

Univerzita Karlova

Filozofická fakulta

Ústav translatologie



Diplomová práce

Bc. Soňa Kyselová

**Srovnání (a historická podmíněnost) výstupů ze strojových
překladačů**

Comparing Machine Translation Output (and the Way it Changes over Time)

Praha 2018

Vedoucí práce: PhDr. et Bc. Tomáš Svoboda, Ph.D.

Poděkování:

Děkuji PhDr. et Bc. Tomáši Svobodovi, Ph.D. za podněty a připomínky, které mi poskytl během psaní této diplomové práce.

Prohlášení:

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 4. 1. 2018

.....
Soňa Kyselová

Klíčová slova

strojový překlad, online překladače, kvalita, hodnocení, textový typ, vývoj v čase

Keywords

machine translation, online machine translation systems, quality, evaluation, text-type, development over time

Abstrakt

Předmětem této diplomové práce je strojový překlad, jehož výzkumu se lingvistika (a později i translatologie) věnuje poměrně dlouho a který se v posledních letech dostal i do popředí zájmu širší veřejnosti.

Cílem této práce je postihnout vývoj kvality výstupů z veřejně dostupných překladačů v čase. Teoretická část se proto nejprve zabývá strojovým překladem obecně, tedy základními definicemi, stručnou historií a možnými přístupy, poté jsou představeny veřejně dostupné online překladače a metody hodnocení kvality strojového překladu. V závěru této části je popsán metodologický model pro část empirickou.

V empirické části se na vzorku překladů vyhotovených s využitím veřejně dostupných online překladačů ověřuje, jak se vybrané překladače vypořádají s jednotlivými textovými typy a zda dochází ke zlepšení kvality výstupů v čase. Za tímto účelem je provedena jak translatologická analýza překladů, která zohledňuje textový typ, sémantickou, lexikální, stylistickou a pragmatickou rovinu, tak hodnocení na stupnici za účelem zhodnocení celkové použitelnosti překladu.

V závěru práce jsou porovnány a shrnuty výsledky empirické studie. Na základě tohoto srovnání jsou vyvozeny závěry a nastíněny obecné tendence vyplývající z empirické části práce.

Abstract

This diploma thesis focuses on machine translation (MT), which has been studied for a relatively long time in linguistics (and later also in translation studies) and which in recent years is at the forefront of the broader public as well.

This thesis aims to explore the quality of machine translation outputs and the way it changes over time. The theoretical part first deals with the machine translation in general, namely basic definitions, brief history and approaches to machine translation, then describes online machine translation systems and evaluation methods. Finally, this part provides a methodological model for the empirical part.

Using a set of texts translated with MT, the empirical part seeks to check how online machine translation systems deal with translation of different text-types and whether there is improvement of the quality of MT outputs over time. In order to do so, an analysis of text-type, semantics, lexicology, stylistics and pragmatics is carried out as well as a rating of the general applicability of the translation.

The final part of this thesis compares and concludes the results of the analysis. With regard to this comparison, conclusions are made and general tendencies stated that have emerged from the empirical part of the thesis.

Obsah

1	ÚVOD	10
2	TEORETICKÁ VÝCHODISKA.....	12
2.1	STROJOVÝ PŘEKLAD	12
2.1.1	<i>Základní definice.....</i>	<i>12</i>
2.1.2	<i>Stručná historie strojového překladu.....</i>	<i>13</i>
2.1.3	<i>Přístupy ke strojovému překladu.....</i>	<i>19</i>
2.2	VEŘEJNĚ DOSTUPNÉ PŘEKLADAČE.....	26
2.2.1	<i>Online překladače obecně</i>	<i>27</i>
2.2.2	<i>Google Translate</i>	<i>28</i>
2.2.3	<i>Bing Translator</i>	<i>30</i>
2.3	HODNOCENÍ STROJOVÉHO PŘEKLADU.....	31
2.3.1	<i>Metody hodnocení strojového překladu.....</i>	<i>32</i>
2.3.2	<i>Přístupy k hodnocení strojového překladu.....</i>	<i>37</i>
2.3.3	<i>Typické chyby.....</i>	<i>38</i>
3	METODOLOGICKÝ ZÁKLAD	42
3.1	VÝCHOZÍ HYPOTÉZY	42
3.2	MODEL PRO HODNOCENÍ	43
3.2.1	<i>První stupeň: Translatologická analýza.....</i>	<i>44</i>
3.2.2	<i>Druhý stupeň: Posouzení na stupnici</i>	<i>50</i>
3.3	EMPIRICKÝ MATERIÁL	52
3.3.1	<i>Popis výchozích textů</i>	<i>52</i>
3.3.2	<i>Sběr materiálu.....</i>	<i>54</i>
4	VÝSLEDKY EMPIRICKÉHO VÝZKUMU.....	56
4.1	APELATIVNÍ TEXTY	57
4.1.1	<i>Hodnocení překladů na základě translatologické analýzy.....</i>	<i>57</i>
4.1.2	<i>Vývoj počtu chyb ve třech vybraných kategoriích</i>	<i>58</i>
4.1.3	<i>Celkové zhodnocení překladu pomocí stupnice</i>	<i>62</i>
4.2	INFORMATIVNÍ TEXTY	62
4.2.1	<i>Hodnocení překladů na základě translatologické analýzy.....</i>	<i>62</i>

4.2.2	<i>Vývoj počtu chyb ve třech vybraných kategoriích</i>	64
4.2.3	<i>Celkové zhodnocení překladu pomocí stupnice</i>	67
4.3	UMĚLECKÉ TEXTY	67
4.3.1	<i>Hodnocení překladů na základě translatologické analýzy</i>	67
4.3.2	<i>Vývoj počtu chyb ve třech vybraných kategoriích</i>	69
4.3.3	<i>Celkové zhodnocení překladu pomocí stupnice</i>	72
5	SHRNUTÍ VÝSLEDKŮ	74
6	ZÁVĚR A VÝHLED K BUDOUCÍMU VÝZKUMU V DANÉ OBLASTI	78
7	BIBLIOGRAFIE	83
8	PŘÍLOHY	88
9	SEZNAM TABULEK, GRAFŮ A OBRÁZKŮ	93

Seznam zkratek:

ALPAC	Automatic Language Processing Advisory Committee
BLEU	Bilingual Evaluation Understudy
BT	Bing Translator
CAT	Computer-Assisted/Aided Translation
CBMT	Corpus-based Machine Translation
EBMT	Example-Based Machine Translation
FAHQT	Fully Automatic High Quality Translation
FAMT	Fully Automatic Machine Translation
GT	Google Translate
HAMT	Human-Assisted/Aided Machine Translation
HMT	Hybrid Machine Translation
HT	Human Translation
MAHT	Machine-Assisted/Aided Human Translation
METEOR	Metric for Evaluation of Translation with Explicit Ordering
MFF	Matematicko-fyzikální fakulta
MT	Machine Translation
NIST	National Institute of Standards and Technology
NMT	Neural Machine Translation
NS	Normostrana
PBSMT	Phrase-Based Statistical Machine Translation
PEMT	Post Edited Machine Translation
PER	Position-Independent Word Error Rate
RBMT	Rule-Based Machine Translation
RNN	Recurrent Neural Network
SMT	Statistical Machine Translation
SP	Strojový překlad
TER	Translation Error Rate
TM	Translation Memory
TPU	Tensor Processing Units
VT	Výchozí text
WER	Word Error Rate

1 Úvod

Strojový překlad (SP) je disciplína na pomezí lingvistiky a informatiky. Technologie SP však proniká i do běžného života. Strojové překladače patří v dnešní době k rozšířeným nástrojům, které jsou denně používány miliony lidí po celém světě pro informativní účely nebo při vyhledávání na internetu. Strojově jsou překládána jednotlivá slova, fráze, celé texty nebo webové stránky. Strojový překlad je využíván i profesionálními překladateli, neboť je tato funkce integrována do většiny moderních nástrojů počítačem podporovaného překladu.

Tato technologie je již od dob svého vzniku provázena otázkou kvality. V této souvislosti je nutné poznamenat, že hodnocení strojového překladu je mezioborová činnost, která zahrnuje různé oblasti, jako je například translatologie, počítačnická, formální, obecná nebo kontrastivní lingvistika.

Tématem této diplomové práce je právě hodnocení kvality strojového překladu. Práce si klade za cíl podrobit výsledný produkt strojového překladu analýze a na základě z ní získaných výsledků zhodnotit, jak se vybrané strojové překladače vypořádají s různými textovými typy při překladu z němčiny do češtiny, který z nich bude produkovat kvalitnější výstupy a jestli dojde ke zvýšení kvality překladů v čase.

Tato práce je jedna z prvních závěrečných prací na Ústavu translatologie Filozofické fakulty Univerzity Karlovy, která se zabývá strojovým překladem, a první prací, jež se zaměřuje na hodnocení a vývoj kvality u výstupů ze strojových překladačů. Dalším specifikem této práce je skutečnost, že se věnuje hodnocení strojového překladu u méně často zkoumané jazykové dvojice němčina/čeština.

Z těchto důvodů se bude teoretická část práce nejprve snažit poskytnout úvod do obecné problematiky strojového překladu. Na začátku dojde k vymezení základních termínů, se kterými je možné se v této oblasti setkat, pak bude následovat exkurz do historie strojového překladu, dále budou popsány možné přístupy ke SP, nastíněny principy jejich fungování a stručně charakterizovány vybrané veřejně dostupné strojové překladače. Nakonec se teoretická část zaměří na metody hodnocení kvality výstupů ze SP.

Teoretické kapitoly vychází především ze zahraniční odborné literatury, neboť vědeckých prací psaných v českém jazyce je k tomu tématu poměrně málo. S touto problematikou souvisí skutečnost, že česká terminologie není v některých případech zcela ustálená.

V metodologické části budou nejprve formulovány výchozí hypotézy, které pak budou verifikovány na základě výsledků zjištěných v případové studii. Dále zde bude představen analytický model pro hodnocení, které bude probíhat ve dvou fázích. Následně bude charakterizován empirický materiál, budou tedy popsány výchozí texty a proces, jakým způsobem byly získávány překlady.

Poté bude následovat empirická část, v jejímž rámci budou prezentovány výsledky případové studie. Tyto výsledky budou představeny především z hlediska kvantitativního s cílem zaujmout stanovisko k hypotézám definovaným v metodologické části práce.

Praktická část bude zakončena závěrem, v němž budou na základě teoretických úvah rozvedených v této práci představena zjištění, která vyplývají z analytické fáze této práce. Nakonec dojde ke stručnému nástinu možností směřování budoucího výzkumu v této oblasti.

2 Teoretická východiska

2.1 Strojový překlad

Tato kapitola nabídne stručný úvod do problematiky strojového překladu. Abychom se vyhnuli terminologickým nejasnostem, budou v první části vysvětleny pojmy, které se v této oblasti často objevují. Následovat bude stručný exkurz do historie poměrně krátkého, leč bohatého vývoje SP. Kapitola bude zakončena popisem jednotlivých typů strojových překladačů.

2.1.1 Základní definice

Machine Translation (MT), tedy strojový překlad (SP), je souhrnným již tradičně užívaným označením pro počítačové systémy, které jsou schopny automaticky vyprodukovat překlad textu z jednoho přirozeného jazyka do jazyka druhého. Pod stejným termínem se skrývá i název vědní disciplíny, jež se zabývá výzkumem v této oblasti (KRÁL, 2003, str. 15). V odborné literatuře se někdy můžeme setkat i s názvy jako *mechanical translation* nebo *automatic translation* čili mechanický nebo automatizovaný překlad, avšak tyto termíny se v angličtině používají velmi zřídka. Zajímavé je, že jejich cizojazyčné ekvivalenty se v některých jazycích prosadily, jako například francouzské *traduction automatique* (HUTCHINS & SOMERS, 1992, str. 3). Než se začneme zabývat strojovým překladem jako takovým, je potřeba vysvětlit některé termíny, které se v souvislosti se SP používají.

Alcina (2008, str. 81n) popisuje několik typů počítačových technologií v překladatelském odvětví, které rozlišuje mimo jiné podle míry automatizace v překladatelském procesu, tedy podle stupně zapojení lidského překladatele. Termín tradiční lidský překlad (*traditional human translation*) je chápán jako překlad bez jakéhokoliv stupně automatizace. Protipólem tohoto označení je plně automatizovaný vysoce kvalitní strojový překlad (*fully automatic high quality translation*, FAHQT). Překlad je v tomto případě proveden bez lidského zásahu.

Mezi předchozími dvěma termíny se na pomyslné škále znázorňující stupeň automatizace nachází strojem podporovaný překlad vyhotovený člověkem a člověkem podporovaný strojový překlad. V rámci strojem podporovaného překladu vyhotoveného člověkem (*machine-aided human translation*, MAHT) dochází k podpůrnému využívání počítačů pro překlad. O člověkem podporovaném strojovém překladu (*human-aided machine*

translation, HAMT) se hovoří v případě, kdy je pro dosažení přijatelných výsledků nutné, aby člověk před vložením výchozího textu do překladače provedl předběžnou úpravu textu nebo aby pak výstup z překladače opravil. Hutchins a Somers (1992, str. 150) označují tyto úpravy jako *pre-editing* a *post-editing*.

Hranice mezi MAHT a HAMT je někdy velmi tenká a obě tato odvětví lze souhrnně označit termínem počítačem podporovaný překlad, tedy *computer-assisted translation* nebo také *computer-aided translation*, CAT (HUTCHINS & SOMERS, 1992, str. 3). Počítačem podporovaný překlad funguje na bázi překladové paměti (*translation memory*, TM) a na rozdíl od strojových překladačů přeloží pouze text, který vykazuje určitou míru shody (SVOBODA, 2012, str. 80). Král (2003, str. 15) upozorňuje na skutečnost, že s rozmachem osobních počítačů se škála nástrojů pro překlad znatelně rozšířila, proto mezi ně můžeme zařadit v zásadě jakékoliv počítačové systémy, které překladateli podstatným způsobem ulehčují práci.

2.1.2 Stručná historie strojového překladu

Do povědomí široké veřejnosti se strojový překlad dostal teprve v několika posledních desetiletích. Úvahy o využití strojů pro překlad však sahají do vzdálené minulosti. Kdybychom chtěli pátrat po prvních impulzech k automatizaci překladu, museli bychom se vrátit dokonce o několik století nazpět. Jednu z prvních myšlenek o automatizaci překladu totiž formuloval německý mnich Johannes Becher již v 17. století ve Špýru. Ve svém spisu *Charakter. Pro Notitia Linguarum Universali* z roku 1661 se zabývá možností automatizace překladu pomocí matematického metajazyka. Becherův překlad měl fungovat tak, že každému výrazu, včetně jeho různých gramatických tvarů, by byl přidělen číselný kód. Takovému kódovému slovníku zamýšlel vytvořit pro různé jazyky, překlad z jednoho jazyka do druhého by pak probíhal „mechanicky“ přes tyto kódy (FREIGANG, 2001, str. 1).

Podobnými myšlenkami na usnadnění překladu se zabývali i další. První skutečně mechanická zařízení pro překlad však byla vynalezena až v první polovině 20. století. Francouzský vědec G. Artsrouni patentoval v roce 1933 zařízení, které mimo jiné fungovalo jako mechanický slovník a pracovalo na principu děrné pásky (FREIGANG, 2001, str. 1). Vynálezce jej nazval *Mechanical Brain* (SHUK MAN, 2015, str. 237). Ve stejném roce podal přihlášku na podobný patent i ruský vědec Petr Smirnov-Trojanskij. Roku 1939 představil vylepšenou verzi svého překladače a hodlal dále pokračovat ve vývoji. Jeho návrhy a představy

však zůstaly v Sovětském svazu bez odezvy a jeho práce upadla na mnoho let v zapomnění (REIMANN, 1991, str. 69).

K zásadnímu průlomů ve vývoji strojového překladu dochází až po druhé světové válce díky vynálezu počítačů. Za skutečné průkopníky SP bychom mohli považovat britského krystalografa Andrew D. Bootha a Američana Warrena Weavera z Rockefellerovy nadace, kteří spolu v letech 1946–47 diskutovali o možnosti využití počítačů pro překlad přirozených jazyků (HUTCHINS W. , 2006, str. 375n). Podnětem pro zahájení výzkumu v oblasti strojového překladu v USA bylo Weaverovo memorandum pro Rockefellerovu nadaci z července 1949, ve kterém používá analogii mezi překladem a dekódováním:

„I have a text in front of me which is written in Russian but I am going to pretend that it is really written in English and that it has been coded in some strange symbols. All I need to do is strip off the code in order to retrieve the information contained in the text.“ (ARNOLD et al., 1994, str. 13)

Weaver ve svém memorandu nastínil možný vývoj SP a navrhl využití několika konkrétních metod, jako je například kryptografie, statistika, teorie informace, logika a jazykové univerzálie (HUTCHINS W. , 2006, str. 376).

Během několika následujících let započal výzkum v oblasti strojového překladu na různých univerzitách, například v Seattlu, Los Angeles nebo na Massachusetts Institute of Technology (MIT). Právě na MIT byl roku 1951 pověřen první akademický pracovník výzkumem strojového překladu na plný úvazek. Byl jím izraelský filosof, matematik a lingvista Yehoshua Bar-Hillel, který se snažil SP všemožně podporovat a propagovat. Roku 1952 dokonce zorganizoval první mezinárodní konferenci o strojovém překladu (HUTCHINS & SOMERS, 1992, str. 6). O několik let později však vyjádřil pochybnosti o tom, zda je plně automatický vysoce kvalitní strojový překlad vůbec někdy uskutečnitelný (ARNOLD et al., 1994, str. 13).

Éru mezi lety 1954 až 1966 označuje Hutchins (2006, str. 376) jako období velkých očekávání. Roku 1954 byla uskutečněna první veřejná prezentace strojového překladu jazykového páru ruština-angličtina, která se konala na Georgetownské univerzitě ve spolupráci se společností International Business Machines Corporation (IBM). Úspěch tohoto veřejného experimentu pomohl odstartovat financování výzkumu SP americkou vládou. Ve stejném roce započal výzkum strojového překladu na Harvardské univerzitě pod vedením Anthonyho

Oettingera, jenž napsal první disertační práci v tomto oboru (KOERNER & ASHER , 1996, str. 446).

Hutchins (HUTCHINS W. , 2003, str. 503n) uvádí, že se ve zmíněném období začaly objevovat tři základní přístupy ke strojovému překladu. Prvním byl model přímého překladu (*direct translation model*), v jehož rámci byla vyvinuta programovací pravidla pro převod z jednoho určitého zdrojového jazyka do jednoho určitého cílového jazyka s minimální analýzou a syntaktickou reorganizací. Druhý, interlingvální model (*interlingua model*), překládal z výchozího jazyka do jazykově neutrální reprezentace a teprve z ní pak do cílového jazyka a již zahrnoval syntaktickou a sémantickou analýzu. Posledním modelem byl model transferový (*transfer model*). Překládaný text byl nejprve převeden do abstraktní podoby výchozího jazyka a z ní poté do abstraktní podoby cílového jazyka.

Model přímého překladu je popsán jako systém strojového překladu první generace, který se využíval od 50. let až do poloviny 70. let 20. století (HUTCHINS W. , 2003, str. 504). Systémy založené na interlingválním a transferovém přístupu byly charakteristické pro druhou generaci a prosadily se hlavně v 70. a 80. letech minulého století. Tyto přístupy byly založené na specifikaci určitých pravidel, například pro morfologii, syntax nebo sémantiku, proto jsou v dnešní době souhrnně označovány jako systémy založené na pravidlech (*rule-based systems*). Pravidlovými systémy a dalšími druhy strojových překladačů se budeme blíže zabývat v oddílu 2.1.3.

Během studené války byly do strojového překladu vkládány velké naděje, a proto z počátku docházelo k investicím horentních sum do jeho vývoje. Většinou se jednalo o financování systémů pro překlad mezi jazykovou dvojicí angličtina-ruština (Stein 2013: VI). Vystřízlivění z euforie přichází v polovině šedesátých let. Ukázalo se totiž, že s tehdejšími technickými možnostmi se vývoj strojového překladu dostal do slepé uličky. To roku 1966 potvrzuje i zpráva vládní poradní komise *Automatic Language Processing Advisory Committee* (ALPAC), která vznikala při Národní akademii věd Spojených států amerických (ARNOLD et al., 1994, str. 13):

„We have already noted that, while we have machine-aided translation of general scientific text, we do not have useful machine translation. Furthermore, there is no immediate or predictable prospect of useful machine translation.“ (ALPAC, 1966, str. 32)

Výstupy ze strojového překladače byly co do kvality shledány jako neuspokojivé, protože vyžadovaly rozsáhlou post-editaci. V konečném důsledku byl tedy SP nejen pomalejší,

ale i méně přesný a dvakrát tak nákladný jako překlad vyhotovený člověkem. Toto zjištění znamenalo pro výzkum a financování strojového překladu v celosvětovém měřítku velký útlum (HUTCHINS W. , 2014, str. nečíslováno). Místo investic do strojového překladu doporučovala zpráva financovat výzkum nástrojů, jako jsou například automatické slovníky, které by překladateli urychlily práci (ALPAC, 1966, str. 34). V USA, Evropě i Sovětském svazu následně došlo ke stagnaci v oblasti výzkumu SP, pouze v Kanadě, Francii a Německu se tento trend neprosadil (GARNELL, 2014, str. 25).

Avšak ani v tehdejším Československu neměla zpráva ALPAC takové dopady jako ve světě. Hutchins (2012, str. nečíslováno) se domnívá, že to bylo z části způsobeno tím, že byla zastíněna politickými událostmi roku 1968 a z části proto, že si byli výzkumníci stále jistí pokrokem v této oblasti. V Československu měl totiž strojový překlad bohatou tradici.

Již na konci 50. let začaly dvě skupiny výzkumníků, jedna z Matematicko-fyzikální fakulty (MFF) a jedna z Filozofické fakulty (FF) Univerzity Karlovy, provádět experimenty v oblasti SP, které byly založeny na lingvistické analýze a explicitním popisu jazyka (KIRSCHNER, 2000, str. 349). Za průkopníky strojového překladu u nás jsou považováni Petr Sgall a Eva Hajičová (HUTCHINS J. , 2012, str. nečíslováno).

V roce 1957 se konal první experiment, kdy byla na počítači SAPO přeložena jedna věta mezi češtinou a angličtinou. Co se týče větších projektů, mohli bychom zmínit systém APAČ Zdeňka Kirschnera, který byl vyvíjen během 80. let na MFF a měl sloužit pro překlad z angličtiny do češtiny. Systém pracoval s transdukčním slovníkem o velikosti 1500 výrazů (TÝNOVSKÝ, 2007, str. 10).

Na podobném principu byl později založen i SP pro překlad mezi češtinou a ruštinou nazvaný RUSLAN. Systém byl vyvíjen na Karlově univerzitě od roku 1985 a měl překládat dokumentaci pro operační systémy. RUSLAN využíval pravidlový přístup a transdukční slovník. Výzkum byl zastaven v roce 1990 kvůli nedostatečnému financování, které měly na svědomí politické změny. S rozpadem sovětského bloku totiž zanikla potřeba strojového překladu mezi ruštinou a češtinou (KUBOŇ, 2012, str. 284).

Myšlenka překladu mezi příbuznými jazyky byla oživena roku 1998, kdy byl zahájen vývoj systému Česílko, který byl původně určen pro překlad v rámci jazykového páru čeština-slovenština. Týnovský (2007, str. 10) uvádí, že jeho doménou je především lokalizace softwarových systémů, která probíhá ve dvou krocích. Nejdříve je proveden lidský překlad

z typově odlišného jazyka do tzv. pivotního jazyka, což je v tomto případě čeština. Z tohoto pivotního jazyka je pak proveden strojový překlad do příbuzných jazyků.

V 90. letech u nás začala nová éra strojového překladu. Systémy využívající lingvistická pravidla byly postupně nahrazeny strojovými překladači založenými na korpusech a začal se prosazovat především statistický strojový překlad. Čeští výzkumníci v čele s Janem Hajičem, Ondřejem Bojarem a Petrem Homolou hrají v této oblasti výzkumu nezanedbatelnou roli (HUTCHINS 2012, nečíslováno).

Významným českým přínosem pro výzkum SP bylo vytvoření *Pražského závislostního korpusu* (*Prag Dependency Treebank*, PDT), který je jedním z prvních velkých korpusů s ručními větnými rozbory, a *Prague Czech-English Dependency Treebank* (PCEDT), jež obsahuje rozbory anglických textů a jejich českých překladů. Zmíněné korpusy se zakládají na teorii *funkčního generativního popisu*, jež definuje několik rovin jazykového popisu. Syntaktická struktura jednotlivých vět je zachycena pomocí závislostního stromu, tedy tzv. *dependency tree* (BOJAR, 2012, str. 69n).

Existuje celá řada modelů SP, které zpracovávají strukturu vět. Bojar (2012, str. 100n) v této souvislosti zmiňuje unikátní český systém TectoMT Zdeňka Žabokrtského a Martina Popela. Tento systém provádí rozbor a syntézu až na hloubkovou syntaktickou rovinu, strukturu stromu však téměř nemění. Díky tomu vybočuje ze synchronních gramatik.

Vrátíme-li se zpět k vývoji ve světě po uveřejnění zprávy ALPAC, zjistíme, že ani tam se výzkum v oblasti strojového překladu nezastavil dlouhodobě. Již v 70. letech pomalu dochází k jeho obnovení. Peter Toma z Georgetownské univerzity představil systém Systran, který začalo od roku 1970 používat Letectvo Spojených států amerických, od roku 1975 NASA a rok později byl tento systém instalován i v Komisi Evropských společenství pro překlad mezi angličtinou a francouzštinou (HUTCHINS W. , 2014, str. nečíslováno). Ve stejném roce vyvinuli kanadští výzkumníci seskupení okolo projektu TAUM (Traduction Automatique à l'Université de Montréal) úspěšný strojový překladač pro zprávy o počasí nazvaný Meteo (GARNELL, 2014, str. 25).

Na konci sedmdesátých let, konkrétně roku 1978, zahájila Komise Evropských společenství velmi ambiciózní projekt nesoucí název Eurotra. Díky tomuto projektu měl vzniknout překladač, který by zajistil strojový překlad mezi všemi oficiálními jazyky Společenství. Toho času se jednalo o angličtinu, dánštinu, francouzštinu, holandštinu, italštinu a němčinu, brzy nato k oficiálním jazykům přibyla řečtina a roku 1986 portugalská

a španělština. Eurotra měl překládat mezi všemi devíti jazyky, tedy mezi 72 jazykovými kombinacemi (HUTCHINS & SOMERS, 1992, str. 239).

V průběhu 80. let došlo k rozvoji výzkumu v oblasti SP především v Evropě a Japonsku a začal se prosazovat transferový model. Jako nejvýznamnější projekty této doby založené na transferovém přístupu bychom mohli jmenovat systém Ariane skupiny GETA z francouzského Grenoble, systém SUSY (Saarbrücken), systém Mu (Kyoto) a již zmiňovaný projekt Evropských společenství Eurotra (HUTCHINS W. , 2006, str. 378n).

Konec 80. let byl pro strojový překlad důležitým milníkem. Společnost IBM totiž přišla s projektem Candide, který fungoval na bázi statistického modelu (viz oddíl 2.1.3.2). Tento přístup se následně stal dominantním (HUTCHINS W. , 2014, str. nečíslováno). Kromě dynamického rozvoje statistického překladu přinesla 90. léta také přesun od čistě teoretického výzkumu k využití dosažených poznatků v praxi. Toto období je také provázáno vzrůstajícím zájmem o překlad mluveného jazyka (HUTCHINS W. , 2006, str. 381). Příkladem takovýchto snah byl německý projekt Verbmobil, který byl v letech 1993–2000 financován německou vládou (HUTCHINS W. , 2006, str. 381). V roce 1991 byla založena Mezinárodní asociace pro strojový překlad (IAMT), která sdružuje výzkumné ústavy, odborníky z příbuzných oborů a překladatele, vydává různá periodika a pořádá konference a workshopy (WESLEY-TANASKOVIC et al., 1994, str. 287).

Po roce 2000 se objevuje hybridní přístup, který se snaží kombinovat výhody statistického modelu strojového překladu s dalšími přístupy (STEIN, 2013, str. VII). Díky neustálému vývoji SP bylo nutné začít se v novém tisíciletí intenzivněji zabývat otázkou kvality, a tak vznikají modely pro ruční a automatické hodnocení strojového překladu. Nejpoužívanější metrikou pro automatické hodnocení překladu je BLEU neboli *Bilingual Evaluation Understudy* (ZHANG & LUI, 2014, str. 109). Hodnocením kvality výstupů ze strojových překladačů se bude blíže zabývat kapitola 2.3.

