

JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH
FILOZOFICKÁ FAKULTA
ÚSTAV ROMANISTIKY

BAKALÁŘSKÁ PRÁCE

**HODNOCENÍ AUTOMATICKÉHO PŘEKLADU SMT SYSTÉMŮ (GOOGLE
TRASLATE, BING) Z FRANCOUZŠTINY DO ČEŠTINY: KOLOKACE
Z OBLASTI „BANKOVNICTVÍ“**

Vedoucí práce: doc. PhDr. Jan Radimský, Ph.D.

Autorka práce: Tereza Sekalová

Studijní obor: Francouzský jazyk pro evropský a mezinárodní obchod

Ročník: třetí

2015

Prohlašuji, že svoji bakalářskou práci jsem vypracoval/a samostatně, pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledky obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice květen 2015

.....
Tereza Sekalová

PODĚKOVÁNÍ

Ráda bych poděkovala doc. PhDr. Jan Radimský, Ph.D. za jeho trpělivost, cenné rady a odborné vedení během celé práce.

ANOTACE

Cílem této bakalářské práce je testovat a vyhodnotit úspěšnost překladu vybraných terminologických kolokací z oblasti „bankovníctví“ pomocí volně dostupných překladačů (Google Translate, Bing Translator).

Práce se skládá z nezbytných teoretických poznatků z oblasti strojového překladu a z hodnocení překladů. Teoretické poznatky zahrnují historický vývoj strojového překladu, typy překladů (pravidlový strojový překlad, statistický strojový překlad, hybridní strojový překlad a počítačem podporovaný překlad) a přibližují základní informace o Googlu Translate a Microsoft Bing Translatoru.

Hodnocení překladů bylo vykonáno pomocí analýzy strojově přeložených kolokací a jejich následným porovnáním s dostupnými terminologickými databázemi (zejm. IATE) a s vybranými odbornými slovníky.

ABSTRACT

The aim of this bachelor's thesis is to test and evaluate the translation quality of selected terminological collocations in the area of "banking" using freely available translators (Google Translate, Bing Translator).

The thesis consists of the necessary theoretical knowledge of machine translation and translation evaluation. The theoretical knowledge includes historical development of machine translation, translation types (rule-based machine translation, statistical machine translation, a hybrid machine translation and computer aided translation) and approaches some basic information about Google Translate and Microsoft Bing Translator.

The evaluation of the translations was done by analysing machine-translated collocations and their subsequent comparison with the available terminology databases (esp. IATE) and with specialized dictionaries.

Seznam zkratek

RBMT – pravidlový strojový překlad

SMT – statistický strojový překlad

HMT – hybridní strojový překlad

CAT – počítačem podporovaný překlad

GT – Google Translate

BT – Microsoft Bing Translator

MT – strojový překlad

Obsah

Úvod	9
1. Historie strojového překladu	11
2. Rozdělení systémů strojového překladu	13
A. Způsoby užití strojového překladu	13
B. Pravidlový strojový překlad (RBMT)	14
i. <i>Přímý model překladu (Direct approach)</i>	15
ii. <i>Nepřímý model překladu (Indirect approach)</i>	15
C. Statistický strojový překlad (SMT)	16
i. <i>Word-based models</i>	17
ii. <i>Phrase-based models</i>	17
D. Hybridní strojový překlad (HMT)	17
E. Počítačem podporovaný předklad (CAT)	18
3. Online překladače	20
a) Historie a vývoj online překladačů	20
b) Google Translate	20
c) Microsoft Bing Translator	21
4. Hodnocení strojového překladu	22
5. Evaluace Google Translate	27
I) Souhrn jevů u GT	27
a) <i>Překlad anglický</i>	27
b) <i>Hyperonyma a hyponyma</i>	29
c) <i>Překlad doslovný</i>	30
d) <i>Absence překladu části kolokace</i>	30
e) <i>Zvolení jiného pádu</i>	32
f) <i>Chybná syntax</i>	32
6. Evaluace Microsoft Bing Translatorů	33

I) Souhrn jevů u BT	33
a) Překlad doslovný	33
b) Hyponyma	35
c) Absence překladu části kolokace	36
d) Zvolení jiného pádu	37
e) Chybná syntax	37
f) Překlad anglický	38
Závěr	39
RESUMÉ	41
Seznam odborné literatury a dalších zdrojů	43
Přílohy	45

Úvod

První záznamy o strojovém překladu se objevily již po 2. světové válce (Weaverovo memorandum (Weaver 1955), Georgetown-IBM experiment (Hutchins 2010)) a od té doby se strojový překlad značně rozvinul. Prvotní myšlenka systémů založených na lingvistických pravidlech byla revolučně nahrazena statistickým strojovým překladem, kde se brzy ukázalo, že dostatečné množství paralelních jazykových dat je schopno v relativně krátkém čase překonat stávající pravidlové systémy. V současné době se výzkum strojového překladu silně zaměřuje na takzvaný hybridní překlad, který vznikl kombinací výhod pravidlových a statistických systémů. Vzestup kvality strojového překladu byl doprovázen také rozšířením jeho využití, které proniklo až do obchodní a podnikatelské sféry. Dnes jsou SMT systémy bezplatně dostupné na internetu pro širokou veřejnost.

Předmětem předkládané bakalářské práce s tématem „*Hodnocení automatického překladu SMT systémů (Google Translate, Bing) z francouzštiny do češtiny: kolokace z oblasti bankovníctví*“ je testovat a vyhodnotit úspěšnost překladů těchto výše uvedených překladačů.

Toto téma považuji za velice aktuální, jelikož oba překladače jsou dostupné online a denně pomáhají s překladem až 200 milionům lidí po celé planetě.¹

V prvních třech kapitolách práce si přiblížíme nezbytné teoretické poznatky z oblasti vývoje strojového překladu. První kapitola pojednává o historii strojového překladu a popisuje revoluční vývoj MT. Druhá kapitola rozděluje systémy strojového překladu na pravidlový strojový překlad (RBMT), který je nejstarším existujícím přístupem a je založen na bázi jazykových pravidel. Dalším přístupem je statistický strojový překlad (SMT), který funguje na základě existence vícejazyčných paralelních korpusů, čímž způsobil v oblasti MT revoluci. Třetím přístupem je hybridní strojový překlad (HMT), který kombinuje paralelní korpus s lingvistickými znalostmi. V poslední části této kapitoly si přiblížíme také počítačem podporovaný překlad (CAT), který se postupně stává standardem v poskytování překladatelských služeb a výrazně ulehčuje profesionálním překladatelům práci. Třetí kapitola je věnována online překladačům a to konkrétně Google Translate a Microsoft Bing Translatoru, které budou hrát klíčovou

¹ <http://googletranslate.blogspot.cz/>

roli pro vyhodnocení úspěšnosti překladu. Čtvrtá kapitola stručně popisuje metodologii hodnocení strojového překladu a nastiňuje roli databáze IATE.

V páté kapitole se zaměříme na vyhodnocení jevů, které se vyskytly v průběhu překladu kolokací z francouzštiny do češtiny pomocí překladače Google Translate, stejně jako v následující šesté kapitole, kde se zaměříme na jevy, které se objevily při překladu pomocí Bing Translatoru.

Práce je zakončena závěrem, který shrnuje a porovnává výkony obou výše zmíněných překladačů a doplňuje ji resumé ve francouzském jazyce.

1. Historie strojového překladu

V následující kapitole se budeme věnovat historii strojového překladu a popíšeme si, jak se strojový překlad vyvíjel od svých počátků. Budeme vycházet z publikace *Machine Translation: a concise history* od W. J. Hutchinse (2010)., *Machine Translation: past, present, future* od W. J. Hutchinse (1986) a z *Statistical machine translation* od P. Koehna (2009).

První záznamy o strojovém překladu se objevily již na počátku 50. let v USA. Prvním průkopníkem se stal Warren Weaver, který navrhl využít schopností počítačů také pro účely mezijazykového překladu a představil rovněž možná úskalí, se kterými se tento obor bude potýkat. V roce 1949 vzniká první publikace týkající se strojového překladu s názvem Weaverovo memorandum (Weaver 1955) a záhy je uspořádána i první konference. Počítače, které byly v této době dostupné, byly již samočinné a sloužily zejména ke kryptografickým účelům v době války. Hlavním bodem zájmu byl tehdy rychlý a levný překlad z ruštiny do angličtiny v kontextu studené války, jelikož překladatelé nebyli tehdy lehce dostupní.

Prvním pokusem o realizaci strojového překladu byl Georgetown-IBM experiment, na němž bylo strojově přeloženo zhruba 60 vět z ruštiny do angličtiny. Tento experiment spočíval v překládání textů slovo od slova za případného upravení slovosledu za pomoci naprogramovaných pravidel. Takovýto přístup nazýváme jako přímý překlad (direct translation). Výstup tohoto experimentu byl velice těžkopádný a kostrbatý, nicméně se díky němu strojový překlad stal předmětem dalšího výzkumu a přilákal tak i různé sponzory.

Práce se však nevyvíjely, jak bylo zprvu předpokládáno. Zpráva poradního výboru pro automatické zpracování jazyka (Automatic Language Processing Advisory Committee, ALPAC) z roku 1966 zhodnotila tehdy dostupné systémy pro strojový překlad jako v praxi nepoužitelné. Prvotní úspěch vycházel z faktu, že zadávané věty byly velice jednoduché a postačovalo tedy jen omezené množství jazykových pravidel a slovní zásoby. Výzkum automatického překladu se po vydání této zprávy výrazně omezil, což znamenalo konec rozvoje tohoto oboru přibližně po následujících 10 let.

V 70. letech se v Evropě, Rusku a také Japonsku vývoj strojového překladu nepozastavil, ale pokračoval bez přerušení. V Kanadě se podařilo vytvořit jako jeden z prvních komerčně úspěšných programů pro překlad předpovědí počasí – Meteo. Základem úspěchu byly jednoduché věty, omezená gramatika a slovní zásoba – opakovaly se stále ty samé fráze. Obsahoval už ale i náznaky morfologické a syntaktické analýzy. Tento přístup se nazývá pravidlový (*rule-based*).

Velmi populárním se pak stal také systém Systran, který navázal v USA na předchozí práci a jeho výsledky se využívaly v projektech Apollo, Sojuz a především v Evropském společenství pro překlady anglických textů. Kvalita překladu byla pouze přibližná, nicméně dostačující pro porozumění obsahu zprávy. Ve svých začátcích ho dokonce používal i Google Translator.

Během 80. let se pravidlové (či transferové) systémy se dále zlepšovaly a vyvíjely a kromě morfologické a syntaktické analýzy se začala využívat i sémantická rovina jazyka. Vědci totiž v průběhu svých výzkumů narazili také na skutečnost, že počítače postrádají „skutečnou znalost“ okolního světa. Protože už se nedařilo nijak výrazně zlepšovat výsledky strojového překladu založeného na znalosti jazyka, začaly se využívat první překlady založené na statistických metodách, což bylo podmíněno existencí velkých paralelních vícejazyčných jazykových korpusů. Vznikl tak přístup založený na korpusech, tedy na člověkem přeloženém vzoru (*corpus-based approach*). Přestože byl statistický strojový překlad na počátku svého vývoje, jeho úspěšnost se rovnala téměř 50 %.

