

**FILOZOFICKÁ FAKULTA UNIVERZITY PALACKÉHO
V OLOMOUCI**

KATEDRA SLAVISTIKY

**POČÍTAČEM PODPOROVANÝ PŘEKLAD- HISTORIE A
SOUČASNOST**

(bakalářská diplomová práce)

VYPRACOVALA: Eliška Balcarová

VEDOUcí PRÁCE: doc. PhDr. Zdeňka Vychodilová, CSc.

OLOMOUC 2012

Prohlašuji, že jsem práci vypracovala samostatně a uvedla všechny použité
prameny.

V Olomouci, 19. 4. 2012

podpis

Děkuji doc. PhDr. Zdeňce Vychodilové, CSc., za konzultace, rady a připomínky, které mi během psaní diplomové práce poskytla.

podpis

OBSAH

ÚVOD	5
1. Historie strojového překladu	6
1.1 50. a 60. léta	7
1.2 70. a 80. léta	8
1.3 Přelom 20. a 21. století.....	9
2. Teorie strojového překladu.....	10
3. Strojový překlad (SP).....	12
3.1 Problematika slovníků.....	13
3.2 Analýza textu.....	14
3.2.1 Analýza skrz syntézu.....	15
4. Počítačem podporovaný překlad	17
4.1 Překladová paměť a terminologická databáze.....	17
4.2 Segment, jazykový pár a tagy	19
4.2.1 Shoda překladu.....	19
4.3 Terminologie, kontrola pravopisu a korektura	20
4.4 Export, import projektu a normostrana	21
5. Překlad jako technická operace	21
5.1 Textový editor	22
5.2 Role slovníku při překladu	22
5.3 Grafika.....	22
6. Překladače	23
6.1 Struktura překladačů	24
7. CAT nástroje	25
ZÁVĚR.....	29
PEŽIOME	30
BIBLIOGRAFIE	35
PŘÍLOHA 1.....	37
ANOTACE.....	40

ÚVOD

Předmětem mé bakalářské práce je počítačem podporovaný překlad (strojový překlad), jeho historie a současnost. Cílem mé práce bylo zmapovat historii a vývoj strojového překladu. Zabývala jsem se jednotlivými obdobími vývoje strojového překladu. Při studiu odborné literatury se ukázalo, že je velmi důležitá i teorie překladu. Bez ní bychom nepochopili podstatu problematiky strojového překladu. Teorie počítačem podporovaného překladu má samozřejmě svou terminologii. Abych porozuměla všemu, studovala jsem počítačové slovníky a počítačové zkratky. Při zpracovávání tématu jsem využila především kompilační metody. Čerpala jsem z časopiseckých, knižních a internetových zdrojů.

Mohla by se naskytnout otázka, proč jsem zvolila toto technické téma. Odpověď je jednoduchá. V budoucnu se chci věnovat kariéře překladatele a tato práce mi umožnila nahlédnout na možnosti překladatelských softwarů.

Práce sestává ze sedmi základních kapitol. Nejdříve jsem zmapovala historii strojového překladu. Historii jsem rozdělila na tři podkapitoly z chronologického hlediska. Dále jsem se věnovala teorii strojového překladu. V kapitole strojový překlad jsem čerpala převážně z ruských zdrojů. K překladu neodmyslitelně patří práce se slovníkem, proto jsem následující kapitolu nazvala problematika slovníků. V kapitole analýza textu jsem čerpala opět z ruského zdroje, z knihy Osnovy obščego i mašinnogo perevoda. Následně se věnuji počítačem podporovanému překladu, jako takovému. Vysvětluji zde podstatu a smysl používání počítače při překladu. Další kapitoly jsou věnovány termínům, které je nutné znát, abychom byli schopni použít překladatelské programy. V této části mé práce jsem čerpala z časopiseckých zdrojů, a to zejména z časopisů ToP. Kapitola s názvem překlad jako technická operace je věnována technickým schopnostem, které by měl uživatel počítače mít. Dále jsem se věnovala překladačům, jejich struktuře a konstrukci. V poslední kapitole se věnuji jednotlivým překladatelským softwarům. V této části jsem čerpala převážně z internetových zdrojů.

Součástí mé práce je slovníček termínů a zkratk figurujících v oblasti překladatelských softwarů a počítačů. Tento slovníček jsem sestavila po přečtení nejrůznějších recenzí na překladatelské softwary, dále obsahuje termíny, které se vyskytují v mé práci.

1. Historie strojového překladu

Nahlédla jsem do historie a zjistila, jaké byly začátky počítačem podporovaného překladu. Průkopníky historie byli ve 30. letech **G. Artstrouni** a **P. S. Trojanskij**, kteří se podíleli na vynalezení překladače na principu děrované papírové pásky. „Historie strojového překladu (SP) má kořeny ve 40. letech minulého století. Již v roce 1946 byla vyřčena myšlenka o možnosti používání počítače pro překlad, v době, kdy se objevily první výpočetní stroje. V roce 1947 americký matematik a teoretik komunikace **Warren Weaver** poprvé zformuloval koncepci strojového překladu v dopisech s **Andrew D. Boothem** a **Norbertem Wienerem**. Warren Weaver psal: „Mám před sebou text v ruském jazyce, ale budu předstírat, že je napsaný anglicky a že byl zakódován do zvláštních znaků. Vše, co potřebuji, je rozšifrovat je tak, abych získal informaci, která je obsažena v textu. V kontextu poválečné doby byla podobnost překladu a dešifrování přirozená, jestliže bereme v úvahu úspěchy, kterých bylo dosaženo v šifrování v době 2. světové války. SP byl založen na koncepci dvou etap. Výchozí věta se překládala do zprostředkovatelského jazyka, který byl založen na zjednodušeném anglickém jazyce, a poté byl výsledek tohoto překladu představen prostřednictvím cílového jazyka. Memorandum Warrena Weavera vyvolalo zájem o problematiku SP. Stroje se v té době samozřejmě odlišovaly od těch současných. Byly velké a drahé, zabíraly celé místnosti a k jejich obsluze byla potřeba celá řada inženýrů a programátorů. Původně se tyto počítače používaly pro matematické výpočty a pro potřeby válečných úřadů, ale sloužily také matematickým a fyzikálním fakultám na univerzitách. Díky studené válce byl ve Spojených státech amerických kladen důraz na rusko-anglický překlad a v Sovětském svazu na anglicko-ruský překlad.“ (přeložila E. B.)¹ Šlo o konflikt, který vypukl po 2. světové válce. Sovětský svaz a USA měly politické spory a hospodářsky si konkurovaly, což bylo ukončeno až v roce 1991, kdy se Sovětský svaz rozpadl. V roce 1952 se konala první konference SP na Massachusettské technologické univerzitě. Impulzem pro aktivní činnost v oblasti SP byl známý Georgetownský experiment, který se uskutečnil v roce 1954. Byl to první vykonaný experiment překladu na světě, který proběhl na Georgetownské univerzitě v USA. Nehledě na jeho primitivnost (slovník se 150 jednotkami, gramatika se 6 pravidly, překlad několika jednoduchých vět), měl tento experiment velkou odezvu.

¹ http://www.dialog-21.ru/trends/?id=1744&forum_id=13&f=1

1.1 50. a 60. léta

V 50. a začátkem 60. let se začaly zajímat o SP mnohé země, např. V USA byla sestavena celá řada výzkumných skupin. Výzkumy ale probíhaly i v Anglii, Bulharsku, Itálii, Číně, Francii nebo Japonsku. Investoři vkládali do těchto výzkumů značné prostředky, nicméně výsledky nepřicházely. Jednou z hlavních příčin nízké kvality strojového překladu byly minimální parametry přístrojů, jako je malá paměť a zdlouhavý přístup k informaci, neplnohodnotné využití programovacího jazyka na vysoké úrovni. Další příčinou byl nedostatek teoretických základů, které byly nutné pro vyřešení jazykových problémů. Proto se první systémy strojového překladu omezily na doslovný překlad textů bez jakékoli syntaktické celistvosti. V polovině 60. let byly představeny dva systémy SP vytvořené v USA. Sloužily pro překlad z ruského jazyka do anglického. Byl to systém MARK 1 a systém SERNA zhotovený na Georgetownské univerzitě. V roce 1959 vystoupil filozof Yohoshua Bar-Hillel s tvrzením, že plně automatický strojový překlad nemůže být nikdy vysoce kvalitní. Jako příklad uvedl větu s anglickým slovíčkem *pen*, kdy toto slovíčko nelze přeložit jako pero, ale jako dětská ohrádka. V tomto případě je překlad podmíněn znalostí vnějších jazykových skutečností. A právě tato znalost je příliš obsáhlá a různorodá na to, aby byla zavedena do počítače. Bar-Hillel neodmítal nápad SP jako takový, ale považoval za perspektivní spíše vývoj strojových systémů, které by byly orientované na překladatele, který je využívá. Toto vystoupení se nepříznivě projevilo na vývoji SP ve Spojených státech amerických. Neblaze proslulá komise **ALPAC** završila toto první období vzrůstu tím, že označila překlady pomocí těchto systémů za nekvalitní a neziskové. Komise ve svých závěrech doporučila prohloubit teoretické výzkumy, aby byly sestrojeny dokonalejší systémy. Usnesení ALPACu se negativně odrazilo na dalším rozvoji SP. Zájem o problematiku už nebyl tak silný, financování ze strany vlád se pozastavilo a některé výzkumy byly zrušeny úplně. Poté se však objevily nové myšlenky a výpočetní technika se aktivně rozvíjela. Možnosti počítačů byly stále vyšší, už nebyly brány jen jako nástroj k výpočtům, ale jako prostředek zpracování informace.

