalita významu frazeologických jednotek při jejich automatickém překladu.

Práce je standardně členěna na teoretickou (s. 10–43) a rozsáhlejší praktickou část (s. 44–108), jež se autorce podařilo úspěšně propojit v koherentní celek. Teoretická část je rozdělena do čtyř kapitol. První z nich se tematicky zaměřuje na představení systémů strojového překladu (MT), jejich historii, typologii (*pravidlové, statistické, hybridní a neurální*) a problémy spojené s automatizovaným překladem (např. *bridging* nebo rozřešení významové dvojznačnosti, tzv. *ambiguity*). Postihnouti zejména sémantických úskalí strojového překladu autorku velmi záhy přivádí k formulování hypotézy, a sice že „(...) se předpokládá, že vzhledem k povaze frazeologismů, (...), nebudou vybrané nástroje strojového překladu ve většině případů schopny pro danou frazeologickou jednotku ve španělském jazyce vyhledat správný ekvivalent v jazyce českém,“ (s. 22). V první kapitole diplomandka neopomíná uvést ani čtyři vybrané systémy pro on-line strojový překlad, s nimiž bude pracovat při výzkumu (Google Translate, DeepL Translate, Bing Microsoft Translator, MyMemory). Druhým tematickým blokem jsou dvě kapitoly věnované vymezení, funkcím, vlastnostem a typologii frazeologických jednotek obecně a zooapelativ konkrétně, a to jak z perspektivy české literatury (Čermák, Mrhačová), tak z pohledu španělského (Corpas Pastor). V samém závěru teoretické části pak autorka vysvětluje *absolutní, relativní, analogickou a nulovou ekvivalenci* (vychází zejména ze Strakové či Vlahova a Florina), jejichž znalost bude aplikována při praktickém výzkumu. K teoretickému oddílu práce nemám výhrad (snad by mohlo být lépe popsáno, jak autorka chápe pojem *obraznost*, se kterým hojně pracuje, ale to je připomínka spíše okrajová).

Úctyhodná výzkumná část je rovněž velmi propracovaná. Kladně hodnotím její přehlednost, logickou strukturu a názornost. V úvodu do tohoto oddílu (s. 44–50) je představena metodologie, k níž nemám žádné připomínky (selekce 48 frekventovaných frazémů z relevantních slovníků – Junceda, Doubek, jejich vyhledání v kontextu v jazykových korpusech ČNK a CREA, následný překlad pomocí zvolených systémů MT, zachycení výsledků kontrastivní analýzy v tabulkách a celková syntéza znázorněná graficky). V pořádku je i stanovení dílčích hypotéz na s. 44, k jejichž podstatě ale níže položím doplňující otázku (ot. č. 5). Místy se vyskytují zavádějící formulace jako „lze říci, že určité skupiny budou vykazovat *lepší správnost* překladu,“ (s. 44, zvýraznění je oponentky), ale naopak vítám, že autorka věnovala čas důkladnému posouzení ekvivalentů mezi španělským a českým jazykem, kde jsem identifikovala jen minimální spornost (např. *být ve svém přirozeném prostředí* dle mého názoru nutně neznamená *být ve svém žvilu* – s. 54, ze spojení *pekelná doba* není odvoditelný význam *psí počasí* – s. 59, frazém *como abeja en flor* bych posuzovala samostatně, jelikož se u něj nejedná o absolutní ekvivalenci k překladu *jako ryba ve vodě* – s. 54), k zamyšlení vybízí také otázka, kam se v automatickém překladu ve všech případech vytratila spojka *když* u frazému *cuando el gato no está, los ratones bailan* a všude byla nahrazena slučovací *a* – s. 55.