V 90. letech se s rozmachem levných osobních počítačů začaly mezi uživateli šířit jednoduché překladové systémy (také prozatím založené na znalosti jazyka) a strojový překlad začal nacházet své využití i ve sféře obchodní a podnikatelské a stal se tak postupem času součástí trhu. S příchodem internetu bylo následně možné zpřístupnit korpusové a statistické systémy široké veřejnosti.

2. Rozdělení systémů strojového překladu

V této kapitole si nejprve přiblížíme způsoby užití strojového překladu, kde vysvětlíme, v jaké oblasti je možné použít určitý typ překladu. Dále popíšeme základní principy pravidlového strojového překladu, který je označován jako nejstarší překladatelský přístup a je založen na principu lingvistických pravidel. Vysvětlíme si také statistický strojový překlad, který funguje na bázi paralelních korpusů. Zabývat se budeme také hybridním strojovým překladem, který kombinuje obě dvě předchozí varianty, a v poslední řadě nastíníme počítačem podporovaný překlad tzv. CAT.

A. Způsoby užití strojového překladu

Od samého začátku se badatelé téměř výhradně soustředili na překlad vědeckých a technických dokumentů, kde potíže s kulturními odlišnostmi a nejrůznějšími spornými kontexty jsou méně pravděpodobné než u literárních textů a různých domén sociologie. Ve vědě a technice byla poptávka po překladech vždy vyšší, než byl počet kvalifikovaných překladatelů v této oblasti a je nutné poznamenat, že tato poptávka neustále roste.

V zásadě existují dva typy poptávek. Zaprvé je to „tradiční“, neboli obvyklá poptávka po překladech se zveřejnitelnou kvalitou (např. vyprodukování vícejazyčných dokumentů pro velké společnosti). V tomto případě může strojový překlad ušetřit mnoho času a samozřejmě i peněz, což je pro zákazníka ideální kombinace. Strojový překlad tedy vyprodukuje jakýsi koncept neboli hrubý náčrt, který je dále upravován profesionálním překladatelem. Tento způsob užití strojových překladů se nazývá „human-aided machine translation“ (HAMT) [automatický překlad s interagujícím překladatelem]. Nicméně velmi často je potřeba text pouze rychle a přibližně přeložit tak, aby byl pro člověka srozumitelný, tedy nemusí nutně být gramaticky a stylisticky správný. Takovýto překlad se nazývá „machine translation for assimilation“ [přípodobňovací strojový překlad]. Oproti tomu „machine translation for dissemination“ [strojový překlad vhodný k šíření] je jeho pravým opakem a je tedy třeba, aby byl vyhotoven v publikovatelné kvalitě. V nedávné době se rozvinulo použití strojového překladu také v internetové komunikaci, kde rovněž není zapotřebí dosáhnout vysoké

kvality překladu. Tento překlad byl nazván jako „machine translation for communication“ [strojový překlad pro komunikaci]. (Hutchins 2010)

Tato oblast zahrnuje také s tím spojený rozvoj tzv. „computer-based translation“ [počítačový překlad]. Tento výstup se zprvu používal jen pro tvorbu elektronické slovníků a glosářů, později přibýlo užití v překladatelských pracovištích a databázích.

B. Pravidlový strojový překlad (RBMT)

Pravidlový strojový překlad (rule-based machine translation) byl prvním systémem, který se užíval pro strojový překlad. Systém aplikuje různá lingvistická pravidla, což mu umožňuje přizpůsobovat slovosled cílového jazyka a také analyzuje kontextovou rovinu textu. Prvním pokusem o realizaci pravidlového strojového překladu byl Georgetown-IBM experiment (1954) a jednou z první společností, která vyvíjela tyto systémy, byl Systran. Metodologie systémů založených na jazykových pravidlech spočívá ve třech fázích: analýza zdrojového textu (sémantická a syntaktická), převod a vygenerování cílového textu. Zdrojový text je tedy rozdělen do segmentů, které jsou následně vyhledány ve slovníku. Takto vznikne hrubý překlad. Dále je třeba vyřešit problém víceznačných segmentů. Systém tedy vybírá nejvhodnější překlad segmentů, které mají více významů. Systém tedy musí vyřešit rozdíly mezi zdrojovým a cílovým jazykem jako je například shoda v rodě nebo čísle. Jedním z hlavních problémů překladu je výběr správného významu slova. Za účelem odstranění tohoto problému je možné využít metodu odstranění dvojznačnosti jednotlivých slov. Systém si tedy automaticky vygeneruje kontext, ve kterém je slovo použito, což umožňuje vybrat správný překladový výraz v cílovém jazyce (Costa-Jussà et al. 2012)

Největším stinnou stránkou systémů založených na lingvistických pravidlech je časová náročnost na vytvoření takového systému. Stejně tak je potřeba velké množství jazykovědných zdrojů. Z toho vyplývá, že ve výsledku je vytvoření takového systému velice finančně nákladné. Pro zlepšení takového systému by bylo potřeba změnit lingvistická pravidla, což opět vyžaduje další jazykovědné znalosti. Mimoto, změna jednoho pravidla ještě nemusí vést k celkovému zlepšení systému a tak se znovu dostáváme k časové náročnosti.

i. Přímý model překladu (Direct approach)

Model tzv. přímého překladu (direct approach) je nejstarším přístupem ke strojovému překladu a je tedy znám jako systém první generace strojového překladu. Přímá metoda překladu tedy spočívá v absenci jakékoliv „prostřední etapy“ v průběhu překladu. Zdrojový text je tedy přímo převeden do cílového jazyka. Je nutné si uvědomit, že počítače, které byly vyvinuty v této době (1950) byly velice primitivní a nebylo tedy možné naprogramovat jakékoliv složitější operace pro překlad. Přímý překlad se tedy nejvíce podobal doslovnému překladu. (Hutchins and Somers 1992)

System tedy nejprve morfologicky analyzoval zdrojový text a následně jej zpracoval v příslušném dvojjazyčném slovníku. Kompletně zde tedy chyběla analýza syntaxe či sémantické stránky textu.

ii. Nepřímý model překladu (Indirect approach)

Neuspokojivé výsledky překladu přímého modelu vedly k vývoji nepřímého modelu překladu (indirect approach), který byl obohacen o jakési mezistádium v průběhu překladu ze zdrojového jazyka do jazyka cílového.

Prvním takovým přístupem byl model interlingvální (interlingua approach). Cílem bylo vytvořit samostatné mezistádium překladu, kde zdrojový text byl převeden do jisté abstraktní formy. V této fázi již nebylo nutné znovu zpracovávat zdrojový text, jelikož k dalšímu procesu překladu se využívala již zmíněná abstraktní forma, která byla tedy zcela nezávislá na zdrojovém textu a která byla dále zpracována v cílový text. Pro přidání nové jazykové kombinace do systému tedy bylo třeba přidat analyzační a generační modul. Hlavním cílem této metody bylo vytvořit abstraktní formu jazyka pro všechny přirozené jazyky, nicméně toho se nikdy docílit nepodařilo.

Druhou nepřímou metodou překladu byl tzv. transferový přístup (transfer approach). Tato metoda opět obsahuje jakýsi mezistupeň překladu, ale v případě transferové metody je tato abstraktní forma jazykově závislá. Bilingvní transfer spočívá v převodu abstraktní formy zdrojového jazyka do abstraktní formy cílového jazyka, která je dále zpracována generačním modulem. Nicméně abstraktní model je specifický pro každý

jazyk a není tedy možné, jako v případě interlingválního modelu, vytvořit jakousi univerzální formu. Další nevýhodou takového systému je fakt, že přidání nového jazyka do systému vyžaduje nejen přidání modulu pro analýzu a generačního modulu, ale také transferové moduly všech jazykových kombinací v systému. (Hutchins and Somers 1992)

C. Statistický strojový překlad (SMT)

Statistický strojový překlad je metoda strojového překladu, která je motivovaná především existencí paralelních korpusů. Za méně než dvě století dokázal statistický strojový překlad zaujmout dominantní postavení na trhu strojových překladů, jelikož už nebylo potřeba velké množství jazykových znalců a následně desítky let práce pro vytvoření překladatelského systému. Vyvinutí statistického strojového překladu tedy znamenalo revoluci v oblasti strojových překladů. SMT systémy zpracovávají překlady přirozeného jazyka tím způsobem, že přezkoumávají mnoho vzorů textů, které již byly přeloženy kvalifikovanými překladateli. Tímto způsobem vzniká algoritmus, na jehož základě systém nadále překládá texty. (Lopez 2008)

Statistický strojový překlad tedy překlad vnímá jako „znalostní problém“ a proto při procesu překladu aplikuje algoritmus na velké množství již přeložených textů. Tyto soubory textů vytvoří tzv. paralelní korpus. Jedná se o soubor textů ve zdrojovém jazyce a jejich překladů do cílového jazyka. Dokumenty obvykle pochází z internetu či z jiných volně dostupných zdrojů. Z těchto textů vznikají dvě tabulky segmentů. První tabulka je vygenerována ve zdrojovém jazyce a druhá tabulka v cílovém jazyce. Tyto segmenty (věty či slova) se spárují mezi sebou, to znamená, že segmenty zdrojového jazyka se spárují s odpovídajícími segmenty jazyka cílového. (Yulian 2014)

Z těchto dat si překladový systém extrahuje slovník, který je obohacený o pravděpodobnosti jednotlivých možných překladů. Nejčastěji se jedná o frázový překlad, kdy překladová tabulka obsahuje nejen překlady jednotlivých slov, ale zejména krátkých vět či slovních spojení. Je tomu tak mimo jiné proto, že slova zdrojového a cílového jazyka si obvykle neodpovídají v poměru 1:1, a překlad celých víceslovných frází proto obvykle vede k úspěšnějšímu výstupu překladu (viz dále zmíněná problematika word-based models).

i. Word-based models

První užití tzv. word-based models (model založený na slovech) sahá do roku 1980 a byl tedy jedním z prvních způsobů statistického překladu. Ačkoliv se dnes již zmíněná metoda nepoužívá, její podstata zůstala zachována. Metoda spočívá v překladu slov jako samostatných jednotek za pomoci statistiky, která zaznamenávala četnost výskytu určitých překladů jednotlivých slov. To znamená, že se v průběhu překladu nebere v potaz zasazení jednotlivých slov do kontextů, což obvykle vede ke změně významu. (Koehn 2009) Hlavním problémem tohoto postupu tedy byl fakt, že většina slov má více významů a je tedy třeba vybírat pro danou situaci vždy ten sémanticky nejvhodnější. Další úskalí tohoto přístupu spočívalo v tom, že jedno slovo zdrojového jazyka se může rovnat i více slovům v jazyce cílovém. Z těchto důvodů byl „model založený na slovech“ neúspěšný.

ii. Phrase-based models

Přístup zvaný phrase-based model (model založený na větách) je dnes nejúspěšnějším překladatelským přístupem vůbec. Jedná se o přístup, který pracuje s tzv. segmenty (věty či slovní spojení). Z těchto segmentů se po spárování vytvoří paralelní korpus, který je v případě phrase-based models obohacován pomocí víceslovných jednotek (segmentů) a tudíž se snižuje riziko ambiguit. (Koehn 2009)

D. Hybridní strojový překlad (HMT)

Hybridní strojový překlad je specifická metoda strojového překladu, která spočívá v kombinování více způsobů strojového překladu do jedné metody. Rozvoj hybridního strojového překladu je způsoben zejména tím, že předchozí techniky překladu nedosahovaly uspokojivých výsledků.