„Následujících deset let se na systémech SP v USA pracovalo na univerzitě Brigham Young University ve státu Utah. Tato činnost byla financována Mormonskou církví, která měla zájem o překlad Bible. Zvláštní pozornost pak zasluhuje práce ruských jazykovědců v této oblasti, jako je **I. A. Melčuk** nebo **Ju. D. Apresjan**. Tito dva jazykovědci se dopracovali k lingvistickému procesoru **ETAP**. V roce 1960 byla v Leningradu založena laboratoř Vědecko-výzkumného institutu matematiky a mechaniky pro experimenty

strojového překladu, která byla později přestavěna na laboratoř matematické jazykovědy Leningradské státní univerzity.“ (přeložila E. B.)²

1.2 70. a 80. léta

„V 70. letech vešel strojový překlad do epochy „renesance“ současně s rozvojem výpočetní techniky. Objevily se mikropočítače, sítě prošly vývojem, zvýšily se i zdrojové paměti. Pozornost se začínala ubírat jinam. Výzkumníci si stanovili cíl vyvíjet systémy SP, které jsou reálné, a předpokládá se účast člověka při různých stádiích procesu překladu. Systémy SP se mění z nepřítelů na nenahraditelného pomocníka profesionálních překladatelů, který pomáhá šetřit čas a lidské zdroje. Objevila se celá řada prakticky pracujících systémů. Zkušenosti s SP bylo čím dál tím víc, staly se předmětem jednání konferencí. Jedna taková konference se uskutečnila v roce 1981 v Londýně a nesla název „Praktická aplikace strojového překladu.“³ V 80. letech byl vytvořen trh s komerčními systémy. O obrození strojového překladu svědčí i fakt, že Komise pro evropskou spolupráci (**CEC**) kupuje anglicko-francouzskou verzi **SYSTRAN** a také systém pro překlad z ruského do anglického jazyka. Systém SYSTRAN je nyní jedním z nejrozšířenějších systémů na světě. Je založen na již zmíněném systému GAT a zdokonalen ve smyslu hodnoty a využití glosářů. Kromě toho CEC objednává francouzsko-anglickou a italskou verzi. Díky CEC je v této době vytvořen projekt **EUROTRA**, který je založen na propracování skupin **SUSY** a **GETA**. Na systémech SP se začíná pracovat i v Japonsku. Tyto systémy jsou založeny na technologii interlingva. „Interlingua, též interlingva, je umělý mezinárodní jazyk, založený volně na románských jazycích, zejména latině (do jisté míry jde o zjednodušenou latinu), publikovaný (ve své definitivní verzi) v roce 1951 Mezinárodní asociací pro pomocný jazyk.“⁴ Panamerická zdravotnická organizace (**OPS**) si v USA objednává systém **SPANAM** v španělském a anglickém jazyce. Vojensko-letecké síly USA financují systémy strojového překladu v Jazykovědném výzkumném centru na Texaské univerzitě v Houstonu. Skupina **TAUM** dosahuje viditelných úspěchů při práci se systémem **METEO** v Kanadě. Tento systém se používal především pro překlad v oblasti meteorologických předpovědí. Celá řada projektů,

² http://www.dialog-21.ru/trends/?id=1744&forum_id=13&f=1

³ http://www.dialog-21.ru/trends/?id=1744&forum_id=13&f=1

⁴ <http://cs.wikipedia.org/wiki/Interlingua>

kteřé započaly v 70. a 80. letech, se stala plnohodnotnými obchodními systémy.“ (přeložila E. B.)⁵

V letech 1978 až 1993 se v USA za výzkumy v oblasti strojových překladů utratilo 20 milionů dolarů, v Evropě 70 milionů a v Japonsku 200 milionů.

Jednou z nových technologií se stala **TM** neboli translation memory, která pracuje na principu hromadění. To znamená, že je v procesu překladu zachován výchozí segment a jeho překlad, díky kterému se ve výsledku vytvoří základní lingvistické údaje. Jestliže je segment identický nebo podobný výchozímu segmentu, v překládaném textu se objeví znovu. Je zobrazen společně s překladem a uvedením shody v procentech. Překladatel se rozhoduje, jestli zvolí upravit, odmítnout nebo přijmout. Výsledek je v systému uložen. V konečném výsledku nemusíme překládat tutěž větu dvakrát. Dnes je nejvýznamnějším obchodním systémem **TRADOS**, který byl založen v roce 1984. Trados je založený na technologii TM.

1.3 Přelom 20. a 21. století

90. léta s sebou přinesla prudký vývoj trhu informačních technologií. Síť internet je široce využívána a stává se nejinternacionálnější a mnohojazykovou. To vše mělo vliv na další vývoj systému strojového překladu. Objevují se nové technologie, které jsou založeny na využití neuronových sítí. „Neuronová síť je jedním z výpočetních modelů používaných v umělé inteligenci. Jejím vzorem je chování odpovídajících biologických struktur. Umělá neuronová síť je struktura určená pro distribuované paralelní zpracování dat. Skládá se z umělých (nebo také formálních) neuronů, jejichž předobrazem je biologický neuron. Neurony jsou vzájemně propojeny a navzájem si předávají signály a transformují je pomocí určitých přenosových funkcí. Neuron má libovolný počet vstupů, ale pouze jeden výstup. Neuronové sítě se používají mimo jiné i pro rozpoznávání a kompresi obrazů nebo zvuků, předvídání vývoje časových řad (např. burzovních indexů), někdy dokonce k filtrování spamu. V lékařství slouží k prohlubování znalostí o fungování nervových soustav živých organismů. Například perceptronová síť vznikla původně jako simulace fyziologického modelu rozpoznávání vzorů na sítnici lidského oka.“⁶

⁵ http://www.dialog-21.ru/trends/?id=1744&forum_id=13&f=1

⁶ http://cs.wikipedia.org/wiki/Neuronová_síť

V červenci roku 1990 byl poprvé na světě představen na výstavě PC Forum v Moskvě první obchodní systém strojového překladu pod názvem **PROMT** (PROgrammer's Machine Translation). V roce 1992 společnost PROMT vyhrála konkurz na dodávku systémů strojového překladu do NASA. PROMT byla na tomto konkurzu jedinou neamerickou společností. Ve stejném roce PROMT vyrábí celou soustavu systémů pod novým názvem **STYLUS** pro překlad z anglického, německého, francouzského, italského a španělského jazyka do ruského jazyka. O rok později je vytvořen první systém strojového překladu pro Windows na světě. V roce 1997 byl podepsán kontrakt s francouzskou firmou **Softissimo** o vytvoření systémů překladu z francouzského do německého a anglického jazyka a naopak. V téže roce společnost PROMT vyrobila systém, který je založen na technologii Gigant pro podporu několika jazykových směrů. V roce 1998 je uvedena na trh celá řada programů pod novým názvem PROMT 98.

Automatické programy pro překlad se specializují na určité oblasti, jako jsou například technické manuály, nebo slouží jako prostředek pro rychlý, hrubý překlad, kterého využívaly zejména nadnárodní společnosti. Například systém SYSTRAN byl užíván NATO nebo Evropským společenstvím. Později vznikají první **CAT nástroje** (Computer Aided Translation, v překladu počítačem podporovaný překlad) pro mikropočítače, což byly pomůcky dostupné širší veřejnosti překladatelů. Při automatizaci překladu je kladen důraz na okamžitou odezvu, vysokou rychlost a nízké náklady, tedy co největší využití předností automatického strojového překladu.

2. Teorie strojového překladu

Vznikl překlad, který je specifický tím, že je při něm používán počítač, takzvaný strojový (automatický, automatizovaný) překlad. Když jsem četla literaturu, ukázalo se, že je velmi důležitá i teorie, která studuje problematiku strojového překladu. Teorie strojového překladu našla svůj smysl v početných deskriptivních pracích tvůrců, takzvaných „papírových“ algoritmů strojového překladu. Mezi tyto tvůrce můžeme **zařadit N. D. Andrejeva, I. Bar- Hillela, I. A. Melčuka, I. I. Revzina** nebo **V. Ju. Rozencvejga**. Teorie však našla smysl i v pragmatických průmyslových systémech, které jsou pod vedením **Ju. N. Marčuka**, dále pak Celosvazové skupiny „Statistika řeči“ pod vedením **R. G. Piotrovského**.

„Algoritmus je matematické pravidlo nebo postup použitý pro řešení určitého problému. Název pochází od středověkého arabského matematika jménem Ál Ghoritmí.“⁷

Co se týče „papírových“ algoritmů, Ju. N. Marčuk se domnívá, že se strojovému překladu v Sovětském svazu moc nepoškodilo. Idea strojového překladu ihned získala pozornost skupiny teoretiků, kteří v ní uviděli možnost, jak se zmocnit vedoucího postavení v jazykovědě. Žádné velké úsilí nevynaložili, jejich důkazem byly pouze obecné teorie.

Ju. N. Marčuk píše, že strojový překlad, jako každý překlad, je zajímavý tím, že je možné ihned ověřit správnost a sestavený algoritmus a provedení překladu ukáže efektivitu tvůrčích úvah.

Nicméně teoretikové si zvolili jinou cestu a přišli se sloganem „ strojový překlad bez překladu, bez strojů, bez algoritmů“. Takto se pokoušeli hájit ideu realizování překladu prostřednictvím deklaračních znalostí. Jazykovědec popíše pouze fakty jazyka a postavení algoritmů je na matematicích a programátorech. Toto nepochopení nové koncepce ze strany jazykovědců dalo možnost dotyčné skupině teoretiků zaujmout vedoucí postavení a zabránit rozvoji praktických prací v oblasti strojového překladu.

Tato spekulace byla ve značné míře i ve výpočetní technice. Teoretikové přesvědčivě hovořili o tom, že v té době moderní počítače nejsou schopné realizovat takové univerzální idey jako je „text- význam- text“, ale dokonalejší stroje dokážou udělat vše. V dnešní době máme mnohem dokonalejší počítače, než si umíme představit, ale do realizace nereálných idejí „papírových algoritmů“ máme stejně daleko, jako tomu bylo dříve.

Ju. N. Marčuk považuje za škodlivé i to, že reálné úsilí výzkumníků je narušováno nutnými směry v teorii a praxi strojového překladu, což mělo za následek ztrátu lingvistického zabezpečení moderních systémů překladu v rámci konkrétních jazykových dvojic. Proto je dnes velmi důležité uvědomit si a překonat tu ztrátu, jak v teorii počítačové jazykovědy, tak i v praxi při srovnávání charakteristik shod v překladu v konkrétních jazykových dvojicích.

⁷ Nádběla, J. 2004: Velký počítačový slovník. Kralice na Hané (Computer Media s.r.o.), s. 29

V dnešní době již existuje pozitivní přístup k teorii a praxi strojového překladu a to dovoluje vytvářet efektivní systémy strojového překladu a nahradit vyprovokovanou ztrátu. Tak, jak stoupá důležitost internetu, možnost vyhledávání internetem usnadňuje překladatelům práci. V roce 1977 byla spuštěna **AltaVista Translation Service** (zkráceně **Babelfish**) první internetová služba, která zdarma poskytovala SP. Později se objevují další internetové portály, které zdarma nabízejí překladatelské služby on-line. Překladatelské programy se stávají standardem překladatelské práce ve specializovaných oblastech.

3. Strojový překlad (SP)

Termínem strojový překlad (dále SP) je označován překlad textů z mateřského, přirozeného jazyka do cizího za pomoci osobního počítače (dále PC). V překladu za pomoci PC existuje mnoho různých forem organizace spolupráce mezi PC a člověkem. Původně se odborníci domnívali, že práce překladatelských systémů plně nahradí překladatele, člověk zůstane pouze editorem, který překlad postrediguje, tedy upraví po práci, kterou odvede stroj. Další způsob uspořádání práce, který vznikl později, byl překlad s úpravou před. Editor opravil chyby, zjednodušil výchozí text před tím, než přišla řada na stroj. Překlada předcházela práce editora. Tak, jak se postupně rozvíjela spolupráce člověka a PC, vznikl ještě jeden způsob uspořádání práce. Práce byla rozdělena mezi člověka a stroj. Vzniklo tzv. interaktivní editorování, kdy se člověk vměšuje do práce překladatelského systému v různých stádiích. Šlo o částečně automatizovanou úpravu, člověk musel řešit složité situace. Velmi blízko k těmto systémům SP měly automatické slovníky. V tomto případě člověk překládal a stroj pouze provedl vyhledávání ve slovnících. V práci překladatele zaujímala práce se slovníkem 40% celkového času. „Vytvoření překladatelských systémů potřebovalo řešení rozsáhlého kruhu problémů v různých oblastech vědění. Bylo potřeba:

- a) určit, pro jaké jazyky je systém sestavován;
- b) určit okruh překládaných textů;
- c) stanovit požadavky na kvalitu výsledného překladu;
- d) vybrat způsob rozdělení práce mezi člověka a PC;
- e) určit souhrnnou strategii, to znamená typ překladatelského systému;
- f) vypracovat různé druhy zabezpečení systému.

