**UNIVERZITA PARDUBICE**
**FAKULTA FILOZOFICKÁ**

# BAKALÁŘSKÁ PRÁCE

**2008** **Bc. Kateřina Haššáková**

**University of Pardubice**
**Faculty of Arts and Philosophy**

# CAT-Tools – modern technologies in translating

**Kateřina Haššáková**

**Bachelor Paper**
**2008**

Univerzita Pardubice
Fakulta filozofická
Katedra anglistiky a amerikanistiky
Akademický rok: **2006/2007**

# ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Kateřina HAŠŠÁKOVÁ**

Studijní program: **B7310 Filologie**

Studijní obor: **Anglický jazyk pro hospodářskou praxi**

Název tématu: **CAT-Tools - modern technologies in translating**

Zásady pro vypracování:

Tato bakalářská práce bude zaměřena na zkoumání využívání moderních technologií v současném odborném překladatelství a zaměří se především na využívání CAT (computer assisted translation), tj. software pro počítačem podporovaný překlad.
V první části práce bude autorka definovat CAT nástroje, provede jejich srovnávací analýzu, ohodnotí jejich efektivitu z hlediska funkčních (odborných) stylů, překladatelských teorií a postupů a bude charakterizovat cílového uživatele těchto nástrojů.
Ve druhé části bude proveden výzkum jehož předmětem bude zkoumání využívání CAT nástrojů překladateli.
Autorka bude ve své práci využívat jak sekundární zdroje z oblasti moderních technologií, teorie překladu, stylistiky, lingvistiky a metodologie výzkumu tak zdroje primární, tj. recenze CAT nástrojů provedené překladateli, primární data získaná od vybraných překladatelů, atd..

**Declaration of authorship:**

I,…….., declare that the present bachelor paper and the work presented in it are my own.

I further declare that where I have consulted work of others, this is always clearly attributed. Where I have quoted from the work of others, the quotation is always marked and the source given. With the exception of such quotations, this paper is entirely my own work.

I understand that any attempt at plagiarism in this work - violation of this declaration – may result in further disciplinary actions against me.

I understand that this declaration of authorship must be read, understood and signed in order for this paper to be accepted.

Date: Signed:

## Abstract

The aim of the paper is to introduce CAT-Tools. It is particularly focused on MT and MAHT. It is also to identify the distinctive features of MT and MAHT and to describe how these systems work with different text complexity and regularity. Moreover it is to show whether MT and MAHT are suitable for translation of technical text. Lastly, software used in the Czech Republic is introduced together with the final user of CAT-Tool systems.



## Abstrakt

Cílem této práce je představení CAT-Tool nástrojů se zaměřením převážně na MT (automatický překlad) a na MAHT (počítačem podporovaný překlad). Zároveň definuje klíčové rozdíly mezi MT a MAHT a popisuje, jak tyto systémy pracují s texty, které se od sebe liší obtížností a pravidelností jednotlivých znaků. Dále se zabývá vhodností MT a MAHT pro překlad odborných textů. V závěru práce jsou představeny softwary, které se používají v České republice společně s definováním konečného uživatele těchto softwarů.

**Table of contents**

# Introduction

Translating is one of the oldest ‘crafts’ worldwide. The first written record is nearly as old as the writing itself. As an example the multilingual signs which were made by king Sargon 3 000 B.C. can be taken, also king Chammuraphy’s edicts which were published in more than one languages at the beginning of 2000 B.C., and last but not least well preserved cuneiform boards that contain something like primitive dictionary.

As every craft has been a subject to development during the centuries, it has not been different with the development of translating process. Translation was mentioned in the works of the Old Greek authors as was for instance philosopher Marcus Tulius Cicero (106-43 B.C.). Starting with these ancient authors over the Sophronius Eusebius Hieronymus (around 348-420 A.C. - Saint Jeronimus), to Richard F. Burton (1821-1890) and Matthew Arnold (1822-1888), who were considering the questions of translation of the Bible into the modern languages. If we should continue and include all the famous philosophers or authors of famous literature, who were directly or indirectly involved in the translation ‘craft’ or mentioned it in their works, the list of the names would be quite long as the ages passed along.

Translation as a process has been defined in many pieces of literature or educational works. In our paper we understand the translation as:

> What is generally understood as a translation involves the rendering of source language (SL) text into the target language (TL)[1] so as to ensure that the surface of meaning of the two will be approximately similar and the structure of the SL will be preserved as closely as possible but not so closely that the TL structures will be seriously distorted. (Bassnnett 1991 in Vondráčková 2007, p. 4)

The process of translation is very clearly and exactly defined, but what cannot be found in the definition is that the translation is most of all a piece of art. The human translators are not only blindly copying what was once written in one language into another language. To create a perfect translation requires the ability of putting ‘the heart and the soul’ into it.

---

-[1] In this paper by the target language we understand Czech language.

Therefore, it can be seen as unbelievable that the 'art of translation', which has been developing for such a long time, could be connected with something that exact as computer. The combination of using technical device in the human craft of translation is still in the process of expanding. Firstly, it was Machine Translation (MT) which was developed at the early 20th century and few decades later Machine Assisted Human Translation (MAHT) joint the group of what is nowadays called CAT-Tools[2] (computer assisted translation).

The aim of this paper is to evaluate the role of MT and MAHT in the translation process, if they can help and support translators' effort. Moreover the crucial distinctions between MT and MAHT are provided to show how these two systems work with particular complexity and regularity of text. The objective of the paper is also to find out whether MT or MAHT can be considered as suitable systems for the translation of the technical (scientific) text. Besides that, the comparison of software used in the Czech Republic is introduced and the final user is defined.

In the first chapter the characteristics of the technical translation together with the distinctive features between technical and literal translation is defined. Secondly, the definition of CAT-Tools and its brief history is provided including MT and MAHT description, their utilization and suitability. As well as the MT the MAHT systems are described in another subchapter of CAT-Tools and also their use and suitability for the translation is introduced. The examples that are provided for the better understanding are based on the language pair English-Czech as the survey was held in the Czech Republic and is based on the needs of Czech translators. The third and the last chapter provides the results of the survey and data interpretation together with evaluation of advantages and disadvantages of software used in the Czech Republic.

[2] The full name of CAT-Tools is no longer used among translators. This approach is also followed in this paper.

# 1. Characteristics of the Technical Translation

This chapter gives us a brief outline of the language functions, text types and functional style which are considered to belong among the most important features investigated in the process of translation. The technical and literal translations differ therefore, it is need to identify the crucial and distinctive features of each of them to show which type of translation is more suitable for CAT tools. As they can be used in written texts we concentrate on this type of texts only.

## *1.1 Language Functions, Text Types and Functional styles*

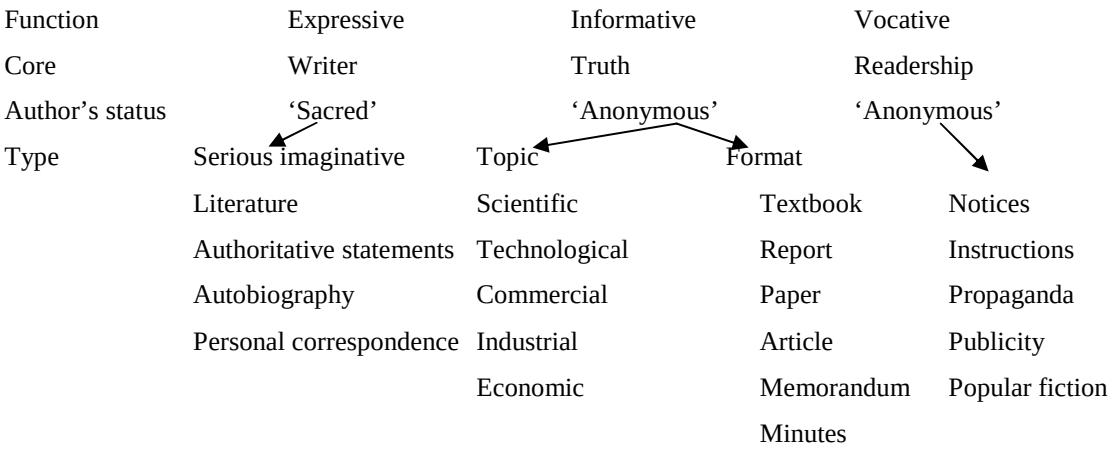
Language functions are firmly bound together with text types and functional styles. According to Newmark(1988, p. 30) all translations are based implicitly on a theory of language. In his Textbook of translation he uses the functional theory of language by Buhler which was later adapted by Jacobson and which is the most suitable for translating. The essential function of the language is communicative. It is to express feelings, ideas and emotions. It enables the speaker tell stories and remember past. Buhler's theory concerns three main functions of language. Jacobson (in Newmark 1988, p. 42) adds three others: aesthetic, phatic and metalingual. However, for our purpose the three firstly mentioned are discussed in detail since they corresponds to the certain text types which are concerned to be proper for CAT-Tools.

1) Expressive
2) Informative
3) Vocative

Expressive (emotive) function enables speakers to express their emotions and attitudes through linguistic means paralinguistic features or non-verbal communication

Informative function allows speakers present new information. Language used represents reality or refers to it.

Vocative (directive) function is used by speakers to express the attitude they want to cause in the addressee. The utterance should make the hearer act.




In the scheme above are shown the language functions together with their particular types of the text. It means that the language devices that are used are characteristic for each text type. The user's manual is written differently from Walter Scott's novel Ivanhoe. The most important thing is to distinguish, what function and aim the text have. Knittlová (1990, p. 7) states that the main functional styles were formed according to the functions of the language. Every style is characteristic by its language devices.

- Administrative style concerns formal language and can be further divided into a] commercial/business documents and correspondence. The purpose of these is to inform or persuade, b] legal documents where the function is directive and informative. They are exact and convey precise meaning.
- The purpose of popular scientific style is to inform. As an example textbooks and manuals can be presented
- The aim of publicistic style is to inform, instruct or persuade. It can be presented as a mixture of the styles (newspapers, magazines and articles can be presented).
- Scientific style concerns the communication of scientific knowledge (textbooks, research, scientific studies).
- Colloquial style is simple communicative style
- The purpose of block language is to put maximum information into minimum space

- Literary/belletristic style arouses emotions and the purpose is to entertain. It can be again mixtures of the styles. The language is figurative (fiction, epics, lyrics, drama).