HMT systémy tedy užívají individuální přístup k překladům za užití kombinace pravidlových a statistických systémů za účelem dosažení co nejlepšího výstupu. Tímto přístupem překonává systémy založené na pravidlech, jejichž největším problémem je tzv. překladatelská dvojnáčnost. Pokud má slovo nebo věta více možných významů, statistický přístup může lépe vyhodnotit situaci a najít co nejvhodnější překlad.

Další výhodou statistického přístupu je výrazně větší pokrytí než u pravidlového. Pravidlové systémy naopak obvykle dobře vystihují gramatickou stránku textu, která se obvykle řídí danými pravidly. Statistické systémy se gramatická pravidla snaží pouze odhadovat z dat, bez lingvistických znalostí. Kombinací těchto dvou přístupů vzniká tzv. hybridní překlad.

Mezi nejznámější systémy založené na hybridním způsobu překladu patří např. PROMT, SYSTRAN nebo také široké veřejnosti známý Microsoft Bing Translator.

E. Počítačem podporovaný předklad (CAT)

Počítačem podporovaný překlad, neboli CAT z anglického „computer-aided translation“ se postupně stává standardem v poskytování kvalitních překladatelských služeb. Tato technologie překladu spočívá v opětovném použití již dříve přeložených textů, které se ve specializovaných softwarech ukládají do tzv. překladatelské paměti (translation memory). Pokud se překlad shoduje nebo i pouze podobá již dříve přeloženému překladu, překladatel má možnost předchozí výstupy použít. Tento systém tak zabraňuje pozměňování terminologie v nejrůznějších oborech, jako je např. technika, medicína, právo a mnohé další. Další výhodou může být také cena, jelikož jak vyplývá z předchozího popisu CAT, překladatel nemusí překládat identické věty vícekrát.

Další výhodou je, že překladatelskou paměť si tvoří překladatel sám a není tedy naplněna automaticky. Naplněním této paměti vznikne tzv. terminologická databáze, která obsahuje konkrétní termíny ve zdrojovém a cílovém jazyce. Při vytváření této databáze se často pracuje s ucelenými větami, kolokacemi či jinými víceslovnými soubory, které se nazývají segmenty. (Pollack and Kaufmann 1997)

Hlavní rozdíl mezi strojovým překladem a počítačem podporovaným překladem spočívá v „činiteli překladu“. U strojového překladu se o proces překládání stará sám počítač s tím, že později bude pravděpodobně text upraven kvalifikovanou osobou. Kdežto u systému CAT obstarává překlad překladatel, za užití překladatelské paměti, která slouží ke zvýšení produktivity překladatele a tím pádem i k ulehčení celé práce. Jinými slovy, strojový překlad se snaží lidskou práci zcela nahradit, zatímco překlad podporovaný

počítačem má za úkol tuto práci pouze ulehčit a zefektivnit. (Mathur, Cettolo and Federico 2013)

Mezi nejpoužívanější CAT nástroje patří zejména SDL TRADOS, MemoQ, Déjà Vu či Wordfast.

3. Online překladače

V této kapitole se zaměříme na nejznámější a zároveň nejvíce používané překladače v České republice. Jsou jimi Google Translate od společnosti Google Inc., který je volně dostupný na webové adrese <https://translate.google.cz/>, a Microsoft Bing Translator od firmy Microsoft Corporation, který je rovněž zdarma dostupný na webové adrese <http://www.bing.com/translator>.

a) Historie a vývoj online překladačů

Prvním volně dostupným překladačským systémem byl Babel Fish, který byl uveden do provozu v roce 1997. Díky tomuto systému se stal strojový překlad bezplatným a volně dostupným pro širokou veřejnost. Původně obsahoval deset jazykových kombinací včetně pěti hlavních evropských jazyků. Nejvyhledávanějšími jazykovými dvojicemi byla angličtina – španělština, později také angličtina – francouzština a angličtina – němčina. Za krátkou dobu po spuštění systému Babel Fish vstoupily na trh další obdobné systémy a v roce 2000 bylo v této oblasti zainteresovaných více než deset různých společností jako např. Transparent Language Inc., Alis Technologies, Apex Network a další. V roce 2010 se počet společností zvýšil na 56. (Gaspari and Hutchins 2007)

b) Google Translate

Google Translate je systém provozovaný společností Google Inc² od roku 2006 a je zdarma dostupný na webu své mateřské společnosti pro širokou veřejnost. Tato překladačská služba byla původně založena na systému Systran, který je jednou z nejstarších firem provozujících strojový překlad. Od roku 2007 se Google Translate zakládá na principech statistického strojového překladu, což znamená, že vyhledává při překládání texty, které jsou podobné zdrojovému textu ve velkém množství dokumentů, a vybírá ten nejadekvátnější překlad.

Vedle webového rozhraní podporuje Google Translate také aplikace pro mobilní telefony s operačními systémy iOS a Android. Většina z nynějších 90 podporovaných

² <http://googletranslate.blogspot.cz/>

jazyků pro překlad je pro uživatele k dispozici také formou mluveného slova. Překladač může být použit pro překlad samostatných slov, ale také rozsáhlých textů. Aktuálně Google Translate podporuje 90 jazyků.

c) Microsoft Bing Translator

Microsoft Bing Translator³ je online dostupná překladatelská služba, provozovaná společností Microsoft. Systém byl původně vyvinut pro interní potřeby společnosti Microsoft a sloužil k překladu velkého množství technických textů. Pro širokou veřejnost funguje tento systém od roku 2007. Jedná se o systém založený na principu statistického strojového překladu a zároveň je obohacený o lingvistické prvky. Výstupem je tedy tzv. hybridní překlad, který jsme blíže vysvětlili výše v kapitole Hybridní strojový překlad.

V současné době překladač umožňuje překlad 45 jazyků⁴ a jeho webové rozhraní povoluje překlad textu o maximálně 5000 znacích. Společnost Microsoft zpřístupnila tento překladač také uživatelům mobilních telefonů s operačním systémem Windows Phone formou bezplatné aplikace.

³ <http://blogs.msdn.com/b/translation/>

⁴ <http://www.bing.com/translator/help/#Languages>

4. Hodnocení strojového překladu

Účelem této bakalářské práce je testovat a následně vyhodnotit úspěšnost překladu dvou zdarma dostupných online překladačů – Googlu Translate a Microsoft Bing Translatorem. Úspěšnost budu testovat na 145 kolokacích, které byly zpracovány a přeloženy do češtiny K. Virtovou (2007). Tyto kolokace byly vyexcerpovány z korpusu textů na téma „Bankovníctví – Právo usazování a volný pohyb služeb“ ze systému EUR-Lex a databáze CELEX pomocí programu ParaConc. Hlavním cílem bude vyhodnotit typy chyb a jevů, které se u obou překladačů vyskytují a v druhé řadě nastínit, jak může uživatel ovlivnit správnost překladu zadáním vhodného vstupu.

Pro vyhodnocení kvality překladů použiji odborné specializované slovníky a také databázi IATE, což je mnohojazyčná terminologická databáze EU. V této databázi jsou evidovány termíny z nejrůznějších oborů a oblastí jako je právo, zemědělství, informační technologie a mnoho jiných. Tato databáze byla zpřístupněna veřejnosti v roce 2007 a obsahuje 8,4 milionů termínů, včetně asi 540 000 zkratk a 130 000 slovních spojení, a zahrnuje všech 24 úředních jazyků EU.⁵

⁵ <http://iate.europa.eu/brochure/index.html>

Seznam zkoumaných kolokací

<i>kolokace</i>	<i>český překlad</i>
accord de libre-échange	dohoda o volném obchodu
accord de taux futurs	dohoda o budoucích úrokových sazbách
actif constituant des créances	aktiva představující pohledávky
actif corporel	hmotná aktiva
actifs donnés en garantie	zástava
actif hypothéqué	aktiva zatížená hypotékou
actif incorporel	nehmotná aktiva
action en nullité, en annulation ou en inopposabilité	žaloba na neplatnost, neúčinnost nebo odporovatelnost
actions préférentielles cumulatives	kumulativní prioritní akcie
actions propres	vlastní akcie
adéquation des fonds propres	kapitálová přiměřenost
autorité de contrôle	orgán dohledu
autorité de contrôle des banques	orgán bankovního dohledu
Banque agricole de Grèce	Řecká zemědělská banka
Banque asiatique de développement	Asijská rozvojová banka
banque centrale	centrální banka
Banque centrale européenne	Evropská centrální banka
Banque de développement des Caraïbes	Karibská rozvojová banka
banque d'investissement	investiční banka
Banque européenne d'investissement (BEI)	Evropská investiční banka (EIB)
Banque interaméricaine de développement	Meziamerická rozvojová banka
Banque internationale pour la reconstruction et le développement	Mezinárodní banka pro obnovu a rozvoj
Banque mondiale	Světová banka
banque multilatérale de développement	mezinárodní rozvojová banka
Banque nationale d'investissement pour le développement industriel	Národní investiční banka pro průmyslový rozvoj
Banque nordique d'investissement	Severská investiční banka
bénéfice de l'exercice	zisk za účetní období
blanchiment de capitaux	praní peněz
catégorie d'actions	třída akcií, druh akcií
certificats de dépôt	vkladní list
code d'identification de banque (BIC)	identifikační kód banky
comité européen des contrôleurs des comptes	Evropský výbor orgánů bankovního dohledu
compétence de contrôle	kontrolní pravomoc
comptes annuels	roční účetní závěrka
comptes consolidés	konsolidovaná účetní závěrka
comptes de l'établissement du bénéficiaire	účet instituce příjemce
comptes de profits et pertes	výkaz zisků a ztrát, výsledovka
comptes agrégés	sdužená účetní závěrka
Conseil de l'Union européenne	Rada Evropské unie
contrat à terme de taux d'intérêt	forwardová termínová smlouva
contrat de taux de change à terme	termínovaný (futures) měnový kontrakt

contrats d'échange	swapy
contrat d'option	opce
contrats financiers à terme sur taux d'intérêt	úrokové futures (futures na úrokové sazby)
contrat sur taux de change	devizový (měnový) kontrakt
contrat sur taux d'intérêt	úrokový kontrakt, dohoda o úrokové míře
contrôle des comptes	audit účetních závěrek, audit účetnictví
contrôle des comptes annuels	audit ročních účetních závěrek
contrôle légal	povinný audit
contrôle légal des comptes	povinný audit
contrôle légal des documents comptables	audit účetních dokladů (dokumentů)
contrôle prudentiel	obezřetnostní dohled
contrôleur des comptes	auditor
contrôleur légal (des comptes)	oprávněný auditor, statutární auditor
créance expressément garanties	pohledávka krytá výslovnými zárukami
créance sur la clientèle	pohledávka za klienty
dépôt d'épargne	úsporný vklad
dépôt en liquide	hotovostní vklad
dépôt indisponible	nedisponibilní vklad
dette envers des établissements de crédit	závazek vůči úvěrovým institucím
dette envers la clientèle	závazek vůči klientům
dette publique	veřejný dluh, státní dluh
dette représentée par un titre	závazek z dluhových cenných papírů
document relatif à compte annuel	roční účetní doklad
effet de commerce	obchodní směnka
effet du trésor	pokladní poukázka
effet en portefeuille	směnka držaná v portfoliu
effet public	státní cenný papír
échange de taux d'intérêt	úrokový swap, swap úrokových sazeb
échange de taux d'intérêt variable	bázový swap
échange des actions	směna akcií
échange d'informations	výměna informací
éléments d'actif	aktiva, aktivní položky
éléments de passif	pasiva
endos d'effets réescomptés	indosament eskontovaných směnek
établissement de monnaie électronique	instituce elektronických peněz
exercice comptable	účetní období
exercice de la surveillance	dohled
expert comptable	auditor
fonds de pension	penzijní fondy
fonds de retraite	penzijní fondy
Fonds européen d'investissement	Evropský investiční fond
Fonds monétaire international (FMI)	Mezinárodní měnový fond
fonds propres	vlastní kapitál
fonds propres complémentaires	dodatkový kapitál
fonds reçus en échange de monnaie électronique	peněžní prostředky přijaté výměnou za elektronické peníze