Při výběru jazyků evidovaných v systému byl rozhodující počet jazyků, výběr vstupního a výstupního jazyka. Původně se tyto otázky řešily následujícím způsobem: byly dva jazyky, přičemž se překládalo z cizího jazyka na jazyk země, kde je systém sestaven. To byl překlad s cílem získání informace. Poté začal být zájem o překlad z mateřského jazyka do cizího s cílem rozšíření informace. Následovala výroba systémů s více než dvěma jazyky. Zpočátku šlo o překlady vědecko-technického nebo obchodního charakteru a zřídka o překlady publicistické. Umělecká literatura nebyla brána v potaz. Samostatnou kapitolou pak byly systémy zaměřené na velmi specifickou tematiku. Příkladem byl systém METEO, který překládal pouze texty předpovědi počasí. „, (přeložila E. B.)⁸

3.1 Problematika slovníků

K vyčlenění jednotek překladu slouží slovník. Je to prostředek k identifikaci překladu. Mezi úseky textu výchozího jazyka a úseky textu cílového jazyka jednoduše vzniká shoda. V minulosti někteří profesionální překladatelé pohrdali dvojjazyčnými slovníky, dokonce se domnívali, že překladu škodí. Slovník je takový soupis shod, s kterým se setkáváme při překladu v praxi. Velice zajímavé je prohlášení jednoho z tehdejších teoretiků: „Slovník můžeme definovat jako nástroj, který umožňuje přeložit text a přitom nezná ani jeden jazyk.“ Nedůvěra ke slovníkům je paradoxně vyvolána jejich přednostmi, a to především tím, že nabízí nadměru variant překladu, přičemž praxe překladu požaduje, aby bylo každé slovo přeloženo pouze jedním způsobem. Můžeme se představit slovník, ve kterém je každá nabídka překladu oddělena lomítkem. Na prvním místě se nachází překlad a za ním další varianty překladu. Už takovýto překlad umožňuje pochopit význam a člověk je schopen překladu, aniž by uměl jazyk. Nicméně takovýto překlad je špatný, je v něm mnoho zbytečného, a tak ruší informaci. Mikrokontext odpovídajících jednotek znamená, že slovník zachycuje nejen slova, ale i slovní spojení. Je to jeden ze způsobů vyloučení zbytečné informace. Dalším způsobem vyloučení zbytečné informace je makrokontext. Jedno slovo může mít dva významy, význam doslovný a význam terminologický, což znamená, že se v prvním případě používá jako obyčejné slovo a v druhém případě jako termín.

⁸ Kulagina, O.S. 1991: O sovremennom sostojaniji mašinnogo perioda. In: Matematičeskije voprosy kibernetiki, Moskva (Nauka), s. 8-9

3.2 Analýza textu

Podmínkou úspěšnosti SP je analýza textu, jež je základním předpokladem odstranění homonymie a mnohoznačnosti slov. „Homonymie je vedle polysémie případ víceznačnosti/polyfunkčnosti izolované jazykové formy, avšak na rozdíl od ní bez významové příbuznosti. Homonymie se projevuje ve všech jazykových rovinách nesoucích význam: homonymní mohou být gramatické prostředky, slovotvorné prostředky, slova, spojení slov, věty. Nejčastěji se o homonymii hovoří v souvislosti s lexikem. Lexikální homonyma mohou vznikat hláskovým vývojem domácích slov, užitím homonymních stavebních prvků, tvarovým připodobněním slova ke slovu formálně blízkému, tvarovou shodou přejatého slova se slovem domácím, tvarovou shodou slov přejatých z různých jazyků.“⁹ Odstranění homonymie požaduje interpretaci textu, to znamená odhalení toho, co máme na mysli. Ustálení všech obrátů ve slovníku je nemožné, proto je dobré jít takovou cestou, kdy předpokládáme vyčlenění elementárních významů ve struktuře zprostředkovacího jazyka, které by nám dovolilo určit jednotky bez pomoci mikrokontextu. „Jedna z častých situací při analýze spočívá v tom, že pro některé úseky textu ve slovníku vůbec neexistuje shoda. To může být objasněno různými příčinami: neúplnost slovníku nebo tím, že dané slovo není zavedeno v žádném slovníku. To se týká neologismů a reálií. Další příčinou může být to, že daný úsek není slovem, může se jednat o vzorce, vlastní jména, schémata. Zde je z pohledu teorie překladu užitečné rozlišovat dva případy:

A. Slovo je generováno z kontextu. K tomu dochází, když fráze obsahuje nadbytečné informace. Vycházíme-li z kontextu fráze, není složité doplnit vynechaná slova, jak říká V.A. Zvegincev. Takové cvičení je velice důležité pro rozvíjení předpokladu doplnění do kontextu jak u překladatelů, tak i všeobecně u lidí, kteří se učí cizí jazyk. Předpokladem jsou elementární smysly, které umožňují nahradit vynechané slovo. Podobné nahrazení je důležité při synchronním tlumočení.

Př.

Мы пошли в разные ...

Эта ... была мной исхожена вдоль и поперок.

⁹ Karlík, Nekula, Pleskalová 2002: Encyklopedický slovník češtiny, Praha (Nakladatelství Lidové noviny), s. 169

Обсудим вопрос с разных

B. Slovo není nahrazeno v kontextu. Potom jsou možné následující případy:

a) Fáze analýzy nechává kontext nezměněný a takto se přenáší do zprostředkovacího jazyka a dále, ve fázi syntézy je kontext v cílovém jazyce.

b) Ve fázi analýzy vzniká doslovná shoda mezi písmeny a skupinami písmen překládaného slova a odpovídajícími fonémy. Podle tohoto principu se nejprve analyzují vlastní jména, geografické názvy, názvy úřadů, novin, parníků, slova označující tzv. domácí reálie a také neologismy, to znamená slova, která patří do tzv. bezekvivalentní lexiky.“ (přeložila E.B.)¹⁰

„Foném je minimální funkční (abstraktní) jednotka fonologického systému mající rozlišovací platnost, může rozlišit význam slova.“¹¹

3.2.1 Analýza skrz syntézu

Metody analýzy skrz syntézu jsou důležité pro řešení otázek o homonymii, ve skryté formě se vyskytují při jakékoliv analýze. „ Zpracování přirozeného jazyka, počítačové zpracování přirozeného jazyka (Natural language processing, NLP, též počítačová lingvistika neboli computational linguistics) je obor na pomezí lingvistiky, informatiky (umělé inteligence), popř. též akustiky a dalších. Zkoumá problémy analýzy či generování textů nebo mluveného slova, které vyžadují určitou (ne absolutní) míru porozumění přirozenému jazyku strojem.

Významné úkoly v NLP:

- Syntéza řeči (text-to-speech)
- Rozpoznávání řeči
- Generování přirozeného jazyka (en:Natural language generation)
- Strojový překlad (en:Machine translation)
- Odpovídání na otázky (en:Question answering)
- Získávání informací (en:Information retrieval)

¹⁰ Revzin, Rozencvejk 1964: Někotoryje specijalnyje voprosy teriji obščego i mašnnogo perevoda. In: Osnovy obščego i mašinnogo perioda, Moskva (Vysšaja škola), s. 180-222

¹¹ Lotko, E. 2003: Slovník lingvistických termínů pro filology, Olomouc (Univerzita Palackého v Olomouci), s. 36

- Extrakce informací (en:Information extraction)
- Korektura textu
- Výtah z textu (en:Automatic summarization)¹²

„Znalost transformační historie pomáhá vyčlenit dva druhy opakovaného záporu v ruském jazyce.

1. Opakování se přenáší do cílového jazyka jednorázovým záporem. Zápor tohoto druhu nemá vliv na smysl.

2. Dvojnásobný zápor, který vznikl transformací a zachoval smysl věty.

Znalost transformační historie je důležitá pro překlad přechodníkových konstrukcí v ruském jazyce. Věc se má tak, že v ruštině může mít přechodníkový obrat nejružnější příslovečný význam. Výsledkem transformace jsou vedlejší věty.“ (přeložila E.B.)¹³ Pravidel je zde nespočet. Řeč ještě nebyla o tom, jak ve SP zjišťujeme syntaktické vztahy, to znamená spojitost mezi slovy všeobecně. Pro teorii překladu je tento proces důležitý, protože simuluje metodu testování a chyb, které se často vyskytují v lidské analýze textu v málo známém jazyce.

Teď si řekneme něco málo o nezávislosti analýzy na syntéze. Objevily se tzv. binární algoritmy překladu a mohlo se zdát, že překlad závisí na cílovém jazyce, že fáze syntézy je závislá na analýze. Ve skutečnosti tomu tak nebylo. Zprostředkovací jazyk byl systém shod mezi dvěma danými jazyky a vlastnosti tohoto zprostředkovacího jazyka zcela závisely na dvojici jazyků. Bylo pohodlné sestavit takové binární zprostředkovací jazyky, kdy pravidla překladu z jazyka 1 do jazyka 2 byla zcela odlišná od pravidel překladu z jazyka 2 do jazyka 1. A tak se ve SP přišlo na to, že pravidla překladu z ruštiny do angličtiny mohou být velice jednoduchá tehdy, pokud budou pravidla překladu z angličtiny do ruštiny velmi složitá. V problematice SP bylo velmi rychle pochopeno, jakou výhodu mají tato sestavená pravidla, při kterých není analýza textu výchozího jazyka závislá na vlastnostech cílového jazyka a naopak. Ukázat to můžeme na jednoduchém příkladu. Jestliže máme sto jazyků a chceme text přeložit z každého do každého, potřebujeme devět set různých souborů pravidel, protože se

¹² http://cs.wikipedia.org/wiki/Zpracování_přirozeného_jazyka

¹³ Revzin, Rozencvejk 1964: Někotoryje specijalnyje voprosy teoriji obščego i mašinogo perevoda. In: Osnovy obščego i mašinogo perevoda, Moskva (Vysšaja škola), s. 180-222

jedná o kombinace jazyků ve dvojicích. Pokud využijeme nezávislou analýzu a nezávislou syntézu, tak bude postačující provést sto analýz a sto syntéz. Nezávislou analýzou je zde myšlen přechod od textu výchozího jazyka ke zprostředkovacímu jazyku, jehož výsledkem je pochopení textu. Naopak nezávislá syntéza znamená přechod od zprostředkovacího jazyka k textu cílového jazyka. Jedná se o tvůrčí moment, vyjádření myslí.