As looking closer into the language functions, text types and functional styles, it is seen that all of them are in connection. According to the Newmark (1988, p. 42) only few texts are purely expressive, informative or vocative. Most of them include all three functions. However, strictly, the expressive function has no place in a vocative or informative text. If it is present, it is there only unconsciously. Most informative texts have vocative function that is restricted to a separate section of recommendation, opinion or value judgment. Cognitive means of each functional style are different, therefore the first essential step for any translator is to identify them and, based on the identification, to distinguish the style, function and purpose of the text that is being translated.

## *1.2 Technical versus Literary Translation*

We provide only brief outline of follows since it is not the purpose of the paper to describe the differences between them. As the styles, texts, and the language that is used in the technical and literary translations differ, also the approach to the translation of each of them is not the same. Kufnerová (2003, p. 25) states that according to stylistic features of the original text, these text and translation of them can be divided into two main branches: technical (scientific) and literary translation. Generally, the difference between technical and literary translation is based on the language function predominating in the original text. In the technical translation the language is subject matter oriented, however, in the literary translation the language has mainly the expressive value. The core of the technical style is in the semantic use of vocabulary. The words have to be understood as they are in the dictionary. Morphemes, words, phrases or sentences in the technical text express the general and conventional (denotative) meaning conveyed by words. The register used within the text is highly specific without any kind of expressivity. The terminology used in the technical text is highly repetitive either within the single document or across the documents. The pragmatic level which, '…concerns the meaning, not as generated by the linguistic

system but as conveyed and manipulated by participants in a communicative situation…'(Baker1992, p. 217) is excluded.

However, the literary texts are full of personal components and figurative language: i.e. original metaphors, language puns, unusual collocations, words with connotative meaning (Newmark 1988, p. 40). The involvement of the personal touch of the writer is crucial. But not only register and the use of the vocabulary are different. There are plenty of syntactical differences as well.

The technical text is significant with the use of passives, present and present perfect tenses, imperative structure of sentences with the use of second person singular, cohesive devices such as conjuncts, e.g. first, second (Newmark 1988, p. 40). Also the format is very specific, e.g. paragraphing, numbering, spacing, schemes. In technical texts the already known information is logically followed by the new information which is the most important one.

The literary text can have unconventional syntax, phatic structures, direct speech. All these make the text 'vivid'.

The differences between the technical and literary translations are great. It starts with the particular use of language for each of them. It is specific terminology versus figurative language. It continuous over the sentence structure that is also different. It is mostly passive versus active. The layout is also different. It is picture of the mountains versus scheme of the stock rates.

The translation of technical text is a translation of a material reality. It goes over the diagrams, graphs, schemes figures, illustrations and references. It is the field where the translator has to be always up to date because the technology development is still in progress and the value of the information is priceless.

# 2. CAT-Tools (Computer Assisted/Aided Translation)

According to Feder (2001, p. 49 in Palacz 2003, p. 6) the term Computer Assisted Translation is a general term, which defines any kind of involvement of technical device (e.g. electronic dictionaries) in the process of translating. The very important thing is to take in consideration the word *assisted or aided* and the distinction between the two terms, which are related and are often interchanged, should be stated at this point. MT (Machine translation) and MAHT (Machine Assisted Human Translation) are two different systems, they work with different levels of text complexity and the outcome is also different.

## 2.1 Brief History of CAT- Tools

With the history of MT two very important names have to be mentioned. George Artsrouni and Petr Smirnov-Troyanskii were two scientists who were, at the beginning of 1930s, working independently on the research of using the computers for the purpose of translation. Unfortunately their work together with the results remained unknown up till 1950s. (Craciuescu at.al 2004, p. 8)

The beginning of the real boom can be marked by the year 1949 when Warren Weaver publicized his ideas concerning the involvement of the computer into the process of translation. His ideas were that interesting and a number of Universities in US started their own researches in the field of machine translation. Nowadays very well known company IBM together with Georgtown University developed a trial version which was, after the demonstration, considered as successful. The machine translation software which was developed after was based on bilingual dictionaries together with rules for the word order. Since the syntactic and semantic functions of the language are fields which are complex, developers became disappointed and sceptical. Nevertheless IBM introduced MARK II with the hope that this system would be able to cover all already mentioned points, however it was again another blind shot (Hutchinson 1996, p. 2). Since the filed of machine translation became interesting for many and it was in connection with insufficiency of each system which was developed, US Government introduced the ALPAC (Automatic Language Processing Advisory Committee) in 1966

to evaluate every single system. ALPAC (1966, p.32) published a highly critical report concerning the machine translation systems. Although the inefficiency, slowness and expensiveness of these were mentioned on the first place of the report the need for the computer assistance in the field of translation was stressed. However the negative critics caused that all researches and development concerning this theme in US were stopped. Fortunately in other countries it continued and two systems were 'born' German SYSTRAN which was used EU Commission and Canadian TAUM-METEO for translation of weather forecast from French to English (Craciuescu at. al 2004, p. 9).

From 1980s over the 1990s the need for the machine translation increased stimulated by the boom of business. France, Germany, Canada and Japan were countries which continued in development of new machine translation systems such as Logos (language pair - German to French and vice versa). Also the personal computers became more accessible, companies such as ALPS, Weidner, Globalink (North America and Europe), Sharp, NEC, Mitsubishi, Sanyo (Japan) needed these programs for their own businesses. Some of the most important projects were GETA-Ariane (Grenoble), SUSY (Saarbrücken), MU (Kyoto), and Eurotran (the European Union).

The approach to the research and development from the 1990s changed a bit. The strategy of the machine translation, before the mentioned year, was based only on grammatical rules. The approach after the year 1990s concerned translation based on the bodies of the text. The language is dynamic element that changes according to the use and the user. It is element that is influenced and bound according to the culture and socio-economical surrounding. Therefore the machine translation is still in the progress.

Since the MT research and development of a system that is available to provide Fully Automatic High Quality Translation where quite disappointing for the ALPAC in 1960s, the researchers focused on something that would be helpful in translation process. It is interesting that the failure of one was an impulse for the rise of the other one. The period 1992 to1994 was crucial for the MAHT when IBM and other major software companies released their first MAHT tools which gained the popularity among translators. One example for all can be Atril company with their Déjà Vu software (Craciuescu at. al 2004, p. 11).

It is necessary to state at the beginning what is included into the MAHT tools. MAHT tools are simply called workbench packages or integrated translation systems.

They are equipped with translation memory systems, terminology management, dictionary, and automatic translation facilities. It also includes so called authoring tools such as spell, grammar, style checkers, electronic dictionaries, encyclopedias, terminological database, email and internet services, and morphological analysis tools. MAHT tools posses the distinctive feature of integrating these tools resulting in certain synergy effect. (Feder 2001, p. 50 in Palacz 2003, p. 9) Another feature that was recently developed was working with a various number of file formats.

According to what is mentioned above it can be though that MAHT software is full of features and immediately the user buys it, it will without any obstacles provide a perfect translation. (Sandin 2006, p. 18) 'We should remember that all TM tools are "empty" when we take them out of the box. It is our own work that enables us to build up the database.'

This is the first hint for the translator to guess what the MAHT tools are. Simply it can be said that the MAHT tools work on the basis of the translators own works. The MAHT software is as good as the translator him/herself.

## *2.2 Machine Translation (MT)*

MT is a system which automatically translates from source language into target language without the intervention of the human translator. It does not mean that after the translation is done it is locked and no other editions are possible to do with it. Basically MT can be divided into two groups (Bologna 2004, p. 1):

- MT-enabled (Unassisted)
  produces automatic translation with no human post-editing, the translation is unpolished but can be useful for the translation of materials that are in huge volume therefore time consuming and would be for human translator practically impossible to translate them or due to the need of immediate turn-around requirements.
- MT-enhanced (Assisted)
  produces automatic translation with the intervene of translator.

### 2.2.1 MT approaches

With MT there are three approaches that can be classified as follows:

- the direct strategy
- the transfer strategy
- the pivot language strategy

#### 2.2.1.1 Direct Strategy

Craciunescu at al (2004, p. 21) say in their studies of MT that the direct strategy involves minimum of linguistics theory. The translation is done in the direct way from the source language to the target language. Every single word in the source language has its direct link which corresponds with the unit in the target language. The sentence in the source language is analyzed, broken down into small and easily translatable units and finally the sentence with the same structure is produced in the target language. MT works with huge dictionaries and corpora of texts that have been previously translated. Even if the system has highly developed mechanism for morphological analysis and the dictionary is complex the need of post edition is inevitable.

**Example sentence which shoes the stages of the direct translation done by MT**

The translation is done in the direct way from the source language to the target language

**SL text**

The translation is done in the direct way from the source language to the target language.

**Break down in source language**

The translation is done in the direct way from the source language to the target language.

**Lexical transfer (**use of equivalents of the TL)

Překlad je udělaný v přímé cestě ze zdroje jazyk do cílového jazyk.

**Adaptation in target language (TL)**

Překlad je tvořen přímo ze zdrojového jazyka do cílového jazyka.

In the example it can be seen that the translation is almost correct. Czech language is synthetic language and uses the declension and conjugation whereas English is not, the MT is not able to do the declension. Therefore the words ***source*** and ***language*** do not have the correct form in the provided translation. The word ***source*** is translated as a noun but actually it should be an adjective ***zdrojový***. The second step has to be performed by the translator since the MT is not able to do declension of the inflectional ending ***ový*** for the accusative case ***ového*****.** The same process should be done with the word ***language***, here again the inflectional ending ***a*** for the accusative is missing.

### 2.2.1.2Transfer Strategy

It involves three stages.

- analysis stage
- transfer stage
- generation stage

**Analysis Stage –** It is description of the source document from the linguistic point of view, use of a source language dictionary.