V současné době je strojový překlad již samozřejmou součástí naší uživatelské zkušenosti. Při vyhledávání na internetu jej pro informativní účely využívají miliony lidí po celém světě. Mezi podporované jazyky jsou neustále přidávány další, například překladač Google Translate v současnosti nabízí již 103 jazyků včetně latiny a esperanta¹. A právě již

¹ Online [cit. 2017-08-11] z: http://technet.idnes.cz/google-translate-neural-networks-dlq-/sw_internet.aspx?c=A170418_224649_sw_internet_pka

zmíněný Google Translate je z části představitelem takzvaného neuronového strojového překladače² (viz oddíl 2.1.3.4), což je v současné době nejnovější přístup k SP.

2.1.3 Přístupy ke strojovému překladu

Odborná literatura³ tradičně rozlišuje tři způsoby, jakými je možné přistupovat ke strojovému překladu:

- 1) *Rule-Based Machine Translation* (RBMT), pravidlové systémy;
- 2) *Corpus-Based Machine Translation* (CBMT), korpusové systémy;
- 3) *Hybrid Machine Translation* (HMT), hybridní systémy.

Jak již bylo uvedeno v kapitole 2.1.2, nejnovějším přístupem k SP je *Neural Machine Translation* (NMT), tedy *neuronový strojový překlad*. V této kapitole budou popsány všechny čtyři zmíněné přístupy. Kapitola se však v žádném případě nesnaží podat detailní popis těchto systémů. Pro účely této práce postačí, když jednotlivé přístupy stručně charakterizujeme, krátce popíšeme, jak fungují a nastíníme jejich výhody a nevýhody.

2.1.3.1 Pravidlový strojový překlad (RBMT)

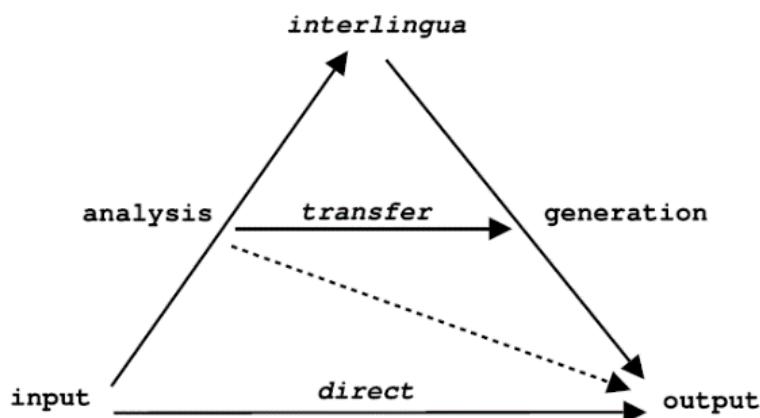
Rule-Based Machine Translation (RBMT), tedy *strojové překladače založené na pravidlech*, jsou také někdy označovány jako teoretické (MAČURA, 2011, str. 218). Bojar (2012, str. 24) definuje RBMT jako systémy, které využívají formální gramatiku, ručně vyrobené slovníky a jejichž postup se dá rozdělit na analýzu, transfer a syntézu. Nejprve proběhne analýza výchozího textu na různých úrovních (například na syntaktické, morfologické atd.), v rámci transferu dojde k přeložení struktur získaných analýzou do struktur cílového jazyka, nakonec je provedena syntéza textu v cílovém jazyce (TÝNOVSKÝ, 2007, str. 10n).

Během analýzy výchozího textu a syntézy cílového textu aplikuje RBMT morfologická, syntaktická a sémantická pravidla. Systém využívá tato předem definovaná lingvistická pravidla a díky nim se například snáze vyrovná s problémy v oblasti pořádku slov (OLADOSU et al., 2016, str. 121).

² Online [cit. 2017-08-11] z: <https://research.googleblog.com/2016/09/a-neural-network-for-machine.html>

³ Viz Mačura (2011, str. 218), Okpor (2014, str. 160nn), EuroMatrix (2007, str. 18nn)

Jak již bylo naznačeno v oddíle 2.1.2, jako překladače založené na pravidlech jsou v dnešní době souhrnně označovány tyto koncepty: model přímého překladu (*direct translation model*), interlingvální model (*interlingua model*) a *transfer model*, tedy transferový model (HUTCHINS W. , 2003, str. 504). Všechny tyto tři modely přehledně znázorňuje trojúhelník Bernarda Vauquoise (Obrázek 1).



Obrázek 1: Vauquoisův trojúhelník (GIMÉNEZ, 2008, str. 4)

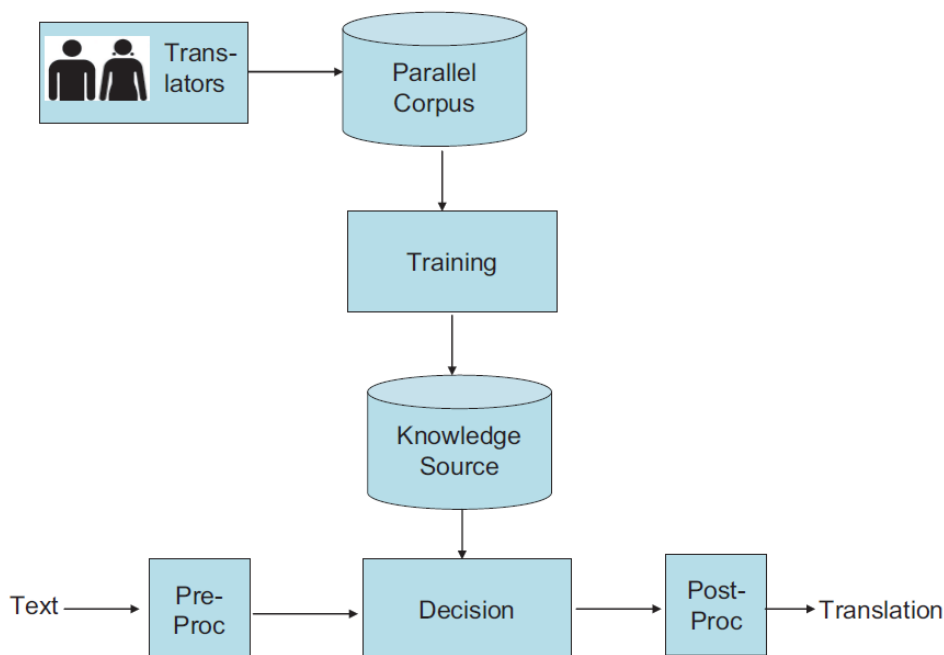
V dolní části trojúhelníku vidíme model přímého překladu, nejstarší a nejméně oblíbený koncept, který funguje na úrovni jednotlivých slov. Systém provede pouze stručnou syntaktickou a sémantickou analýzu a překládá slovo za slovo pouze s minimálními gramatickými úpravami. Na vrcholu trojúhelníku je znázorněn interlingvální model, sofistikovanější lingvistický koncept druhé generace. Výchozí text je transformován do interlingvy, jazykově neutrální a nezávislé reprezentace, z níž je poté generován cílový text. Tento přístup využíval například komerční systém KANT. Transferový model, který můžeme vidět ve střední části trojúhelníku, má s interlingválním společné to, že překlad vytvoří na základě reprezentace, která simuluje význam výchozí věty. Na rozdíl od předchozího přístupu je však tento model závislý na jazykovém páru zapojeném v překladu (OKPOR, 2014, str. 161n).

Hlavní výhodou pravidlových systémů je predikovatelnost výstupu. To znamená, že máme přesnou kontrolu nad tím, co se jakým způsobem analyzuje a jak se to překládá. Chyby lze proto poměrně snadno odstranit pomocí úprav. Bojar (2012, str. 25) poukazuje na skutečnost, že RBMT mohou díky specializovanosti jednotlivých komponent vykazovat dobré výsledky například u dvojic typologicky vzdálených jazyků jako angličtina a japonština nebo u jazykových párů s pevným, ale velmi rozdílným slovosledem. Co se týče nevýhod těchto

systemů, je nutné zmínit především velkou finanční a časovou zátěž (COSTA-JUSÁ et al., 2012, str. 249).

2.1.3.2 Korpusové systémy (CBMT)

Pro korpusové systémy (*Corpus-Based Machine Translation*, CBMT) je v odborné literatuře možné nalézt i označení empirické systémy nebo také anglický název *data driven machine translation* (OKPOR, 2014, str. 162). Jak již prozrazuje název, CBMT pracuje s dvojjazyčným paralelním korpusem obsahujícím originály a překlady, které jsou využívány pro získávání znalostí potřebných pro překlad. Práce jazykových expertů je tedy v tomto případě nahrazena automatickou extrakcí znalostí z dat obsažených v korpusu (TÝNOVSKÝ, 2007, str. 11). Obrázek 2 znázorňuje schéma strojového překladač zakládajícího se na korpusu.



Obrázek 2: Schéma CBMT (EuroMatrix, 2007, str. 19)

V rámci strojového překladač na základě korpusu odborná literatura tradičně rozlišuje dva přístupy:

- *statistický strojový překlad (Statistical Machine Translation, SMT);*
- *strojový překlad založený na příkladech (Example-Based Machine Translation, EBMT).*

Nápad na vytvoření strojového překladač založeného na příkladech představil japonský profesor Makoto Nagao na Světovém sympoziu NATO o umělé a lidské inteligenci již v roce

1981 (ANASTASIOU, 2010, str. 17), o tři roky později popsal metodu, která by se dala shrnout pomocí následujících třech kroků⁴:

- 1) *matching*: rozdělení věty výchozího jazyka na fragmenty a nalezení jim podobných příkladů v databázi s příklady;
- 2) *alignment*: přeložení fragmentů do cílového jazyka analogicky podle podobného příkladu;
- 3) *recombination*: poskládání přeložených fragmentů do věty v cílovém jazyce.

EBMT nepracuje se slovy, nýbrž s frázemi a snaží se napodobit lidského překladatele tím, že usiluje o překlad jednoduchých frází výchozího textu pomocí analogie s již existujícím překladem podobných frází z databáze příkladů (TÝNOVSKÝ, 2007, str. 13n).

Anastasiou (2010, str. 47) upozorňuje na skutečnost, že přestože EBMT a překladová paměť (TM) vykazují několik společných rysů, mezi něž lze uvést například *matching* fragmentů nebo využívání již existujících překladů, neměli bychom tyto dva koncepty zaměňovat. TM je nástroj pro překladatele, který je součástí systémů s překladovou pamětí (viz CAT v oddílu 2.1.1). Překladová paměť slouží pouze jako pomůcka, záleží tedy jen na překladateli, zda navrhanou shodou přijme, nebo ne. Naproti tomu u strojového překladu založeného na příkladech proběhnou automatické procesy, během nichž je cílový text rovnou vygenerován.

Mezi hlavní výhody EBMT patří skutečnost, že funguje na základě korpusu, není tedy potřeba pro něj definovat pravidla manuálně, jako je tomu u RBMT, tím pádem není vytváření EBMT tak časově náročné. Tyto systémy mají samozřejmě i své nevýhody. Často je zmiňovaná například výpočetní náročnost (OKPOR, 2014, str. 164) nebo také možnost sestavení výstupu z příkladů, které spolu vzájemně nesouvisí, tento jev je označován jako *boundary friction* (TÝNOVSKÝ, 2007, str. 20).

Jak již bylo zmíněno výše, ke korpusovým systémům patří také statistický strojový překlad (SMT). Jako SMT označuje Bojar (2012, str. 26) systémy, které nesou tyto typické rysy:

- a) *trénování*⁵ dat: slovníky pro SMT nejsou vytvářeny ručně, statistické systémy získávají potřebné informace automaticky díky paralelním i jednojazyčným korpusům;

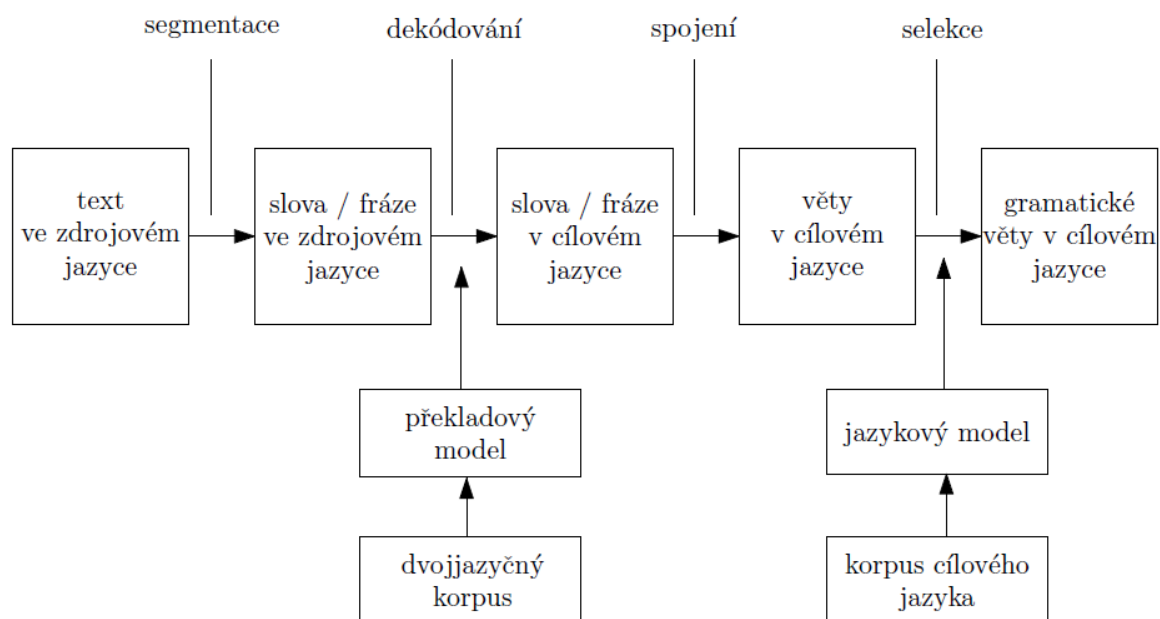
⁴ Podrobněji viz Nagao (1984, str. 178nn)

⁵ S výrazem *trénování* se v oblasti SP můžeme setkat poměrně často. V tomto kontextu se jedná o běžně používaný termín převzatý z angličtiny (*to train*). Tréninkový proces umožňuje přechod od poskytnutých trénovacích dat

- b) práce s mnoha variantami a přesná procedura volby vítězného kandidáta: statistické strojové překladače pracují s několika variantami výstupu, ty pak ohodnotí a na základě tohoto hodnocení vyberou vítězného kandidáta.

V rámci hodnocení využívá statistický SP metodu podmíněné pravděpodobnosti $p(e | f)$. Ta vyjadřuje, jak velká je pravděpodobnost, že reprezentace e v cílovém jazyce je překladem vstupní reprezentace f . Statistický model zvolí ze všech cílových reprezentací e tu, která je nejpravděpodobněji překladem zdrojové věty f (BOJAR, 2012, str. 25n).

SMT pracuje se dvěma modely, které je možné natrénovat pomocí korpusu, jedná se o *jazykový model cílového jazyka* a *překladový model*. K natrénování prvního jmenovaného stačí jednojazyčný korpus. V rámci tohoto modelu je větě cílového jazyka přiřazena tím vyšší pravděpodobnost, čím více se v nějakém ohledu blíží plynulé gramatické větě daného jazyka. Druhý model je možné natrénovat pomocí dvojjazyčného korpusu. Čím si je dvojice vět výchozího a cílového jazyka významově blíží, tím vyšší pravděpodobnost je této dvojici přiřazena (TÝNOVSKÝ, 2007, str. 11n). Obrázek 3 ukazuje, jak jsou tyto dva modely zapojeny během překladu prováděného statistickým strojovým překladačem.



Obrázek 3: Schéma statistického systému (TÝNOVSKÝ, 2007, str. 15)

V rámci statistického strojového překladu je možné rozlišit systémy založené na slovech (*word-based*), systémy založené na frázích (*phrase-based*) a systémy založené na syntaxi

(korpusu) k překladovému modelu. Korpusy je však nutné před zahájením samotného tréninku předzpracovat. Toto předzpracování se skládá z následujících kroků *tokenizace*, *true casing* a *očistění*. Více viz Adamec (2015, str. 22nn).

(*syntax-based*). Původní statistické strojové překladače byly založeny na korespondenci mezi jednotlivými slovy výchozího a cílového jazyka. Frázové modely již pracují se sekvencemi po sobě následujících slov. Statistické překladače založené na syntaxi se při analýze nezaměřují na jednotlivá slova nebo fráze, nýbrž na celé syntaktické jednotky (AHMED & HANNEMAN, 2005, str. 2nn).

Nejběžnější variantou SMT je frázový statistický překlad (*phrase-based statistical machine translation*, PBSMT), který pracuje se slovy jako izolovanými a nedělitelnými jednotkami (BOJAR, 2012b, str. 489). PBSMT byl postupně vylepšován, a tak vznikl například hierarchický frázový překlad (*Hierarchical Phrase-based SMT*). S nástupem frázových SMT se začaly používat delší fráze, takže se PBSMT začal blížit EBMT a v mnohých případech se rozdíl stíraly (TÝNOVSKÝ, 2007, str. 15).

Oproti pravidlovým přístupům se dá SMT snáze vytvořit, není třeba definovat žádná lingvistická pravidla a systém navíc funguje nezávisle na jazykových párech. Přes tyto nesporné výhody musí statistický strojový překlad čelit i několika výzvám. Vyžaduje množství paralelních textů, je výpočetně náročný a při překladu mezi jazykovými páry s velmi rozdílným pořádkem slov a morfologií mohou vyvstat problémy (COSTA-JUSÁ et al., 2012, str. 253).

Co se týče budoucnosti SMT, bývá zmiňováno, že se statistický strojový překlad začíná blížit svým limitům. V této souvislosti se často hovoří o tzv. *bodu degradace*⁶ (SVOBODA, 2015, str. 258). Úspěšnost statistického SP do značné míry závisí na velikosti paralelních korpusů a také na jejich kvalitě. Pokud bude v těchto korpusech větší zastoupení dat přeložených strojově bez posteditačních úprav, hrozí, že se kvalita výstupů ze SP bude snižovat. Budoucnost tedy pravděpodobně patří spíše hybridním systémům nebo neuronovému strojovému překladu. NMT pracuje s paralelními daty, proto ho lze zařadit ke korpusovým systémům. V této práci je však popsán samostatně v oddílu 2.1.3.4.

2.1.3.3 Hybridní strojový překlad (HMT)

Každý z výše jmenovaných přístupů k SP má své výhody a nevýhody. Této skutečnosti jsou si vývojáři SP dobře vědomi, a proto vytváří hybridní strojové překladače (*Hybrid Machine Translation*, HMT), které se snaží využívat výhod jednotlivých přístupů tak, že kombinují dva

⁶ Franz Josef Och, bývalý hlavní vývojář Google Translate, v tomto případě hovořil o tzv. *Google Time Loop*, poznamenává však, že se snaží tomuto problému předejít, a proto vyvinuli algoritmus, který má rozpoznávat strojově přeložené texty (RUBENS, 2014).

nebo dokonce více systémů. Například Lingstat je hybridní SP, který kombinuje statistický a lingvistický přístup, METIS-II propojuje dokonce přístupy tři, konkrétně SMT, RBMT a EBMT (EuroMatrix, 2007, str. 41).

Při vytváření NMT nejčastěji dochází ke kombinaci statistického a pravidlového přístupu (OLADOSU et al., 2016, str. 121). V rámci hybridizace může být systém řízen RBMT, v tomto případě je pak možné do pravidlového SP zavést korpus, nebo výstup posteditovat pomocí statistiky. Hybridní systém může být samozřejmě také postaven na korpusovém přístupu, pravidla se do něj pak dají integrovat dvěma způsoby, buď jejich použitím při předzpracování a/nebo závěrečných úpravách (*pre-processing*, *post-processing*), nebo implementací pravidel či slovníků přímo do CBMT (COSTA-JUSSÁ & FONOLLOSA, 2015, str. 5n).

2.1.3.4 Neuronový strojový překlad (NMT)

Nejmladším přístupem k SP je neuronový strojový překlad (*Neural Machine Translation*, NMT), který byl představen roku 2013 (KALCHBRENNER & BLUNSOM). Jak jsme již poznamenali v oddílu 2.1.3.2, kdybychom vycházeli z tradičního rozdělení, bylo by možné NMT zařadit ke korpusovým systémům. K trénování totiž potřebuje paralelní data a využívá pokročilých metod z oblasti statistiky a strojového učení. Přísně vzato by tedy šlo na neuronový SP nahlížet jako na podtyp SMT. Ve většině odborných článků je však vnímán spíše jako další podtyp korpusových systémů na úrovni SMT⁷. Pro účely této práce bude NMT věnován samostatný oddíl.

Pokud bychom hledali co nejjednodušší odpověď na otázku, co to vlastně NMT je, mohla by znít takto:

„Neural Machine Translation is the approach of modelling the entire MT process via one big artificial neural network.“ (MANNING & SOCHER, 2017)

Dá se tedy říci, že neuronová síť je naprogramovaný model, který simuluje biologické struktury neuronů v živých organismech a je schopen učit se prostřednictvím tzv. *deep learning* z dostupných dat.

Pro lepší představu bychom mohli porovnat NMT s frázovým statistickým překladem. Kdybychom tak učinili, zjistili bychom, že zatímco SMT pracuje s mnoha malými

⁷ Uvedené informace byly získány na základě emailové konzultace s Mgr. Martinem Popelem z Ústavu formální a aplikované lingvistiky MFF UK.

subkomponenty, snaží se neuronový SP vytvořit a natrénovat velkou jednotnou neuronovou síť, která je schopná pojmout celou větu a vygenerovat její správný překlad (BAHDANAU et al., 2014, str. 1). Tento přístup se tedy snaží větu analyzovat jako celek, a pak ji jako celek také přeložit.

Většina neuronových strojových překladačů je založena na tzv. *encoder-decoder* modelu, který využívá rekurentní neuronové sítě (*recurrent neural network*, RNN) k tomu, aby byl schopen pojmout sekvenci výchozího jazyka a předpovědět sekvenci v cílovém jazyce. Zdrojová věta je nejdříve zakódována do vektoru s pevnou délkou. Během dekódování je pak generován překlad ze zakódovaného vektoru (BAHDANAU et al., 2014, str. 1).

Přestože se jedná o poměrně nový přístup, odborníci se shodují, že podává slibné výsledky. Ovšem aby NMT správně fungoval, musí mít k dispozici velký paralelní korpus. Pokud nemá dostatek dat k tréninku, kvalita se zhoršuje. A právě absence dostatečně velkých paralelních korpusů je problém, se kterým se musí potýkat mnoho jazykových párů, a to nejen malé jazyky jako je baskičtina, ale dokonce i jazykové páry některých velkých jazyků, například kombinace ruština a němčina (ARTETXE & CHO, 2017, str. 1).

Koehn a Knowles (2017, str. 28) formulovali ještě další výzvy, se kterými se NMT bude muset do budoucna vyrovnat. Neuronový strojový překlad vykazuje horší kvalitu při překladu velmi dlouhých vět, problémy se mohou objevit i při překladu méně frekventovaných slov. V některých případech obětuje adekvátnost (*adequacy*) na úkor plynulosti (*fluency*). Co se týče výhod neuronového SP, je možné jmenovat právě větší plynulost generovaného překladu či užívání širšího kontextu za účelem vyšší přesnosti (MANNING & SOCHER, 2017, str. 16).

2.2 Veřejně dostupné překladače

V empirické části této práce budou analyzovány a porovnány výstupy ze dvou veřejně dostupných překladačů. Právě z tohoto důvodu se bude tato kapitola zabývat online překladači, nejdříve obecně, a pak zcela konkrétně těmi, které byly vybrány pro účely této diplomové práce. Jedná se o:

- Bing Translator (<https://www.bing.com/translator>);
- Google Translate (<https://translate.google.cz/>).

Volba těchto překladačů bude zdůvodněna v metodologické části práce. Překladače Google Translate a Bing Translator patří k nejznámějším a pravděpodobně i nejužívanějším

online strojovým překladačům. Při popisu těchto systémů není možné jít do hloubky, protože detailní informace o tom, jakým způsobem fungují, jsou součástí know-how, které vývojáři neprozrazují.

2.2.1 Online překladače obecně

Již od 90. let se na internetu začaly objevovat překladače online. Jedním z prvních byl Babel Fish, který byl spuštěn 9. prosince roku 1997. Vznikl díky spolupráci společnosti Systran Software Inc. a v té době populárního vyhledávače AltaVista. Babel Fish byl unikátní v tom, že byl na rozdíl od staršího komerčního překladače CompuServe zdarma a dostupný všem uživatelům internetu (GASPARI & HUTCHINS, 2007, str. nečíslováno). V roce 2003 byl překladač Babel Fish prodán společnosti Yahoo! a v roce 2012 nahrazen překladačem Bing Translator, který vyvinula a vlastní společnost Microsoft (POIBEAU, 2017, str. 227).

Asi nejznámějším překladačem dostupným na internetu je Google Translate. S výzkumem v oblasti SP začal Google již po roce 2000. Překladač od Googlu nejdříve využíval při překladu technologii od společnosti Systran, ten byl ale postupně nahrazován systémem, který si Google vyvinul sám. Nejdříve byl spuštěn pro ruštinu, arabštinu a čínštinu a od podzimu roku 2007 již byl schopen překládat mezi 25 jazykovými páry. Co se týče dalších překladačů, mohli bychom jmenovat například již zmíněný Bing Translator, Babylon, Systranet společnosti Systran nebo jeho největšího konkurenta na trhu Prompt (POIBEAU, 2017, str. 228).

Volně dostupné překladače si mezi širokou veřejností získávají čím dál větší oblibu. Uživatelské přízně se těší především proto, že jsou schopné generovat překlady velmi rychle a bezplatně⁸. Přestože kvalita výstupů SP nemusí být ve všech ohledech dostačující, strojový překlad poskytne alespoň základní přehled o tom, o čem pojednává originál, což je pro uživatele online překladačů v mnoha případech dostačující. S online systémy lze navíc vytvářet překlady mezi mnoha jazykovými páry.

Veřejně přístupné strojové překladače mají samozřejmě i své stinné stránky. Za zmínku určitě stojí skutečnost, že online SP v žádném případě nezaručují zachování důvěrnosti vložených informací. Každý uživatel by si měl být vědom toho, že všechna data, která poskytne,

⁸ Pokud překladačelé používají Google Translate v nástroji CAT, mají k dispozici 1 000 000 znaků za 20\$. Více na: <https://cloud.google.com/translate/pricing> [cit. 2017-11-16]

budou zpracována a uchována veřejně dostupnou službou SP. Volně dostupné překladače jsou většinou zaměřené obecně, jen málo z nich zatím nabízí možnost filtrování podle toho, jak je text zaměřen obsahově. Jednu z výjimek tvoří například překladač My Memory⁹, který nabízí možnost výběru z více než 30 tematických okruhů.

Problém může nastat i tehdy, pokud je překládaný text příliš dlouhý, některé překladače jsou totiž omezeny na určitý počet znaků. Například překladač Dictionary.com využívající pro překlad technologii od Microsoftu, umožňuje přeložit najednou pouze 300 znaků, delší text je tedy třeba vkládat po částech. Tyto překladače navíc ve většině případů¹⁰ vyžadují připojení k internetu, v off-line režimu není možné je používat. Je důležité si uvědomit, že kvalita výstupů SP do jisté míry závisí i na kvalitě originálu, pokud bude výchozí text defektní, nemůžeme očekávat valnou kvalitu ani u cílového textu.

2.2.2 Google Translate

Jak již bylo uvedeno výše, překladač Google Translate původně při překladu používal technologii založenou na pravidlech od Systranu. Výzkumný tým Googlu vedený Franzem Josefem Ochem však brzy začal pracovat na vlastní technologii, která měla při překladu uplatňovat statistický přístup (LEE, 2017, str. 106). Tento nový systém postupně nahradil původní technologii u všech jazykových párů.

