

# Porovnanie humánneho prekladu a posteditácie strojového prekladu

## Comparing Human Translation and Machine Translation Post-Editing

**Diana Pavlíková**

Univerzita Komenského v Bratislave

[dianapavlikova@protonmail.com](mailto:dianapavlikova@protonmail.com)

### Abstract

As machine translation becomes a part of the translation market around the world, the post-editing of machine translation appears to be an increasingly used alternative to human translation. The aim of this article is to compare human translation with the post-editing of machine translation through an experiment on students at master's degree level in the specialization of Translating and Interpreting in a language combination with English at Comenius University and to discover which process provides better and more time-efficient translations. The results acquired through a quality assessment of the submitted translations were analysed based on a temporal aspect of effort and quality. At the end of the article, the findings of the analyses are interpreted and the results are explained to reveal the strengths and weaknesses of human translation and the post-editing of machine translation.

**Keywords:** machine translations, post-editing of machine translations, human translations, post-editing, LQA

### Úvod

Strojový preklad sa stal súčasťou prekladateľského trhu na celom svete. V kombinácii s posteditáciou strojového prekladu môže byť v istých jazykových kombináciách a najmä odvetviach odborného prekladu alternatívou k humánnemu prekladu (Tatsumi, 2009; Plitt a Masselot, 2010; Green a kol., 2013).

V tomto príspevku budeme porovnávať humánny preklad s posteditáciou strojového prekladu a zisťovať, ktorý z týchto dvoch procesov je úspešnejší na úrovni rýchlosti práce a kvality výsledného prekladu.

Na úvod zhrnieme základné teoretické východiská strojového prekladu a posteditácie strojového prekladu. Predstavíme tiež Memsource Translate, funkciu spoločnosti Memsource<sup>1</sup> na spravovanie nástrojov strojového prekladu, pričom predstavíme hodnotenie kvality prekladu, ktoré sme vo výskume využívali. Väčšia časť príspevku sa bude venovať vykonanému experimentu so študentmi 1. a 2. ročníka z Univerzity Komenského v magisterskom stupni štúdia v odbore filológia so zameraním na prekladateľstvo a tlmočníctvo v jazykovej kombinácii s anglickým jazykom; a analýze výsledkov prostredníctvom kvalitatívnej a časovej analýzy.

## 1 Strojový preklad

V súčasnosti si už nedokážeme predstaviť prekladateľa, ktorý by nepracoval na počítači. Technológie ako nástroje CAT, terminologické databázy či pamäte uľahčujú translačný proces. Čoraz viac sa však spomína strojový preklad, „mechanický a automatický proces prekladania textu z jedného prirodzeného jazyka do druhého“ (Munková, 2013, s. 16).

Hlavným cieľom tohto úplného zautomatizovania translačného procesu počítačom je urýchliť prekladateľom prácu a naplniť tak rastúci dopyt na prekladateľskom trhu (Hutchins a Somers, 1992). Za posledných desať rokov totiž dopyt po prekladateľských službách vzrástol tak významne, že preyšuje ponuku poskytovateľov prekladateľských služieb (Hudecová a kol., 2021).

Strojový preklad je vo všeobecnosti považovaný za rýchly, lacný (napr. Prekladač Google je pre bežných používateľov dokonca zadarmo) a dostupný (strojový preklad sa okrem online prekladačov ako Prekladač Google nachádza aj v neplatených verziách niektorých nástrojov CAT).

Aj napriek týmto výhodám strojový preklad nie je dokonalý a nedá sa naň stopercentne spoľahnúť, čo sa dokázalo aj v článku *Porovnanie strojového a humánneho prekladu terminológie*, v ktorom sa porovnával humánny preklad a neurónový strojový preklad Prekladača Google v štyroch rôznych odborných oblastiach. Aj keď boli zaznamenané pozitívne výsledky, objavila sa aj chybovosť (v priemere 10,5%), najmä pri preklade

---

<sup>1</sup> Aktuálne sa spoločnosť Memsource už nazýva Phrase a nástroj Memsource Translate zasa Phrase Translate. Vzhľadom k tomu, že premenovanie nastalo krátko pred publikovaním príspevku, používajú sa ešte staré názvy.

viacvýznamových termínov. Záverom bolo, že ak chceme dosiahnuť vysoko kvalitný preklad pomocou strojového prekladu, strojový preklad si vyžaduje posteditáciu (Pavlíková, 2021).

## 1.1 Neurónový strojový preklad

Neurónový strojový preklad (NMT, z angl. *Neural Machine Translation*) má aktuálne najväčší potenciál na prekladateľskom trhu, keďže ide o technológiu, ktorá sa oproti starším typom strojového prekladu „dokáže naučiť priamo zmapovať vstupný text a prekódovať ho do výstupného textu“ (Wu a kol., 2016, s. 1).

Geneová (2019) predstavila ďalšie dôvody, prečo je neurónový strojový preklad efektívnejší oproti iným typom strojového prekladu. Systém neurónového strojového prekladu:

- sa dokáže naučiť zložité vzťahy medzi oboma jazykmi, s ktorými pracuje;
- berie do úvahy celé vety, pričom chápe vzťahy medzi slovami aj napriek väčšej vzdialenosti v rámci vetnej jednotky;
- hodnotí plynulosť na úrovni celej vety.

Aj napriek svojej efektívnosti však neurónový strojový preklad vykazuje isté nedostatky, ako napríklad neúplnosť prekladov pri dlhších textoch, pomalé tréningovanie systému na veľkom korpuse a problém s prekladom menej frekventovaných slov (Wu a kol., 2016).

### 1.1.1 Neurónový strojový preklad a posteditácia strojového prekladu

V praxi sa strojový preklad často kombinuje s posteditáciou strojového prekladu. „Na zabezpečenie publikovateľnej kvality prekladu musí byť výstup strojového prekladu ešte stále posteditovaný ľuďmi“ (Tatsumi, 2010, s. 3).

Posteditácia strojového prekladu sa využíva aj v kombinácii s neurónovým strojovým prekladom a v súčasných štúdiách sa dokazuje zvýšenie produktivity posteditorov využívajúcich posteditáciu neurónového strojového prekladu v porovnaní s prekladateľmi využívajúcimi prekladateľskú pamäť (Sánchez-Gijón, Moorkens a Way, 2019; Läubli a kol., 2019).

V praxi však posteditácia neurónového strojového prekladu dokáže vytvoriť problémy. Keďže výstupy neurónového strojového prekladu sú čoraz plynulejšie a prirodzenejšie, posteditor si musí dávať pri práci väčší pozor, aby správne identifikoval a opravil chyby. Okrem toho systém neurónového strojového prekladu pracuje v rámci kontextu jednej

vety. Pri práci na dlhšom texte musia teda posteditori sledovať a upravovať konzistentnosti naprieč celým textom (Geneová, 2019).

Zdá sa však, že kombinácia posteditácie s neurónovým strojovým prekladom je aj napriek týmto problémom výhodná, keďže neurónový strojový preklad má oproti starším typom strojového prekladu, ako sme už spomenuli, značné výhody a posteditácia pomôže výstupný text pozdvihnúť na publikovateľnú úroveň porovnateľnú s humánnym prekladom.

## 2 Nástroje strojového prekladu

So strojovým prekladom sa prekladatelia stretávajú pri používaní nástrojov CAT. Okrem zabudovaných prekladačov v nástrojoch CAT existujú aj voľne dostupné prekladače pre bežných používateľov jazyka. V našom experimente sme použili konkrétne dva nástroje strojového prekladu, a to Memsourc Translate, nástroj spoločnosti Memsource, ktorý pomáha používateľovi vybrať ten najvhodnejší prekladač na konkrétny projekt, a Prekladač Google ako nástroj neurónového strojového prekladu volený na jazykové kombinácie s anglickým a slovenským jazykom.

### 2.1 Memsource Translate

Memsource Translate od spoločnosti Memsource slúži na spravovanie nástrojov strojového prekladu. Spoločnosť Memsource vedie online cloudový nástroj CAT, ktorý v súčasnosti patrí medzi tri najpoužívanejšie prekladové systémy, najmä vďaka svojej dostupnosti cez webový prehliadač a jednoduchému používateľskému rozhraniu (Dengová, 2020).

Cieľom nástroja Memsource Translate je uľahčiť používateľom výber z veľkého množstva nástrojov strojového prekladu dostupných na trhu. Aktuálne Memsource Translate vyberá spomedzi troch online prekladačov – Amazon Translate, Microsoft Translator a Prekladač Google. Algoritmus Memsource Translate volí spomedzi týchto troch nástrojov najvhodnejší primárne na základe jazykového páru, no vďaka aktualizáciám z roku 2020 dokáže na základe kľúčových slov rozlišovať aj viaceré odborné oblasti a vybrať teda vhodný prekladač aj podľa toho, či ide o právny, technický alebo medicínsky text (porov. Frívaldský, 2020 a Pavlíková, 2021).

Pri projektoch s angličtinou ako zdrojovým jazykom a slovenčinou ako cieľovým jazykom s textami z viacerých oblastí Memsource Translate primárne volí Prekladač

Google, ktorý funguje na základe neurónového strojového prekladu, čo znamená, že ho primárne pre túto jazykovú kombináciu vyhodnocuje ako najvhodnejší.

### 3 Posteditácia strojového prekladu

Posteditácia strojového prekladu (PEMT, z angl. *Post-editing of Machine Translation*) je jedným z procesov, ktorý slúži na úpravu výstupu strojového prekladu.

V slovenskom jazyku Absolon (2018) rozlišuje tri pojmy spojené s týmto procesom: pojem *posteditácia* zastupuje výsledný produkt, pojem *posteditovanie* samotný proces používania strojového prekladu na vytvorenie vhodného prekladu v cieľovom jazyku a *posteditácia strojového prekladu* je združujúci termín pre produkt (posteditácia) a proces (posteditovanie). V praxi sa však často táto terminológia nedodržiava, keďže pojem *posteditácia strojového prekladu* je zdĺhavý. Aj my preto vo zvyšku článku používame výraz *posteditácia* v zmysle posteditácie strojového prekladu.

Posteditácia strojového prekladu nemá len viacero možných pojmov, ale aj samotných definícií, spomedzi ktorých predstavíme obsiahlejšiu definíciu od uznávaného odborníka z tejto oblasti, Mesa-Laoa (2013, s. 6): „Posteditáciu strojového prekladu možno zadefinovať ako kontrolu vopred preloženého textu strojovým prekladačom, pri ktorej posteditor opraví možné chyby, aby boli dodržané stanovené kritériá kvality pri čo najmenšom počte úprav.“

Práve stanovenie adekvátnych kritérií kvality a dodržiavanie zásad (ako napríklad spomínané dožadovanie sa čo najmenšieho počtu úprav), ktoré sa v definícii spomínajú, sú pri posteditácii kľúčové, no stále v prekladateľskom svete nejednotné (Hu a Cadwell, 2016).

#### 3.1 Výhody a nevýhody posteditácie strojového prekladu

Posteditácia strojového prekladu prišla na prekladateľský trh ako rýchlejšia, lacnejšia a efektívnejšia náhrada humánneho prekladu. Efektivita posteditácie oproti prekladu sa potvrdila vo viacerých výskumoch. Podľa Robertovej (2013) dokáže prekladateľ vďaka posteditácii navýšiť priemerný počet preložených slov za deň z 2 000 na 3 500 slov. Guerberofová Arenasová (2010) dokonca uvádza nárast až na 5 000 slov. Dodáva však, že nárast počtu slov sa môže líšiť v závislosti od typu textu či skúseností posteditora.

