Faculty of Arts

2ND year Arts

English Dept.

CAT TOOLS: (Computer-Aided Translation)

AY 2022-2023

CAT TOOLS: (Computer-Aided Translation)

An Introduction

(A COMPILED COURSE-BOOK FOR TEACHING AND EDUCATIONAL PURPOSES ONLY)

2nd YEAR ARTS

Compiled by Dr. Heba Abdelraheim Alkady

A Y (2022-2023)

CAT TOOLS:

(Computer-Aided Translation)

(A COMPILED COURSE-BOOK FOR TEACHING AND EDUCATIONAL

PURPOSES ONLY)

The Concept of Translation

Translation can be defined as the process of rendering words/phrases/ a text from one language into another (Merriam-Webster Dictionary) or a product of such a process. Such a definition is naturally very simplified, and the process is far more intricate in reality. Such is the opinion of Bassnett (2013), who argues that translation is not a straightforward process that can be carried out by everyone, as it "involves complex negotiation between languages" (p. 3). When translating, TRLs must both interpret the ST and reformulate the meaning in another language.

This can be a difficult process, as different languages have different syntax, structures and vocabulary, and the TL often has no equivalent for an idea articulated in the SL.

In today's increasingly globalized world, where the demand for translation growing, as well as with new developments in technology, translations have become more sophisticated, increasingly relying on computer technology to facilitate the process. Computer translations have become a necessity for TRLs in today's globalized information age, to translate large amounts of

text quickly (Esselink 2000; Lagoudaki 2006, as cited in Bowker & Fisher, 2010). Depending on the technology used, translation can be divided into three types: human translation, machine / mechanical translation (MT), and computer-assisted / computer-aided translation (CAT).

This study will mostly be focused on CAT and the analysis of its tools, although CAT and MT will also be briefly compared.

It may be difficult to provide a comprehensive definition for the term, translation because theorists differ on the precise nature and the processes involved in translation. However, most definitions of translation may be classified under one of two headings. While the first is concerned with the process of translation, the second focuses on translation as a product, and describes what a good translation should be like. Example of definitions that concentrate on translation as a process is that suggested by Catford (1965) and another by Hatim and Mason (1990). Catford defines translation as "a process of substituting a text in one language for a text in another". Thus, he considers translation as an operation performed on

language, and so builds his definition on the concept of equivalence. Like Catford, Hatim and Mason consider translation as "a communicative process which takes place within a social context". They argue further that translation is a useful test case for examining the whole issue of the role of language in social life and creating a new act of communication of a previously existing one".

On the other hand, Nida and Taber (1969) and Newmark (1981) in their definitions, focus on translation as a product. Nida and Taber define translation as "producing in the receptor language the closest natural equivalent of the source language message, first in terms of meaning and secondly, in terms of style". Thus, they attach greater importance to the forms understood and accepted by the target audience for which a translation is designed, and also emphasize the importance of understanding the original text before constructing a text in the target language. Newmark, in the same manner, defines translation as "a craft consisting in the attempt to replace a written message and/or a statement in one language by the same message and/or statement in another language".

Translation, he continues, is rendition of the meaning of a text into another language in the way that the author intended the text. It would be observed that Newman, like Nida and Taber also stresses the importance of understanding the intention of the original text author before translating it into a target language. Since the comprehension of the intention of the original text author is crucial for producing an accurate text in a target language, we subscribe, therefore, to the definitions here above among many others.

Understanding the need for Translation

To facilitate effective communication among peoples of different cultures and languages, translation must take place from native to foreign languages and vice-versa in matters of socialization, education, administration, diplomatic relations, religion, commerce and etcetera. Furthermore, nearly all companies today use ordinary translators to decode all types of documents such as press releases, reports and other material relevant to their business. Nevertheless, at some point these companies will need to interpret a document of a technical nature. Using the same translation services for a technical paper that we use for general translations might cause a fundamental mistake. Sometimes, the person we rely upon to provide the translation service may not have the requisite skills or training to undertake a technical translation. This is because the technique of translation service is quite different from any other translation service. The major differences are in the intensity of complication in the translation and the level of accuracy required. It is therefore very necessary to understand both aspects when contracting out tasks and projects to translators.

Difference between translation and interpreting

The professions of translation and interpreting are notably different but the two overlap in some areas. A translator interprets the written word and the result of his/her interpretation is usually in written form. He/she has time to deliberate, conduct research, proof-read, revise, consult colleagues and submit his/her written translation to client. An interpreter, on the other hand, interprets the spoken word and mostly does not have the extra time nor a second chance to revise the result of the interpretation. Many translators, though, will have done some interpreting but this will probably have been secondary to written translation.

The Importance of Translation

The importance of translation can be brought to a sharp focus if we consider the fact that effective communication among different cultures can only be achieved through translation. It is by way of translation that peoples from different cultural backgrounds are introduced to various languages and ways of thought. Further, to train translators and interpreters, it is very necessary to understand the relationship between language and culture because "translation takes place in the context of the relations between two cultures, two worlds of thought and perception". Newmark opines that there is a cultural value in translation". Culture is also heavily indebted to translation for its intellectual development, and it is through translation that the complexities of language and of specific texts are

clearly demonstrated. In addition to that, good writing and bad writing are easily exposed in the process of translation. Translation is an important source of diffusion of every kind of knowledge; it facilitates the understanding of development of every aspect of culture in other civilizations; and thus enriches people's knowledge and understanding of their own culture.

The Annals of Translation

Translation activity is as old as time. There is no doubting the fact that every intellectual or scientific awakening has always been preceded by or begun with active translation. Inthe ancient times, strenuous translations from Greek language to Latin had been carried out following the decline of the Greek civilization and emergence of the Roman. Europe remained under the darkness of the middle ages till the fall of Constantine and transfer of its treasures through the relics of the Greek and the Roman to the then great western nations. Extensive translation exercise thus became active, rendering these antiques into such nascent European languages as English, French and Spanish, most of which developed out of Latin, and from elements of regional dialects or languages of varying impacts and power depending on region. The annal of renaissance period began with those translation works stabilizing, digesting and assimilating the translated, and then impressing it on the shaping of European mentality.

Historical Background of English – Arabic Translation

Arab civilzation likewise, had in history experienced many great translations. Between the 7th and the 8th centuries and beyond, the heritages of such early nations as Persia and Rome, that preceded the Arab nation, had been translated

into Arabic via Sanskrit, Syriac and later directly through the Greek translation. The Arabs, therefore, started learning foreign languages following stability of Islam in conquered regions, hence, they were able to render many alien heritages into Arabic. The Abbasid period particularly witnessed a widespread of vivacious translation activities as Iraq was populated by a sizeable number of medical practitioners, philosophers and astrologers. Translation during this era passed through several stages: it began by translating materials on Medicine, Astrology, Mathematics, Philosophy and Logic, and culminated in the translation of various books on sciences and literatures. This activity eventually resulted in the diffusion of knowledge, development of thinking styles and growth of Islamic sects, whose methodology is based on dialectic. All of these climaxed in the flourishing of Syntax and the evolution of Rhetoric. It is however, pertinent to mention that the later translation experience, which started by the middle of the 19th century, when Rifa'at Tahtawi founded in Egypt the language academy for translation of European languages (English inclusive), was an upshot of those early experiences in translation. Thenceforth, translation became a science having its own fundamental principles, after it had been an art relying on the competence and genius of a translator.

Full Translation and Partial Translation

There are several kinds of translation depending on perspectives of translation theorists and their classification. Though, in the Arabic – English translation module, you have been taken through some types of translation including literal, idiomatic, interlinear and free translation. These are indeed the most popular

types nonetheless, in the present unit; you will be exposed to other types such as full and partial translation. Catford (1965) shall be our reference point here as he succinctly distinguishes between these two kinds of translation, depending on the extent to which the source language text is submitted to the translation process. According to him, in a full translation, the entire text is subjected to the translation process and every part of the source text is replaced by target text material. In partial translation, on the other hand, he observes that some parts of the source language text are left untranslated and they are simply rendered to and incorporated in the target text.

Total Translation and Restricted Translation

Catford (1965) also draws a line between total and restricted translations. These have to do with the levels of language involved in translation. Total translation, according to him, is "the replacement of source language grammar and lexis by equivalent target language grammar and lexis with consequential replacement of source language phonology by (non-equivalent) target language phonology". Restricted translation, in contrast, means "the replacement of source language textual material by equivalent target language textual material at only one level". It needs mentioning that most translation works are concerned with the later as using the former may pose a greater difficulty in translation process and output.

It is pertinent to mention here that, there exist other types of translation aside from the "Normal" translation, which is the well known, traditional lingual process of exchanging and adapting a text from one language to another. The guidelines for normal translation are usually fulfilled by qualified translators, though sometimes, if the translation requires deeper and more special knowledge which

can be provided only by professionals in a field, we prefer them to general translators. These other types include ‘Certified Translation’ and ‘Localization’. In cases, when we need to submit documents to authorities, we might need certified translation of financial reports and certificates. In Nigeria for example, the simplest description of "Certified translation" is the translation made by departments of languages in various universities. Others call it "Official translation". Official translation is normal translation, which has been proofread and printed on official paper with a bilingual certification and official stamp. Localizing, on the other hand, is a special type of translation. In this type, the safeguarding of the global content, the form and the desired effect is often more important than the exact meaning of words. It therefore requires different view and way of thinking than "Normal translation". The translator's creativeness comes to play during localization, while it is not typical during "Normal" translation procedures. Localization is often used in advertisement spots, in formulating slogans and for designing various digital contents.

Techniques and strategies of English - Arabic Translation

Translation Techniques can broadly be divided into two: Direct Translation Techniques and Indirect Translation Techniques (Oblique). Direct Translation Techniques are used when theoretical and structural elements of the source language can be transposed into the target language.

Direct translation techniques include:

- Borrowing
- Calque
- Literal Translation

Borrowing is the taking of words directly from one language into another without translation. Many English words are "loaned" into other languages; for example software in the field of technology and funk in culture. English also borrows numerous words from other languages; like abattoir and résumé, which were loaned from French; hamburger and kindergarten borrowed from German; and musk and sugar, which were loaned from Sanskrit. Examples of borrowing from English into Arabic include Radio راديو, Television تلفاز Computer كمبيوتر Brandy براندي Bank البنك etc. Loaned words are often printed in italics to indicate that they are "foreign".

A calque or loan translation is a phrase borrowed from another language and translated literally word-for-word. It is often seen in specialized or internationalized fields such as quality assurance (assurance qualité taken from English). Examples of calque that have been absorbed into English include standpoint and beer garden from German Standpunkt and Biergarten. Some calques can become widely accepted in the target language (such as standpoint and beer garden). The meaning of other calques can be rather incomprehensible for most people, especially when they have to do with specific vocations or subjects such as science and law. For example Solución de compromiso is a Spanish legal term taken from the English compromise solution and although Spanish attorneys comprehend it, the meaning is not readily understood by the

layman. A fruitless calque can be exceedingly unnatural, and can cause unwanted humor, and is often interpreted as signifying lack of expertise of the translator in the target language.

CAT Tools

Translators around the whole world have noticed a huge innovation in the translation field in recent years. The CAT Tools, Computer-Assisted Translation Tools, Computer-Aided Translation Tools, or Machine-Assisted Translation Tools, have become a very useful instrument for every translator. CAT Tools are not a novelty on the online market. According to Austermühl, in 1997 more than a half of asked translators answered that they were using electronic dictionaries and roughly thirty percent were using systems with translation memory.1

The answer to the question – why to use a CAT Tool – is very simple – the translation becomes much more effective, accurate and easier. CAT Tools use a synoptical user interface to keep the source text and the translated text right next to each other and along with the translation memory, which saves already translated terms and is very useful for any kind of translations, is each CAT Tool representing an enormous help for every translator.2

Although the frontal CAT Tool brands are paid, many Computer-Assisted Translation Tools are possible to be used for free, and which might be misguided, these are not any worse than the paid ones.

Before this thesis had been written, several free and paid Computer Assisted Translation Tools were, after a long consideration, chosen – two paid and two free-for-use CAT Tools.

Although it may seem that there is plenty of free CAT Tools available on the internet, the reality is very different. A large part of the free CAT Tools were found inappropriate for this thesis due to the fact that they did not operate properly, problems during the installation occurred, or some of them were not so-called "standalone" applications but applications that are integrated into some other application, such as the WordFisher which is automatically installed into Microsoft Word. Based on these facts only two standalone free Computer Assisted Translation Tools have been chosen for this thesis. These applications were OmegaT and Across; they were chosen due to their popularity and high user usage.

As for the paid CAT Tools, there were not found any problems with their choosing. The consideration of choosing two application was based on user's popularity. A Microsoft-based MemoQ was chosen as the first paid CAT Tool. This application is very popular among its users, either for its easy usage or the user-friendly user interface. Probably the most famous CAT Tool – SDL Trados Studio, version 2015, was chosen as the second paid application. Both applications were downloaded from their official website and the trial 30 days long version was installed. And even though it was only a trial version, all features were included.

This work is divided into two main parts – the theoretical and the practical.

The theoretical part describes how Computer-Assisted Translation Tools actually work, which principles they use, where they are useful and, on the contrary, where they should not be used.

In the practical part, every single CAT Tool is described in details according to the same syllabus and tested with the same text. Many other factors, such as the user interface, installation, availability on several operation systems, how the translation memory works and many others will be described in this chapter as well.

Translation

The explanation of the term “translation” could be found in Tom McArthur’s Concise Oxford Companion to the English Language, as some kind of transformation of a source- language text by equivalent expressions of the target language.

The term “translation” is very often misunderstood with two different terms – localization and globalization. These three concepts have, according to Pym’s study in Translation Technology and its Teaching (with much mention of localization), different meanings. While translation, as stated above, is a transformation of a source language into the target language in any kind of project, the localization is based mainly on a translation of software engineering. The term globalization is more connected with localization than with translation because it is only expanded localization.

The first mention of translation that is possible to be found is the Epic of Gilgamesh, which was written 2000 BC, and translated roughly during the same time into Southwest Asian languages. As for Europe, the first translation mention was situated in Greece, in Alexandria, and it was a translation of Jewish Scriptures, the Septuagint. 6

History of Computer Assisted Translation Tools

Surprisingly, the history of Computer Assisted Translation Tools is dated back to the age of the Cold War. For the first time in history, huge amounts of money were invested into translation technology due to the need for quick translations of information obtained from the intelligence services. The U.S. Air Force used this tool in order to understand the general meaning of Russian documents. It was thought that translators could be fully replaced by translating machines in the near future but this idea was later considered as fault because the results were not satisfying. Therefore, during the 60s and 70s was suggested a new process – the translator machine should not translate itself, but only facilitate human translator with the translation. During the late 70s translating memories came up, and later on, in the 80s, so-called Translation Support System was developed by an American company. However, its price was too high and only very rich companies carrying out many translations could afford to buy one. One of these companies was for example the American company IBM.

Later on, one of the nowadays frontal Computer Assisted Translation Tools companies was inspired by TSS (Translation Support System) – Trados. In 1984 had Trados invented “text tools” and started a whole new era of Computer Assisted Translations.

Computer Assisted Translation Tool

As was said before, the expression CAT Tool stands for the Computer Assisted/Aided Translation Tool. The concept “CAT Tool” is often misunderstood and considered by many people as one out of many automatic translation tools, such as Google translator or any other. This idea is completely fault because the concept of a CAT Tool stands for an absolutely different meaning. These two concepts are connected only via the word “translation” but there is one very important difference. While “automatic translators”, such as earlier mentioned above Google translator, work on the purpose that the user inputs some text into some text field and the translation is made by the computer, however CAT tools work in a completely different way. There is a necessity of a human translator that translates the source text from a source language into a target language while the CAT Tool is only helping him/her with the translation.8

Nowadays “automatic translators” are very helpful and the level of translations is still on an ascent. Sometimes, it does not make any sense, though, and a correction or a whole reconstruction is necessary. Therefore, the human interference is needed and always will. When not talking about technical translation, computers simply cannot understand irony, slang and also any other linguistic aspect may cause troubles.

As was said before, the CAT Tool is a tool which helps the translator to translate larger amounts of text. A tool without which many translators would not be able

to work at all. The whole translation gets much easier, faster and the global overview is much superior. 9

Basically, the CAT Tool is an application having a synoptical user interface so the translator can see the source and the target texts at the same time. CAT Tools divide the whole text into segments (sentences) and each segment is displayed on one independent line. Therefore, segments which are not translated yet can be easily found. Several applications also allow users to import their own glossaries and other documents helping with the translation and improving the efficiency of their work.