**Transfer Stage –** It is transformation of the results of the analysis stage and establishment of the linguistic and structural equivalents between the two languages, use of a bilingual dictionary from source language to target language

**Generation Stage** – It is generation of a document in target language on the basis of a linguistic data of a source language by means of a target language dictionary

In other words it lets the linguists build up the grammatical rules which are afterwards used by the system. Therefore when translating it analyses the source language document, transfer the grammatical structure to the target language and generates the translation.

Transfer strategy was developed by GETA (Groupe d'Etude pour la Traduction Automatique / Machine Translation Study Group) in France. GETA has supported other projects concerning the transfer strategy. Some of them are in the process of experimenting some of them have been already and successfully used, for example Canadian TAUM-MÉTÉOTAUM is an acronym which was created according to the

initial letters of Traduction Automatique de l'Université de Montréal (University of Montreal Machine Translation). It is also one of the oldest and effectively working systems. Its purpose is to translate weather forecasts. Firstly it was only from English to French but since 1989 it has been possible the other way round. TAUM- MÉTÉO is 80% effective in forecasting since the forecast are linguistically limited. It works with 1,500 lexical units, most of them proper nouns. It uses repetitive sentences with simple and stereotypical syntax and the match from structure to structure is perfect (Craciuescu at al 2004, p. 8).

### 2.2.1.3 The Pivot Language Strategy

Pivot language is based on developing an 'independent language' that serves as a conveyor of information that is communicating. It works as a neutral and central axis for SL and TL and is distinct for both. The translation itself then is done via the analyses of the source text, where the conceptual representation is found and in the generation module the diverse units are matched to their equivalents in target language. The system of using the pivot language is not focused on the direct translation the aim is to reformulate the source text from the important information. Concerning the pivot strategy it is necessary to mention Dutch DLT project (Distributed Language Translation project). This project was run between the years 1985 to 1990 and as a pivot language was used Esperanto. In this project were included translations of 12 Eroupean languages (Craciuescu 2004, p. 8)

At this point should be said that even if the system is able to translate a specific text with limited vocabulary and restricted syntax in a highly sufficient way, machine translation still can not be taken as finished product. According to Christian Boitet, director of GETA who says in an interview given to the journal Le français dans le monde Nº314

> ….it allows translators to concentrate on producing a high-quality target text. Perhaps then "machine translation" is not an appropriate term, since the machine only completes the first stage of the process. It would be more accurate to talk of a tool that aids the translation process, rather than an independent translation systém (in Craciuescu 2004, p.10).

## *2.3 Figurative language should be avoided when the Machine Translation is used*

MT is nowadays used for translation of company emails. If the sender of the email wants to be sure that the receiver will understand the email when he or she type it into the MT and click the button 'translate' the figurative language should not be present in the text. As it is mentioned above translations provided by the MT are nearly perfect but still need to be edited by the human translator since some errors can occur.

Features that definitely make the language more luxuriant should be avoided. MT is able to translate metaphors, slang, puns, jokes, idioms, regional or national dialects, however they will be translated literary. Unit by unit. Also sentences that are longer and more complex can be after the process of translation ambiguous or incomplete. As example a simple proverb is used to show what happens when it is translated with the use of MT.

**Example sentence 1:**

'You say that your sales will increase 10 times by the end of this year? Don´t count your chickens until they hatch.'

| **Don´t count your chickens until they hatch** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **SL:** | **Do** | **not** | **count** | **your** | **chickens** | **until they** | **hatch** |
| **TL:** | **Nespočítejte** | | | **vaše** | **kuřata** | **dokud oni ne-** | **líhnout** |
| | | | **počítat** | **vaše** | **kuřata** | **až do oni** | **poklop** |
| | | | **počet** | **tvůj** | **drůbež** | | **líheň** |

The MT offers three possible translations of which the user can choose the most appropriate one.

Firstly the pronoun *your* is translated as *vaše* it can be seen that this pronoun is translated either as a formal addressing of the second person singular or as second person plural. The word *Nespočítejte* has the perfect aspect however in the Czech would by used the imperfect aspect *nepočítejte.* Pronoun *they* is translated as *oni,* which is

undoubtedly correct however in Czech sentence the subject in this case would be hidden. Preposition *until* is divided into two parts. Negative prefix *un-* and the root *till.* The negative prefix *un-* becomes a part of the infinitive *hatch - ne-líhnout.* The infinitive *ne-líhnout* needs to be further change into the perfect aspect with the use of preposition *vy-* and the future form and the proper ending for the 3rd person plural has to be use to achieve the correct form of the verb *nevylíhnou.* Furthermore Czech language has reflexive pronoun *se*. This reflexive pronoun has not got a direct equivalent in English. It is not used in the source language and therefore it cannot be translated into the target language. Translation which is done with the use of MT is literary exact. However in the target language it would not convey the hidden meaning of the idiomatic expression.The pragmatic level is completely omitted. Also the cultural baground is not taken in concern. Even if the MT would be that developed and did all the necessary grammatical syntactical and semantical changes translated sentence "Nepočítej kuřata dokud se nevylíhnou" would not be understood by the reader as a proverb and in the context it would not make a sense. Therefore in this case the intervention of human translator is inevitable. *Neřikej hop dokud nepřeskočíš* would be the suitable translation of this proverb into the Czech language.

## *2.4 What Is the MT Good for?*

Machine translation is widely used for the translation of the texts that are highly repetitious. MT contains terminological dictionaries which can be tailored to fit the needs owing to the fact that these texts also contain terminology which is not set into complex sentence units. Moreover the translation of the text is done on the daily basis. For example the translation of weather forecast. MT allows the translation during few seconds and the translated information is still up to date. The same translation will take to the human translator much longer due to the fact that human translator is able to translate approximately 250 words per hour. Craciunescu at al. (2004, p. 8) confirm that MT allows the computer to translate the text without human intervention and the information which is obtained is not ambiguous and set into real time.

The high volume of emails which is received by the international companies is the best material for the translation via MT. It would be time consuming and extremely expensive to use the services of human translators to translate the emails and instant

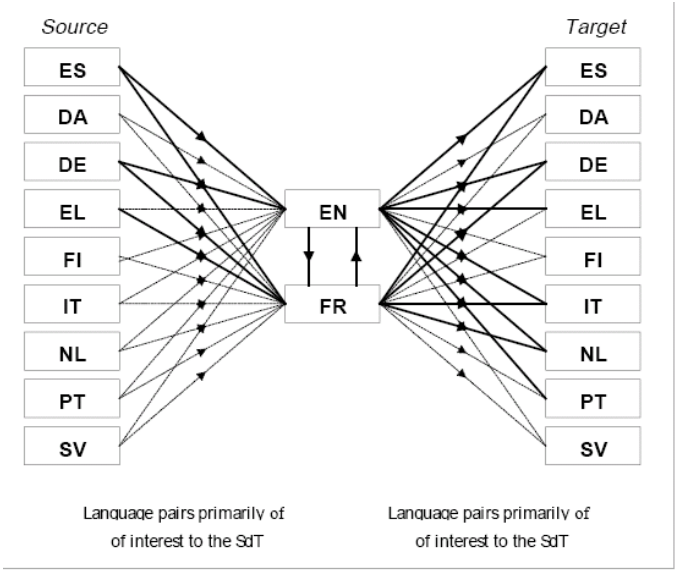
messages on the daily basis. It also excludes the third party in case that the information is private and confidential. This works especially for companies, which work with international vendors who receive data in foreign language.

'Get the gist of it' from this simple phrase, by adding the suffix -ing the gerund 'gisting' was formed. By other words, it is the way to obtain the general idea of a text. Nowadays it is used in a business communication. Gisting is used by the companies to get the gist of the purpose of the document in foreign language and on this basis they are able to decide whether they need the official translation or not (Drugan, p. 11).

## *2.5 Languages Supported by MT*

It is necessary to mention also the language pairs that are supported by the MT. This is probably the most important thing.

When purchasing the MT software every translator should be aware of the fact that not every language is already in pair with another. In other words when there is a SL it does not automatically mean that there is particular TL. For example not every software package that is offered on the market supports Czech language. According to John Beaven (1998, p.1), representative of Systran company[3], there are 110 language pair combinations possible with 17 languages (Below are shown the 17 already existing pairs).




[3] Systran company –developers of the MT Systran

Concluding the part of MT and summarize the theoretical pieces of knowledge it is necessary to say that MTs can be without hesitation used for translation of technical documents which consist of simple and specific sentences containing terminology. Also it is possible to use them for the translation of high volume of the previously mentioned documents since the human translation would be expensive. They are not suitable for translation of literary works because the pragmatic level of the language is completely excluded. Also it is necessary to translate between or among those languages that are already in pair otherwise MT is useless.

## *2.6 Machine Assisted Human Translation*

As it is defined at the beginning of the paper CAT is general term and MT and MAHT are the subsets of it. Machine Assisted Human Translation (MAHT), according to Feder, '…is a type of translation activity where a human remains the pivotal part of the translation process.' (2001, p. 49 in Palacz 2003, p. 6) In other words, the translation itself is done by the human translator and the computer system only supports it. It is the main characteristic of MAHT. These characteristic features are Translation Memory, Terminology Database, Alignment Tools and Workspace. These four creates together something what is generally called translator's workbench. That is, what it sets apart from MT. It is the example of perfect cooperation between human and technology. Technologies are nowadays quite developed and it can be said that everything is a subject of business. It includes MAHT systems as well. Every single product on the market tries to fit the needs of the user of the system.

## *2.7 Translation Memory (TM ) and Terminology Database(TMD)*

Translation memory and terminology database are two of the key features that support the human translator and are the most helpful in the translation process.

> …translation memory is a multilingual text archive containing segmented, aligned, parsed and classified multilingual texts, allowing storage and retrieval of aligned multilingual text segments against various search conditions. Unlike other language engineering systems a TM does not come

provided with linguistic data. It is a shell to be populated with translation equivalents. (EAGLES 1995)

TM tools are programs that store previous translations and, upon translating texts, offer suggestion to the translator, based on material stored in a memory database.