4. Počítačem podporovaný překlad

Ted' bych se ráda věnovala překladům podporovaným počítačem, jako takovým. Počítačem podporovaný překlad či překlad podporovaný počítačem je termín, užívaný v češtině pro zkratku CAT (z anglického computer-aided translation či computer-assisted translation). K tomuto překladu využíváme software, který se nazývá CAT nástroje. Mezi nejznámější CAT nástroje patří Déjà Vu, Transit, Trados nebo OmegaT. Při práci s těmito softwary pracujeme s překladovou pamětí. Text si rozdělíme na jednotlivé části vět neboli segmenty, které přeložíme. Překlad se uloží do databáze. Jeden pár zdrojového segmentu a jeho překladu se nazývá překladová jednotka (translation unit). Jestliže se pak překladová jednotka objeví v textu znovu, program nabídne překladateli to, co bylo uloženo dříve. Tyto programy se využívají především v odborných textech, ne v beletrii.

4.1 Překladová paměť a terminologická databáze

Co se týče překladové paměti, „v širším slova smyslu se jedná o počítačový nástroj, obvykle v podobě komerčně dostupného softwaru, s jehož pomocí překladatel překlad provádí, dal by se také nazvat překladatelské studio.“¹⁴ Databáze překladové paměti zahrnuje jazyk výchozí a jazyk cílový. Překladové jednotky jsou děleny do dvojic, jedná se o větné celky. Při překladu je překladová paměť prohledávána, pokud dojde ke shodě, jednotka je uživateli nabídnuta. Jsou zde dvě možnosti. Překladatel jednotku zkontroluje, odsouhlasí shodu a pokračuje dál, nebo je shoda pouze částečná. V tomto případě překladatel jednotku upraví a až poté odsouhlasí. Nemusí tu být však shoda žádná. V takovém případě uživatel přeloží segment sám a překladovou jednotku uloží. Překladová paměť se neustále navyšuje, čímž se zvyšuje efektivita užívání programu.

Těmito programy se překládají texty odborných funkčních stylů, kde se často opakují odborné termíny. Jestliže používáme tuto překladovou paměť, překlady jsou rychlejší a

¹⁴ Král, P. 2005: Český překlad II (1945-2004). Praha (Univerzita Karlova), s. 388

přesnější. Jsou zde však i úskalí, která se dělí na principálně-teoretická a uživatelsko-praktická. První problematika se týká toho, jak počítač pracuje s překladovou pamětí. Počítač na základě překladové paměti segmenty pouze srovnává, srovnává jejich znaky. Na jazykový význam už důraz neklade. Počítač pracuje mechanicky, nebere v potaz kontext textu. Tyto problémy se objevují zejména u jazyků, které mají velké typologické odlišnosti. Při překladu pak může dojít k tomu, že bude nalezena úplná shoda a přesto se nebude jednat o ekvivalent. Příkladem je překlad z angličtiny do češtiny: *low initial engine speed / nízké startovní otáčky motoru – nízkými startovními otáčkami motoru*. Jsou zde i další problémy. „ Za mnohé jmenujme např. realizaci funkční větné perspektivy. V češtině je **réma**, tj. jádro výpovědi, obvykle na konci věty, takže jedné anglické větě mohou ve dvou různých kontextech odpovídat různě tematicky strukturované české překlady. (Př. *The control levers are at the front / Na přední části jsou umístěny ovládací páky – Ovládací páky jsou umístěny na přední části*).“¹⁵ Pozor musíme dát i na to, jestli se jedná o kontext, nadpis nebo popis obrázku neboli **situační kontext**. Příkladem může být anglická věta *How to log into the collaborative platform*. V textu tuto větu přeložíme jako *Jak se přihlásit k platformě pro spolupráci*, kdežto jako nadpis bychom to přeložili jako *Přihlášení k platformě pro spolupráci*. Nalezená shoda může být vhodná pro větu v textu, ale už se nemusí hodit jako popis tabulky. Věta může mít dva naprosto odlišné významy neboli **pragmatický kontext**. V takovéto situaci musíme dát pozor na překlady do cílového jazyka. Příkladem je slovo *obloha* v jídelním lístku, která bývá často chybně přeložena jako *sky*. Pokud se spoléháme na překladovou paměť příliš, překlad nebude mít dobrou kvalitu. Uživatelsko-praktická problematika může spočívat v pohodlnosti překladatele. Musíme být kritičtí, pečlivě vybírat shody.

Jak jsem se dočetla na wikipedii, dalším nástrojem, který je obsažen v CAT nástrojích, je terminologická databáze (terminology database, zkráceně termbase), kterou uživatel (překladatel) naplňuje ručně, nenaplňuje se tedy automaticky na základě dosavadního překládání. Terminologická databáze obsahuje konkrétní termíny ve zdrojovém i cílovém jazyce. Terminologická databáze se využívá zejména u technických textů při překládání přesných termínů, například chemického názvosloví, kdy program při překládání automaticky nahrazuje termíny typu carbon dioxide termínem oxid uhličitý nebo greenhouse effect termínem skleníkový efekt. Zbytek segmentu (věty) překladatel však musí přeložit sám.

¹⁵ Král, P. 2005: *Český překlad II (1945-2004)*. Praha (Univerzita Karlova), s. 390

4.2 Segment, jazykový pár a tagy

Segment je nejmenší jednotka, na kterou CAT program rozdělí původní text. Nejčastěji je to jedna věta nebo jeden řádek. Tato samostatná jednotka se překládá najednou. Segment se překládá do cílového jazyka také jako segment. Není brán ohled na délku jednotky.

„**Jazykový pár** je souhrn všech segmentů vytvořených z jednoho souboru, Takto pracuje například TRANSIT. Z každého souboru na přeložení vytvoří dvojici jazykových párů (pro zdrojový a cílový jazyk), obsahující určitý počet segmentů podle výše uvedených kritérií. TRADOS z každého dodaného souboru vytvoří nový soubor. Projekt pak obsahuje tolik souborů TRADOS, kolik bylo původních souborů. V RANSITU se všechny dodané soubory, přeměněné na jazykové páry, spojí v jeden projekt.“¹⁶

Tag má podobnou funkci jako **HTML**. „HTML je zkratka Hypertext markup Language (hypertextový značkovací jazyk). HTML je v dnešní době stále nejoblíbenější jazyk pro vytváření jednoduchých www stránek. HTML je značkovací jazyk, který používá definované značky (tagy) k vytváření a formátování dokumentů pro webové stránky.“¹⁷ Tag je značka, která odděluje obsah textu od jeho formy. To znamená, že udává velikost písma, barvu textu, podtržení, vzhled písma. Tagy jsou párové a nepárové. Párové tagy označují začátek a konec podtrženého textu. Nepárové tagy jsou rozdělovací znaménko nebo mezera. Tagy jsou důležité pro exportování překladu do původního formátu. Pokud některý z tagů smažeme, konečný překlad už není možné exportovat.

4.2.1 Shoda překladu

Dále je třeba se zmínit o rozsahu shod a pojmech s nimi souvisejících. **Perfect match** (PM) znamená naprostou shodu zdrojového segmentu se segmentem překladové paměti. Nicméně se překlad musí zkontrolovat, jestli segment logicky zapadá do překládaného textu. Čím delší segment je, tím větší je pravděpodobnost, že se bude shodovat s překladem. Překlad může obsahovat například číslice nebo jednotky. Jestliže tyto segmenty označíme jako PM, CAT program je přeloží sám. Dále sem patří segmenty, které se liší pouze jedním znakem. Pokud máme uloženo v překladové paměti 500 t a potřebujeme přeložit 505 t, CAT program

¹⁶ Vedral, J. (2010). Základní pojmy používané při překladech pomocí CAT nástrojů. Časopis ToP 96, s. 30

¹⁷ Pilař, A. 1999: Počítačové zkratky. Praha (Grada Publishing), s. 105

automaticky nahradí číslou 500 číslem 505, což se nazývá *partially translated*. Zde platí totéž, co u PM. Čím delší je segment, tím větší bude shoda. Dalším pojmem v CAT programu je *internal repetition* (vnitřní opakování). Jedná se o segment, který se v textu velmi často opakuje. Nejprve ho přeložíme my a poté už ho překládá CAT program. Samozřejmě to musí být stoprocentně shodné segmenty.

Překladatel nebo CAT program může snížit shodu překladu. To se nazývá *penalizace*. Když máme segment, který se liší v překladu pouze tečkou na konci, je zřejmé, že to neovlivní text. Můžeme tedy nastavit nulovou penalizaci za interpunkční znaménka. Segment se přeloží stejně bez ohledu na čárky, středníky, pomlčky nebo odrážky. Dále pak můžeme nastavit penalizaci za velikost písma, barvu písma nebo font.

Fuzzy match znamená, že je shoda obsahu segmentu se segmentem v překladové paměti vyšší než padesát procent a nižší než sto procent. Obsah překladu se může lišit pouze jedním slovem. CAT program nabídne překlad, který je podobný nebo stejný jako segmenty v překladové paměti. Nabízený překlad pak stačí potvrdit a přepsat pár písmen, aby byl překlad logický. Jestliže je segment krátký, může se stát, že nám CAT program nabídne úplný nesmysl. Pokud je shoda nižší než padesát procent, překladatel musí segment přeložit sám.

Pojem **konkordance** znamená, že překladatel nezná překládaný segment. Funkce konkordance prohledá celou překladovou paměť a zobrazí všechny možnosti, které obsahují hledaná slova. „Segment překladu tedy může mít stav: (před)přeložený = překlad pomocí *perfect match*, *partially translated*, *internal repetition*, *fuzzy* = překlad segmentu čeká na překladatele, ale CAT nabídne překlad, nepřeložený = překladatel musí segment přeložit, CAT mu sám překlad nenabízí.“¹⁸

4.3 Terminologie, kontrola pravopisu a korektura

Dalšími oblastmi, s nimiž se může překladatel při své práci setkat, jsou terminologie, kontrola pravopisu a korektura. Překladatel dostane zakázku s určitou terminologií, tvořivost překladatele zde není žádána. Toto platí obzvláště v takových případech, kdy na projektu pracuje více překladatelů nebo je projekt překládán do více jazyků. Udržet jednotnou terminologii je nemožné. K zakázkám se tedy přidávají slovníky se závaznou terminologií.

¹⁸ Vedral, J. 2010: Základní pojmy používané při překladech pomocí CAT nástrojů. Časopis ToP 96, s. 30

Po dokončení překladu překladatel musí zkontrolovat správnost pravopisu. CAT programy v sobě většinou mají korektor pravopisu. Ten kontroluje překlapy slovo od slova. Neopraví však gramatické a syntaktické chyby. Dále překladatel musí zkontrolovat gramatická pravidla (jazyková korektura), jestli je dodržen závazný slovník a terminologie (srovnávací, terminologická korektura, korektura konzistence), správné umístění tagů v textu (tag checking) a zda byl přeložen celý text projektu.