Pros and cons of Computer Assisted Translation Tools

The main advantages of the CAT Tools involve:

The source text is presented as segments, which are sentences; these segments are presented in such a way that the translation is easier and faster. Translated or not translated segments are easily seen which helps the translator to have an overview of the whole translation.

Translation units

Each segment is saved as a group with the source text which is later presented as a so-called translation unit.

Later editable texts

The possibility of returning to each translation unit at any time.

Translation memory

The function of saving translation units into a database (so-called translation memory) which allows the user to use them in the same or any other text again at any time. Many CAT Tools allow the user to import their own translation memory, which might be very useful for translators translating larger amounts of text more often because they basically create their own “sentence dictionaries”.

Quality control

CAT tools prevent from making mistakes, such as typo mistakes, punctuation, and many others. For example, when the source text includes some numbers, such as 1950, and the translator makes a typo when writing the numbers, such as 190, the CAT Tool highlights the possible error.

Lucidity

Possibility for searching words to check the way they were translated before.

Formatting

It is often required to keep the formatting type as it is, which might be sometimes very difficult. Not with the CAT tool. The source text can be simply exported and afterwards the translated text is imported back. This is extremely useful if working with PDF files, presentations and other because the formatting is not necessary to be bothered with. Many CAT Tools also provide a possibility of editing the formatting. In such a case the translator is able to see the formatting all the time.

Backup

Every single segment (every single sentence) is being saved when translated. Therefore, you cannot lose more than one sentence in case the computer crashes.

Updates

In a case of just finished translation, and the client updates the source text it is no problem when using the CAT Tool. The updated version of the source text is imported and the CAT Tool highlights changed parts while the unchanged parts are kept.

Progress

The CAT tool creates a progress chart and therefore you always know what progress you made and how fast you are working. This feature becomes very useful when translating several documents where it is necessary to meet some deadlines.

Terminology

Many CAT tools provide a terminology function. If a client provides the terminology list, it is possible to easily import it into the CAT tool. Then, when running into some term which is included in the terminology list, the CAT tool notifies you to use this term.11

Other functions

Many CAT tools are way further in development and along with the local platform they dispose of a server platform. This functionality is used for having an up-to-date version of the source text from the translator's customer. In this case, the server should be installed on the customer's computer and on the translator's computer as well. Then, if the source text is changed by the customer, it changes automatically in the translator's computer so no useless work is done.12

The main disadvantages of the CAT Tools:

Price of paid Computer Assisted Translation Tools

As is described later in this thesis, CAT Tools are very expensive applications to buy. However, there is a possibility to use a CAT Tool which is basically not that classy as the paid ones, but works as it should and makes the translation easier.

Being paid less for the translation

Some translation agencies do pay their translators less in case they use Computer Assisted Translation Tools for requested translations. As agencies added, the paycheck is reduced since the translator has the translation much easier when using the translation memory. This is a very bad approach and even though the

translator invests not a small amount of money to buy a CAT Tool, he/she will be even punished afterwards.

CAT Tools testing

In the following chapter, the chosen CAT Tools will be tested. First, a few words about the whole testing concept will be said, after that the first tested CAT Tool, OmegaT, will be tested, followed by the second free-to-use CAT Tool, Across. Once these two are tested and described in details, two chosen paid Computer Assisted Translation Tools –

MemoQ and SDL Trados Studio 2015 - will be described and tested as well.

Introduction

After a long search and decision-making process, two best fitting free-to-use Computer Assisted Translation Tools were chosen. These are:

- OmegaT,
- Across.

To provide a more complex overview of available CAT Tools on the current market, two paid Computer Assisted Translation Tools were also chosen along

with the free-to-use CAT Tools. These tools were picked according to their user's satisfaction. The chosen tools are:

- MemoQ,
- SDL Trados Studio 2015.

These tools will be tested via the same testing document (file .docx) and afterwards described in details based on key aspects for choosing a proper free-to-use Computer Assisted Translation Tool. The key aspects described include:

- Installation,
- Availability on various platforms,
- User interface and the user's guide,
- Translation progress,
- Compatible files,
- Translation memory,
- Final evaluation.

Reasons for choosing these factors:

There is no doubt that many other factors could be involved as well but these points were founded as the most important.

Installation was chosen as the first factor. This factor is highly important because in many applications the user can meet many problems throughout the installation or it might not be done at all. Also, some installations were not provided via an installation wizard and therefore the whole installation is much harder to get through.

Next, the availability on various platforms, namely on Linux, Windows, and Mac OS X, will be covered as well. If a translator is about to choose one of the huge variety of CAT Tools, the availability on his/her operation system will be definitely one of the main aspects for choosing it.

In case the translator works on one project only, the translator progress feature is not that important. On the other hand, if the translator works on several projects and must meet deadlines, it is well assorted when all deadlines are visible right when the application is turned on.

When a new user starts using some to him/her unknown application, the level of the user interface is essential. Many applications have their user interface not

sophisticated enough, so there is a necessity of reading the user's guide. Therefore the user's guide is involved in this point as well.

Whether the translator translates technical texts or not, the glossary function will be handy in any case. The glossary may seem as a not that interesting or useful function, but the opposite is the truth. This feature saves translators many hours of searching on the internet or in a dictionary for some definitions.

Also, the compatible files are highly important when choosing the proper CAT Tool.

The main factors cannot devoid the best CAT Tool's feature – the translation memory. The function of translation memory, how the recurring text is highlighted or in some cases even automatically translated, and other aspects will be covered at this point.

Afterwards, the final evaluation of the whole application will be carried out.

OmegaT

OmegaT is one of the best free-to-use CAT Tools and is written in Java. This application is not divided into accounts so users do not have to expect any necessary payments for a premium account or premium functions. Since OmegaT is spread around the whole world and a large group of translators uses this handy application, a great amount of documentation and help videos is available on the internet.

Installation

The installation of OmegaT is very intuitive and even a non-experienced user will handle it. On the official website of OmegaT, there are many versions available to download. The web page is not made that well and it might be a little bit problematic to find the required version which fits the user's demands. Once the proper version is selected and downloaded, the installation itself is very fast and easy. A simple clicking "Next" button will do all the work.

Availability on various platforms

Since Microsoft Windows is the most spread operational system in the world, application support for this system has become very usual. OmegaT is no exception. However, OmegaT does not support only Microsoft Windows, but Linux contributions Mac OS X are supported as well.

User interface and the user's guide

Once you run the OmegaT tool, you will be surprised how well the user's guide is prepared (the user's guide is accessible also by pressing F1 button or clicking Help -> User's guide in the main menu). This guide goes through the whole application and is available in several languages and almost every step is described in details. The guide is separated into several paragraphs which provide perfect lucidity. The user interface is not that modern as we might be used to from some other applications, which might be the reason why OmegaT developers highlight using the user's guide so much.

Translation progress

The first disappointment comes with the translation. It is no longer intuitive and it might be way better. Once the user imports documents to translate, the formatting of the imported document is gone and is replaced by XML tags which are disturbing. Also, when a translation file is loaded into OmegaT, and the translation begins, the source text is replaced by the target text, as can be seen in Figure 1, and the source text might be seen again only if the user clicks on desired segment, as can be seen in Figure 2. Therefore, any later check is very slow.

OmegaT is also capable of reading PDF files, one problem was faced, though. In case the PDF file consists of a text which is structured in one column, no problem occurs. However, the problem arises when the text is structured into two or more columns, as stated in Figure 3. There is the text divided into two columns.

Translation memory

The translation memory is one of the main features of every computer assisted translation tool and that is no exception for OmegaT. The translation memory function works in such a way that the user does not have to translate recurrent expressions each time but only once. Translation memory in OmegaT works as expected – every recurring expression is automatically translated when the user clicks on some segment, then it is possible to edit it by the user's requests.

Special features

Any extra features that would exalt OmegaT against its competition were not found.

Final evaluation

Though OmegaT is a well sophisticated free-to-use CAT Tool, it is obvious immediately that it is an old application. On the contrary, using OmegaT is very simple and as stated in the user's guide – "Learn how to use application OmegaT in 5 minutes!" – says it all.

The presence of many European but also other languages is very pleasant. Also the fact that OmegaT is free of any advertisement and the user interface is clear and synoptical ranks OmegaT in the front line of the best free-to-use Computer Assisted Translation Tools.

Across

Another tested CAT Tool was Across. Across is a free-to-use application but it disposes of two possible variants. Users can use either the basic non-paid variant, or it is possible to create an account and have the Premium variant. In Figure 5 can be seen the "Order…" button by which the user can upgrade his/her account to premium. The premium account offers several more functions than the ordinary account; it cannot be bought online, though. In case a user would like to buy this premium account, it is necessary to contact the reseller via an email or telephone.

Installation

The installation itself is very easy and fast. When the installation is finished, the application can be run. Across is asking for the username and password but it is not necessary to fill any data. The application fills the CD key by itself and the translation may begin.

Availability on various platforms

The CAT Tool Across is prepared simply for Windows users, not for any others. This might be a disadvantage; it is, however, possible to run Across under some virtual boxes, such as VirtualBox or Parallels with some Windows contribution. Therefore, Mac and Linux users as well will be forced to follow this option. Also, several more negative aspects will arise when running Across on Mac – namely

the shortcuts. Since the Mac keyboard does not contain some of Windows' keyboard keys, such as Alt or Ctrl, several shortcuts cannot be used.

User interface and the user's guide

As almost every application, even Across contains the user's guide. It is not necessary, though. Everything throughout the whole application is on a perfect and intuitive level and no problems should occur when translating.

Translation progress

Every step that the Across' user does is visible right after starting up the application. On the main dashboard there are several boxes that demonstrate the progress, such as:

 Pending projects, Remaining words, Unassigned tasks, Unconfirmed tasks, Projects due in the next 7 days, Projects overdue, Projects finished in the last 7 days.

Translation memory

This function works absolutely perfectly with no troubles. Whenever some structure is translated and afterwards completely the same structure is gone across in the project, a warning window appears (see Figure 8) and the user is asked whether the same translation should be used or not.

Special features

Across has many beneficent features by which it is able to be a huge step before its competitor. Several of them are listed below:

 Translator's connection with their customers –Across has a so-called Across Language Server for this purpose that allows the translator to connect with his/her customer, therefore the order details are always up to date. This function is available for both accounts, basic and premium. The database – During the installation of Across several databases are installed as well. Afterwards, while translating, Across is saving the translator's translation memory and terminology into these databases. These data can be used at any time or can be stored, etc. Management of terminology – If the translator prefers some terms, e.g. some technical terms, they can be stored and afterwards easily used. Also, it is possible to store terms that should not be translated and so these can be neglected.

Final evaluation

In case a translator does not want to spend higher amounts of money on some paid applications but expects a first-class product, Across will be the best option. Everything works as is expected and the user interface is highly sophisticated. Also, the presence of wizards is very pleasant so that the user does not have to search for settings but everything is presented in setup wizards.

MemoQ

MemoQ is the third tested CAT Tool. This Computer Assisted Translation Tool provides a 45-days-long “testing version” called MemoQ Translator Pro, which is a fully active version of MemoQ. When this 45 day period ends, the user is supposed to buy a full-licensed MemoQ product. MemoQ Translator Pro edition with its price of $770 belongs rather in a more expensive class of Computer Assisted Translation tools. Before the user starts the installation itself, he/she is supposed to choose which product will be used on the MemoQ website. Currently, there are six possibilities – MemoQ Translator Pro; MemoQ Project Manager; MemoQ Cloud Server; MemoQ Server; MemoQ Translator Free; Language terminal. In contrast to Across, which provides all these functions in one single application (of course, only the translator function is free of charge), the MemoQ user has to download and install three different applications.15

Installation

At first, the future user needs to visit the MemoQ web page and select the desired application to download. As mentioned earlier, there are currently six applications to select from. Once the required application is downloaded, the installation itself is extremely easy. A simple pressing the “Next” button is far enough.

Availability on various platforms

MemoQ is a Microsoft application, therefore, it is absolutely clear that OS Windows does support MemoQ. However, a problem appears in case the MemoQ user has a Mac OS X. Since MemoQ is highly connected with OS Windows (at first by .Net, second via Microsoft Word which is crucial to display the source text) it is very problematic to run this application on different platforms than Windows. On every new Intel-based Macs there is a possibility to install an application called Parallel Desktop which is an application that allows a Mac user to run Windows applications. Of course, Parallel Desktop is just an example and there are many others similar to this one. Linux has the same conditions as Mac OS X has – runnable under some virtual machine only

User interface and the user's guide

A brisk user may notice that the user interface is very similar to Microsoft Office products and since MemoQ is a Microsoft product it is understandable. Microsoft Office products are well known for their intuitiveness and easy handling. However, this product does not remain within that range. The upper menu does look similar to Microsoft Office's menus but that is the only similarity. The menu is no longer intuitive, the document is no longer easy to control and the whole visualization is not that classy as the price is. The user's guide is, however, processed uncommonly because there is no integrated user's guide within the environment. The user is supposed to click the "Help" icon which allows the user to visit MemoQ's website where all necessary information can be found

There are several available guides, such as for securing a proper installation along with the activation of the product, a guide for absolute MemoQ beginners and others. These guides are accompanied by several demonstrating videos which can be seen after logging in.

Translation progress

The translation progress can be found in the footer where a small sample of the translated page is presented. Formatting is included in this small sample, therefore, the translator has an overview of the translation during the whole work. The whole document is divided into segments. Each segment stands for one sentence which makes a very synoptical overview of the whole document.

Glossary

MemoQ does not contain a pre-imported glossary function and the user is supposed to import one himself/herself. In a case the translator translates non-technical text, this imperfection may cause troubles because importing a glossary of basic language might be more problematic than importing a glossary of technical language.

Compatible files

As for compatible files is CAT Tool MemoQ an absolute leader. This application supports almost every single file which is an extreme advantage. Since this application is a product of Microsoft Office, it is clear that Microsoft Office files

are going to be supported. The following list will explain more about compatible files:

 Source documents: o Adobe documents – Adobe FrameMaker; Adobe InCopy; Adobe Indesign; Adobe PDF; Adobe Photoshop, o AuthorIT, o DITA, o FreeMind (mind maps), o HTML, o Microsoft Office – Microsoft Word (2003 – 2013); Microsoft Powerpoint (2003 – 2013); Microsoft Excel (2003 – 2013), o Microsoft Visio, o MS Help, o OpenDocument, o Plain text, o Rich Text Format, o Scalable Vector Graphics, o TXM filter, o Typo3 pages, o XML; YAML. Software localization formats: o .NET resource files; HTML; Java; JSON; Multilingual XML; Multilingual Excel/CSV; PO Gettext files; Regex text files; XML; YAML, Bilingual documents: o MemoQ Bilingual Document; MemoQ XLIFF; SDL TradosTag; SDL Trados / Wordfast Classic bilingual RTF; SDL Worldserver; SDLXLIFF; Two-column .RTF; Wordfast Professional TXML; XLIFF; Interoperability Now! XLIFF, Project files: o Handoff packages; SDL Studio package; STAR Transit project; TIPP package, Translation memories: o Translation Memory eXchange; MemoQ translation memories,

 Term bases: o Delimited files; Microsoft Excel term bases; SDL Multiterm XML term bases; TBX files; Translation Memory eXchange; MemoQ term bases, Segmentation rules: o Segmentation Rule eXchange.

Translation memory

The translation memory in MemoQ works perfectly. First, the translation memory has to be created by the user. Then, whenever the translator translates some segment and presses “Ctrl + Enter” buttons, the selected line will be confirmed

and written into the translation memory. Afterwards, if the document contains some recurrent phrases, they will be automatically translated.

Special features

CAT Tool MemoQ disposes of many other features than just the translation environment along with the translation memory.

 Term Basis – the translator imports a file containing a glossary and afterwards, whenever such expression is found, it is highlighted in the right panel. Quality assurance – MemoQ automatically checks whether the translation does not contain any undesired parts, such as recurring numbers, typos, incorrect terminology or formatting. Extensive review – the translator has an opportunity to involve customers, other translators or reviewers to the process of translation review.

Final evaluation

MemoQ is a paid Computer Assisted Translation Tool, which is visible right away. Any free-to-use Computer Assisted Translation Tool is not offering such great variety of compatible files that can be loaded.