As the translation is being proceeded from the source text to the target text the translation memory finds the Translation Unit in the source language which is stored alongside with the Translation Unit in the target language. TM works with the segments of text, therefore as a TU, which is stored in TM, is usually referred to as a sentence. Anyway the TU can be even a single word or number or a bigger segment of the text, for example a whole paragraph. These TUs can be easily created thanks to so-called alignment tool which allows the translator to couple the SL and TL TUs from previously translated text and saved them into the translation memory. As it has been once translated and saved into TM always it occurs in any other translating document. The translation memory retrieves, matches and automatically translates the unit and provide the translation in the target text. For a better understanding of the process see the Figure 1 below.

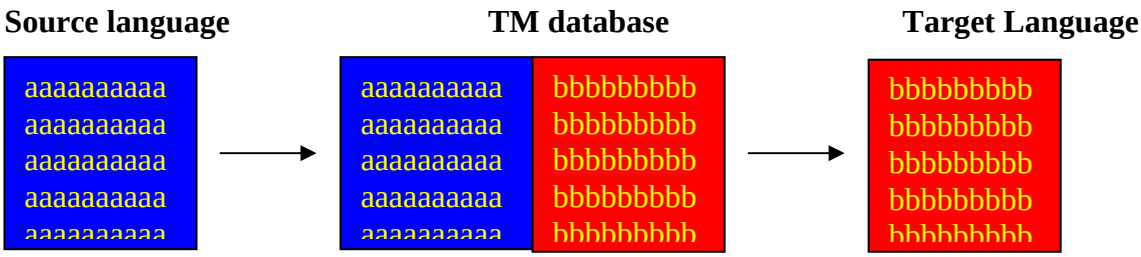



**Figure 1 translation memory process**

This use of TM can be considered as highly efficient way of translating. It saves the time and the effort of the human translator to remember how the term or phrase has been once translated. Therefore it can be said it ensure the consistency of the units previously translated in other or in the same document.

In the figure 1, there is shown the 100% exact or perfect match. It means that the TU, which was retrieved by TM was exactly the same as the one which is in the translating text. The occurrence of the exact match is around 70%. The rest of 30% is

called fuzzy match in (the similarity of the TUs is less than 100%). The fuzzy match is entered into the document and it the decision of accepting or rejecting is done by the translator. (Figure 2, 3)

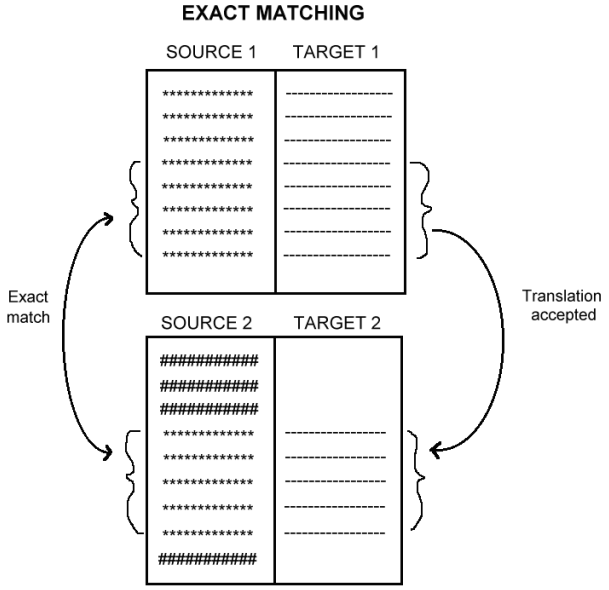



**Figure 2 exact matching**

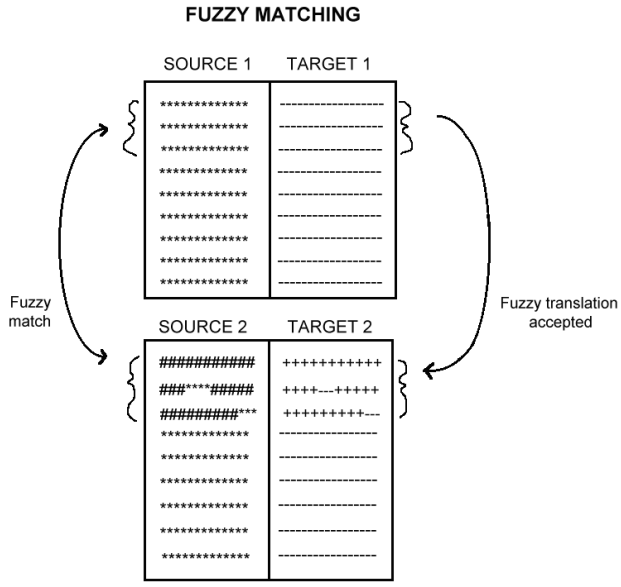



**Figure 3 Fuzzy matching**

For better imagination how it looks like on screen see Figure 4.

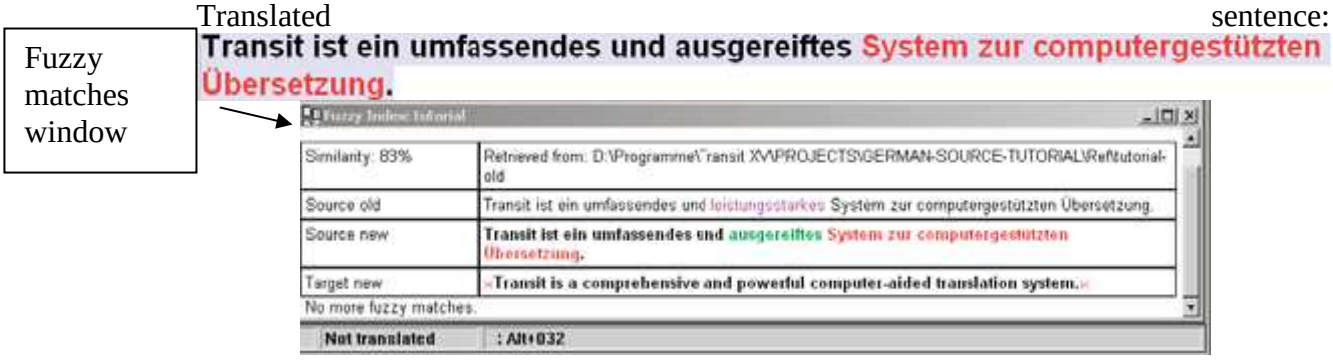



**Figure 4 Fuzzy logic Windows in Transit**

TM is also very useful when the document is being translated alongside with the source document that is being written. It is valid especially for documents that are in high volume and the documents in the source language and in the target language have to be released together. It can be very exhausting and time consuming for the translator

to find all the updates that were done through the whole source document. According to Webb (1998, p. 15) TM allows the translator to run the source document through the fuzzy matches and find all changes in segments or in terminology that have been done .

Together with TM is Terminology Database that allows the translator to full use of TM. Terminology Database is integrated and information is stored in it.

> Using an integrated terminology database allows a translator to perform fuzzy matching for a specific term or to use a term in the database suggested by TM. Without a terminology database that is compatible with translation memory, the TM user cannot easily obtain suggested translations for individual words without opening a separate electronic dictionary or looking through a conventional dictionary. (Webb 1998, p. 13)

## *2.8 Alignment tools and Workspace*

Alignment tools are instruments that are necessary for building of TM and TMD. They link SL and TL segments of the document to each other. Alignment tools allow the translator to take the advantage of using his or her translations which were performed outside the MAHT system.

> Current sentence alignment algorithms often rely on the identification of some kind of 'anchor' that connects SL and TL units. These anchors are strings that are identical or very similar on the SL and TL sides of the corpus. Besides formatting tags and punctuation marks, anchors typically comprise figures (numbers, dates, etc.), proper nouns and so-called 'cognates' (i.e., SL and TL words that share 'obvious' phonological or orthographic and semantic properties, with the result that they are likely to be used as mutual translations. (Simard et al., 1992 p. 71)

The alignment is not automatic. There are segmentation rules that have to be set by the user. User has to specify what should serve as a segment marker (question marks, full stops, commas, etc…). However not always these markers have to work produce proper SL and TL segment pair. If the translator does not check the alignment and when it is need correct it, the improper pair will be stored into TM and will be always retrieve.

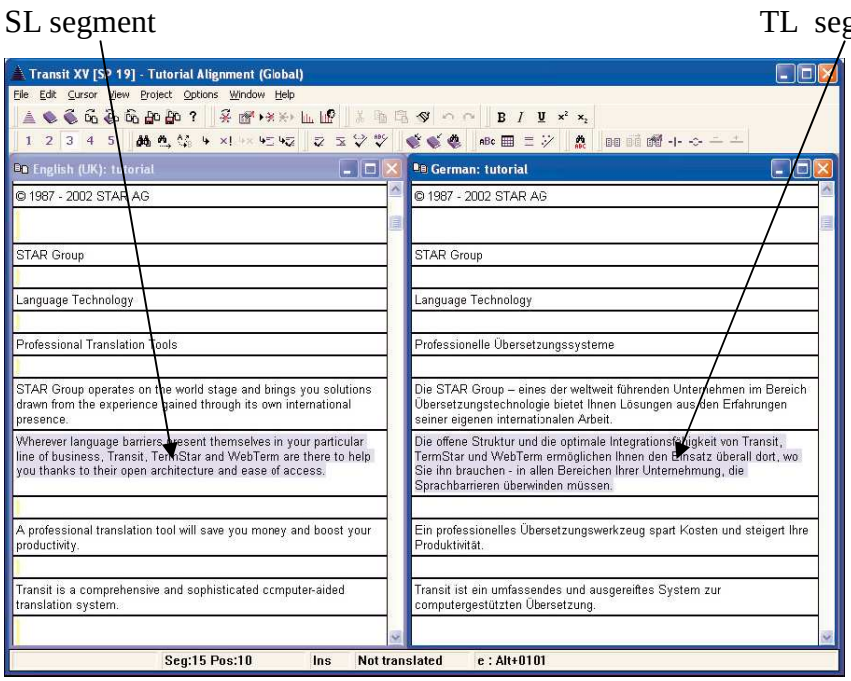



**Figure 5 Transit Align window**

Every MAHT provide an integrated workspace in which the translations are performed by the translator. Also the editing of aligned segments of SL and TL is done in this integrated workspace. It provides the access to the TM and TMD when the data is updated or edited.