fonds remboursables	splatné peněžní prostředky, splatné prostředky
chèque de voyage	cestovní šek
chèque sur support papier	papírový šek
chèque transfrontalier	přeshraniční šek
immobilisations corporelles	dlouhodobá hmotná aktiva
immobilisations financières	finanční aktiva, dlouhodobá finanční aktiva
instance de contrôle	kontrolní orgán
instrument de monnaie électronique	elektronický peněžní prostředek
instruments dérivés hors bourse	mimoburzovní zúčtování (OTC) obchodovaných derivátových nástrojů, OTC derivátové nástroje
juste valeur	reálná hodnota
monnaie électronique	elektronické peníze
monnaie étrangère	cizí měna
numéro international de compte bancaire (IBAN)	mezinárodní číslo bankovního účtu
obligation de couverture par les fonds propres	povinné kapitálové krytí
obligation en circulation	emitovaný dluhopis
office de chèque postal	poštovní banka, banka poštovního úřadu
opération de bourse à terme	opce
option d'achat	kupní opce
option sur devises achetées	zakoupená měnová opce
option sur taux d'intérêt achetée	zakoupená úroková opce
passif subordonné	podřízená pasiva
perte de l'exercice	ztráta za účetní období
pour compte propre	na vlastní účet
prime d'émission	emisní ážio
protection des dépôts	ochrana vkladů, pojištění vkladů
relevé de compte	výpis z účtu
revenu d'actions	výnos z akcií
succursale d'entreprise d'assurance	dceřinná pojišťovna mateřského podniku, pobočka pojišťovny
succursale d'entreprise d'investissement	pobočka investičního podniku
succursale d'établissement de crédit	pobočka úvěrové instituce
succursale d'établissement de crédit et d'établissement financier	pobočka úvěrové a finanční instituce
système de dépôt centralisé	centrální vkladový systém
système de garantie des dépôts	systém pojištění vkladů
système d'indemnisation des investisseurs	systém pro odškodnění investorů
Système européen de banques centrales	Evropský systém centrálních bank
taux de change	směnný kurz
taux d'inflation	míra inflace
taux d'intérêt	úroková sazba; úroková míra
taux d'intérêt de référence	referenční úroková sazba
titre à revenu fixe	cenný papír s pevným výnosem, cenný papír
titre à revenu variable	cenný papír s proměnlivým výnosem

titre couvert par des créances hypothécaires	cenný papír krytý hypotékou
titre de créance	dluhový nástroj, dluhový cenný papír
titre négociable	obchodovatelný cenný papír, převoditelný cenný papír
transaction sur actions	transakce s akciemi
union douanière	celní unie
union économique et monétaire	Hospodářská a měnová unie
Union européenne	Evropská unie
union monétaire	měnové unie
union régionale	regionální unie
valeur comptable	účetní hodnota
valeur d'acquisition	pořizovací cena
valeur du marché	tržní hodnota
valeur hypothécaire	hypoteční hodnota
valeur immobilisée	dlouhodobé aktivum
valeur mobilière	převoditelný cenný papír, cenný papír
valeur supérieure du marché	vyšší tržní cena
valeur vénale de l'immeuble	tržní hodnota nemovitosti
virement transfrontalier	přeshraniční převod, přeshraniční platba

5. Evaluace Google Translate

V následující kapitole si rozebereme úspěšnost překladů pomocí Google Translate a vyhodnotíme typy jevů, ke kterým dochází v průběhu překladu tímto SMT překladačem.

Při zadávání vybraných kolokací do webového rozhraní GT si můžeme povšimnout, že jsme schopni zjistit, zda GT považuje zadaný text za kolokaci – tedy pokud najedeme kurzorem myši na napsaný text a zobrazí se nám označený vcelku, pak Google Translate vnímá tento text jako kolokaci. Tato funkce se může při evaluaci tohoto překladače ukázat jako užitečná, neboť vnímá-li překladač text jako kolokaci, měl by tento fakt teoreticky zabránit doslovným či přímo nesmyslným překladům.

V druhé řadě je důležité si uvědomit, že ne každá kolokace se bude vyskytovat v jazykových korpusech pouze v singuláru. Z tohoto důvodu jsem zadávala do překladače všech 145 kolokací v druhém sledu také v plurálu. Nepřesnosti lze očekávat také v oblasti flexe, jelikož systém vždy zvolí překlad, který se ve statistice vyskytuje nejvýznamněji a nemusí tedy nutně syntakticky zapadat do větné struktury.

I) Souhrn jevů u GT

V následujících podkapitolách si ukážeme, které překladatelské jevy se vyskytují při používání volně dostupného překladače Google Translate. Tyto jevy jsem zařadila do šesti kategorií, které jsou řazeny sestupně podle četnosti výskytu jednotlivých jevů.

a) Překlad anglický

První zkoumanou kategorií je překlad anglický. Bude se tedy jednat o ty překlady, které GT nepřeložil česky, ale anglicky a to buď jen z části, nebo celé. Tento jev se vyskytl u 7 případů z celkových 145, což činí cca 5%.

kolokace	překlad GT	změna čísla?	chápe text GT jako kolokaci?	databáze IATE	český překlad (Virtová 2009)
accord de taux futurs	Forward Rate Agreement		ANO	dohoda o úrokové míře	dohoda o budoucích úrokových sazbách

Banque agricole de Grèce	Zemědělské Bank of Greece		ANO	x	Řecká zemědělská banka
Banque interaméricaine de développement	Inter-American Development Bank		NE		Meziamerická rozvojová banka
Banque nordique d'investissement	Nordic Investment Bank		NE		Severská investiční banka
code d'identification de banque (BIC)	Bank Identifier Code		NE		identifikační kód banky
endos d'effets réescomptés	zpět z rediscounting		NE	x	indosament eskontovaných směnek
numéro international de compte bancaire (IBAN)	International Bank Account Number		NE	x	mezinárodní číslo bankovního účtu

Google Translate je překladatelský nástroj, který používá jako „mezijazyk“ angličtinu (Boitet and Blanchon 2010), což nesmírně zmenšuje množství požadovaných existujících paralelních korpusů. Kdyby totiž GT překládal pokaždé přímo z jazyka A do jazyka B, potřeboval by pro fungování $90 \cdot 89 / 2 = 4\,005$ jazykových korpusů (bereme-li v potaz že GT v současné době podporuje 90 jazyků), což je nesmírně obsáhlé množství dat. Pokud GT používá angličtinu jako „mezijazyk“, stačí tedy, aby měl v paměti 90 jazykových korpusů, což systémovým programátorům výrazně usnadňuje práci. Tato skutečnost způsobuje mimo jiné i „ztrátu rodu.“ Pokud překládáme např. z češtiny do francouzštiny větu „Jsem učitelka.“ GT ji přeloží jako „Je suis l'enseignant.“ Tímto procesem se nám tedy vytratí ženský rod z původní české věty, jelikož v angličtině existuje pouze slovo „teacher“, které se používá pro maskulinum i femininum.

Pro výše vyobrazenou kategorii anglického překladu z toho vyplývá, že při procesu překládání používá GT angličtinu jako „mezijazyk“ a proto se může stát, že výsledný výstup, který by měl být v češtině, je ve finále anglický. Tento fakt by mohl také vypovídat o tom, že GT nenašel danou kolokaci v cílovém jazyce ve statistické paměti nebo byl tento pojem pokaždé přeložen do češtiny s větším či menším rozdílem.

b) *Hyperonyma a hyponyma*

Následující kategorie zahrnuje takové kolokace, které GT přeložil pomocí hyponyma (slova podřazeného významu) nebo pomocí hyperonyma (slova nadřazeného významu). V této skupině se vyskytuje 6 případů ze 145 (4%).

kolokace FR	překlad GT	změna čísla?	chápe text GT jako kolokaci?	databáze IATE	český překlad (Virtová 2009)
actif corporel	hmotný majetek		ANO	x	hmotná aktiva
effet de commerce	cenný papír		ANO	x	obchodní směnka
éléments de passif	Závazky		x	x	pasiva
immobilisations corporelles	dlouhodobý majetek		ANO	hmotná aktiva	dlouhodobá hmotná aktiva
valeur vénale de l'immeuble	tržní hodnota budovy		NE	x	tržní hodnota nemovitosti
valeurs immobilisées	dlouhodobý majetek		ANO	x	dlouhodobé aktivum

V případě „effet de commerce“ se jedná o zvolení hyperonyma – tedy slova nadřazeného významu, jelikož GT toto slovní spojení přeložil jako „cenný papír“, přičemž správný překlad je „obchodní směnka“, což je jeden z mnoha druhů cenných papírů. V dalších pěti případech zvolil GT hyponymum tedy slovo podřazeného významu. Např. kolokaci „actif corporel“ překládá jako „hmotný majetek“, nikoliv jako „hmotná aktiva“, přičemž majetek je jedním z možných typů aktiv. Je však nutné dodat, že SMT překladače pracují pouze s texty, které byly již dříve přeloženy kvalifikovanými překladateli, kteří vždy pracují s konkrétními kontexty, kde dle vlastního uvážení dosazují příslušné termíny podle souvislostí, které z textu vyplývají.

c) Překlad doslovný

V kategorii doslovného překladu jsou řazeny takové kolokace, které byly přeloženy GT víceméně „slovo od slova“ – každé slovo v kolokaci bylo přeloženo samostatně a nikoliv jako celek. Překlad doslovný se vyskytl u 5 případů z celkových 145 (3,4%).

kolokace	překlad GT	změna čísla?	chápe text GT jako kolokaci?	databáze IATE	český překlad (Virtová 2009)
comptes agrégés	souhrnné účty		NE	x	sdružená účetní závěrka
contrat sur taux de change	smlouvy o směnných kurzech	ANO	NE	x	devizový (měnový) kontrakt
contrat sur taux d'intérêt	smlouvy o úrokových sazbách	ANO	NE	dohoda o úrokové míře	úrokový kontrakt
document relatif à compte annuel	doklady týkající se ročních účetních závěrek		NE	x	roční účetní doklad
obligation de couverture par les fonds propres	povinnost pokrytí vlastního kapitálu		NE	x	povinné kapitálové krytí