Also, the possibility for the user to import his/her own glossary file, Term Basis file and others, is very pleasant.

The translation memory does not work as expected, nor does it work as is usual for MemoQ's competitors. The translation memory could be used as usually, by launching a new translation memory file into which a new terminology is written.

However, the translator could also own his/her translation memory file already and in this case importing such file is quite sufficient.

MemoQ is definitely a great Computer Assisted Translation Tool and within its paid field it reaches top places. On the other hand, the navigation menu is not very synoptical and it takes a lot of time to get used to it. Also, with its price above $700, it is a very expensive option and many free-to-use Computer Assisted Translation Tools are on a similar quality level.

Trados SDL

Trados SDL, the fourth tested application, is one of the favorite and most popular Computer Assisted Translation Tools. The era of SDL Trados (the first company name was Trados GmbH) started back in 1992 by its product MultiTerm. Since that time a long time passed along with a gigantic Trados SDL innovation.16

It is necessary to point out that Trados SDL is another paid Computer Assisted Translation Tool. But as well as MemoQ, SDL Trados Studio 2015 provides SDL Language Cloud, which is a free 30-day long license

The user has three possibilities when choosing a product – SDL Trados Studio 2015 Professional (for a single user or a whole network), SDL Trados Studio 2015 Freelance and finally SDL Trados Studio 2015 Freelance Plus.

SDL Trados Studio 2015 Professional, as was already said, is available either for a single user or for a whole network. The network version provides a possibility of

sharing the license key with several other users, which fits mainly translation companies. On the other hand, the single-user version is, as obvious from the name, for one user only. Since the professional version is the highest version available, it offers the greatest possibilities to the user. The professional version includes functions for additional translation, project management or translation reviewing. This version is for users available either by upgrading from SDL Trados Studio 2014 Freelance Plus, SDL Trados Studio 2014 Freelance, SDL Trados Studio 2014 Professional or SDL Trados Studio 2011 Professional or by buying the full application. Both ways are when compared to the other three tested applications in this thesis, incredibly expensive. For an upgrade from a previous SDL Trados Studio version the user pays at minimum 795 €. In case the user does not dispose of any SDL Trados Studio 2014 version or the SDL Trados Studio 2011 Professional version, there is no other way than buying a full application, which is even more expensive – 2595 €.

The Freelance version of SDL Trados Studio 2015 is provided in two possible variants. Either in Freelance or Freelance Plus. Except for one difference, these two versions are completely the same. The difference is in the possibility of installing the Freelance Plus version on two different PCs simultaneously, which is not provided by the Freelance version. As for the price of these two versions, it is no longer that high, as for the SDL Trados Studio 2015 Professional version. The basic SDL Trados Studio 2015 Freelance costs 695 €, in a case of an upgrade from a previous version the price is reduced to 245 € at the minimum. The Freelance Plus version is a little bit more expensive – the full version is available for 855 € and when upgraded from a previous version, there is a price reduction and the final price is from 160 €.

However, the listed upgrade prices are slightly confusing because it depends on the version the upgrade is done from. For example, if the user decides to upgrade to the SDL Trados Studio 2015 Professional from the SDL Trados Studio 2014 Professional, the price is 795 €, while when upgrading from the SDL Trados Studio 2014 Freelance, the price is more or less three times higher – 2145 €.

Installation

It is again necessary to approach the website of SDL Trados Studio and download an installation file. However, the user is not let to download the file unless a form is filled, which consists of the user's first name, last name, email address, phone number, company name (this part is quite misguided since the future user does not necessarily have any company), and a question whether the user is a student, translator etc. Once the form is filled and the privacy statement is checked, clicking the Submit button leads the user to the download page.

The installation itself is very easy and simple pressing the "Next" button does all the necessary work.

Availability on various platforms

Since the first versions of SDL Trados Studio had been working via Microsoft Word (there was a plugin installed into Microsoft Word and the translation was getting done directly in Microsoft Word), it is highly connected to Microsoft, and therefore, availability on any version of Microsoft Windows is doubtless.

However, availability on other platforms, such as Mac OS X and Linux is not supported at all. Linux and Mac OS X users are advised to use a third-party software by SDL Trados themselves. As is written on their web page: “Although SDL Trados Studio is not officially supported on an Apple Mac computer, we have had many success stories from our users running SDL Trados software via Microsoft Windows on their Apple Macs. For each of these solutions, a copy of Microsoft Windows will need to be purchased from an independent retailer or the Microsoft Windows shop.”

This approach might seem odd, however, for Linux and Mac OS X system users, there is no other way for using SDL Trados Studio. Maybe in the future even such users will be complied and SDL Trados Studio will be supported on Mac OS X and Linux as well.

User interface and the user’s guide

In SDL Trados Studio there is no user’s guide presented, however, the “Tutorials” button is presented, which serves as a redirect on a Youtube channel where the user can find more than one hundred tutorial videos. The user interface is simply perfect. There is nothing the user would have to look for, every single step is accomplished via “Installation Wizards”. The translation windows are synoptical, having the modern appearance. Small icons are usually used for shortcuts, such as “Home page” or “Projects overview”.

Translation progress

When SDL Trados Studio is started, the first displayed page is a page of the current open project with all its files. Each file’s translation progress is displayed

in percentage along with a small progress line. Another progress overview can be seen in the Projects view tab where all projects are shown. In this window a small calendar highlights working days, days until the projects must be done and other useful information.

Glossary

Of course, the Glossary function does exist in SDL Trados Studio. The glossary is not pre- included though. However, the user can either open an existing glossary file or start a new one along with starting a new project. Then, selected expressions can be stored in the glossary file. Such file can be afterwards saved and reopened whenever necessary, for example in another translation project.

Compatible files

Neither in this case SDL Trados Studio is losing its credit since the list of supported files is very extensive and all of them are listed below:

 Adobe

o Adobe FrameMaker - versions 8 and 9; Adobe InDesign CS2-CS3 INX; Adobe InDesign CS4 ICML and IDML; Adobe PageMaker,

 Text files, such as the Microsoft Windows' Notepad,

 HTML files,

 Java Resources,

 Microsoft Office

o Microsoft Excel 2000 – the latest; Microsoft PowerPoint 2000 – the latest; Microsoft Word 2000 – the latest; Microsoft Visio, Open Office documents,

 PDF files,

 QuarkXPress 7 and 8,

 QuickSilver,

 RTF,

 SDL Edit,

 SDL XLIFF,

 SGML,

 Software file formats

o .exe; .dll; .rc,

 Star Office,

 TRADOStag,

 World Server,

 XHTML,

 XLIFF,

 XML,

o Microsoft .Net Resources; OASIS DITA Compliant; W3C ITS Compliant.

Translation memory

In SDL Trados Studio the translation memory works as expected. In case of earlier translations, the user is allowed to insert already created translation memory file. On the other hand, if the user does not have his/her translation memory file created yet, it is a matter of ten seconds until one is created because when starting a new project, the Project Wizard offers such an option. Afterwards, when translating and this Computer Assisted Translation Tool finds a recurring expression, it is automatically translated. What is very pleasant is the SDL Trados Studio's ability to help with the whole translation via the translation memory since not only completely identical expressions are translated automatically. Trados displays the likeness of each segment (if there is any), and if the likeness is high enough, the translation memory copies the whole segment from the primary one and displaying the likeness in percentage. A hundred percent likeness is highlighted by green color while lower percentages have a different color, therefore, the translator immediately knows the current status.

Special features

Of course, in such a great CAT Tool, as SDL Trados Studio without any doubt is, there are many special features, and some of them are highlighted below.

 Terminology Management – this very special feature serves mainly in larger translation/localization companies. This function provides a possibility of creating one approved terminology and share it across the whole company. Then, due to this feature, the global text looks consistent of voice. Quality reports – since the development of SDL Trados Studio is mainly aimed at larger companies, quality

control function cannot be missing. The testing can be done either by some automatic linguistic models or the user can create his/her own one.

Final evaluation

When evaluating a Computer Assisted Translation Tool, the target group of customers should be taken into consideration. This CAT Tool has a target group in large companies translating huge amounts of texts daily. Therefore, SDL Trados Studio is a high-class product.

Trados is a well-structured, highly synoptical and unbelievably easy to use Computer Assisted Translation Tool. Even an unexperienced user will handle most of its functions which are supported by tutorial videos on Trados Youtube channel.

However, with its price for the cheapest product of almost 700 € Trados takes place within the most expensive CAT Tools. Simply, SDL Trados Studio is not generally developed for an ordinary translator, which is obvious even from the highly sophisticated user interface and many useful functions that many other CAT Tools do not dispose of.

Conclusion

As was already stated, even though it seems there are many free-to-use CAT Tools on the internet available, the reality is different. It was not an easy task to find these two tools because many of them had a problematic installation or were not possible to be installed at all. Therefore, OmegaT and Across were chosen. On the contrary, choosing paid Computer Assisted Translation Tools, since it was based on the users' satisfaction, was a matter of a little while.

All four applications were tested and afterwards described in details in two main chapters. The whole description was aimed mainly at the installation, availability on various platforms, user interface and the user's guide, translation progress, glossary, compatible files, translation memory and in the end the final evaluation of each application.

OmegaT is an elder application and it is apparent. There are no setup wizards, no colorful windows, everything is just grey and simple. This simplicity is the best

feature of OmegaT. Nobody will ever have problems with finding some buttons – because there are no presented. There are just text fields and that is all. However, the main advantage of OmegaT is definitely its support among three operational systems (Microsoft Windows, Linux, and Mac OS X) and when compared to other application, even the brand new ones, OmegaT comes as the winner.

Across is one of the newer applications with colorful windows, setup wizards, easy to find and intuitive features, and all is topped by an excellent workflow. Across is a way further than OmegaT, mainly for its many functions. For example, its offer of connection with the translator's customer might be found very pleasant. But Across is supported only under the Microsoft Windows operational system and usage on other systems can be reached only using some third-party software for running a virtual version of Microsoft Windows.

MemoQ probably gained its popularity due to its similarity to Microsoft Office applications. Since these applications are heavily used among the majority of computer users, they are familiar with their usage. However, as was said when describing Across, even MemoQ does support only Microsoft Windows operational system. Other operational system users are forced to use either other Computer Assisted Translation Tool, or to use for example Oracle VM VirtualBox to virtually run Microsoft Windows.

MemoQ is, however, one of the best CAT Tools as for the supported files. Users of this application will probably never have troubles with importing their favorite types of files.

SDL Trados Studio 2015 belongs to the top class Computer Assisted Translation Tools, as by its quality or the astronomically high price. With its perfectly sophisticated user interface ensures maximal satisfaction from the user's point of view. It is necessary to highlight its classy translation memory design. While the other CAT Tools via the translation memory function make sure that the recurrent (identical) text is automatically translated, SDL Trados Studio 2015 provides, along with the automatic translation of identical text also an automatic translation of only a similar text while displaying the dissimilarities of both texts.

However, not even SDL Trados Studio 2015 is supported on Linux or Mac OS X. Users are only advised to use some third-party software to virtually install and run Microsoft Windows.

All tested and described tools - OmegaT, Across, MemoQ and SDL Trados Studio 2015 are great applications and with no doubt help many translators to save their time. One of them is better than the others, though.

For the purpose of the evaluation which application is better, each application will be evaluated for every described feature (installation, availability on various platforms, user interface and the user's guide, translation progress, glossary, compatible files and translation memory). The minimum of point obtained is zero while the maximum is three.

Of course, the final evaluation is based on my personal opinion. However, all these applications were tested via the same text file and compared according to the same syllabus. Some of them did better, some worse.

As for the first aspect, the installation, all Computer Assisted Translation Tools obtained the same, maximum, amount of points – 3. In each case the installation was very easy to carry out and intuitive.

The second part, the availability on various platforms, was won by CAT Tool OmegaT. In this case, each point meant one operational system and as was said before, OmegaT supports all three considered operational systems – Microsoft Windows, Mac OS X, and Linux. Since OmegaT's competitors are able to work only under Microsoft Windows, each one of them obtained only one point.

Another evaluated aspect was the user interface and user's guide. This part was quite even, however SDL Trados Studio 2015 came out as the winner since its user interface was considered as the best along with its guide videos. If some of these tools have a good userinterface, it does not dispose of a well-elaborated user's guide or vice versa. The only exception is SDL Trados Studio 2015 and therefore, only this tool obtained 3 points.

The fourth aspect was the translation progress. Since OmegaT does not actually provide any translation progress, it obtains 0 points. However, the progress of Across and MemoQ is roughly on the same level and therefore both applications received 2 points. SDL Trados Studio 2015 was chosen as the most useful tool as for the demonstration of the translation progress.

The fifth point was the glossary function. And since all tested CAT Tools dispose of a function via which a glossary file can be loaded and used afterwards, every tool earns three points.

Next, as the sixth aspect compatible files were evaluated. As was mentioned earlier in this thesis, MemoQ is the absolute leader as for compatible files, therefore, this tool obtains three points. Two tools were evaluated by two points, Across and SDL Trados Studio 2015. And since OmegaT does not provide sufficient variety of compatible files when compared with its competitors, it earns only one point.

Last but not least, translation memory was evaluated. And SDL Trados Studio 2015 was considered as the best also in this category by obtaining three points. Across also disposes of a sufficiently well elaborated translation memory. However, the level of the translation memory is lower than at SDL Trados Studio 2015, which is the reason of earning only two points. The translation memory is, however, not quite well elaborated in OmegaT and MemoQ, therefore, these two tools earn only one point.

On the last, fourth, place ended OmegaT. OmegaT achieved a respectful score of 13 out of 21 possible points. When its age and the fact that it is used for free is taken into consideration, the score is not bad at all.

The second worst CAT Tool, according to this thesis, is MemoQ. MemoQ represents the second part of this thesis – paid CAT Tools. But even though this

application is paid, the final result is not that high. MemoQ achieved only 15 out of 21 possible points and places just by a single point after the free-to-use CAT Tool – OmegaT.

The second place is occupied by another free-to-use Computer Assisted Translation Tool – Across. Across is way further than OmegaT, either by its modern user interface or its additional features. Across ended up with 16 out of 21 possible points.

SDL Trados Studio 2015 becomes the best Computer Assisted Translation Tool according to the research carried out and many tests done in this thesis. This CAT Tool achieved 18 out of 21 possible points and with its overall quality will definitely not disappoint any future user.

References

1. Austermühl F (1990), Electronic Tools for Translators. Abington: St. Jerome Publishing

2. “The usefulness of CAT tools” (2013) on Translating is an Art [online] Available from https://pbtranslations.wordpress.com/2013/02/11/the-usefulness-of-cat-tools/ [February 11, 2013]

3. “Statistical Approaches to Computer-Assisted Translation” in ACL Anthology [online] Available from http://aclweb.org/anthology/J/J09/J09-1002.pdf [March 2009]

4. “Why translate with a CAT tool ?” on Metatexis [online] Available from http://www.metatexis.com/cat.htm [February 15, 2010]

5. “Frequently Asked Questions” on Across [online] Available from http://www.across.net/en/support/faq/

6. “OmegaT compatibility HowTo” on OmegaT [online] Available from http://www.omegat.org/en/howtos/compatibility.html

7. “OmegaT” on OmegaT [online] Available from http://www.omegat.org

8. “MemoQ Translation Products” on MemoQ [online] Available from https://www.memoq.com/translation-products

9. McArthur T (1998), Concise Oxford Companion to the English Language. Published online: Oxford University Press

10. Bassnett S (2002), Translation Studies. Published online: Taylor & Francis e-Library

11. Cohen J.M (1986), Encyclopedia Americana, “Translation”. Volume 27. [online] Available from https://babel.hathitrust.org/cgi/pt?id=wu.89094370806;view=1up;seq=6

12. “Long term memories: Trados and TM turn 20” on JoSTrans [online] Available from http://www.jostrans.org/issue04/art_garcia.php

13. “CAT Tools History, Advantages & Disadvantages” on Arab Translation Group [online] Available from https://arabtranslatorsgroup.wordpress.com/2013/11/23/cat- tools-history-advantages-disadvantages/

14. Pym A. (2006), Translation Technology and its Teaching (with much mention of localization). Tarragona

List of used symbols

CAT Tool, Translation memory, Segment, Operational system, Platform, Translation, Computer Assisted Translation, Localization, Globalization

WHAT IS MACHINE TRANSLATION?

