

Posudek magisterské diplomové práce

Univerzita Palackého v Olomouci, Filozofická fakulta, KAA

Autor práce: Bc. Jana Baštinčová

Název práce: Analýza strojového překladu idiomatických sloves

Vedoucí práce: PhDr. Pavel Král

Autor posudku: Mgr. Michal Kubánek

Hodnocení: A = 1, B = 1-, C = 2, D = 2-, E = 3, F = nedostatečně

Hodnocení

1. Přínos práce, aktuálnost a důležitost tématu.

Jak zjistila i samotná autorka v průběhu závěrečné fáze psaní své diplomové práce, kdy došlo k nasazení nové verze neurálního strojového překladu, je zvolené téma velmi aktuální. Je zřejmé, že tato technologie vstupuje do nové fáze svého kvalitativního vývoje, přičemž výsledky této sondy mohou být mimoděk použity jako doklad fáze předchozí.

A

2. Kvalita přehledu poznatků o tématu. Použití odborné literatury a dalších zdrojů.

Autorka předkládá v úvodní části nazvané poněkud vágně „Rešerše dostupné literatury“ vymezení tří základních konceptů, které jsou pro práci stěžejní: idiomatické sloveso, novinový článek a strojový překlad. Při výkladu o idiomatických slovesech vychází snad až příliš z díla Jitky Šutové, na druhou stranu chybí práce anglistů (např. Dušková) nebo lexikologů a frazeologů (např. Čermák). Pojednání o stylistických rysech novinových článků je vzhledem ke zvolenému tématu jen málo relevantní, stejně jako stylistika podle Galperina (naopak chybí Čechová, Jelínek či Hoffmannová). Totéž platí ještě ve větší míře o typických vlastnostech digitální žurnalistiky (multimediální prvky, hypertextualita, vlastnosti titulku,...). Také pojednání o posledním z trojice pojmů bych směřoval spíše k představení studií, které se zabývaly podobným tématem, a k jejich zjištění a závěrům. Exkurz do historie strojového překladu nepovažuji za důležitý. Naopak rozsah informací o předmětu výzkumu, tedy strojovém překladači Google Translate, má rozsah jedné strany a vychází výhradně z jednoho zdroje, kterým je přímo blog jeho vývojářů.

Struktura přehledu poznatků o tématu je místy chaotická. Autorka jakoby neví, čím začít. Příkladem je odstavec na str. 11: „Zde se dostáváme k rozdílu mezi angličtinou a češtinou. Ze všeho nejdříve je však nutné definovat termín částice. Jeremy Harrison tvrdí, že částice je to, co „bud“ mění smysl slovesa, nebo sloveso obohatí o další význam.“ Problematické užití tohoto termínu se věnuje kapitola 1.1.3, ponecháme jej proto mezitím stranou.“

V úvodu výše uvedené věty je také nevhodně použité označení „termín“. V kapitole 1.1.3 se již mluví o „pojmu částice“, což je správné označení. Opačný problém je vidět na str. 10, kde autorka píše: „Pro tento účel poslouží definice Jitky Šutové, která pojmem idiomatická slovesa označuje...“, odkud je zřejmé, že autorka má na mysli „termín“ jako prostředek označování pojmů.

Autorka z nějakého důvodu používá při označení autorů kombinaci křestního jména a příjmení: „Jak uvádějí Hatim a Munday ve své knize Translation: An Advanced Resource Book“ (str. 6), ovšem dále pak „Basil a Munday však uvádějí, že je tento extrémní jen pouhou fabulací“ (str. 40) a „Basil a Munday uvádějí, že je to proto, že mu chybí velmi podstatné dovednosti“ (str. 41).

Pro práci v dané oblasti je vhodnější závorkový systém odkazování. Autorka často oba systémy kombinuje, navíc v textu redundantně uvádí název práce, ze které čerpá: „O historii idiomatických sloves hovoří například Anna H. Live ve své stati The Discontinuous Verb in English (2015).“ (str. 11).

D

3. Formulace záměrů práce. Metody práce. Jejich popis a adekvátnost.

Záměr práce je jasně představen hned v prvním odstavci a práce ve své praktické části adekvátně směřuje k jeho naplnění. Autorka také představuje důvody, proč ji tato oblast zajímá. Obsah kapitoly 5, která představuje metodu práce, hodnotím kladně: je dostatečně detailní, aby si čtenář mohl udělat náležitou představu o průběhu získávání dat.