4.4 Export, import projektu a normostrana

Jak píše Jiří Vedral ve svém článku Základní pojmy používané při překladech (Vedral 2010), CAT nástroje dovedou převést větší množství formátů do jednoho jediného. Dojde k oddělení formy od obsahu a převedení textu ze zdrojového souboru na segmenty. Program tedy umí importovat určitý formát dat a exportovat dokument v totožném vzhledu (za předpokladu dodržení rozmístění všech tagů), ale v jiném jazyce. Normostrana (NS) je standardizovaná strana textu. Podle této normostrany se počítá cena překladu, která zahrnuje znak, slovo a řádek. Standardizovaná strana textu- např. *§ 24 odst. 2 vyhlášky 37/1967 Sb.*, k provedení zákona o znalcích a tlumočnících- hovoří o délce 1800 znaků (symbolů a mezer, někdy mylně používaného úhazu), což odpovídá třiceti řádkům o šedesáti znacích nebo přibližně 250 slovům běžného textu. Před nástupem počítačů se počet NS počítal podle počtu řádků a znaků na řádce psacího stroje, dnešní textové editory ukážou počet slov, řádků, znaků s mezerami (děleno 1800 = počet NS) a bez mezer (dělíme 1500). To byly základní pojmy, které jsou nezbytné pro používání CAT programu.

5. Překlad jako technická operace

V této kapitole bych se ráda věnovala tomu, jakou roli hraje technika v překladu pomocí počítače. V dnešní době se stal překlad jevem, který slučuje sociální, kulturní, politické a náboženské aspekty s aspekty jazykovými, logickými a technickými. Jak píše Daniela Mügelová ve své knize *Komunikácia, tlmočenie, preklad alebo prečo spadla Babylonská veža* (Mügelová 2009), cílem překladu v dnešní době není zaměnit slovo za slovo, cílem je poskytnout komplexní produkt, který vyhovuje po stránce formální, obsahové, grafické, technické a jazykové. Překlad tedy není jen jazyková operace, ale i technická. Překladatelé by měli reagovat na změny, které se dějí v technologické oblasti zpracování textů a v oblasti technologických nástrojů, které mohou při překladu využít. Nedostatek použití technologických nástrojů se projevuje právě v oblasti překladu technických a odborných textů.

Mezi technické schopnosti překladatele patří používání počítače při překladu, ovládání textových, grafických a tabulkových editorů, používání internetového vyhledávače, práce s elektronickými slovníky a terminologickou databází, práce se softwarovými nástroji na automatický strojový překlad, práce s nástroji CAT nebo práce s lingvistickými softwary. Technika v překladu je v dnešní době nenahraditelná.

5.1 Textový editor

Jedním z nástrojů pro práci překladatele je textový editor. Slouží pro realizaci textu jakéhokoli charakteru. První textové editory byly založeny na principu editace textu v jednom řádku. Obsahovaly ty nejjednodušší operace. Princip textových editorů byl velmi podobný práci na psacím stroji. Moderní editory prošly velkými změnami, zejména se změnila vizualizace uživatelské rozhraní textových editorů.

5.2 Role slovníku při překladu

Jednou ze základních činností překladatele je práce se slovníkem. V moderní době se vedle slovníků v knižní podobě již na trhu vyskytují elektronické slovníky, které šetří čas. Do elektronického slovníku můžeme dopisovat nová slova, elektronický slovník je ihned zařadí podle abecedy. Mezi výhody počítačového slovníku patří rychlost vyhledávání, velké množství slov nebo nulová váha. Mezi nevýhody pak patří někdy chybný obsah, pro použití potřebujeme počítač a omezené informace ve slovníkovém hesle. Slovníkové heslo obsahuje základní informace o slově, jeho gramatických kategoriích, výslovnost a příklady použití.

Mezi základní typy knižních slovníků patří slovníky **monolingvální** (vysvětlovací/výkladové), **bilingvální** (překladové), **tezaurické** (organizace slovní zásoby synonymicky nebo podle témat), **piktorální** (obrázkové), **frekvenční**, **slovníky zkratek**, **synonymické**, **fonetické**, **ortografické**, **etymologické**, slovníky regionální lingvistiky, specializované slovníky (medicína, právo a další). Mezi výhody knižního slovníku patří komplexní informace, slovník jako kniha je náš majetek. Naproti tomu je vyhledávání v knižním slovníku pomalé a slovník nemůžeme aktualizovat.

5.3 Grafika

Překladatel by měl ovládat i grafické soubory, protože překlady mohou obsahovat schémata, nákresy nebo tabulky. Grafické elementy mohou být v počítači zobrazeny dvěma základními formáty, dvěma typy zobrazení. Rastrový a vektorový formát. „Bitmapová grafika (rastrová grafika) je obrazová informace, která je uložena jako množina bodů. Každý obrazový bod (pixel) má nastavenou barvu. Takto vytvořený soubor má větší velikost,

obvykle se používá komprese typu JPEG nebo GIF.¹⁹ Zatímco vektorová grafika využívá k definování obrázků matematicky definované vektory. Jak píše Josef Nádběla ve svém Velkém počítačovém slovníku, jednotlivé body obrázku nejsou popsány, ale jsou popsány prvky jako celek (obdélník, kružnice, křivka). Výsledný soubor je menší. Tento program využívají např. programy CAD (Computer Aided Design, tedy program, který podporuje návrhy, kresby, schémata, tvorbu technické dokumentace), CorelDraw!, ZonerCallisto.

6. Překladače

„Překladač je program, který jako vstupní data zpracovává text zapsaný v nějakém jazyce. Tomuto jazyku budeme říkat zdrojový jazyk a vstupnímu programu budeme říkat zdrojový program.“²⁰ Překladače můžeme rozdělit na interpretační a kompilační. „Kompilační překladač čte zdrojový program zapsaný ve zdrojovém jazyce a překládá (transformuje) jej na ekvivalentní cílový program zapsaný v cílovém jazyce.“²¹ Interpretace zdrojového programu je taková činnost, při které jsou čteny vstupní data s je prováděn výstup výsledků.

„Interpretační překladač čte zdrojový program, analyzuje jej a zajišťuje provádění operací, které odpovídají jednotlivým příkazům zdrojového programu.“²²

V době, kdy se objevily první počítače, vznikla snaha realizovat automatickou činnost. Podstatou této automatické činnosti by bylo pochopení originálu a tvorba překladu. Mezi první překladače patří **Systran**, automatický strojový překladač, který se začal vyvíjet roku 1968. Společnost Systran realizovala mnoho projektů pro Evropskou komisi a Ministerstvo obrany v USA. „Systran v dnešní době poskytuje automatickou překladovou technologii pro servery Yahoo!, Alta Vista (Babel Fish) a jazykové kombinace v jazykových nástrojích Google. Komerční verze běží pod operačními systémy Microsoft Windows (a Windows Mobile), Linux a Solaris.“ (přeložila E.B.)²³ Systran zvládá překlady dvou odlišných jazykových skupin a to se děje díky pokročilému zpracování textů. Nepracuje pouze na povrchové úrovni jazyka, ale především zasahuje do významové úrovně.

¹⁹ Nádběla, J. 2004: Velký počítačový slovník. Kralice na Hané (Computer Media s.r.o.), s. 63

²⁰ Melichar, Češka, Ježek 1999: Konstrukce překladačů I. Část, Praha (Vydavatelství ČVUT), s. 7

²¹ Opakovaná citace, s. 7

²² Opakovaná citace, s. 8

²³ Mügelová, D. 2009: Komunikácia, tlumočenie, preklad alebo prečo spadla Babylonská veža. Bratislava (Enigma Publishing), 298 s.

Dalšími typy překladačů jsou **PC Translator**, **Free language Translator** nebo **Wordmaster XP**, nejpoužívanějším je však díky své popularitě **Google Translate**. Tyto nástroje jsou v dnešní době velmi pokročilé, ale zatím nejsou schopné nahradit roli překladatele. Nejsou vhodné pro překlad uměleckých textů, pomohou však uživateli rychle pochopit podstatu textu. Jsou vhodné pro překlad manuálů, internetových stránek nebo obchodní korespondence.

6.1 Struktura překladačů

Překladače mají různou strukturu, můžeme však najít několik společných rysů, mezi které patří lexikální analyzátor, syntaktický analyzátor, zpracování sémantiky nebo generátor cílového programu. „Lexikální analyzátor čte postupně znaky zdrojového programu a vytváří z nich lexikální symboly programu jako např. čísla, identifikátory, klíčová slova (např. begin, do, for) a jednoznakové a víceznakové omezovače.“²⁴ Identifikátor je kombinace čísel a znaků. Výsledkem lexikální analýzy je posloupnost znaků, symbolů, které předávají štafetu syntaktické analýze. „Syntaktický analyzátor analyzuje syntaktickou strukturu programu a přitom zjišťuje, zda je program zapsán syntakticky správně. Výstupem syntaktického analyzátoru je informace o syntaktické struktuře programu nebo signalizace chyby.“²⁵ Zpracování sémantiky zahrnuje program ve vnitřní formě, což znamená, že jsou všechny znaky převedeny do číselné formy pomocí lexikálního analyzátoru. Tento program ve vnitřní formě je vstupem pro generátor cílového programu. To byla struktura kompilačního překladače, struktura interpretačního překladače je podobná. První tři rysy se neliší, ke změně dochází u čtvrtého rysu. Namísto generátoru cílového programu je zde interpretace. Dochází zde k interpretování požadovaných operací. „Je-li přesně specifikován vstupní a výstupní jazyk, představuje konstrukce překladače dva zásadní úkoly: 1. Specifikaci překladu, tj. určení korespondence mezi vstupním a výstupním jazykem 2. Implementaci překladu.“²⁶ Toto velmi povrchné nahlédnutí do problematiky překladačů vystřídají překladačské softwary.

²⁴ Opakovaná citace, s. 9

²⁵ Opakovaná citace, s. 10

²⁶ Koukolíková, J. 1990: Aplikace prologu v konstrukci překladače. In: Překladače programovacích jazyků, Praha, s. 124

7. CAT nástroje

Počítačem podporovaný překlad neslouží k tomu, aby pracoval za překladatele. CAT nástroje se snaží pomoci překladateli v jeho práci. Efektivita těchto softwarových nástrojů je závislá na šikovnosti překladatele a jeho technických možnostech. **Metatexis** je program, který má překladovou paměť i terminologickou databázi. Moderní databázové programy dokážou rozpoznat kromě překladu i gramatické kategorie, příklady použití, nebo dokonce obrázek. Jestliže se překladatel na program spoléhá až moc, překlad může ztratit na kvalitě. Proto je dobré využívat technologické nástroje jen tehdy, když je opravdu potřebujeme.