Google Translate (GT) v současné době podporuje 103¹¹ jazyků včetně například mezinárodního umělého jazyka esperanta a každý den přeloží více než 140 bilionů slov. Překladač byl původně založen na frázovém statistickém přístupu. V listopadu 2016 byl u Google Translate spuštěn neuronový strojový překlad pro několik prvních jazyků. Technologie NMT u Google Translate je postupně rozšiřována o další jazyky.

Na oficiálních stránkách s informacemi o Google Translate¹² bylo v době vzniku práce uvedeno, že NMT není pro český jazyk podporován. Na webové stránce idnes.cz v rubrice Technet byl však 19. dubna 2017 uveřejněn článek¹³ o tom, že od 18. dubna 2017 má být

⁹ Překladač nabízí mimo jiné filtrování dle specializovaných kategorií, jako je například chemie, informatika, letectví nebo kosmonautika. Dostupné z: <https://mymemory.translated.net/> [cit. 2017-11-17]

¹⁰ Například Google Translate nabízí možnost stažení jazyků k používání off-line. Tato možnost je však dostupná pouze v aplikaci Překladač Google, nikoliv v počítači. Více na: <https://support.google.com/translate#topic=7011755> [cit. 2017-11-17]

¹¹ Údaje o GT jsou platné k 18.11.2017 a pochází z: <https://translate.google.com/intl/en/about/languages/>

¹² <https://translate.google.com/intl/en/about/languages/> [cit. 2017-11-29]

¹³ https://technet.idnes.cz/google-translate-neural-networks-dlq-sw_internet.aspx?c=A170418_224649_sw_internet_pka [cit. 2017-11-29]

neuronový strojový překlad nově dostupný i pro češtinu. Článek explicitně neuvádí, odkud tyto informace pocházejí, autor ale během emailové komunikace odkazuje na rozhovor s produktovým manažerem Google Translate Barakem Turovskym. Tezi o spuštění NMT pro češtinu potvrzuje i fakt, že se v uživatelském rozhraní online překladače při najetí myší na slovo již nepodbarví jen dané slovo originálu a odpovídající slovo v překladu jako dříve, ale celé věty. NMT totiž překládá po celých větách a ne pouze po slovech nebo frázích. Také uživatelé diskuzní platformy *Google Translate Forum* uvádějí, že NMT bylo pro češtinu již spuštěno a oficiální stránka s informacemi o Google Translate není aktualizovaná¹⁴. V rámci empirické studie tedy budeme pracovat s hypotézou, že neuronový SP byl pro češtinu skutečně spuštěn v dubnu 2017.

Dříve, v rámci statistického frázového přístupu, probíhal překlad pomocí Google Translate minimálně přes jeden pivotní jazyk, a to většinou angličtinu. U jazyků, pro které je spuštěn NMT, již probíhá překlad přímo mezi daným jazykovým párem. V této souvislosti se hovoří o tzv. *zero-shot translation*¹⁵. Jedná se o schopnost systému překládat mezi dvěma libovolnými jazyky včetně jazykových párů, pro které systém nebyl trénován. Neuronové sítě překládají po celých větách a ne pouze po částech, což je velmi náročné na výpočetní výkon. Google proto používá vlastní hardware zvaný TPU, tedy *Tensor Processing Units* (WU et al., 2016, str. 2).

Při překladu nabízí GT místo ručního výběru jazyka také možnost rozpoznání jazyka. Dále je možné nechat si pomoci tlačítkem *Poslech výslovnosti* přehrát zadaný text ve zvoleném jazyce. Překladač umožňuje přeložit fráze, texty i celé webové stránky. Pole pro zadávání originálního textu je kapacitně omezeno na 5 000 znaků.

Pomocí obrazovkové klávesnice a okna pro rozpoznávání ručně psaného textu umožňuje GT pohodlnější zápis textu na dotykových obrazovkách. Po stisknutí tlačítka s hvězdičkou se překlad uloží do slovníků frází, což je seznam oblíbených překladů, který je vázaný na účet Google, přidávání tedy funguje až po přihlášení. Po kliknutí na přeložený segment se zobrazí alternativní překlady. Kliknutím na alternativní výstup dojde k záměně za dříve nabídnutý překlad. Pokud je uživatel s výstupem nespokojen, může pomocí tlačítka *Vylepšit překlad* nebo *Navrhněte úpravu* vložit vlastní verzi a tím službě pomoci se zlepšovat.

¹⁴ <https://productforums.google.com/forum/#!/topic/translate/CTUVkzsDF5Y;context-place=forum/translate> [cit. 2017-11-30]

¹⁵ Více na: <https://research.googleblog.com/2016/11/zero-shot-translation-with-googles.html> [cit. 2017-11-30]

Google nabízí také aplikaci do mobilního telefonu, jejíž pomocí je možné překládat mezi všemi 103 jazyky, off-line mezi 59 jazyky, dále umožňuje překlad z fotoaparátu, okamžitý obousměrný překlad mluveného slova ve 32 jazycích nebo překlad textu napsaného rukou¹⁶.

2.2.3 Bing Translator

Microsoft začal s vývojem statistického strojového překladače pro interní účely již roku 2002. O pět let později byl jeho překladač zpřístupněn široké veřejnosti na webové stránce společnosti. Původně nesl název Windows Live Translator. V roce 2009 představil Microsoft vyhledávač Bing společně s překladačem nazvaným Bing Translator (BT). Jak již bylo uvedeno, roku 2012 byl Babel Fish společnosti Yahoo! nahrazen právě tímto překladačem, do kterého pak byly integrovány všechny jazyky, se kterými Babel Fish pracoval (SIN-WAI, 2016, str. 211).

Bing Translator podporuje v současné době více než 60 jazyků včetně například klingonštiny, uměle vytvořeného jazyka fiktivní rasy Klingonů ze seriálu Star Trek. Microsoft pro své překladače také vyvinul neuronový strojový překlad, ten je zatím spuštěn celkem pro 21¹⁷ jazyků včetně češtiny, která byla nově přidána spolu s dalšími devíti jazyky v polovině listopadu 2017. V této době byl sběr materiálu pro empirickou studii již ukončen. Vývojáři z Microsoftu se domnívají, že díky tomuto přístupu je překlad ve většině případů plynulejší a bližší lidskému překladu než výstup ze statistického překladače. Zvýšení kvality by mělo být znatelné i v případě překladu pomocí hybridního přístupu, tedy v situaci, kdy je NMT spuštěno jen pro jeden z jazyků, z nebo do kterého je překládáno.

Stejně jako je tomu u Google Translate, není možné vložit do pole pro originální text více než 5 000 znaků. BT je schopný automaticky rozpoznat jazyk, ze kterého překládá. Bing Translator umí podobně jako GT předčítat originální i přeložený text, momentálně se však vypořádá pouze s rozšířenějšími jazyky, čeština zatím chybí. Systém dále nabízí možnost ohodnotit kvalitu překladu tím, že ho uživatel označí jako dobrý, nebo ho nahlásí jako

¹⁶ Více informací na: <https://play.google.com/store/apps/details?id=com.google.android.apps.translate&hl=cs> [cit. 2017-11-18]

¹⁷ Údaje o BT jsou platné k 18.11.2017 a pochází z: <https://blogs.msdn.microsoft.com/translation/2017/11/15/microsoft-translator-accelerates-use-of-neural-networks-across-its-offerings/> Jedná se o tyto jazyky: angličtina, arabština, bulharština, čeština, čínština, dánština, francouzština, hindština, italština, japonština, korejština, němčina, nizozemština, norština, polština, portugalština, rumunština, ruština, španělština, švédština a turečtina.

nesprávný. Výstup z překladače je také možné pomocí tlačítka rovnou sdílet na různých sociálních sítích.

Překladač je dostupný také jako aplikace do mobilního telefonu, kterou lze stáhnout pod názvem Microsoft Translator. Tato aplikace je schopná strojově přeložit obsah zadaný pomocí klávesnice, mluvené slovo nebo text, na nějž uživatel míří fotoaparátem.

2.3 Hodnocení strojového překladu

Výzkum strojového překladu byl téměř od svého počátku¹⁸ doprovázen výzkumem v oblasti kvality MT. Při posuzování strojového překladu je totiž kvalita ústředním pojmem. A právě vyhodnocováním strojového překladu (*evaluation of machine translation*) se bude zabývat tato kapitola. Vzhledem k odlišnému pohledu na účel strojového překladu jsou přístupy k měření kvality SP velmi různorodé. První část této podkapitoly se bude zabývat metodami hodnocení kvality. V další části se zaměříme na možné přístupy k hodnocení strojového překladu. Nakonec budou v podkapitole 2.3.3 popsány typické chyby, které se objevují ve výstupech ze SP.

Jak již bylo naznačeno výše, hodnocení SP je vědeckým tématem, v jehož rámci se sice z velkého množství různých metod prosadilo několik typů, zatím ale stále neexistuje žádný standardizovaný nebo jednotný přístup:

„What may be surprising is that despite some forty years of research on MT there is still no generally accepted methodology for the evaluation of systems.“ (HUTCHINS & SOMERS, 1992, str. 161)

Vyhodnocování MT je bezpochyby úlohou velmi složitou. Jazyk je velmi variabilní, neexistuje tedy pouze jeden správný překlad. Záleží také na tom, za jakým účelem výstup ze strojového překladače vzniká. Právě v této souvislosti vyvstává otázka, jak dobrá kvalita je „dost dobrá“. Dále bychom neměli zapomenout na to, že lidští hodnotitelé často tíhnou k subjektivním soudům. Některé studie jsou prováděny osobami, které se v problematice strojového překladu vyznají jen velmi málo, anebo vůbec, nejsou tedy schopny posoudit, co je možné a co je nerealistické (HUTCHINS & SOMERS, 1992, str. 161).

Zamysleme se krátce nad otázkou, proč je důležité se kvalitou strojového překladu zabývat. Jako uživatelé si na základě hodnocení například můžeme snáze vybrat, který systém

¹⁸ Srovnej Miller a Beebe-Center (1956) nebo Pfafflin (1965)

zakoupíme. Pomocí evaluace různých systémů můžeme třeba zjistit, kolik času při zapojení MT ušetříme, z kterého softwaru získáme kvalitnější výsledky nebo nejvhodnější výstup k posteditaci. Jako vědci a tvůrci systémů můžeme prostřednictvím hodnocení kvality MT zjišťovat, který přístup a která konfigurace vykazuje větší potenciál. Bojar (2012, str. 105) upozorňuje na skutečnost, že je třeba si vždy pečlivě vybrat metodu, abychom skutečně měřili to, co je pro konkrétní účel důležité. Navíc bychom měli mít na paměti, že každá metoda má své nepřesnosti a specifická úskalí.

2.3.1 Metody hodnocení strojového překladu

Co se týče metod měření kvality strojového překladu, rozlišujeme metody automatické a metody ruční. Většina metod automatického hodnocení se opírá o existenci jednoho nebo dokonce více referenčních překladů, se kterými je výstup ze strojového překladače srovnáván (BOJAR, 2012, str. 105). Ruční hodnocení kvality provádí bilingvní hodnotitelé (evaluátoři), v tomto případě tedy osoby, které ovládají jak zdrojový, tak cílový jazyk, případně monolingvní hodnotitelé, kteří ovládají cílový jazyk a mají k dispozici referenční překlad (KOEHN, 2010, str. 218).

2.3.1.1 Automatické hodnocení

Automatické metody pro hodnocení strojového překladu jsou souhrnně označovány jako *metriky kvality strojového překladu (MT evaluation metrics)*, přestože se z matematického hlediska o metriky v pravém slova smyslu nejedná, neboť nespĺňují například axiom trojúhelníkové nerovnosti (BOJAR, 2012, str. 112). Automatické metriky mají za úkol hodnocení zrychlit a také se vyrovnat s hlavními nevýhodami ručního hodnocení, mezi něž mimo jiné patří subjektivita a nereprodukovatelnost, jak je popsáno v oddílu 2.3.1.2. Existuje celá řada automatických metod, v této práci ale nebude možné obsáhnout všechny, proto zde budou krátce představeny jen ty nejužívanější.

Jednou z nejnámějších, nejpoužívanějších a nejstarších metrik je BLEU, tedy *Bilingual Evaluation Understudy*. Tato metrika byla vyvinuta již roku 2001 K. Papinenim a jeho týmem pro IBM, představena však byla až o rok později (PAPINENI et al., 2002, str. 311nn). Je založena na myšlence, že čím více se strojový překlad blíží profesionálnímu lidskému překladu,

tím je lepší. BLEU měla být alternativou k ručnímu hodnocení, kterou je možno používat v případech, kdy jsou vyžadována rychlá a častá hodnocení (EuroMatrix, 2007, str. 51).

BLEU pomohla mnoha výzkumným týmům ve zrychlení cyklu vývoje překladového systému, dále se také prokázalo, že se výborně hodí pro automatické ladění systému. Mezi hlavní výhody této metriky patří relativní jazyková nezávislost a jednoduchost (BOJAR, 2012, str. 113). Jedná se o metriku tzv. *modifikované n-gramové přesnosti*. Způsob, jakým BLEU funguje, se dá nejlépe vysvětlit použitím n o velikosti 1, tedy pomocí tzv. *unigramu*. Metrika prochází větu ze strojového překladače slovo po slově a pro každé slovo počítá množství výskytů daného slova v referenčních překladech. Poté, co zjistí počet výskytů v referenčních překladech, označí ve strojovém překladu tolik výskytů daného slova, kolik výskytů se objevilo v referenčních překladech. Následně vypočítá podíl: označená slova/celkový výskyt (ADAMEC, 2015, str. 19).

Metrika BLEU tedy vyhodnocuje překlady podle toho, kolik n -gramů z výstupu je potvrzeno referenčním překladem (CER et al., 2010, str. 555). Pro každé n vypočítá přesnost překladu n -gramů. Právě díky jejich použití je metrika schopná správně ohodnotit plynulost hypotézy (MACHÁČEK, 2012, str. 13). Prostřednictvím tzv. *pokuty za stručnost (brevity penalty, BP)* penalizuje systémy, které vyprodukují příliš krátký výstup (BOJAR, 2012, str. 114). Metrika také počítá se skutečností, že pro každou větu bude k dispozici více referenčních překladů.

NIST (*National Institute of Standards and Technology*) je metrika, která je založena na stejném principu jako BLEU. Také pracuje s metodou porovnávání n -gramů, avšak na rozdíl od metriky BLEU, která přisuzuje všem shodným výskytům stejnou váhu, počítá NIST, jak informativní daný n -gram je (OLADOSU et al., 2016, str. 123).

Akronym METEOR je odvozen od *Metric for Evaluation of Translation with Explicit Ordering* (CER et al., 2010, str. 555). Tato metrika byla vyvinuta na Carnegie Mellon University (EuroMatrix, 2007, str. 53). METEOR se snaží vyvarovat některých chyb, kterých se dopouští metrika BLEU. Ta například penalizuje použití synonym či sémanticky blízkých slov, protože nevykazují přesnou shodu. METEOR sice také využívá n -gramovou přesnost a shodu mezi referenčním a strojovým překladem, ale navíc zohledňuje shodu mezi slovy

založenými na stejném kořenu, mezi parafrázemi a synonymy, k čemuž využívá databázi WordNet¹⁹ (KOEHN, 2010, str. 228).

Následující tři metriky jsou založeny na principu editační vzdálenosti, to znamená, že metrika počítá, kolik editací je nutné ve větě ze strojového překladače provést, abychom dostali referenční větu. WER (*Word Error Rate*) je založena na Levenshteinově²⁰ vzdálenosti. Počítá se množství úprav, jako je smazání, substituce nebo přidání slova. Nevýhodou této metriky je skutečnost, že vyžaduje bezchybný pořádek slov ve větě (OLADOSU et al., 2016, str. 123).

Věta ze strojového překladače se může od referenčního překladu lišit například pořadím celých bloků slov. Přestože lidský hodnotitel by tento překlad mohl označit jako kvalitní, metrika WER by ho penalizovala. Tento problém se snaží vyřešit metrika *Translation Error Rate* (TER). Metrika je sice podobná té předchozí, ale k povoleným úpravám navíc kromě smazání, přidání a substituce připojuje i přesun celých bloků slov. Všechny tyto úpravy mají při hodnocení stejnou váhu (MACHÁČEK, 2012, str. 14).

Position-Independent Word Error Rate (zkráceně PER) také patří k metrikám založeným na editační vzdálenosti, na rozdíl od WER a TER však nezohledňuje pořadí slov. Tato skutečnost může být v některých případech i nevýhodou, protože PER se o pořádek slov ve větě nestará vůbec, takže ani nepenalizuje špatný pořádek slov (POPOVIĆ & NEY, 2011).

2.3.1.2 Ruční hodnocení

Existuje celá řada různých metod pro ruční hodnocení kvality, neustále vznikají další, které se snaží odstranit hlavní nevýhody těch stávajících, mezi něž patří především nízká mezianotátorská shoda, malá reprodukovatelnost (BOJAR, 2012, str. 105), vysoké náklady (OLADOSU et al., 2016, str. 122) a časová náročnost (POPOVIĆ & NEY, 2011, str. 667). I přes všechny tyto nedostatky jsou metody ručního hodnocení hojně využívány, protože automatické metriky nejsou vždy přesné a často ani nejsou schopny poskytnout dostatečný vhled pro analýzu chyb (LAVIE, 2010, str. 13).

Hutchins a Somers (1992, str. 163) se domnívají, že nejvhodnějšími kritérii pro hodnocení překladu jsou:

¹⁹ <http://wordnet.princeton.edu/> [cit. 2017-10-13]

²⁰ Srovnej Levenshtein (1965)

- (a) věrnost (*fidelity*) nebo přesnost (*accuracy*), kdy se zjišťuje, zda výchozí a cílový text obsahují stejnou informaci;
- (b) srozumitelnost (*intelligibility*) či jasnost (*clarity*) hodnotí, jak jednoduché je pro čtenáře porozumět překladu;
- (c) styl (*style*), v jehož rámci se hodnotí, jestli je jazyk překladu přiměřený svému obsahu a záměru.

Každé z kritérií však může být nezávislé, překlad věrný originálu může být například těžko pochopitelný anebo srozumitelný výstup nemusí být přeložen přiměřeným stylem (HUTCHINS & SOMERS, 1992, str. 163).

Stejná kritéria uvádí i Schwarzl (2001, str. 87) pouze s rozdílem, že při hodnocení stylu nezohledňuje obsah a záměr překladu, nýbrž skutečnost, zda překlad používá stejný jazyk (registr) jako originál. Arnold et al. (1994, str. 161n) je přesvědčen, že srozumitelnost (*intelligibility*) je tradičním kritériem pro hodnocení kvality překladu. Dalším důležitým faktorem je podle něj také přesnost, tedy *accuracy*, autorem též označovaná jako *fidelity*. Podle Arnolda et al. znamená vysoká srozumitelnost (*intelligibility*) obvykle také vysokou přesnost (*accuracy*), protože tato dvě kritéria spolu často úzce souvisí.

Koehn (2010, str. 218n) navrhuje zaměřit se v rámci ručního hodnocení kvality strojového překladu na posuzování plynulosti (*fluency*) a adekvátnosti (*adequacy*), tyto dva pojmy definuje následujícím způsobem:

- (a) plynulost (*fluency*): zkoumá, jestli je výstup přeložen plynulým jazykem, jestli je gramaticky správný a idiomatický;
- (b) adekvátnost (*adequacy*): kritérium hodnotí, jestli výstup zprostředkovává stejný význam jako originál a zda nebyla nějaká část informace ztracena, přidána nebo překroucena.

Dá se říci, že Koehnovo pojetí adekvátnosti (*adequacy*) se do jisté míry kryje s kritériem věrnosti (*fidelity*) či přesnosti (*accuracy*) u Hutchinse a Somerse, stejně tak bychom mohli srozumitelnost (*intelligibility*) či jasnost (*clarity*) souhrnně označit jako plynulost (*fluency*).

Obrázek 4 je příkladem evaluačního nástroje pro měření adekvátnosti a plynulosti strojového výstupu. Tyto škály jsou velmi vágní, a proto je pro hodnotitele často velmi těžké zůstat při hodnocení konzistentní.

Adequacy		Fluency	
5	all meaning	5	flawless English
4	most meaning	4	good English
3	much meaning	3	non-native English
2	little meaning	2	disfluent English
1	none	1	incomprehensible

Obrázek 4: Stupnice pro hodnocení plynulosti a adekvátnosti (KOEHN, 2010, str. 219)

Hodnocení pomocí plynulosti (*fluency*) a adekvátnosti (*adequacy*) se používalo například v rámci soutěže strojového překladu WMT (*Workshop on Statistical Machine Translation*), kde se však příliš neosvědčilo, protože shoda mezi anotátory byla nízká a hodnotitelé obě škály nevyužívali, neboť výsledky vykazovaly zbytečně těsnou korelaci (BOJAR, 2012, str. 106). V soutěži se lépe osvědčila metoda *uspořádávání hypotéz* (*hypothesis ranking*), v jejímž rámci mají lidé uspořádat překlady z různých překladačů od nejlepšího po nejhorsí, přičemž je povoleno označit případně některé systémy za stejně dobré.

K ověření toho, jestli čtenář výstupu ze strojového překladače rozumí, je možné použít *test srozumitelnosti vět* (*sentence comprehension*). Kdybychom čtenářům položili otázku „Rozumíte této větě?“, výsledky by se pravděpodobně daly těžko interpretovat, metoda má proto následující dva kroky: editaci naslepo a kontrolu adekvátnosti. Nejdříve dostane hodnotitel pouze výstup z překladače bez originálu a referenčního překladu, tento výstup musí opravit podle toho, co si myslí, že má daná věta říkat. V druhém kroku dostane druhý hodnotitel opravenou větu, originál i referenční překlad a má rozhodnout, zda je opravený překlad v pořádku, nebo vyjadřuje něco jiného než vstup (BOJAR, 2012, str. 107). Srozumitelnost překladu lze také ověřit pomocí odpovídání na kontrolní otázky na základě informací, které čtenář získá z překladu (SVOBODA, 2015, str. 251).

Další metodou pro ruční hodnocení kvality je měření času potřebného k *posteditaci* textu, tedy jak dlouho bude trvat, než opravíme překlad do přijatelné kvality. Před použitím této metody je potřeba stanovit, jakou kvalitu má mít opravený text, jestli má být publikovatelný, jestli je určen k interní komunikaci atd. Dále bychom měli brát v potaz, že chyby nemají vždy stejnou váhu, a proto se i čas potřebný k jejich opravě může lišit (EuroMatrix, 2007, str. 45n).

Alternativním přístupem k posteditaci je tzv. *značkování chyb* (*error flagging*). Hodnotitelé dostanou překlad a mají za úkol v něm vyznačit chyby podle předem dohodnuté

klasifikace (BOJAR, 2012, str. 108). Tuto klasifikaci chyb pro ruční hodnocení kvality SP představil Vilar et al. (2006).

Měření kvality můžeme také provést pomocí metody nazvané *cloze test*. Tato metoda byla poprvé použita k hodnocení MT již roku 1965 (BISHOP & CROOK). Hodnotitel dostane větu ze strojového překladače, v které chybí slovo, jež je nahrazeno mezerou. Hodnotitel se má snažit odhadnout, které slovo bylo vynecháno. Úspěch evaluátorů v odhadování chybějícího slova koreluje s tím, jak srozumitelný je výstup ze strojového překladače (EuroMatrix, 2007, str. 46).

Aiken (2010, str. nečíslováno) uvádí jako další způsob metodu *Round-trip Translation*, která se v literatuře objevuje též pod názvem *back-and-forth translation* nebo *reverse translation*, kdy je proveden zpětný překlad výstupu SP do jazyka originálu. Čím víc se opětovně přeložený text blíží originálu, tím je původní překlad kvalitnější.

U určitých textových typů se provádí *test náležitě odezvy*, tzv. *performance test*. Hodnotitelé dostanou výstup ze strojového překladače a mají jednat dle pokynů, které jsou v něm obsaženy. Kvalita výstupu se pak hodnotí na základě toho, jestli byly úkony provedeny správně. Tato metoda je relevantní například pro návody k použití (SVOBODA, 2015, str. 251).

Kvalitu SP je například také možné měřit pomocí *hodnocení na stupnici (rating technique)*, kdy jsou překlady hodnoceny na škále s různými hledisky. Stupnice může například obsahovat pouze dvě hlediska, kdy hodnotitelé rozhodují, jestli je překlad srozumitelný nebo ne, ale i deset hledisek či dokonce více (SVOBODA, 2015, str. 251). Další možností pro hodnocení kvality SP je lingvistická analýza (CORPAS PASTOR & MORENO ORTIZ, 1997, str. 708nn). Jednotlivé metody lze samozřejmě také kombinovat, jako tomu bude v empirické části této diplomové práci.

2.3.2 Přístupy k hodnocení strojového překladu

K hodnocení strojového překladu lze přistupovat různými způsoby. Svoboda (2015, str. 250) s odkazem na dizertační práci Bauerové (2002, str. 95nn) předkládá přehled možných přístupů, které jsou členěny do následujících dichotomií:

- 1) glass-box evaluation *x* black-box evaluation (posuzování, kdy hodnotitel dílčí automatické procesy MT zná a kdy je nezná);

- 2) mikroevaluace *x* makroevaluace (posouzení celkové použitelnosti systému versus hledání nedostatků systému a potenciálu k jeho zlepšování);
- 3) typologická (zabývá se tím, jaké jevy může MT zpracovat) *x* deklarativní (zkoumá výkon SP, například z hlediska srozumitelnosti nebo věrnosti originálu) *x* operativní evaluace (zaměřuje se na ekonomické hledisko, např. na nároky týkající se posteditace);
- 4) formální *x* funkční analýza (použití měřitelných parametrů *x* využití systému v konkrétním kontextu, jako je podpora poskytovatele nebo oborová specifická);
- 5) standard performance *x* optimum performance (minimální zásahy ze strany uživatele *x* využití všech možností systému);
- 6) hodnocení adekvátnosti (zkoumá splnění specifických požadavků ze strany systému) *x* hodnocení pokroku (testování, jestli se systém zlepšuje) *x* diagnostické hodnocení (zkoumání, kde selhal).

Svoboda (2015, str. 250) dodává, že v odborné literatuře lze najít ještě další dichotomie:

- 7) interní hodnocení (určené pro vlastní potřeby vývojářů) *x* komparační hodnocení (srovnání více modelů pro výběr toho nevhodnějšího);
- 8) hodnocení odborníkem *x* posouzení neoborným hodnotitelem;
- 9) použití již existujícího modelu *x* vytvoření vlastní metody.

2.3.3 Typické chyby

Strojový překlad se stává běžnou součástí naší uživatelské praxe na internetu. V překladačích se dají přeložit nejen slova, fráze, texty, ale i celé webové stránky. Téměř každý uživatel internetu má jistě nějakou zkušenost se SP. Stejně tak se každý z těchto uživatelů určitě alespoň jednou v životě setkal i s nějakým nekvalitním překladem vytvořeným pomocí SP. Chceme-li se zabývat otázkou kvality výstupů ze strojových překladačů, určitě je vhodné mít přehled o tom, které jevy jim obvykle činí potíže. V této části proto budou popsány typické chyby, které prozrazují rukopis strojového překladače.

Velmi tvrdým oříškem pro SP je mnohoznačnost výstupu. Například ve větě *Ženu holi stroj*, kterou uvádí Bojar (2012, str. 6), má každé slovo více než jeden význam. Tvar *ženu* lze chápat jako akuzativ od podstatného jména žena, nebo jako oznamovací způsob první osoby jednotného čísla od slovesa hnát. Sedmý pád od podstatného jména hůl a oznamovací způsob

třetí osoby jednotného čísla od slovesa holit mají stejný tvar (*holí*). *Stroj* zde může být jako podstatné jméno, nebo rozkazovací způsob od slovesa strojít. Věta se dá tedy vyložit dokonce čtyřmi způsoby:

- (1a) *Běžím za strojem a popoháním jej holí.*
- (1b) *Žena je holena strojem.*
- (1c) *Oblékej ženu pomocí hole.*
- (1d) *Oblékej ženu (mnoha) holí. Jako je Pán prstenů, může být žena holí.*

Homonymní výrazy a víceznačné věty mohou při překladu způsobit problémy i lidskému překladateli. Ten má však oproti SP tu výhodu, že je schopen zohlednit například kontext nebo aplikovat své znalosti ohledně dané problematiky a na základě toho větu přeložit správně. Soudobé strojové překladače však zatím kontext zohledňují málo.