## *2.9 Suitable Text Types for the Use of MAHT*

The use of TM depends on two important factors internal (reoccurrence of words, phrases or sentences in the same document) and external (reoccurrence of words phrases or sentences among documents) repeatability of a texts or text type. This reoccurrence can be seen in the technical types of texts, therefore these can be considered as the most suitable, however, it is still the subject of discussion what is the best text type for using the translation memory (Feder 2001 in Palacz 2003, p. 18).

However the higher is the content repetitive the more is suitable for the use of TM in the process of translation. Those categories of functional styles of texts listed below belong among the text types with the reoccurrence of the words, phrases, paragraphs and terminology.

- Business/Commercial(annual reports)
- Legal(patents, contracts)

- Technical(users manuals)
- Scientific

.

## 2.10 File Formats Supported by MAHT

An inevitable part of efficient use of MAHT tools is to be aware that the documents received from the client for the translation are in some kind of format. Therefore it is necessary for the translator to use software that supports the particular file format. Mostly they are MS Word, MS Excel or MS Power Point which are nowadays the basic ones which are used for documents. However it can happen that the received document is in other format. The leading companies in CAT-Tools want to meet the needs of their clients, therefore the range of the formats supported by the software are being widen. It is no more important what type of format was originally used. The text of the document is imported into the software programme and the translator works in the well known environment of the programme. When the translation is finished the document is, with the help of alignment tool, exported in the required format.

## 2.11 Languages Supported by MAHT

MATH tools that are on the market and are up to date support the majority of the existing languages. There are no boundaries even for right-to-left languages as Arabic, Farsi, Hebrew and their variants. If there is a limit it is caused only by the availability of the spellcheckers and fonts supported by the programmes. On the other hand this availability can cause problem with minority languages that have some special characters, for example Czech.

To summarize all the above mentioned features of MAHT it can be said that the nearly 50 years of continual work of the developers of the system can examined as successful. They developed system that is able to store the translations and repeatedly retrieve the terminology used in them thanks to TM and TMD. Alignment tools allows the translator work with number of file formats which can be imported into the MAHT and exported out off them with unchanged layout of the translated document. However,

they can be used only for the technical translation due to the fact of the repetitiveness of terminology used either within one document or across the documents. Finally, the number of languages that MAHT are able to work with is high and includes also languages with very special characteristics.

# 3. CAT-Tools as a Support

The CAT-Tools are widely advertised on the market as a support that relieve the translator of the majority of the routine work and allow them to concentrate on the translating itself. The growth of the market and expansion of companies to foreign countries requires the communication in various languages. Contracts are signed on daily basis and the content of the contract need to be comprehensible for both involved parties. Alongside the market growth, the demand for more and more translations increases. Translators nowadays do not only need to translate the document, but also they need to translate it quite fast. Furthermore, translators, to meet the needs of their clients, are required to work with various file formats and do quite much adjustments of the layout of the already translated document to keep it as similar as the original is. For this purpose CAT- Tools are being developed. The questions are whether the CAT-Tools when they are involved into the process of translation are anyhow helpful and also what are the required skills and equipment that the translators must have to be able to take the full advantage of using CAT-Tools.

To take the opinion about the two questions mentioned above a survey among those who are supposed to use CAT-Tools was held. As a sample, fifty translators were chosen randomly from the Companies Register in Czech Republic and from companies itself, however, only 45 were participated in the survey. These 45 translators were further, during the survey, divided for those who work as employees for certain companies and for those who work as freelancers. The translators were also asked for additional interview, however only ten of them were wiling to participate in it.

The method used for the survey was questioning, therefore the instrument was questionnaire with thirty-two questions and interview with those translators who agreed on it.

The types of the questions were different to guarantee reliable and valid information. Firstly, it was dichotomous type of question that asked for YES/NO answer. This type of screening question is used for instance when the researcher wants to make sure that only those people that are interested in the topic participate in the survey. Secondly, it was multiple choice question. This type of question consists of three or more exhaustive, mutually exclusive categories. Multiple choice questions can

ask for single or multiple answers. Thirdly, filter question was used in order to determine if the respondent is qualified or experienced enough to answer subsequent question. Lastly it was open-end question. Open-end question seeks to explore the qualitative, in-depth aspects of a particular topic or issue. It gives the person the chance to respond in detail (Trochim 2006). In our questionnaire open-end question type was mostly used. Questionnaires were spread via email. Respondents were asked to fill the questionnaire and send it back. With ranslators, who were willing to participate in the interview the appointment was set. Interview was based on the questionnaire and questions were prepared before the interview took place. The task of the interview was to check the responses and to gather additional information. For more exact record of the interview the Dictaphone was used. We do not provide the transcription of the interviews since the purpose of the interview were to verify the answers filled into the questionnaires and the role of the interview was only additional.

Data were prepared and checked in Microsoft Excel. Following is the interpretation of the results of the survey. The data itself is presented in the appendix.

## *3.1 Use of CAT-Tools*

CAT-Tools in any form are widely used among translators. To put them on a scale electronic dictionaries are favourites, they are used by all the translators. MT systems are used for translation of the documents by 40.5 % of translators. MAHT systems are more common and are used by 67.5 % of translators used for their translations. Translators tend to use two of these systems together. It is either MT together with dictionary or MT with MAHT.

There are also those, who tried to use MAHT systems but for some reason they quit. First and the most striking reason is the difficulty of the use of MAHT and the equipment required by these systems. Secondly there are purchase costs that are for the translator who just starts with translating or whose capital is not high enough pretty high.

Further the four MAHT system used by the Czech translators have been identified. It is TRANSIT from the production of the STAR company, SDL TRADOS from SDL Technologies, Déjà Vu from ATRIL and WORDFAST from WORDFAST LLC. Concerning the MT only PC TRANSLATOR from LANGSOFT has been

identified as the most widely used. Concerning MAHT the four above mentioned systems are representatives of the latest development. Every system more or less contains the same set of functions but a number of characteristic features. Among the very same functions belong TM, Terminology management, alignment, and workspace.

Apart the four representatives some others are known among the translators but they are not used in the Czech Republic: e.g. Systran and Reverso are MTs and are widely used abroad however is not possible to use them for Czech translation because they do not support Czech language. Others are Metatexis, Multicorpora, Terminotix.

Among widely used electronic dictionaries are Oxford and Cambridge dictionary. The use of various electronic and internet dictionaries is related to the technical field which the translators are specialized in. Business and commerce, law, medicine, constructions and IT are the fields where the use of high quality dictionaries should be used. Nowadays it is easy to find these types of dictionaries on the internet and translators have already built up a net of the best of them. The most important thing is that the use of the electronic dictionary is much faster than looking up each word in a paperback edition.

## *3.2 File Formats and Text Types*

When MAHT or MT is involved in the process of translation it is used only for the translation of the technical texts with repetitive content within the text or across the texts, which are considered as the most suitable. Some of the translators tried to use them for translation of literary texts, however the outcome was without any quality and more than 80% had to be edited by the translator. MTs are used for ‘gisting’ but only in very low number of translators. Translations that are created by MTs are translations of very short sentences, with specific terminology which cannot be misunderstood. Nevertheless the posteditioned is always done by the translator before the text is confirmed as translated.

Among the most common translated text types with the support of MAHT or MT belong contracts, business correspondence between parties of business, presentations, advertisements, internal company rules, IT and user’s manuals, product guides, technical reports, engineering drawings, various ratios, …etc..

These text types are often bound to a particular file format in which they are created. Contracts, business correspondence and company rules are mainly produced in Microsoft Word. Advertisements and users manual can be produced in Adobe Acrobat PDF format which is considered as the most problematic for translation. It is difficult to extract the text or do any changes with the length of the sentences and keep the layout of the page unchanged. For engineering drawings AutoCAD is used. Translators who translate the web pages have to work with HTML or XHTML formats. Presentations are mostly created in MS PowerPoint. When the translator used the MAHT system it is necessary that this system is able to import the format of the document. If not it is useless for the translator. Even though the translators have already established a deal with their clients about receiving certain file formats, it can sometimes happen that the document which needs to be translated is in format that cannot be imported. At that moment translation becomes much time consuming and exhausting. The document has to be printed out and new identical document has to be produced. This problem concerns mainly PDF files which cannot be imported into the MAHTs. From this format the text is extracted via Ghost View software into the MSWord. However as it is said by many translators the PDF format is used for more than a quarter of the documents they sometimes have to work with. Developers from the STAR company are aware of this fact and the year 2008 should be the year when new feature that is able to extract the text from the PDF files is going to be launched.

## *3.3 Advatages and Disadvatages of TM, TMD, Aligment tools*

The set of provided advantages and drawbacks concerns those MAHT systems that have been identified by translators. Some of the translators work in projects that consist of one or small files at a time which means that they do not need to operate in more documents that are being translated and also do not need to concentrate on the loads of formats. On the other hand some of them work in projects that consist of many more files and have to swap among the formats. MAHT system has the advantage of importing the documents. The translator works in the environment of the MAHT system and do not need to have the knowledge of various formats. There is one feature which

the translators would appreciate the most, however, none of the mentioned software has not got it. It is the on-line link to any kind of central internet TMD. Translation agency Skřivánek s. r. o. has been developing a system called ACROSS. This system will have the ability of access to the internet TMDs and it will allow the translator to take the advantage of running through various TMs that are up to date.

Another two advantageous features are TM and Terminology database with fuzzy logic. Fuzzy logic is probably one of the most appreciated features that are offered by MAHT systems. Those translators who work in a project that is being written and translated at the same moment, claimed that thanks to fuzzy logic it is easy and quick to find updated parts of the document. Alignment tools are also evaluated by the translators as useful and time saving, not only because they do not need to concentrate on the layout. In the alignment tools it is possible to import the old translations (source and a target document) and match them segment by segment and saved them into the TM. The disadvantage of this however can be that when in the previous translation which is either done separately and afterwards imported or is created within the programme can be mistake. This mistake is then constantly repeated in other translations up till the moment when someone spots it.

MAHT systems are also able to import electronic dictionaries but they have to be compatible with the MAHT system.