Zde můžeme sledovat, že jednotlivá slova v kolokaci jsou přeložena víceméně samostatně, čemuž také napovídá fakt, že Google Translate těchto 5 případů nevnímá jako kolokaci. Např. „*obligation de couverture par les fonds propres*“ GT přeložil jako „*povinnost pokrytí vlastního kapitálu*“ – zde je patrné, že překladač přeložil každé slovo zvlášť, ale stejně jako ve zbylých případech to nebrání v porozumění překladu a jedná se spíše o odchýlení od odborné terminologie.

d) Absence překladu části kolokace

Pod tuto kategorii spadají takové překlady, které jsou neúplné, neboť je na první pohled zřejmé, že překladač přeložil jen část zadané kolokace. Tento jev se vyskytl u 5 případů ze 145 (3,4%).

kolokace	překlad GT	změna čísla?	chápe text GT jako kolokaci?	databáze IATE	český překlad (Virtová 2009)
autorité de contrôle	vedoucí		x	donucovací orgán	orgán dohledu
expert comptable	účetní		x	x	auditor
instrument de monnaie électronique	elektronický peněžní		NE	x	elektronický peněžní prostředek
prime d'émission	pojistné		x	emisní ážio	emisní ážio
titre couvert par des créances hypothécaires	na které se vztahuje hypotéky		NE	cenný papír krytý hypotékou	cenný papír krytý hypotékou

V každém z výše zobrazených případů tento jev posouvá význam přeložené kolokace. Např. „expert comptable“ není pouze „účetní“, ale „auditor“ nebo také „účetní znalec“ (Pohorská 2002). Stav překladu nám napovídá, že překladač přeložil pouze slovo „comptable“, ale nikoliv už slovo „expert“. Takovýto překlad lze považovat za chybný, jelikož profese účetního znalce je poněkud odlišná od činnosti účetního, což ve finále vede ke změně významu celé kolokace. Stejně tak je tomu u kolokace „autorité de contrôle“ jejíž správný překlad je „orgán dohledu“, ale GT ji přeložil jako „vedoucí“. Z toho vyplývá, že překladač přeložil pouze slovo „autorité“ jako „vedoucí“ a druhou část kolokace „de contrôle“ zcela vynechal, čímž se opět význam kolokace mění, jelikož „vedoucí“ je obecné pojmenování pro osobu, která jistým způsobem řídí a organizuje např. chod firmy. Oproti tomu orgán dohledu je úřad, který vykonává dohled nad finančním trhem.⁶ Obdobným případem je slovní spojení „prime d'émission“, které GT přeložil jako „pojistné“, čemuž by odpovídal francouzský termín „prime“. Nicméně „prime d'émission“ je definováno jako „emisní ážio“ (Pohorská 2002), což je rozdíl mezi emisním kurzem a jmenovitou hodnotou akcií.⁷ V posledních dvou případech – tedy „instrument de monnaie électronique“ a „titre couvert par des créances hypothécaires“ jde o jasnou absenci přeložení substantiva v hlavě kolokace. GT tedy nepřeložil pojmy „instrument“ a „titre“, z čehož vyplývá, že se kolokace stává nesrozumitelnou.

⁶ <http://business.center.cz/business/pojmy/p2908-organ-dohledu.aspx>

⁷ <http://business.center.cz/business/pojmy/p90-emisni-azio.aspx>

e) *Zvolení jiného pádu*

Tato třída obsahuje takové kolokace, které GT přeložil v jiném mluvnickém pádu než v nominativu. Tento jev se vyskytl ve 3 případech ze 145 (2%)

kolokace FR	překlad GT	změna čísla?	chápe text GT jako kolokaci?	databáze IATE	český překlad (Virtová 2009)
banque multilatérale de développement	mezinárodní rozvojovou bankou		ANO	x	mezinárodní rozvojová banka
instance de contrôle	kontrolními orgány	ANO	ANO	x	kontrolní orgán
système d'indemnisation des investisseurs	o systémech pro odškodnění investorů	ANO	ANO	system pro odškodnění investorů	system pro odškodnění investorů

Tento jev je zapříčiněn tím, že se daná kolokace vyskytuje ve výchozí statistice nejhojněji v jiném než prvním pádu a proto by nás takováto forma překladu neměla překvapit.

f) *Chybná syntax*

V této kategorii jsou zastoupeny kolokace, jejichž slovosled je přeházený, a tudíž je syntakticky nesprávný. Tento jev se vyskytl u 2 kolokací ze 145 (1,4%).

kolokace FR	překlad GT	změna čísla?	chápe text GT jako kolokaci?	databáze IATE	český překlad (Virtová 2009)
dette représentée par un titre	Dluh je zastoupeno cennými papíry		NE	x	závazek z dluhových cenných papírů
options sur taux d'intérêt achetée	Možnost úroková sazba zakoupili		NE	x	zakoupená úroková opce

Chybná větná skladba se vyskytla ve dvou případech. Tento jev by se dal přirovnat k doslovnému překladu, kdy překladač přeloží jednotlivá slova samostatně, čímž je porušen vzájemný syntaktický vztah větných členů. Napovídá tomu také fakt, že GT nepovažuje tato slovní spojení za kolokace.

6. Evaluace Microsoft Bing Translatoru

V následující kapitole si rozebereme úspěšnost překladů pomocí Microsoft Bing Translatoru a vyhodnotíme typy jevů, ke kterým dochází v průběhu překladu tímto SMT překladačem.

Na rozdíl od Google Translatoru nelze u BT rozpoznat, zda vnímá zadaný text jako kolokaci, což nám zabraňuje vytvořit si hypotézu o tom, zda je překlad přijatelný či ne. Další rozdíl spočívá v tom, že Bing Translator funguje na základě hybridního strojového překladu, což znamená, že kombinuje prvky statistického strojového překladu s prvky pravidlového statistického překladu. Teoreticky by tedy BT měl dosahovat kvalitnějších překladů než GT, jelikož obsahuje i naprogramovaný lingvistický modul.

U Bing Translatoru se ukázalo jako bezpředmětné zadávání kolokací také v množném čísle, pro případné dosažení kvalitnějšího výstupu překladu. Tento fakt lze připisovat zásahu lingvistického modulu, který přeloží kolokaci do plurálu, aniž by měnil strukturu kolokace.

I) Souhrn jevů u BT

V následujících kategoriích se zaměříme na překladatelské jevy, které se vyskytují při používání volně dostupného překladače Microsoft Bing Translatoru. Tyto jevy jsem zařadila do šesti kategorií, které jsou řazeny sestupně podle četnosti výskytu jednotlivých jevů.

a) Překlad doslovný

V kategorii doslovného překladu jsou řazeny takové kolokace, které BT přeložil víceméně „slovo od slova“, jelikož doslovný překlad nedbá na přirozené gramatické struktury a zabývá se každým slovem zvlášť, bez ohledu na kontext. Překlad doslovný se vyskytl u 20 případů z celkových 145 (13,8%) a tvoří tedy nejobsáhlejší kategorii vysledovaných jevů vůbec.

kolokace	překlad Bing	databáze IATE	překlad (Virtová 2007)
action en nullité, en annulation ou en inopposabilité	akce v neplatnosti, odporovatelnosti nebo neúčinnosti	x	žaloba na neplatnost, neúčinnost nebo odporovatelnost
banque multilatérale de développement	multilaterální rozvojové banky	x	mezinárodní rozvojová banka
certificats de dépôt	depozitní certifikát	vkladní list	vkladní list
comptes agrégés	souhrnné účty	x	sdužená účetní závěrka
comptes de profits et pertes	účty zisků a ztrát	x	výkaz zisků a ztrát, výsledovka
contrat d'option	opční smlouva	x	opce
contrat sur taux de change	Smlouva o směnný kurz	x	devizový (měnový) kontrakt
contrôle des comptes	řízení účtů	povinný audit	audit účetních závěrek, audit účetnictví
contrôle légal	právní kontrola	zákonný audit, povinný audit	povinný audit
contrôle légal des comptes	právní kontrola účtů	povinný audit	povinný audit
dette représentée par un titre	dluh prokázaný potvrzením	x	závazek z dluhových cenných papírů
document relatif à compte annuel	dokument týkající se roční účetní závěrky	x	roční účetní doklad
effet de commerce	vliv obchodu	x	obchodní směnka
effet en portefeuille	efekt v portfoliu	x	směnka držaná v portfoliu
effet public	veřejná efekt	x	státní cenný papír
instance de contrôle	kontrolní instance	x	kontrolní orgán
obligation en circulation	povinnost zbývajících	x	emitovaný dluhopis
option d'achat	možnost koupě	x	kupní opce
titre couvert par des créances hypothécaires	název se vztahuje hypotéky	cenný papír krytý hypotékou	cenný papír krytý hypotékou

Stejně jako v kapitole 5.1.2 *Překlad doslovný* můžeme i zde sledovat, že BT přeložil jednotlivé části kolokací nezávisle na sobě, což v několika případech vede ke změně významu daného víceslovného pojmenování. Jedná se zejména o kolokace s hlavou „effet“ – tedy „effet de commerce“ „effet en portefeuille“ a „effet public“, kdy Bing Translator překládá slovo „effet“ jako „vliv“ či „efekt“, což je jeden z jeho možných významů. Stejný problém se dále vyskytuje i u kolokací „obligation en circulation“, „option d'achat“ a „titre couvert par des créances hypothécaires.“ BT zde opět zvolil nejznámější význam substantiv, které tvoří hlavu kolokace. „Obligation“ lze jistě přeložit jako „povinnost“, avšak v obchodní francouzštině má tento výraz odborný

překlad „dluhopis“. V případě „option d’achat“ BT překládá slovo „option“ jako „možnost“, ale správným překladem je odborný termín „kupní opce“, která je definována jako „právo koupit nebo prodat konkrétní množství aktiv (akcií, deviz, obligací) za stanovenou opční cenu.“⁸ Dalším nesprávně disambiguovaným pojmem je „action en nullité, en annulation ou en inopposabilité“, kterou BT přeložil jako „akce v neplatnosti, odporovatelnosti nebo neúčinnosti“. Zde je opět patrný doslovný překlad slova „action“, které BT přeložil jako „akce“ a nikoliv jako „žaloba.“ Posledním takovým případem je „titre couvert par des créances hypothécaires“, který BT přeložil jako „název se vztahuje hypotéky“. Z tohoto výstupu je znovu patrný chybný překlad hlavy kolokace, kdy je substantivum „titre“ přeloženo jako „název“ a nikoliv jako „cenný papír.“ V druhé řadě se zde vyskytuje také chybně zvolená syntax.

Ve zbývajících případech se jedná podobně jako u Google Translate o odchýlení od odborné terminologie, které však nijak významně neovlivňuje celkové porozumění textu.

b) Hyponyma

Následující kategorie zahrnuje takové kolokace, které BT přeložil pomocí hyponyma, tedy pomocí slova podřazeného významu. V této skupině se vyskytuje 6 případů ze 145 (4%).

kolokace	překlad Bing	databáze IATE	překlad (Virtová 2007)
actif corporel	dlouhodobý hmotný majetek	x	hmotná aktiva
actif hypothéqué	Majetek zastavený	x	aktiva zatížená hypotékou
actif incorporel	Dlouhodobý nehmotný majetek	x	nehmotná aktiva
éléments de passif	závazky	x	pasiva
immobilisations corporelles	hmotného investičního majetku	hmotná aktiva	dlouhodobá hmotná aktiva
immobilisations financières	Dlouhodobý finanční majetek	x	finanční aktiva, dlouhodobá finanční aktiva

V této kategorii můžeme sledovat, podobně jako u Google Translate, použití podřazeného výrazu. Dominující jsou kolokace od hlavy „actif“ a „immobilisations“, které jsou v odborné literatuře překládány jako „aktiva.“ BT překládá ve všech

⁸ <http://business.center.cz/business/pojmy/p396-opce.aspx>

případech tyto pojmy jako „majetek“, který tvoří součást aktiv.⁹ Co se týče případu kolokace „éléments de passif“, BT ji překládá jako „závazky“, které opět tvoří složku již zmíněných pasiv.