Problematicky se při překladu mohou chovat i zájmena. Pokud mají ve větě zástupnou funkci, je nutné je překládat s ohledem na rod a číslo slova, které nahrazují. Pokud dva výrazy odkazují ke stejnému objektu, mluví se o koreferenci nebo anafoře (BOJAR, 2012, str. 10). Situaci, kdy překladač nezohlednil rod slova, na které se odkazuje, ilustruje následující příklad:

- (2) *He saw a pen. It was red.*
GT: *Viděl pero. Byla červená.*
BT: *Uviděl pero. Byla rudá.*

Mezi další jevy způsobující SP potíže patří například idiomy. Bojar (2012, str. 9) však upozorňuje, že pokud má systém seznam či ukázky idiomů k dispozici, je schopen je použít. Bohužel už je většinou nedokáže upravit tak, aby do věty správně pasovaly. Příkladová věta ukazuje situaci, kdy překladač nerozpoznal, že se jedná o idiom, neboť ho s největší pravděpodobností k dispozici neměl:

- (3) *Ich kenne die Gegend wie meine Westentasche.*
GT: *Znám tu oblast, jako je záda mé ruky.*
BT: *Znám tu oblast jako kapsu na vestu.*

Strojové překladače občas dělají chyby při překladu negace. To může mít více důvodů. Bojar (2012, str. 12) vysvětluje ztrátu negace ve větě číslo čtyři tím, že použitý česko-anglický systém neznal konkrétní tvar českého slovesa *nepotřebovat*, nenašel ho totiž ve svém slovníku, a proto hledal alespoň tvar základní. Ten však zní *potřebovat*, a tak má celá věta po překladu zcela opačný smysl:

- (4) *Pan Klaus řekl, že Evropa nepotřebuje ústavu.*
Mr Klaus said that Europe needs a constitution.

Absence negace v některých výstupech ze SP může souviset i s českou zdvojenou negací:

- (5) *I have no dog. = Nemám žádného psa.*

Systém se z trénovacích dat na větách se zdvojenou negací naučí, že anglické spojení *I have* odpovídá českému *Nemám*.

Co se týče syntaxe, je nutné dát si pozor na to, že některé strojově přeložené výstupy nevyhovují požadavkům aktuálního členění větného (SVOBODA, 2012, str. 88):

- (6) *Diese Abmachungen wurden nach der Gründung der Tschechoslowakischen Republik nicht eingehalten.*
Tyto smlouvy nebyly splněny po založení Československé republiky.

Rovněž u číselných údajů je potřeba být ve střehu. Svoboda (2012, str. 88) uvádí, že v této oblasti se chyby vyskytují poměrně často, proto je nutné provést v takovýchto případech následnou manuální opravu:

- (7) *7.0A @ 230Vac 15.9 @ 120Vac Geräuschpegel:*
7.0A @ 230 @ 120 VAC 15/09 hluku:

Jak je možné vidět v příkladu osm, také překlad německých funkčních verbálních spojení (Funktionsverbgefüge) činí strojovým překladačům problémy (SVOBODA, 2012, str. 89):

- (8) *Diese Vorstellung brachte Trtfk auch in einem 1969 in deutscher Sprache veröffentlichten Aufsatz zum Ausdruck:*
Tento nápad Trtfk také přinesl v 1969 eseje publikované v německém jazyce k vyjádření:

Svoboda (2012, str. 89n) uvádí ještě další problematické jevy, které se často objevují ve výstupech ze strojových překladačů. Zmiňuje například neharmonizovanost terminologie, kterou však lze vyřešit pomocí nástroje CAT, kde je obvykle možné termíny spravovat. Dále se ve výstupu mohou místy vyskytovat výpustky a občas nedochází k zohlednění gramatického čísla. Jak již bylo zmíněno v kapitole 2.2, některé systémy používají při překladu tzv. pivotní jazyk, kterým je většinou angličtina. Při překladu z němčiny do češtiny tedy proběhne nejdříve překlad z němčiny do angličtiny, a pak teprve z angličtiny do češtiny. Pod vlivem angličtiny se mohou v překladu objevovat různé chyby:

- (9) *Die Anlagen können für Heizungs-und Prozesswärme sowie für die Stromerzeugung (Dampf- und ORC-Prozess) eingesetzt werden.*

Rostliny lze použít k vytápění a tepelné zpracování, stejně jako pro výrobu energie (pára a ORC procesu).

V příkladu devět došlo k chybnému překladu anglického homonymního slova plant (rostlina, nebo podnik/závod/provoz) do češtiny.

3 Metodologický základ

3.1 Výchozí hypotézy

V této kapitole budou představeny výchozí hypotézy, k jejichž ověření, či vyvrácení dojde v empirické části práce.

Za prvé se budeme zabývat otázkou, který z vybraných strojových překladačů produkuje kvalitnější výstupy. Vzhledem k velké oblibě, rozšíření a vyspělé technické podpoře překladače Google Translate lze předpokládat, že co do kvality překladu dosáhne tento překladač při hodnocení lepších výsledků než překladač Bing Translator. V oddílu 2.2.2 pracujeme také s hypotézou, že u Google Translate byl NMT pro český jazyk spuštěn v dubnu 2017. V této souvislosti se hovoří dokonce o „skokovém zlepšení překladu“²¹. Tato skutečnost by se tedy měla projevit při hodnocení posledních dvou sad textů přeložených překladačem Google Translate, konkrétně u výstupů z 20. dubna a 13. července 2017.

Druhá hypotéza se týká jednotlivých textových typů. Co se týče kvality, vycházíme z předpokladu, že lepší kvalitu by měly vykazovat překlady informativních textů než překlady apelativních nebo uměleckých textů. Tato hypotéza se opírá o skutečnost, že korpusy veřejně dostupných překladačů obsahují velké množství textů mezinárodních organizací, jako je EU nebo OSN, což znamená, že jde především o informativní texty administrativního rázu. Avšak s nástupem pokročilejších technologií a přístupů, jako je NMT, je potřeba si uvědomit, že tato skutečnost již nemusí být relevantní.

Za třetí se budeme zabývat historickou podmíněností výstupů ze SP. Zaměříme-li se na časový aspekt experimentu, je možné se domnívat, že kvalita překladů se bude postupem času zvyšovat, neboť po technické stránce neustále dochází ke zlepšování systémů. Dále je potřeba zmínit, že také sami uživatelé mají možnost se podílet na zvyšování kvality tím, že mohou překlady vyhotovené pomocí veřejně dostupných překladačů hodnotit nebo upravovat. Svou roli by samozřejmě mohly hrát i korpusy, které jsou neustále obohacovány novými daty, na druhou stranu je ale nutné mít na paměti, že korpusy mohou obsahovat i velmi nekvalitní lidské nebo strojové překlady. Jak bylo zmíněno v oddílu 2.1.3.2 v této souvislosti se často hovoří

²¹ https://technet.idnes.cz/google-translate-neural-networks-dlq-/sw_internet.aspx?c=A170418_224649_sw_internet_pka [cit. 2017-12-11]

o tzv. bodu degradace. Obohacování korpusů tedy nemusí být zárukou zvyšování kvality překladů.

3.2 Model pro hodnocení

Metoda analýzy přeložených textů má dva stupně. Hodnocení prvního stupně probíhá pomocí lingvisticky orientované translatologické analýzy podle modelu Kathariny Reißové (viz oddíl 3.2.1). Druhým stupněm je pak posouzení srozumitelnosti a použitelnosti strojového překladu na stupnici o sedmi úrovních, kterou vytvořil A. Torrens a která bude blíže popsána v oddílu 3.2.2. Tento dvoustupňový model evaluace již byl použit v podobném experimentu²², který se rovněž zabýval hodnocením kvality překladů vyhotovených strojovými překladači. V rámci translatologické analýzy dojde ke srovnání výchozího a cílového textu, bude stanoven typ textu, dále budou hodnoceny sémantické, lexikální, gramatické a stylistické kategorie, přihlédnuto bude též k relevantním pragmatickým kategoriím. Analýza bude čistě evaluativního rázu, texty nebudou dále posteditovány, posteditační úsilí tedy nebude měřeno.

Druhá část hodnocení se zabývá celkovou použitelností přeloženého textu. Pro tyto účely je použita stupnice dle Antonia Torrense (1994, str. 394), která má pro hodnocení strojového překladu k dispozici sedm úrovní od 0 do 6 s odpovídajícím procentuálním hodnocením (0–120 %).

Pokud bychom měli tento model pro hodnocení kvality výstupů ze strojových překladačů klasifikovat podle kritérií Reginy Bauerové (2002), které byly popsány v oddílu 2.3.2, jednalo by se o *black box* hodnocení, tedy stav, kdy hodnotitel nemá vzhled do toho, jak systémy fungují. Přesto však lze díky znalosti základních principů fungování systémů původ některých chyb odhalit. Metoda je kombinací mikroevaluace a makroevaluace, neboť se zaměřuje jak na konkrétní problematické kategorie, tak na hodnocení výkonů jako celku. V rámci experimentu je provedena formální evaluace, která zkoumá kvalitu, a z části i funkční evaluace zaměřená na funkční hledisko systému (podpora jazykového páru, maximální objem vloženého textu atd.). Dále bychom mohli model přiřadit do kategorie standardní výkon,

²² Jedná se o experiment, který probíhal v rámci semináře *Translatologická analýza* na Ústavu translatologie Filozofické fakulty Univerzity Karlovy v letním semestru akademického roku 2013/2014. Podrobně je popsán v článku *Hodnocení kvality strojového překladu s případovou studií k překladu ve směru DE–CS při využití veřejně dostupných překladačů* (SVOBODA, 2015, str. 261nn).

protože uživatelské zásahy do vybraných online překladačů nejsou možné. Hodnocení bylo modifikováno individuálně, jde tedy o *home-made* model.

3.2.1 První stupeň: Translatologická analýza

Paralelou k literární kritice je kritika překladu. Kritika překladu by měla být vždy objektivní, aby nedocházelo k paušálním soudům a závěrům typu: „povedený překlad“, „nečitelný překlad“ atd. Aby bylo možné této objektivitě dosáhnout, je třeba stanovit si kritéria, podle kterých se bude překlad hodnotit. Za tímto účelem vznikají různé modely určené pro translatologickou analýzu. Každá dobrá kritika překladu by tedy měla pracovat s metodou translatologické analýzy. V empirické části této diplomové práce je aplikován model Kathariny Reißové, tato kapitola proto nastiňuje základní principy tohoto modelu.

Teoretický model translatologické analýzy od Kathariny Reißové byl poprvé představen v publikaci *Möglichkeiten und Grenzen der Übersetzungskritik: Kategorie und Kriterien für eine sachgerechte Beurteilung von Übersetzungen* z roku 1971. Publikace doposud nebyla přeložena do češtiny, užívaná terminologie v českém prostředí je proto často rozkolísaná.

Reißová hovoří o kritice překladu (*Übersetzungskritik*), kterou není možné provést bez srovnání s originálem (REIß, 1986, str. 11). Model byl vytvořen proto, aby při kritice překladů nedocházelo k subjektivnímu hodnocení a paušálním soudům. Každá kritika, ať už pozitivní, či negativní, by měla být vždy dobře zdůvodněna a doložena příklady. Nejde však pouze o to, upozornit na nevhodná řešení, je dobré vyzdvihnout i ta zdařilá. Co se týče negativní kritiky, měla by být konstruktivní, a proto je nutné navrhnout lepší řešení (REIß, 1986, str. 12).

Model translatologické analýzy dle Kathariny Reißové se skládá ze tří kategorií. Nejprve jen nutné určit typ textu, aby bylo možné zjistit, zda byla zvolena vhodná metoda překladu. Reißová pojímá textový typ jako literární kategorii překladatelské kritiky (1986, str. 52). Dalším krokem je posouzení jazykových kategorií. Nakonec je nutné se zabývat pragmatickými kategoriemi, tedy hodnocením mimojazykových determinantů.

3.2.1.1 Literární kategorie (*Die literarischen Kategorien*)

Jak již bylo naznačeno, první kategorie se zabývá určením textového typu. Stejně jako překladatel před zahájením své práce na překladu určí pomocí textové analýzy textový typ originálu, musí si i kritik vyjasnit, jaký typ textu se mu dostal do ruky, aby mohl překlad

posuzovat podle odpovídajících měřítek. Právě typ a druh²³ textu jsou totiž podle Reißové základem pro stanovení kritérií, podle kterých má být volena metoda překladu a překlad hodnocen:

„Zatímco typ textu [*Texttyp*] zásadně ovlivňuje metodu překladu a pořadí toho, co má zůstat zachováno v cílovém textu, určuje druh textu [*Textart*], které vnitrojazykové instrukce je při překladu potřeba přednostně dodržet.“ (REIß, 1986, str. 34), překlad S.K.

Co se týče diferenciacie textových typů pro účely translatické analýzy, došlo v této oblasti k mnoha pokusům o kategorizaci textů na základě specifických znaků²⁴, dle Reißové (1986, str. 26) však těmto rozdělením často chybí jednotná koncepce nebo se hodí spíše pro lingvistické účely, proto se rozhodla, že texty rozdělí podle toho, jakou funkci v jednotlivých textech plní jazyk.

Při své klasifikaci vychází Reißová z funkcí jazyka podle Karla Bühlera²⁵ (1934, str. 28), na jejichž základě definuje tři textové typy: texty zaměřené na obsah (*inhaltsbetonte Texte*), které odpovídají zobrazovací funkci jazyka a logické dimenzi, texty zprostředkovávající formu (*formbetonte Texte*), ve kterých je dominantní výrazová funkce jazyka a estetická dimenze a *appellbetonte Texte*, tedy texty s apelativní funkcí spadající do dialogické dimenze jazyka (REIß, 1986, str. 32nn).

Kromě těchto tří typů vymezuje Reißová ještě čtvrtý textový typ, do kterého spadají audio-mediální texty (REIß, 1986, str. 34). Jedná se o psané texty, se kterými je však příjemce konfrontován pomocí nejazykového média, tedy mluvenou, případně zpívanou formou. Na podnět Bernda Spillnera (1980, str. 75) změnila Reißová označení *audio-medial* na *multi-medial*, což souvisí s enormním rozvojem nových technologií a využíváním dalších kanálů.

K textům zaměřeným na obsah jsou řazeny například druhy textů, jako je novinová zpráva, komentář, reportáž, návod k použití, obchodní korespondence, učebnice a odborné publikace z oblasti přírodních věd, techniky a humanitních věd. Tyto texty mají za úkol hlavně informovat a poučit. Při překladu informativních textů je důležité především zachování věcné správnosti, invariantem je tedy obsah, kritik se musí přesvědčit o tom, zda byly informace a obsah správně předány (REIß, 1986, str. 34nn).

²³ Reißová (1986, str. 34) rozlišuje typ textu (*Texttyp*) a druh textu (*Textart*). Pokud jako reprezentanta první skupiny uvedeme informativní texty, je možné do druhé kategorie zařadit například komentář nebo reportáž, tedy jednotlivé texty, které spadají pod tento textový typ.

²⁴ Viz například Otto Kade (1964), Georges Mounin (1967) nebo Rudolf Walter Jumpelt (1961).

²⁵ Z tohoto modelu jazykových funkcí vychází ve své teorii i Peter Newmark, který se domnívá, že Bühlerova klasifikace rozšířená Romanem Jakobsonem se dá výhodně použít pro účely překladu (NEWMARK, 1988, str. 39).

Pod texty zaměřené na formu spadají texty, ve kterých autorem použité formální elementy vytváří specifické estetické působení na čtenáře (REIß, 1986, str. 38). Do této kategorie tedy řadíme různé žánry s estetickou ambicí, jak poezii, tak prozaické texty, jako je například fejeton, esej, novela, román, povídka nebo anekdota. Nejdůležitějším kritériem při překladu tohoto typu textů je ekvivalence estetického působení (REIß, 1986, str. 42).

Hlavním úkolem apelativních textů není jen předání informace, ale dosažení určitého cíle, úmyslu, mimojazykového efektu (REIß, 1986, str. 44). Účelem těchto textů je pohnout nebo přimět čtenáře či posluchače k tomu, aby jednal, či nejednal určitým způsobem. Při překladu je proto nutné vytvořit odpovídající apel, který u čtenáře překladu vyvolá stejnou reakci jako u čtenáře originálu. K tomuto textovému typu patří reklama, propaganda, polemika nebo satira (REIß, 1986, str. 45).

Audio-mediální nebo multimediální texty jsou realizovány prostřednictvím mimojazykového média v akustické, grafické nebo optické podobě (REIß, 1986, str. 49). Jedná se hlavně o texty zprostředkované pomocí rádia, televize nebo informačních technologií. Při překladu je důležité zohlednit mimojazykové elementy a zvolit takovou metodu překladu, která umožní vytvoření stejného účinku na příjemce cílového textu jako na příjemce výchozího textu.

Reiřová si je vědoma toho, že v textu není vždy zastoupena pouze jedna funkce, jednotlivé funkce se často kříží a překrývají (ZEHNALOVÁ et al., 2015, str. 63), při translatologické analýze je však nutné rozeznat tu dominantní.

3.2.1.2 Jazykové kategorie (*Die innensprachlichen Instruktionen*)

Poté, co kritik určí, o jaký typ textu se jedná, může přejít k další skupině kritérií, a to k jazykovým kategoriím. Jedná se o vnitrojazykové kategorie, v jejichž rámci se posuzují sémantické (*semantische*), lexikální (*lexikalische*), gramatické (*grammatische*) a stylistické instrukce (*stilistische Instruktionen*).

V závislosti na sémantických instrukcích se zkoumá, jak se překladatel vyrovnal s polysémií a homonymií, případně zda nedošlo k výpustkám, k mylné interpretaci výrazů užitých v originále, k doplněním oproti původnímu textu nebo jestli se v textu neobjevují nekryjící se významové jednotky či posuny. Zohlednění sémantických instrukcí je rozhodující pro zachování obsahu. Pro posouzení sémantické ekvivalence je nutné zaměřit se na kontext,

protože jen díky němu můžeme poznat, co chtěl autor originálu skutečně vyjádřit (REIß, 1986, str. 58).

Co se týče lexikálních instrukcí, je hlavním hodnoticím kritériem adekvátnost (REIß, 1986, str. 61). Kritik tedy má při hodnocení na lexikální rovině za úkol posoudit, jestli byly instrukce obsažené v originálu adekvátním způsobem přeneseny do cílového jazyka. Měl by se zaměřit především na odborné termíny, idiomy, jména, nepřeložitelná slova, falešné přátele, metafory, přísloví a slovní hříčky.

Hlavním hodnoticím kritériem u gramatických instrukcí je správnost (REIß, 1986, str. 63n). V rámci kritiky je posuzováno, zda jsou využity morfologicky a syntakticky relevantní struktury pro cílový jazyk. Důležité je, aby překlad odpovídal úzu cílového jazyka a nedocházelo k odchýlkám od správných gramatických tvarů.

Posledním prvkem, který se v rámci této kategorie posuzuje, je stylistika. Kritik musí prozkoumat, zda cílový text vykazuje korespondenci na úrovni stylistiky (REIß, 1986, str. 66). Je hodnoceno, zda překladatel zohlednil rozdíly mezi hovorovými a spisovnými výrazy a jestli dodržel úzus daného textového typu. Pokud došlo ve výchozím textu k vědomému mísení stylů, pak by se měl překladatel snažit docílit v překladu téhož.

3.2.1.3 Mimoslovníkové determinanty (*Die außersprachlichen Determinanten*)

Poslední kategorií tohoto modelu jsou mimoslovníkové determinanty, v jejímž rámci jsou hodnoceny pragmatické kategorie daného textu. Jedná se o mimoslovníkové skutečnosti, jako je například místo, čas, autor, příjemce atd., které jsou nějakým způsobem zohledněny v textu výchozí kultury. Dalo by se tedy říci, že se jedná o skutečnosti, pro které Nida (1964, str. 243) užívá pojmy *communicative context* a *cultural context of the source language* a které by se do jisté míry daly identifikovat s *národní a dobovou specifičností* podle Levého (1998, str. 119nn). Reiřová (1986, str. 71nn) rozlišuje na této rovině následujících sedm kategorií: užší situační kontext (*engerer Situationsbezug*), věcný faktor (*Sachbezug*), faktor času (*Zeitbezug*), faktor místa (*Ortsbezug*), faktor příjemce (*Empfängerbezug*), závislost na mluvčím (*Sprecherabhängigkeit*) a afektivní implikace (*affektive Implikationen*).

Kategorie užšího situačního kontextu se nevztahuje na text jako celek, ale na jednotlivé pasáže a situace (REIß, 1986, str. 71n). Jedná se například o citoslovce nebo aluze na různá literární díla, dějinné nebo aktuální události. Tyto pasáže se často objevují u divadelních her

nebo v románových dialozích. Při jejich překladu je důležité správné pochopení konkrétní situace v rámci textu.

U překladatele je předpokládána nejen jazyková vybavenost, ale také znalost tématu, o kterém překládaný text pojednává. V rámci věcného faktoru je hodnocena překladatelova faktická obeznámenost s tématem. Faktor času se zabývá zohledněním doby vzniku originálu a překladu. Reißová (1986, str. 74) upozorňuje na skutečnost, že jazyk je neustále se měnící organismus, který je utvářen dobou, což se musí projevit i v překladu. S tím souvisí fenomén zastarávání překladu. Z tohoto důvodu musejí být starší texty světové literatury s časovým odstupem překládány znovu.

Co se týče faktoru místa, zkoumá kritik nakládání s reáliemi, vlastními jmény a specifiky výchozí země, kultury či jazykového společenství. To, jakým způsobem se s nimi překladatel vyrovná, by se mělo odvíjet od textového typu (REIß, 1986, str. 79). Dále by se měla kritika zabývat faktorem příjemce, přičemž zohledňuje recipienta výchozího textu. Předposlední kategorií je závislost na mluvčím, kdy je posuzován specifický styl autora, který je podmíněn jeho původem, vzděláním nebo například obdobím jeho tvorby. Nakonec by se měl kritik zabývat afektivními implikacemi, tedy tím, zda překladatel správně rozpoznal a interpretoval jevy, jako je humor, ironie, sarkasmus nebo pohrdání a jestli byly odpovídajícím způsobem převedeny do cílového jazyka (REIß, 1986, str. 86).

3.2.1.4 Kritika modelu Kathariny Reißové

Model translatické analýzy představený Katharinou Reißovou musel od doby svého vzniku mnohokrát čelit kritice. V této souvislosti je potřeba zmínit, že některé kritické výtky jsou oprávněné. Modelu je vyčítáno, že je zaměřen hodně lingvisticky, to však souvisí s dobou jeho vzniku. Dalším problémem je, že Reißová ve své publikaci zmiňuje pouze příjemce originálu a zcela opomíjí tzv. sekundární komunikaci, která zahrnuje i další účastníky komunikační situace, kromě příjemce překladu se jedná také například o redaktora, korektora nebo editora. Jak upozorňuje Popovič (1975, str. 69n), kritika by vliv dalších příjemců zohledňovat měla, zejména pak již zmíněného redaktora či editora:

„Na tvorbe prekladového textu sa okrem priamych účastníkov komunikácie istou mierou zúčastňuje aj externý činiteľ – redaktor, resp. upravovateľ definitívneho znenia textu prekladu. Ako ukazuje moderná prax, pri publikovaní prekladov, bez tohto činiteľa sa nedá celkom zaobísť. Redaktor je totiž sprostredkujúcim faktorom medzi

vládnucími jazykovými normami a textom prekladu a vykonáva direktívy, ktoré si vplyvom čitateľských návykov osvojila vydavateľská inštitúcia.“

Popovič (1975, str. 122n) ďalej poukazuje na skutočnosť, že Reißová vystihuje pouze zmeny na makrostylistické rovině textu (*Ort-, Sach-, Zeitbezug*). Pro účely literárního překladu je však nutné model doplnit také o mikrostylistickou rovinu, Popovič proto definuje následující posuny: *výrazové zosilňovanie, výrazová zhoda, výrazové zoslabovanie*.

Jako problematické se může jevit také označení první skupiny hodnotících kritérií jako literární kategorie. Tato kategorie zahrnuje stanovení jazykové funkce daného textu a určení textového typu. Její označení pomocí adjektiva literární může působit matoucím dojmem, protože nezahrnuje jen texty s literární ambicí, ale i další textové typy. Další bod, který by se dal modelu vytknout, je například absence kategorie, která by se zabývala typografií a celkovým grafickým zpracováním textu.

Přestože je model translatologické analýzy Kathariny Reißové kritizován, je zjevné, že ho lze s určitým doplněním aplikovat na široké spektrum textů od uměleckých až po odborné, což prokazuje jeho univerzálnost. Kritika ze strany současných autorů navíc potvrzuje skutečnost, že je tento model, který byl představen před více než čtyřiceti lety, stále diskutovanou problematikou a může být za předpokladu zohlednění některých problematických bodů uplatněn i v současnosti.

3.2.1.5 Vyhodnocování translatologické analýzy

Pomocí modelu Kathariny Reißové, který byl představen v předcházejících kapitolách, je hodnoceno šest textů patřících ke třem různým textovým typům. Tyto texty byly přeloženy veřejně dostupnými překladači v pěti časových obdobích. Proces sběru materiálu je popsán a jednotlivé texty jsou představeny v kapitole 3.3. Model je aplikován flexibilně, proto je pro každý textový typ přizpůsoben tak, že se při hodnocení zaměřuje právě na ty kategorie, které jsou pro daný textový typ relevantní. Výstupy budou hodnoceny v rámci jednotlivých textových typů.

Vyhodnocení výsledků prvního stupně proběhne na základě bodového ohodnocení jednotlivých jevů v rámci stanovených kategorií (sémantická, lexikální, gramatická, stylistická, pragmatická, zdařilá řešení). Kategorie mají různou váhu, chyby jsou tedy různě bodově ohodnoceny, a to negativně. Gramatická a lexikální kategorie má váhu jedna, chyby z těchto kategorií jsou ohodnoceny jedním negativním bodem. Sémantická kategorie má dvojnásobnou

váhu, významové chyby jsou hodnoceny dvěma zápornými body. Za stylistické prohřešky jsou systémy penalizovány půlbodem, za pragmatické bodem. Velmi zdařilá řešení mají váhu půl bodu a jsou hodnocena kladně. Každý přeložený text je ohodnocen bodovým součtem, jsou tedy sečteny body ze všech hodnocených kategorií. Čím nižšího negativního skóre překladač při hodnocení dosáhl, tím vyšší je kvalita jím produkovaných překladů.

Hodnocené jevy budou vyznačeny barevně v elektronických verzích dokumentů (sémantické jevy – zelená, lexikální jevy – modrá, gramatické jevy – červená, stylistické jevy – žlutá, pragmatické – oranžová, zdařilá řešení – signálně zelená). Ukázky takto barevně zpracovaných textů je možné nalézt v příloze v kapitole 8. Všechny originály, strojově vyhotovené překlady a jejich verze s grafickým hodnocením jsou dostupné v elektronické podobě na přiloženém CD.

Dále pro každý textový typ vybereme tři kategorie, které jsou pro daný typ relevantní a při překladu musí být zohledněny. Počty chyb z těchto kategorií budou pro lepší přehlednost zpracovány graficky. Jednotlivé grafy pak zachytí vývoj počtu chyb v rámci vybrané kategorie během celého zkoumaného období, tedy od června 2016 do července 2017. Na vodorovné ose grafu budou datумы, kdy byly jednotlivé překlady provedeny (30. 6. 2016, 2. 10. 2016, 11. 1. 2017, 20. 4. 2017 a 13. 7. 2017), svislá osa bude ukazovat počet chyb. Protože jsme formulovali hypotézu, že se bude kvalita překladů zlepšovat, lze v případě jejího potvrzení počítat s klesavou tendencí grafů.