The environment of the MAHT was already mentioned however it is also one of the features and advantage of all of them. It is necessary to say that the environment is the same. It does not matter which system the translator use. They work on the basis of windows, which is user friendly and everyone knows how to operate in it.

MAHT have many functions therefore to take full advantage of MAHT requires also the full knowledge of the system. Most of the translators claim that to use the system properly is difficult. Some of them after they once tried the MAHT immediately quit because it meant for them to spend many hours just with learning how to use it. They also claim that the training with the professional lecturer is necessary and although they took it they were not able to use the system properly. Those who did not give up claimed that they had to work with the programme constantly for more than a week to gather all necessary knowledge and the translating took them much longer then without the MAHT. Nevertheless for the use of MAHT there is no need for some extra

knowledge of the use of the computer. As an advantage concerning the equipment, translators appreciate that the recommended configuration (600 MHz or faster Intel Pentium processor, RAM – 128 MB, Hard disk capacity – 100 MB, SVGA graphics card, 1024 x 768 pixels resolution) is not unreachable. Nowadays computers have even higher configuration therefore the MAHT performance is not the issue.

## 3.4 Costs and benefits

The fact whether the translator work as a freelancer or as a freelancer who does the translation for a translation agency or as an employee is inevitably bound with the purchase costs of CAT-Tools. Not only translators but also the companies which demand the translation use the CAT-Tools. We are talking about translation agencies and companies which have their own translation department. Translation agencies sometimes require the translator to have some MAHT software when he or she wants to work for them on the basis of freelancing. The reason is very simple. Company can easily check number of the words, spelling and store the translation in format which can be recycled. Also there are 16, 65 % of freelancing translators who do not work either in a translation department or for a translation agency. Half of them claimed that they are not required to use MAHT by the client and also their clients do not own them. 65, 2% of translators work for translation agencies and 22, 5 % confirmed that they use MAHTs. 18 % of translators work in translation departments and all of them use MAHTs. Freelancers who are not working for translation agencies and use MAHT had to purchase it. The price of some of MAHTs starts at 20 000,- and it is the price what is claimed as one of the reason for not using MAHTs by those who refused to work with them. Working for translation agency can have some advantages. Some of the agencies land their software or pay the half of the purchasing price. Translators claimed that to purchase the MAHTs becomes worthy when the average of the translated pages per month is more that 50 standard pages (average price for a standard page – 350,- after taxation) . The situation is best for the employees of translation departments because the software is purchased by the company.

As the competition on the market is strong every software company wants to meet the current needs of the customer. The software is still under process of

progressive development. Trados, Transit, Déjà Vu and Wordfast provide their customers with service packs for their versions of software. All of them offer special discounts, trial versions of newly developed features which are free of charge. Wordfast can be downloaded from the internet for free. Transit is sent on CD via post office and if the person decides to purchase it he or she is sent the code which decodes the CD for further use. Translators appreciate free training of the purchased software which is offered for Transit and Trados. On special workshops, as was for example 'Jeronimus' days', the companies present their software, demonstrate its features and are willing to answer any questions. Also when there is any kind of workshop event, the companies offer discount when the software is purchased within the worshop. After the purchase there is an on-line help in case of any problems. Wordfast is probably the cheapest 'software' therefore it is used by most of the translators. However it has to be mentioned that it differs from the rest three. Rather than software package it is MSWord template. When it is downloaded into computer it is automatically started with MSWord. All of the translators who work with it said that to make it work is difficult but the user's manual describes the use of it step by step.

To summarize and generalize the above mentioned results of the survey which was held among the 45 users of CAT-Tools in the Czech Republic it can be said that CAT-Tools are often involved in the process of translating. They offer many advantages that are considered as helpful and timesaving. However the popularity of each of the subdivisions as MAHT, MT and any other electronic device which is involved in the process of translation has to be treated separately. The electronic dictionaries are used on daily basis among the translators and when they are compatible with MAHTs together they create an invincible synergy in the process of translation. MTs are used only marginally since it is not possible to provide a high quality translation without the postedition by the translator. Concerning MAHTs only, there are two groups of users. The first group contains certain number of translators who learnt how to use MAHTs and are satisfied with them and claimed them as a perfect support within the process of translation. The amount of translation is high and the time period for each translation is short. Translators quite often work on more projects at a time hence the MAHT is considered by them as a timesaving and effective support. They

appreciate every single update of the software and are willing to spend time with training and practising and they are ready to spend money for this kind of support.

On the other hand there is a majority group of translators who find the MAHTs difficult to use and, for the purpose of the translation they do on daily basis, needless. The amount of the translation they provide per month is low and also the content of the documents is easy to translate or that much repetitive that they can provide the translation without any kind of support. Therefore generally it can be said that the final user of CAT –Tools can be anyone who does technical translation which is considered to be the most suitable for the involvement of CAT-Tools in the process of translation. However, the same cannot be said about on of the subdivisions of CAT-Tools. MAHTs are very specific software packages and ensuing from the above mentioned facts the final user of MAHT definitely falls within the first group.

# Conclusion

The aim of the paper was to describe CAT-Tools and their subdivisions MAHTs and MTs and also their involvement in the process of translation. Furthermore the distinctive features of MAHTs and MTs in processing of the text with different complexity and regularity were provided and accompanied with examples. Furthermore it was to prove whether MAHTs and MTs can be anyhow relieving and effective for the translator. The objective of the paper was to find out whether MAHTs and MTs are proper devices for the translation of the technical text. Additionally the final user of CAT-Tools was identified.

First chapter was aimed to description of the technical translation in comparison with the literary translation. A brief outline of the language functions, text types and functional styles were introduced in order to identify the key features of the technical and literary translation and to show the probable suitability of each of the translations for the use of CAT-Tools. Finally the distinctive differences between each type of the translation were put into contrast.

Second chapter was devoted to definition of what CAT-Tools are, also to the brief history of development and further progress of MTs and MAHTs. Firstly, MT was introduced together with the strategies that are applied to the process of translation. Based on the theoretical findings translation of technical documents was identified as a proper for the use of MT in the process of translation. Since MTs are still not able to produce high quality translation without postedition it is recommended to use them only for the translation of simple sentences with specific use of terminology and vocabulary. Secondly, MAHT was defined together with its characteristic features (TM, TMD, Alignment tools and Workspace), which set the MAHT apart from MT. It was shown that the MAHTs systems can be considered as software package that is developed to support the work of the translator by offering him or her number of functions that can relieve the translator from the routine work and let him or her to concentrate on the translation itself. However, as well as MT, MAHT system is proper for the translation of technical documents since the content of these documents is highly repetitive.

Third chapter contained the survey held among the Czech translators who have any experience with the use of CAT-Tools. The market growth goes hand in hand with the demand for translation. The requirements on the translators are nowadays higher.

The faster the translation is performed the better. However translators do not only translate the documents they also have to work with different layouts of the documents and keep the format of them unchanged. It means quite many additional acts that are involved in the translation process. CAT-Tools have been being developed in order to assume these additional acts and provide the translator with other functions that make the translating faster and easier.

The aim of the survey was to prove that CAT-Tools support the work of the translator and make it more efficient. As an instrument of the survey was chosen questionnaire since it is considered as the most suitable for gathering the required data.

It was found out that CAT-Tools are used among the Czech translators quite often. CAT- Tools were divided into three groups: electronic dictionaries, MT, MAHT. The first mentioned group is used by every translator and is considered as easy to use. MTs are used for translation only marginally and the translation is always postedited by the translator. The main function of MTs is gisting for the main idea of the document. Concerning the last group, which is MAHT the translators are not unified in the opinion. There are those who appreciate the existence of MAHTs since they offer quite many advantages and are labelled as supportive and efficient. Also the service, which is provided after the purchase by the companies, is considered as sufficient and easily reachable. These translators are willing to invest an indispensable sum of money to take the full advantage of MAHT.

However, the price and the difficulty of use of MAHTs is considered as a disadvantage and it is the main reason given by translators who do not use the MAHTs and do not even think of it.

From the survey follow that CAT-Tools are generally considered as supportive systems, however, their subdivision MAHT have still a long way ahead before it will get into subconscious of all translators as a profitable system despite the fact that it is necessary to invest a considerable sum of money.

# Resumé

Překladatelství je jedním z nejstarších odvětví na světě. První doložené prameny jsou skoro stejně tak staré jako písmo samo. Všechna odvětví se neustále vyvíjí a není to mu jinak s překladatelstvím. Zmínky o překladech jsou dochovány již v dílech řeckých autorů. Překlad byl zmiňován a rozvíjen během staletí a ani v současné době tomu není jinak. Překlad ovšem není jen převod textu z jednoho jazyka do jiného. Je to tvorba nového díla, do kterého překladatel vkládá část samu sebe. Může tedy znít neuvěřitelně, že něco tak lidského je možné spojit s něčím tak přesným a umělým jako je počítač. Myšlenka spojení počítače s překladatelem není tak stará. Na začátku 20. století byl v USA započat vývoj strojového překladu, tzv. Machina Translation. K velkému zklamání všech zúčastněných vědců, byl ale tento vývoj v roce 1966 zastaven organizací ALPAC z důvodů neschopnosti programu vytvořit dostatečně kvalitní automatický překlad. Nebylo tomu tak všude. Tento výzkum pokračoval v Evropě, kde byl Němci vyvinut program SYSTRAN, který používala Evropská komise, a TAUM-METEO, který je používán Kanaďany pro překlad zpráv o počasí z francouzštiny do angličtiny. V USA byl výzkum směřován opačným směrem. Protože se nepodařilo vyvinout dostatečně dobrý program pro automatický překlad, vědci se zaměřili na vývoj takového softwaru, který by překlad netvořil, ale jen podporoval samotného překladatele v jeho činnosti. V devadesátých letech největší softwarové společnosti v čele s IBM představily první systém, který podporoval překlad tzv. Machine Assisted Human Translation tool. Postupně byly vytvořeny další společnosti, jejichž cílem bylo vyvinout další a výkonnější software pro počítačem podporovaný překlad.