Mezi nejznámější CAT nástroje patří **Déjá Vu**, **Trados**, **Transit**, **Wordfast**, **Cafetran**, **OmegaT** nebo **SDLX**. Jedná se o překladatelské softwary, které lze zakoupit. Ceny jsou různé, pohybují se v řádech deseti tisíců. Rozdíl je v jednotlivých verzích. Například software Transit Freelance PRO stojí v současné době okolo třiceti tisíc korun a verze Professional stojí už o něco víc- kolem šedesáti tisíc korun s tím, že verze Freelance jsou určeny pro překladatele na volné noze. Software Trados se pohybuje okolo čtyřiceti tisíc korun a Wordfast byl do roku 2002 bezplatný, poté však získal licenci. Ceny softwaru Wordfast jsou však pětina ve srovnání s cenami jiných CAT softwarů.

Transit NXT je jednoduchý překladatelský software založený na překladatelské paměti a fuzzy logice. Nejedná se o strojový překlad, software za nás text nepřeloží. Zapamatuje si však každý segment, který přeložíme, a pokud se v budoucnu objeví identický segment, je nabídnut. **Fuzzy logika** znamená, že je shoda segmentů pouze částečná. Transit NXT najde rozdíl mezi překládanými větami a na nás potom je, abychom opravili, co je špatně. Výhoda tohoto softwaru je v tom, že se nemusíme starat o formátování dokumentu. Transit NXT vše zachová, my nemusíme nic měnit a můžeme se v klidu věnovat překladu. Další výhodou je vyhledávání výrazů, Transit NXT má zabudovanou morfologickou analýzu, a proto je schopen vyhledat všechny výrazy v různých gramatických tvarech. V neposlední řadě je zde nespočet kontrolních funkcí. Slovníky z předchozích překladů můžeme kdykoli rozšířit.

OmegaT je dalším příkladem překladatelského softwaru. Tato aplikace se skládá z jednoho velkého okna, které má ještě tři podokna. Vlevo se nachází editor, kam vpisujeme svůj překlad. Nahoře napravo je jedno z podoken s názvem přibližný překlad. Zde se objevují nalezené shody překladu s překladovou pamětí. Pod tímto oknem se nachází glosář, náš vlastní slovník, který jsme vytvořili. Již ve zmíněném okně editor je zobrazen výchozí text,

který je rozdělen do jednotlivých segmentů, tedy vět nebo odstavců. Segmenty překládáme jeden za druhým, překlady jsou zaznamenávány do překladové paměti. Po přeložení všech segmentů OmegaT vygeneruje přeložený dokument do cílové složky za pomoci překladové paměti. Každý projekt má svůj adresář, ve kterém je uložen zdrojový text, to znamená text, který překládáme, překladová paměť, glosář a nakonec cílová složka, kterou OmegaT sama vytvoří, a automaticky je zde uložen výsledný překlad. Nový projekt vytvoříme tak, že v menu vybereme projekt a uložíme název projektu. Dále musíme určit jazyk zdrojových souborů a jazyk přeložených souborů, segmentaci vybereme buď podle vět, nebo podle odstavců podle toho, co nám vyhovuje víc. Překlad se píše mezi tagy tak, že přepíšeme zdrojový text. Pro další segment může OmegaT najít v překladové paměti přibližnou shodu, to je shoda větší než 30%. OmegaT má několik vyhledávacích funkcí, mezi které patří i Google překladač. Stačí jej aktivovat.

CafeTran je program s mnoha funkcemi, jeho ovládání je jednoduché a překlad je podporován on-line překladatelskými systémy. Design je uzpůsoben tak, aby bylo možné pracovat rychle. V CafeTran najdeme funkce, které nám ušetří čas, a práce se stává produktivnější. Mezi tyto funkce patří vytvoření druhé překladové paměti pro vybrané termíny, které shledáme během překladu důležitými. Funkce automatického překladu je závislá na velikosti překladatelské paměti. Čím větší paměť máme, tím přesnější bude výsledek. Funkce fuzzy matches (v překladu nepřesné shody) nabídne vždy tři shody pro překládaný segment. Výše shody je minimálně 33%, výši však lze nastavit v poli úpravy. Další užitečnou funkcí je automatické dokončení, které automaticky doplňuje slova při psaní. Funkce historie slov extrahuje slova ze zdrojového segmentu a pamatuje si je spolu s jejich kontextem. V záložce nastavení najdeme tzv. segment timer (v překladu časovač segmentu). Časovač automatizuje proces kontroly překladu a zároveň projekt automaticky ukládá. Další možností nastavení je Google Translate. Vybereme-li tuto možnost, CafeTran přidá do pracovních postupů on-line službu Google Translate.

Wordfast je počítačem podporovaný program, jehož hlavním cílem je pomoc s překladem dokumentů. Wordfast spojuje technologii segmentace a překladové paměti. Program pracuje v rámci MS-Word, ale i MS Office jako je Excel a PowerPoint. HTML (značkovací jazyk pro hypertext) lze také přeložit. Wordfast nabízí pokročilé funkce terminologie. Najdeme zde odkazy na webové terminologické databáze. Neomezený počet uživatelů může sdílet stejnou překladovou paměť. Nástroje a rozšíření funkcí umožňuje

pokročilým uživatelům provádět složité úkoly, optimalizovat překladové paměti, analyzovat projekty nebo sledovat používání terminologie mezi různými překladateli. Překladová paměť Wordfastu je otevřená a snadno čitelná. Co se týče glosáře, program je schopen používat tři slovníky současně. Velikost slovníku je omezena na 250 000 záznamů. Wordfast podporuje všechny evropské, latinské jazyky, čínštinu, japonštinu, korejštinu, arabštinu, hebrejštinu, různé formy hindštiny a menšinové jazyky. Fuzzy shoda má výchozí hodnotu 75%, lze ji však nastavit od 50 do 99%. Hodnota nižší než 75% se nedoporučuje, protože pak software nabízí příliš mnoho shod.

Dalším překladatelským softwarem je **Déjà vu X2**, jehož nejznámější verze se nazývá Déjà vu X2 Professional. Tato verze má řadu kvalitních funkcí, mezi které patří DeepMiner. Tato funkce umožňuje získat více informací z terminologie a překladové paměti. Provádí náročné křížové analýzy těchto databází a vytváří vylepšený překlad. Další funkce Autowrite zvyšuje produktivitu textů, které obsahují jen minimální nebo žádné shody s databází. Autowrite pracuje přímo s naší databází, proto není nutné vytvářet speciální slovníky. Déjà vu X2 Professional je schopný provádět několik operací najednou u všech souborů. Tento software bere v úvahu kontext jednotlivých segmentů, proto se nemusíme starat o větší množství shod v překladové paměti. Déjà vu X2 Standard je cenově dostupná verze, která je vhodná pro překladatele, kteří začínají, a kteří pracují s krátkými překlady. Díky funkci SmartView si můžeme přizpůsobit okno, jak potřebujeme. Nechybí zde ani průvodce, který nám poradí, jak snadno vytvořit projekt a databázi.

SDLX nabízí produkt SDLX Professional, který je nejčastěji využíván ve velkých firmách a agenturách a produkt SDLX Standard určený pro překladatele na volné noze. Zdrojové a cílové soubory jsou v tomto programu uspořádány do projektů. Text ve zdrojovém souboru je rozdělen na segmenty, tedy na věty. Výsledný překlad je uložen v souboru s názvem cílový soubor. Překladová paměť se skládá z frází a vět ve výchozím jazyce. Každý segment (věta) má odpovídající překlad do cílového jazyka. Jestliže se vyskytne nový zdrojový soubor, všechny segmenty vyskytující se v překladové paměti jsou přeloženy. Uživatel může překlad přijmout nebo ho změnit.

SDL Trados Studio 2011 Professional je překladatelský software, který poskytuje nástroje pro vytvoření, editaci a kontrolu kvalitních překladů v krátkém čase. Je založen na technologii SDL (Structured Content Technologies) neboli technologie se strukturovaným obsahem. SDL software používá více než 80% překladatelů, rychlost překladů projektů je

rychlejší o 40%, což samozřejmě záleží na typu překladu. Tato technologie zajišťuje opětovné použití již přeloženého textu, proto nebudeme muset přeložit stejnou větu dvakrát. Mezi inovativní funkce patří například PerfectMatch (100% shoda), AutoSuggest (automatická návrh) nebo QuickPlace (rychlé umístění). Překladačské projekty mají různé formáty, SDL Trados Studio 2011 Professional obsahuje soubor filtrů, díky kterým může přijmout celou řadu typů dokumentů. Uživatel se tak nemusí starat o kompatibilitu.

Co se týče výhledů do budoucna, známá americká společnost Microsoft pracuje na hlasovém překladači, který bude mluvit jako člověk. Jedná se o program, který převede mluvenou řeč z jednoho jazyka do druhého a napodobí hlas mluvčího. Virtuální překladač by měl být schopen kromě překladu i simulovat hlavu a mimiku mluvčího. Podstatou bude zkombinování systému rozpoznávání řeči se strojovým překladačem a hlasovým syntezátorem. Problémem pak bude fáze překladu z jednoho jazyka do druhého. O podobné systémy se pokouší i jiné firmy, stroje jsou však stále pozadu za živými překladači a domnívám se, že i v budoucnu budou existovat oblasti, ve kterých nebude možné člověka v roli překladače nahradit sebedokonalejším počítačem.

ZÁVĚR

Cílem mé bakalářské práce bylo zmapovat historii a vývoj strojového překladu. Myslím si, že se mi podařilo vytvořit historický přehled strojového překladu, který má kořeny ve 40. letech minulého století. Zabývala jsem se obdobím 2. světové války, dále pak 50. a 60. léty 20. století, která jsou charakteristická rozsáhlými výzkumy v oblasti strojového překladu. Následně jsem věnovala pozornost 70. a 80. létům 20. století, pro která jsou typické první funkční systémy. Přelom 20. a 21. století pak charakterizuje vývoj nových informačních technologií a vznikají první CAT nástroje. Došla jsem k závěru, že překlad, který je specifický tím, že je při něm používán počítač, se neobejde bez teorie, která studuje problematiku strojového překladu. Dále jsem zjistila, že základní podmínkou pro překlad je analýza textu, a tedy i práce se slovníkem. Vytvořila jsem přehled výhod a nevýhod knižních a elektronických slovníků. Podařilo se mi vyhledat termíny, které je nutné znát, abychom mohli pracovat s překladatelským softwarem. Zjistila jsem, že překlad není jenom jazyková operace, ale i technická. Proto musí mít uživatel překladatelského softwaru technické schopnosti.