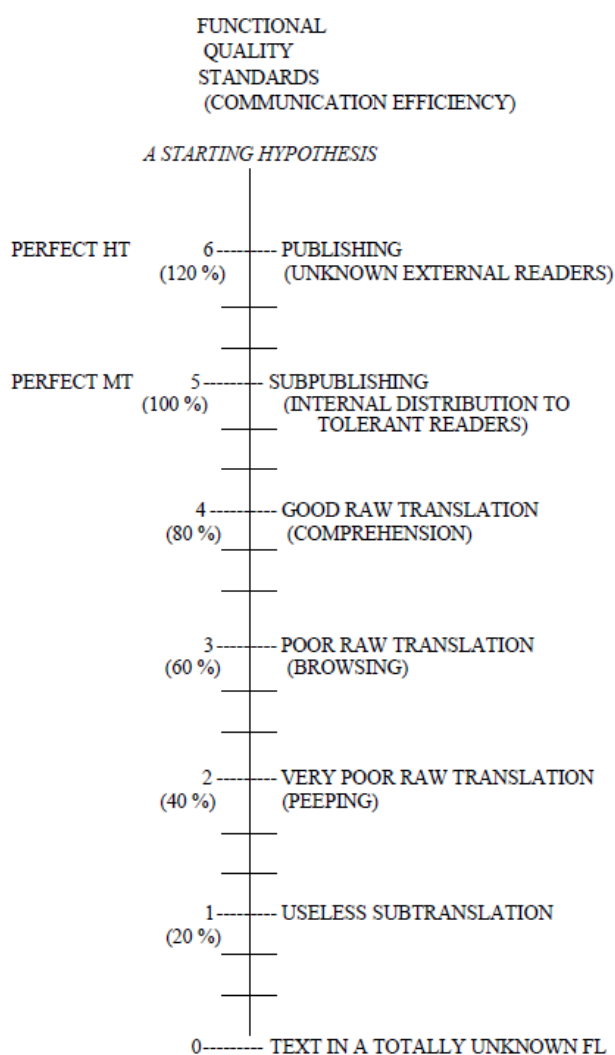
U apelativních textů se budeme nejdříve zabývat pragmatickou kategorií *afektivní implikace*, konkrétně bude hodnoceno, jak se překladačům povedl přeložit apel na čtenáře, dále se zaměříme na sémantické a gramatické chyby. Co se týče informativních textů, budeme zkoumat vývoj počtu sémantických chyb, dále se zaměříme na věcné chyby, tedy pragmatickou kategorii nazvanou *věcný faktor*, a poslední graf bude znázorňovat vývoj počtu gramatických prohřešků. V překladech uměleckých textů bude nejdříve sledován vývoj chyb stylistických, pak sémantických a nakonec chyb na rovině gramatiky.

3.2.2 Druhý stupeň: Posouzení na stupnici

Po translatologické analýze bude následovat druhý stupeň hodnocení, konkrétně hodnocení na stupnici vytvořené Antoniem Torrensem (1994, str. 394). Model translatologické analýzy podle Kathariny Reißové, který je užit v rámci prvního stupně hodnocení, pracuje

s detailním systémem různých hodnoticích kategorií, nehodnotí však celkovou použitelnost výstupu. Tato druhá fáze byla do hodnocení zařazena proto, aby tento nedostatek napравиła.

Jak ukazuje Obrázek 5, stupnice je rozdělená na sedm úrovní od nuly do šesti, přičemž nulou by byl ohodnocen text v neznámém cizím jazyce. Jeden bod, jemuž odpovídá 20 %, by obdržel nepoužitelný překlad, dva body a 40 % velmi špatný hrubý překlad a 3 body s 60 % špatný hrubý překlad. Čtyři body odpovídající 80 % by dostal dobrý hrubý překlad, který umožňuje porozumění. Text perfektní strojové kvality, který by byl publikovatelný interně pro tolerantní čtenáře, by byl ohodnocen pěti body a 100 %. Šest bodů (120 %) by získal text dosahující perfektní lidské kvality, jež by bylo možné publikovat pro neznámé externí čtenáře.



Obrázek 5: Stupnice podle A. Torrense (1994, str. 394)

V rámci druhého stupně hodnocení bude pomocí této stupnice každý text ohodnocen procentuálně jako celek. Hodnocení se zaměří především na použitelnost textu a dodržení funkce. V tomto případě bude platit, že čím vyšší je získaná procentuální hodnota, tím vyšší je

celková použitelnost a srozumitelnost překladu pro čtenáře a tím vyšší je i jeho kvalita. Vzhledem k tomu, že bude hodnocení provedeno intuitivně, nemusí se jeho výsledky nutně shodovat s výsledky translatologické analýzy. Na závěr dojde ke sloučení obou fází hodnocení.

3.3 Empirický materiál

V této kapitole je představen empirický materiál, který byl použit v rámci experimentu. Nejprve je stručně popsán výběr výchozích textů a uvedena jejich stručná charakteristika. Dále je představen proces a časový rámec sběru cílových textů a odůvodněna volba vybraných strojových překladačů.

3.3.1 Popis výchozích textů

Korpus vytvořený pro empirickou studii obsahuje šest textů. Texty byly zvoleny s ohledem na typologii, kterou pro svůj model translatologické analýzy vytvořila Katharina Reiřová. Ke každému textovému typu náleží dva texty z korpusu. Při výběru textů bylo rozhodující, aby se jednalo o originály vytvořené v němčině, a ne překlady z jiných jazyků. Pro účely této práce byly vybrány tyto texty:

Číslo textu	Název výchozího textu (VT)	Textový typ podle Reiřové	Počet slov / znaků VT	Počet NS ²⁶ VT
1.	Rede von Bundeskanzlerin Merkel im Rahmen der Plenarsitzung beim World Humanitarian Summit (Merkelová)	Appelbetonte Texte/ apelativní texty	329/2351	1,31
2.	Schuldenberatung (Poradentví)	Appelbetonte Texte/ apelativní texty	305/2379	1,32
3.	EU-Kommission verlängert Zulassung von Glyphosat (Glyfosát)	Inhaltsbetonte Texte/ informativní texty	289/2221	1,23
4.	Verordnung zur Durchführung eines Monitorings auf das Virus der Geflügelpest bei Wildvögeln (Monitoring)	Inhaltsbetonte Texte/ informativní texty	323/2287	1,27
5.	Allein unter Dieben (Schmeißer)	Formbetonte Texte/ umělecké texty	373/2422	1,35
6.	Caravaggio-Fieber. Nichts für unser Wohnzimmer, Chérie! (Caravaggio)	Formbetonte Texte/ umělecké texty	333/2267	1,26

Tabulka 1: Textový korpus sestavený pro studii

²⁶ 1 normostrana (NS) = 1800 znaků včetně mezer

a) Apelativní texty

Co se týče apelativních textů, jako první byl do korpusu vybrán projev kancléřky Merkelové na plenárním zasedání Světového humanitárního summitu z května 2016 získaný ze stránek bundesregierung.de. Druhým textem je reklamní leták týkající se dluhového poradenství, které nabízí *Betreuungsverein Bremen* e.V.

U obou textů je dominantní funkce apelativní, protože autor textu se snaží změnit postoje adresáta k dané skutečnosti nebo jej přimět k nějakému jednání – cílem je ovlivnění chování nebo myšlení adresáta textu. Za tímto účelem používá kancléřka ve svém projevu oslovení, řečnické otázky, modální slovesa muset či potřebovat a spřežky zaprvé, zadruhé, aby zdůraznila jednotlivé argumenty. Sdělování informací sleduje u apelativních textů jiný cíl než jen pouhé informování příjemce. Tyto textové typy nepřinášejí jen nové poznatky, jejich cílem je změnit názor a přesvědčení příjemce textu. Toho se snaží dosáhnout i reklamní leták k dluhovému poradenství, a to prostřednictvím krátkých úderných nadpisů, bodových výčtů argumentů a oslovování čtenáře pomocí vykání.

b) Informativní texty

Texty číslo tři a čtyři patří k informativnímu textovému typu. Jedná se o zprávu o prodloužení povolení pro využívání herbicidu glyfosátu ze stránky *Deutsche Welle* (dw.com) a nařízení k provedení monitorování viru ptačí chřipky u divokých ptáků, které vydalo německé Spolkové ministerstvo spravedlnosti a ochrany spotřebitele a bylo zveřejněno na internetových stránkách gesetze-im-internet.de.

Zpráva o prodloužení povolení glyfosátu má informativní charakter a obsahuje informace o aktuálním dění. Nalezneme v ní konkrétní údaje, které jsou úplné, věcně správné a doložené citáty od zainteresovaných osob. Fakta jsou pouze konstatována a podána objektivně, autor se zdržuje subjektivního hodnocení. Nařízení o monitoringu divokých ptáků bychom mohli přiřadit k administrativnímu stylu, tedy stylu informativní a věcné komunikace. Text je zaměřen především na fakta a na jejich sdělení. Autor je zastřen a užívá zcela neutrální spisovný jazyk, bez citového zabarvení. V textu se objevují zkratky a odkazy na různé zákony.

c) Umělecké texty

K uměleckým textovým typům patří texty číslo pět a šest z datového korpusu. V prvním případě se jedná o úryvek z dětské knížky o rodině zlodějů s názvem *Allein unter Dieben* od spisovatele Franka Schmeißera, ve druhém o fejeton z online vydání z novin *Frankfurter Allgemeine* s názvem *Caravaggio-Fieber* o nález Caravaggiova obrazu při opravách na půdě.

Úryvek z dětské knížky o zlodějích nese znaky vyprávěcího slohového postupu. Obsahuje dějová slovesa, konkrétní podstatná jména, přirovnání pro snadnější představu popisované situace a přímou řeč, která zvyšuje živost a působivost vypravování. Fejeton o Caravaggiově obraze je méně rozsáhlý literárně publicistický útvar se subjektivním nádechem a lehkou stylistickou formou, který se zaměřuje na aktuální kulturní téma, tedy nález Caravaggiova obrazu na půdě. Aby vzbudil zájem u čtenáře, nese vtipný název a střídá vrstvy jazyka.

3.3.2 Sběr materiálu

Sběr empirického materiálu určeného k analýze probíhal dvanáct měsíců. Šest předem zvolených textů v německém jazyce bylo dohromady pětkrát s přibližně tříměsíčním časovým odstupem vloženo do dvou vybraných strojových překladačů a přeloženo do češtiny. První překlad proběhl 30. června 2016, další 2. října 2016, třetí 11. ledna 2017, předposlední překlad byl vyhotoven 20. dubna 2017 a poslední 13. července 2017. Studie je provedena v rámci jazykové dvojice němčina-čeština, což je u podobných experimentů méně častá kombinace.

Překlad textů provedly dva veřejně dostupné online překladače. Jedná se o již zmíněné překladače Google Translate a Bing Translator. Vybrané systémy musely splňovat několik kritérií. Bylo nutné, aby se jednalo o online překladače, které jsou zdarma a běžně přístupné široké veřejnosti. Dále bylo potřeba, aby systémy podporovaly jazykovou dvojici němčina-čeština a poskytovaly plnou kvalitativní funkčnost. Tato kritéria splňuje kromě zvolených překladačů také například systém Babylon. Pro účely tohoto experimentu však byly vybrány pouze dva výše uvedené překladače, neboť se jedná o nejrozšířenější a nejužívanější online překladače.

Všech šest textů bylo do vybraných strojových překladačů dohromady vloženo pětkrát. Každý online překladač tedy během celého zkoumaného období vyhotovil 30 překladů. Pomocí představného modelu pro evaluaci kvality výstupů z veřejně dostupných online překladačů bylo v rámci empirické studie dohromady ohodnoceno 60 textů.

4 Výsledky empirického výzkumu

V této kapitole budou představeny dílčí výsledky empirické studie, která se zabývala srovnáním a historickou podmíněností výstupů z veřejně dostupných strojových překladačů. Veřejně dostupné strojové překladače Bing Translator (BT) a Google Translate (GT) produkovaly v pěti časových obdobích po dobu dvanácti měsíců překlady šesti textů náležících ke třem různým textovým typům. V rámci empirické studie bylo zkoumáno, jak se vybrané překladače vypořádaly s různými textovými typy a jestli dochází ke zlepšování kvality výstupů ze SP v čase.

Jak již bylo naznačeno v metodické části práce, překlady byly vyhodnocovány v rámci jednotlivých textových typů. Hodnocení proběhlo dvoustupňově, nejdříve byla provedena translatologická analýza zaměřená na sémantické, lexikální, gramatické, stylistické a pragmatické kategorie, pak došlo k procentuálnímu zhodnocení celého překladu za účelem zjištění celkové použitelnosti, srozumitelnosti a dodržení funkce.

Výsledkům hodnocení každého textového typu je věnována jedna podkapitola. V každé podkapitole jsou nejdříve představeny výsledky získané na základě translatologické analýzy. Dále byly pro každý textový typ vybrány tři kategorie, které jsou pro něj obzvláště relevantní. Počty chyb jsou pro každou kategorii zpracovány graficky, aby bylo možné zachytit jejich vývoj během celé studie. Na konci každého oddílu jsou prezentovány výsledky druhé fáze hodnocení.

Celkové shrnutí výsledků empirické studie podává kapitola 5, ve které rovněž došlo ke zhodnocení toho, zda se výzkumné hypotézy formulované v kapitole 3.1 podařilo potvrdit, nebo vyvrátit.

4.1 Apelativní texty

4.1.1 Hodnocení překladů na základě translatologické analýzy

Projev kancléřky Merkelové a reklamní materiál k dluhovému poradenství byly vybrány jako zástupci apelativních textů. Tabulka 2 srovnává bodové hodnocení získané v rámci prvního stupně hodnocení strojového překladu projevu kancléřky pomocí translatologické analýzy. Překladač Google Translate si při překladu po celé zkoumané období vedl lépe, z původních 133,5 negativních bodů se na konci studie dostal na 122,5 bodů, došlo tedy ke snížení chybovosti o 11 bodů.

Překladač Bing Translator při překladu apelativního projevu nedocílil tak dobrých výsledků. Na začátku zkoumaného období získal 152,5 negativních bodů, během následujících třech překladů došlo ke zvýšení o 8,5 bodu a chybovost se ustálila lehce nad hranicí 160 bodů. Na závěr studie se chybovost sice mírně snížila, přesto zůstal počet záporných bodů vyšší než na začátku. Při překladu apelativního projevu kancléřky překladačem Bing Translator tedy došlo k celkovému zvýšení chybovosti o 5 negativních bodů.

Merkelová		
Datum překladu	Bing Translator	Google Translate
30. 6. 2016	152,5	133,5
2. 10. 2016	161	133,5
11. 1. 2017	161	133,5
20. 4. 2017	163	122,5
13. 7. 2017	157,5	122,5
Zlepšení o:	-5	11
Celkový průměr (ø):	159	129,1

Tabulka 2: Text Merkelová: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)

Tabulka 3 ukazuje bodové hodnocení, které překladače obdržely v rámci hodnocení překladu reklamního materiálu týkajícího se dluhového poradenství. Při hodnocení prvního překladu bylo zjištěno, že si BT vedl hůře než GT. Získal totiž 202,5 bodů, tedy o celých 14 negativních bodů více než jeho konkurent. Během zkoumaného období však došlo ke značnému snížení chybovosti, takže poslední překlad vyhotovený překladačem Bing Translator obdržel 165 záporných bodů. Porovnáme-li výsledek prvního a posledního hodnocení translatu, zjistíme, že došlo ke snížení dokonce o 37,5 bodů.

Reklamní text přeložený překladačem Google Translate sice na začátku obsahoval méně chyb než překlad vyhotovený BT a při analýze obdržel 188,5 záporných bodů, v průběhu studie

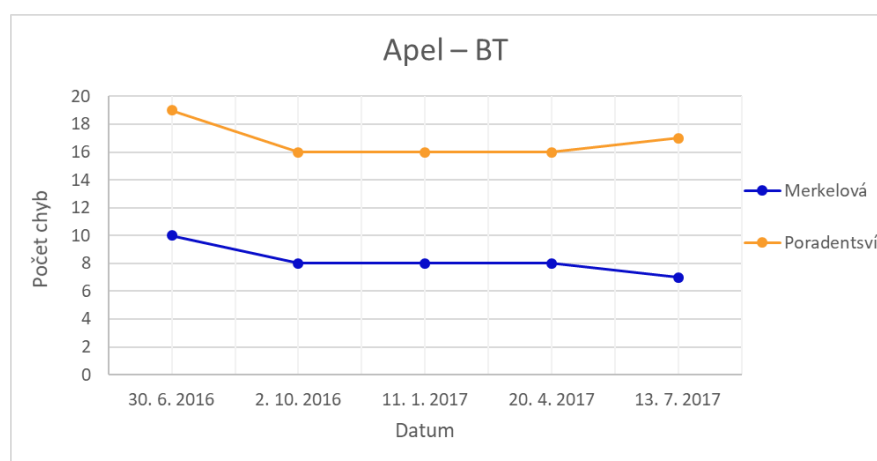
však nedošlo k tak výraznému zlepšení jako u BT. Překladač GT dosáhl při čtvrtém a pátém hodnocení výsledku 185, což znamená, že chybovost se během trvání experimentu snížila pouze o 3,5 bodu. Porovnáme-li poslední výsledek obou překladačů, zjistíme, že co se týče chybovosti, vykazuje BT o 20 záporných bodů lepší výkon.

Poradenství		
Datum překladu	Bing Translator	Google Translate
30. 6. 2016	202,5	188,5
2. 10. 2016	184,5	188,5
11. 1. 2017	184,5	188,5
20. 4. 2017	184,5	185
13. 7. 2017	165	185
Zlepšení o:	37,5	3,5
Celkový průměr (ø):	184,2	187,1

Tabulka 3: Text Poradenství: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)

4.1.2 Vývoj počtu chyb ve třech vybraných kategoriích

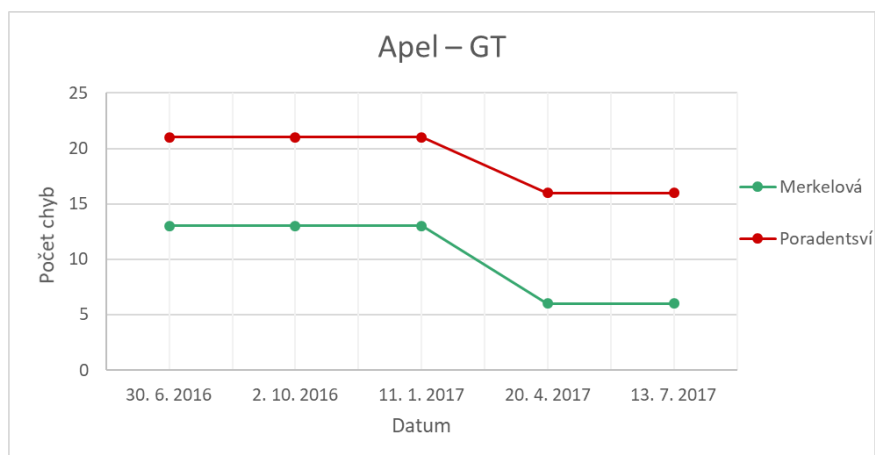
Pro kategorii apelativních textů byly vybrány tyto tři kategorie: afektivní implikace, v rámci této kategorie budeme zkoumat vývoj počtu chyb při vyjadřování apelu na čtenáře, dále se zaměříme na chybovost v oblasti sémantiky během celého zkoumaného období a nakonec se budeme zabývat vývojem počtu chyb na rovině gramatické.



Graf 1: BT (apelativní texty): Počet chyb v rámci celého zkoumaného období (apel)

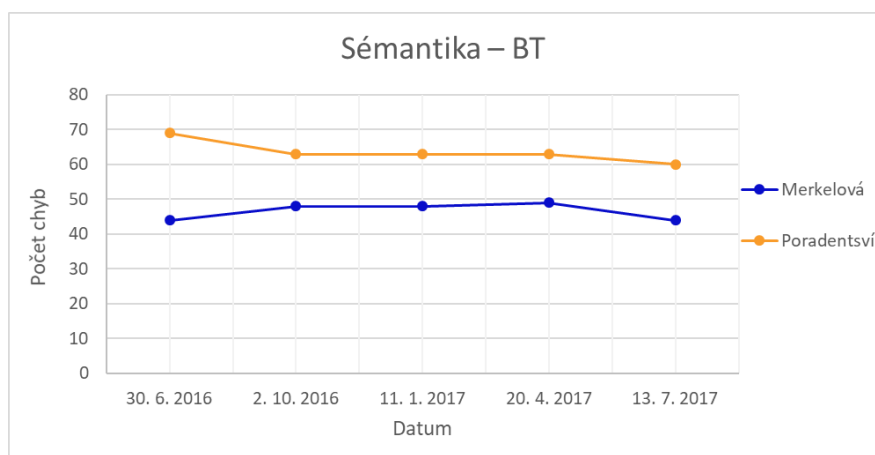
Graf 1 ukazuje, jak se překladač Bing Translator vyrovnal s vyjádřením apelu na čtenáře. V projevu kancléřky Merkelové se v této oblasti objevilo méně chyb než v reklamním textu, navíc zde došlo k mírnému zlepšení, z původních deseti chyb klesla chybovost na konci zkoumaného období o tři na sedm. Převod apelu u propagačního textu k dluhovému poradenství

činil překladači od Microsoftu větší obtíže. První překlad obsahoval dokonce 19 chyb, v dalších třech zkoumaných obdobích klesl počet chyb na 16. V posledním zkoumaném období již k dalšímu poklesu nedošlo, naopak se počet chyb zvýšil o jednu na 17.



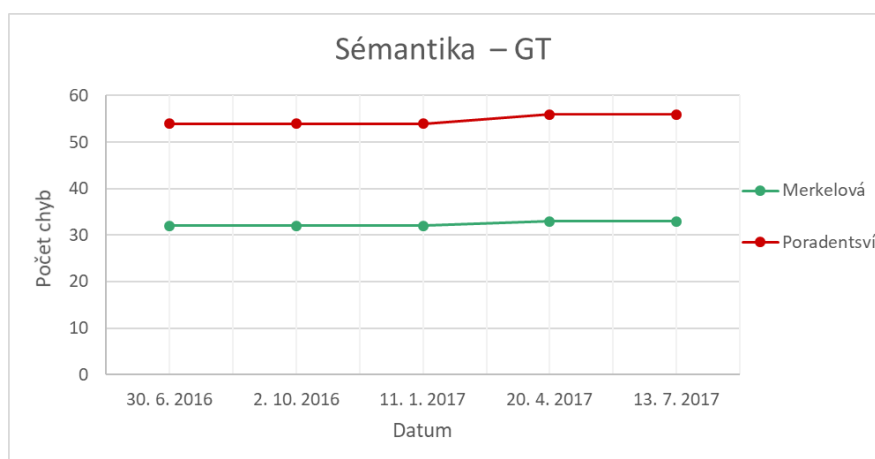
Graf 2: GT (apelativní texty): Počet chyb v rámci celého zkoumaného období (apel)

Při hodnocení apelativních textů, které vyhotovil Google Translate (Graf 2), bylo zjištěno, že v pragmatické kategorii produkoval překladač při převodu apelu více chyb než překladač BT, avšak v průběhu studie došlo u obou textů ke snížení chybovosti a GT tak dosáhl při posledním hodnocení v této kategorii textů lepšího výsledku než Bing Translator, a to u obou textů. Chybné vyjádření apelu překladačem GT v posledních dvou hodnoceních bylo zaznamenáno 16krát u reklamního textu a šestkrát u projevu. V prvním případě tedy došlo ke snížení počtu chyb o pět a ve druhém dokonce o sedm.



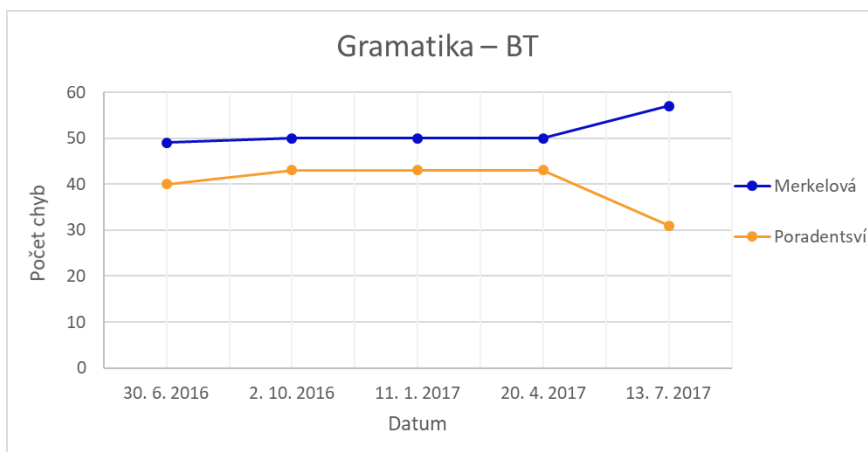
Graf 3: BT (apelativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)

Graf 3 znázorňuje vývoj počtu sémantických chyb u výstupů z překladače Bing Translator. Chybovost v reklamním textu o poradenství vykazuje mírnou klesavou tendenci, z původních 69 chyb při prvním hodnocení došlo k poklesu o devět na 60 u posledního překladu. Na konci zkoumaného období se totiž systém dopouštěl méně výpustek než při prvním překladu. V překladu z června 2016 bylo zjištěno 28 výpustek, v posledním překladu již jen 20. Překlady projevu kancléřky obsahovaly méně chyb na sémantické rovině než převody reklamního textu. Při prvním překladu projevu se BT dopustil 44 sémantických chyb, v dalších třech hodnocených obdobích došlo k mírnému zvýšení chyb, které byly způsobeny více výpustkami slov. V posledním překladu bylo zjištěno snížení počtu výpustek, a to dokonce o deset, a zvýšení počtu posunů a mylných interpretací slov, takže poslední výstup obsahoval celkem 44 sémantických chyb. Na konci zkoumaného období tedy BT docílil v oblasti sémantiky stejného počtu chyb jako na začátku.



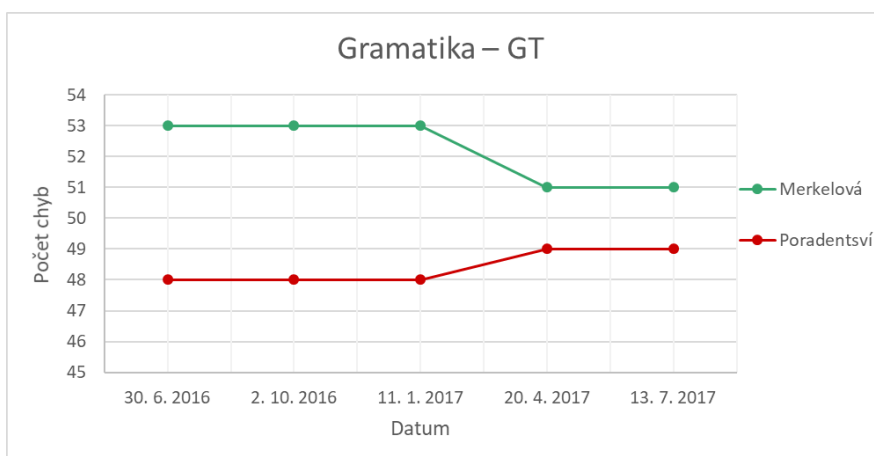
Graf 4: GT (apelativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)

Jak je možné vidět díky grafickému znázornění (Graf 4), překladač Google Translate produkoval při překladu sémantických jevů u obou textů méně chyb než překladač Bing Translator. Graf 4 však ukazuje, že chybovost v oblasti sémantiky zůstává u překladů projevu Merkelové téměř konstantní a u textu týkajícího se finančního poradenství se dokonce mírně zvyšuje. Při překladu politického projevu dosáhl GT na rovině sémantiky nejnižšího počtu chyb a tím svého nejlepšího výkonu v kategorii sémantiky v rámci všech textů, které přeložil.



Graf 5: BT (apelativní texty): Počet chyb v rámci celého zkoumaného období (gramatika)

Při překladu projevu kancléřky produkoval Bing Translator při prvních čtyřech překladech podobné množství chyb, jejich počet se pohyboval přibližně okolo 50, při posledním hodnocení nastalo zvýšení počtu gramatických chyb o osm, v posledním hodnocení totiž došlo k mírnému zhoršení slovosledu a interpunkce. Jak ukazuje Graf 5, text o finančním poradenství obsahoval méně gramatických chyb než druhý apelativní text. Poslední překlad reklamního textu navíc obsahoval o 12 gramatických chyb méně než jeho předchozí překlad, došlo ke zlepšení především v oblasti deklinace a konjugace.



Graf 6: GT (apelativní texty): Počet chyb v rámci celého zkoumaného období (gramatika)

Projev Merkelové přeložený překladačem od Googlu vykazoval více gramatických chyb než propagační materiál k dluhovému poradenství, stejně jako tomu bylo u překladače Bing Translator. Graf 6 však dokládá, že vývoj počtu chyb probíhal u obou výstupů z GT zcela opačně než u překladů z BT. U politického projevu se počet gramatických chyb při čtvrtém překladu mírně snížil z 53 na 51 a v reklamním textu naopak mírně zvýšil ze 48 na 49.