Po rýchlejšej posteditácii je dopyt aj vďaka rastu lokalizačného priemyslu a globalizácie, keďže mnohé podniky v súčasnosti prechádzajú na mnohojazykovú expanziu, vďaka

čomu agentúry poskytujúce lokalizačné služby zaznamenávajú 30 až 50 % ročný nárast v počte objednávok (Allen, 2003).

Tento vzrastajúci dopyt po posteditácii potvrdila aj prieskumná agentúra Common Sense Advisory v roku 2016 predpoveďou, že dopyt po posteditačných službách bude rásť rýchlejšie ako v rámci ktoréhokoľvek iného segmentu jazykového priemyslu. Prostredníctvom predošlých prieskumov totiž zistili, že poskytovatelia jazykových služieb, ktorí implementovali používanie strojového prekladu medzi rokmi 2013 a 2015, rástli takmer 3,5-krát rýchlejšie (Common Sense Advisory, 2016).

V neposlednom rade rastu posteditácie pomáha zmena požiadaviek od klientov na cieľový text. V súčasnosti sa požiadavky na kvalitu prekladu zmiernili a mnohé spoločnosti si objednávajú preklad na interné účely, ktorý sa nikam nepublikuje, a nevyžaduje teda vysokú kvalitu humánneho prekladu (Bubnic, 2022).

Aj napriek spomenutým výhodám a popularite posteditácie strojového prekladu ide stále o nový proces v prekladateľskom odvetví (prvýkrát sa spomenul v 50. rokoch minulého storočia), a preto má aj nedostatky. Podľa Geneovej (2019) existuje niekoľko výziev, ktoré sa pri posteditácii strojového prekladu ešte musia vyriešiť:

- neexistencia všeobecných zásad a pravidiel pre posteditáciu,
- nejednotné kompetencie a požiadavky na posteditora,
- otázka platu pre posteditorov.

Okrem toho Doherty a Gaspari (2013) ako ďalšiu nevýhodu uvádzajú nedostatok učebných a tréningových materiálov pre posteditorov, čo spôsobuje problémy uvedené vyššie. Keďže sa posteditácia strojového prekladu často nevyučuje ako samostatný odbor, chýbajú učebné materiály a všeobecné zásady pre posteditáciu, ktoré by dokázali lepšie stanoviť jednotné kompetencie a požiadavky na posteditora a zároveň aj vyriešiť otázku platu pre posteditorov.

### 3.2 Typy posteditácie

Ako sme už uviedli, posteditácia býva vo všeobecnosti rýchlejšia ako humánny preklad. Existujú však viaceré typy posteditácie, ktoré môžu okrem rýchlosti ovplyvniť aj iné faktory, ako napríklad počet úprav či cena.

Vo všeobecnosti sa pri posteditácii spomínajú dva základné typy. Allen (2003) však predstavil tri typy posteditácie:

1. **Rýchla posteditácia strojového prekladu** (známa aj ako *neúplná* alebo *čiasťočná*, z angl. *Light/Rapid/Fast PEMT*): využíva sa pri textoch na interné účely. Hlavnou úlohou posteditora je urobiť čo najmenej úprav a ponechať zachovaný najmä obsah textu, pričom si nemusí všímať štylistické nedokonalosti.
2. **Úplná posteditácia strojového prekladu** (známa aj ako *publikovateľná*, z angl. *Full/Conventional PEMT*): využíva sa pri preklade textov, ktoré sa budú publikovať, a teda vyžaduje cieľový text vysokej kvality porovnateľnej s humánnym prekladom. Očakáva sa preto viacero úprav a okrem zachovania obsahu textu je dôležité aj dosiahnutie správnej syntaxe, gramatiky a interpunkcie.
3. **Minimálna posteditácia strojového prekladu** (z angl. *Minimal PEMT*): rovnako ako pri rýchlej posteditácii, aj pri tomto type sa posteditor snaží urobiť čo najmenší počet úprav, no pracuje s textami, ktoré sa využívajú aj na externú komunikáciu.

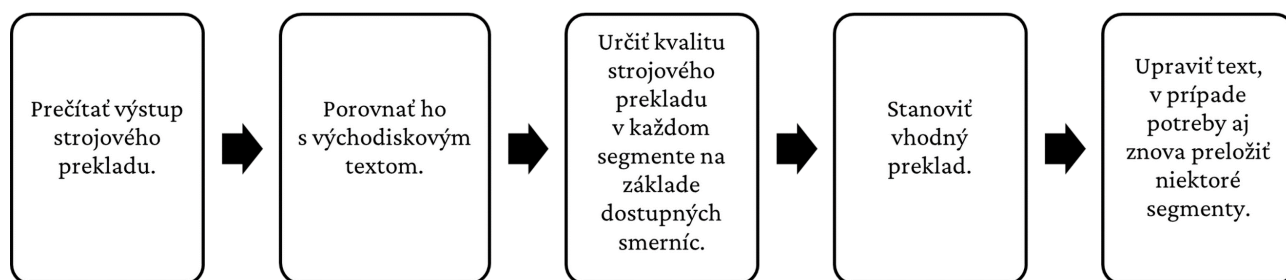
V našom experimente budeme pracovať s úplnou posteditáciou strojového prekladu, pričom sa teda naša vzorka bude snažiť vytvoriť výstupný preklad porovnateľnej kvality s humánnym prekladom.

### 3.3 Proces posteditácie

Doherty a Gaspari (2013) uvádzajú, že sa pri posteditácii, konkrétne pri úplnej posteditácii strojového prekladu, pracuje s tromi textami:

1. východiskový text,
2. nespracovaný výstup strojového prekladu,
3. posteditovaný výstup strojového prekladu/cieľový text.

Keďže ide o prácu s tromi textami, celkový posteditačný proces zahŕňa podľa modelu organizácie TAUS (2010) až päť krokov:



Obrázok 1 – *Model posteditačného procesu*<sup>2</sup>

Podľa tohto modelu musí posteditor pred samotnou úpravou textu vykonať štyri kroky. Celý tento proces prebieha na úrovni segmentov. Posteditori totiž väčšinou pracujú v nástrojoch CAT, ktoré text na základe vopred stanovených pravidiel text segmentujú. Aj keď segmentácia pomáha ponechať text prehľadný, posteditor si musí dávať pozor, aby text vnímal ako celok a upravil všetky nekonzistencie nielen na úrovni segmentu, ale aj celého textu.

Podľa Kringsa (2001) posteditor pri práci vykonáva viaceré typy úkonov, ako napríklad úkony spojené s východiskovým textom (čítanie východiskového textu a všímanie si jeho prvkov), so strojovým prekladom, s prácou so zdrojmi či písaním na počítači. Najväčšia časť úkonov sa však zaraďuje do kategórie úkonov spojených s tvorbou cieľového textu, kde posteditor pracuje s dostupnými posteditačnými smernicami, aby vytvoril adekvátny cieľový text. Geneová (2019, s. 10) zhrnula najčastejšie zmeny, ktoré vo výstupe strojového prekladu posteditor vykonáva: „oprava interpunkcie, doplnenie veľkých začiatkových písmen, zmena slovosledu, úprava gramatických zhôd v rode a čísle, zmeny vo formátovaní, pridávanie vynechaných výrazov alebo odstraňovanie nadbytočných slov a opätovné preloženie slov alebo fráz nehodiacich sa do cieľového textu.“

### 3.4 Posteditor

Posteditácia strojového prekladu vytvorila na prekladateľskom trhu nové pracovné pozície v rámci skupiny poskytovateľov jazykových služieb. Mesa-Lao (2013) uvádza, že množstvo prekladateľských agentúr stále využíva na posteditáciu prekladateľov na voľnej nohe. No nie každý prekladateľ zvládne prácu posteditora. Skúsení prekladatelia si totiž ťažšie zvykajú na vytváranie cieľového textu s nižšou úrovňou kvality v porovnaní s prekladom (Allen, 2003).

<sup>2</sup> Model posteditačného procesu preložila autorka príspevku.



Nielenže sa od seba posteditačný a translačný proces odlišujú, keďže pri preklade sa pracuje iba s dvoma textami namiesto troch, posteditori musia mať odlišné kompetencie a vynaložiť inú mieru úsilia pri práci. Posteditori musia totiž mať znalosti o strojovom preklade a vedieť identifikovať chyby, ktoré strojový preklad robí. Oproti prekladateľom musia byť taktiež rozhodnejší a ľahšie prekonávať neistotu pri problémoch, aby dokázali v porovnaní s prekladateľmi rýchlejšie vytvoriť cieľový text podľa klientových požiadaviek a očakávaní (Torrejón a Rico, 2012).

### 3.5 Posteditačné smernice

Posteditačné smernice slúžia ako návod na posteditáciu, vďaka ktorému dokážu posteditori splniť očakávania klienta a vytvoriť cieľový text adekvátnej kvality. V súčasnosti sa na trhu prejavuje nejednotnosť týchto smerníc, keďže každá agentúra a organizácia si zvykne vytvoriť vlastné smernice na interné použitie (Hu a Cadwell, 2016).

V rámci experimentu sme využili voľne dostupné posteditačné smernice organizácie TAUS úplnej posteditácie strojového prekladu.

#### **TAUS Machine Translation Post-editing Guidelines (2016)**<sup>3</sup>

- Snažte sa o gramaticky, syntakticky a sémanticky správny preklad.
- Postarajte sa o to, aby bola kľúčová terminológia správne preložená a v preklade neboli použité termíny, ktoré klient označil za nevhodné.
- Postarajte sa o to, aby v preklade neboli vynechané údaje alebo informácie navyše.
- Upravte akýkoľvek nevhodný alebo kultúrne neprijateľný obsah.
- Využite čo najviac nespracovaného výstupu strojového prekladu.
- Riadte sa základnými pravidlami pravopisu aj pri úprave interpunkcie a znamienok.
- Uistite sa, že bolo použité správne formátovanie.

Tieto posteditačné smernice zahrňujú viaceré kategórie, na ktoré si musí dávať posteditor pozor. Okrem vytvorenia adekvátneho textu po gramatickej, syntaktickej a sémantickej stránke sa v nich uvádza aj výber vhodnej terminológie, zadávanie interpunkcie a znamienok podľa pravidiel pravopisu či správne formátovanie.

---

<sup>3</sup> Posteditačné smernice od organizácie TAUS preložila autorka príspevku. Ich pôvodné znenie je dostupné na webovej stránke: <https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines>.

### 3.6 Výučba a súčasné využívanie posteditácie strojového prekladu na Slovensku

Posteditácia strojového prekladu sa v prekladateľskom svete skloňuje čoraz viac. O'Brienová (2002) uvádza, že vďaka výučbe posteditácie strojového prekladu by sa absolventi naučili pracovať so strojovým prekladom a osvojili by si posteditačné zručnosti a kompetencie. Okrem toho by výučba pomohla uspokojiť rastúci dopyt po prekladateľských službách predstavením rýchlejšej alternatívy a v neposlednom rade by pomohla zjednotiť posteditačné smernice.