Machine Translation (henceforth, MT), sometimes called Automatic Translation (AT) has been defined as “the process that utilizes computer software to translate text from one natural language to another” (Systran 2004). This definition involves accounting for the grammatical structure of each language and using rules and assumptions to transfer the grammatical structure of the source language (text to be translated) into the target language (translated text). This definition also stresses the fact that machine translation is not simply substituting words for other words, but like human transla- tion it involves the application of complex linguistic rules especially in morphology, syntax and semantics. This definition was widely accepted till the more dramatic developments in the history of MT took place recently when the statistical approaches to MT have started to

gain ground as will be shown later in the paper. The European Association for Machine Translation puts it simply as “the application of computers to the task of translating texts from one natural language to another.” The association adds that “One of the very earliest pursuits in computer science, MT has proved to be an elusive goal, but today a number of systems are available which produce output which, if not perfect, is of sufficient quality to be useful in a number of specific domains.” (European Association For Machine Trans- lation 2004). Doug Arnold et al. (1994), in their book entitled Machine Translation: Introductory Guide, define it as “the attempt to automate all, or part of the process of translating from one human language to another.” Technology Review in the Christian Science Monitor (April 22, 2004) asserts that “Universal translation is one of 10 emerging technologies that will affect our lives and work in revolutionary ways within a decade.” This translation can be “unidirectional” translating in one direction as in the case of English into Arabic, “bi-directional” translating in both directions as from English into Arabic and from Arabic into English or even multidirectional translation back and forth between more than two languages or language pairs. Another aspect can be added to this definition which is the presence of a com- puter system as an initiative for translation. Hahn (2004) distinguishes between “Au- tonomous” (Unassisted MT) initiative with the computer system where the target is what has been termed in the literature as Fully Automatic High Quality Translation (FAHQT) on one hand and Machine Aided Translation (MAT) where the user is asked to perform post editing and answer disambiguation/clarification questions on the other. Machine Aided Translation is human translation supported by a lot of help from computer systems. This help includes translation memory, lexical data, domain information and organizational support. The human

cleans up after the translation in order to get better results. The text in Unassisted Machine translation results in what has been termed in the literature as “gisting” which is the gist of the source unpolished. (see Napier 2000). Another distinction is made between Human Aided Machine Translation (HAMT) or Computer Aided Translation (CAT) where the machine uses human help and Machine Aided Human Translation (MAHT) where the human uses machine help.

THE EARLY DAYS

W. John Hutchins (1995), an authority and widely published historian on MT, traces the beginnings of machine translation to the 17th century when the use of mechani- cal dictionaries to overcome the barriers of language was first suggested. But it was not until the twentieth century that concrete proposals were made with patents is- sued independently by George Artsouni in France in 1933 for a storage device on paper tape which could be used to find the equivalent of any word in another lan- guage and by Peter Smirnov-Troyanskii in Russia in 1937 who envisioned three stages of editing for mechanical translation; an editor knowing the source language to undertake the analysis of words, a

machine to transform sequences into equiva- lents in another language and another editor to do analysis in the target language. The inspiration for a machine that translates from one language to another, maintains Napier (2000), stemmed originally from code cracking of the second world war. Napier quotes a memo drafted by Warren Weaver who was the vice presi- dent of the Rockefeller Foundation and director of its natural science division and he was also an exceptional mathematician and coauthor of a text entitled "The Math- ematical Theory of Communication." The memo which ushers in the beginnings of Machine Translation reads as follows:

I have a text in front of me which is written in Russian but I am going to pretend that it is really written in English and that it has been coded in some strange symbols. All I need do is strip off the code in order to retrieve the information contained in the text. Weaver, making use of the developments in computing and Shannon's information theory, outlined the prospects of machine translation and suggested different methods. This prompted research in some American universities a few years later, especially at the University of Washington at Seattle, the University of California at Los Angeles, and the Massachusetts Institute of Technology. The first Machine Trans- lation Conference was convened in 1952, and according to The Economist Technology Quarterly (2002), the first public demonstration of a MT system which came as a result of collaboration between IBM and Georgetown University took place in 1954. This early system which was based on "a simple bilingual dictionary with a few rules to determine word order, caused a surge of enthusiasm and funding." According to Wikipedia (2004) the demonstration was widely reported in the newspapers even though it was no more than "what today would be called a toy system." With 250 words

and translating “49 carefully selected Russian sentences into English in the field of Chemistry,” that demonstration pushed the perception of the imminence of MT and more importantly it stimulated research funding not only in the US but worldwide.

Optimism, High Expectations and Disillusion

It was a time of optimism and high expectation for machine translation as most MT historians put it. MT research tried also to overcome some of the basic problems they were facing in MT, chief among which is the problem of the “limitations of simple dictionary-based systems” using more developed approaches involving analysis of source texts using grammatical rules. The Economist (2002) cites the Atlantic Monthly in 1959 declaring that “Today, the computer, or the electronic brain, is well along toward picking up the burden of machine translation.” However, the optimism

was turned into “disillusion” in the middle of the 1960s. The government sponsors of MT in the US formed a committee to evaluate and examine the prospects of this field. The Automatic Processing Advisory Committee (ALPAC) submitted its famous report in 1966 and the report concluded, as Hutchins (1995)

states, that “Machine Translation was slower, less accurate, and twice as expensive as human translation and that there is no immediate or predictable prospect of useful machine transla- tion.” The report added that there was no need for further investment in MT re- search. Instead, the report recommended that research should focus on the development of systems to assist human translators rather than to replace them. The report had a profound influence on MT and funding for pure MT research “dried up.” MT was for a decade after that perceived as a “complete failure.” It is revealing at this point to compare the findings of the ALPAC report to some of the claims made by Systran, a known machine translation software developer, a few years later. Translation and Interpretation (2004), a website, interestingly reports, based on information from Systran, that MT is much faster than human translation. Systran estimates that humans can translate 2000-3000 words a day while Systran’s MT software is estimated to translate 3700 words per minute. Systran also claims that its software has a better memory than human translators in that it can store documents that have been translated and use phrases that have already been trans- lated. Systran also claims that if MT is used to provide a first draft to be polished by a human translator, MT can save both time and money. The decade after the ALPAC report was called by Hutchins “the quiet decade.” Work continued in the US, apparently for political reasons related to the perception of the Soviet danger, on English translation of Russian scientific and technical mate- rials. In Canada, there was a big demand on English-French translation. Systran sold a Russian to English translation system to the US Air force in 1970 and the same system was later adopted by the European Commission. It was in this

decade when, according to the Economist (2002), that demand for translation systems started to emerge in business communities.

Revival

New developments worldwide in the fields of technology, political and socioeconomic trends starting in the 1980s contributed to a revival of Machine Translation advancement and research. These developments include the strides made in infor- mation technology, a rapid fall in the cost of computing power, globalization and increasing demand from multinational companies and governments for translation. These developments are by no means the prime mover of research and development behind MT; they just helped increase the pace of development. Translation and Interpretation (2004) asserts that research and development of MT has been going on since the 1950s "engaging some of the best minds in computing, linguistics and artificial intelligence," and cites an often quoted statement by Steve Silberman from an article of his entitled "Hello, World" which appeared in the publication Wired (May 2000) in which he says:

The dream of translation by computer is older than the high tech industry itself. Before email, before word processing, before command-line interfaces, machine translation – or MT – was one of the first two computer applications designed to act upon words

instead of numbers (the other was code breaking)… But it turns out that really good MT is so hard to pull off that the task exhausted the top-end computing resources of every generation attempting it. Regardless, machine translation R&D is going stronger than ever, fired up by the globalization of the Net. Today, all

over the world, software designers, programmers, hardware engineers, neural-network experts, AI specialists, linguists, and cognitive scientists are enlisted in the effort to teach computers how to port words and ideas from language to language." ("Hello, World," Wired, May 2000) Since the 1980s, many new operational MT systems appeared, were expanded and were driven by the commercial markets. These systems have included: The Georgetown system developed in the mid 1960s; the French multilingual system TITUS; the Chinese English CULT system; the Spanish-English SPANAM by the Pan American Health Organization and the tailor made systems developed by the New York based Smart Corporation; the Systran Russian-English system which was adopted by the US Air Force and the European Community; the System of Logos Corporation. In Europe, the Commission of the European Communities (CEC) sup- ported a lot of work on the English-French version of the Systran. Another system is that developed in Germany and called SUSY (Saarbrucker Ubersetzungssystem), the French German System (ASCOF) and (SEMSYN) for the translation of Japanese sci-entific articles into German. A more ambitious and reputable system developed in this era is the EUROTRA project of the European Communities. This project aimed the development multilingual transfer system for translating among all the Commu- nity languages. In the 1980s, according to Hutchins (1995), Japan maintained the greatest commercial activity where most computer companies developed software for computer aided translation mainly for the Japanese English Japanese markets. According to WTEC Hyper Librarian (1994), machine translation in Japan is viewed as an "important strategic technology that is expected to lay a key role in Japan's increasing participation in the world economy." Examples of Japanese MT industry include systems from Oki (PENSEE),

Mitsubishi (MELTRAN), Sanyo, Toshiba (ASHITACI, HICATS) and Fujitsu (ATLAS). The most sophisticated system com- mercially available was METAL, a German-English system originated from research at the University of Texas at Austin with the support of Siemens which obtained commercial rights for marketing it. (Lehmann 2000: 162) New factors in the 1980s caused serious revival of research and interest in MT to the point where the spirit of the age in translation is captured by Steve Silberman (AAAI Report 2005) when he strongly asserts that "a renewed international effort is gearing up to design computers and software that smash language barriers and create a borderless global marketplace." Most important of these factors referred to are the sharp falls in the cost of computing power, greater demand from governments and multinational corporations and the growing spread of globalization. Instead of work on rule manipulation of the syntax of different languages, according to the Economist (2002) the focus shifted to the development of interlingua systems which depended on translating the source text into an intermediate language or symbolic representa- tion from which it could be translated into any of other several languages.

Dramatic Developments

The most dramatic development in MT took place in the 1990s as computers became more powerful and storage capacities much larger and cheaper. The new development shifts from grammar based approaches to what has been called "statistical approaches" emanating from the study of "corpus linguistics." Statistical translation systems do not depend on underlying grammatical rules any longer. Put very simply, as Farah (2003) put it in an article for the New York Times

(reprinted in the International Herald Tribune), traditional MT relied heavily on bilingual programmers to enter the vast wealth of information on the lexicon and syntax that is needed by the computer to translate from one language to another. A team from International Business Machines (IBM) in the 1990s tried feeding a computer an English text and its translation in a different language, then by the use of statistical analysis, the computer learns the second language. The example given by Farah (2003) is revealing:

Compare two simple phrases in Arabic: “rajl kabir” and “rajl tawil.” If a computer knows that the first phrase means “big man” and the second means “tall man,” the machine can compare the two and deduce that rajl means “man,” while kabir and tawil mean “big” and “tall,” respectively. Phrases like these, called N-grams (with “N” repre- senting the number of terms in a given phrase), are the basic building blocks of statis- tical machine translation. Researchers on MT are focusing more on the quality of MT especially after the introduction of Kevin Knight’s newly developed software package called Egypt/Giza at the Information Science Institute at the University of Southern California. The package made statistical translation accessible to researchers across the US. Farah summarizes the impact of the new development saying:

Today, researchers are racing to improve the quality and accuracy of the translations. The final translations generally give an average reader a solid understanding of the original meaning but are far from grammatically correct. While not perfect, statistics- based technology is also allowing scientists to crack scores of languages in a fraction of the time, and at a fraction of the cost, that traditional methods involved. A team of computer scientists at Johns Hopkins led

by David Yarowsky is developing machine translations of such languages as Uzbek, Bengali, Nepali – and one from "Star Trek." Mackin (2003), in an article interestingly entitled "Romancing the Rosetta Stone," reports on work on translation using statistical approaches, which is being pursued at the University of Southern California. Mackin quotes the USC computer scientist Franz Joseph Och boasting: "Give me enough parallel data, and you can have a translation system in hours.' This boast comes after Och's software was judged best in head to head tests against seven Arabic translation systems (five research and two commercial off the shelf products) and 14 Chinese systems (nine research and four commercial) by the US Commerce Department's National Institute of Standards and Technology. Makin asserts that Och's boast is a reminder of the Greek scientist Archimedes' historic boast when he said "Give me a place to stand on, and I will move the world" after offering a mathematical explanation for the lever. The new approach for translation uses huge volumes of "matched bilingual texts" which are the encoded equivalents to gigabites and gigabites of the Rosetta Stone inscriptions. Elaborating on this model in a workshop at Johns Hopkins University, Och (Makin2003) asserts that the new approach uses statistical models to find "the most likely translation for a given input." Unlike the older approaches used for developing MT commercial systems largely depending on encoding the grammar and lexicon of a foreign language then analyzing and consequently producing English sentences based on hard rules, the new approach tries based on a statistical model to find the English sentence that is the most likely translation of the foreign input sentence. The new approach ignores or "rather rolls' over explicit rules of grammar and traditional dictionary lists of the lexicon in order to have the computer itself match up patterns between Arabic texts and

English translations. Och's work (Makin 2003) is an improve- ment on earlier work on the statistical approach that started back in the late 1980s and early 1990s by Peter F. Brown and his colleagues at IBM's Watson Research Center. Much of the work also, as Makin (2003) states, was expanded originally at Aachen University of Technology (Rheinisch-Westfaelisschen Hochschule Aachen) where Och himself did post-doctoral work. Up to now, it can be generalized that the quality of machine translation leaves much to be desired. It has been improving and it has been being improved by MT research and commercial enterprise because of its immense commercial prospects. But, especially in the case of Arabic, more work is still needed in the area of semantic representation systems without which it is difficult to achieve high quality transla- tion as Aljlayl (2004) concludes while referring to the work of Aljlayl and Frieder (2001), Beesely (1998), and El-Dessouki et al (1998). Evidence reported in MT sources generally indicates that MT has scored real success in two major fields. The first of which is the field of "restricted language" or restricted subject matter where the syntax of the discourse is simplified, vocabulary is predictable and words tend to have one meaning because of the technical nature of this restricted variety. This is the language of technical documents, manuals, maintenance manuals and weather reports. Translation and Interpretation (2004) gives the example of the Meteo system which was developed in Montreal for translating Canadian Bulletins between English and French on a daily basis since 1977. The other field is what has been referred to in the literature as machine aided translation (MAT), computer aided translation (CAT) or "interactive translation" (IT). This kind of translation is intended for pro- fessional translators fluent in two languages and they seek a draft which would save time. Translation.net (1994-2003) lists the tools of CAT to

include translation memory, terminology management, data banks and data bases.

3. OTHER IMPORTANT MT RELATED SOFTWARE DEVELOPMENT

Whenever the notion of Machine Translation is used, it is directly associated with the translation of the written language. The question in order here is whether it is really the case that Machine Translation is restricted to the written variety of language(s) targeted. The answer is negative, because MT work and research have been concerned with so many other aspects than the written language. AAAI (2005) reports on a num- ber of MT related items that could have dramatic impact on the lives of the target populations. These included Speech to speech technology developed by NEC and tested in Papero (Partner –Type Personal Robot) which can translate "verbally between two languages in colloquial tongue (see News 2004); E-translators for universal trans- lation (see Lamb 2004); Speech recognition (see Rupley 2004), Robotalk and pocket

translators (see Twist 2004); the "Verbmobile" a computer that translates between German, English and Japanese developed in Germany and it brings humanity "one step closer to the concept of artificial intelligence" (see Dw-world.de 2001); Comput- erized language translators for the military in several languages (see Associated Press 2002), Paula, a virtual interpreter, a digital English to American Sign Language (ASL) translator (see Associated Press/ National Geographic 2002), Avatars converting spoken English into text taken by avatars and the meaning signed on a screen (see Wurtzel, BBC 2002) Babylon and one way phrase translation system developed by the army, among many other developments.