4.1.3 Celkové zhodnocení překladu pomocí stupnice

Tabulka 4 ukazuje výsledky celkového procentuálního hodnocení překladu apelativních textů, kdy bylo prováděno hodnocení na stupnici (0–120 %) a zkoumáno, jaká je celková použitelnost textu a zda došlo k zachování funkce. Nejlepšího průměrného výsledku dosáhl překladač Google Translate při překladu projevu kancléřky Merkelové, jeho konkurent Bing Translator obdržel v rámci celkového průměru při překladu stejného textu o 12 procent méně, což je mezi překlady apelativních textů nejhorsí výkon. Druhé nejlepší průměrné procentuální skóre (55 %) získal při hodnocení apelativních textů opět Google Translate, a to za překlad textu týkajícího se finančního poradenství. Bing obdržel při hodnocení stejného textu průměrně o jedno procento méně.

Datum překladu	Merkelová		Poradenství	
	Bing Translator	Google Translate	Bing Translator	Google Translate
30. 6. 2016	55 %	60 %	50 %	55 %
2. 10. 2016	50 %	60 %	55 %	55 %
11. 1. 2017	50 %	60 %	55 %	55 %
20. 4. 2017	50 %	70 %	55 %	55 %
13. 7. 2017	55 %	70 %	55 %	55 %
Celkový průměr (̄):	52 %	64 %	54 %	55 %

Tabulka 4: Procentuální hodnocení překladů apelativních textů na stupnici podle Torrense (0–120 %) z hlediska jejich použitelnosti/funkce (čím vyšší procentuální hodnota, tím vyšší kvalita překladu)

4.2 Informativní texty

4.2.1 Hodnocení překladů na základě translatologické analýzy

Informativní texty reprezentují v datovém korpusu vytvořeném pro účely této empirické studie dva texty, článek o prodloužení povolení glyfosátu a nařízení o monitoringu divokého ptactva. Jak ilustruje Tabulka 5, za první překlad informativního článku o glyfosátu byl Bing Translator penalizován 170 zápornými body. Přestože se jeho překlady v průběhu studie co do počtu chyb výrazně zlepšily, a to dokonce o 33 bodů, zůstal překladač při posledním překladu 37 negativních bodů za překladačem od Googlu.

Google Translate produkoval od začátku empirického výzkumu informativní texty s méně chybami než jeho konkurent od Microsoftu. Za první překlad byl penalizován pouhými 116,5 body, při druhém hodnocení bylo zaznamenáno nepatrné zhoršení o jeden bod, při třetím

hodnocení bylo zjištěno, že text vykazuje naprostou shodu s prvním výstupem. Při dubnovém i červencovém hodnocení došlo ke zlepšení o 16,5 negativních bodů oproti prvnímu výstupu. Při posledních dvou hodnoceních překladu zprávy o glyfosátu získal GT pokaždé pouze 100 negativních bodů, což je nejlepší výkon celé empirické studie.

Glyfosát		
Datum překladu	Bing Translator	Google Translate
30. 6. 2016	170	116,5
2. 10. 2016	136,5	117,5
11. 1. 2017	139,5	116,5
20. 4. 2017	137,5	100
13. 7. 2017	137	100
Zlepšení o:	33	16,5
Celkový průměr (̅):	144,1	110,1

Tabulka 5: Text Glyfosát: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)

První překlad textu o monitoringu ptactva od překladače Bing Translator byl ohodnocen 167 negativními body (Tabulka 6). Během následujících třech hodnocení došlo ke zhoršení o 8,5 bodů, překladač tedy třikrát po sobě obdržel 175,5 negativních bodů. Při poslední analýze bylo zjištěno výrazné snížení chybovosti o 25 bodů oproti předchozímu výkonu. Zaměříme-li se na výsledek prvního a posledního hodnocení, zjistíme, že v rámci celého zkoumaného období chybovost výstupů BT klesla o 17 bodů. I přes toto poměrně výrazné zlepšení obdržel BT při poslední analýze o tři záporné body více než GT.

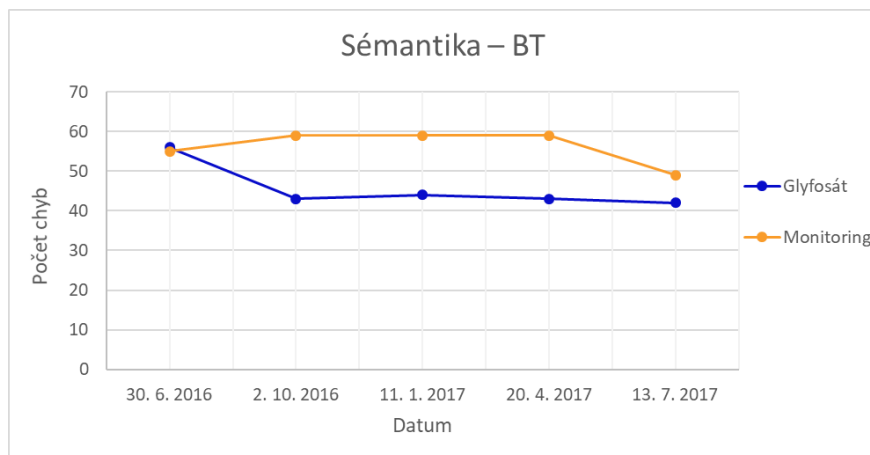
Překladač Google Translate podal při prvních třech hodnoceních shodný výkon, za který získal 151 záporných bodů. Třetí a čtvrtý překlad vykazoval mírné snížení počtu chyb, překladač získal 147 bodů, což je o čtyři negativní body méně než při předchozích hodnoceních. Přestože snížení počtu negativních bodů nebylo u výstupů z GT nijak markantní, vedl si překladač ve všech zkoumaných obdobích o něco lépe než Bing Translator.

Monitoring		
Datum překladu	Bing Translator	Google Translate
30. 6. 2016	167	151
2. 10. 2016	175,5	151
11. 1. 2017	175,5	151
20. 4. 2017	175,5	147
13. 7. 2017	150	147
Zlepšení o:	17	4
Celkový průměr (̅):	168,7	149,4

Tabulka 6: Text Monitoring: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)

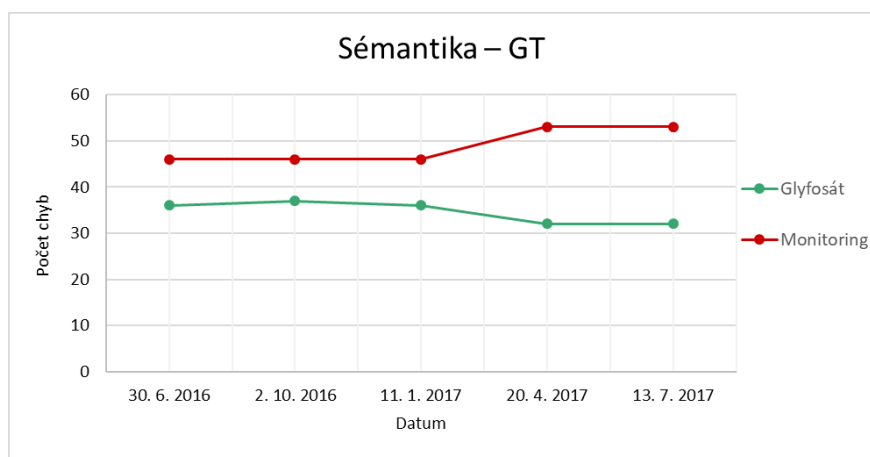
4.2.2 Vývoj počtu chyb ve třech vybraných kategoriích

V rámci informativních textů bude nejprve graficky znázorněna chybovost v oblasti sémantiky během celého zkoumaného období. Další dva grafy nastíní vývoj počtu pragmatických chyb, které byly zahrnuty do kategorie věcný faktor a poslední dva grafy ilustrují hodnocení gramatických jevů během celého zkoumaného období.



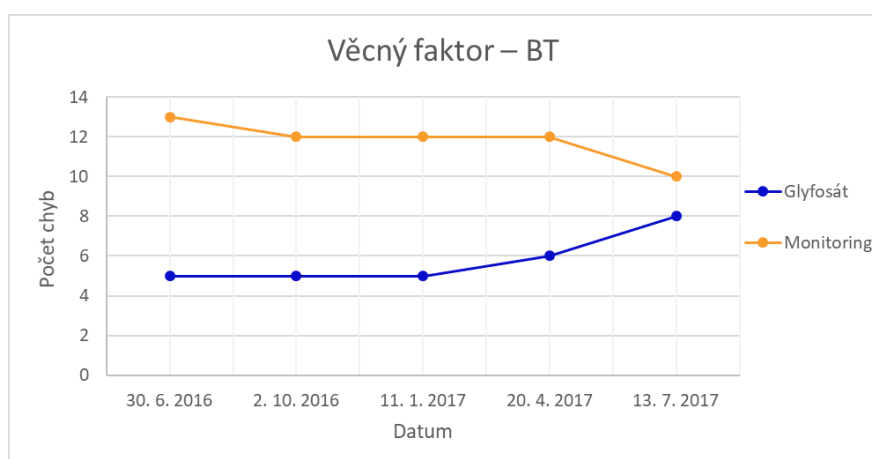
Graf 7: BT (informativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)

Zaměříme-li se na sémantické chyby ve výstupech z překladače Bing Translator (Graf 7), zjistíme, že při prvním překladu vykazovaly oba texty téměř stejné množství chyb. U zprávy o glyfosátu došlo při druhém překladu k poklesu chyb, výstup produkovaný v říjnu 2016 totiž obsahoval oproti originálu méně doplněných slov než překlad z června 2016. V textu o monitoringu se v prvních čtyřech obdobích nacházelo přibližně stejné množství sémantických chyb, ke zlepšení došlo až při posledním překladu, neboť výstup také obsahoval méně doplnění. Po sémantické stránce došlo u obou textů ke zlepšení stejného jevu, tedy ke snížení počtu chyb v sémantické kategorii doplnění oproti originálu.



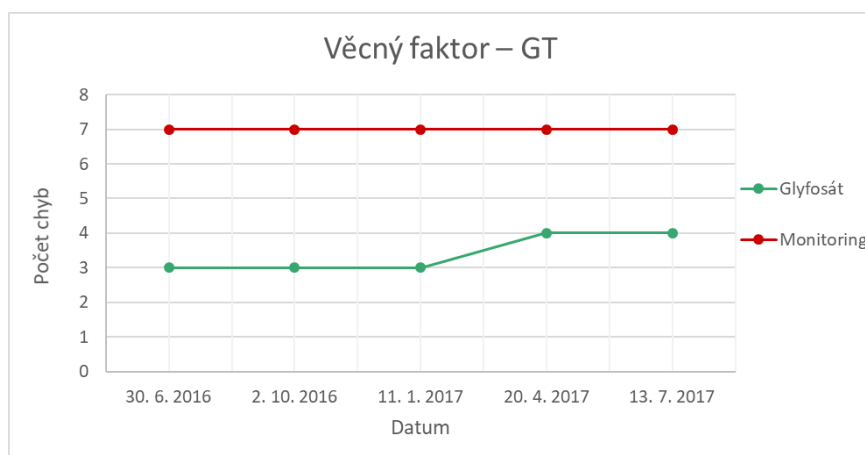
Graf 8: GT (informativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)

Graf 8 ilustruje vývoj počtu chyb na rovině sémantiky u informativních textů přeložených překladačem Google Translate. Překlad zprávy o glyfosátu obsahoval během celého zkoumaného období méně sémantických chyb než překlad textu o monitoringu ptactva. U glyfosátu navíc došlo při posledních dvou překladech v kategorii sémantiky k mírnému zlepšení, klesl totiž počet chybných doplnění oproti originálu. U překladů textu o monitoringu se objevil zcela opačný trend, v posledních dvou překladech došlo ke zvýšení počtu chybných doplnění oproti originálu a celkový počet negativních jevů v oblasti sémantiky se zvýšil o šest.



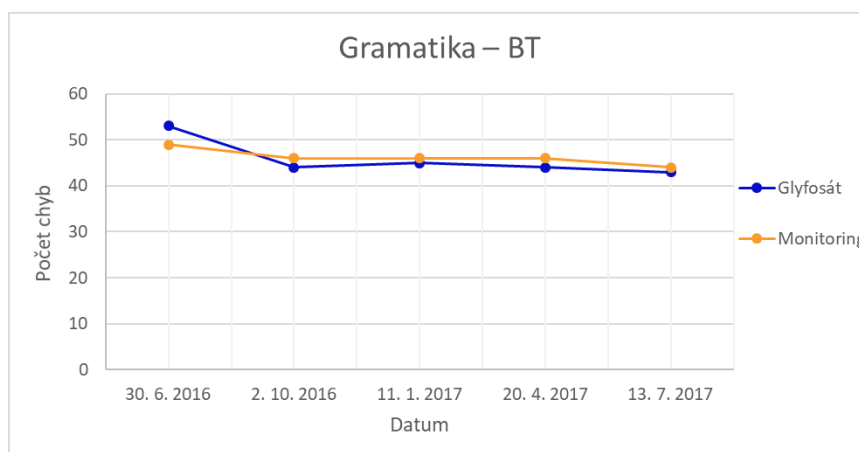
Graf 9: BT (informativní texty): Počet chyb v rámci celého zkoumaného období (věcný faktor)

Jak znázorňuje Graf 9, v kategorii věcného faktoru došlo při překladu textu o monitoringu k mírnému snížení počtu věcných chyb z 13 na 10. Při překladu druhého informativního textu (Glyfosát) sice překladač Bing Translator produkoval v rámci kategorie věcného faktoru méně chyb, avšak postupem času se jejich počet zvyšoval. Došlo k celkovému vzrůstu věcných chyb z původních pěti na osm.



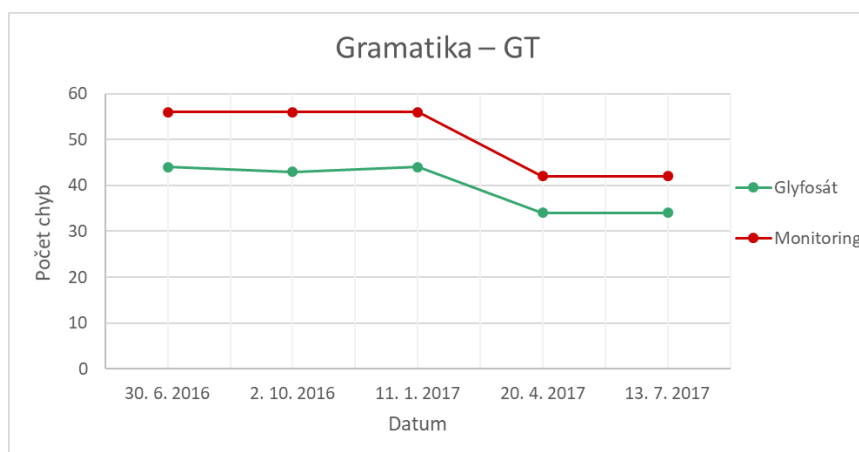
Graf 10: GT (informativní texty): Počet chyb v rámci celého zkoumaného období (věcný faktor)

Porovnáme-li výkony překladače Google Translate a Bing Translator v kategorii věcného faktoru (Graf 9 a Graf 10), zjistíme, že při překladu obou textů vyprodukoval Google méně věcných chyb. V textu o monitoringu nedošlo u GT z věcného hlediska k žádnému zlepšení, počet chyb zůstal konstantní po celou dobu experimentu. Při překladu druhého textu došlo v této kategorii během posledních dvou zkoumaných období k nepatrnému zvýšení počtu chyb ze tří na čtyři (Graf 10).



Graf 11: BT (informativní texty): Počet chyb v rámci celého zkoumaného období (gramatika)

V kategorii gramatiky byly počty chyb u výstupů z překladače Bing Translator u obou textů poměrně vyrovnané (Graf 11). Počet gramatických chyb se u prvního překladu pohyboval okolo 50, ve druhém překladu u obou textů počet gramatických chyb mírně klesl a až do konce zkoumaného období se ustálil okolo hranice 45 chyb. U zprávy o glyfosátu došlo k mírnému zlepšení v deklinaci a konjugaci, u textu o monitoringu proběhlo výraznější zlepšení v oblasti slovosledu.



Graf 12: GT (informativní texty): Počet chyb v rámci celého zkoumaného období (gramatika)

Graf 12 ukazuje, že se Google Translate při překladu informativních textů v obou případech zlepšil v gramatice. První tři překlady textu o monitoringu obsahovaly 56 gramatických, při čtvrtém překladu v dubnu 2017 došlo k poklesu chyb o 12, a to především v oblasti konjugace a deklinace. Při prvním a třetím překladu textu o glyfosátu vyprodukoval GT pokaždé 44 gramatických chyb, při druhém 43. V rámci dubnového hodnocení bylo zjištěno zlepšení a chybovost v oblasti gramatiky se snížila o 10 chyb oproti předchozímu překladu. Ubyla nesprávná interpunkce a chyby v deklinaci a konjugaci.

4.2.3 Celkové zhodnocení překladu pomocí stupnice

Co se týče informativních textů, při procentuálním hodnocení použitelnosti a dodržení funkce překladu docílil nejlepšího výsledku Google Translate při překladu zprávy o glyfosátu, získal \bar{x} 67 %, což je zároveň nejlepší výsledek celého druhého stupně hodnocení. Druhé nejlepší průměrné procentuální hodnocení překladů informativních textů obdržel také GT. Jednalo se o překlad textu o monitoringu ptactva, za který obdržel \bar{x} 63 %. Jak ilustruje Tabulka 7, z hlediska použitelnosti a funkce překladu podal Bing Translator při převodu obou textů horší výkon než GT. Při hodnocení textu o glyfosátu získal v průměru o sedm procent méně než GT a při hodnocení textu o monitoringu ptactva dokonce o deset procent méně.

Datum překladu	Glyfosát		Monitoring	
	Bing Translator	Google Translate	Bing Translator	Google Translate
30. 6. 2016	55 %	65 %	55 %	65 %
2. 10. 2016	60 %	65 %	50 %	65 %
11. 1. 2017	60 %	65 %	50 %	65 %
20. 4. 2017	60 %	70 %	50 %	60 %
13. 7. 2017	65 %	70 %	60 %	60 %
Celkový průměr (\bar{x}):	60 %	67 %	53 %	63 %

Tabulka 7: Procentuální hodnocení překladů informativních textů na stupnici podle Torrense (0–120 %) z hlediska jejich použitelnosti/funkce (čím vyšší procentuální hodnota, tím vyšší kvalita překladu)

4.3 Umělecké texty

4.3.1 Hodnocení překladů na základě translatologické analýzy

Posledním zkoumaným textovým typem jsou texty umělecké, které jsou zastoupeny úryvkem z dětské knížky od F. Schmeißera a fejetonem o nálezu Caravaggiova obrazu. Překlad uměleckého textu určeného pro děti způsoboval oběma překladačům velké problémy. Tabulka

8 obsahuje údaje, které dokládají, že oba překladače produkovaly při překladu Schmeißerova textu výstupy s velkým množstvím chyb a při hodnocení obdržely nejvyšší záporné skóre v rámci celé studie. Google Translate získal v prvních třech překladech 227 záporných bodů, pak došlo ke zvýšení chybovosti o 3,5 bodu, takže při hodnocení posledních dvou výstupů vzrostlo negativní skóre na 230,5 bodů.

První čtyři výstupy od překladače Bing Translator obsahovaly více chyb než překlady vyhotovené GT. Například za první překlad obdržel překladač od Microsoftu 251 negativních bodů, tedy o 24 negativních bodů více než Google Translate. Avšak na rozdíl od GT se Bing Translator během zkoumaného období zlepšoval a při posledním hodnocení obdržel dokonce jen 187,5 záporných bodů. Oproti prvnímu překladu tedy klesla chybovost o 63,5 negativních bodů, což je největší zlepšení dosažené za celou studii.

Schmeißer		
Datum překladu	Bing Translator	Google Translate
30. 6. 2016	251	227
2. 10. 2016	239,5	227
11. 1. 2017	239,5	227
20. 4. 2017	239,5	230,5
13. 7. 2017	187,5	230,5
Zlepšení o:	63,5	-3,5
Celkový průměr (ø):	231,4	228,4

Tabulka 8: Text Schmeißer: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)

Jak ukazuje Tabulka 9, ani při překladu fejetonu se chybovost u GT nesnižovala. Při prvních třech hodnoceních dosáhl v rámci translatologické analýzy 180 negativních bodů. V rámci hodnocení dubnového překladu došlo ke zvýšení chybovosti o 8,5 bodů, a tím i ke zvýšení negativního skóre na 188,5. Shodně vypadala i situace při posledním hodnocení.

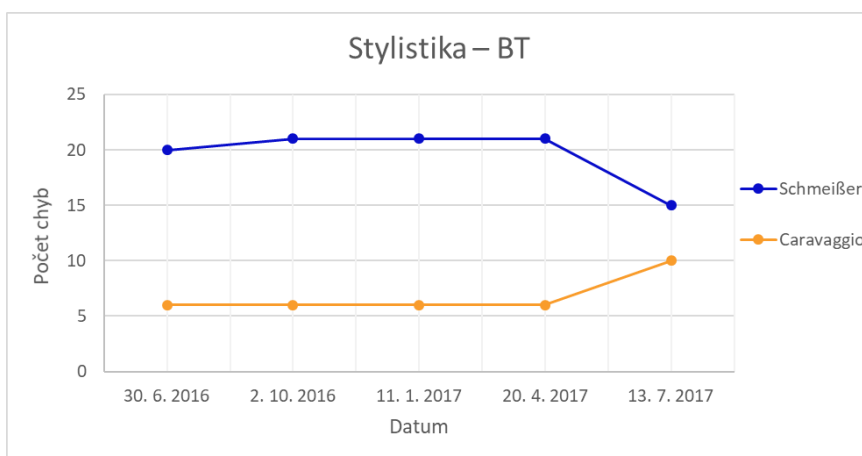
Tabulka 9 také ozřejmuje, že při překladu uměleckého textu o Caravaggiově obraze si Bing Translator vedl o něco lépe než Google Translate. Jeho první překlad byl ohodnocen 173 zápornými body, což je o sedm bodů nižší chybovost než u prvního překladu vyhotoveného GT. V následujících třech hodnoceních obdržely výstupy z BT 168 záporných bodů, došlo tedy ke zlepšení o pět bodů. Při posledním hodnocení z července 2017 lze pozorovat zvýšení chybovosti o čtyři záporné body, to znamená, že oproti prvnímu překladu došlo k celkovému zlepšení o jeden negativní bod.

Caravaggio		
Datum překladu	Bing Translator	Google Translate
30. 6. 2016	173	180
2. 10. 2016	168	180
11. 1. 2017	168	180
20. 4. 2017	168	188,5
13. 7. 2017	172	188,5
Zlepšení o:	1	-8,5
Celkový průměr (φ):	169,8	183,4

Tabulka 9: Text Caravaggio: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)

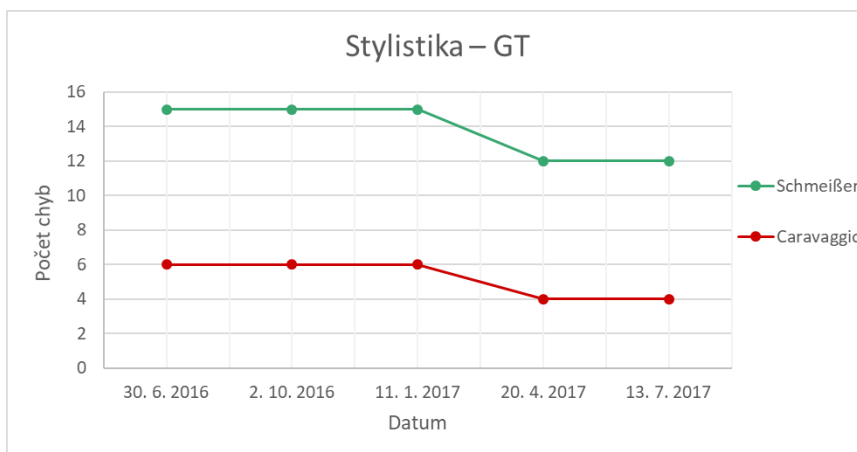
4.3.2 Vývoj počtu chyb ve třech vybraných kategoriích

U textů uměleckých se nejdříve zaměříme na to, jak se oba překladače s texty vyrovnaly po stylistické stránce, dále bude pomocí grafů znázorněn vývoj chybovosti v rámci sémantické kategorie. Třetím hodnoceným jevem je gramatika, poslední dva grafy tedy ukazují vývoj počtu gramatických chyb.



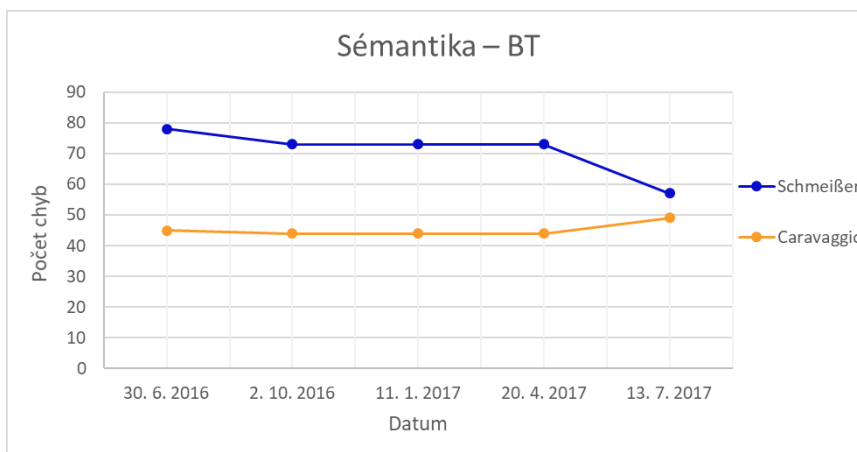
Graf 13: BT (umělecké texty): Počet chyb v rámci celého zkoumaného období (stylistika)

Co se týče stylistiky, v prvním překladu úryvku z dětské knížky vyhotoveném překladačem Bing Translator bylo zjištěno 20 stylistických prohřešků (Graf 13), při následujících třech hodnoceních bylo zaznamenáno nepatrné zhoršení. Při posledním překladu došlo k prudkému poklesu z 21 stylistických chyb na 15. Při překladu fejetonu o Caravaggiově obrazu produkoval BT výrazně méně stylistických chyb, zato však v této kategorii nedošlo k žádnému zlepšení, chybovost se naopak zvýšila ze šesti stylistických chyb na deset.



Graf 14: GT (umělecké texty): Počet chyb v rámci celého zkoumaného období (stylistika)

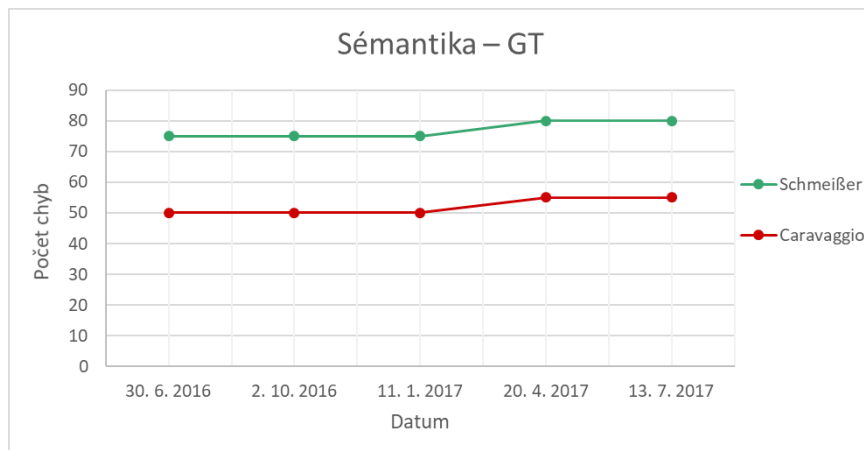
Porovnáme-li Graf 13 a Graf 14, je možné si všimnout, že výstupy z překladače Google Translate obsahovaly méně stylistických prohřešků než překlady vytvořené pomocí BT. První tři překlady beletristického textu pro děti od F. Schmeißera obsahovaly každý 15 stylistických chyb, v dubnovém překladu bylo zaznamenáno snížení počtu chyb o tři, stejného výsledku dosáhl GT i při posledním překladu v červenci 2017. Klesavou tendenci počtu stylistických chyb je možné pozorovat i u překladů fejetonu o Caravaggiu. Počet chyb v kategorii stylistické instrukce klesl při hodnocení posledních dvou překladů na čtyři.