Čoraz viac univerzít začína ponúkať výučbu posteditácie v rámci predmetov venovaných počítačom podporovanému prekladu, strojovému prekladu či lokalizácii (napr. Kabát, 2022). Na Slovensku však na svojich stránkach spomína posteditáciu iba Univerzita Konštantína Filozofa v Nitre. Čo sa týka využívania posteditácie strojového prekladu v praxi, aj napriek vysokému dopytu po strojovom preklade vo svete, na Slovensku „posteditácia strojového prekladu ako produkt zatiaľ predstavuje iba malé percento trhu“ (Absolon, 2018, s. 95).

Absolon (2018) predpokladá zvýšenie dopytu po posteditácii aj na domácom trhu, ale očakáva, že tento postup nebude navonok prezentovaný, keďže v súčasnosti stále mnoho klientov vníma strojový preklad negatívne. Tento predpoklad sa začína premieňať do ponuky služieb prekladateľských agentúr. Malé percento z nich však vyslovene v ponuke strojového prekladu spomína posteditáciu a zamieňajú ju za klientom známejšie pojmy ako *revízia*.

## 4 Hodnotenie kvality prekladu

Hodnotenie kvality prekladu (TQA, z angl. *Translation Quality Assessment* aj LQA, z angl. *Linguistic Quality Assessment*) je „proces hodnotenia celkovej kvality hotového prekladu pomocou modelu s vopred určenými hodnotami, ktoré je možné priradiť k množstvu parametrov používaných na účely hodnotenia“ (Korkas, n. d.).

Tento systém hodnotenia kvality sa čoraz viac využíva v prekladateľských agentúrach a veľkých spoločnostiach na hodnotenie kvality prekladu či posteditácie strojového prekladu. Túto kontrolu vykonávajú revízori či korektori pomocou šablóny (zvyčajne v dokumente Excel), do ktorej zadávajú chyby v preklade, možné opravy problémov a komentáre (Kabát, 2022). Hlavným cieľom hodnotenia kvality prekladu je „identifikovať medzery v kvalite a navrhnúť riešenia pre prekladateľov, aby preklad zodpovedal kvalitatívnym potrebám a očakávaniam klienta“ (Finnegan, 2018).

## 4.1 Šablóna na hodnotenie kvality organizácie TAUS

Šablóna organizácie TAUS slúži na hodnotenie humánneho alebo strojového prekladu a posteditácie strojového prekladu akéhokoľvek typu vstupného textu.

Táto šablóna používateľovi pomáha spočítať, identifikovať a kategorizovať chyby v preklade, ohodnotiť kvalitu prekladu známkou, navrhnúť možné riešenia v komentároch, a celkovo tak zlepšiť kvalitu prekladu (TAUS, n. d.).

Šablóna na hodnotenie kvality organizácie TAUS obsahuje úvodný hárok, do ktorého sa uvádzajú základné informácie o projekte, návod, ako sa šablóna používa, popis typológie chýb a úrovni závažnosti chýb a priestor na uvádzanie chýb v jednotlivých segmentoch. Revízor alebo korektor v tomto priestore uvádza východiskový segment, jeho preklad, návrh, ako by mal opravený segment vyzeráť, pričom zvolí typ chyby a úroveň závažnosti chyby. V prípade potreby prekladateľovi nechá komentár. Jednotlivé závažnosti chýb majú rôzne bodové ohodnotenie, ktoré sa spočítava do penalizačných bodov. Po vyplnení šablóny korektor uvidí výsledné skóre, na základe ktorého vie vyhodnotiť, či bol preklad úspešný (podľa počtu chýb a výsledného percenta úspešnosti), alebo mal veľa chýb.

Typológia chýb v šablóne organizácie TAUS obsahuje 8 kategórií, do ktorých je zahrnutých 33 typov chýb obsahujúcich morfológické, lexikálne, štylistické, syntaktické, terminologické, formátovacie, lokalizačné a kultúrne problémy.

Tabuľka 1: *Kategórie a typy chýb v šablóne na hodnotenie kvality organizácie TAUS*<sup>4</sup>

Kategórie	Typy chýb
Správnosť prekladu	Nadbytočná informácia v preklade Chýbajúca informácia v preklade Nesprávny preklad Nadmerný preklad Nedostatočný preklad Nepreložený text Nesprávne schválená zhoda z prekladateľskej pamäte
Plynulosť prekladu	Interpunkcia Pravopis Gramatická chyba Slovná zásoba Nekonzistentnosť obsahu Nekonzistentnosť v odkazoch/bibliografii Kódovanie znakov

<sup>4</sup> Typológiu chýb preložila autorka príspevku.

<b>Terminológia</b>	Nekonzistentnosť s glosárom Nekonzistentnosť v rámci terminológie
<b>Štýl</b>	Nevhodný štýl Štýl porušujúci pokyny spoločnosti Nekonzistentný štýl Štýl porušujúci pokyny tretích strán Neidiomatický štýl
<b>Formátovanie</b>	Dĺžka Lokálne formátovanie Tagy (značky) Text sa nezobrazuje Skrátenie/rozšírenie textu
<b>Lokalizácia</b>	Formát adresy Formát dátumu Formát meny Formát mier a váh Klávesové skratky Formát telefónneho čísla
<b>Pravdivosť prekladu</b>	Špecifický kultúrny odkaz
<b>Iné</b>	Iné

Závažnosť jednotlivých chýb sa označuje piatimi úrovňami:

1. **Závažná chyba:** chyba, ktorá by mohla mať zdravotné, bezpečnostné, právne alebo finančné dôsledky alebo poškodzuje dobré meno spoločnosti, spôsobuje nefunkčnosť aplikácií/zariadení, negatívne ovplyvňuje či skresľuje funkčnosť výrobku alebo služby.
2. **Veľká chyba:** chyba, ktorá môže používateľa zmiast' alebo brániť správne použitiu výrobku či služby z dôvodu veľkej zmeny významu alebo preto, že sa chyba vyskytuje vo viditeľnej alebo dôležitej časti obsahu.
3. **Malá chyba:** chyba, ktorá nevedie k nesprávne významu či obsahu, ale znižuje štylistickú kvalitu, plynulosť alebo jasnosť prekladu.
4. **Neutrálna chyba:** dodatočná informácia alebo zmena, ktorá sa musí ešte vykonať a nepočíta sa preto za chybu. Ide napríklad o zmenu termínu v glosári, o ktorej prekladateľ počas práce nevedel.
5. **Pochvala:** sa používa na vyzdvihnutie šikovného riešenia prekladateľa.

Šablóna organizácie TAUS je jedným z príkladov, ako šablóna na hodnotenie kvality vyzerá. Niektoré prekladateľské agentúry si však šablóny prispôbujú aj na základe typov

textov, ktoré sa u nich prekladajú najčastejšie, či na základe typu jazykovej služby (preklad, posteditácia strojového prekladu, lokalizácia atď.).

## 5 Porovnanie humánneho prekladu a posteditácie strojového prekladu

Na porovnanie humánneho prekladu a posteditácie strojového prekladu sme vykonali experiment, v ktorom sme rozdelili 20 študentov prekladateľstva a tlmočníctva na dve skupiny. Jedna skupina text prekladala a druhá ho posteditovala, pričom obe skupiny boli vopred oboznámené s procesom posteditácie strojového prekladu.

### 5.1 Metodika

Na experimente sa zúčastnilo 20 študentov 1. a 2. ročníka magisterského stupňa odboru filológia so zameraním na prekladateľstvo a tlmočníctvo v jazykovej kombinácii s anglickým jazykom. Mali teda skúsenosti s prekladom, a preto sme sa pri príprave študentov na experiment zamerali skôr na posteditáciu strojového prekladu.

Experiment sa konal na hodine predmetu Lokalizácia herných a neherných softvérov na Katedre anglistiky a amerikanistiky Filozofickej fakulty Univerzity Komenského, pod dohľadom autorky príspevku, cez MS Teams (pre pandemickú situáciu rozvinutú počas letného semestra akademického roku 2020/2021).

Študenti si mali vopred naštudovať pokyny na preklad a posteditáciu, ktoré sme im poskytli, aby sa bližšie oboznámili s požiadavkami na preklad a posteditáciu. Okrem toho dostali aj všeobecné pokyny, v ktorých bol opísaný postup experimentu.

V deň experimentu sme 20 študentov rozdelili na dve skupiny po desať študentov na základe abecedného poradia podľa priezviska. Prvá skupina text prekladala a riadila sa pokynmi na preklad. Druhá skupina text posteditovala a riadila sa pokynmi na posteditáciu. V rámci posteditácie strojového prekladu študenti vykonávali úplnú posteditáciu a riadili sa posteditačnými smernicami od organizácie TAUS z roku 2016 (uvedené vyššie).

Obom skupinám sme prostredníctvom školskej platformy Moodle poskytli súbory, ktoré si po začatí experimentu nahrali do nástroja Memsource. Prvá skupina dostala súbor vo formáte .docx na preklad a druhá skupina dostala súbor vo formáte .xliff na posteditáciu, ktorý obsahoval nespracovaný strojový preklad získaný z nástroja Memsource pomocou zvoleného Prekladača Google. Okrem súborov na preklad a posteditáciu

sme študentom poskytli terminológiu vo formáte .tbx, ktorú si mali nahrať do vlastnej terminologickej databázy. Terminológia obsahovala základné termíny k zvolenej téme, aby sme študentom uľahčili prácu pri vyhľadávaní.

Ako text na preklad/posteditáciu sme študentom vybrali úryvok z americkej brožúry o kompostovaní, konkrétne návod na postavenie domáceho vermikompostéra. Išlo o odborný text marketingového charakteru, keďže priamo oslovoval čitateľa. Návod obsahoval 262 slov a nástroj CAT Memsource ho rozdelil na 26 segmentov.

Po nahratí súborov do nástroja Memsource študenti pracovali na posteditácii alebo preklade. Po skončení súbor vyexportovali a odovzdali ho v bilingválnom súbore späť cez platformu Moodle. Študenti boli v pokynoch taktiež upozornení na to, že sa bude zaznamenávať ich čas, takže boli vyzvaní, aby po dokončení odovzdali výsledný súbor čím skôr, a nedošlo tak k ďalšiemu skresleniu časových výsledkov.

## 5.2 Cieľ a výskumné otázky

Cieľom výskumu bolo porovnať humánny preklad s posteditáciou strojového prekladu na danej vzorke študentov, a zistiť tak, ktorý proces bol úspešnejší na úrovni rýchlosti práce a kvality výsledného prekladu. Pri hodnotení kvality chceme odpovedať na nasledujúce otázky:

- Ktorá skupina bola úspešnejšia na základe priemernej chybovosti?
- Ktorá skupina bola úspešnejšia na základe priemerného množstva penalizačných bodov?
- Aké typy chýb prevažne vykonávali prekladatelia a aké typy chýb vykonávali posteditori?

Okrem kvalitatívnej analýzy sa pozrieme aj na časový aspekt úsilia u oboch skupín. V rámci tejto časti budeme teda vyhodnocovať:

- Ktorá skupina vykonala preklad rýchlejšie?

Odpovede na tieto otázky nám v rámci kvalitatívnej a časovej analýzy výsledkov naznačia, ako sa humánny preklad a posteditácia strojového prekladu od seba líšia a aké sú ich silné či slabé stránky.