THE CASE OF ARABIC

It must be made very clear at the outset that Arabic is one of the major languages on which some experimentation in MT was made in the very early days of MT and specifically in the US. When we bear in mind that one of the earliest motives behind MT was breaking the codes and monitoring various fronts of the enemy lines, espe- cially those related to scientific, technical and military literature, we can explain why there was primary focus on Russian in the field of MT in its earliest days. Scanty evidence reported in the literature of the history of MT shows that Arabic ranked very high on the list of languages for which MT tools were to be researched and developed in the US fifty or so years ago. Muriel Vasconcellos (2000) narrated the story of the development of the Georgetown MT project and detailed the role of Leon Dostert in it. The project which depended on what they called General Analysis Technique (GAT) prevailed after the Pentagon demonstration changed and the ini- tials reinterpreted it to stand for “Georgetown Automatic Translation.” Georgetown got the largest grant to be

given for Machine Translation from the CIA using the National Science Foundation as a front for this grant which accounted for 93.5% of the research money made available for MT research and development. That was in the late 1950s. Vasconcellos (2000: 92-930) writes: Even though the CIA (three words which one dared not speak aloud, Vasconcellos wrote earlier) grant was for research on Russian, Dostert believed that insights could be gained from linguists specialized in other languages. With this thought in mind, he invited A F. R. "Tony" Brown, then a professor of Semitic linguistics at the institute, to consider how he might address the MT task. Brown, using a French Corpus, wrote a program which he called "Simulated Linguistic Computer." Eventually he took over the programming of GAT. Arabic, which had been designated a priority language by the US government was taken by Nancy Kennedy, a graduate student at the institute [emphasis mine].... All these people came together at the Friday morning seminar which was the highlight of the week. The various groups and individuals would take turns presenting their work and answering questions from colleagues. Sometimes discussions got rather heated. Yngve (2004) reports that Arabic was one of the languages besides English, German and French which were the subjects of the COMIT project in the late 1950s. Attempts at parsing algorithms in a three step scheme were made. Even though certain aspects of Arabic were researched in the early days of machine translation, the language has always been considered "due to its morphological, syntactic, phonetic and phonologic properties one of the most difficult languages for written and spoken language processing" (Boualem 2003). The same author continues to state in a recent call for papers for a major conference on Arabic Language Processing:

Research on written Arabic language processing started in the 1970s, even before the problems of Arabic text editing were completely solved. The first studies focused primarily on lexicons and morphology. In the past ten years, the internationalization of the WWW and the proliferation of communication tools in Arabic have led to the need for a large number of Arabic NLP applications. As a result, research activity has extended to address more general areas of Arabic language processing, including syn- tactic analysis, machine translation, document indexing, information retrieval, etc. Research on Arabic speech processing has made significant progress due to more improved signal processing technologies, and to recent advances in the knowledge of the prosodic and the segmental characteristics of Arabic and the acoustic modeling of Arab schemes. These results should make it possible to further progress in more innovative areas, such as Arabic speech recognition and synthesis, speech translation and automatic identification of a speaker and his/her geographic origin discrimina tion, etc. In a very interesting paper entitled “Toward Corpus Based Machine Translation of Arabic,” Guidere (2002) distinguishes between two approaches for the study of machine processing of Arabic at large. These are the “particularistic” approaches which delineate the linguistic idiosyncrasies of Arabic and use them for a local processing approach specific to the internal linguistic system of Arabic. Such approaches have been concerned with the morphological and semantic aspects of Arabic lan- guage, especially the trilateral root system. Sakhr, the only and foremost Arabic speaking group working systematically on Arabic, put the particular aspects of Ara- bic at the top of its priorities in developing software solutions. Sakr (2004) asserts: The Arabic language differs tremendously in terms of its characters, morphology and diacritization from other languages, and to

claim otherwise would be a mistake. Furthermore, to import solutions from these other languages would only be at the expense of the unique features of the Arabic language. Systran, a company that has been developing software applications for different languages in the world including Arabic specified the following points which they called “facts that help in translating Arabic” (TranslationSoftware4u.com 2004):

Facts that help in Translation of Arabic

– Arabic is written from right to left in a horizontal form. – Arabic writing sits on the line. – There are no capital letters in Arabic. – Punctuation is similar to English except for comas which sit on the line instead of under the line.

– Arabic uses gender for all known nouns, no neutral ones. – Space is left between words in a sentence.

– Some letters change shape depending on whether they are at the start, in the middle or at the end of the word.

– There are 29 letters in Arabic – with 3 letter sounds which do not even exist in the English language.

– Arabic does not distinguish between vowels and consonants; the use of a small sign on the top or under the letter indicates the pronunciation The other approach which is viewed as complementary to the “particularist” approach is called the “universalist” approach which explores possibilities of applications of

methods already tried on other languages like English or French with or without adaptation. These latter approaches focus on the syntactic aspects of the linguistic system in general. Regardless of which approach has been used, Guidere (2002) inter- estingly concludes that the few systems of machine translation available to and from Arabic primarily concern the Arabic English pair and “in reality constitute improved versions of electronic dictionaries.” Guidere continues to assert that other available applications by well known companies have a restricted coverage of Arabic linguistic phenomena and they are essentially based on specialized dictionaries. These applications are technical translation aids rather than machine translation software packages. As this discussion shows, there is some kind of polarization between work based on the “particular approach” as claimed by Sakhr on one hand and the “universalist” approach already followed by most of the companies producing software application for Arabic.

. MODERN WORK ON ENGLISH/ARABIC/ENGLISH MACHINE TRANSLATION

It is to be pointed out here at the beginning of this section that apart from one company in the Arab world, there has been no serious work on the development of software packages of machine translation or other applications of natural language processing for Arabic in its areas of influence, i.e., where it is a first language. The Arab universities, educational institutions, research centers and other relevant organizations have seriously lagged behind in the area of Information Technology research and Arabic language related natural language processing in particular. They left this kind of research and development to the profit driven markets in the West which benefit greatly from research work at

universities, research centers and educational institutions (see Makin 2003, referred to earlier, and the outstanding work of Och at USC on a statistically based translation system of English/Arabic). There has been individual work, of course, but it has never been coordinated and the individual work has also been done by Arabs working in Western educational institutions (see Aljlayl et al 2004, Aljlayl et al 2001, El-Dessouki et al 1989). The one company referred to earlier is Sakr Software, a branch of al-Alamiah Group and it was established in Kuwait in 1982 and then relocated after the Iraqi occupation of Kuwait in 1990 to Egypt. The company boasts in its literature of appointing no less than 100 specialists in information technology to work on Arabic language software applications. On their website, Muhammad Al-Sharekh, Chairman of the Board of Sakhr Software, summarizes their objectives and research work as follows: Sakhr Software has directed its every effort over the past twenty years towards enlisting computer and communication technology to serve the unique needs of the Arabic language, and not vice versa. The Arabic language differs tremendously in terms of its characters, morphology and diacritization from other languages, and to claim otherwise would be a mistake. Furthermore, to import solutions from these other languages would only be at the expense of the unique features of the Arabic language. This was apparent to us from the start, so we undertook the development of the Auto- matic Morphologizer, the Automatic Diacritizer and the electronic dictionaries. Our ardent dedication to the research and development of these tools formed the corner- stone of all future Sakhr products and Internet solutions, including Automatic speech Recognition, Machine Translation and Electronic Publishing solutions. Our investment in Natural Language Processing, or NLP, has yielded outstanding results, and is expected to yield even greater results

as the demand for Natural Language Processing in the exchange of information grows bigger. Indeed, it is worth stating here that Sakhr, whose products go under other names internationally (See Armedia 2004), has been successful in the creation of a number of "CORE" software which makes the processing of Arabic possible and easier. It makes the task of developing software packages for various applications such as machine translation, publishing, and development of electronic dictionaries possible. Sakhr has done to Arabic the basic work that was done earlier to European languages and mainly to English. Following is a list of the software developed by Sakhr to make Natural Language Processing of Arabic feasible. The following list was reported by Armedia (2004).

1. Arabic Optical Character Recognition (A-OCR). Work on developing OCR started in 1993.
2. 2. Multi-Mode Morphological Processor (MMMP) This package is claimed to be a morphological analyzer-synthesizer of Arabic. The analyzer identifies all pos- sible stem forms of a word, i.e., extracting its basic form stripped from affixes. Unlike the English Stemmer, the MMMP analyzer does not stop at the stem level but proceeds to extract the root and the Morphological Pattern (MP) of the word. Decomposing Arabic words into their morphological primitives is a basic requirement for full text indexing, search, dictionary organization and look up, as well as for spelling and grammatical checking. Even more important, the MMMP enables deeper processing of Arabic at the syntax and semantic levels. The MMMP synthesizer works in a reverse mode to generate linguistically-correct final word forms. The synthesizer is a key tool for generating the required

output in machine translation systems and other text generation applications, such as summarizers and style checkers.

3. 3. Multi-mode Syntactic Processor (MMSP) MMSP parses the Arabic sentence into its constituents Verb, Subject, Object, Adverb, Predicate …etc. 4. Arabic Automatic Diacritizer (AAD). AAD handles unvowelized Arabic texts. In other words it provides the Arabic diacritical markings which indicate syntactic functions of the word in the sentence. It is claimed to stimulate "the mental process exercised by Arabic native speakers in interpreting undiacritized text and substituting missing vowels. The Automatic Diacritizer provides different options for diacritization: full, mandatory, or case ending diacritics. The AAD is the entry point for rendering written Arabic text suitable for serious computation.".

   4. Arabic Text Fragmenter (ATF) ATF automatically divides the text into sentences. It serves as a basic front-end processor, which prepares narratives for sentence- based processors such as parsers and for machine translation.

   5. Arabic Automatic Indexer (AAI) examines the content of a document to identify key words and phrases. It enables the creation of book indices for Arabic books. AAI has different levels of indexing and has an HTML version for the Internet.

6. .Arabic Text to Speech (TTS) and Automatic Speech Recognition (ASR) engines. The TTS engine converts any Arabic computer readable text into a human sounding synthetic voice. The ASR engine recognizes Arabic utterances and commands from different speakers and different accents.

8. The Summarizer. It is used for summarizing Arabic and English documents. It extracts the main ideas, based on linguistic analysis of the document, to make it possible for the user to preview these ideas instead of reading the whole document,

9. Johaina. A news search engine which translates news from different sources into Arabic. It has both a navigation and monitoring service. (see Armedia 2004-Johaina)

10. IBSAR. An integrated bilingual solution for the blind or visually impaired, in the Arab speaking countries, Ibsar for Windows works with any PC to provide access to most of today's software applications and the Internet. With its Arabic/English Text To Speech (TTS) engine and the computer's soundcard, information is read aloud, providing access to a wide variety of information.

IBSAR also reads docu- ments, printed books with Sakhr's Optical Character Recognizer and enables users to print on normal and Braille printers. (see Armedia 2004-Ibsar). According to the company's website, IBSAR has the following features:

• Provides self-learning for the blind using a computer.

• Maintains the user's privacy and independence.

• Reads the screen's output and any pressed key on the keyboard

. • Spells every word in a program.

• Integrated Spell checker.

• Reads and navigates any Web page in IE directly.

• Collects links within web pages and then uses the keyboard to select the desired link. • Searches for any (Arabic/English) information on the Web.

• Writes, reads and sends email messages using Microsoft Outlook or popular web- mail services like Hotmail and Yahoo.

• New Integrated Arabic. Speech synthesizer to speed up reading.

• Integrated OCR 7.1.

• New enhanced Integrated Arabic diacritizer.

• Reads all the details within the active dialogs and windows.

• Supports Microsoft Excel and Microsoft Word.

• Reads tables in Microsoft Word. • Supports fast printing using a Braille printer (full control in Word, Notepad, etc).

• Integrated Dictionary. 11. IDRISI general search engine. It is claimed to have the following features. (see ArmediA 2004-idrisi)

• Integrates with INSO filters that support all Microsoft Office formats and other stan- dard formats such as HTML, TXT, RTF and WRI.

• Supports multiple code pages, whether from different platforms (such as Macintosh) or different languages (such as French).

• Complies with Windows NT security schemes, with control of access to data.

• Is capable of handling, indexing and updating data automatically allowing end-users to get the latest and most accurate information without manually rebuilding and re- indexing collections.

• Is a fully bilingual search engine, supporting both Arabic and English features.

• Provides users with customizable search and result templates. These templates cover all the options that an end user may need, from simple or compact results to advanced detailed information about the number of hits.

12. Arab DOX. An English Arabic French Document management system. (see Aramedia 2004-ArabDOX)

13. Sakhr Corrector. The corrector automatically detects and corrects Arabic spell-ing mistakes as well as grammatical mistakes. (see AramediA 2004-corrector)

14. Sakr Categorization Engine. Sakhr Categorization Engine and organizes valuable information into a topic, tree or taxonomy. (see AramediA 2004-categorization)

ACCELERATION OF WORK ON ARABIC

Apart from Sakhr, the interest in developing translation related software packages for Arabic has been sharply growing for three major reasons. The first is related to glo- balization, developments in information technology and the international giant strides in communication. The second is commercial and is represented by the many companies in the industrial world which were led by the tremendous prospects for business in the Arab world to invest in the development of different translation related software applications. As will be shown in this part of the paper, most of these countries are either American or Japanese. Systran was founded in 1968 in San Diego in the US. On its home page (Systran 2004-company), Systran maintains that it is “the leading provider of the world’s most scalable and modular translation architecture. Its core technology powers

revolutionary translation solutions for the Internet, PCs and network infrastructures that facilitate communication in 36 language pairs and in 20 specialized domains. Their expertise, they assert, “spans over 30 years of building customized translation solutions for large corporations, portals, ISPs, governments and public administrations through open and robust architectures.” They have done work for major corporations in America including Ford Motor Company, Cisco Systems, NCR, DaimlerChrysler Corporation, PricewaterhouseCoopers, Dow Corning Corporation, and others. They also did work for Google, AOL, Altavista, Apple’s Sherlock Internet Search, CompuServe, Lycos, OracleMobile.com, and others. They contracted the European Commission and the US Intelligence Community. ATA (ATA Software 2004) is another major company producing software targeting Arabic English Machine Translation and it considers itself as a world leader in English Arabic Machine Translation. The company asserts, on its home page (ATA software 2004-company), that it is a London based company specializing in Arabic Software production. It was established in 1992 by programmers and specialists whose experi- ence goes back to the 1980s. In 1995 it released Al-mutarjim Al-arabey. Another known company, especially in the area of electronic dictionaries, notably handheld electronic translators and pocket multilingual talking dictionaries, is ECTACO, an American based company established in 1990. ECTACO asserts, on its homepage (About ECTACO-2004-company), that it is “the world leader in the development and production of electronic handheld dictionaries. LingvoSoft™ is a registered trademark and an ECTACO division delivering translation software.” Since its foundation in 1990, the company maintains that it has produced 7 generations of electronic dictionaries of Language Teacher® and Partner® brands for over 45

languages. Another company with products in the Arabic market is CIMOS, a French based company that was established in 1997. The company, according to its profile (CIMOS 2004-profile), is interested in Translation and Localization Multilingual Processing Development of linguistic Software. Their products have included Automatic Translator Software (English, French, Arabic), Universal Semantic Analyser General and Specialized dictio- naries (English, French, Arabic), Topic and thesaurus dictionaries (English, French, Arabic), and Tools for NLP (Natural Language Processing), Tools for NLU (Natural Language Understanding), Translation and Localization Services. CIMOS produced morphological Analyzer, grammatical analyzer, automatic vocalizer and a number of electronic dictionaries. It is worth stating that the bulk of the work has so far focused on Arabic/English/ Arabic. More recently, and specifically after the events of Sept the 11th, the demand for translating Arabic increased. Many companies "immediately made Arabic translation a priority. The Newsletter of Systran Translation Solutions (TranslationSoftware4u. com) (2004) quotes Everett Jordan, Director of the National Virtual Translation Cen- ter, an organization jointly sponsored by the FBI and CIA under the USA Patriot Act. When talking about their English/Arabic/English translation software they say: Linguistics technology is beginning to play an increasingly important role when it comes to ensuring national security. Because of the enormous volume of multilin- gual intelligence information that must be analyzed with limited human resources, technologies that can assist in sifting, sorting, and finding critical information are essential in ensuring that threats are detected as quickly as possible. Whereas the US Government cannot endorse any one product over another, we are pleased to

see that companies are responding to the government's call for solutions to these difficult issues.

SOME AVAILABLE SOFTWARE APPLICATIONS FOR ARABIC

The translation related software applications available can be classified in the follow- ing major categories:

.1. Unidirectional, bi-directional and multidirectional general machine translation sys tems.

.2. Translational systems directed to Web translation.

.3. Computer aided translation systems

..4. Unidirectional, bi-directional and multidirectional electronic dictionaries.