Při zkoumání jednotlivých překladatelských softwarů jsem došla k závěru, že základ mají všechny stejný, fungují na stejném principu. To znamená, že mají překladovou paměť a terminologickou databázi. Liší se pak v jednotlivých funkcích. Například Transit NXT má zabudovanou morfologickou analýzu, díky které je schopen vyhledat všechny výrazy. OmegaT má několik vyhledávacích funkcí, které je potřeba pouze aktivovat. Mezi užitečné funkce CafeTranu patří automatické dokončení, které automaticky doplňuje slova při psaní. Překladatelské softwary se dále liší cenou nebo velikostí slovníku a ne všechny programy podporují všechny jazyky.

Překladatelské softwary, kterých je dnes na trhu nespočet, velice usnadňují práci překladatele. Je pravděpodobné, že se v budoucnu budu zabývat překlady, a proto pro mě bylo velice poučné zjistit, jak tyto softwary fungují.

Dílčím cílem mé bakalářské práce byl pohled na možnosti překladatelských softwarů. Vyhledala jsem informace o jednotlivých programech a vzájemně je porovnávala, proto se domnívám, že splnit cíl pohled na možnosti překladatelských softwarů se také podařilo. Tato bakalářská práce může posloužit jako zdroj informací jak studentům překladatelských oborů, tak profesionálním překladatelům.

РЕЗЮМЕ

Темой моей дипломной работы является перевод, поддерживаемый компьютером- история и настоящее время. Речь идет о переводческих системах программного обеспечения, которые облегчают работу переводчика. Я выбрала переводческие программы согласно критериям- наиболее используемые и самые известные. Я выбрала эту тему, потому что в будущем я бы хотела заниматься карьерой переводчика и эта дипломная работа предоставила мне возможность посмотреть на альтернативы переводческих систем программного обеспечения.

Сначала я занималась историей машинного перевода, связанного с такими именами как француз **Ж. Арцуни** или советский ученый **П. С. Троянский**. Оба этих ученых изобрели транслятор на принципе бумажной перфорированной ленты. Следующим важным событием было изобретение электронного компьютера во время второй мировой войны. Американец **У. Уивери** и англичанин **А. Д. Бут** совместно предложили использование компьютера для перевода. Так называемым меморандумом Уивера стартовало в 1949 г. исследование в сфере машинного перевода.

В дальнейшем я занималась качеством машинных переводов, первые попытки которых были буквальными, а соответственно могли быть не качественными. Расходы на машинный перевод были в два раза выше расходов на перевод человеком. Сообщение специальной правительственной комиссии **ALPAC** (Automatic Language Processing Advisory Committee) повлекло за собой ослабление больших размеров во всемирном масштабе. Позднее ситуация меняется и машинный перевод опять находится на этапе развития. Автоматические системы для перевода служат теперь как быстрое средство для приблизительного перевода. Такие переводы использовали транснациональные корпорации, институты и организации. Позднее появляются микрокомпьютеры и для них возникают первые **программные средства КЭТ** (CAT= computeraidedtranslation), пособия доступные для более широкой общественности переводчиков. Мне удивило, сколько денег было истрачено за исследования. Наибольшие инвестиции в исследование вложила Япония, около 200 миллионов долларов.

Что касается современности, акцент кладут на невысокие расходы, высокую скорость и немедленную реакцию. Интенсивное развитие рынка принесло с собой дальнейшее развитие систем машинного перевода. В Москве это была, например, система под названием **PROMT**. После прочтения необходимой литературы я узнала, что к тому, чтобы машинный перевод был надлежащим образом употреблен,

необходима также и теория, изучающая проблематику машинного перевода. Я упомянула таких авторов работ, как **Н. Д. Андреев**, **И. Бар-Хиллел** или **И. А. Мельчук**. Также я вспоминала и известный эксперимент машинного перевода с русского языка на английский в Нью Йорке в Джорджтаунском университете в 1954 г.

В следующей главе я занималась сущностью перевода, поддерживаемого компьютером. Сокращенное слово КЭТ обозначает перевод, к которому мы используем софт, так называемые КЭТ программные средства. Самыми известными КЭТ системами программного обеспечения являются **DéjàVu**, **Transit**, **Trados** или **OmegaT**. Более подробно я буду заниматься ими в конце работы. Сущностью этих КЭТ программных средств является работа с переводной памятью. Текст разделен на сегменты (предложения), после того как мы делаем перевод, он записан в базу данных. Этот сегмент, касающийся источника и его перевода, называется переводной единицей. Если данная переводная единица покажется в переводимом тексте снова, софт предложит переводчику тот вариант, который был записан ранее. Эти переводческие софты служат, прежде всего, к переводу специальных текстов, где часто появляются специальные термины. В дальнейшем я занималась проблематикой программ. Так как компьютер работает механически, он кладет акцент на значение или контекст. Поэтому это очевидно, что софт не работает за нас, но только помогает.

В следующей части моей работы я занималась терминами, которыми необходимо владеть, для работы с программой. Я упомянула сегмент, лингвистическую пару, **perfectmatch** (полное совпадение), **fuzzymatch** (частичное совпадение) или конкордацию. Чтобы объяснить эти термины, мне понадобилось воспользоваться журналом ТоП (устный и письменный переводы).

Перевод нужно всегда проверять. КЭТ программы имеют встроенные корректоры правописания, которые контролируют опечатки и другие различные ошибки. Однако эта корректура не исправляет грамматических и семантических ошибок.

В следующей главе я исходила из публикации Даниелы Мюгловой «Коммуникация, устный перевод, письменный перевод или почему упала Вавилонская башня?» Автор здесь употребляет символику библейской повести о Вавилонской башни. Вавилонская башня была, согласно библейской книге Бытия, огромным зданием, которое должно было достигать неба. Оно должно было быть построено на месте, именованном напоследок Бабел, это еврейское название для Вавилона. Согласно

Библии, после всемирного потопа все люди были единомышленны и говорили на одном языке. Люди решили построить город и башню до неба, чтобы возвеличить род человеческий. Башня была построена, чтобы восславить человека, а не Бога. Бог тогда разделил до сих пор единый человеческий язык на несколько языков и разогнал людей по всей земле, так как он не хотел, чтобы собор принадлежал кому-то кроме него. Значение этой легенды она переносит на современные мультикультурные отношения, которым без перевода не обойтись. По мнению автора, перевод является не только языковой операцией, но и технической. В этой главе я занималась различными средствами работы переводчика. К ним относится работа со словарем, управление редактором, касающимся текста или графических файлов.

Кроме использования чешских источников я занималась и русскими публикациями, к которым относятся «Математические вопросы кибернетики» О. С. Кулагиной и другие. Главы, которые меня заинтересовали, впоследствии надо было переводить и обработать. Здесь я частично возвращаюсь к основам машинного перевода. Системы сначала занимались специфическими сферами перевода, художественная литература, к примеру, вообще не переводилась. Первоначально переводы осуществлялись с целью приобретения информации и только позднее это стало касаться расширения информации. Необходимой частью перевода является использование словаря. Практика перевода требует, чтобы одно слово было заменено другим, однако словарь предложит вам больше вариантов. В дальнейшем я занималась микроконтекстом и макроконтекстом. Микроконтекст обозначает, что словарь регистрирует не только слова, но и словосочетания. Макроконтекст обозначает, что одно слово может иметь два значения- буквальное и терминологическое.

Следующая глава называется анализ текста. Перед анализом текст должен быть лишен омонимии и многозначности, то есть, полифункциональности изолированной языковой формы. В отличие от полисемии, омонимия не имеет смысловой родственности. Очень частой является ситуация, когда в словаре не существует аналога для некоторых частей текста. Эти части текста либо являются неологизмами, либо реалиями. Неологизмы возникают из-за потребности наименовать новые вещи или заменить, до сих пор существующие наименования более точными. В рамках технической стороны данной работы я занималась и трансляторами, их конструкцией и структурой.

На конец я занялась отдельными переводческими софтверами. Сущностью этих программ является автоматическая сегментация, которая предоставляет обзорную последовательность. Нам не надо беспокоиться, что мы пропустим строку или некоторую часть текста. Исходный и целевой текст находятся рядом, значит, мы не должны перескакивать между разными окнами. К важным критериям перевода специальных текстов относится высокая степень точности. Она предоставляет возможность повторения предложений, фраз и терминов. Главной функцией переводческих софтверов является переводная память, которая служит к быстрому решению уже однажды реализованных переводческих единиц. Эту функцию мы оценим больше с увеличивающимся количеством переводов. Принцип этих программ основан на предложениях исходного языка. Они разделены в сегментах на переводные единицы. Во время перевода находится аналога в переводной памяти. Переводчик либо примет предложенную переводную единицу, либо он поправит ее, и возникнут предложения целевого языка. Целью моей работы было взглянуть на возможности переводческих программ. Системы программного обеспечения можно разделить по финансовым критериям. Я приведу несколько примеров для интереса. Лицензия пользователя **Wordfast 4.20** фирмы **Champollion** стоит 90 евро, лицензия пользователя на **Translator's Workbench 6,5** фирмы **Trados** стоит 695 евро, лицензия пользователя на **DéjàVu X** фирмы **Atril** стоит 490 евро. Чешский переводческий софтвер **Transit XV** фирмы **Star** стоит 24 990 чешских крон.

Переводческие софты или КЭТ программные средства я выбрала самые известные. **Déjà vu, Trados, Transit, Wordfast, Cafetran, OmegaT** и **SDLX**. Программы **Wordfast** развивает частная компания **Wordfast LLC**, находящаяся в Соединенных Штатах Америки. Первая версия появилась в 1999 г. В то время на рынке находилось лишь несколько софтверных пакетов с переводной памятью. В настоящее время **Wordfast** употребляет более чем 25 000 переводчиков по всему миру практически во всех языках. Достоинствами этой программы являются простота и несложное управление.

SDL Trados был основан в Штутгарте в Германии в 1984 г. В его основании принимали участие **Йохен Хюммел** и **Ико Книппхаузен**. Этот переводческий софтвер в настоящее время употребляет более чем 185 000 профессиональных переводчиков. После 25 лет исследований и развития была издана версия **SDL Trados Studio 2009**, которая способствовала появлению последующего поколения систем программного

обеспечения переводной памяти. Затем в 2010 г. фирма Trados разработала новую версию **SDL Passolo 2011**.

CafeTran является программой со многими функциями, его управление довольно простое и перевод поддерживается онлайн переводческими системами. Дизайн разработан таким образом, чтобы было возможно быстро работать. Функции в программе **CafeTran** экономят время и работа становится более продуктивной.

Коренным автором программы **OmegaT** является **Кэйт Годфри (Keith Godfrey)**. Он относится также к изобретателям OmegaT. Софтверные программы, использующие переводную память эти люди поддерживают с целью развития и дистрибуции. К другим изобретателям относятся, например, **Золтан Бартко (ZoltanBartko)**, **Дидье Брейль (Didier Briel)** или **Ким Брюнинг (Kim Bruning)**.