## 5.3 Kvalitatívna analýza

Hodnotenie kvality odovzdaných prekladov od študentov sme vykonali pomocou šablóny na hodnotenie kvality prekladu od organizácie TAUS.

V rámci procesu penalizácie sme použili bodový systém zadaný v šablóne ako východiskový, pričom organizácia TAUS ponúkala aj iné alternatívy.

Tabuľka 2: Bodový systém penalizácie podľa šablóny organizácie TAUS

<b>Úroveň závažnosti</b>	<b>Počet penalizačných bodov</b>
Závažná chyba	10
Veľká chyba	5
Malá chyba	1

Úspešnosť odovzdaného prekladu sme vyhodnocovali podľa počtu penalizačných bodov, ktoré prekladateľ/posteditor získal.

V kvalitatívnej analýze sa teda zameriame na to, aká bola úspešnosť prekladateľov a posteditorov na základe penalizačných bodov, koľko a aké typy chýb v akých úrovniach závažnosti robili a ako sa v týchto kategóriách od seba odlišovali, pričom uvedieme aj konkrétne príklady najčastejších chýb v rámci oboch skupín.

### 5.3.1 Chybovosť v preklade

V rámci kvalitatívnej analýzy sme najprv skúmali chybovosť, a teda koľko chýb urobili prekladatelia a posteditori. Chyby sme rozdelili na základe úrovne závažnosti chýb. Ako závažnú chybu sme hodnotili najmä chyby v číselných údajoch a konverzii jednotiek, ktoré by mohli viesť k neúspešnosti návodu. Medzi veľké chyby sme radili typy chýb, ktoré menili význam východiskového textu a taktiež by mohli viesť k nefunkčnosti návodu (išlo najmä o kategórie správnosti prekladu, plynulosti prekladu, terminológie a lokalizácie). Malé chyby sa väčšinou týkali kategórie štýlu, formátovania a plynulosti prekladu, do ktorej sme zaradili aj chyby pri písaní interpunkcie a matematických symbolov.

Tabuľka 3: *Priemerná chybovosť na základe úrovne závažnosti chýb*

<b>Úroveň závažnosti</b>	<b>Priemerná chybovosť</b>	
	<b>Prekladatelia</b>	<b>Posteditori</b>
Závažná chyba	0,4	0,0
Veľká chyba	4,7	7,2
Malá chyba	2,7	2,8
<b>Spolu</b>	<b>7,8</b>	<b>10,0</b>

Po spočítaní chýb u každého prekladateľa a posteditora a vytvorení priemeru sme väčšiu chybovosť zaznamenali u posteditorov, ktorí robili v priemere 10 chýb. Prekladatelia mali priemernú chybovosť 7,8. Pri bližšej analýze týchto výsledkov si môžeme všimnúť, že v kategórii závažnej chyby, ktorá najviac ovplyvňuje úspešnosť prekladu, mali posteditori priemernú chybovosť 0, no prekladatelia 0,4. Prekladatelia teda v porovnaní s posteditormi vykonali viac závažných chýb, čím mohli negatívne ovplyvniť kvalitu prekladaného návodu. Posteditori však vykonávali značne viac veľkých chýb, priemerne 7,2, pričom prekladatelia mali priemernú chybovosť 4,7. Posteditori teda taktiež robili chyby, ktoré negatívne ovplyvňovali funkčnosť návodu na výrobu vermikompostéra. V poslednej kategórii malých chýb, kde sme prevažne hodnotili nezrovnalosti v tagoch, interpunkcii a matematických symboloch, boli na tom posteditori a prekladatelia podobne.

### 5.3.2 Úspešnosť prekladu

Úspešnosť prekladu sme v oboch skupinách hodnotili na základe počtu získaných penalizačných bodov.

Tabuľka 4: *Priemerný počet penalizačných bodov*

<b>Skupina</b>	<b>Priemerný počet penalizačných bodov</b>
Prekladatelia	30,2
Posteditori	38,8

Na porovnanie výsledkov sme vytvorili priemer v oboch skupinách. Keďže penalizačné body úzko súviseli s chybovosťou v preklade, aj v úspešnosti prekladu na tom boli lepšie prekladatelia. V priemere získali 30,2 penalizačných bodov, pričom posteditori v priemere získali 38,8 penalizačných bodov.



### 5.3.3 Typológia chýb

Poslednou časťou kvalitatívnej analýzy je vyhodnotenie počtu chýb na základe typológie chýb šablóny organizácie TAUS, pričom sa pozrieme aj na podkategórie chýb a jednotlivé príklady.

Šablóna organizácie TAUS obsahuje osem základných kategórií chýb. Pri analýze sme však v dvoch kategóriách, pravdivosť prekladu a iné, nezaznamenali žiadnu chybu, a preto budeme ďalej pracovať iba so šiestimi hlavnými kategóriami chýb.

Tabuľka 5: *Priemerná chybovosť na základe typológie chýb*

Kategória typológie chýb	Priemerná chybovosť	
	Prekladatelia	Posteditori
Správnosť prekladu	2,8	3,5
Plynulosť prekladu	3,4	4,4
Terminológia	0,7	0,9
Štýl	0,1	0
Formátovanie	0,2	0,9
Lokalizácia	0,6	0,3
Pravdivosť prekladu	0	0
Iné	0	0

Zo šiestich kategórií vykázali posteditori väčšiu chybovosť až v štyroch kategóriách, a to správnosť prekladu, plynulosť prekladu, terminológia a formátovanie. Najväčší rozdiel sme zaznamenali v kategórii plynulosti prekladu, v ktorej mali prekladatelia priemernú chybovosť 3,4 a posteditori až 4,4. Táto kategória však ponímala široké spektrum chýb zastúpené jednotlivými podkategóriami, takže vyhodnocovaniu konkrétnych chýb sa budeme venovať v hlbšej analýze. Môžeme však konštatovať, že prekladatelia sa preukázali plynulejším výsledným prekladom.

Vo zvyšných dvoch kategóriách vykázali väčšiu priemernú chybovosť prekladatelia. V kategórii štýlu sme zaznamenali u prekladateľov priemernú chybovosť 0,1, pričom posteditori v tejto kategórii neurobili žiadnu chybu. Aj kategóriu lokalizácie zvládli posteditori lepšie s priemernou chybovosťou 0,3, pričom prekladatelia mali priemernú chybovosť 0,6.

Aby sme zistili, v čom a ako sa odlišovali chyby, ktoré vykonávali posteditori a prekladatelia, v ďalšej časti príspevku sa pozrieme na podkategórie jednotlivých kategórií chýb a konkrétne príklady.

### 5.3.3.1 Správnosť prekladu

Kategória správnosti prekladu upozorňuje na prípady, v ktorých preklad neodzrkadľuje východiskový text, a teda správne neprenáša jeho význam či účel. Táto kategória zahŕňa sedem typov chýb: nadbytočná informácia v preklade, chýbajúca informácia v preklade, nesprávny preklad, nadmerný preklad, nedostatočný preklad, nepreložený text a nesprávne schválená zhoda z prekladateľskej pamäte.

Tabuľka 6: *Priemerná chybovosť v kategórii správnosti prekladu*

Typy chýb	Priemerná chybovosť	
	Prekladatelia	Posteditori
Nadbytočná informácia v preklade	0,1	0
Chýbajúca informácia v preklade	0,4	0,3
Nesprávny preklad	1,7	3,2
Nadmerný preklad	0	0
Nedostatočný preklad	0,3	0
Nepreložený text	0,3	0
Nesprávne schválená zhoda z prekladateľskej pamäte	0	0

Pri bližšej analýze kategórie správnosti prekladu badáme, že aj keď je celková priemerná chybovosť tejto kategórii vyššia u posteditorov, v jednotlivých typoch chýb je tomu tak iba pri nesprávnom preklade, kde mali posteditori priemernú chybovosť 3,2 a prekladatelia 1,7.

Vyššiu priemernú chybovosť prekladatelia vykázali až v štyroch podkategóriách, a to nadbytočná informácia v texte, chýbajúca informácia v texte, nedostatočný preklad a nepreložený text. Posteditori pritom pochybili iba v dvoch podkategóriách, a to chýbajúca informácia v texte a nesprávny preklad. Môžeme teda zhodnotiť, že v kategórii správnosti prekladu prekladatelia vykázali rozmanitejšiu chybovosť.

### 5.3.3.1.1 Nadbytočná informácia v texte

Tabuľka 7: *Príklady chýb v podkategórii nadbytočnej informácie v texte*

Skupina	Prekladatelia	Segment	3
Východiskový text	You can compost food scraps indoors using a worm bin!		
Preklad	Pomocou vermikompostéru môžete kompostovať zvyšky z jedál môžete aj v interiéri!		
Navrhovaný preklad	Pomocou vermikompostéru môžete kompostovať zvyšky z jedál aj v interiéri!		

Pri hodnotení kvality prekladu sme zaznamenali iba jednu chybu typu nadbytočnej informácie v texte. Urobil ju jeden prekladateľ v treťom segmente, pričom išlo skôr o duplicitné slovo v jednej vete, než nadbytočnú informáciu. Prekladateľ mal vo svojom preklade uvedené dvakrát slovo *môžete*, pričom malo byť uvedené iba raz. Táto chyba bola zrejme urobená z nepozornosti, keď prekladateľ menil slovosled a zabudol vymazať predošlé usporiadanie slov.

U posteditorov sme chybu tohto typu nezaznamenali.

### 5.3.3.1.2 Chýbajúca informácia v texte

Tabuľka 8: *Príklady chýb v podkategórii chýbajúcej informácie v texte*

Skupina	Prekladatelia	Segment	8
Východiskový text	Dimensions should be approximately 12 x 12 x 12 inches (one cubic foot) but do not have to be precise.		
Preklad	Rozmery by mali byť približne 30 x 30 x 30 cm, no nemusí sa to presne zhodovať.		
Navrhovaný preklad	Rozmery by mali byť približne 30 x 30 x 30 cm (s objemom zhruba 28 litrov), no nemusí sa to presne zhodovať.		
Skupina	Posteditori	Segment	15
Východiskový text	These holes will provide oxygen to the worms and other decomposer organisms in the bin.		
Preklad	Tieto otvory poskytnú kyslík dážďovkám a iným organizmom v koši.		
Navrhovaný preklad	Tieto otvory poskytnú kyslík dážďovkám a rozkladacím organizmom v nádobe.		

Pri hodnotení kvality prekladu sme u prekladateľov zaznamenali v tejto podkategórii štyri chyby, pričom tri z nich uvádzame vo vybranom príklade. Niektorí prekladatelia totiž vynechali informáciu v zátvorke v ôsmom segmente. Keďže rozmery sú tam uvedené, pravdepodobne usúdili, že objem je nadbytočná informácia. V návodoch je však nutné dbať na číselné údaje a bez konzultácie so zadávateľom prekladu či klientom nie je vhodné odstraňovať napríklad objem nádoby, ako sa stalo v tomto prípade.

U posteditorov sme v tejto podkategórii zaznamenali tri chyby, pričom okrem rovnakej chyby, akú robili prekladatelia, jeden posteditor vynechal preklad termínu *decomposer*, čo robili aj niektorí prekladatelia. V kategórii chýbajúcej informácie v texte sa od seba prekladatelia a posteditori teda veľmi neodlišovali a robili rovnaké chyby.