Stejně jako u GT je nutné dodat, že použití hyponym i hyperonym je u lidských překladatelů běžným postupem, jelikož tito odborníci pracují vždy s konkrétními větnými souvislostmi.

c) Absence překladu části kolokace

V následující kategorii se budeme zabývat překlady, které jsou neúplné, neboť je evidentní, že Bing Translator přeložil jen část zadané kolokace. Takovýto překlad se objevil u 4 případů ze 145 (2,8%).

kolokace	překlad Bing	databáze IATE	překlad (Virtová 2007)
expert comptable	účetní	x	auditor
systeme de garantie des dépôts	system	system pojištění vkladů	system pojištění vkladů
titre à revenu fixe	pevné příjem	x	cenný papír s pevným výnosem, cenný papír
titre à revenu variable	variabilní výnos	x	cenný papír s proměnlivým výnosem

Stejně jako Google Translate si ani BT neporadil s překladem kolokace „expert comptable“ a přeložil ji jako „účetní.“ Jedná se tedy o stejný případ, jako v kapitole 5.1.3 *Chybí překlad jednoho slova v kolokaci.* V druhé řadě se jedná o kolokaci „systeme de garantie des dépôts“, který BT přeložil pouze nespecifikovaným výrazem „system.“ Zde je patrná absence překladu druhé části kolokace – tedy „pojištění vkladů.“ Dále se zde objevily dvě kolokace od hlavy „titre“ a to „titre à revenu fixe“ a „titre à revenu variable“), které byly pomocí BT přeloženy jako „pevné příjem“ a „variabilní výnos.“ Překladač tedy zcela vynechal pojem „cenný papír“, bez kterého je kolokace nesrozumitelná.

⁹ <http://business.center.cz/business/pojmy/p1294-aktiva.aspx>

d) Zvolení jiného pádu

Tato kategorie zahrnuje takové kolokace, které BT přeložil v jiném než prvním pádě. Tento jev byl zaznamenán u 5 případů ze 145 (3,4%).

kolokace	překlad Bing	databáze IATE	překlad (Virtová 2007)
banque multilatérale de développement	multilaterální rozvojové banky	x	mezinárodní rozvojová banka
fonds remboursables	splatných peněžních prostředků	x	splatné peněžní prostředky, splatné prostředky
immobilisations corporelles	hmotného investičního majetku	hmotná aktiva	dlouhodobá hmotná aktiva
option sur devises achetées	zakoupené měnové opce	x	zakoupená měnová opce
virement transfrontalier	přeshraniční bezhotovostní platby	x	přeshraniční převod, přeshraniční platba

Tento jev je stejně jako u Google Translatoru zapříčiněn tím, že se daná kolokace vyskytuje ve výchozí statistice nejčastěji v jiném než prvním pádu a proto bychom opět neměli považovat takový překlad nutně za chybný. U všech v tabulce zapsaných kolokací můžeme sledovat, že Bing Translator použil druhý mluvnický pád (genitiv).

e) Chybná syntax

Další podkapitolu tvoří skupina kolokací, kterou BT přeložil syntakticky chybně, což znamená, že slovosled jednotlivých částí daného slovního spojení je nesprávný nebo přímo nesmyslný. Tento jev se vyskytl u 4 případů ze 145 (2,8%).

kolokace	překlad Bing	databáze IATE	překlad (Virtová 2007)
comptes de l'établissement du bénéficiaire	účet příjemce, je instituce	x	účet instituce příjemce
instrument de monnaie électronique	nástroj elektronické peníze	x	elektronický peněžní prostředek
option sur taux d'intérêt achetée	Úroková opce nakoupené	x	zakoupená úroková opce
succursale d'établissement de crédit et d'établissement financier	úvěrové a finanční instituce pobočky	x	pobočka úvěrové a finanční instituce

Chyba ve větné skladbě je pro lidské překladače neznámá, zejména pokud je překladačem rodilý mluvčí. V případě použití volně dostupných překladačů však dochází ke zpřeházení větných členů, čímž se nevyhnutelně poruší jejich syntaktické vztahy a výsledný výstup je pak nesrozumitelný. Překvapivé je, že tento jev se vyskytl čteněji u BT, kde bychom teoreticky mohli očekávat zásah lingvistického modulu, než u Google Translatoru.

f) Překlad anglický

Poslední podkapitolou je překlad anglický, který se vyskytl pouze ve dvou případech ze 145 (1,4%).

kolokace	překlad Bing	databáze IATE	překlad (Virtová 2007)
Banque agricole de Grèce	Zemědělské Bank of Greece	x	Řecká zemědělská banka
chèque de voyage	Traveller's šek	x	cestovní šek

Tento jev se vyskytl ve větší míře u GT, který však ve většině případů ponechal celou kolokaci v angličtině. Bing Translator se snažil přeložit alespoň část slovního spojení.

Závěr

Hlavním cílem této bakalářské práce bylo testovat a vyhodnotit úspěšnost překladů vybraných terminologických kolokací z oblasti „bankovníctví“ pomocí volně dostupných překladačů (Google Translate, Bing Translator). Abychom tento záměr splnili, vybrali jsme 145 kolokací z tematického okruhu „bankovníctví“, na kterých byla provedena analýza. Pro zvýšení přehlednosti jsme hodnotili výkony obou překladačů zvlášť.

Práce je rozdělena do šesti kapitol. V první kapitole, která je věnována historii strojového překladu, je podrobně popsána jeho evoluce od konce 2. světové války, kdy se strojový překlad revolučně přesunul od pravidlově založených systémů (RBMT) ke statistické analýze (SMT) a využívání paralelních jazykových korpusů, které tvoří již dříve přeložené texty kvalifikovanými překladateli.

Druhá kapitola pojednává o rozdělení systémů strojového překladu a je dále rozdělena do pěti podkapitol. V první podkapitole se dozvídáme o různých způsobech a doménách využití strojového překladu. Následující, tedy druhá podkapitola, popisuje pravidlový strojový překlad (rule-based machine translation), který byl prvním přístupem ke strojovému překladu a je založen na bázi lingvistických pravidel a analyzuje tedy zdrojový text pouze z jazykovědného pohledu. Ve třetí kapitole je charakterizován statistický strojový překlad (statistical machine translation), který se zakládá na existenci rozsáhlých paralelních jazykových korpusů a jehož vyvinutí znamenalo revoluci ve světě MT. Čtvrtá podkapitola je věnována hybridnímu strojovému překladu (hybrid machine translation), který kombinuje obě předchozí metody strojového překladu. Pátá podkapitola popisuje počítačem podporovaný překlad (computer-aided translation), který spočívá v možnosti využití tzv. překladatelské paměti, která překladateli pomáhá ušetřit čas.

Ve třetí kapitole věnované online překladačům, se dozvídáme nejprve o historii a vývoji volně dostupných překladačů a následně zde nalezneme veřejně dostupné informace o dvou nejnámějších a nejpoužívanějších online překladačích – Google Translate a Microsoft Bing Translatoru.

Čtvrtá kapitola stručně pojednává o metodologii hodnocení strojového překladu a podává obecné informace o databázi IATE. Nalezneme zde také kompletní výčet zkoumaných kolokací v podobě tabulky.

Pátá a šestá kapitola jsou věnovány evaluaci online překladačů Google Translate a Microsoft Bing Translatoru. Z celkových výsledků analýzy (viz Přílohy) je zřejmé, že Google Translate byl o něco úspěšnější, jelikož přeložil korektně 63,4% ze 145 zkoumaných kolokací což je o 9,6% více než Bing Translator, jehož úspěšnost byla 53,8%. Co se týče jevů, které se v průběhu analýzy vyskytly, jejich povaha byla u obou překladačů téměř totožná. U Google Translate jsme se setkali nejhojněji s překladem anglickým, který byl zapříčiněn tím, že GT používá angličtinu jako mezijazyk za účelem zmenšení požadovaného množství paralelních korpusů. U Bing Translatoru jsme se s tímto jevem setkali pouze ve dvou případech. Dalším analyzovaným jevem bylo zvolení hyperonyma či hyponyma. Tento jev vypovídá o tom, že SMT systémy čerpají z korpusu textů, které již byly již dříve přeloženy kvalifikovanými překladateli, kteří vždy pracují s určitými kontexty, podle nichž následně volí nejvhodnější překlad. Dalším studovaným jevem byl překlad doslovný, který se vyznačoval samostatným překladem jednotlivých slov v kolokaci. Tento problém se vyskytl výrazně více u Bing Translatoru než u Google Translate, který se projevil lepší znalostí odborných termínů. Následujícím jevem byla absence překladu části kolokace, kde oba překladače přeložily pouze část dané kolokace a došlo tak k významovému posunu. Dále jsme se setkali se zvolením jiného pádu než nominativu, což způsobuje fakt, že daná kolokace se pravděpodobně vyskytuje ve statistikách nejhojněji v jiném pádu a proto by nás takový překlad opět neměl překvapit. Posledním analyzovaným jevem byla chybná syntax, která se projevila přeházeným až nesmyslným slovosledem. Překvapivé je, že tento jev způsobil těžkosti spíše Bing Translatoru, kde by se dal očekávat zásah lingvistického modulu.

RESUMÉ

Le présent mémoire de licence a pour but principal d'évaluer et de tester la qualité de la traduction des collocations terminologiques sélectionnés dans le domaine de la "sphère bancaire" à l'aide des traducteurs disponibles gratuitement (Google Traduction, Bing Translator).

Tout d'abord, le premier chapitre qui est consacré à l'histoire de la traduction automatique, illustre en détail son évolution depuis la fin de la Deuxième Guerre mondiale, la première expérience pratique de la traduction automatique (Georgetown-IBM), l'apparition de la traduction automatique à base de règles et consécutivement l'invention de la traduction automatique statistique.

Le deuxième chapitre est consacré aux différents procédés de la traduction automatique. D'abord il semble nécessaire de décrire les principales possibilités d'utiliser la traduction automatique, ensuite introduire le fonctionnement de la traduction automatique à base de règles et ses trois principales approches (approche directe, approche interlangue, approche par transfert). Par la suite, en présentant la traduction automatique statistique, apprendre d'utilisation des modèles construits à partir de corpus bilingues. Après en développant la traduction automatique hybride, il se montre que cette approche combine les corpus bilingues avec l'intervention des règles linguistiques. Enfin, en mentionnant la traduction assistée par ordinateur il se trouve que cette méthode est fréquemment utilisée par les sociétés de traduction, grâce à la possibilité de créer leur propre mémoire de traduction à l'aide des logiciels spécialisés.