Graf 15: BT (umělecké texty): Počet chyb v rámci celého zkoumaného období (sémantika)

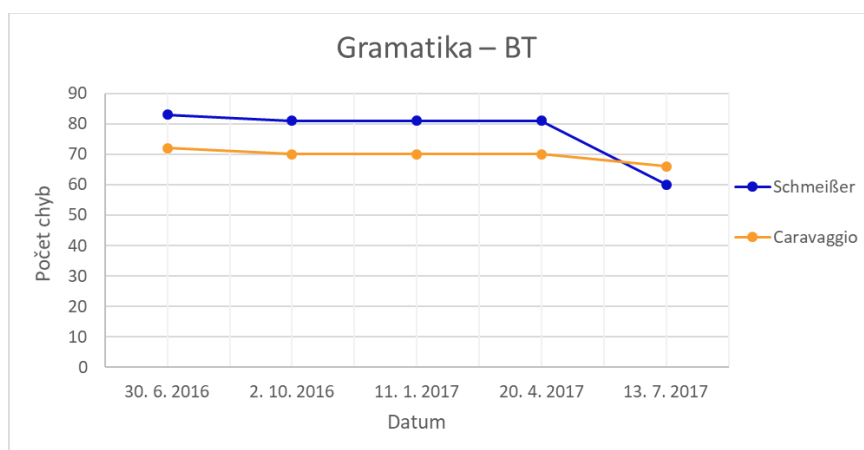
Graf 15 znázorňuje vývoj počtu sémantických chyb v rámci celého zkoumaného období u uměleckých textů vyhotovených překladačem Bing Translator. První čtyři překlady úryvku z dětské knihy obsahovaly výrazně více sémantických chyb než překlady fejetonu. Je nutné poznamenat, že právě první čtyři překlady Schmeißerova textu obsahovaly nejvyšší počet sémantických chyb ze všech textů vyhotovených BT. Při hodnocení však bylo v oblasti sémantiky zaznamenáno velké zlepšení. Počet sémantických chyb se z původních 78 snížil

v posledním překladu na 57, došlo totiž k poklesu chybných doplnění oproti originálu. Při překladu publicistického textu nastala situace opačná, počet sémantických chyb se zvýšil, a to z původních 45 na 49 chyb v posledním výstupu.



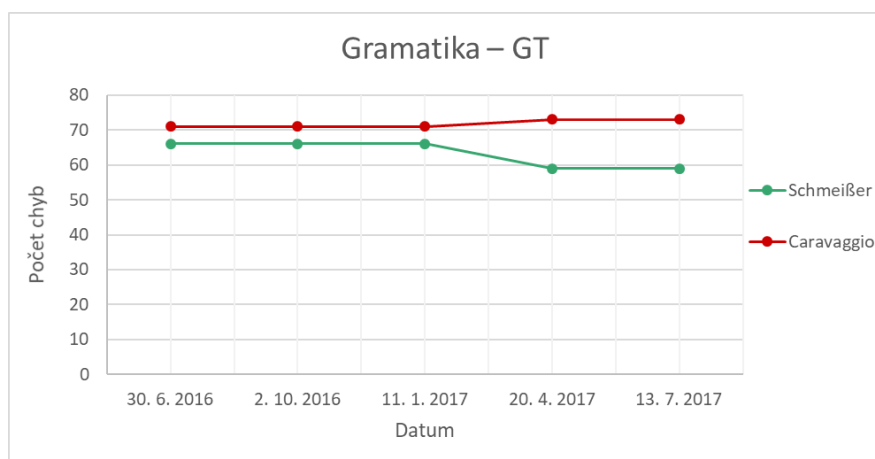
Graf 16: GT (umělecké texty): Počet chyb v rámci celého zkoumaného období (sémantika)

Také Google Translate dosáhl při překladu dětské beletrie v oblasti sémantiky svého nejhorsího skóre. První tři texty obsahovaly každý 75 sémantických chyb, v předposledním a posledním překladu bylo zjištěno dokonce 80 chyb (Graf 16). Přibyly chyby v kategorii mylných interpretací a homonymií. Překlad fejetonu sice obsahoval výrazně méně sémantických chyb než prvně zmíněný zástupce uměleckých textů, avšak ani u něj nedošlo po sémantické stránce ke zlepšení. Poslední dva překlady obsahovaly 55 sémantických chyb, což je o pět více než u překladů z předcházejících období.



Graf 17: BT (umělecké texty): Počet chyb v rámci celého zkoumaného období (gramatika)

Také v kategorii gramatiky docílil překladač Bing Translator u uměleckých textů svého nejhoršího výsledku. První tři překlady Schmeißerova textu obsahovaly více než 80 gramatických chyb. Při hodnocení posledního překladu bylo zjištěno 60 gramatických chyb, došlo tedy ke snížení počtu chyb o více než 20. Proběhlo totiž zlepšení v oblasti slovosledu, konjugace, deklinace a kongruence. Graf 17 ukazuje, že ke snížení chybovosti došlo i u posledního překladu publicistického textu, snížení však bylo minimální, počet gramatických chyb se snížil z počátečních 72 na 66 v posledním překladu.



Graf 18: GT (umělecké texty): Počet chyb v rámci celého zkoumaného období (sémantika)

Pokud srovnáme Graf 18 s ostatními grafy ilustrujícími vývoj gramatických chyb u výstupů z překladače Google Translate, vyplyne, že i GT vyprodukoval při překladu uměleckých textů po gramatické stránce nejvyšší počet chyb ze všech svých překladů. U fejetonu došlo k mírnému zvýšení počtu gramatických chyb, v prvních třech překladech bylo zjištěno 71 chyb, v posledních dvou 73. Úryvek z dětské knížky obsahoval v prvních třech zkoumaných obdobích 66 gramatických chyb, pak došlo k poklesu, poslední dva překlady tedy vykazovaly 59 gramatických prohřešků.

4.3.3 Celkové zhodnocení překladu pomocí stupnice

Nejhorší průměrné procentuální skóre v rámci celého druhého stupně hodnocení obdržel překladač Bing Translator při překladu uměleckých textů, jednalo se o 45 %. Také překladač Google Translate získal při překladu uměleckých textů své nejhorší hodnocení z celé druhé fáze. Jak dokládá Tabulka 10, jeho překlady úryvků z dětské knížky byly průměrně ohodnoceny

pouze 47 % a překlady fejetonu průměrně 50 %, což je i přes poměrně nízké hodnocení nejlepší výkon v kategorii uměleckých textů.

Datum překladu	Schmeißer		Caravaggio	
	Bing Translator	Google Translate	Bing Translator	Google Translate
30. 6. 2016	40 %	45 %	45 %	50 %
2. 10. 2016	45 %	45 %	45 %	50 %
11. 1. 2017	45 %	45 %	45 %	50 %
20. 4. 2017	45 %	50 %	45 %	50 %
13. 7. 2017	50 %	50 %	45 %	50 %
Celkový průměr (̄):	45 %	47 %	45 %	50 %

Tabulka 10: Procentuální hodnocení překladů uměleckých textů na stupnici podle Torrense (0–120 %) z hlediska jejich použitelnosti/funkce (čím vyšší procentuální hodnota, tím vyšší kvalita překladu)

5 Shrnutí výsledků

V rámci této kapitoly dojde ke shrnutí dílčích výsledků empirické studie, které byly představeny v předcházející kapitole. Tyto výsledky byly konfrontovány se vstupními hypotézami formulovanými v kapitole 3.1, aby mohla být provedena jejich verifikace.

Z první fáze hodnocení vyplynulo, že Google Translate produkuje výstupy lepší kvality než Bing Translator. Chyby na předem definovaných rovinách byly hodnoceny zápornými body, proto platí, že čím nižšího negativního skóre překladač dosáhl, tím vyšší je kvalita jeho překladů. Při zprůměrování výsledků získaných během celého zkoumaného období se zjistilo, že kvalita výstupů vyhotovených GT je sice vyšší než u BT, ale ne tolik, jak se očekávalo. Jak ilustruje Tabulka 11, celkový rozdíl činí pouze 11,6 negativních bodů.

Překladač	Bing Translator			Google Translate		
	Apelativní	Informativní	Umělecké	Apelativní	Informativní	Umělecké
Textový typ						
Ø (textový typ)	171,6	156,4	200,6	158,1	129,75	205,9
Ø (překladač)	176,2			164,6		

Tabulka 11: Celkové porovnání zprůměrovaných bodových hodnocení získaných v první fázi hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)

Shrnutí výsledků druhé fáze hodnocení obsahuje Tabulka 12. V druhém stupni hodnocení došlo k ověření celkové použitelnosti překladu a dodržení funkce. Texty byly hodnoceny intuitivně pomocí procent (0–120 %), přičemž vyšší procentuální hodnocení znamená vyšší kvalitu. Po zprůměrování výsledků za jednotlivá období bylo zjištěno, že i během druhé fáze hodnocení si Google Translate vedl lépe než překladač od Microsoftu.

Překladač	Bing Translator			Google Translate		
	Apelativní	Informativní	Umělecké	Apelativní	Informativní	Umělecké
Textový typ						
Ø (textový typ)	53 %	56,5 %	45 %	59,5 %	65 %	48,5 %
Ø (překladač)	51,5 %			57,7 %		

Tabulka 12: Celkové porovnání zprůměrovaných procentuálních hodnocení získaných v druhé fázi hodnocení (čím vyšší procentuální hodnota, tím vyšší kvalita překladu)

Pokud porovnáme první a druhou fázi hodnocení, zjistíme, že obě fáze přinesly totožný výsledek. Na základě výzkumu bylo zaznamenáno, že Google Translate produkoval kvalitnější překlady textů vybraných pro účely této empirické studie než překladač Bing Translator. Na základě toho zjištění tedy můžeme konstatovat, že došlo k potvrzení první hypotézy.

Mimo jiné jsme pracovaly s hypotézou, že u překladače Google Translate bylo NMT spuštěno pro češtinu v dubnu 2017. Podívejme se tedy na porovnání průměrného bodového

hodnocení překladače Bing Translator a Google Translate krátce po zavedení NMT u překladače Google Translate. Jak znázorňuje Tabulka 13, překlady apelativních a informativních textů přeložené neuronovým překladačem od Googlu 20. dubna 2017 získaly v rámci první fáze hodnocení výrazně nižší počet záporných bodů. To znamená, že překlady jmenovaných textových typů překladačem GT byly kvalitnější než překlady stejných textů od BT. Je však zajímavé, že i přes výrazně lepší výsledky při překladu informativních a apelativních textů si neuronový systém vedl při převodu uměleckých textů hůře než Bing Translator.

Překladač	Bing Translator			Google Translate (NMT)		
	Apelativní	Informativní	Umělecký	Apelativní	Informativní	Umělecký
Ø (typ textu, 20.4.)	173,75	156,5	203,75	153,75	123,5	209,5
Ø (překladač, 20.4.)	178			162,25		

Tabulka 13: Porovnání bodového hodnocení BT a GT krátce po spuštění Google Neural Machine Translation, duben 2017 (čím nižší číselná hodnota, tím vyšší kvalita překladu)

Podíváme-li se na zprůměrované výsledky bodového hodnocení překladů provedených v červenci 2017 (Tabulka 14), tedy přibližně po třech měsících od spuštění NMT u GT, zjistíme, že u překladů vyhotovených překladačem Bing Translator došlo ke zlepšení, počet negativních bodů oproti minulému hodnocení klesl, a to u všech textových typů (u uměleckých textů dokonce o 24 negativních bodů).

Zato u překladače GT se bodové hodnocení nezměnilo. Pokud porovnáme celkové bodové ohodnocení překladů provedených 13. 7. 2017, které bylo vypočítáno aritmetický průměrem ze všech bodů získaných daným překladačem ve zmíněném období, je možné konstatovat, že při posledním překladu si Bing Translator vedl nepatrně lépe. V souvislosti se spuštěním neuronového SP u Googlu můžeme poznamenat, že u překladů dvou třetin textů sice došlo v dubnu 2017 ke zlepšení, avšak v rámci jazykového páru němčina/čeština není nijak výrazné. V tomto případě tedy rozhodně nelze hovořit o „skokovém vylepšení překladu“²⁷.

Překladač	Bing Translator			Google Translate (NMT)		
	Apelativní	Informativní	Umělecký	Apelativní	Informativní	Umělecký
Ø (typ textu, 13.7.)	161	143,5	179,75	153,75	123,5	209,5
Ø (překladač, 13.7.)	161,42			162,5		

Tabulka 14: Porovnání bodového hodnocení BT a GT po spuštění Google Neural Machine Translation, červenec 2017 (čím nižší číselná hodnota, tím vyšší kvalita překladu)

²⁷ https://technet.idnes.cz/google-translate-neural-networks-dlq-sw_internet.aspx?c=A170418_224649_sw_internet_pka [cit. 2017-12-29]:

Co se týče jednotlivých textových typů, Tabulka 11 ukazuje, že nejlepší kvalitu vykazovaly překlady informativních textů, druhé místo obsadily texty apelativní a nejhorší hodnocení získaly texty umělecké. Z tohoto hlediska vypadaly výsledky u obou překladačů stejně. Celkově nejlepšího výsledku v rámci překladu jednotlivých textových typů dosáhl Google Translate, a to při překladu informativních textů (ø 129,75 negativních bodů). Stejný překladač získal i nejhorší hodnocení, konkrétně při překladu uměleckých textů (ø 205,9 záporných bodů). Na základě těchto zjištění můžeme druhou výzkumnou tezi rovněž považovat za prokázanou.

V rámci studie byla zkoumána historická podmíněnost výstupů. Vycházeli jsme z předpokladu, že v průběhu studie dojde ke zvýšení kvality překladů. Tabulka 15 znázorňuje, o kolik negativních bodů kleslo hodnocení mezi prvním překladem (30. 6. 2016) a posledním (13. 7. 2017) u jednotlivých textů. Nejvyšší pokles a tím pádem i největšího zlepšení dosáhl překladač Bing Translator při překladu dětské beletrie. Během zkoumaného období došlo k celkovému snížení o 63,5 záporných bodů. I druhé a třetí největší zlepšení zaznamenal překladač Bing Translator. Záporné bodové hodnocení u apelativního textu týkajícího se daňového poradenství kleslo v průběhu experimentu o 37,5 bodů a při hodnocení informativního textu o glyfosátu došlo ke snížení o 33 negativních bodů.

Ostatní bodové poklesy již nebyly tak výrazné. U některých textů dokonce došlo i ke zhoršení, respektive ke zvýšení počtu záporných bodů. Jednalo se o překlady apelativního textu, konkrétně o projev kancléřky Merkelové překladačem BT (zvýšení o 5 negativních bodů), a dále o překlady obou uměleckých textů vyhotovených překladačem Google Translate. U dětské beletrie došlo ke zvýšení o 3,5 záporných bodů, u fejetonu dokonce o 8,5. Třetí hypotéza se tedy potvrdila jen z části. Tabulka 15 ozřejmuje, že překladač Bing Translator dosáhl obecně většího zlepšení než překladač Google Translate.

Textový typ	Text	Bing Translator	Google Translate
Apelativní	Merkelová	-5	11
	Poradenství	37,5	3,5
Informativní	Glyfosát	33	16,5
	Monitoring	17	4
Umělecké	Schmeißer	63,5	-3,5
	Caravaggio	1	-8,5

Tabulka 15: Pokles počtu negativních bodů mezi prvním překladem v červnu 2016 a posledním v červenci 2017 (čím vyšší číslo, tím méně chybných jevů)

Spíše než výrazné zlepšení v oblasti kvality textů bylo možné pozorovat změny, které probíhaly u všech textových typů přeložených daným překladačem ve velmi podobných intervalech. U překladače Bing Translator většinou nevykazovaly výstupy přeložené během druhého, třetího a čtvrtého zkoumaného období žádné velké změny. Často se jednalo pouze o změny v oblasti formátování textu, které ale v rámci této studie nebyly hodnoceny.

Velmi podobně vypadala situace také u překladače Google Translate. V rámci prvních třech překladů většinou nedošlo téměř k žádným zlepšením nebo zhoršením, výraznější změny se objevily až u dubnového překladu, což s největší pravděpodobností souvisí se spuštěním neuronového strojového překladu. Pátý překlad byl ve většině případů identický s předposledním, z čehož by bylo možné vyvodit závěr, že bude-li díky zavedení NMT u překladače Google Translate docházet ke zlepšování, jak je předpokládáno, bude se tak dít v delších časových horizontech, než je období čtvrt roku.

6 Závěr a výhled k budoucímu výzkumu v dané oblasti

Tato diplomová práce se zabývala hodnocením kvality výstupů z veřejně dostupných překladačů Google Translate a Bing Translator v rámci jazykové dvojice němčina/čeština. Cílem bylo na základě srovnání jednotlivých překladů sestavit pořadí systémů podle kvality produktu, zjistit, jak se vypořádají s různými textovými typy a jestli dochází ke zlepšení kvality překladů v čase.

První část práce nabízí teoretický úvod do problematiky strojového překladu. Nejdříve byly definovány základní pojmy, pak byla nastíněna poměrně krátká, leč velmi bohatá historie vývoje strojového překladu, a to nejen ve světě, ale i u nás. Následně byly stručně popsány možné přístupy ke strojovému překladu, přičemž se kapitola věnovaná této problematice snažila podat základní informace o tom, jak jednotlivé systémy fungují a jaké jsou jejich hlavní výhody a nevýhody. Další podkapitola se věnovala online překladačům nejprve z obecného hlediska, pak se zaměřila na překladače vybrané pro účely empirické studie, jejíž výsledky byly vyhodnoceny v praktické části této práce. Dále byly s ohledem na cíl této práce popsány metody hodnocení kvality strojového překladu, a to jak ruční, tak automatické. Poté došlo ke shrnutí možných přístupů k hodnocení SP a byly popsány typické chyby, které se objevují ve strojově vyhotovených překladech.

V metodologické části byly nejdříve představeny vstupní hypotézy, které se zabývaly otázkou, zda bude Google Translate podávat lepší výsledky než Bing Translator, který textový typ bude při hodnocení vykazovat lepší výsledky a jestli se kvalita translátů bude v průběhu experimentu zvyšovat. Za účelem verifikace těchto výzkumných hypotéz došlo k sestavení analytického modelu, který se skládal ze dvou fází. V první fázi byl použit kritický model translatologické analýzy podle Kathariny Reißové, který byl ve druhé fázi doplněn procentuálním hodnocením vycházejícím ze stupnice podle A. Torrense. Tato kombinace se ukázala jako vhodný metodologický základ, protože první fáze se zaměřovala především na jednotlivé lingvistické kategorie, zatímco v rámci druhého stupně analýzy proběhlo celkové zhodnocení použitelnosti a dodržení funkce textu. Studie mimo jiné prokázala, že ačkoliv se metoda translatologické analýzy používá při hodnocení lidských překladů, je možné ji aplikovat i na strojově vyhotovené transláty.

Předpokladem analýzy v empirické části práce bylo získání překladů vybraných textů v pěti časových obdobích po dobu dvanácti měsíců. Z tohoto důvodu byl nejprve vytvořen

datový korpus výchozích textů, do nějž bylo vybráno šest textů. Každému textovému typu náležely dva z těchto textů. Texty z datového korpusu byly přeloženy pomocí veřejně dostupných překladačů Bing Translator a Google Translate, které patří k nejznámějším a nejpoužívanějším online překladačům. Vybrané překladače přeložily všechny texty celkem pětkrát s přibližně tříměsíčním odstupem, aby bylo možné se v rámci experimentu zaměřit také na to, jakou roli hraje časový aspekt. Vlastní analýza získaných textů měla za cíl posoudit, zda byly výzkumné hypotézy potvrzeny, či vyvráceny.

V kapitole 4 byly představeny výsledky, které byly získány na základě empirické studie. V rámci translatologické analýzy došlo k hodnocení jevů na různých rovinách, v praktické části práce pak byly tyto jevy popsány především z hlediska kvantitativního. Na základě počtu získaných negativních bodů bylo zkoumáno, který z vybraných překladačů produkuje kvalitnější výstupy. Dále byl u každého textu v rámci jednotlivých textových typů sledován vývoj počtu chyb v různých kategoriích během celého zkoumaného období. Pak byly prezentovány výsledky druhého stupně analýzy.

Nakonec došlo ke sloučení první a druhé fáze hodnocení a shrnutí získaných výsledků, které byly konfrontovány se vstupními hypotézami. Po shrnutí výsledků všech hodnocení se ukázalo, že se první hypotéza potvrdila. Překladač Google Translate obecně produkoval kvalitnější transláty než Bing Translator, i když rozdíl nebyl tak veliký, jak se očekávalo.

Dále jsme vycházeli z hypotézy, že u Google Translate byl v dubnu 2017 spuštěn neuronový strojový překlad. Odborná veřejnost v této souvislosti často hovořila o výrazném zlepšení kvality překladu. V rámci experimentu bylo zjištěno, že při překladu informativních a apelativních textů došlo po zavedení NMT ke zlepšení, avšak zvýšení kvality při překladu z němčiny do češtiny rozhodně nebylo výrazné. Při překladu uměleckých textů překladačem Google Translate bylo dokonce zjištěno mírné zhoršení.

Při ověřování druhé hypotézy bylo zjištěno, že po srovnání s výsledky rovněž došlo k jejímu potvrzení. Co do kvality se oba dva překladače obecně nejlépe vyrovnaly s texty informativními, o něco hůře si vedly při překladu apelativních textů a nejhorších výsledků dosáhly při překladu textů uměleckých.

Zaměříme-li se na třetí výzkumnou hypotézu, která se zabývala otázkou, zda dochází ke zvýšení kvality překladů v čase, dojdeme k závěru, že se potvrdila jen z části. U některých textů totiž zlepšení zaznamenáno bylo a u některých ne. Například při překladu obou

uměleckých textů překladačem Google Translate byla kvalita výstupů na konci studie horší než na začátku. V průběhu experimentu tedy ve většině případů nenastalo výrazné zlepšení kvality, rozhodně však můžeme potvrdit, že u všech textových typů došlo k vývoji v čase.

Výhodou zvoleného analytického modelu bylo omezení subjektivity díky použití modelu translatologické analýzy, který má pevně stanovená kritéria hodnocení. Nevýhodou tohoto přístupu byla náročnost sběru dat a především jejich vyhodnocování. Při první fázi hodnocení bylo problematické přiřadit chybu pouze jedné kategorii v případech, kdy ji bylo možné přiřadit více kategoriím. Záporným rysem druhé fáze hodnocení je poměrně nízká reprodukovatelnost vzhledem ke skutečnosti, že byla prováděna intuitivně.

Vzhledem k tomu, že téměř všechna odborná literatura k tomuto tématu je napsaná v angličtině, bylo pro účely této práce potřeba k některým často používaným anglickým termínům vytvořit odpovídající protějšky v českém jazyce. V oddílu 2.1.2, který se zabýval historií strojového překladu, sice bylo zmíněno, že SP má u nás bohatou tradici, avšak velká část příspěvků k tomuto tématu je v současné době vydávána v angličtině. Například odborné statě o výzkumech prováděných na Ústavu formální a aplikované lingvistiky MFF Univerzity Karlovy, který je v oblasti výzkum SP velmi aktivní, jsou produkovány převážně v anglickém jazyce.

Specifikem této diplomové práce byla skutečnost, že se zabývala hodnocením strojového překladu mezi méně často zkoumanou jazykovou dvojicí němčina/čeština, konkrétně ve směru z němčiny do češtiny. Velká část prací, které se při výzkumu SP zaměřují na český jazyk, se věnuje jazykovému páru angličtina/čeština.

Pokud bychom chtěli obecně shrnout chyby, které se objevovaly v překladech, na úrovni sémantiky bychom mohli jmenovat mylné interpretace slov způsobené nezohledňováním kontextu a významové posuny. Po lexikální stránce někdy nedošlo k rozpoznání frazému, proto byl přeložen po jednotlivých slovech. Dále měly překladače občas problémy s rozeznáváním vlastních jmen a názvů. V kategorii stylistiky někdy docházelo v rámci jednoho textu k neopodstatněnému užívání slov různých stylistických rovin.

Co se týče gramatiky, byly systémy v rámci hodnocení často penalizovány za chybnou konjugaci, deklinaci a kongruenci. Také poměrně často docházelo k nerespektování aktuálního členění větného a pravidel interpunkce nebo nedodržování českého úzu při psaní uvozovek, takže se v některých překladech projevila interference z angličtiny. Překlady vyhotovené

překladačem Google Translate pomocí NMT už tuto interferenci neobsahovaly. Dále bylo zjištěno, že některé překlady obsahují špatně převedenou negaci. Systémy v takových případech nerespektovaly zdvojenou českou negaci (*žádný chtěl*) nebo naopak používaly zcela nesprávně negaci několikanásobnou (*neměl nikdo nevěnoval*). V překladu se objevily i nespisovné výrazy (*aby jste, pořádněj*). V této souvislosti je nutné si uvědomit, že kvalita výstupů závisí také na kvalitě korpusových dat. Zmíněné případy dokládají, že korpusy obsahují i texty nerespektující jazykovou normu. Z tohoto důvodu je potřeba počítat s tím, že ve strojově vyhotovených překladech může dojít k výskytu takovýchto tvarů.

Měli bychom mít na paměti, že máme k dispozici pouze výstupní data. Můžeme se sice snažit na základě jejich analýzy a obecných teoretických poznatků o fungování překladačů odhadnout, proč tyto chyby vznikly, avšak jejich pravou povahu nemůžeme odhalit. Google a Microsoft totiž poskytují jen velmi omezený vhled do toho, na jakém principu jejich překladače fungují. Na svých stránkách ve většině případů zveřejňují jen velmi obecné informace, aby nedošlo k prozrazení jejich obchodního tajemství. Například na stránkách věnovaných překladači Google Translate nebylo možné dohledat ani aktuální informace o tom, pro jaké jazyky je NMT v současné době spuštěno nebo kdy ke spuštění u jednotlivých jazyků došlo. Vzhledem k nedostatku spolehlivých informací o použitých strojových překladačích si tato práce neklade ambice systémy zlepšovat.

Nyní se zaměříme na výhled k budoucímu výzkumu v oblasti zkoumání kvality strojového překladu. Tyto úvahy vychází z předcházejících kapitol a stručně nastiňují směřování případného dalšího výzkumu související problematiky. Úvahy se dotýkají pouze některých vybraných aspektů této diplomové práce, nelze totiž poskytnout vyčerpávající výčet všech možných směrů případného budoucího výzkumu.

Při psaní této diplomové práce bylo zjištěno, že existuje jen velmi málo zdrojů pojednávajících o vývoji SP u nás, ačkoliv hrál československý a později český výzkum v této oblasti vždy nezanedbatelnou roli. Rozhodně by bylo vhodné zaměřit se v rámci budoucího výzkumu také na historii tohoto oboru jak v bývalém Československu, tak i v současné době.

Dále bylo nutné se potýkat s nedostatkem zdrojů k tématu typických chyb, které se objevují ve strojově vyhotovených překladech. Více zdrojů k této problematice by mohlo pomoci usnadnit interpretaci některých chyb. Bylo by zajímavé, kdyby se výzkum zaměřil nejen na typické chyby obecně, ale i na časté chyby vyskytující se při překladu z češtiny do

jednotlivých jazyků, neboť každý jazyk má svá specifika a tím pádem dochází v rámci různých jazykových dvojic ke specifickým chybám.

Co se týče hodnocení, model použitý v tomto výzkumu byl zvolen tak, aby bylo možné jej aplikovat při hodnocení různých textových typů. Další studie by se mohly zaměřit na jednotlivé textové typy. Avšak v tomto případě by bylo vhodné model upravit, aby více odpovídal potřebám zvoleného textového typu, nebo sestavit model nový. Při hodnocení uměleckých textů by mohl být použit například model Gideona Touryho, který se primárně věnoval překladu literárnímu.

Místo použitého analytického modelu by bylo možné v dalším podobně zaměřeném výzkumu hodnotit kvalitu translátů na základě měření času potřebného k jejich posteditaci (*Post-Editing Machine Translation*). Při použití této metody hodnocení je však nutné si uvědomit, že dovednosti potřebné k posteditaci se v některých aspektech liší od dovedností překladatelských.