.5. Other translation related software packages.

.1. Unidirectional, bi-directional and multidirectional general machine translation systems:

– Tarjim: A bi-directional English/Arabic English Machine translation system developed by Sakr and available on Ajeeb website (http://tarjim.ajeeb.com/ajeeb/). It also translates web sites and web pages from English into Arabic and vice versa. – Al Mutarjim Al Araby V.2.0, V3.0 A machine translation system For English/ Arabic developed by ATA Software Technology. – AL-Wafi Translator V.2.0 (discontinued), a smaller version of Al Mutarjim Al Arabi.

– Al-Wafi v4.00. English/Arabic translation system developed by ATA Software. – Al-Misbar. ATA Software English/Arabic uni-directional translation system as well as a website and internet (URL) translation system. It is an online free ser- vice. (see ATA Software –2004)

– MutarjimNet. A translation system for companies and institutions that provides a translation network for their employees. (ATA Products 2004-Products). – Al-Nakel El-Arabi: Machine translation system for English Arabic. General dictionaries with over 100,000 words and phrases. Special dictionaries on banking, commerce, computers, law, petroleum, gas production and trade. Available are English/Arabic bi-directional Al-Nakel, Single directional English Arabic, Single directional Arabic/English and bi-directional French and Arabic. (see Aramedia 2004-nakel) advertised also by CIMOS (2004-index).

– TranslateNet. Arabic/English/French, also translates web sites. Developed by CIMOS. – English to Arabic Translation/Arabic to English Translation. A bi-directional translation system developed by Systran.

– Weinder WCC. Machine translation system for English/Arabic, developed in the USA, commercially available since 1980; Micro CAT for PCs and Macro CAT for Minicomputers. – Multilingual Machine Translation, developed for a dozen pairs of languages. It goes under the name of Apptek products. (see Aramedia 2004-nakel) – SYSTRAN 5.0. A Machine translation system developed by Systran for bi-direc- tional machine translation of a number of world language pairs including Ara- bic/English and English/Arabic. (Systran 2004-Systransoft) SYSTRAN'S features include (SYSTRAN 2004-mainentry): • Translates directly from your Office desktop with Microsoft® Office plugging for Word, Outlook, Excel, and PowerPoint files. • Real–time translation of Web pages with plugging for Internet Explorer™. • PDF plugging allows you to easily translate business–critical PDF documents into Word. • Includes 5 specialized dictionaries – Business, Colloquial, Industries, Sciences, Life, covering 20 domains. • Integrated Translation Memory – supports TMX standard.

2. Translational system directed to Web translation:

– Sakhr Enterprise Translation (SET) (see Armedia 2004-Set) Web based English Arabic bi-directional Machine Translation Solution. It is targeted for companies and institutes that have a high flow of information to be translated. It translates documents and web pages and it can create memory databases and glossaries to assure consistency of translation. SET is claimed by Armedia to have the follow-ing features:

1. Automatic translation from English ⇔ Arabic.

2. Ability to save your translated sentences in a database called translation memory for future reuse. This will enhance the translation quality especially when you translate files in the same domain.

3. Ability to add your own glossaries to enhance and customize translation output.

4. High-speed translation for the common file formats: html, rtf, and txt.

5. Web pages translation, of any size.

6. Lookup for meaning of words in a bi-directional dictionary, Arabic ⇔ English.

7. Integrated bilingual spell checker to correct spelling mistakes before translation.

8. Administrator has full control on the account and can define access rights for each user.

9. Ability to submit any document for high-quality human translation. Our professionals translate this document in a timely manner with a competitive price.

10. Easy access from anywhere, just type your login name and password.

11. Simple, user-friendly web interface with a detailed help section.

12. Benefit from daily updates and ongoing developments on our engines and glossaries. Sakhr's Easy Lingo Instant Translator.

3. Computer aided translation systems:

– CAT Translator Workbench, CAT Translator Enterprise. A Translator Workbench developed by Sakhr and meant to be a Computer-Aided Translation system that sup- ports bi-directional, bilingual translation between English and Arabic. Sakhr Translator Workbench System is said to apply Natural Language Processing (NLP) technologies to both Arabic and English as source languages. (see Armedia 2004-catrans)

.4. Unidirectional, bi-directional and multidirectional electronic dictionaries.

– Al-Wafi School Dictionary V1.00. English /Arabic/English.

– Al-Wafi English/Arabic/English developed by Ata Software. – Pocket Electronic Translators:

– ECTACO English/Arabic/English Talking Partner for Pocket PC.

– ECTACO language Teacher English/Arabic/English

– ECTACO Police Speech Guard PD-4

– ECTACO Military Speech Guard GI-4a Arabic, English – ECTACO Medical Speech Guard MD-4

– ECTACO Partner EAF430T Arabic – French – English – Bidirectional Trilingual Talking Arabic/English/French Dictionary. Talks English and French. It has medical, technical, legal, business terms, as well as slang, idioms, and general expressions. (see Aramedia 2004-talkdic2)

– Sakhr's Bilingual Dictionary Al-Qamoos. Arabic /English/Arabic Dictionary with Arabic Synonym Dictionary, English Synonym Dictionary, Arabic Ant- onym Dictionary, English Antonym Dictionary. (see Aramedia 2004-diction)

– E. W. LANE Arabic English Lexicon: The special feature of this lexicon is the fact that entry in Arabic is by root. The different derivatives of the root are translated into English. (see Armedia 2004-ewlane)

– Sakhr's Al-Qamoos Multilingual Dictionary. Arabic/English/French/German/ Turkish (see Aramedia 2004-dictionaries)

– World Translator™ It provides the most advanced Multi-Language Translating Dictionary in over 100 languages. (Multimedia 2004-world trans dictionary)

– World Translator™ Multi-Language Translating Dictionary Packages Contents: Arabic ⇔ Arabic Dictionaries Arabic ⇔ Arabic, English & French Dictionaries English ⇔ Farsi Dictionary English ⇔ French, German, Spanish, Italian Brazilian Portuguese

– Standard English ⇔ French, German, Spanish, Italian Brazilian Portuguese

– Professional English ⇔ French, German, Spanish, Italian Brazilian Portuguese – Advanced Multi-language European Dictionaries Package (25 Languages, 30 Dictionaries) Bosnian, Bulgarian, Croatian, Czech, Danish, Dutch, Finnish, French, German, Greek, Hungarian, Icelandic, Italian, Latin, Norwegian, Polish, Polish (law), Polish (Business), Portuguese (Brazilian), Portuguese (Portugal), Romanian, Russian, Russian (Aero). Russian (Business), Russian (Mine), Serbian (Cyrillic), Serbian (Latin), Slovenian, Spanish, Swedish.

– Sakhr's Multilingual Islamic Dictionary. A dictionary intended for Muslim scholars and those interested in Islamic religion. It has the following features (see Aramedia 2004-dictionary): The program includes more than 55,000 items and

meanings, translated into six languages: English, Malay, Indonesian, Turkish, French and German. It is outstanding in its simple display, fast search and accurate meaning of items included in the fields of Faith, Jurisprudence, Economy, History, Islamic art and Architecture, in addition to sections tackling many other fields supported by multimedia technology. It displays an indexed dictionary of 35,000 words from the Holy Qur'an with their meaning. It also cross-references these words to verses where they are found with recitation. It includes the meaning of more than 2,500 difficult words mentioned in the Prophetic Hadiths, linking them to the Hadiths containing them. It also includes terms on the Science of Hadith and their meanings. It presents a brief explanation of many juristic terms including terms of Zakah, transaction, Islamic economy in addition to an Economic Subject Tree linked to the Qur'anic Verses. The history section includes historical terminology and the most important battles, conquests, treaties and historical places. It provides an Islamic World Atlas focusing on the significant information on each Islamic country.

.5. Other translation related software packages.

– Sakhr software, for example, has been working in Arabic NLP for about 12 years, and in English NLP for about 4 years now. 3 years ago, Sakhr initiated their Bidi English ⇔ Arabic MT project (Translate Model). Sakhr has now produced ver- sion 1 of the MT engine that has been used to develop Tarjim.com site, which is the first Web on-line translation English to Arabic on the Internet of which the average accuracy is ~ 60%.

Machine translation is a long chapter in the history of the collaboration between linguists and computer scientist and it has gone through ups and downs depending on many non-linguistic and non-computer related variables. It has

depended on the political atmosphere prevailing at certain times, on economic factors, on availability of funding, on how convinced the decision makers in both governments and the private sectors are concerning the feasibility of this project. More recently, with the revolutions in information technology, the strides in computer develop- ment at both levels of hardware and software, and with the political atmosphere overwhelmed by globalization and concern for security, MT has been given real boosts and with the collaboration of research institutes and industry, attemps are being made to perfect MT. Moreover, MT has prospects for dramatic impact on the lives of the deaf and the blind. Undoubtedly, the bulk of the work has been on the translation of English into other languages and the translation of some languages of interest to the US in particular into English. Pashtu, for example, does not rank high among world languages in terms of number of speakers, but it is a language of concern for the US being one of the languages of Afghanistan. Arabic is an important language in all aspects mentioned. It is the sixth language in the world as to the number of speakers, it is of high concern for the US at the political level, and it is potentially a tremendous market. Despite these factors, work on Arabic lagged behind work on major languages in the world. Work on Arabic started effectively in the 1980s and accelerated, along with MT work on other languages, in the 1990s. It has been shown to work in English/Arabic/English. Work on MT for Arabic started in the West by research institutes at American and some European universities and by Western profit driven companies. The establishment of the one Arab company,

Sakr – Al-Alamiyyah Group, with serious interest in developing soft- ware applications for Arabic, has had a significant bearing on MT as well as on devel- oping different programs for the natural processing of Arabic for different purposes. Major software applications for unidirectional, bi-directional and multidirectional translation developed by Western companies as well as Sakhr were listed. The listing remains partial because of the expanding market for some of these applications. If we take electronic dictionaries, for example, we will find there is a multiplicity of these dictionaries developed by Western companies, and a growing number of these being developed by Japanese companies.

REFERENCES

AAAI (2005): "Machine Translation" <http://www.aaai.org/AITopics/html/machtr.html> About ECTACO (2004): "About ECTACO) <http://www.ectaco.com/company/index.php3?refid=831> <http://arabic.dictionaries.ectaco.com/> Aljlayl, M., Frieder, O. and D. Grossman (2004): "On Arabic-English Cross Language Informa- tion Retrieval: A Machine Translation Approach. Aljlayl, M. and O. Frieder (2001): "Effective Arabic-English Cross-Language Information Retrieval via Machine Readable Dictionaries and Machine Translation," ACM Tenth Confer- ence on Information and Knowledge Management, Atlanta, Georgia. Aramedia (2004): "Arabic Software, Desk Top Publishing, Machine Translation, Document man- agement, NLP, OCR, ASR, TTS and Multimedia" <http://www.aramedia.com/nlp.htm> <http://aramedia.com/speech.htm> <http://aramedia.com/set.htm> http://aramedia.com/johaina.htm

Cultural Problems of English –Arabic Translation

The cultural factor in translation is indisputable as no meaningful communication can take place unless the message transmitted through texts is well understood by the communicants. But understanding can be achieved only if the information contained in language units is accompanied by background knowledge of facts referred to in the message. People belonging to the same linguistic community are members of one culture. They share in common many traditions, habits, and ways of life. They also have common knowledge about the geography, history, and climate of their country, its political, economic, social and cultural institutions, as well as accepted morals and taboos of their society. All these cultural identities are the basis of the communicants' presumptions which enable them to produce and to decode messages in their linguistic form. It is therefore

understandable that in inter lingual communication or translation this common knowledge may be critically limited thereby inhibiting understanding and affecting translation output. Researchers have observed that general problems of translation include lack of exact equivalent in the target language due to cultural difference, semantic differentiation, grammatical structure, dearth of vocabulary etc. Isabel Alousque, while explaining the implication of cultural domains for translation observes that: "Words encoding cultural information are difficult to translate since they involve cultural knowledge and a cultural background. Literal translation may not fully render the meaning of culture bound words because they do not have the same semantic range in the source and the target languages". These cultural problems could be divided into political, social, religious, material and ecological (Nida1964). There are differences between Western and Arabic cultures, which may cause problem in English - Arabic translation. Therefore, finding translation equivalents for cultural terms would require bridging the cultural gaps between English and Arabic cultures so as to meet readers' expectations.

. Linguistic Problems of English – Arabic Translation

It is particularly necessary for English – Arabic translators to be conscious of the differences between English and Arabic linguistic systems, as differences between the two systems may cause problems in translation, and affect the quality of the Arabic translation (i.e. the Arabic output). Studies have identified morphology and

syntax (such as adjectives, adverbials, prepositions, conjunctions etc) as some of the areas of linguistic problems in English – Arabic translation because both languages differ in morphological and syntactic systems. Unlike Arabic, some English words may belong to more than one word class such as "fire", which can be both a noun and a verb. On the other hand, some grammatical features which are expressed by morphology in Arabic such as passivation are expressed in English by syntactic features. One word in Arabic can frequently be a complete sentence in English. For example, the Arabic "I know him" contains the verb, the subject and the object ـه. Further, the morphological endings of words can overturn the functions of nouns. For example, in "Zayd met 'Amr", Zayd is the subject, while in "Amr met Zayd" Amr is the subject. The syntax in the two sentences is the same (verb, subject and object) but the morphological difference reverses the functions of the two nouns.

Qualities of a Translator

Translation as an art and a science is a difficult task and requires constant practice to acquire expertise. It demands sound knowledge of various branches of the source language as well as that of the target language.

The task of the translator goes beyond reliance on dictionary; rather he should be able to present the real and appropriate meaning in a language construction in vogue, so that he can express the ideas of the source language.

The flair, skill and experience that are required by a good translator are similar to the qualities that are needed by an ‘original’ writer; hence, writing and translating

often go hand in hand. Experienced members of the Translators Association have therefore produced a profile of a good translator to include the following:

- the translator needs to have a strong interest in language and a feeling for it.

- s/he must have a good knowledge of the source language and of the regional culture and literature, as well as a sound knowledge of any special subject that is dealt with in the work that is being translated.

- the translator should understand the meaning of the source text and be able to present the meaning in the target language through proper grammar and a good writing style.

- s/he should have linguistic competence in both source language and target language so that s/he can analyze grammatical and lexical relationships in texts and deal with complex structures to determine the intended meaning.

- the translator should be familiar with other works of the original author. - s/he must be an experienced and innovative writer in the target language and nearly will always be a native speaker of it.

- the translator should always be capable of switching between different styles in the language depending on subject of the work being translated. - the translator should always aim at conveying the message of the original work and not just rendering accurate meaning of the words.

- s/he should be competent to produce a text that reads well, and at the same time echo the tone and style of the original work

– as if though the original author were writing in the target language.

Applicatons of MT

Translate the following

“Robert and the Dog” (The Original Version)

Robert’s new employer was a young medical doctor just returned from abroad. He was cheerful, exuberant and polite. It was obvious to Robert that he had not been in the country for a long time. Because he did not once lose his temper, he did not shout at Robert, he called him by his first name and always asked him about his wife, children and other members of his family. Robert, accustomed to moving from household to household, thought he had at last found fulfillment. The more so as the young doctor appeared to be a bachelor. Stewards, including Robert, prefer to serve a bachelor. Because every bachelor is as wax in the hands of his steward. The later determines what is to be spent on grocery, how much found is to be served at meat times, what is to be done with the remnants of food. In short, he holds the bachelor’s life in his hands. And hat is tremendous power. Robert quickly settled into his new situation and took full control of the house. Experience had taught him never to occupy the servants quarters which were attached to a situation rather messy. So it was that Robert’s family lived in the filth and quagmire of Ajegunle which the wags termed The Jungle. In his one bedroom apartment in The Jungle, Robert was king. And he always repaired there nightly to exercise his authority over his wife and six children. The experience he had gained in running his household helped him a great deal in organising the life of every new employer. Robert was particularly happy in his new situation because the young man was carefree and happy. There was, as has been said, no wife breathing down. Robert’s neck and limiting his abundant authority. There

were no children whose nappies and numerous clothes had to be washed. He did not have to cook several meals a day. The young man ate but once a day, except for the cup of coffee and toast early in the morning. Trouble began when the young man announced after six months that his wife was about to join him. Robert's face fell visibly at the announcement. But he did not worry very much at the expected curtailment of his wide powers. Who knew, the lady might not be an ogre after all.