### 5.3.3.1.3 Nesprávny preklad

Tabuľka 9: Príklady chýb v podkategórii nesprávneho prekladu

Skupina	Prekladatelia	Segment	10
Východiskový text	Fine screen to keep out pests (optional)		
Preklad	Jemné sito proti škodcom (nepovinné)		
Navrhovaný preklad	Jemná sieťka proti škodcom (nepovinné)		
Skupina	Posteditori	Segment	12
Východiskový text	\{b\>Tray (optional)\<b\}		
Preklad	\{b\>Zásobník (voliteľný)\<b\}		
Navrhovaný preklad	\{b\>Podnos (voliteľný)\<b\}		

V kategórii správnosti prekladu robili prekladatelia najviac chýb typu nesprávneho prekladu. Väčšina chýb sa týkala najmä nesprávne preloženého termínu *screen* v kontexte návodu na zostavenie vermikompostéra. Mnohí prekladatelia ho preložili ako *sito*, niektorí ako *plochu*, *triedič* či *filter*, pričom šlo o *sieťku* proti hmyzu, ktorá mala slúžiť ako prevencia proti škodcom.

Posteditori mali taktiež problémy s prekladom niektorých termínov. Zatiaľ čo prekladatelia si nevedeli poradiť iba so slovom *screen*, u posteditorov bol tento problém rozšírenejší a mnohé slová boli v kontexte preložené nesprávne: slovo *bin* preložené ako *kôš*, aj keď by malo ísť buď o nádobu, alebo vermikompostér (závisí od kontextu), slovo *directions* preložené ako *smery*, aj keď išlo o inštrukcie alebo postup, ako si vyrobiť vermikompostér, a v neposlednom rade aj prípad z uvedeného príkladu, a to slovo *tray*

preložené ako *zásobník*, aj keď išlo o podnos alebo tácku, ktorá slúži na zachytávanie dážďovkového čaju z vermikompostéra. Väčšina týchto nesprávnych prekladov je zapríčinená práve nespracovaným strojovým prekladom, ktorý posteditori následne neopravili.

#### 5.3.3.1.4 Nedostatočný preklad

Tabuľka 10: *Príklady chýb v podkategórii nedostatočného prekladu*

Skupina	Prekladatelia	Segment	26
Východiskový text	Worms can tolerate temperatures a bit beyond this range, but they will be less active.		
Preklad	Dážďovky dokážu znášať aj teploty mierne pod 13°C, no v tom prípade nebudú také aktívne.		
Navrhovaný preklad	Dážďovky dokážu znášať aj teploty mierne mimo tento rozsah, no v tom prípade nebudú také aktívne.		

Chyby v kategórii nedostatočného prekladu vykonávali iba prekladatelia. Posteditori totiž pracovali so strojovým prekladom, ktorý preložil každú informáciu v texte.

V uvedenom príklade prekladateľ chybne upozornil iba na to, že dážďovky znesú nižšie teploty, ako sa odporúča, pričom vynechal informáciu, že dážďovky znesú aj teploty mierne nad najvyššiu odporúčanú hodnotu.

#### 5.3.3.1.5 Nepreložený text

Tabuľka 11: *Príklady chýb v podkategórii nepreloženého textu*

Skupina	Prekladatelia	Segment	24
Východiskový text	Worms prefer temperatures between 55°F and 80°F (13°C and 27°C).		
Preklad	Dážďovky uprednostňujú teplotu medzi 55 °F a 80 °F (13 °C and 27 °C).		
Navrhovaný preklad	Dážďovky uprednostňujú teplotu medzi 55 °F a 80 °F (13 °C a 27 °C).		

Aj v tejto podkategórii vykonávali chyby iba prekladatelia, pretože opäť sa dá skonštatovať, že posteditori pracovali so strojovým prekladom, ktorý preložil všetky informácie v texte, a nebolo teda možné nájsť v ňom nepreloženú časť textu bez toho, aby ju tam manuálne vkladali.

Niektorým prekladateľom sa podarilo nechať nepreložený text v preklade a vo všetkých prípadoch išlo o rovnakú chybu. V 24. segmente nepreložili text v zátvorke. Je pravdepodobné, že ho skopírovali z východiskového textu bez toho, aby si všimli, že tam je spojka *and*. Vo výslednom preklade tak ostala spojka v anglickom jazyku.

### 5.3.3.2 Plynulosť prekladu

Kategória plynulosti prekladu upozorňuje na problémy s obsahom, ale aj formou prekladu. Táto kategória zahŕňa sedem podkategórií: interpunkcia, pravopis, gramatická chyba, slovná zásoba, nekonzistentnosť obsahu, nekonzistentnosť v odkazoch/bibliografii a kódovanie znakov.

Tabuľka 12: *Priemerná chybovosť v kategórii plynulosti prekladu*

Typy chýb	Priemerná chybovosť	
	Prekladatelia	Posteditori
Interpunkcia	1,9	1,7
Pravopis	0,4	0,1
Gramatická chyba	0,9	2,5
Slovná zásoba	0,2	0,1
Nekonzistentnosť obsahu	0	0
Nekonzistentnosť v odkazoch/bibliografii	0	0
Kódovanie znakov	0	0

Celkovo je priemerná chybovosť v kategórii plynulosti prekladu vyššia u posteditorov. Ak sa však pozrieme na rozpísané podkategórie, vidíme, že vyššia priemerná chybovosť u posteditorov sa ukázala iba v podkategórii gramatickej chyby. Posteditori v nej vykázali priemernú chybovosť 2,4, pričom prekladatelia mali priemernú chybovosť iba 0,9. V podkategóriách interpunkcia, pravopis a slovná zásoba spravili viac chýb prekladatelia. Aj v tejto kategórii teda prekladatelia vykázali väčšiu rozmanitosť chýb.

### 5.3.3.2.1 Interpunkcia

Tabuľka 13: *Príklady chýb v podkategórii interpunkcie*

<b>Skupina</b>	<b>Prekladatelia</b>	<b>Segment</b>	<b>22</b>
<b>Východiskový text</b>	If you drilled holes in the bottom on the bin, place a tray under the bin to catch any “leachate”—this is a waste product of the composting process made of excess moisture.		
<b>Preklad</b>	Ak ste do dna nádoby vyvrtali otvory, umiestnite pod nádobu podnos, aby ste zachytili dážďovkový čaj – odpadový produkt kompostovacieho procesu, ktorý je výsledkom prebytočnej vlhkosti.		
<b>Navrhovaný preklad</b>	Ak ste do dna nádoby vyvrtali otvory, umiestnite pod nádobu podnos, aby zachytil dážďovkový čaj – odpadový produkt kompostovacieho procesu, ktorý je výsledkom prebytočnej vlhkosti.		
<b>Skupina</b>	<b>Posteditori</b>	<b>Segment</b>	<b>24</b>
<b>Východiskový text</b>	Worms prefer temperatures between 55°F and 80°F (13°C and 27°C).		
<b>Preklad</b>	Dážďovky dávajú prednosť teplotám medzi 13 ° C a 27 ° C.		
<b>Navrhovaný preklad</b>	Dážďovky preferujú teploty medzi 13 °C a 27 °C.		

Okrem bežných problémov s čiarkami vykazovali prekladatelia chyby aj v písaní jednotiek a pomlčiek. V uvedenom príklade u prekladateľov vidíme nesprávne zapísaný spojovník namiesto pomlčky. Každý prekladateľ, ktorý sa rozhodol napísať túto vetu v rovnakom usporiadaní, v akom je pôvodný text, napísal namiesto pomlčky spojovník. Táto chyba mohla byť spôsobená aj tým, že študenti nie sú zvyknutí po sebe kontrolovať pomlčky. Ak totiž napíšu spojovník v dokumente Word s medzerami okolo znamienka, spojovník sa automaticky zmení na pomlčku. V nástrojoch CAT to takto však nefunguje. Rovnako ako u prekladateľov, aj u posteditorov sme zaznamenali problémy s písaním čiarok či pomlčiek. Ani v skupine posteditorov nikto nenapísal správne pomlčku. Môžeme teda zhodnotiť, že interpunkcia robí študentom vo všeobecnosti problém, nie je to však nič, čo by nemohol opraviť korektor pri revízii. Okrem správneho zápisu pomlčiek mali posteditori problém so správnym zápisom jednotiek. V uvedenom príklade máme nesprávny zápis jednotky stupňa Celzia, ktorý bol spôsobený strojovým prekladom a nebol následne opravený.

### 5.3.3.2 Pravopis

Tabuľka 14: *Príklady chýb v podkategórii pravopisu*

<b>Skupina</b>	<b>Prekladatelia</b>	<b>Segment</b>	<b>10</b>
Východiskový text	Fine screen to keep out pests (optional)		
Preklad	Tekná plocha, ktorá zabráni prístupu škodcom (nepovinné)		
Navrhovaný preklad	Jemná sieťka, ktorá zabráni prístupu škodcom (nepovinné)		
<b>Skupina</b>	<b>Posteditori</b>	<b>Segment</b>	<b>15</b>
Východiskový text	These holes will provide oxygen to the worms and other decomposer organisms in the bin.		
Preklad	Tieto otvory poskytnú kyslík dážďovkám a iným heteretrofným organizmom vo vermikompostéri.		
Navrhovaný preklad	Tieto otvory poskytnú kyslík dážďovkám a iným heterotrofným organizmom vo vermikompostéri.		

Podkategória pravopisu zahŕňala chyby spojené s nesprávnym zápisom slov, ako napríklad preklepy, nesprávne písanie i/y či zmenu hlásky pri spodobovaní. V našom prípade sa týkala najmä preklepov, ktoré robili vo väčšej miere prekladatelia. V uvedenom príklade vidíme zle napísané slovo *tenká*, ktoré mohlo byť opravené kontrolou pravopisu.

U posteditorov sme zaznamenali iba jeden preklep, keďže nespracovaný strojový preklad neobsahoval žiadne preklepy. Nesprávne zapísané slovo *heterotrofný* urobil samotný posteditor prepisom strojového prekladu.

### 5.3.3.3 Gramatická chyba

Gramatika zahŕňa syntax a morfológiu. Medzi gramatické chyby teda patrí nesprávne použitie pádu, zle zvolený tvar pri skloňovaní či nesprávna konštrukcia vety. My sme v rámci tejto podkategórie sledovali aj spisovnosť použitých slov.

Gramatické chyby sa u prekladateľov spájali najmä s nesprávnymi predložkovými väzbami a používaním nespisovných slov či výrazov. V uvedenom príklade môžeme vidieť nespisovný tvar slova *zohnať*. V iných prípadoch prekladatelia tiež používali frázy, ktoré sa v slovenčine považujú za bohemizmy, ako napríklad výraz *o veľkosti* namiesto výrazu *s veľkosťou*.