Le troisième chapitre traite des informations générales sur les traducteurs disponibles gratuitement (Google Traduction, Bing Translator). D'abord nous allons nous concentrer sur l'histoire des traducteurs gratuitement accessibles et ensuite développer les informations fondamentaux sur ces deux traducteurs.

En ce qui concerne le quatrième chapitre, il traite de la méthodologie de la recherche de cette mémoire de licence et contient la liste des 145 collocations étudiées.

Le cinquième chapitre de cette mémoire de licence introduit les différents phénomènes qui se sont produits au cours du processus de la traduction à l'aide du traducteur Google Traduction. Parmi ces phénomènes nous pouvons compter notamment la traduction anglaise, la traduction littérale, l'absence de traduction d'une partie de la collocation, l'hyperonymie et l'hyponymie, la problématique des cas grammaticaux et les erreurs dans le domaine de la syntaxe. Le même processus a été fait dans le cas de Bing Translator dans le sixième chapitre de la présente mémoire de licence.

Seznam odborné literatury a dalších zdrojů

Odborné publikace

Bowker, L., 2002, Computer-aided Translation Technology: A Practical Introduction, University of Ottawa Press. 185 s. ISBN 0-7766-3016-4

<

http://books.google.cz/books?id=ly29mc6dO0C&pg=PA3&hl=cs&source=gbs_toc_r&cad=3#v=onepage&q&f=false>

Boitet, C.; Blanchon, H.; Seligman, M.; Bellyncq, V., 2010, MT on and for the Web GETALP, UPMF, Grenoble, France, ISBN: 978-1-4244-6896-6

<<http://www-clips.imag.fr/geta/herve.blanchon/Pdfs/NLP-KE-10.pdf>>

Costa-Jussà, M.R., M. Farrús, J.B. Mariño, J.A.R. Fonollosa 2012. "Study and comparison of rule-based and statistical Catalan-Spanish machine translation systems". Computing and Informatics, volume 31, issue 2, pages 245--270. ISSN 1335-9150.

<<http://www.cai.sk/ojs/index.php/cai/article/view/940/430>>

Gaspari, F., and Hutchins, J., 2007, Online and free! Ten years of online machine translation: origins, developments, current use and future prospects. In: *Proceedings of the Machine Translation Summit XI: September 10-14, 2007, Copenhagen, Denmark*. [Allschwil]: EAMT (The European Association for Machine Translation), s. 199-206. ISBN 9788790708160 8790708164.

<www.hutchinsweb.me.uk/MTS-2007.pdf>

Hutchins, J., and Somers, H., 1992. *An Introduction to Machine Translation*. London: Academic Press.

<<http://www.hutchinsweb.me.uk/IntroMT-TOC.htm>>

Hutchins, John. 2007. *Machine Translation: A concise history. Computer aided translation: Theory and practice*. China: Chinese University of Hong Kong.

<<http://www.hutchinsweb.me.uk/CUHK-2006.pdf>>

Koehn, P., 2010, *Statistical Machine Translation*. 1st ed. Cambridge New York: Cambridge University Press. 446 s. ISBN-13 978-0-511-69132-4.

Lopez, A. 2008. Statistical machine translation. *ACM Comput. Surv.*, 40, 3, Article 8 (August 2008), 49 pages
<<https://alopez.github.io/papers/survey.pdf>>

Mathur, P., Cettolo, M., Federico, M., 2013, Online learning approaches in Computer Assisted Translation. *ACL Workshop on Machine Translation (WMT)*, Sofia, Bulgaria
<www.statmt.org/wmt13/pdf/WMT37.pdf>

Microsoft Corporation. *About Bing Translator*. [Online]. ©2014.
<<http://www.bing.com/translator/help/#Languages>>

Pollack, Martha E, Kaufmann, M., 1997, *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*, Nagoya, Japan, ISBN:1-555860-480-4
<<http://ijcai.org/Past%20Proceedings/IJCAI-97-VOL2/PDF/026A.pdf>>

Weaver, W. 1955 "Translation (1949)" In *Machine Translation of Languages*. MIT Press, Cambridge, MA.

Yulian, H. 2014 Statistical machine translation based on translation rules // *Journal of Chemical & Pharmaceutical Research*;2014, Vol. 6 Issue 7, p1628
<<http://jocpr.com/vol6-iss7-2014/JCPR-2014-6-7-1628-1635.pdf>>

Slovníky

Pohorská R., Vlasák V., *Česko-francouzský a francouzsko-český finanční slovník*, Vysoká škola ekonomická v Praze, Praha, 1997

Pohorská R., *Francouzsko-český finanční slovník*, FRAUS, Praha 2002

Vlasák V., Lyer S. *Česko-francouzský slovník (R-Ž)*, Státní pedagogické nakladatelství, Praha, 1993

Webové stránky

http://iate.europa.eu/about_IATE.html

<http://googletranslate.blogspot.cz/>

<http://blogs.msdn.com/b/translation/>

<http://www.bing.com/translator/help/#Languages>

<http://iate.europa.eu/brochure/index.html>

<http://business.center.cz/business/pojmy/p1294-aktiva.aspx>

<http://business.center.cz/business/pojmy/p396-opce.aspx>

<http://business.center.cz/business/pojmy/p2908-organ-dohledu.aspx>

<http://business.center.cz/business/pojmy/p90-emisni-azio.aspx>

Přílohy

Seznam správně přeložených kolokací pomocí Google Translatoru

kolokace	překlad GT	změna čísla?	chápe text GT jako kolokaci?	databáze IATE	český překlad (Virtová 2007)
accord de libre-échange	dohoda o volném obchodu		ANO	dohoda o volném obchodu	dohoda o volném obchodu
actif constituant des créances	aktiva představující pohledávky		ANO	x	aktiva představující pohledávky
actif incorporel	nehmotné aktivum		ANO	x	nehmotné aktivum
action propre	vlastní akcie	ANO	ANO	x	vlastní akcie
actions préférentielles cumulatives	kumulativní prioritní akcie		ANO	x	kumulativní prioritní akcie
adéquation des fonds propres	kapitálová přiměřenost		ANO	kapitálové požadavky	kapitálová přiměřenost
autorité de contrôle des banques	orgán bankovního dohledu		ANO	x	orgán bankovního dohledu
Banque asiatique de développement	Asijská rozvojová banka		ANO	Asijská rozvojová banka	Asijská rozvojová banka
banque centrale	centrální banka		ANO	centrální banka	centrální banka
Banque centrale européenne	Evropská centrální banka		ANO	Evropská centrální banka	Evropská centrální banka
Banque de développement des Caraïbes	Karibská rozvojová banka		NE	Karibská rozvojová banka	Karibská rozvojová banka
banque d'investissement	investiční banka		ANO	investiční banka	investiční banka
Banque européenne d'investissement (BEI)	Evropská investiční banka		NE	Evropská investiční banka (EIB)	Evropská investiční banka (EIB)
Banque internationale pour la reconstruction et le développement	Mezinárodní banka pro obnovu a rozvoj		NE	Mezinárodní banka pro obnovu a rozvoj	Mezinárodní banka pro obnovu a rozvoj
Banque mondiale	Světová banka		ANO	Světová banka	Světová banka
Banque nationale d'investissement pour le développement industriel	Národní investiční banka pro průmyslový rozvoj		NE	x	Národní investiční banka pro průmyslový rozvoj

blanchiment de capitaux	praní špinavých peněz		ANO	praní peněz	praní peněz
catégorie d'actions	třída akcií		NE	druh akcií	třída akcií, druh akcií
certificat de dépôt	vkladové certifikáty	ANO	ANO	vkladní list	vkladní list, depozitní certifikát
compétence de contrôle	přezkoušení odborné způsobilosti		ANO	přezkoušení odborné způsobilosti	přezkoušení odborné způsobilosti, kontrolní pravomoc
compte annuel	roční účetní závěrka	ANO	ANO	účetní závěrka	roční účetní závěrka
compte consolidé	konsolidovaná účetní závěrka	ANO	ANO	x	konsolidovaná účetní závěrka
compte de l'établissement du bénéficiaire	účty instituce příjemce	ANO	NE	x	účet instituce příjemce
comptes de profits et pertes	výkaz zisků a ztrát		ANO	x	výkaz zisků a ztrát, výsledovka
Conseil de l'Union européenne	Rada Evropské unie		NE	Rada Evropské unie	Rada Evropské unie
contrat d'échange	swapy	ANO	x	swap	swap
contrat d'option	opční smlouva		ANO	x	opce, opční smlouva
contrat financier à terme sur taux d'intérêt	futures na úrokové sazby	ANO	NE	x	úrokové futures (futures na úrokové sazby)
contrôle des comptes annuels	audit roční účetní závěrky		NE	x	audit roční účetní závěrky
contrôle légal	statutární audit	ANO	NE	zákonný audit, povinný audit	povinný audit, statutární audit
contrôle légal des comptes	statutární audit		NE	povinný audit	povinný audit, statutární audit
contrôle légal des documents comptables	statutární audit účetních dokladů		NE	zákonný audit	audit účetních dokladů (dokumentů)
contrôle prudentiel	obezřetnostní dohled		NE	kontrola orgánem dohledu	obezřetnostní dohled
contrôleur des comptes	auditoři	ANO	x	auditor	auditor
contrôleur légal (des comptes)	statutární auditor		NE	statutární auditor	oprávněný auditor, statutární auditor
dépôt en liquide	vklad v hotovosti		NE	x	hotovostní vklad
dette envers des établissements de crédit	závazky vůči úvěrovým institucím	ANO	NE	x	závazky vůči úvěrovým institucím
dette envers la clientèle	závazky ke klientům	ANO	ANO	x	závazek vůči klientům
dette publique	veřejný dluh		ANO	veřejný dluh	veřejný dluh, státní dluh
effet public	státní cenné papíry	ANO	ANO	x	státní cenný papír

