

POSUDEK DIPLOMOVÉ PRÁCE

<i>Autorka práce:</i>	Bc. Klára Kopecká
<i>Název:</i>	<i>Kategorizace úprav strojového překladu při post-editaci: jazyková kombinace angličtina - čeština</i>
<i>Akademický rok:</i>	2020/2021
<i>Vedoucí diplomové práce:</i>	PhDr. Bc. Tomáš Svoboda, Ph.D.
<i>Oponent:</i>	Mgr. Věra Kloudová, Ph.D.

Diplomová práce Kláry Kopecké je nezanedbatelným přínosem pro výzkum aktuálních trendů v praxeologii překladu. Práce se zabývá posteditací výstupů ze strojového překladu, na textovém materiálu ve směru angličtina–čeština. Je rozdělená na část teoretickou a empirickou. V teoretické části se autorka zabývá vymezením pojmů stěžejních pro zkoumané téma, tedy nástroje CAT, strojový překlad či posteditace strojového překladu. V empirické části představuje dataset, který analyzovala, a shrnuje výsledky své analýzy jako kategorizaci úprav provedených při posteditaci výstupů ze strojových překladačů a následné revizi těchto již posteditovaných textů.

Tematické zaměření práce je moderní (posteditace), vlastní přístup ke zpracování materiálu pak zásadně inovativní (místo sledování chyb, jichž se překladače dopouštějí, elicitace úprav v translátu provedených posteditory a revizními pracovníky). Cíl práce je přísně aplikovaný, je jím „poskytnout základní přehled [úprav při posteditaci] pro případné kurzy posteditace strojového překladu“ (str. 78). Diplomantka svůj záměr neztrácí ze zřetele, z hlediska metodologie je tedy práce vhodně zacílená (např. odhlédnutí od výzkumu firmy Memsources na daném datasetu či nehodnocení kvality strojového translátu).

Diplomantka využila zdrojů (dataset) poskytnutých společnostmi, u níž pracuje, což přineslo možnost zaměřit se na samotnou analýzu jazykového materiálu bez nutnosti vynakládat úsilí na jeho sběr. Vlastní analýza proběhla „manuálně“, z ryze kvalitativního hlediska. Postup analýzy nelze zpochybnit, plně (byť nepřiznaně) odpovídá heuristickému přístupu, který ve své práci *Umění překladu* už před téměř 60 lety doporučoval Jiří Levý (tedy probírat se materiálem, který jaksi sám naznačí vhodnou kategorizaci relevantních fenoménů). Mám však výhrady k výsledku tohoto analytického přístupu, a to k vlastním kategoriím, které takto vykrytalizovaly. A snad nikoli ke kategoriím jako takovým, ale spíše k jejich některým označením a výsledné systematizaci do jistého předběžného modelu. V posudku zde níže naznačím alternativní způsob pojmenování a částečně i systematizace těchto kategorií. Koneckonců z práce vzešlý model/systém kategorií má být podnětem pro uchopení jednoho aspektu výuky posteditace v ČR, takže je to téma závažné, přičemž robustně zpracovaná práce Kláry Kopecké takovou polemiku jistě snese. Nabízí se úvaha, že by taková výměna názorů mezi vedoucím práce a diplomantkou bývala měla proběhnout v rámci konzultací během zpracování práce, to se však z časových důvodů nezdařilo.

Poznámky k pojmenování a systematizaci z analýzy vzešlých kategorií:

- Skupiny PŘESNOST a PLYNULOST jsem zpočátku považoval za ekvivalenty kategorií *accuracy* a *fluency*, které se nejčastěji používají jako binární opozice při kvalitativním hodnocení automaticky generovaných translátů v oblasti NLP. Přítomnost dalších skupin (LEXIKUM, STYL apod.) mi ale naznačuje, že vymezení kategorií bylo patrně do jisté míry motivováno jinak. Pokud tomu tak je, bývalo by bylo vhodné od těchto (poněkud zavádějících, ale především zcela redukcionistických ekvivalentů – tedy přesnost a plynulost) upustit.
- PŘESchyb / Chybný překlad: Označení chybný překlad je nepřesné, neboť se „chybné překlady“ vyskytují i v jiných skupinách/kategoriích. Navrhuji Sémantický posun.
- PŘESnepř/Nepřeloženo: Opět nepřesné označení, vždyť někdy (u překladatelů) je převzetí originálního výrazu záměrné, čili jedná se o překladatelský postup (srov. např. Newmarkova taxonomie překladatelských postupů).
- PLYNinterp/Interpunkce: Zařazení interpunkce do skupiny „plynulost“ je sporné. Posouzení interpunkce se zpravidla vyznačuje větší exaktností než „plynulost“, mnohem častěji (i když samozřejmě ne vždy) se interpunkční jevy nacházejí na škále správnosti, a nikoli vhodnosti/idiomatickosti. Proto by bylo vhodnější např. zařadit interpunkci do gramatické kategorie, jež patří ke kategoriím, „které [podléhají] pevně stanoveným pravidlům“ (str. 77).
- STYLneobr / Stylistická neobratnost: Kategorii navrhuji sloučit s kategorií STYLneidiom / Stylistická neidiomatickost. U stroje snad kategorie (ne)obratnosti postrádá opodstatnění, navíc si diplomantka uvědomuje vysokou míru subjektivity spojenou s rozlišením obou kategorií (srov. str. 69).

K této polemice na rovině popisu kategorií a jejich systematizace mě vede vědomí, že v translatologii existuje několik desítek modelů tzv. translatologické analýzy, zaměřených lingvisticky, behaviorálně, kvantitativně, kvalitativně, korpusově, sociologicky, kulturně, sémiologicky, diachronně, synchronně, deskriptivně, preskriptivně, redukcionisticky i holisticky, pozitivisticky i hermeneuticky, s nejrůznější metodologií i důrazy, zpracovávaných automaticky i manuálně, ale k tomuto know-how diplomantka nesáhla. Tím – ale jen tím! – se práce – ale jen ve své druhé části – drobně přibližuje absolventským pracím na pracovištích zaměřených na aplikovanou lingvistiku či formální zpracování přirozeného jazyka, které (žel) toto bohatství translatologie, nashromážděné za desítky let, mnohdy ignorují, pohybují se v jakémsi translatologickém vakuu.

Závěrem: S povděkem kvituji, že diplomantka předložila práci v tištěné podobě, což nemusela. Z formálního a jazykového hlediska nelze práci vytknout téměř nic, i typografické zpracování je na výborné úrovni. Dvě drobnosti k odbornému názvosloví: Druhá část diplomové práce by se neměla nazývat „praktická“, nýbrž empirická, a „revizory“ pojdme raději (v souladu s českým zněním normy ISO 17100) nazývat revizními pracovníky.

K diskusi navrhuji následující otázky:

- 1) Co si myslíte o předložené alternativní systematizaci a alternativních pojmenováních pro Vámi vytyčené kategorie?
- 2) Lze z datasetu zjistit podíl segmentů, které byly posteditory / revizními pracovníky převzaty ze strojového translátu bez úpravy? Pokud ano, jaký byl tento podíl?
- 3) Projevila se na výsledcích práce okolnost, že „[i]nstrukce [posteditorům] neobsahovaly požadovanou úroveň posteditace“ (str. 47)?

Shrnutí

Celkově konstatuji, že diplomantka ve své diplomové práci vyšla z aktuální a relevantní literatury a náležitě si osvojila potřebné teoretické zázemí. Práce je svým přístupem výrazně inovativní. Přesto, že polemizuji s některými aspekty výsledného zpracování analyzovaného materiálu, k vlastní metodě práce výhrady nemám a mé připomínky tedy mají primárně charakter příspěvku do odborné diskuse.

Práci doporučuji k obhajobě a navrhuji hodnocení *výborně*.

V Praze dne 2. září 2021



Tomáš Svoboda