Tabuľka 15: *Príklady chýb v podkategórii gramatickej chyby*

<b>Skupina</b>	<b>Prekladatelia</b>	<b>Segment</b>	<b>11</b>
Východiskový text	Purchase at a hardware store.		
Preklad	Zohnáte ho v železiarstve.		
Navrhovaný preklad	Zoženiete ho v železiarstve.		
<b>Skupina</b>	<b>Posteditori</b>	<b>Segment</b>	<b>22</b>
Východiskový text	If you drilled holes in the bottom on the bin, place a tray under the bin to catch any “leachate”—this is a waste product of the composting process made of excess moisture.		
Preklad	Ak ste do spodnej časti koša vyvrtali otvory, umiestnite pod kôš podnos, aby zachytili všetok „dážďovkový čaj“ – ide o odpadový produkt z procesu kompostovania vyrobený z prebytočnej vlhkosti.		
Navrhovaný preklad	Ak ste do dna nádoby vyvrtali otvory, umiestnite pod nádobu podnos, aby zachytil dážďovkový čaj – odpadový produkt kompostovacieho procesu, ktorý je výsledkom prebytočnej vlhkosti.		

Posteditori robili v porovnaní s prekladateľmi viac gramatických chýb. Väčšina z nich však súvisela s nespracovaným strojovým prekladom, ktorý bolo treba upraviť tak, aby sa v ňom gramatické chyby nevyskytovali. V uvedenom príklade ponechali viacerí posteditori slovo *zachytiť* v množnom čísle, aj keď z kontextu vyplýva, že sloveso malo byť v jednotnom čísle.

#### 5.3.3.2.4 Slovná zásoba

Tabuľka 16: *Príklady chýb v podkategórii slovnej zásoby*

<b>Skupina</b>	<b>Prekladatelia</b>	<b>Segment</b>	<b>14</b>
Východiskový text	Drill at least 10 quarter-inch holes in the lid.		
Preklad	Do veka vyvrtajte minimálne 10 otvorov s rozmermi 0,6 centimetra.		
Navrhovaný preklad	Do veka vyvrtajte minimálne 10 otvorov s priemerom 0,6 centimetra.		
<b>Skupina</b>	<b>Posteditori</b>	<b>Segment</b>	<b>14</b>
Východiskový text	Drill at least 10 quarter-inch holes in the lid.		
Preklad	Do veka vyvrtajte najmenej 10 otvorov s rozmerom približne 6mm.		
Navrhovaný preklad	Do veka vyvrtajte najmenej 10 otvorov s priemerom približne 6 mm.		

V podkategórii slovnej zásoby robili prekladatelia a posteditori rovnakú chybu. Namiesto výrazu *priemer* volili na vyjadrenie priemeru daných otvorov výraz *rozmer*, ktorý by mohol čitateľa návodu zmiast.

### 5.3.3.3 Terminológia

Kategória terminológie upozorňuje na chyby v terminológii. Táto kategória zahŕňa dva typy chýb: nekonzistentnosť s glosárom a nekonzistentnosť v rámci terminológie.

Tabuľka 17: *Priemerná chybovosť v kategórii terminológie*

Typy chýb	Priemerná chybovosť	
	Prekladatelia	Posteditori
Nekonzistentnosť s glosárom	0	0,2
Nekonzistentnosť v rámci terminológie	0,7	0,7

Na základe tabuľky s typmi chýb v kategórii terminológie si môžeme všimnúť, že v podkategórii nekonzistentnosti v rámci terminológie na tom prekladatelia a posteditori boli rovnako s priemernou chybovosťou 0,7.

V podkategórii nekonzistentnosti s glosárom sme zaznamenali chyby iba v skupine posteditorov, čo znamená, že posteditori nedostatočne opravovali strojový preklad, aby terminológia sedela s poskytnutým glosárom.

#### 5.3.3.3.1 Nekonzistentnosť s glosárom

Tabuľka 18: *Príklady chýb v podkategórii nekonzistentnosti s glosárom*

Skupina	Posteditori	Segment	24
Východiskový text	Worms prefer temperatures between 55°F and 80°F (13°C and 27°C).		
Preklad	Červy dávajú prednosť teplotám medzi 13 °C a 27 °C.		
Navrhovaný preklad	Dážďovky dávajú prednosť teplotám medzi 13 °C a 27 °C.		

Nekonzistentnosti s glosárom sa vyskytli iba u posteditorov. Obe zaznamenané chyby sa týkali pojmu *worm*, ktorý bol nesprávne ponechaný ako *červ*, aj keď podľa glosára sa mal preložiť ako *dážďovka*. Opäť išlo o chybu spôsobenú strojovým prekladom, ktorú následne posteditori nezmenili, pretože sa ním buď nechali ovplyvniť, alebo chybu neopravili z nepozornosti či nedostatočnej finálnej kontroly.

### 5.3.3.3.2 Nekonzistentnosť v rámci terminológie

Tabuľka 19: Príklady chýb v podkategórii nekonzistentnosti v rámci terminológie

Skupina	Prekladatelia	Segment	23
Východiskový text	After you make or buy a bin, store it in a good location for both you and the worms.		
Preklad	Keď už ste hrdým majiteľom zakúpeného alebo vyrobeného domáceho kompostéra, umiestnite ho na miesto, ktoré vyhovuje vám ale je vhodné aj pre dážďovky.		
Navrhovaný preklad	Keď už ste hrdým majiteľom zakúpeného alebo vyrobeného domáceho vermikompostéra, umiestnite ho na miesto, ktoré vyhovuje vám, ale je vhodné aj pre dážďovky.		
Skupina	Posteditori	Segment	20
Východiskový text	Use non-toxic glue to attach fine screen over holes (optional).		
Preklad	Pomocou netoxického lepidla pripevnite jemné sito cez otvory (voliteľné).		
Navrhovaný preklad	Pomocou netoxického lepidla pripevnite jemnú sieťku cez otvory (voliteľné).		

V podkategórii nekonzistentnosti v rámci terminológie mali prekladatelia problém najmä s termínom *vermikompostér*, namiesto ktorého uvádzali nesprávne termín *kompostér*.

Posteditori mali okrem rovnakých nekonzistentností ako u prekladateľov aj problém pri termínoch *screen* či *tray*. V uvedenom príklade môžeme vidieť, že posteditor preložil *screen* ako *sito*, aj keď v predošlých segmentoch ho prekladal ako *obrazovka*. Táto nekonzistentnosť bola prítomná už v strojovom preklade a nebola posteditorom opravená.

### 5.3.3.4 Štýl

Kategória štýlu upozorňuje na stylistické problémy. Táto kategória zahŕňa päť typov chýb: nevhodný štýl, štýl porušujúci pokyny spoločnosti, nekonzistentný štýl, štýl porušujúci pokyny tretích strán a neidiomatický štýl.

Pri bližšej analýze kategórie štýlu si môžeme všimnúť, že z piatich typov chýb boli zaznamenané chyby iba v jednej podkategórii, a to nekonzistentný štýl. V tejto podkategórii pochybili iba prekladatelia.

Tabuľka 20: *Priemerná chybovosť v kategórii štýlu*

Typy chýb	Priemerná chybovosť	
	Prekladatelia	Posteditori
Nevhodný štýl	0	0
Štýl porušujúci pokyny spoločnosti	0	0
Nekonzistentný štýl	0,1	0
Štýl porušujúci pokyny tretích strán	0	0
Neidiomatický štýl	0	0

#### 5.3.3.4.1 Nekonzistentný štýl

Tabuľka 21: *Príklady chýb v podkategórii nekonzistentného štýlu*

Skupina	Prekladatelia	Segment	20
Východiskový text	Use non-toxic glue to attach fine screen over holes (optional).		
Preklad	Jemný triedič prelepíte cez diery netoxickým lepidlom (nepovinné).		
Navrhovaný preklad	Jemnú sieťku prelepte cez diery netoxickým lepidlom (nepovinné).		

V rámci podkategórie nekonzistentného štýlu sme zaznamenali u prekladateľov jednu chybu. Prekladateľ napísal sloveso *prelepíť* v oznamovacom spôsobe a nie v rozkazovacom, ako ho uvádzal vo zvyšku prekladu a ako by ho bolo vhodné uvádzať práve v texte s návodom.

#### 5.3.3.5 Formátovanie

Kategória formátovania upozorňuje na problémy s formátovaním. Táto kategória zahŕňa päť podkategórií: dĺžka, lokálne formátovanie, tagy (značky), text sa nezobrazuje a skrátenie/rozšírenie textu.

V kategórii formátovania vykázali väčšiu priemernú chybovosť posteditori, ale pochybili iba v podkategórii tagov, pričom prekladatelia opäť vykázali rozmanitejšiu chybovosť a okrem podkategórie tagov pochybili aj v podkategórii lokálneho formátovania.

Tabuľka 22: *Priemerná chybovosť v kategórii formátovania*

Typy chýb	Priemerná chybovosť	
	Prekladatelia	Posteditori
Dĺžka	0	0
Lokálne formátovanie	0,1	0
Tagy (značky)	0,1	0,9
Text sa nezobrazuje	0	0
Skrátenie/rozšírenie textu	0	0

### 5.3.3.5.1 Lokálne formátovanie

Tabuľka 23: *Príklady chýb v podkategórii lokálneho formátovania*

Skupina	Prekladatelia	Segment	1
Východiskový text	{1>INDOOR COMPOSTING<1} with a worm bin		
Preklad	Interiérové kompostovanie s vermikompostérom		
Navrhovaný preklad	{1>INTERIÉROVÉ KOMPOSTOVANIE<1} s vermikompostérom		

V rámci podkategórie lokálneho formátovania sme zaznamenali chybu iba u prekladateľov. Jeden prekladateľ totiž nesledoval formátovanie východiskového textu a výraz *interiérové kompostovanie* napísal s malými písmenami, aj keď bolo pôvodne napísané s veľkými.

### 5.3.3.5.2 Tagy (značky)

V podkategórii tagov sme u prekladateľov zaznamenali iba jednu chybu, keď sa jeden prekladateľ rozhodol v prvom segmente tagy úplne vynechať.

Posteditori mali s tagmi väčší problém. Okrem úplného vynechávania tagov sme zaznamenali aj nesprávny zápis tagov s nadbytočnými medzerami. U posteditorov sa tieto chyby s tagmi vyskytovali už v strojovom preklade.

Tabuľka 24: *Príklady chýb v podkategórii tagov*

Skupina	Prekladatelia	Segment	1
Východiskový text	{1>INDOOR COMPOSTING<1} with a worm bin		
Preklad	Interiérové kompostovanie s vermikompostérom		
Navrhovaný preklad	{1>INTERIÉROVÉ KOMPOSTOVANIE<1} s vermikompostérom		
Skupina	Posteditori	Segment	1
Východiskový text	\{1\>INDOOR COMPOSTING\<1\} with a worm bin		
Preklad	\{1\> VNÚTORNÉ KOMPOSTOVANIE \<1\} s vermikompostérom		
Navrhovaný preklad	\{1\>VNÚTORNÉ KOMPOSTOVANIE\<1\} s vermikompostérom		

### 5.3.3.6 Lokalizácia

Kategória lokalizácie upozorňuje na prípady, v ktorých preklad nezodpovedá konvenciám pre špecifickú krajinu. Táto kategória zahŕňa šesť podkategórií: formát adresy, formát dátumu, formát meny, formát mier a váh, klávesové skratky a formát telefónneho čísla.

Tabuľka 25: *Priemerná chybovosť v kategórii lokalizácie*

Typy chýb	Priemerná chybovosť	
	Prekladatelia	Posteditori
Formát adresy	0	0
Formát dátumu	0	0
Formát meny	0	0
Formát mier a váh	0,6	0,3
Klávesové skratky	0	0
Formát telefónneho čísla	0	0

V kategórii lokalizácie prekladatelia a posteditori pochybili iba v podkategórii formátu mier a váh, v ktorej vykázali väčšiu chybovosť prekladatelia.