Tato práce je zaměřena na CAT-Tools, tedy na systémy pro počítačem podporovaný překlad. Cílem této práce je zhodnocení systémů používaných pro automatický překlad a pro počítačem podporovaný překlad a zjištění, zdali jsou tyto systémy skutečným přínosem pro překladatelské odvětví. Tato práce se dále zabývá popisem klíčových rozdílů mezi automatickým překladem a počítačem podporovaným překladem a zároveň se snaží zjistit, zdali jsou CAT-Tools nástroje vhodné pro překlad odborných textů. V neposlední řadě se autor snaží identifikovat konečného uživatele těchto nástrojů.

První kapitola je věnována charakterizování odborného překladu. Je zde stručně popsána jazyková funkce prostě sdělná, umělecká a funkce vokativní tedy výzva něco činit. Dále jsou zde popsány funkční styly společně s příslušnými typy textů pro jednotlivé styly. Tyto funkce jsou v překladu jako takovém považovány za klíčové. Při překládání je pro každého překladatele důležité, aby byl schopen rozlišit od sebe jednotlivé styly a jejich znaky a zároveň cíle jednotlivých textů, aby mohl důsledně volit jazykové prostředky, které v překladu použije. Vzhledem k tomu, že tato práce není zaměřena na rozdíly mezi literárním a odborným překladem, je zde uveden jen stručný popis jednotlivých znaků, typických pro literární a odborný text. Je zde věnováno více prostoru odbornému textu, který je považován za vhodný pro použití CAT nástrojů. Důvodem je především denotativní význam slov použitých v odborném textu, s kterým jsou počítačové systémy schopny pracovat a jsou ve své podstatě unifikovány pro daný odborný text.

V druhé části práce je popsán vývoj CAT-Tool nástrojů od jejich počátku na začátku 20. století až po současnost. Je zde představen automatický překlad (Machine Translation) a jeho jednotlivé strategie, na jejichž základě může automatický překlad fungovat.

První z nich je strategie přímého překladu, která se zakládá na skutečnosti, kdy každé slovo ve zdrojovém jazyce má přímé spojení s ekvivalentem v cílovém jazyce, tzn. věta zadaná ve zdrojovém jazyce je rozdělena na jednotlivého morfémy, které se dají přeložit. K těmto morfémům najde jejich ekvivalenty v cílovém jazyce a vytvoří z nich větu s naprosto stejnou strukturou.

Druhá je strategie převodu. V tomto případě není překlad prováděn slovo od slova, ale spíše význam za význam. Dokument ve zdrojovém jazyce je zanalyzován z lingvistického hlediska, poté jsou vytvořeny ekvivalenty mezi oběma jazyky s použitím bilingvního slovníku a závěrem je vytvoření dokumentu v cílovém jazyce na základě lingvistických dat cílového jazyka s použitím slovníku v cílovém jazyce.

Poslední strategii je strategie klíčového jazyka. Klíčový jazyk je nezávislý jazyk, který slouží pro přenos informace. Je centrální osou jak pro zdrojový jazyk tak pro cílový jazyk. Překlad opět není prováděn přímo. Tato strategie se zakládá na přeformulování důležitých informací zdrojového dokumentu. Jako klíčový jazyk slouží převážně Esperanto.

Automatický překlad je v dnešní době používán jen velmi okrajově a jeho použití má svá pravidla. Aby bylo dosaženo srozumitelného překladu, je nutné vynechat ve zdrojovém dokumentu použití metaforického jazyka. Systém automatického překladu není schopen rozlišit konotativní význam od denotativního, takže jakákoli slovní hříčka či přísloví bude přeložena doslovně. V některých případech pak tento překlad bude nesrozumitelný nebo nepochopitelný. Tato část kapitoly tedy popisuje znaky, které by měly být vynechány při použití automatického překladu. Dále jsou zde uvedeny činnosti, ke kterým je nejčastěji automatický překlad používán, společně s vymezením typů textů, které jsou vhodné pro použití automatického překladu.

Další část této kapitoly je věnována nástrojům pro počítačem podporovaný překlad tzv. MAHT ( Machina Assisted Humann Translation). MAHT jsou soubory programů, které poskytují překladateli nástroje, které ulehčují práci s překládaným textem. Mezi tyto nástroje patří především překladová paměť (Translation Memory), která slouží k ukládání již hotových překladů. Terminologická databáze (Terminology Database), která obsahuje terminologii použitou v předešlých překladech, a která společně s překladovou pamětí, zajišťuje konsistenci překládaného textu. Nástroje pro uspořádání formátu textu a tvorbu terminologické databáze (Alignment tools) společně s pracovní plochou ( Workspace) umožňují editaci a tvorbu terminologické databáze a její update a zároveň povolují práci s různými formáty souborů. Tyto čtyři nástroje tvoří základ všech typů softwarů, které jsou vyvinuty pro podporu překladu a jsou zde detailně popsány. Dále se autorka v této kapitole věnuje určení vhodných typů textů, jako jsou například smlouvy a obchodní korespondence, na které lze uplatnit použití těchto nástrojů. V dnešní době jsou dokumenty tvořeny a přenášeny převážně v elektronické podobě a proto jsou zde popsány jednotlivé formáty souborů, se kterými jsou tyto softwary schopny bez problémů pracovat a ulehčit tak překladateli práci s těmito formáty.

Třetí, poslední kapitola této práce se zabývá výzkumem, který byl proveden mezi uživateli CAT-Tool nástrojů. CAT-Tool nástroje jsou dnes na trhu poměrně běžným artiklem a jsou firmami propagovány jako nástroje, které ulehčí překladateli od takových úkonů, které ve své podstatě nemají s překládáním moc společného a urychlí překládání jako takové. Trh se dnes velice rychle rozrůstá a neustálá poptávka po překladech roste společně sním. Mezi požadavky na překladatele dnes běžně patří

rychlost, pružnost, a znalost práce s různými formáty dokumentů. Překladatelé pracují na několika různých překladech zároveň, nebo překládají rozsáhlý dokument, který je neustále měněn. Překlad takových dokumentů, nebo rozsáhlejších projektů, by byl v dnešní době bez použití počítače velmi náročný nejen časově ale i finančně. Cílem výzkumu bylo zjistit zdali CAT-Tools nástroje jsou schopny ulehčit a zrychlit překládání, jaké jsou jejich výhody a nevýhody, a zároveň jaké jsou požadavky na překladatele v případě, že se rozhodne tyto nástroje používat.

Dále tato kapitola popisuje metodologii výzkumu a nástroje, které byly při výzkumu použity. Jsou zde teoreticky popsány jednotlivé typy otázek, které byly použity v dotazníku a způsob výběru respondentů, kteří se na výzkumu podíleli. Dále jsou zde interpretována data výzkumu. Interpretace těchto dat je rozdělena do tří podkapitol.

V první podkapitole jsou interpretována data týkající se používaní CAT-Tool nástrojů mezi překladateli. Je zde procentuelně vyjádřen poměr uživatelů automatického překladu a počítačem podporovaného překladu společně s úkony, které jsou těmito nástroji prováděny. Zároveň jsou zde identifikovány MAHT nástroje, které jsou mezi překladateli v České republice nejčastěji užívány. Druhá podkapitola je zaměřena na formáty souborů a typy textů, se kterými překladatelé nejčastěji pracují. Třetí podkapitola uvádí výhody a nevýhody MAHT nástrojů, jmenovitě překladové paměti, terminologické databáze a nástrojů pro tvorbu terminologické databáze. Je zde popsáno i technické vybavení, které je potřeba pro bezchybné fungování těchto nástrojů, které je, bez větších obtíží, v dnešní době dostupné. V poslední podkapitole jsou zhodnoceny náklady na pořízení CAT-Tool nástrojů, jejich dostupnost pro jednotlivé skupiny překladatelů, společně s výhodami, které vyplývají z pořízení těchto nástrojů. Zároveň je zde zhodnocena péče, která je poskytována po zakoupení nástrojů.

Výzkum prokázal užitečnost CAT-Tool nástrojů při překládání odborných textů. Výhody, které plynout z použití CAT-Tool nástrojů jako např. konsistence terminologie v rámci překládaného textu a zároveň konsistence v rámci několika textů je hodnocena uživateli jako jedna z největších výhod těchto nástrojů. Tato výhoda a další jiné, nicméně na některé překladatele dostatečně nezapůsobily. Tito překladatelé se spíše kloní k názoru, že CAT-Tool nástroje jsou velmi složité a je nutné věnovat příliš mnoho

času na nastudování všech funkcí, které CAT-Tool nástroje nabízejí. Zároveň jsou tyto nástroje příliš drahé.

I přesto, že jsou CAT-Tool nástroje prokazatelně pro překladatele výhodné, mají před sebou ještě dlouhou cestu, aby se dostaly do podvědomí všech překladatelů jako nástroje, které i přes svoji náročnost na obsluhu, ulehčí překládání odborných textů.

# Bibliography

ALPAC.Language and machines:computers in translation and linguistics. *A report by Automatic Language Processing Advisory Comitee,Division of Behavioural Science,National Academy of Science, National Research Council.* [online]. Washington, D.C.,1966, no. 1416 [cit. 2008-04-20]. Dostupný z WWW: <www.books.nap.edu/openbook.php?record_id=9547apage=32>.

BAKER, Mona. *In other words : A coursebook on translation*. London : Routledge, 2001. 304 s. ISBN 0415030862.

BAVEN, John. Future MT developments. *Europa* [online]. 1998, no. 1 [cit. 2008-03-10]. Dostupný z WWW: <www.ec.europa.eu/translation/reading/articles/pdf/1998_01_tt_beaven2.pdf>.

BASSNETT, Susan. *Translation Studies*. London: Routledge, 1991. In: Vondráčková, Dana. *Rosten´s "O Kaplan! My Kaplan!" As a challenging literary piece for its translator Antonín Přidal*, 2007. 37 s. Faculty of Philosophy, Department of English and American Studies, University of Pardubice, Pardubice. Bakalářská práce. In courtesy of supervisor Mgr. Irena Reimannová

BOLOGNA, Sandra. *Babble Not-Machine Translation for the Technical Communicator*. *Newsletters* [online]. 2004, no. 12 [cit. 2008-04-24]. Dostupný z WWW: <www.wintranslation.com/newsletters/December2004Newsletter.htm>.