Provedená studie ukázala, že ve zkoumaném období sice docházelo při překladu mezi němčinou a češtinou ke změnám, avšak o velkém zvýšení kvality translátů ve většině případů nelze hovořit. Určitě by proto bylo zajímavé zaměřit se na vývoj kvality strojově vyhotovených překladů v rámci delšího období, než je jeden rok, jak tomu bylo u této studie.

Vzhledem k zavedení neuronového strojového překladu jak u překladače Google Translate, tak u překladače Bing Translator lze v rámci budoucího výzkumu v dané oblasti počítat se studii, které se budou zabývat srovnáním kvality překladů, které vyprodukuje tyto systémy za použití NMT.

Třebaže v rámci prováděné studie došlo k potvrzení prvních dvou hypotéz a částečnému potvrzení třetí, rozhodně se nejedná o vyčerpané a uzavřené téma. Výzkum strojového překladu, jehož další možné směřování bylo stručně nastíněno v předchozích odstavcích, je vzhledem ke svému širokému poli působnosti žádoucí, a to především v oblasti kvality a jejího vývoje u různých textových typů.

7 Bibliografie

- ADAMEC, R. (2015). *Testování strojového překladu*. Plzeň: Západočeská univerzita v Plzni.
- AHMED, A., & HANNEMAN, G. (2005). Syntax-Based Statistical Machine Translation: A review. *Computational Linguistics*. Získáno 10. listopadu 2017, z <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.517.4686&rep=rep1&type=pdf>
- AIKEN, M. (2010). The Efficacy of Round-trip Translation for MT Evaluation. *Translation Journal* 14 (1), On-line: <http://mt-archive.info/TranslationJ-2010-Aiken.pdf>.
- ALCINA, A. (2008). Translation technologies: Scope, tools and resources. *Target: International Journal on Translation Studies*. Vol. 20, Issue 1, 79–102.
- ALPAC, A. (1966). *Language and machines: Computers in Translation and Linguistics – A Report by the Automatic Language Processing Advisory Committee*. Washington, D.C.: National Academy of Sciences.
- ANASTASIOU, D. (2010). *Idiom Treatment Experiments in Machine Translation*. Saarbrücken: Universität des Saarlandes.
- ARNOLD et al., D. (1994). *Machine Translation: An Introductory Guide*. Manchester: NCC Blackwell.
- ARTETXE, M., & CHO, K. (2017). Unsupervised Neural Machine Translation. *Conference Paper at the International Conference on Learning 2018*. Získáno 12. 11. 2017, z <https://arxiv.org/pdf/1710.11041.pdf>
- BAHDANAU et al., D. (2014). Neural machine translation by jointly learning to align and translate . *Proceedings of the 2014 International Conference on Learning Representations*.
- BAUER, R. (2002). *Linguistische Evaluation maschineller Übersetzungssysteme*. Leipzig: Marcus-Schr.
- BISHOP, H., & CROOK, M. (1965). Evaluation of machine translation. *Final report*.
- BOJAR, O. (2012). *Čeština a strojový překlad: Strojový překlad našincům, našinci strojovému překladu*. Praha: Ústav formální a aplikované lingvistiky.
- BOJAR, O. (2012b). Strojový překlad . *Vesmír* 91, 488–490.
- BÜHLER, K. (1934). *Sprachtheorie: die Darstellungsfunktion der Sprache*. Jena: Gustav Fischer.
- CER et al., D. (2010). The best lexical metric for phrase-based statistical MT system optimization. *HLT-NAACL. North American Association for Computational Linguistics-Human Language Technologies*, 555-563.
- CORPAS PASTOR, G., & MORENO ORTIZ, A. (1997). Corpas Pastor, G. & Moreno Ortiz, A. (en prensa) "Technical and Linguistic Analysis of a Commercial Machine Translation System (Power Translator). *La palabra vertida: investigaciones en torno a la traducción. Actas de los VI Encuentros Complutenses en torno a la Traducción*, 705-713.
- COSTA-JUSÁ et al., M. (2012). Study and Comparison of Rule-based and Statistical Catalan-Spanish Machine Translation Systems. *Computing and Informatics*, Vol. 31, 245 – 270. Získáno 23. října 2017, z <http://www.cai.sk/ojs/index.php/cai/article/download/940/430>

- COSTA-JUSSÁ, M., & FONOLLOSA, J. (2015). Latest trends in hybrid machine translation and its applications. *Computer Speech & Language* 32, 1, 3-10. Získáno 10. listopadu 2017, z https://ac.els-cdn.com/S0885230814001077/1-s2.0-S0885230814001077-main.pdf?_tid=81f738ec-ce18-11e7-84a4-00000aacb35d&acdnat=1511199211_f0dac96f034f697ee3e911ada36d9506
- EuroMatrix. (2007). *1.3: Survey of Machine Translation Evaluation*. Public Distribution: Project funded by the European Community under the Sixth Framework Programme for Research and Technological Development.
- FREIGANG, K.-H. (2001). Automation of Translation: Past, Presence, and Future. *Revista tradumàtica*. Získáno 8. srpna 2017, z <http://www.mt-archive.info/RevTradumatica-2001-Freigang>
- GARNELL, X. (2014). *Multilingual Information Management: Information, Technology and Translators*. Waltman, Massachussets: Chandos Publishing.
- GASPARI, F., & HUTCHINS, J. (2007). Online and Free! Ten Years of Online Machine Translation: Origins, Developments, Current Use and Future Prospects. *Proceeding of the MT Summit XI, 10-14 September 2007, Copenhagen, Denmark*, 199-206.
- GIMÉNEZ, J. (2008). *Empirical Machine Translation and its Evaluation*. Barcelona: Universitat Politècnica de Catalunya.
- HUTCHINS, J. (2012). *Historical sketch of machine translation in Eastern and Central Europe*. Hamburg: Crosslingual Language Technology in service of an integrated multilingual Europe. Získáno 1. listopadu 2017, z <http://www.hutchinsweb.me.uk/Hamburg-2013.pdf>
- HUTCHINS, W. (2003). Machine translation: General Overview. V R. MITKOV, *The Oxford Handbook of Computational Linguistics* (stránky 501-511). Oxford: University Press.
- HUTCHINS, W. (2006). Machine Translation: History. V K. BROWN, *Encyclopedia of Language & Linguistics, Second Edition, volume 7* (stránky 375-383). Oxford: Elsevier.
- HUTCHINS, W. (2014). Machine translation: History of Research and Applications. V C. SIN-WAI, *The Routledge Encyclopedia of Translation Technology* (stránky 120-135). London/New York: Routledge. Získáno 11. 8. 2017, z <http://www.hutchinsweb.me.uk/Routledge-2014.pdf>
- HUTCHINS, W., & SOMERS, H. (1992). *An Introduction to Machine Translation*. London: Academic Press.
- JUMPELT, R. (1961). *Die Übersetzung naturwissenschaftlicher und technischer Litaratur*. Berlin-Schöneberg: Langenscheidt Bibliothek Bd.1.
- KADE, O. (1964). *Subjektive und objektive Faktore nim Übersetzungsprozess*. Leipzig.
- KALCHBRENNER, N., & BLUNSOM, P. (2013). Recurrent continuous translation models. *Proceedings of the ACL Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1700-1709.
- KIRSCHNER, Z. (2000). Pioneer work in machine translation in Czechoslovakia. V W. HUTCHINS, *Early years in machine translation: memoirs and biographies of pioneers* (stránky 349-360). Amsterdam: John Benjamins.
- KOEHN, P. (2010). *Statistical Machine Translation*. Cambridge: Cambridge University Press.

- KOEHN, P., & KNOWLES, R. (2017). Six challenges for neural machine translation. *Proceedings of the First Workshop on Neural Machine Translation*, 28-39. Získáno 16. 11. 2017, z <http://www.aclweb.org/anthology/W17-3204>
- KOERNER, E., & ASHER, R. (1996). *Concise History of the Language Sciences: From the Sumerians to the Cognitivists*. New York: Pergamon.
- KRÁL, P. (2003). Využití počítačových technologií v překladatelském procesu. *ToP (Tlumočení-Překlad)*(69), 15–15.
- KUBOŇ, V. (2012). Machine Translation among Related Slavic Languages. V C. VERTAN, *Multilingual Processing in Eastern and Southern EU Languages: Low-Resourced Technologies and Translation* (stránky 283-307). Cambridge: Scholars Publishing.
- LAVIE, A. (2010). *Evaluating the Output of Machine Translation Systems*. Association for MT in the Americas (AMTA) Tutorial, Denver, Colorado, USA.
- LEE, C. (2017). *Multilingualism Online*. Londýn/New York: Routledge.
- LEVENSHTAIN, V. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics - Doklady* 10, 707-710.
- LEVÝ, J. (1998). *Umění překladau*. Praha: Ivo Železný.
- MAČURA, M. (2011). Limitation of CAT and MT Technology. V J. ZEHNALOVÁ et al., *Teaching Translation and Interpreting Skills in the 21st Century* (stránky 207-219). Olomouc: Univerzita Palackého v Olomouci.
- MACHÁČEK, M. (2012). *Metriky pro optimalizaci modelů strojového překladau*. Praha: Matematicko-fyzikální fakulta, Univerzita Karlova.
- MANNING, C., & SOCHER, R. (2017). *Natural Language Processing with Deep Learning*. Stanford: Stanford University. Získáno 17. 11. 2017, z <https://web.stanford.edu/class/cs224n/lectures/cs224n-2017-lecture10.pdf>
- MILLER, G., & BEEBE-CENTER, J. (1956). Some Psychological Methods for Evaluating the Quality of Translation, 3 (3) . *Mechanical Translation* 3, 73-80.
- MOUNIN, G. (1967). *Die Übersetzung. Geschichte, Theorie, Anwendung*. München.
- NAGAO, M. (1984). A Framework of a Mechanical Translation between Japanese and English by Analogy Principle. *Artifitial and Human Intelligence*, 173-180.
- NEWMARK, P. (1988). *A Textbook of Translation. About Translation*. Clevedon: Multilingual Matters.
- NIDA, E. (1964). *Toward a Science of Translating*. Leiden: E.J. Brill.
- OKPOR, M. (2014). Machine Translation Approaches: Issues and Challenges. *International Journal of Computer Science Issues*, Vol. 11, Issue 5, No 2, 159-165. Získáno 30. října 2017, z <https://www.ijcsi.org/papers/IJCSI-11-5-2-159-165.pdf>
- OLADOSU et al., J. (2016). Approaches to Machine Translation: A Review. *FUOYE Journal of Engineering and Technology*, Volume 1, Issue 1, 120-126.
- PAPINENI et al., K. (2002). BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting of the ACL*, 311-318.

- PFAFFLIN, S. (1965). Evaluation of Machine Translations by Reading Comprehension Tests and Subjective Judgments. *Mechanical Translation and Computational Linguistics*, 8 (2), 2-8.
- POIBEAU, T. (2017). *Machine Translation*. London/Cambridge: The MIT Press.
- POPOVIĆ, M., & NEY, H. (2011). Towards Automatic Error Analysis of Machine Translation Output. *Computational Linguistics, Volume 37, Number 4*, 657-688.
- POPOVIČ, A. (1975). *Teória umeleckého prekladu: aspekty textu a literárnej metakomunikácie*. Bratislava: Tatran.
- REIMANN, M. (1991). *Maschinelle Übersetzung. Ein Überblick über Theorie und Praxis*. Berlin/Heidelberg.: Springer-Verlag.
- REIB, K. (1986). *Möglichkeiten und Grenzen der Übersetzungskritik. Kategorien und Kriterien für eine sachgerechte Beurteilung von Übersetzungen*. München: Hueber.
- RUBENS, P. (2014). Building Babel: Lost in machine translation. *BBC FUTURE*. Získáno 18. 11. 2017, z <http://www.bbc.com/future/story/20120306-lost-in-machine-translation>
- SHUK MAN, C. (2015). The Teaching of Machine Translation. V C. SIN-WAI, *The Routledge Encyclopedia of Translation Technology*. London/New York: Routledge.
- SCHWARZL, A. (2001). *The (Im)Possibilities of Machine Translation*. Frankfurt am Main: Peter Lang GmbH.
- SIN-WAI, C. (2016). *The Future of Translation Technology: Towards a World Without Babel*. New York: Routledge.
- SPELLNER, B. (1980). Semiotische Aspekte der Übersetzung von Comics-Texten. V W. WILSS, *Semiotik und Übersetzen*, (stránky 73-86). Tübingen: Narr.
- STEIN, D. (2013). Machine Translation: Past, Present and Future. *Translation: Computation, Corpora, Cognition 3.1*, V-XIII. Získáno 9. srpna 2017, z <https://www.blogs.uni-mainz.de/fb06-tc3/files/2015/11/28-143-1-PB.pdf>
- SVOBODA, T. (2012). *Kapitoly z překladatelské praxe: odborný překlad mezi němčinou a češtinou*. Praha: Univerzita Karlova, Filozofická fakulta.
- SVOBODA, T. (2015). Hodnocení kvality strojového překladu. V J. ZEHNALOVÁ, & et al., *Kvalita a hodnocení překladu: modely a aplikace* (stránky 245-272). Olomouc: Univerzita Palackého.
- TORRENS, A. (1994). Machine Translation Evaluation and Quality Benchmarks. *Terminologie et Traduction, 1*, stránky 375–415.
- TÝNOVSKÝ, M. (2007). *Využití lingvistických informací při EBMT*. Praha: Matematicko-fyzikální fakulta Univerzity Karlovy.
- VILAR et al., D. (2006). Error Analysis of Machine Translation Output. *International Conference on Language Resources and Evaluation*, 697–702.
- WESLEY-TANASKOVIC et al., I. (1994). *Expanding Access to Science and Technology: The Role of Information Technologies*. Tokyo.
- WU et al., Y. (2016). *Google's Neural Machine Translation System: Bridging the Gap. between Human and Machine Translation*. Technical Report. Získáno 19. 11. 2017, z <https://arxiv.org/pdf/1609.08144.pdf>

ZEHNALOVÁ et al., J. (2015). *Kvalita a hodnocení překladu: Modely a aplikace*. Olomouc: Univerzita Palackého v Olomouci.

ZHANG, X., & LUI, Q. (2014). Machine Translation. General. V C. SIN-WAI, *The Routledge Encyclopedia of Translation Technology* (stránky 105-119). London/New York: Routledge.

Hodnocené texty:

Apelativní texty (appelbetonte Texte):

Rede von Bundeskanzlerin Merkel im Rahmen der Plenarsitzung beim World Humanitarian Summit (Merkelová), online [25.6.2016], dostupné z:

<https://www.bundesregierung.de/Content/DE/Rede/2016/05/2016-05-23-merkel-plenarsitzung-whs.html>

Schuldenberatung (Poradentví), online [25.6.2016], dostupné z:

<http://www.betreuungsverein-bremerhaven.de/schuldnerberatung-nach-305-inso/>

Informativní texty (inhaltsbetonte Texte):

EU-Kommission verlängert Zulassung von Glyphosat (Glyphosát), online [27.7.2016], dostupné z:

<http://www.dw.com/de/eu-kommission-verl%C3%A4ngert-zulassung-von-glyphosat/a-19364455>

Verordnung zur Durchführung eines Monitorings auf das Virus der Geflügelpest bei Wildvögeln (Monitoring), online [29.6.2016], dostupné z:

<https://www.gesetze-im-internet.de/bundesrecht/wvgeflpestmonv/gesamt.pdf>

Umělecké texty (formbetonte Texte):

Allein unter Dieben (Schmeißer), online [25.6.2016], dostupné z:

http://www.fischerverlage.de/media/fs/308/LP_978-3-7373-5143-0.pdf

Caravaggio-Fieber. Nichts für unser Wohnzimmer, Chérie! (Caravaggio), online [27.6.2016], dostupné z:

<http://www.faz.net/aktuell/feuilleton/kunst/caravaggio-gemaelde-fund-auf-dachboden-bei-toulouse-14176587.html>

8 Přílohy

Z kapacitních důvodů nebylo možné zařadit do přílohy všechny ohodnocené překlady. Tento oddíl proto obsahuje pouze ukázkou grafického hodnocení dvou textů. V obou případech se jedná o překlad článku o prodloužení povolení herbicidu glyfosátu. Pro srovnání je zde možné nalézt jak text vyprodukovaný překladačem Bing Translator, tak Google Translate. Jedná se o překlady, které byly v rámci empirické studie provedeny 13. 7. 2017, tedy až na jejím konci. Ostatní ohodnocené výstupy jsou dostupné v elektronické podobě na přiloženém CD.

Kromě všech šedesáti graficky ohodnocených textů obsahuje CD také překlady bez grafického hodnocení, originály textů vybraných pro experiment do datového korpusu a excelové soubory, které obsahují pro každý přeložený text jednu tabulku s bodový hodnocením jevů zkoumaných v rámci translatologické analýzy.

Bing Translator – 13. 7. 2017 (Glyfosát)

Komise EU rozšiřuje povolení glyfosátu
Evropská Komise se rozhodla pro rozšíření povolení glyfosátu, a proto je v ostré kritice. Studie je nyní prozkoumat potenciální riziko rakoviny u rostlin pesticidů.
Evropská Komise rozšířila schválení glyfosátu kontroverzní ošetřující výrobky od 18 měsíců. To byl sdílen orgánem Bruselské ve středu. Do konce roku 2017 má nová studie EU chemické látky agentury ECHA přezkoumat, zda opravný prostředek představuje nebezpečí pro zdraví. Federální vláda pro životní prostředí a ochranu přírody Německo (Bund) vyjadřuje ostrou kritiku rozhodnutí. "Toto rozhodnutí o glyfosát je urážka mnoha přesvědčit občany," řekl předseda federální Hubert odmítá. "Stále více lidí v Evropě a v Německu požadoval zemědělství bez pesticidů. Politika to musí konečně vzít v úvahu."
Odvětví zemědělství, což je náročné přebírání na 15 let, také kritizoval rozhodnutí Komise. Pracovat Místo s znovu glyfosát s "technickými rozšíření", je te pouze nevyhovující provizorní řešení podle zemědělsko-průmyslové asociace (IVA).
Obě strany si stěžoval, že rozhodnutí Evropské Komise bylo kvůli tlaku ze strany sdružení lobby. Centrální zahradnictví asociace (ZVG) hovořil o "politickému tlaku ze zúčastněných stran a nevládních organizací".
Rozhodnutí Evropské Komise o rozšíření přichází jeden den před uplynutím předchozího povolení de 30. června. Státy EU se dříve, minul se dohodnout na postoji většiny přes opakované pokusy. Minulý pátek, zůstalo konečné hlasování neopodstatněné, s níž Německo se hlasování zdrželo.
Glyfosát je nejběžněji používané herbicidy v EU a je v podezření, že jsou karcinogenní. Na druhé straně je negativní dopad na životní prostředí nepopíratelné a zabíjení všech možných nežádoucích rostlin, například kukuřice květiny. Na obrovské pole osázené monokultur totálního herbicidu ničí potraviny základnu a stanoviště mnoha zvířat, a tak ohrožuje biologickou rozmanitost.

Vysvětlivky:

Sémantické instrukce (vypustka, doplnění)

Lexikální instrukce

Gramatické instrukce

Stylistické instrukce

Pragmatické kategorie

Zdařilá řešení

V pořádku

BT 13.7.2017 Glyfosát				
Instrukce	Jevy	Body	Hodnocení	Celkem
Sémantické	polysémie	-2		84
	homonymie			
	výpustky		12	
	doplnění		9	
	posuny, mylná interpretace		21	
Lexikální	odborné termíny	-1		5
	idiomy			
	jména		5	
	„faux amis“			
	metafory			
	přísloví			
	slovní hříčky			
Gramatické	gramatická správnost	-1	29	43
	slovosled		5	
	interpunkce		9	
Stylistické	stylistické ochuzení	-0,5		0
	pestrost slohu			
	mísení stylů			
Mimojazykové determinanty	faktor místa	-1		8
	faktor času			
	věcný faktor		8	
bez chyb	Bez úprav	0		
Velmi zdařilá řešení		0,5	6	3
Celkem				137

Tabulka 16: Ukázka bodového hodnocení v rámci translátologické analýzy překladu zprávy o glyfosátu, který byl vyhotoven překladačem BT (13. 7. 2017)

Google Translate – 13. 7. 2017 (Glyfosát)

Evropská komise rozšířila schválení glyfosátu

Evropská komise se rozhodla pro obnovu povolení glyfosátu a udeřil s ostrou kritikou. Studie se má nyní vyšetřuje možné riziko rakoviny o přípravku na ochranu rostlin.

Evropská komise prodloužila schválení kontroverzní glyfosátu pesticidů o 18 měsíců. To bylo oznámeno orgány v Bruselu ve středu. Do konce roku 2017 má nové studie Agentury EU pro chemikálie ECHA zkoumat, zda prostředek je zdraví škodlivé.

Federace pro životní prostředí a ochranu přírody Německa (BUND) vyjádřil ostrou kritiku rozhodnutí. „Ty glyfosát rozhodnutí je urážkou přesvědčil mnoho občanů,“ řekl předseda BUND Hubert Weiger. „Stále více lidí v Evropě a Německu nazývá pro zemědělství bez pesticidů. Politici to musí konečně vzít v úvahu.“

Zemědělský průmysl, který volá po nové povolení po dobu 15 let, kritizoval rozhodnutí Komise. Pracovat namísto obnovené schválení glyfosátu s „technickým příponou“ je pouze dočasným řešením neuspokojivé, domnívá se Zemědělský Industry Association (IVA).

Obě strany si stěžovali, že rozhodnutí Komise EU bylo kvůli tlaku ze strany zájmových skupin. Centrální Horticultural asociace (BMA) mluvil o „politický tlak ze strany zúčastněných stran a nevládních organizací“.

Rozhodnutí Komise EU na obnovu přichází den před uplynutím schválení dne 30. června. Dříve to minul země EU, aby se dohodly i přes opakované snahy o postoj většiny. Minulý pátek, závěrečná hlasování byla bezvýhodná, zdržel v Německu se němž.

Glyfosát je nejpoužívanější herbicid v EU a existuje podezření, že jsou karcinogenní. nicméně, je bezesporu prostředí znečišťujících vliv zapříčiňující zabitím všech možných nežádoucích rostlin, jako je chrpa. Na obrovských polí, která jsou osázené monokultur, celková herbicid ničí základní potraviny a biotopy mnoha zvířat, a tím ohrožuje biodiverzitu.

Vysvětlivky:

Sémantické instrukce (výpustka, deplnění)

Lexikální instrukce

Gramatické instrukce

Stylistické instrukce

Pragmatické kategorie

Zdařilá řešení

V pořádku

GT 13.7.2017 glyfosát				
Instrukce	Jevy	Body	Hodnocení	Celkem
Sémantické	polysémie	-2		64
	homonymie			
	výpustky		11	
	doplnění		2	
	posuny, mylná interpretace		19	
Lexikální	odborné termíny	-1		4
	idiomy			
	jména		4	
	„faux amis“			
	metafory			
	příslovní			
	slovní hříčky			
Gramatické	gramatická správnost	-1	23	34
	slovosled		7	
	interpunkce		4	
Stylistické	stylistické ochuzení	-0,5		0,5
	pestrost slohu		1	
	mísení stylů			
Mimojazykové determinanty	faktor místa	-1		4
	faktor času			
	věcný faktor		4	
Bez chyb	Bez úprav	0		
Velmi zdařilá řešení		0,5	13	6,5
Celkem				100

Tabulka 17: Ukázka bodového hodnocení v rámci translátologické analýzy překladu zprávy o glyfosátu, který byl vyhotoven překladačem GT (13. 7. 2017)

9 Seznam tabulek, grafů a obrázků

Seznam tabulek:

Tabulka 1: Textový korpus sestavený pro studii.....	52
Tabulka 2: Text Merkelová: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu).....	57
Tabulka 3: Text Poradenství: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu).....	58
Tabulka 4: Procentuální hodnocení překladů apelativních textů na stupnici podle Torrense (0–120 %) z hlediska jejich použitelnosti/funkce (čím vyšší procentuální hodnota, tím vyšší kvalita překladu) ...	62
Tabulka 5: Text Glyfosát: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu).....	63
Tabulka 6: Text Monitoring: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu).....	63
Tabulka 7: Procentuální hodnocení překladů informativních textů na stupnici podle Torrense (0–120 %) z hlediska jejich použitelnosti/funkce (čím vyšší procentuální hodnota, tím vyšší kvalita překladu)	67
Tabulka 8: Text Schmeißer: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu).....	68
Tabulka 9: Text Caravaggio: Celkové bodové hodnocení zkoumaných jevů v rámci prvního stupně hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu).....	69
Tabulka 10: Procentuální hodnocení překladů uměleckých textů na stupnici podle Torrense (0–120 %) z hlediska jejich použitelnosti/funkce (čím vyšší procentuální hodnota, tím vyšší kvalita překladu) ...	73
Tabulka 11: Celkové porovnání zprůměrovaných bodových hodnocení získaných v první fázi hodnocení (čím nižší číselná hodnota, tím vyšší kvalita překladu)	74
Tabulka 12: Celkové porovnání zprůměrovaných procentuálních hodnocení získaných v druhé fázi hodnocení (čím vyšší procentuální hodnota, tím vyšší kvalita překladu)	74
Tabulka 13: Porovnání bodového hodnocení BT a GT krátce po spuštění Google Neural Machine Translation, duben 2017 (čím nižší číselná hodnota, tím vyšší kvalita překladu)	75
Tabulka 14: Porovnání bodového hodnocení BT a GT po spuštění Google Neural Machine Translation, červenec 2017 (čím nižší číselná hodnota, tím vyšší kvalita překladu)	75
Tabulka 15: Pokles počtu negativních bodů mezi prvním překladem v červnu 2016 a posledním v červenci 2017 (čím vyšší číslo, tím méně chybných jevů)	76
Tabulka 16: Ukázka bodového hodnocení v rámci translatologické analýzy překladu zprávy o glyfosátu, který byl vyhotoven překladačem BT (13. 7. 2017).....	90
Tabulka 17: Ukázka bodového hodnocení v rámci translatologické analýzy překladu zprávy o glyfosátu, který byl vyhotoven překladačem GT (13. 7. 2017)	92

Seznam grafů:

Graf 1: BT (apelativní texty): Počet chyb v rámci celého zkoumaného období (apel)	58
Graf 2: GT (apelativní texty): Počet chyb v rámci celého zkoumaného období (apel).....	59
Graf 3: BT (apelativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)	59
Graf 4: GT (apelativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)	60
Graf 5: BT (apelativní texty): Počet chyb v rámci celého zkoumaného období (gramatika).....	61
Graf 6: GT (apelativní texty): Počet chyb v rámci celého zkoumaného období (gramatika)	61
Graf 7: BT (informativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)	64
Graf 8: GT (informativní texty): Počet chyb v rámci celého zkoumaného období (sémantika)	64
Graf 9: BT (informativní texty): Počet chyb v rámci celého zkoumaného období (věcný faktor)	65
Graf 10: GT (informativní texty): Počet chyb v rámci celého zkoumaného období (věcný faktor)	65
Graf 11: BT (informativní texty): Počet chyb v rámci celého zkoumaného období (gramatika)	66
Graf 12: GT (informativní texty): Počet chyb v rámci celého zkoumaného období (gramatika)	66
Graf 13: BT (umělecké texty): Počet chyb v rámci celého zkoumaného období (stylistika).....	69
Graf 14: GT (umělecké texty): Počet chyb v rámci celého zkoumaného období (stylistika)	70
Graf 15: BT (umělecké texty): Počet chyb v rámci celého zkoumaného období (sémantika).....	70
Graf 16: GT (umělecké texty): Počet chyb v rámci celého zkoumaného období (sémantika).....	71
Graf 17: BT (umělecké texty): Počet chyb v rámci celého zkoumaného období (gramatika)	71
Graf 18: GT (umělecké texty): Počet chyb v rámci celého zkoumaného období (sémantika).....	72

Seznam obrázků

Obrázek 1: Vauquoisův trojúhelník (GIMÉNEZ, 2008, str. 4).....	20
Obrázek 2: Schéma CBMT (EuroMatrix, 2007, str. 19).....	21
Obrázek 3: Schéma statistického systému (TÝNOVSKÝ, 2007, str. 15).....	23
Obrázek 4: Stupnice pro hodnocení plynulosti a adekvátnosti (KOEHN, 2010, str. 219).....	36
Obrázek 5: Stupnice podle A. Torrense (1994, str. 394).....	51