#### 5.3.3.6.1 Formát mier a váh

Prekladatelia mali s lokalizáciou jednotiek väčší problém ako posteditori. V niektorých prípadoch sa rozhodli jednotky vôbec nelokalizovať, čo znamená, že by si čitatelia

návodu na slovenskom trhu museli jednotky konvertovať sami, aby dokázali vermikomposter zostaviť. Okrem toho prekladatelia robili chyby aj v samotnej konverzii jednotiek. Tomuto prípadu podlieha aj chyba v uvedenom príklade, kde prekladateľ údaje prekonvertoval, ale pravdepodobne z nepozornosti uviedol k číselnému údaju nesprávny prefix jednotky, takže uviedol centimetre namiesto milimetrov.

Tabuľka 26: *Príklady chýb v podkategórii formátu mier a váh*

Skupina	Prekladatelia	Segment	14
<b>Východiskový text</b>	Drill at least 10 quarter-inch holes in the lid.		
<b>Preklad</b>	Do veľa vyvrtajte najmenej desať 6 centimetrových otvorov.		
<b>Navrhovaný preklad</b>	Do veľa vyvrtajte najmenej desať 6-milimetrových otvorov.		
Skupina	Posteditori	Segment	8
<b>Východiskový text</b>	\{i\}>Dimensions should be approximately 12 x 12 x 12 inches (one cubic foot) but do not have to be precise.\<i\}		
<b>Preklad</b>	\{i\}>Rozmery by mali byť približne 12 x 12 x 12 palcov (jedna kubická stopa), ale nemusia byť presné.\<i\}		
<b>Navrhovaný preklad</b>	\{i\}>Rozmery by mali byť približne 30 x 30 x 30 cm (s objemom zhruba 28 litrov), ale nemusia byť presné.\<i\}		

Aj keď posteditori mali v tejto podkategórii menšiu priemernú chybovosť, aj v tejto skupine sa niektorí rozhodli jednotky neprekonvertovať vôbec.

Na záver kvalitatívnej analýzy je potrebné odpovedať na tri výskumné otázky, ktoré sa tejto analýzy týkali. V rámci chybovosti v preklade sme vyššiu priemernú chybovosť, a to 10, zaznamenali u posteditorov, pričom priemerná chybovosť u prekladateľov bola 7,8. Pri druhej otázke sme chceli zistiť, ktorá skupina vyprodukovala úspešnejšie preklady. Keďže priemerná chybovosť bola priamo prepojená s penalizačnými bodmi, úspešnejšie preklady vytvorili prekladatelia. Prekladatelia získali v priemere 30,2 penalizačného bodu, pričom posteditori 38,8. Poslednou výskumnou otázkou v rámci kvalitatívnej analýzy sme chceli zistiť, aké typy chýb vykonávali prekladatelia a aké posteditori. Po bližšej analýze jednotlivých kategórií chýb sme zistili, že posteditori robili v priemere viac chýb v kategórii správnosti prekladu, plynulosti prekladu, terminológie a formátovania. Naopak, prekladatelia mali horšie výsledky v kategórii štýlu a lokalizácie. Pri analýze podkategórií sme si však všimli, že prekladatelia robili chyby aj v podkategóriách, v ktorých posteditori nepochybili.

## 5.4 Časová analýza

V rámci experimentu sme sa okrem počtu a typov chýb rozhodli pozrieť aj na časový aspekt úsilia pri preklade a posteditácii, keďže tento aspekt sa dá jednoducho zmerať a porovnať.

Prekladatelia a posteditori boli vopred informovaní o tom, že ich čas strávený na preklade sa bude zaznamenávať, a že sa preto treba venovať iba danej činnosti a potom výsledný dokument odovzdať, keď s ním budú spokojní.

Tabuľka 27: *Priemerný čas strávený na preklade a posteditácii*

Skupina	Priemerný čas strávený na preklade/posteditácii
Prekladatelia	36,8 minúty
Posteditori	25,9 minúty

Pri priemerných časoch skupiny prekladateľov a posteditorov si môžeme všimnúť, že posteditori vo všeobecnosti posteditáciu vykonali rýchlejšie. Oproti prekladateľom mali k dispozícii nespracovaný strojový preklad, ktorý im pravdepodobne urýchlil prácu, a boli tak oproti prekladateľom efektívnejší s priemerným časom 25,9 minúty.

Treba však poznamenať, že rýchlosť prekladateľov na dĺžku textu bola taktiež dobrá, keďže ho dokázali v priemere preložiť za 36,8 minúty.

Na záver by sme teda mohli zhodnotiť, že na danej vzorke sme dokázali vyššiu rýchlosť u posteditorov. Týmto zistením si zároveň odpovedáme na poslednú výskumnú otázku, ktorá sa týkala časovej analýzy. Chceli sme totiž zistiť, ktorá skupina vykonala preklad rýchlejšie a vďaka rozboru nameraných časov môžeme zhodnotiť, že posteditori boli časovo efektívnejší oproti prekladateľom.

## 5.5 Diskusia

Po predstavení výsledkov je dôležité zamyslieť sa, prečo nám dané výsledky vyšli a či sa dali predpokladať.

V rámci vyhodnocovania výsledkov sme začali s kvalitatívnou analýzou, kde sme sa najprv pozreli na priemernú chybovosť a úspešnosť prekladu v oboch skupinách. Keďže chybovosť bola priamo prepojená s úspešnosťou prekladu cez systém penalizačných bodov, v oboch prípadoch sme zistili, že na tom boli horšie posteditori. Ak sme chceli zistiť, prečo nám vyšli takéto výsledky, museli sme sa pozrieť aj na rozbor chybovosti



v typológii chýb. Ten nám totiž odhalil, ktoré chyby robili najčastejšie posteditori a prekladatelia a aké dôvody sa za tým mohli skrývať.

Pri analýze typológie chýb sme pracovali so šiestimi hlavnými kategóriami: správnosť prekladu, plynulosť prekladu, terminológia, štýl, formátovanie a lokalizácia. Až v štyroch kategóriách zo šiestich vykázali vyššiu priemernú chybovosť posteditori. Išlo o kategórie správnosti prekladu, plynulosti prekladu, terminológie a formátovania. Znamená to, že prekladatelia na tom boli horšie iba v kategóriách štýlu a lokalizácie. Každá kategória mala aj podkategórie typov chýb, ktoré nám ďalej odhalili, že aj keď posteditori pochybili častejšie, prekladatelia vykázali väčšiu rozmanitosť chýb.

Ak sa totiž pozrieme na podkategórie správnosti prekladu, môžeme si všimnúť, že z piatich podkategórií, ktorým sme sa bližšie venovali, v priemere viac pochybili posteditori iba v jednej, a to v podkategórii nesprávneho prekladu. V porovnaní s prekladateľmi teda posteditori vykazovali chybovosť iba v jednej podkategórii, a teda robili tie isté chyby, ktoré boli ovplyvnené strojovým prekladom. Množstvo chýb v tejto podkategórii bolo totiž spôsobené práve tým, že posteditori neopravili chyby vytvorené strojovým prekladom. Buď sa na strojový preklad spoľahli viac, ako sa mali, alebo si ich jednoducho nevšimli. Najväčší problém totiž mali práve s nesprávne preloženými slovami, ktoré sa podľa kontextu do daného textu vôbec nehodili. V podkategóriách ako nedostatočný preklad či nepreložený text pochybili iba prekladatelia, pretože strojový preklad vo výstupe nevynechal žiadnu informáciu či nepreložené slovo. Už len v prvej kategórii teda vidíme, že strojový preklad mal jasný vplyv na typy chýb, ktoré posteditori robili, a že ich dokonca v istých ohľadoch mohol zmiast' či mu mohli príliš veriť, a nedodať tak preklad najvyššej kvality.

Rovnaká analýza výsledkov by sa dala aplikovať aj na druhú kategóriu, a to kategóriu plynulosti prekladu, v ktorej taktiež posteditori vykázali vyššiu priemernú chybovosť, no opäť by sa dalo povedať, že bola ovplyvnená strojovým prekladom. Najväčší rozdiel vidieť medzi podkategóriami pravopisu a gramatickej chyby. V podkategórii pravopisu sme vyššiu chybovosť zaznamenali u prekladateľov, pretože strojový preklad ako taký preklepy v cieľovom texte neprodukuje. Ide o čisto ľudskú chybu, ktorá sa tým pádom prejavila v prípade prekladateľov, alebo keď sa posteditori rozhodli prepísať určitý segment v nespracovanom výstupe strojového prekladu. V podkategórii gramatickej chyby však väčšiu priemernú chybovosť vykázali posteditori, pričom sme v nej zaznamenali aj najväčší rozdiel medzi skupinami. Už samotný nespracovaný výstup prekladača totiž obsahoval pár gramatických chýb, ktoré sa však nepodarilo identifikovať viacerým posteditorom. Opäť si ich buď nevšimli, alebo dôverovali strojovému prekladu viac, ako

mali, pretože gramatická správnosť je jedna z najdôležitejších aspektov kvalitného prekladu.

V rámci kategórie terminológie sme zaznamenali najväčší rozdiel v chybovosti v podkategórii nekonzistentnosti s glosárom. V nej totiž pochybili iba posteditori, čo bolo opäť ovplyvnené strojovým prekladom, keďže ten termíny prekladal nesprávne, resp. inak od termínov v glosári a posteditori mali tieto chybné preklady s glosárom zjednotiť. Vo výslednom preklade teda vznikli nielen nekonzistentnosti v rámci terminológie, ktoré posteditori taktiež vo výstupe strojového prekladu neopravili, ale aj nekonzistentnosti s glosárom, ktoré sa u prekladateľov nevyskytovali vôbec.

V kategórii štýlu sa prejavila iba jedna chyba, a to v skupine prekladateľov. Išlo zrejme o chybu z nepozornosti, ktorú si po sebe daný prekladateľ neskontroloval. Posteditori nemali so štýlom problém, zrejme bol výstup strojového prekladu štylisticky akceptovateľný.

V rámci podkategórie formátovania opäť vykázali rozmanitejšiu chybovosť prekladatelia, no posteditori mali vyššiu celkovú chybovosť, najmä čo sa týka podkategórie tagov (značky). Posteditori mali s tagmi väčšie problémy, pretože strojový preklad spôsobil, že sa vo výstupe nachádzali viaceré medzery v tagoch alebo v niektorých segmentoch tagy úplne vynechal. Viacerí posteditori si tieto zmeny nevšimli a neopravili ich. Opäť boli teda ovplyvnení výstupom strojového prekladu, a priemerná chybovosť tak narástla.

V kategórii lokalizácie sme zaznamenali iba jednu podkategóriu, v ktorej robili prekladatelia a posteditori chyby, a to formát mier a váh. Niektorí prekladatelia sa rozhodli merné jednotky nelokalizovať a uvádzali ich v zahraničných hodnotách. V skupine posteditorov sa to však stalo iba raz. Strojový preklad sa v jednom segmente totiž rozhodol neudávať teploty v stupňoch Fahrenheita aj v stupňoch Celzia, ale uviedol ich iba v stupňoch Celzia. Tento segment mohol posteditorov podnietiť jednotky lokalizovať aj v ostatných segmentoch, a zjednotiť tak výsledný preklad. Prekladatelia mali na druhej strane na výber či budú jednotky lokalizovať, alebo nie, a tento výber spôsobil aj prípady, v ktorých teda prekladatelia jednotky nelokalizovali.