Applications of MT

Translate the following

"A Caring Man" - The Original Version I knew he was having affairs on the side. But I hardly thought he would prove so very degenerate, humiliating me into the bargain. You know how it is. You are holding down a full-time job just to make ends meet in these hard, difficult times. So you cannot do the household chores alone. And when you have two little girls into the bargain, housekeeping and mothering become quite a nightmare. So naturally you need help. And your help had better be a girl because then she becomes a friend to your little ones. Dani never did take an interest in these domestic details. He always thought and said it was entirely my responsibility. And I agreed with him. That was how Ayo came to live with us. Ayo was a skinny, spindly-legged girl just approaching puberty when she came to us. I took one look at her and decided there and then that before she

could say one word to me, I'd have to give her a thorough bath. Which I did. And after I gave her some of my old clothes, she looked more presentable. She was not useful to me at first. But I believed I could train her. She did respond to training. And before very long she became indispensable to me and the household. Dani and I were getting on. At fifty, there were strands of grey in my hair as in his. At that age, a woman is perhaps not so attractive to most men. But a husband, well, you would suppose there are things to think back on, memories of twenty-six years of marriage and companionship. I never did worry about the affairs he had on the side. I did not consider myself cheated. I set it down to the warmth of the African male about which no amount of nagging and tears will ever be able to do much. Besides if he was going to be like his father who married eight wives and had well over three dozen children… I had only myself to blame for marrying him. So he had his fun and I had my marriage. And that was that. Problems actually began when I heard that he had had a baby by some woman. I took it badly; any wife would. But I'd have gotten over it by adopting the child, rearing him with my natural children as a part of one happy or not-so-happy family. My friend Toun did so. And it worked. But Dani's mother didn't give me a chance. She was happy

as a bee at my discomfiture and went out of her way to rub salt into the wound her son had inflicted on my. I never did get on well with my mother-in-law, you see. She was secure in the love of her son. I was not. And I've always hated any hint of rivalry. Now with his new baby and her obvious delight, we were perpetually at daggers drawn. And my anger at Dani was duly compounded. I was not about to leave him – I'm not one for a divorce – and I had to take the children into account. But he had to know that I was not exactly ecstatic, to put it

mildly. It helped a lot that I was often away at wok some nights in the month. It meant that I did not have to see him very often. And I could work out my unhappiness alone. He did a lot to make up to me – he was always a caring man – but I did not encourage him. I'm afraid Ayo and the children saw a lot of our rows. Because they were noisy, robust affairs, full of blows, scratches, broken chairs, broken crockery, flying cutlery and fulsome tears. I gave Dani so quarter. I wanted him to feel remorse and to be on his best behaviour at least for some of the time. He was. As a I was soon to find out. One rare Sunday morning after breakfast, Dani and I sat in the lunge listening to classical music as was our habit when a knock came at the door. 'Who is that?' I asked. 'Mama Ayo,' a voice answered. To my surprise, Dani fed upstairs. He was well dressed and I did not understand why he should make so undignified an exit at the sound of Mama Ayo's voice. 'Come in,' I said Mama Ayo and her husband came in. we exchanged pleasantries and I offered them seats. They didn't want a drink or anything. I heard the key click in the lock upstairs. "Mama Ayo spoke next. She was direct and to the point. 'Is Oga in?' She asked. 'No,' I lied. 'He's gone to church.' 'We wanted to see him. Because now that he has started sleeping with Ayo, we won't be pleased if Ayo gets pregnant.' Bland. Crude. Direct. Shattering. I didn't know what to do. I wished I could hide myself twenty feet underground. That the earth under my feet would open and swallow me up. That thunder would strike me dead. I felt ashamed, embarrassed and humiliated all at once. I tried to, but could not speak. It was with difficult that I said the next words to Mama Ayo and her husband. 'I understand how you feel … we will make arrangements to send Ayo back to you. That will stop Oga from …' Already tears were dribbling down my checks. Mama Ayo and her husband quietly withdrew, leaving me to my sorrows.

I grew weak and tired, as though a heavy load had been placed on my shoulders. Stuck in my seat. I felt dizzy and dazed. There was no holding back the hot tears. I do not know for how long I stayed like that but finally, I picked myself up and dragged my unwilling feet up the staircase. The door to our bedroom was still locked. I knocked loud. 'Open the door, you miserable wretch,' I yelled at the top of my voice. There was no sound from the bedroom. I knocked louder still. 'Open the door, or I'll break it myself, you good-for-nothing adulterer. Do you hear me? You shameless, empty gas bag, let me just get my hands on you.' The key turned in the lock and I pushed the door open and rushed into the room. Dani was sitting on the bed. I do not remember the look on his face. Because although I looked at him, I did not see him, so blinding was the hate I felt for him at that moment. 'Why did you run upstairs and lock yourself in?' I asked. 'I didn't run upstairs. I didn't lock myself in,' he lied. 'Liar. You didn't want to stand face to face with your crime.' 'Crime? What crime?' 'You knew what Mama Ayo and her husband came to tell you, didn't you?' 'How was I to know? How could I know?' 'Then why didn't you wait to hear them out?' 'I didn't even know they were the ones at the door. So what did they say?' 'That you've been sleeping with Ayo. That if she gets pregnant, they won't be pleased.' 'And did you believe them?' Dani asked, a hollowness in his voice. 'Have you been sleeping with Ayo?' 'No.' 'So, Mama Ayo and her husband are lying?' 'Do you believe them?' 'Of course. They could not invent such a story.' 'You believe them and disbelieve me?' 'Naturally. You are a born liar. How dare you stand before me to deny what's so obvious? Why did you flee so guiltily upstairs … you good-for-nothing scab. You call yourself a man and you cannot face your own misdeeds. Now I imagine you will tell me the truth with Ayo standing before you. Because I'm going to get Ayo here right now.' And

getting out of the room onto the landing I yelled ‘Ayo, Ayo, Ayo.’ There was no answer. I went to the helpers’ quarters where Ayo normally stayed. She was gone. I returned upstairs to look for Dani. He too had slipped out. I heard the car drive out of the garage. I lay back in bed and cried myself to sleep.

I do not know for how long I slept. But I must have slept for a long time. Because when I woke up it was already night. I found a note from Dani saying he was off to London for a week on a business trip. He stayed away a whole month. During all that time, he telephoned once asking ‘Is everything all right?’ I banged the telephone on him. At the end of the fourth week, he returned, his hair in jerry coils. At fifty-five, he was wearing jerry coils, like a miserable, cocaine-sniffing West Indian on the crumby streets of Brixton. I took one look at him and the image that came to my mind was of a vulture wearing glass beads. I refused to talk to him, of course. I locked him out of our room and he had to make do with the other bedroom. I made sure his sheets were unwashed for weeks. I did not serve him any of his meals. No punishment I could devise looked bad enough for him. And oh, how I hated to see him in those jerry coils! They made me more vicious. Then one day I heard a car drive up to our house. I looked out of the window and saw a brand new Mercedes Coupe, in a lovely wine colour. Dani was at the wheel. He got out and walked up to me and dropped the car keys into my palm. “Yours,’ he said simply, nodding towards the car. Well, jerry coils are one thing, but a Mercedes Coupe in wine colour is something else. The milk of human forgiveness welled up in my breast. Dani was always a caring man, you know.

Arabic Translation of “A Caring Man” ----------

شَأْنٌ أُسْرَوِي

كانَ *دَابو* أَحَدَ السَّمَّاكِينَ النَّاجِحِينَ في بَلْدَةِ *دُكَانَا*، وَبَدَلَ مِنْ أَنْ يَتَّجِهَ إِلَى النُّهَيْراتِ في صَبِيحَةِ ذاتِ يومٍ، صَدَعَ فُجْأَةً بِغِناءٍ يُنشِدُ، فَحَدَثَ ذُعْرٌ في الْبَلْدَةِ.

وَطَبْعاً، إِنَّ كُلَّ واحِدٍ مُؤَهَّلٌ لِتَمَتُّعٍ خَاصٍّ، وَأَيُّ واحِدٍ شاءَ ذلك، لَهُ حَقُّ تَسْلِيَةِ أَصْدِقائِهِ. وَكانَ بَعضُ الأَوقاتِ مُلائِماً لِذَلِكَ وَمُناسِباً، مِثْلَ نِهايَةِ مَوْسِمِ الْحَصادِ، عِنْدَما يَأْخُذُ الرِّجَالُ الْمُجْتَهِدُونَ الْأَقْوِياءُ الوقتَ للرّاحةِ وَلِمُتْعَةِ ثَرائِهِمْ في وَقْتِ الْفَراغِ: يَتَسَلَّوْنَ بِغِناءٍ وَرَقْصٍ، وبِثَرْثَرَةٍ في الْمَلْعَبِ، وَبِأَفْراحٍ في غُرَفِ النَّوْمِ، بَعِيدًا عَنِ مُشِيعِيْنَ الأَخْبارِ وَثَرْثارِيْنَ. لَكِنْ، إِذَا حَدَثَ أَنَّ رَجُلاً عَزِيزاً، ذا مَكانَةٍ، بَعْدَ تَرْتِيبِ أُمورِهِ جَيِّداً، وَبَعْدَ إِصْدَارِ الْأَوَامِرِ إِلَى مَأْمُورِيهِ أَنْ يَجْتَمِعُوا مَعَهُ بُكْرَةَ الْيومِ الثّاني حَتّى يَخْرُجوا في رِحْلَةٍ؛ إِذا حَدَثَ أَنَّ مِثْلَ هذا الرَّجُلِ نَسِيَ نَفْسَهُ فُجْأَةً، وَحَيَّا مُوَظَّفِيهِ وَأَصْدِقائَه بِأُغْنِيةٍ فاسِقَةٍ، حَقِيقَةً، إِنّ هذا الْأَمْرَ لَيْسَ مِمّا يُحْدِثُ الضَّحْكَ لِأَهْلِ بَلْدَةِ *دُكَانَا*.

في صَبِيحَةِ ذاتَ يَومٍ، لَمْ يَكُنْ أَهْلُ بَلْدَةِ *دُكَانَا* مُبْتَهِجِينَ، حَيْثُ شاعَ الْقَلَقُ تَدْرِيجِيًّا في الْبَلْدَةِ، فَأَسْرَعَ كُلُّ واحِدٍ إِلَى بَيْتِ *دَابو*. وَكَأَنَّ *دُوزِيَا* جَعَلَ الأُمُورَ تافِهَةً، فَسُرْعانَ ما حَضَرَ في المكانِ. "إِنّي أَقولَ أَنَّ لِلرَّجُلِ صَوْتٌ قَوِيٌّ. لَقَدْ كانَ فِي مِهْنَةٍ حاطِئَةٍ مُنذُ الزّمانِ،" وقد ضَحِكَ فَوْرَ ما وَجَدَ فُرْصَةً لِذلك.

"شا شا، أَنتَ يا وَاهِنُ، مُتَرَهِّلُ، ابن كَلْبٍ (بَغِيُّ الْأُمِّ)، اُسْكُتْ. هذا ليس بِأَمْرٍ مُضْحِكٍ،" وَقَدْ نَبَّهَ *تار كُول*.

Word-for-Word transLation

This type of translation keeps the SL word order; words are translated out of context according to their most common meaning. Such kind of translation can be used as a preliminary translation step but it is not applied in real translation tasks. The following lines are from The Secret Sharer by Joseph Conrad with their translation into Arabic following the word-for-word method.

On my right hand there were lines of fishing stakes resembling a mysterious system of half-submerged bamboo fences, incomprehensible in its division of the domain of tropical fishes.

Word-for-word translation will be:

على يميني يد كانت خطوط الصيد حصص تشبه غموض نظام نصف مغمور لبامبو أسيجة غير مستوعب في تقسيمه لمجال الاستوائية الأسماك.

Such translation is meaningless to an Arab reader since neither its grammatical structure nor its semantics can help the reader make sense of what he or she reads. An acceptable translation would be:

وكانت على يميني خطوط لأعمدة الصيد تشبه نظاما غامضا من سياج الخيزران نصف المغمورة غير المراعى في تقسيماتها لحركة الأسماك الاستوائية.

As we can see from the above example, word-for-word translation does not take context into consideration. It might be helpful as a first step in the translation but a great deal of editing and modifications have to be made in order to convey the message to the TL reader. The translator has to make sure that the translation follows the word order of the TL and that the choice of semantic equivalents is carried out accurately. Word-for-word translation can also be called interlinear translation

They were new patients to me, all I had was the name, Olson. Please come down as soon as you can, my daughter is very sick. When I arrived I was met by the mother, a big startled looking woman, very clean and apologetic who merely said, Is this the doctor? and let me in. In the back, she added. You must excuse us, doctor, we have her in the kitchen where it is warm. It is very damp here sometimes. The child was fully dressed and sitting on her father's lap near the kitchen table. He tried to get up, but I motioned for him not to bother, took off my overcoat and started to look things over. I could see that they were all very nervous, eyeing me up and down distrustfully. As often, in such cases, they weren't telling me more than they had to, it was up to me to tell them; that's why they were spending three dollars on me. The child was fairly eating me up with her cold, steady eyes, and no expression to her face whatever.

http://www.classicshorts.com/stories/force.html (accessed 17 October 2013)

LiteraL transLation

This type of translation preserves the grammatical structures of the SL where they are translated into their nearest TL equivalents. It takes place when the SL and TL share parallel structures. Words are translated out of context paying no attention to their con- notative meanings. The following example illustrates this point.

To throw dust in the eyes.

يلقي التراب في العيون

The word 'dust' is translated literally as التراب while the equivalent expression in Arabic is الرماد في العيونّيذر where 'dust' is translated into الرماد, ashes. This choice could be justified by the cultural as well as ecological contexts that are different for both languages. If we take the English idiom to throw dust in the eyes, its associative meaning is based on the effect of dust once thrown in someone's eyes blurring their vision and impeding their ability to see. The same sense is expressed in Arabic by using the equivalent ashes rather than dust. Dust is, therefore, not expected to blur vision and hide reality in the Arabic context. Taking the English setting into consideration, you can rarely talk about deserts or dust storms. Another point has to do with religious rituals where some dust is thrown on the buried person in the grave.

1. I am afraid I lost all saved data. We are back to square one.

2. أخشى أني فقدت كل البيانات المخزنة. عدنا للمربع الأول.

.....

.....

3. She decided to throw the baby with the bath water and close the shop forever because she lost few pounds. بعض الجنيهاتْ. تَ أن ترمي الطفل مع ماء الحمام وتغلق الدكان للأبد لأنها خسر ْرتّقر

.....

3 ..... Hold your horses; we still have plenty of time. أمسك خيولك، لدينا الكثير من الوقت.

.....

.....

Provide a literal and Machine translation for the following excerpt.

As doctors often do I took a trial shot at it as a point of departure. Has she had a sore throat? Both parents answered me together, No . . . No, she says her throat don't hurt her. Does your throat hurt you? Added the mother to the child. But the little girl's expression didn't change nor did she move her eyes from my face. Have you looked? I tried to, said the mother, but I couldn't see.

Faithful transLation

This method maintains a balance between the literal meaning of the SL word and the TL syntactic structures. It sounds more reasonable as it takes the context into consid- eration, aiming at producing more precise meaning of the SL texts. Study the following quote by Shaw's Back to Methuselah (1921), Part 1, Act 1:

'You see things; and you say, "Why?" But I dream things that never were; and I say, "Why not?"'

It can be faithfully translated as:

أنت ترى الأشياء وتقول لماذا؟ ولكني أحلم بأشياء لم توجد أبدا وأقول لم لا؟

3.1 Provide a faithful translation of the following quotes.

1. F or if the proper study of mankind is man, it is evidently more sensible to occupy yourself with the coherent, substantial and significant creatures of fiction than with the irrational and shadowy figures of real life. W. Somerset Maugham (1874–1965), British novelist and playwright

http://en.proverbia.net/citastema.asp?tematica=470 (accessed 12 March 2011)

.. 2. Fiction is like a spider's web, attached ever so lightly perhaps, but still attached to life at all four corners. Often the attachment is scarcely perceptible. Virginia Woolf (1882–1941), British novelist and essayist

http://en.proverbia.net/citastema.asp?tematica=470&page=2 (accessed 12 March 2011)

Requiem By Robert Louis Stevenson Under the wide and starry sky, Dig the grave and let me lie. Glad did I live and gladly die, And I laid me down with a will. This be the verse you grave for me: Here he lies where he longed to be; Home is the sailor, home from the sea, And the hunter home from the hill.

http://www.poetry-archive.com/s/requiem.html (accessed 12 March 2011)

Arabic translation

تحت سماء رحبة مرصعة بالنجوم قبري وتدعني أستلقيْلتحفر سعيدا عشت وسعيدا أموت ومعي أرقدت وصية ثيني بهذه الأبياتْفلتر ار لبيته، عاد من البحرّعاد البح والصياد من التلا

Apply Machine translation to the following stanzas and explain the method(s) you have used.