Při popisu metody práce však autorka místy až úporně představuje a obhájí svá rozhodnutí, která na první pohled vypadají jako legitimní a bezproblémová. Obecně je snaha vysvětlit svá rozhodnutí v pořádku, ale zde autorka podle mě své postupy zbytečně problematizuje. Nejlépe je to vidět v „preliminárním výzkumu“, kdy zjišťuje, že novinové články tvoří jen velmi malou část produkce překladatelských agentur. Završením pak je rozhovor s jedním novinářem ve sportovní redakci Lidových novin. Na jeho základě autorka také obhájí volbu překladače Google Translate, protože „byl výslovně jmenován novináři v preliminárním výzkumu“ (str. 51).

C

4. Teoretické aspekty (argumentace, koherence).

Jak již bylo uvedeno v části 2, teoretické zakotvení práce je problematické. Navíc některé pro práci zásadní koncepty jsou nevhodně představené. Autorka ve své definici idiomatičtějšího slovesa píše, že jde o spojení, „které si běžně posluchač sám nedokáže odvodit a které je navíc často při doslovném překladu do češtiny zavádějící“ (str. 11). V lexikologii (frazéologii) se obvykle při popisu významu v těchto případech pracuje s pojmy jako kompoziční funkce, holistický význam a lexikální jednotka (lexém).

Důvody pro volbu termínu idiomatičtější sloveso nejsou příliš přesvědčivé: „Tím prvním je, že tohoto označení používá Jitka Šutová, na jejímž (sic) seznamu frázových sloves je postavena hlavní analýza této práce. Druhý důvod je ryze praktický. Kdybychom chtěli použít pravděpodobně nejčastějšího označení takových sloves, čili frázová slovesa, narazili bychom opět na nejednotnost některých autorů“ (str. 15). Dále se autorka s touto volbou dostává do problémů při definici významu označení *idiomatičtější*. Jako jeden z rysů „idiomatičtější sloves“ autorka uvádí „6. Idiomatičnost. Není však pravidlem, jelikož význam některých idiomatičtější sloves se dá odvodit z jejich komponentů“ (str. 20). To je právě důvod, proč Marks používá termín frázová slovesa. Autorka však v důsledku přesto tvrdí, že idiomatičtější slovesa nemusejí být idiomatičtější.

D

5. Praktické aspekty (návaznost na teoretickou část, výběr textů, textových vzorků, příkladů, způsob analýzy).

Praktická část je dobře promyšlená, autorka vhodně dokládá své postupy výběrem textových ukázek. Rovněž prezentace získaných dat formou tabulek a grafů spolu s textovým doprovodem je podrobná a přehledná. Kladně hodnotím také operativní provedení doplňujícího výzkumu.

A

6. Zpracování výsledků, praktický přínos práce.

Dokud autorka zpracovává své výsledky pouze čistě statisticky, je vše v pořádku. Problém nastává při snaze vztáhnout je k výzkumným otázkám, protože zde chybí operacionalizace použitých výroků (mýtů). Jako příklad poslouží odpověď na obecnou výzkumnou otázku v kapitole 6.3: „Odpověď zní, že statistický strojový překlad si s idiomatičtějšími slovesy dokáže poradit lépe, než si lidé, kteří věří některým mýtům o strojovém překladu, které byly představeny v rešerši dostupné literatury, myslí. (viz 1. mýtus: Strojový překlad je v praxi nepoužitelný).“ Není však zřejmé, jak je „nepoužitelnost“ definována a co si vlastně někteří lidé o kvalitě strojového překladu myslí. Jinými slovy, kolik procent příkladových vět musí být přeloženo chybně, aby byl celkový výsledek označen jako „nepoužitelný“.

Dále: „Z výsledku tohoto výzkumu totiž rovněž vyplývá, že strojový překlad zatím opravdu nedosahuje takové úrovně, aby se jeho výstupy daly porovnat s prací lidských překladatelů“ (str. 69). Porovnání s výkony lidských překladatelů však nebylo součástí výzkumu!