V rámci zhrnutia kvalitatívnej analýzy môžeme povedať, že posteditori mali vyššiu priemernú chybovosť, a teda menej úspešné preklady, no na základe konkrétnych príkladov chýb vieme zhodnotiť, že ich do dostatočnej miery ovplyvnil aj strojový preklad. Ten mal v sebe totiž isté typy chýb, na ktoré si museli dávať posteditori pozor, no niektoré typy chýb neobsahoval vôbec, keďže strojový preklad ako taký väčšinou nevynecháva

časti textu, nepridáva informácie do prekladu či nerobí preklepy. Dalo by sa teda povedať, že posteditor by sa mal sústrediť najmä na opravy tagov, nekonzistentností v texte (najmä termínov a použitých slov), gramatických chýb a nesprávnych prekladov slov a termínov vytvorených strojovým prekladom.

Po kvalitatívnej analýze sme prešli na časovú analýzu, kde sme skúmali časový aspekt úsilia u prekladateľov a posteditorov. Chceli sme zistiť, ktorá skupina bola schopná vytvoriť preklad rýchlejšie. S priemerným časom 25,9 minúty boli rýchlejší posteditori, pričom prekladatelia mali priemerný čas 36,8 minúty. Tento výsledok sa dal predpokladať, keďže vo viacerých nami zmieňovaných článkoch (Krings, 2001; Tatsumi, 2010) sa uvádza, že posteditácia strojového prekladu je rýchlejšia, a teda efektívnejšia ako humánny preklad. V rámci nášho experimentu na danej vzorke sa nám tento výrok potvrdil.

Na záver diskusie však treba dodať, že daný výskum má isté limity, najmä čo sa týka hodnotenia kvality prekladu, ktorá bola vykonaná iba jednou osobou, a to autorkou článku. Aj napriek jej snahe o objektivitu sa na výsledkoch do istej miery z očividných dôvodov musela prejaviť subjektivita. Na dosiahnutie väčšej objektivita by bolo v budúcich výskumoch vhodné na hodnotenie kvality prekladu použiť viacero osôb alebo aspoň takú osobu, ktorá sa priamo výskumom nezaobera, text dopredu nepripravovala a nevedela o znalostiach zúčastnených prekladateľov a posteditorov.

## Záver

V príspevku sme sa venovali porovnávaniu humánneho prekladu a posteditácie strojového prekladu. Naším cieľom bolo porovnať humánny preklad s posteditáciou strojového prekladu na danej vzorke študentov a zistiť, ktorý proces je úspešnejší na úrovni rýchlosti práce a kvality výsledného prekladu. Na splnenie tohto cieľa sme vykonali kvalitatívnu a časovú analýzu.

Na základe výsledkov z oboch analýz, ktoré sme interpretovali v predošlej kapitole, hodnotíme, že posteditácia strojového prekladu má svoje výhody, keďže sa v našom výskume ukázalo, že je časovo efektívnejšia a eliminuje niektoré typy chýb ako preklepy, vynechanie či pridanie informácie. Na druhej strane sme však pozorovali, že strojový preklad stále robí chyby, a preto je posteditácia v tomto procese nevyhnutná. Chyby sa však často opakujú a vyškolený posteditor by mal byť schopný chyby ako napríklad nekonzistentnosti v terminológii opraviť.

Pavlíková, Diana. 2022. Porovnanie humánneho prekladu a posteditácie strojového prekladu. In: L10N Journal. 1(1): s. 64 – 101.

Prostredníctvom výskumu sme taktiež zistili, že študenti bez hĺbkovej prípravy na posteditáciu strojového prekladu vytvárajú posteditáciou preklad nižšej kvality ako študenti, ktorí už majú s humánnym prekladom viacročné skúsenosti počas štúdia. Chybovosť je priamo prepojená s neznalosťou práce so strojovým prekladom. Výučba posteditácie by preto mohla eliminovať nadbytočnú chybovosť v jednotlivých kategóriách gramatickej správnosti a nekonzistentnosti a prekladu termínov. Študenti by sa tiež mohli naučiť, ako k výstupu strojového prekladu správne pristupovať a ako vykonávať posteditáciu efektívnejšie, aby nebola výhodná iba časovo, ale aj kvalitatívne.

## Bibliografia

- Absolon, Jakub. 2018. Strojový preklad a posteditovanie. Nitra: Univerzita Konštantína Filozofa.
- Allen, Jeffrey. 2003. Post-editing. In: Computers and Translation: A translator's guide. Amsterdam: John Benjamins Publishing Company, s. 297 – 318.
- Bubnic, Mauro. 2022. Meeting the Needs of End Users with AI-based NLP Tools. In: Translating Europe Forum 2022. <https://www.youtube.com/watch?v=CXOSCjkcCNc&list=PLLqI-RaiVCGCSrITMx1iG1pRlyYR1ZKSyi&index=15>. Cit. 11. 12. 2022.
- Common Sense Advisory. 2016. Global Market Research Firm Common Sense Advisory Finds Post-edited Machine Translation (PEMT) Among Fastest-growing Segments of the Language Industry. <https://csa-research.com/Blogs-Events/CSA-in-the-Media/Press-Releases/PEMT-Among-Fastest-growing-Segments-of-the-Language-Industry>. Cit. 3. 12. 2022.
- Dengová, Stanislava. 2020. 3 most popular CAT tools compared. <https://www.lexika-translations.com/blog/comparison-of-cat-tools/>. Cit. 17. 1. 2022.
- Doherty, Stephen a Gaspari, Federico. 2013. Effective Post-Editing in Human & Machine Translation Workflows: Critical Knowledge & Techniques. Dublin: Dublin City University.
- Finnegan, Noel. 2018. Linguistic Quality Assessment: How Do You Measure Translation Quality? <https://www.argosmultilingual.com/blog/linguistic-quality-assessment-how-do-you-measure-translation-quality>. Cit. 23. 1. 2022.
- Frívaldský, Dalibor. 2020. Memsource Translate: Interview with Dalibor Frívaldský Memsource Chief Technology Officer. <https://www.memsource.com/blog/memsource-translate-interview-with-chief-technology-officer/>. Cit. 17. 1. 2022.
- Gene, Viveta. 2019. Machine Translation: From Translation to Post-editing. <https://www.gala-global.org/events/events-calendar/machine-translation-translation-post-editing>. Cit. 16. 1. 2022.
- Green, Spence; Heer, Jeffrey a Manning, Christopher D. 2013. The efficacy of human post-Editing for language translation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (ACM). Association for Computing Machinery, s. 439 – 448.
- Guerberof Arenas, Ana. 2010. Project management and machine translation. In: Multilingual: 21(1) s. 1 – 4.
- Hu, Ke a Cadwell, Patrick. 2016. A Comparative Study of Post-editing Guidelines. In: Baltic J. Modern Computing: 4(2), s. 346 – 353.

- Pavlíková, Diana. 2022. Porovnanie humánneho prekladu a posteditácie strojového prekladu. In: L10N Journal. 1(1): s. 64 – 101.
- Hudecová, Elena a kol. 2021. Porovnanie strojového, posteditovaného a ľudského prekladu technickej dokumentácie zo slovenčiny do nemčiny. In: Slovenská Reč: 86(2), s. 192 – 207.
- Hutchins, W. John a Somers, Harold L. 1992. An Introduction to Machine Translation. San Diego: Academic Press.
- Indoor Composting with a Worm Bin. New York: NYC Department of Sanitation. <https://www1.nyc.gov/assets/dsny/downloads/pdf/promotional-materials/indoor-worm-bin-composting-brochure-06340-f.pdf>. Cit. 24. 1. 2022.
- Kabát, Marián. 2022. Všeobecná štylistická príručka pre lokalizáciu softvérových produktov. Bratislava: Stimul.
- Korkas, Vassilis. What is linguistic quality assurance and how it works in localization. <https://lex-iqua.net/what-is-linguistic-quality-assurance-and-how-it-works-in-localization/>. Cit. 23. 1. 2022.
- Krings, P. Hans. 2001. Repairing Texts. Empirical Investigations of Machine Translation Post-Editing Processes. Kent: The Kent State University Press.
- Läubli, Samuel a kol. 2019. Post-editing Productivity with Neural Machine Translation: An Empirical Assessment of Speed and Quality in the Banking and Finance Domain. In: Proceedings of Machine Translation Summit XVII Volume 1: Research Track. Dublin: European Association for Machine Translation, s. 267 – 272.
- Mesa-Lao, Bartolomé. 2013. Introduction to post-editing. In: SEECAT – Speech & Eye-Tracking Enabled CAT. Kodaň: Copenhagen Business School.
- Munková, Daša. 2013. Evalvacia strojového prekladu. Nitra: Univerzita Konštantína Filozofa.
- O'Brien, Sharon. 2002. Teaching Post-editing: A Proposal for Course Content. In: European Association for Machine Translation Conferences/Workshops. Dublin: Dublin City University.
- Pavlíková, Diana. 2021. Porovnanie strojového a humánneho prekladu terminológie. In: Prekladateľské listy 10. Bratislava: Univerzita Komenského, s. 64 – 77.
- Plitt, Mirko a Masselot, François. 2010. A productivity test of statistical machine translation post-editing in a typical localisation context. In: The Prague Bulletin of Mathematical Linguistics, s. 7 – 16.
- Robert, Anne-Marie. 2013. Vous avez dit post-éditrice? Quelques éléments d'un parcours personnel. In: The Journal of Specialised Translation: 19(1), s. 29 – 40.
- Sánchez-Gijón, Pilar; Moorkens, Joss a Way, Andy. 2019. Post-editing neural machine translation versus translation memory segments. In: Machine Translation: 33(1), s. 31 – 59.
- Tatsumi, Midori. 2009. Correlation between automatic evaluation metric scores, post-editing speed and some other factors. In: MT Summit XII – The twelfth Machine Translation Summit International Association for Machine Translation hosted by the Association for Machine Translation in the Americas. Association for Machine Translation in the Americas.
- Tatsumi, Midori. 2010. Post-Editing Machine Translated Text in A Commercial Setting: Observation and Statistical Analysis. Dublin: Dublin City University.
- TAUS. 2010. Post-editing in practice. <https://www.taus.net/think-tank/reports/postedit-reports/postediting-in-practice>. Cit. 23. 1. 2022.
- TAUS. 2016. Machine Translation Post-editing Guidelines. <https://info.taus.net/mt-post-editing-guidelines>. Cit. 23. 1. 2022.
- TAUS. Start tracking errors with DQF-MQM. <https://info.taus.net/dqf-mqm-error-typology-templ>. Cit. 23. 1. 2022.
- Torrejón, Enrique a Rico, Celia. 2012. Skills and Profile of the New Role of the Translatoras MT Post-editor. In: Revista Tradumatica: 10(1), s. 166 – 178.

Pavlíková, Diana. 2022. Porovnanie humánneho prekladu a posteditácie strojového prekladu. In: L10N Journal. 1(1): s. 64 – 101.

Wu, Yonghui a kol. 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. Ithaca: Cornell University.