échange de taux d'intérêt	Úrokové swapy	ANO	ANO	x	úrokový swap, swap úrokových sazeb
échange des actions	Výměna akcií		NE	x	směna akcií, výměna akcií
échange d'informations	výměna informací		ANO	výměna informací	výměna informací
éléments d'actif	aktiva		x	aktiva	aktiva, aktivní položky
établissement de monnaie électronique	instituce elektronických peněz		ANO	instituce elektronických peněz	instituce elektronických peněz
exercice comptable	účetní období	ANO	ANO	účetní období	účetní období
fonds de pension	penzijní fondy		ANO	penzijní fondy	penzijní fondy
fonds de retraite	penzijní fondy		ANO	penzijní fondy	penzijní fondy
Fonds européen d'investissement	Evropský investiční fond		ANO	Evropský investiční fond	Evropský investiční fond
Fonds monétaire international (FMI)	Mezinárodní měnový fond		ANO	Mezinárodní měnový fond	Mezinárodní měnový fond
fonds propres	vlastní kapitál		ANO	vlastní kapitál	vlastní kapitál
fonds propres complémentaires	doplňkový kapitál		NE	x	dodatkový kapitál, doplňkový kapitál
fonds reçus en échange de monnaie électronique	Finanční prostředky získané výměnou elektronických peněz		NE	x	peněžní prostředky přijaté výměnou za elektronické peníze
fonds remboursables	splatné peněžní prostředky		ANO	x	splatné peněžní prostředky, splatné prostředky
chèque de voyage	cestovní šek		ANO	x	cestovní šek
immobilisations financières	finanční aktiva		ANO	x	finanční aktiva, dlouhodobá finanční aktiva
instruments dérivés hors bourse	OTC deriváty		ANO	x	mimoburzovní zúčtování (OTC) obchodovaných derivátových nástrojů, OTC derivátové nástroje
juste valeur	reálná hodnota		ANO	reálná hodnota	reálná hodnota
monnaie électronique	elektronické peníze		ANO	elektronické peníze	elektronické peníze
monnaie étrangère	cizí měna		ANO	cizí měna	cizí měna
option sur devises achetées	zakoupené měnové opce	ANO	ANO	x	zakoupená měnová opce
passif subordonné	podřízené závazky	ANO	ANO	podřízené závazky	podřízená pasiva, podřízené závazky

pour compte propre	na vlastní účet		ANO	x	na vlastní účet
protection des dépôts	ochrana vkladů		ANO	x	ochrana vkladů, pojištění vkladů
relevé de compte	výpisy z účtu	ANO	ANO	x	výpis z účtu
revenu d'action	výnosy z akcií	ANO	NE	x	výnos z akcií
succursale d'entreprise d'assurance	pobočka pojišťovny		NE	x	dceřiná pojišťovna mateřského podniku, pobočka pojišťovny
succursale des entreprises d'investissements	pobočka investičních společností	ANO	NE	x	pobočka investičního podniku
succursale d'établissement de crédit	pobočka úvěrové instituce		NE	x	pobočka úvěrové instituce
succursale d'établissement de crédit et d'établissement financier	pobočka úvěrové instituce a finanční instituce		NE	x	pobočka úvěrové a finanční instituce
système de garantie des dépôts	systém pojištění vkladů		NE	systém pojištění vkladů	systém pojištění vkladů
Système européen de banques centrales	Evropský systém centrálních bank		NE	Evropský systém centrálních bank	Evropský systém centrálních bank
taux de change	směnné kurzy	ANO	ANO	směnný kurz	směnný kurz
taux d'inflation	inflace		x	míra inflace	míra inflace
taux d'intérêt	úrokové sazby	ANO	NE	úroková sazba	úroková sazba; úroková míra
taux d'intérêt de référence	referenční úroková sazba		NE	x	referenční úroková sazba
titre à revenu fixe	cenné papíry s pevným výnosem	ANO	NE	x	cenný papír s pevným výnosem, cenný papír
titre à revenus variables	cenné papíry s proměnlivým výnosem	ANO	NE	x	cenný papír s proměnlivým výnosem
titre de créance	dluhové cenné papíry	ANO	ANO	dluhový cenný papír	dluhový nástroj, dluhový cenný papír
titre négociables	obchodovatelné cenné papíry	ANO	NE	převoditelný cenný papír	obchodovatelný cenný papír, převoditelný cenný papír
transaction sur actions	transakce s akciemi		ANO	x	transakce s akciemi
union douanière	celní unie		ANO	celní unie	celní unie
union économique et monétaire	Hospodářská a měnová unie		NE	Hospodářská a měnová unie	Hospodářská a měnová unie

Union européenne	Evropská unie		NE	Evropská unie	Evropská unie
union monétaire	měnové unie		ANO	měnové unie	měnové unie
union régionale	regionální unie	ANO	ANO	x	regionální unie
valeur comptable	účetní hodnota		ANO	x	účetní hodnota
valeur du marché	Tržní hodnota		ANO	tržní hodnota	tržní hodnota
valeurs mobilières	cenné papíry		ANO	převoditelný cenný papír	převoditelný cenný papír, cenný papír
virement transfrontalier	přeshraniční převody	ANO	NE	x	přeshraniční převod, přeshraniční platba
contrôle des comptes	audit		x	povinný audit	audit účetních závěrek, audit účetnictví
exercice de la surveillance	dohledu		x	provádění dohledu	dohled

Google Translator přeložil korektně 92 kolokací z celkových 145 (**63,4%**).

Seznam správně přeložených kolokací pomocí Bing Translatoru

kolokace	překlad Bing	databáze IATE	překlad (Virtová 2007)
accord de libre-échange	dohoda o volném obchodu	dohoda o volném obchodu	dohoda o volném obchodu
actif constituant des créances	aktiva představující pohledávky	x	aktiva představující pohledávky
actions préférentielles cumulatives	kumulativní prioritní akcie	x	kumulativní prioritní akcie
actions propres	vlastní akcie	x	vlastní akcie
adéquation des fonds propres	kapitálová přiměřenost	kapitálové požadavky	kapitálová přiměřenost
autorité de contrôle	dozorčí orgán	donucovací orgán	orgán dohledu
autorité de contrôle des banques	dozorčí orgán bank	x	orgán bankovního dohledu
Banque asiatique de développement	Asijská rozvojová banka	Asijská rozvojová banka	Asijská rozvojová banka
banque centrale	centrální banka	centrální banka	centrální banka
Banque centrale européenne	Evropská centrální banka	Evropská centrální banka	Evropská centrální banka
Banque de développement des Caraïbes	Karibská rozvojová banka	Karibská rozvojová banka	Karibská rozvojová banka
banque d'investissement	investiční banka	investiční banka	investiční banka
Banque européenne d'investissement (BEI)	Evropská investiční banka	Evropská investiční banka (EIB)	Evropská investiční banka (EIB)
Banque interaméricaine de développement	Meziamerická rozvojová banka	Meziamerická rozvojová banka	Meziamerická rozvojová banka

Banque internationale pour la reconstruction et le développement	Mezinárodní banka pro obnovu a rozvoj	Mezinárodní banka pro obnovu a rozvoj	Mezinárodní banka pro obnovu a rozvoj
Banque mondiale	Světová banka	Světová banka	Světová banka
Banque nationale d'investissement pour le développement industriel	Národní investiční banka pro průmyslový rozvoj	x	Národní investiční banka pro průmyslový rozvoj
Banque nordique d'investissement	Severská investiční banka	Severská investiční banka	Severská investiční banka
blanchiment de capitaux	praní špinavých peněz	praní peněz	praní peněz
catégorie d'actions	třída akcií	druh akcií	třída akcií, druh akcií
code d'identification de banque (BIC)	identifikační kód banky	identifikační kód banky	identifikační kód banky
compétence de contrôle	kontrolní pravomoc	přezkoušení odborné způsobilosti	kontrolní pravomoc
comptes annuels	roční účetní závěrka	účetní závěrka	roční účetní závěrka
comptes consolidés	konsolidované účetní závěrky	x	konsolidovaná účetní závěrka
Conseil de l'Union européenne	Rada Evropské unie	Rada Evropské unie	Rada Evropské unie
contrats d'échange	swapy	swapy	swapy
contrôle légal des documents comptables	povinný audit účetních dokladů	zákonný audit	audit účetních dokladů (dokumentů)
contrôle prudentiel	obezřetnostní dohled	kontrola orgánem dohledu	obezřetnostní dohled
contrôleur des comptes	auditor	auditor	auditor
contrôleur légal (des comptes)	statutární auditor	statutární auditor	oprávněný auditor, statutární auditor
dépôt d'épargne	spořicí vklad	x	úsporný vklad
dépôt en liquide	vklad hotovosti	x	hotovostní vklad
dette envers des établissements de crédit	dluh vůči úvěrovým institucím	x	závazek vůči úvěrovým institucím
dette publique	veřejný dluh	veřejný dluh	veřejný dluh, státní dluh
échange de taux d'intérêt	Úrokový swap	x	úrokový swap, swap úrokových sazeb
échange d'informations	Výměna informací	výměna informací	výměna informací
éléments d'actif	aktiva	aktiva	aktiva, aktivní položky
établissement de monnaie électronique	instituce elektronických peněz	instituce elektronických peněz	instituce elektronických peněz
exercice comptable	účetní období	účetní období	účetní období
exercice de la surveillance	výkon dozoru	provádění dohledu	dohled
fonds de pension	penzijní fondy	penzijní fondy	penzijní fondy
fonds de retraite	penzijní fondy	důchodové fondy	penzijní fondy
Fonds européen d'investissement	Evropský investiční fond	Evropský investiční fond	Evropský investiční fond
Fonds monétaire international (FMI)	Mezinárodní měnový fond	Mezinárodní měnový fond	Mezinárodní měnový fond
fonds propres	vlastní kapitál	vlastní kapitál	vlastní kapitál

fonds propres complémentaires	dodatkový kapitál	x	dodatkový kapitál
fonds reçus en échange de monnaie électronique	finanční prostředky přijaté výměnou za elektronické peníze	x	peněžní prostředky přijaté výměnou za elektronické peníze
instruments dérivés hors bourse	OTC deriváty	x	mimoburzovní zúčtování (OTC) obchodovaných derivátových nástrojů, OTC derivátové nástroje
juste valeur	reálná hodnota	reálná hodnota	reálná hodnota
monnaie électronique	elektronické peníze	elektronické peníze	elektronické peníze
monnaie étrangère	cizí měna	cizí měna	cizí měna
numéro international de compte bancaire (IBAN)	Mezinárodní číslo bankovního účtu	x	mezinárodní číslo bankovního účtu
obligation de couverture par les fonds propres	povinnost krytí kapitálem	x	povinné kapitálové krytí
prime d'émission	emisní ážio	emisní ážio	emisní ážio
protection des dépôts	Ochrana vkladů	x	ochrana vkladů, pojištění vkladů
relevé de compte	výpis z účtu	x	výpis z účtu
revenu d'actions	výnosové akcie	x	výnos z akcií
succursale d'entreprise d'assurance	pobočka pojišťovny	x	dceřinná pojišťovna mateřského podniku, pobočka pojišťovny
succursale d'entreprise d'investissement	pobočka investičního podniku	x	pobočka investičního podniku
succursale d'établissement de crédit	pobočka úvěrové instituce	x	pobočka úvěrové instituce
système d'indemnisation des investisseurs	systému pro odškodnění investorů	systém pro odškodnění investorů	systém pro odškodnění investorů
Système européen de banques centrales	Evropský systém centrálních bank	Evropský systém centrálních bank	Evropský systém centrálních bank
taux de change	Směnný kurz	směnný kurz	směnný kurz
taux d'inflation	míra inflace	míra inflace	míra inflace
taux d'intérêt	úroková sazba	úroková sazba	úroková sazba; úroková míra
taux d'intérêt de référence	referenční úroková sazba	x	referenční úroková sazba
titre de créance	dluhový nástroj	dluhový cenný papír	dluhový nástroj, dluhový cenný papír
titre négociable	obchodovatelný cenný papír	převoditelný cenný papír	obchodovatelný cenný papír, převoditelný cenný papír
union douanière	Celní unie	celní unie	celní unie
union économique et monétaire	Hospodářská a měnová unie	Hospodářská a měnová unie	Hospodářská a měnová unie
Union européenne	Evropská unie	Evropská unie	Evropská unie
union monétaire	Měnové unie	měnové unie	měnové unie
union régionale	regionální unie	x	regionální unie

valeur comptable	účetní hodnota	x	účetní hodnota
valeur du marché	tržní hodnota	tržní hodnota	tržní hodnota
valeur hypothécaire	Hypoteční hodnota	zástavní hodnota nemovitosti	hypoteční hodnota
office de chèque postal	poštovní banka	x	poštovní banka, banka poštovního úřadu

Bing Translator přeložil korektně 78 kolokací z celkových 145 (**53,8%**).