Autorka také mnohdy komentuje konkrétní příklady zavádějícím způsobem, navíc využívá nestandardní terminologii. Jako příklad lze uvést poslední komentář celé práce na str. 80. Autorka srovnává překlad věty: „Then he gave me and the girl a big kiss on the cheek and said he **was off** for the weekend.“ statistickým překladačem: MT1: „Pak mi a dívka **dala** pusu na tvář a řekl, že **byl pryč** na víkend.“ a neurálním překladačem: „MT2: Pak mi dal dívku velký polibek na tvář a řekl, že **je pryč** na víkend.“ Připojuje komentář: „V tomto případě tvar slovesa rodu činitele věty lépe přizpůsobil statistický MT1.“ Pokud je „činitel věty“ rodu mužského (he), pak správný je překlad MT2 (dal). Bohužel podobné komentáře výsledky práce činí méně důvěryhodné.

Autorka místy také nesprávně formuluje své myšlenky: „Pokud bychom však přece jen chtěli obecnou formou polemizovat, tak se dá říci, že překladatel buď dané sloveso „umí“ nebo naopak „neumí“ ve větě identifikovat a přiřadit mu správný význam.“ (str. 64). Zde má zřejmě na mysli možnost „spekulovat“ či „dovozovat“.

D

7. Formální zpracování. Jazyková a stylistická úroveň, rozsah práce, grafická úprava, dodržení publikační normy.

Po formální stránce práce obsahuje především množství chyb v interpunkci: Hlavní část práce, čili hlavní výzkum, si od počátku klade za cíl nabídnout spíše globální pohled na tuto problematiku, než konkrétnější závěry o jednotlivých slovesech. (str. 7); Tato skutečnost byla jedním z důvodů, pro volbu zaměření této práce. (str. 17); Tato práce z logiky věci zaujímá stejný pohled, jako autorka knihy, která posloužila... (str. 18); O idiomatičnosti se vyjadřuje podobně, jako Šutová, na rozdíl od ní ale... (str. 18); může spadat do libovolného žánru z těch, uvedených výše. (str. 23); Kvalita překladů, získaných pomocí nástrojů strojového překladu, je všeobecně velmi nízká. (str. 40) ...

a další přehlédnutí a opomenutí při závěrečné korektuře: „Jak je patrné z některých z definic výše,“ (str. 18); „čili u kterých se překladatel se překladatel nepodařilo“ (str. 63); „Vývodit z této skutečnosti nějaké konkrétnější závěry však není však v naší kompetenci“ (str. 64); „nikoliv jeho správnost jeho tvaru.“ (str. 70)

V českém textu je vhodné přechylovat jména, např. Susan Bassnett(ová).

Nevhodné je rozdělení poměrně malé tabulky 3 na dvě strany 62–63.

D

Další poznámky autora posudku:

Autorka vykonala velký kus práce v praktické části a její výsledky jako surová data jsou cenné. Nicméně celkové uchopení tématu v teoretické části a také zpracování výsledků je spíše podprůměrné.

Doplňující otázky k obhajobě:

1. Jaké dva rozdílné významy může mít přídavné jméno „idiomatický“?
2. Jak by bylo možné vyřešit problémy s operacionalizací výzkumných otázek?
3. V závěru práce navrhuje další možný výzkum. Jednou z cest by bylo prozkoumat, jak ovlivňuje kvalitu překladu frekvence idiomatičného slovesa v jazyce. Vy však píšete, že „[z]jistit frekvenci jednotlivých IS v praxi by však bylo značně obtížné“ (str. 71). Dokázala byste přesto navrhnout způsob, jak frekvenci lexému v jazyce zjistit?

Závěrečné hodnocení:

Práci k obhajobě doporučuji ~~nedoporučuji~~

Celkové hodnocení práce – navrhovaný klasifikační stupeň

C

Poznámka: Závěrečný klasifikační stupeň nelze vypočítat jako prostý průměr hodnocení v jednotlivých oblastech, protože je nutné zohlednit závažnost jednotlivých aspektů hodnocení a přistupovat k práci jako k celku.

Datum: 26. 5. 2017

Podpis:

Na základě kontroly vysokoškolských kvalifikačních prací systémem Theses.cz ani jiným způsobem nebyla u výše uvedené práce zjištěna taková shoda s jinými pracemi, dokumenty nebo texty, která by zakládala odůvodněné podezření z porušování autorských práv.