CRACIUNESCU, Olivia, GERDING-SALAS, Constanza, STRINGER-O\'KEEFFE, Susan. Machine Translation and Computer-Assisted Translation: a New Way of Translating?. *Translation Journal* [online]. July 2004, Volume 8.,No 3 [cit. 2008-03-10]. Dostupný z WWW: <www.accurapid.com/journal/29computers.htm>.

DRUGAN, Joanna. *The effects of CAT tools on translation quality* [online]. 2004 [cit. 2008-05-06]. Dostupný z WWW: <www.leeds.ac.uk/cts>.

EAGLES. Evaluation of Natural Language Processing Systems. *Final Report, EAGLES DOCUMENT EAG-EWG-PR.2,* [online] September 1995, [cit. 2008-03-10].
Dostupný z WWW:
<www.issco-www.unige.ch/ewg95/ewg95.html>.

FEDER, Marcin. Computer assisted translation. A propsal for tool evaluation. In: Palacz, Blazej. A comparative study of CAT tools with translation memory components, Poznan, [online] 2003. 70 s. Adam Mickiewicz University, Poznan, Poland. [cit. 2008-03-10] Diplomová práce. Dostupný z WWW:
<www.transsoft.seo.pl/other/TranslatorTools/TMTCHpt1.pdf>.

WEBB, Lynn E. *Advantages and disadvantages fo translation memory: a cost/benefit analysis*, Monterey, [online] 1998. 70 s. Monterey Institute of International Studies, Monterey, Kalifornia. [cit. 2008-04-24] Diplomová práce. Dostupný z WWW:

<www.tradulex.org/bibliography/Webb.htm>.

HUTCHINSON, John. Computer-based translation systems and tools. *University of East Anglia* [online]. 1996 [cit. 2008-03-10]. Dostupný z WWW: <www.eamt.org/archive/hutchinson_intro.html>.

KNITTLOVÁ, Dagmar. *Funkční styly v angličtině a češtině*. Olomouc : Rektorát Univerzity Palackého v Olomouci, 1990. 100 s.

NEWMARK, Peter. *A Textbook of Translation*. New York : Prentice Hall, 1988. 292 s. ISBN 0139125930.

KUFNEROVÁ, Zlata, et al. *Překládání a čeština*. Praha : H a H Vyšehradská s.r.o, 2003. 264 s. ISBN 8085787148.

TROCHIM, William M.K. *Research Methods Knowledge Base*. [s.l.] : Cornell University, 2005. 270 s. ISBN 1592601464.

SIMARD, Michael, et al. *Using Cognates to Align Sentences in Bilingual Corpora. In:Proceedings of the Fourth International Conference on Theoretical and Methodological issues in Machine Translation(TMI), Centre d\' innovation en technologies de l\'information*. Montreal, Kanada : [s.n.], 1992. 67-81 s.

SANDIN, Donna. Translation Tools: Getting Your Hands Dirty. *The ATA Chronical 35* [online]. 2006, no. 7 [cit. 2008-04-27], s. 18-19. Dostupný z WWW: <http://www.atanet.org//chronicle/feature_article_july2006.php>.

# Appendix I – Letter of Accompany

Dear Sirs / Madames

I am a student of last grade at Department of English and American studies, Faculty of Arts at University of Pardubice. The topic of my Bc. paper is CAT-Tools – modern technologies in translating. In order to gather the needed data, I would like to ask you to fill in the attached questionnaire which concerns the CAT-Tools.

Also I would like to ask you to participate in the additional interview for the purpose of checking the answers.

Thank you very much for your participation.

Kateřina Haššáková

# Appendix II - Questionnaire

1] Do you use CAT tools (Machine translation, electronic dictionaries, Machine assisted human translation systems i.e. TRADOS)?

YES NO

2] Which CAT tools do you use?

Machine translation (MT)

MAHT (Machine assisted human translation)

Electronic dictionaries

Others

3] If you do not use MATH system currently, have you ever tried to use MAHT system?

YES NO

4] If yes why did you quit?

5] Can you name those (MT or MAHT) systems you use?

MT MATH El. dictionaries

6] Do you know any other CAT tool systems?

7] If you use MAHT, what kind of translation do you use them for?

8] If you use MT, what kind of translations or actions do you use them for?

9] Which file formats do you work with? ANSI/ASCII/Unicode text for Windows

_ Apple Macintosh text
_ Corel WordPerfect
_ HTML
_ XHTML
_ XML
_ SGML

_ SVG
_ MS Word for Windows
_ MS Excel
_ MS PowerPoint
_ RTF and RTF for WinHelp
_ QuarkXPress
_ Adobe FrameMaker
_ Adobe PageMaker
_ Interleaf/Quicksilver
_ Adobe InDesign
_ AutoCAD
_ Visio XML
_ Resource files
_ Trados

10] Does the MAHT system you use support all the formats you work with?

YES NO

11] Do you work as an employee of a company or as a freelancer?
Employee

Freelancer

12] If you work as a freelancer, do you work for a translating agency?

Freelancer

Freelancer for agency

13] What kind of technical translation are you specialized at?

IT

Economic

Law

Medicine

Constructions

Others_____

14] What kind of text types do you often work with?

Manuals
Contracts

Business correspondence
Annual reports
Stock news
Minutes
Construction manuals

15] What kind of projects do you work in?
A] projects that consist of one or a small handful of files at a time

B] projects that consist of many more files(updated files)?

16] Is the content of the translating documents repetitive?

YES

Quite often

NO

17] Do your clients use a translation environment tool?

YES NO

18] How many pages per day are you able to translate with the use of MAHT systems?

19] How many pages per day are you able to translate without the use of MAHT systems?

20] How many pages of translation, according to your opinion, is the minimum, the MAHT systems become worthy to purchase.

21] The MAHT software you use for translating was purchased by the client you work for or by you?

Translator

Client

22] How much did you have to pay for it? (approximately)

23] Is it worthy for you to use the… ?

MT- YES NO

MAHT- YES NO

24] What is the support of the parent company of the MAHT system you use like?

Perfect

More or less good

Bad

25] Are there any service packs?

YES NO

26] If yes are they free of charge?

YES NO

27] Does the person who wants to use any CAT tools need to be good at working with computers?

YES NO

28] Is it difficult to use the CAT Tools?

YES NO

29] How long does it take to learn how to use particular CAT tool in the most efficient way? (please name the CAT tool you are talking about)

30] According to your experience, what are the advantages of the CAT tools?

31] Are you able to take all the advantages that CAT tools offer?

YES NOT YET NO

32] According to your experience, what are the disadvantages of the CAT tools?

Thank you very much for your time spent with this questionnaire.

Kateřina Haššáková

## Appendix III – Results chart

| | | | | | |
|---|---|---|---|---|---|
| 1 | YES | NO | | | |
| | 45 | 0 | | | |
| 2 | MT | MAHT | El. dictionaries | Others | |
| | 12 | 33 | 45 | | |
| 3 | YES | NO | | | |
| | 7 | 5 | | | |
| 4 | Difficult to use | expensive | | | |
| | 12 | 10 | | | |
| 5 | MT | MAHT | El. dictionaries | Others | |
| | PC translator | Transit | Oxford | | |
| | | Trados | Cambridge | | |
| | | Deja Vu | Internet dictionaries | | |
| | | Wordfast | | | |
| 6 | MT | MAHT | | | |
| | Reverso | Multicorpora | | | |
| | Systran | Terminotix | | | |
| | | Metatexis | | | |
| 7 | | MAHT | | | |
| | | Technical translation only 33 | | | |
| 8 | MT | | | | |
| | Gisting | 4 | | | |
| | Simple sentences | 9 | | | |
| 9 | HTML | 3 | | | |
| | MSWord | 45 | | | |
| | MSExcel | 45 | | | |
| | Adobe Acrobat | 5 | | | |
| | AutoCAD | 4 | | | |
| | TRADOS | 2 | | | |
| 10 | YES 27 | NO 6 | | | |
| 11 | Employee 8 | Freelancer 37 | | | |
| 12 | Freelancer 8 | Freelancer for agency 29 | | | |
| 13 | IT 3 | Economic 35 | Law 2 | Medicine 0 | Constructions 5 |
| 14 | Manuals 40 | Contracts 40 | Business corr.40 | Annual r. 35 | Stock news 15 |
| | Minutes 25 | Construction | | | |

| | | | | | |
|---|---|---|---|---|---|
| | | manuals 6 | | | |
| 15 | Small projects 36 | Longer project 9 | | | |
| 16 | YES | NO | Quite often | | |
| | 32 | 2 | 11 | | |
| 17 | YES 33 | NO 13 | | | |
| 18 | With Maht | Up to12 | | | |
| 19 | WithoutMAHT | Up to 8 | | | |
| 20 | Min pages | 50 and more | | | |
| 21 | Translator | 20 | | | |
| | Client | 12 | | | |
| 22 | Price | 20 000 and more | | | |
| 23 | MT | MAHT | | | |
| | YES 9 | YES 33 | | | |
| | NO 36 | NO 13 | | | |
| 24 | Perfect | 30 | | | |
| | More or less good | 3 | | | |
| | Bad | 0 | | | |
| 25 | TRADOS | TRANSIT | DEJA VU | WORDFAST | PC-TRANSL |
| | YES | YES | YES | NO | YES |
| 26 | YES | NO | | | |
| | 33 | 0 | | | |
| 27 | YES | NO | | | |
| | 45 | 0 | | | |
| 28 | YES | NO | | | |
| | 40 | 5 | | | |
| 29 | TRANSIT | TRADOS | DEJA VU | WORDFAST | PC-TRANSL |
| | 7-14 days | 7-14 days | 7-14 days | Up to 7 days | Up to 7 days |
| 30 | TM | Expensive | | | |
| | Fuzzy logic | Difficult to use | | | |
| | TMD | Necessity of already translated texts | | | |
| | Efficient | Consistency of mistakes | | | |
| | Translation is faster | | | | |
| | Alignment tools | | | | |
| | Import file formats | | | | |
| | Import dictionaries | | | | |
| | consistency | | | | |