Марк Ланкастер (Mark Lancaster) является основателем и директором LISA (LocalizationIndustryStandardsAssociation), записанной в реестр SDL (провайдер софтвера и услуг для управления содержанием перевода) в 1992 г. после того, как было обращено внимание на необходимость высшего уровня услуг для глобализации софта. SDL предлагает комплексное многоязычное решение для переводческих технологий и переводной памяти. Оно сохраняет свою резиденцию в Великобритании, но широко применяется в Азии, Европе и Северной Америке.

Во время исследования отдельных переводческих софтов, я дошла к заключению, что они имеют одинаковую основу и функционируют на тождественном принципе. Затем они отличаются отдельными функциями. Например, у Transit NXT есть встроенный морфологический анализ, благодаря которому он способен отыскивать все выражения. OmegaT имеет несколько отыскивающих функций, которые нужно только активировать. К полезным функциям программы CafeTran относится автоматическое завершение, которое автоматически дополняет слова во время написания. Переводческие системы программного обеспечения в дальнейшем отличаются ценой или размером словаря, и не все программы поддерживают все языки.

Составной частью моей работы является словарь терминов и сокращенных слов, фигурирующих в области переводческих софтов. Этот словарь я составила после прочтения самых разнообразных рецензий на них. Далее словарь содержит термины, которые появляются в моей работе. В приложении вы найдете картинки с темой «перевод, поддерживаемый компьютером».

BIBLIOGRAFIE

Knižní zdroje

1. Karlík P., Nekula M., Pleskalová J. 2002: Encyklopedický slovník češtiny. Nakladatelství Lidové noviny, Praha.
2. Koukolíková, J. 1990: Aplikace prologu v konstrukci překladače. In: Překladače programovacích jazyků. České vysoké učení technické, Praha.
3. Král, P. 2005: Český překlad II (1945-2004). Univerzita Karlova, Praha.
4. Lotko, E. 2003: Slovník lingvistických termínů pro filology. Univerzita Palackého v Olomouci, Olomouc.
5. Melichar, B., Češka, M., Ježek, K., Richta, K. 1999: Konstrukce překladačů I. Část. Vydavatelství ČVUT, Praha.
6. Mügelová, D. 2009: Komunikácia, tlmočenie, preklad alebo prečo spadla Babylonská veža. Enigma Publishing, Bratislava.
7. Nádběla, J. 2004: Velký počítačový slovník. Computer Media s.r.o., Kralice na Hané.
8. Pilař, A. 1999: Počítačové zkratky. Grada Publishing, Praha.
9. Кулагина, О.С. 1991: О современном состоянии машинного перевода. In: Математические вопросы кибернетики. Наука, Москва.
10. Нелюбин, Л. И. 2009: Введение в технику перевода. Издательство Флинта, Наука, Москва.
11. Ревзин, И. И., Ровзенцевейг, В. Ю. 1964: Некоторые специальные вопросы теории общего и машинного перевода. In: Основы общего машинного перевода, Издательство Высшая школа, Москва.

Periodikum

1. Vedral, J. 2010: Základní pojmy používané při překladech pomocí CAT nástrojů. In: Časopis ToP 96.
2. Ďuricová, A. 2011: Daniela Mügelová: Komunikácia, tlmočenie, preklad alebo prečo spadla Babylonská veža? In: Časopis ToP 102. (Recenze)

Elektronické zdroje

1. CAT. [online]. [cit. 2012-01-21]. Dostupné z: <<http://cs.wikipedia.org/wiki/CAT>>

2. CafeTran. [online]. [cit. 2012-03-04]. Dostupné z: <<http://www.cafetran.republika.pl/handbook.html>>
3. Déjà vu. [online]. [cit. 2012-03-10]. Dostupné z: <<http://atril.com/en/software/deja-vu-x-standard>>
4. Interlingua. [online]. [cit. 2012-2012-01-16]. Dostupné z: <<http://cs.wikipedia.org/wiki/Interlingua>>
5. Neuronová síť. [online]. [cit. 2012-01-16]. Dostupné z: <http://cs.wikipedia.org/wiki/Neuronov%C3%A1_s%C3%AD%C5%A5>
6. OmegaT. [online]. [cit. 2012-03-18]. Dostupné z: <<http://www.omegat.org/cz/omegat.html>>
7. Trados. [online]. [cit. 2012-03-18]. Dostupné z: <<http://www.translationzone.com/en/translator-products/sdl-trados-studio-professional/>>
8. Transit NXT. [online]. [cit. 2012-02-25]. Dostupné z: <<http://www.star-transit.cz/Software/Transit-prekladova-pamet-a-fuzzy-logika/>>
9. Wordfast. [online]. [cit. 2012-02-11]. Dostupné z: <<http://www.wordfast.net/index.php?whichpage=products&lang=engb>>
10. Zpracování přirozeného jazyka. [online]. [cit. 2012-01-21]. Dostupné z: <http://cs.wikipedia.org/wiki/Zpracov%C3%A1n%C3%AD_p%C5%99irozen%C3%A9ho_jazyka>
11. Демос, К., Фраунфельдер, М. История машинного перевода. [online]. [cit. 2011-01-16]. Dostupné z: <http://www.dialog-21.ru/trends/?id=1744&forum_id=13&f=1>

PŘÍLOHA 1

Slovníček zkratk:

Zkratka	Plný název	Český ekvivalent
CAT	Computer Aided Translation	Překlad podporovaný počítačem
FAMT	Fully Automated Machine Translation	Plně automatický počítačový překlad
GIF	Graphics Interchange Format	Grafický formát
HAMT	Human Assisted Machine Translation	Počítačový překlad s lidskou asistencí
HTML	HyperText Markup Language	Programovací jazyk pro vytváření hypertextových dokumentů
JPEG	Joint Photographic Experts Group	Komprimační algoritmus pro kompresi grafických dat
MAHT	Machine Assisted Human Translation	Překlad za pomoci počítače.
MT	Machine Translation	Strojový překlad
NLP	Natural Language Processing	Počítačové zpracování přirozeného jazyka
PM	Perfect Match	Dokonalá shoda

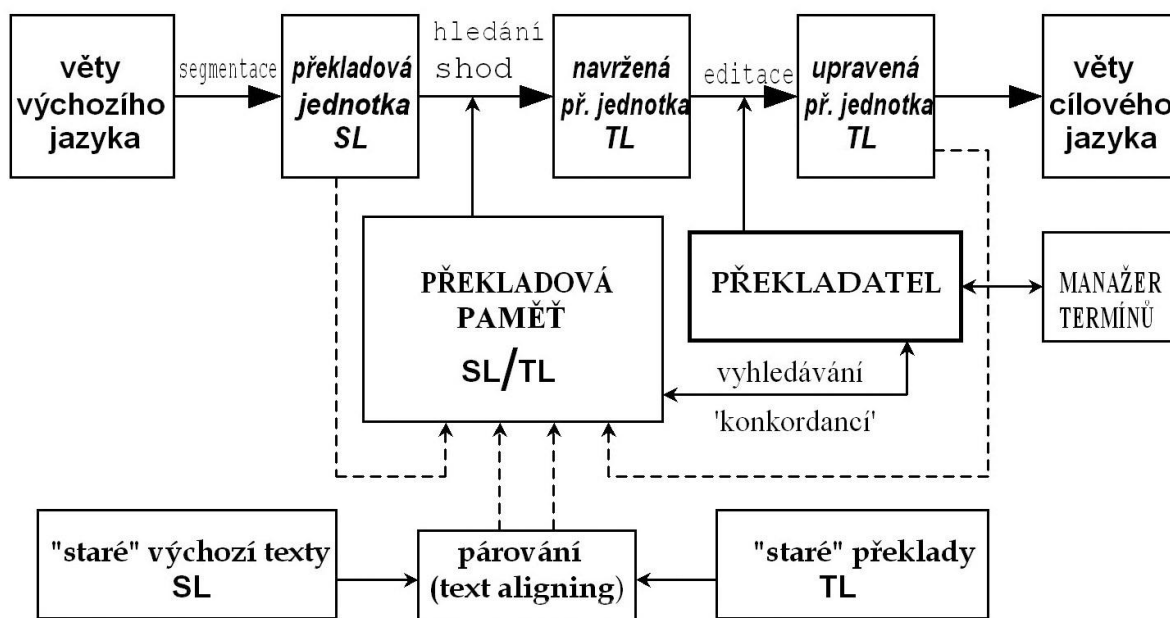
PŘÍLOHA 2

Slovníček termínů:

binární data	Binární neboli dvojková soustava, která vyjadřuje data prostřednictvím jedniček a nul.
bitmapová grafika	Obrazová informace je uložena jako množina bodů. Každý obrazový bod (pixel) má nastavenou barvu. Obvykle se používá komprese typu JPEG nebo GIF.
formát	Struktura souboru vyjádřená v konečném důsledku souborovou příponou jako prvotní znak pro uživatele počítače. Tato přípona je reprezentována tzv. ikonou.
fuzzy match	Technika používaná v CAT, pracuje se shodami, které nejsou stoprocentní.
Java	Objektový programovací jazyk nezávislý na jazykové platformě. Je určen pro vytváření síťových aplikací.
konkordance	Funkce CAT programu, která prohledá celou překladovou paměť a zobrazí všechny možnosti překladu.
rastrová grafika	viz bitmapová grafika
textový editor	Program pro vytvoření, editaci a grafickou úpravu textu, příkladem je Microsoft Word.
vektorová grafika	Grafika, která využívá k definování obrázků matematicky definované vektory, prvky jsou popsány jako celek a výsledný soubor je menší.
zdrojový text programu	Výchozí tvar programu, kdy jsou jednotlivé požadované činnosti popsány v textové formě pomocí programovacího jazyka.

PŘÍLOHA 3

Princip překladové paměti



ANOTACE

Příjmení a jméno autora: Balcarová Eliška

Název katedry a fakulty: Katedra slavistiky, Filozofická fakulta Univerzity Palackého v Olomouci

Název bakalářské práce: Počítačem podporovaný překlad- historie a současnost

Vedoucí práce: doc. PhDr. Zdeňka Vychodilová, CSc.

Počet znaků: 76 746

Počet příloh: 3 (2 082 znaků)

Počet titulů použité literatury: 11 (+ 2 periodika + 11 internetových zdrojů)

Klíčová slova: strojový překlad, historie strojového překladu, teorie strojového překladu, počítačem podporovaný překlad, počítačová terminologie, překladače, CAT nástroje

Charakteristika:

Hlavním cílem této bakalářské práce s názvem „Počítačem podporovaný překlad- historie a současnost“ je pohled na možnosti překladatelských softwarů. Práce je rozdělena na historii a současnost. V historii se věnuji nejdůležitějším průkopníkům strojového překladu, výzkumům a experimentům. Píši zde o prvních systémech, kdo je používal a za jakým účelem. Co se týká současnosti, vysvětlila jsem podstatu CAT nástrojů, proč jsou užitečné a za jakých podmínek je dobré je využít. Zabývala jsem se nejznámějšími překladatelskými softwary, snažila jsem se pohlédnout na jejich možnosti. Součástí mé práce je slovníček zkratk a termínů z počítačové oblasti.