Jeden z drobných komentářů směřuje také k tomu, že poslední pasáž kontrastivní analýzy k *nulové ekvivalenci* (s. 99–101) nezachovává shodné prezentační schéma jako pasáže předchozí a čtenář tak musí opakovaně otáčet stránky, aby se seznámil s překladovými řešeními, ale rozumím tomu, že tak patrně bylo učiněno s ohledem na rozsah práce. Celkově jsou však analýzy zdařilé a úspěšně porovnávají míru adekvátnosti překladu frazémů mezi zvolenými systémy strojového překladu s tím, že jako nejuspěšnější byl vyhodnocen DeepL Translate (s jistými překladovými odchylkami dosáhl téměř 50% úspěšnosti), naopak nejméně

úspěšný byl Microsoft Bing (asi 27% úspěšnost). K obsahu praktické části mám přesto několik průřezových otázek, které mohou posloužit jako opora pro obhajobu:

1. S odkazem ke s. 82 – mohla by autorka upřesnit, co myslela formulací „*Překlady se od českého ekvivalentu liší jak sémanticky, tak i významově*“?
2. Byl by ekvivalentem k frazému *El perro del hortelano (que ni come ni deja comer)* český překlad *Sedí na tom jako žába na kameni*, nebo spíše *jako žába na prameni*?
3. Domnívá se autorka, že výstupy strojového překladu by mohly vést například ke globálnímu sjednocování frazeologismů napříč jazyky, zejména pokud je původní frazém cílového jazyka zastaralý? Odkazuji např. k tomu (viz s. 67), že frazeologismus *de noche, todos los gatos son pardos* byl všemi zvolenými překladači do češtiny (doslovně) přeložen jako *v noci jsou všechny kočky hnědé*, namísto původní české alternativy *potmě je každá kráva černá*.
4. Zamýšlela se autorka nad tím, jak je možné, že Google Translate a MyMemory nejenže vykazují stejná čísla úspěšnosti překladů, ale zároveň se ve výsledném překladu do vysoké míry shodují? Nemohou např. vycházet ze stejných databází?
5. Lze opravdu pravděpodobnost správnosti překladu frazému z výchozího do cílového jazyka odvíjet pouze od míry ekvivalence? Nebude spíše záviset na frekvenci, s jakou se konkrétní frazémy v obou jazycích používají a jak hojně se vyskytují ve zdrojových datech překladačů? (Viz např. *haber gato encerrado vs. být v tom háček*.)
6. S odkazem ke s. 91 – jednalo by se o překlady pouze „částečně“ správné, pokud daný frazém vždy hodnotíme v kontextu celé věty z korpusu a spojení *do jámy lvové* zůstalo zachováno?

Práce je psána velmi pěknou češtinou s občasnými překlipy (*střádání* namísto *strádání*, např. s. 88) a opakujícími se chybami v interpunkci zejména u vztahných vět vložených a za příslovečnými určeními (patrně vlivem interference ze španělského a/či anglického jazyka; např. s. 8, 23 ad.). Cizojazyčná interference je zřejmá také u opakujícího se analytického stupňování, které není češtině vlastní (např. *nejvíce používaných, nejvíce přirozeně* namísto *nejpoužívanějších, nejpřirozeněji*; s. 8, 15, 38, 108 ad.). Autorka dále v seznamu bibliografie nerozlišuje časopisecké příspěvky, kapitoly v knize a monografie a nezachovává jednotnou citační normu (nepravidelně je uváděno např. ISBN; někdy jsou uváděna celá jména, jindy pouze iniciály). Formátování textu je velmi přehledné, barevné tabulky a grafy usnadňují orientaci ve výsledcích výzkumu, španělské resumé neobsahuje zásadní chyby.

**Závěr:** I přes uvedené připomínky se domnívám, že práce Bc. Simony Venkrbcové je velmi zdařilá a splnila požadavky kladené na tento typ prací. Diplomandka prokázala schopnost systematicky analyzovat jazykový vzorek, syntetizovat teoretické poznatky a dostat stanovenému cíli. V práci je zřejmá studentčina svědomitost, zvláště pak oceňuji její snahu o maximální objektivitu při posuzování adekvátnosti strojových překladů. Předloženou práci doporučuji k obhajobě a předběžně ji hodnotím známkou **výborně** s tím, že výsledné hodnocení bude stanoveno na základě obhajoby.

V Českých Budějovicích dne 12. ledna 2023

Mgr. Markéta Šmídová, Ph.D.