To a Skylark By Percy Bysshe Shelley What objects are the fountains Of thy happy strain? What fields, or waves, or mountains? What shapes of sky or plain? What love of thine own kind? what ignorance of pain? With thy clear keen joyance Languor cannot be: Shadow of annoyance Never came near thee: Thou lovest, but ne’er knew love’s sad satiety. Waking or asleep, Thou of death must deem Things

more true and deep Than we mortals dream, Or how could thy notes flow in such a crystal stream?

http://www.netpoets.com/classic/poems/057025.htm (accessed 12 March 2011)

Apply MT

1- عصفور في اليد ولا عشرة على الشجرة.

.....

2- في العجلة الندامة .

.....

3- غاب القط العب يا فار. .....

.....

4- الطيور على اشكالها تقع. .....

.....

5- رب ضارة نافعة. .....

.....

THE NOTE TAKER: And how are all your people down at Selsey? THE BYSTANDER: [Suspiciously] Who told you my people come from Selsey?

وكيف أهلك في سرغيتا ؟ وشو عرّفك إني من سرغيتا؟

Selsey is translated as سرغيتا, a village in Lebanon. This choice might be justified by the translator's attempt to adapt the play to the Lebanese setting; thus choosing

a village in the Lebanese countryside as well as a Lebanese dialect in the translation.

5.1 Apply MT

. 1- اشترت مريم فستانا بخمسين دينار.

.....

2- ..... يا اخي ، اقبل دعوتك للعشاء ولو كان صحن سلطة.

.....

3- ..... عيون هذا الطفل جميلة كعيون البقر.

.....

APPLY MT

. 1. الله يعطيك العافية.

..... صبر عليه صبر

الجمال. 2. ..... ارا

3. ك غدا ان شاء الله.

..... عظم الله

4. اجركم.

.....

5. السلام عليكم.

5.4 Translate the following proverbs from English adapting them to the Arabic culture. 1. The proof of the pudding is in the eating.

.....

2. Love me, love my dog.

.....

3. A Jack of all trades is a master of none.

.....

4. Haste is waste.

.....

5. Still waters run deep.

.....

Provide a free translation and machine translation of the following text.

Now two women who were harlots came to the king, and stood before him. And one woman said, 'O my lord, this woman and I dwell in the same house; and I gave birth while she was in the house. Then it happened, the third day after I had given birth, that this woman also gave birth. And we were together; no one was with us in the house, except the two of us in the house. And this woman's son died in the night, because she lay on him. So she arose in the middle of the night and took my son from my side, while your maidservant slept, and laid him in her bosom, and laid her dead child in my bosom. And when I rose in the morning to nurse my son, there he was, dead. But when I had examined him in the morning, indeed, he was not my son whom I had borne.'

http://www.biblegateway.com/passage/?search=1+Kings+3%3A16-28&version=NKJV (accessed 25 March 2011)

GOOGLE TRANSLATE

O. Jabak

INTRODUCTION

The rapid advancements in technology and communication tools have resulted in the creation of many useful applications in almost all aspects of life. By virtue of these applications, communication among distant and different nations has not only improved, but it has also broken physical borders once and for all. One such application or tool is Google Translate which has been in existence and undergoing continuous improvements for about thirteen years. To be precise, Google Translate was devised by Google in 2006 (Wikipedia) to help translate different kinds of texts from, and into, over a hundred living languages. This virtual multilingual machine translator gathers words, expressions and documents from different languages and retrieves them very fast when prompted to translate any given words or texts. So, the more words, expressions and documents it stores, the faster and better it works.

There is no shred of doubt about the usefulness of Google Translate to people from all different walks of life, especially those who want to find the meaning of individual words and some short expressions and sentences in the target language. Franz Och who was the major scientist and head of machine translation (MT) at Google Inc. at that time, explains how Google Translate integrates statistical MT into its system as follows: “what the system is basically doing (is)

correlating existing translations and learning more or less on its own how to do that with billions and billions of words of text. In the end, we compute probabilities of translation" (Schulz, 2013). This is a direct reference to the law of probability which underpins Google Translate. The end translation produced by Google Translate is just a result of these 'probabilities of translation' with plenty of room for errors and inaccuracies. Och elaborates this point further by suggesting that Google Translate's "current quality improvement curve is still pretty steep" (Helft, 2010). In other words, the quality of translation produced by Google Translate is still poor compared to that produced by professional translators. The quality of Google Translate's outputs is also expected to be poorer when translation from Arabic into English is performed as Arabic and English belong to two widely different families whose linguistic systems and cultures are greatly different.

Since Google Translate is relatively new, there has been relatively little research on the assessment of its translation outputs in the direction of Arabic-English translation because most translation in the Arab world tends to be in the direction of English-Arabic translation. The researcher found one corpus-based study on the evaluation of Arabic-English translation produced by Google Translate and Babylon machine systems (Hadla et al, 2014) which makes the present study a relatively new one, despite the availability of little research on the evaluation of Google Translate's Arabic-English translation. However, there have been some small-scale studies on the assessment of Google Translate in the direction of English-Arabic translation.

To begin with, Al-khresheh et al (2018) conducted a study on the translation of some English proverbs into Arabic by Google Translate to see if the Google Translate's outputs are valid and accurate translations. They selected six famous proverbs in English, fed them into Google Translate and compared the resultant translations with valid Arabic translations of these

English proverbs found in Jabak's (2016) book entitled One Thousand and One English Proverbs Translated into Arabic. The researchers discovered that Google Translate could not render the English proverbs into accurate proverbs in Arabic and that it experienced lexical and syntactic difficulties. This general finding supports the findings of the current study regarding the errors and types of errors made by Google Translate when carrying out translation from Arabic into English.

Nabeel et al (2017) conducted a survey on the history and development of machine translation with regard to Arabic-English translation. The researchers only reviewed earlier research on machine translation and traced its development and the tools or applications which have been added or integrated into it. They did not, however, evaluate Arabic-English translation performed by Google Translate, for example, as this fell beyond the scope of their research study. Even the Arabic examples they provided along with their corresponding English translations produced by machine translation were either individual words or very short random sentences which cannot be used to assess the quality of translation produced by machine translation, unfortunately.

Hadla et al (2014) conducted a corpus-based study on the evaluation of Arabic-English machine translation through Google Translate and Babylon machine systems. The corpus consisted of 1033 Arabic sentences with English model

translations. The researchers fed the Arabic sentences into Google Translate and Babylon to evaluate the translation outputs produced by these machine translation systems. The primary finding of their study was that Google Translate produced better translation outputs than Babylon in terms of precision or accuracy. Another interesting finding was that both machine translation systems did not produce intelligible English translations of Arabic wise sayings or proverbs as these systems translated literally without recognizing the sociocultural aspects of Arabic proverbs. The researchers did not mention the type of Arabic text they fed into these machine translation systems, nor did they mention the kind of analysis they followed when they compared the translation outputs produced by the machine translation systems under study with the model translations or reference translations.

Text 1

The following Arabic text is a short, general text with no technical or specialized expressions. لم يكن ثمة ود واحترام متبادلان على الصعيد الشخصي يمكن أن يشكلا في الاتجاه الصحيح لتحقيق الانسجام في العمل على الأقل .ولهذا ًخطوة ٍ وحذ ٍ بقي الجميع يتعاملون بحرص ر شديدين, p. 54) This text was translated into English by Dickens, et al, (. 2017 Brown (1996, p. 43) cited in Dickins et al (2017, p. 54) as follows: There was neither mutual friendship nor respect on a personal level, which would make possible a step in the right direction towards achieving harmony at work at least. Owing to this, their dealings with each other continued to be motivated by overwhelming greed and extreme caution.

The Arabic text was fed into Google Translate to translate into English, and the result was as follows: There has been no mutual respect and respect on the

personal level that could be a step in the right direction at least to achieve harmony in action. That is why everyone remains very cautious and cautious. By comparing Brown's translation to Google Translate's translation, we can spot some lexical errors in Google Translate's translation (respect, in action, and cautious) and one syntactic error (everyone

remains). These errors make the translation poor and incorrect, with a change in meaning as well.

Text 2

The following Arabic text is somewhere between a general and technical text as it includes some technical words ومما لا شك فيه أن حصاد وإنجازات العمل البترولى خلال الـ عاما الماضية هو بمثابة وسام للعاملين بالبترول ومحصلة للسياسات والمجهودات التي تمت خلال تلك الفترة (Dickens, et al, 2017 , p. 118) This text was translated into English by Dickins (2017, p. 118) as follows: No doubt, the achievements of the petroleum sector during the past 18 years represent a triumph for the workers in this sector and reflect the policies and efforts which have been pursued during this period. The Arabic text was fed into Google Translate to translate into English, and the result was as follows: There is no doubt that the harvest and achievements of oil work during the past 18 years is a medal for oil workers and a result of the policies and efforts that took place during that period. By comparing Dickins' translation to Google Translate's translation, we can spot some lexical errors in Google Translate's translation (harvest, oil work and took place). Although these errors do not affect theoverall meaning of the translation, they render the translation as poor and defective.

Text 3

The following Arabic text is somehow a long, general text despite its political language. في استطاعة أي حزب أن ينجح إذا دافع عن قضية الحرية وحقوق الإنسان، إذا احتضن كل مظلوم، إذا قاوم الفساد، إذا ضرب الأمثلة في ل الكلمات إلى أفعال والوعود إلى حقائق .كل ّ إذا حو ، ه إلى ظهره ,Dickens القدوة الصالحة حزب يقف إلى جانب الشعب يقف الشعب إلى جانبه يحيط به عندما الخناجر وإلى صدره المدافع والسيوف. (ّوجُت et al, 2017, p. 61) This text was translated into English by Dickins (2017, p. 61) as follows: It is possible for any political party to succeed if it defends the issue of freedom and human rights, if it embraces every unjustly treated person, if it opposes corruption, if it sets the highest standards in upright behaviour, if it changes words into deeds and promises into facts. Every party which stands by the people will find that the people stand by it, surrounding it when daggers are aimed at its back and guns and swords at its front. The Arabic text was fed into Google Translate to translate into English, and the result was as follows: Any party can succeed if it defends the cause of freedom and human rights if it embraces every oppressed if it fights corruption by striking examples in good example if it turns words into actions and promises to facts. Each party stands on the side of the people. The people stand beside it, surrounded by daggers, guns and swords. By comparing Dickins' translation to Google Translate's translation, we can spot some lexical errors in Google Translate's translation (every oppressed, examples in a good example, beside and surrounded by daggers, guns and swords) and some syntactic errors (if, and surrounded by daggers, guns and swords). These errors make the translation sound very literal and erroneous.

Text 4

The following Arabic text is a short, prosaic text with some colloquial, religious and cultural expressions. " يفتح الله" . وتصلح بها حالك ، تحل منها ما عليك من دين ، يا رجل ً " ."عشرون جنيها وغدا العيد، وأنت لم تشتر بعد كبش الضحية! وأقسم لو أنني أريد مساع ".)Dickens, et al, 2017, p. 52) هذه النخلة لا تساوي عشرة جنيهات ّفإن ،دتك

This text was translated into English by Montgomery (1994, p. 21) cited in Dickins et al (2017, p. 52) as follows: 'No deal!' Look here my man, with twenty pounds you could settle your debts and make your life a lot easier. The Eid festival is tomorrow and you haven't bought a sacrificial lamb yet. As I would not ordinarily pay more than ten pounds for a date palm like this, I would like to think that I am being of some assistance to you. The Arabic text was fed into Google Translate to translate into English, and the result was as follows: God opens up twenty pounds, man, from which you will be able to make a debt, and you will be able to do it. And tomorrow the feast, you did not buy after the victim's ram! I swear that if I did not want to help you, this palm is not worth ten pounds. By comparing Montgomery's translation to Google Translate's translation, we can spot some lexical errors in Google Translate's translation (God opens up, make a debt, you will be able to do it, after the victim's ram and palm). These lexical mistakes render the translation as unintelligible and inaccurate.

Text 5 The following Arabic text is a relatively long, prosaic text typical of novels or short stories. تململ صابر في سريره دون أن يستبد به النعاس وجعل يطوف بناظريه في أرجاء الحجرة: طاولة مخلعة صغيرة، وكتب متناثرة على حصيرة من القش والقصب، وإبريق من الفخار مملوء بالماء وبعض الملابس Dickens, et al, 2017 (الرثة على أحد الجدران. , p. 55) This text was

translated into English by Brown (1996, p. 38) cited in Dickins et al (2017, p. 52) as follows: Saber fidgeted in his bed without feeling sleepy. Instead he let his eyes roam about the room: a small broken table, books scattered on a straw mat, a clay pitcher full of water and some old clothes hanging on one of the walls. The Arabic text was fed into Google Translate to translate into English, and the result was as follows: Saber restrained himself in his bed without being drowsy and made his way around the room: a small pickled table, books scattered on a mat of straw and cane, a jug of pottery filled with water, and some ragged clothes on a wall. By comparing Brown's translation to Google Translate's translation, we can spot some lexical errors in Google Translate's translation (restrained, made his way, pickled and cane). These lexical mistakes render the translation as both unintelligible and inaccurate.

Text 6

The following Arabic text is a long, prosaic text typical of novels or short stories with some religious and cultural expressions.

وقادته خطواته إلى مسجد كبير، وكان يجلس في داخله شيخ له لحية ق من حوله عدد من الرجال .وكان الشيخ يتكلم عن الله ّتحل ،بيضاء وجميع ، الأشياء ّ والشيطان" :الله خلق كل إلا ً المخلوقات لا تعمل شيئا بأمره." فقال محمد لنفسه: إذن يستطيع الله مساعدته على تحقيق أمنيتي." ." وقال الشيخ :إبليس عدو البشر ..إنه الشر ، ل بلهفة ّ تتوس ً وغادر محمد المسجد بينما كانت دماء شراينه أصواتا Dickens, et al, 2017 وتهفت ضارعة: "يا الله". (, p. 43) This text was translated into English by St. John (1999, pp. 7-8) cited in Dickins et al (2017, p. 52) as follows: His feet led him to a large mosque, and inside it sat a religious teacher with a white beard. Several men were gathered round him and he was talking about God and the Devil. Allah is the Creator of all things, and no creature can do anything unless He wills it. 'So

Allah can help me realize my dream,' said Mohammed to himself. The teacher continued. ‘The Satan is the enemy of Man - he is evil.’ Mohammed left the mosque, and as he did so, the blood in his veins became a mass of imploring voices, calling out woefully: 'Oh God.’ The Arabic text was fed into Google Translate to translate into English, and the result was as follows: His steps led him to a large mosque, where an old man with a white beard was sitting, surrounded by a number of men. The Shaykh spoke of God and the devil: "God created all things, and all creatures do nothing but His command." "Muhammad said to himself:" Then God can help me achieve my wish. And Muhammad said to himself: So God can help him to achieve my wish. " And the sheikh said: Satan is the enemy of mankind. It is evil. "And Mohammed left the mosque while the blood of his veins voices pleading eagerly, and raged and raucous: "O God." By comparing St. John’s translation to Google Translate’s translation, we can spot some lexical errors in Google Translate’s translation (an old man, Shaykh, the sheikh, but, voices, pleading and raucous) and some syntactic mistakes (and, raged, raged and raucous). These types of errors make the translation both literal and inaccurate.

Text 7

The following Arabic text is a relatively long, general text despite its political or legal language.

يجوز في حالة إعلان الأحكام العرفية أو الطوارئ أن يفرض القانون على الصحف والنشرات والمؤلفات التي تتصل بالسلامة العامة وأغراض الدفاع الوطني.) Dickens, et والإذاعة رقابة محدودة في الأمور al, 2017, p. 256) This text was translated into English by Dickins (2017, p. 256) as follows: In the event of the declaration of martial law or a state of emergency, a

limited censorship on newspapers, pamphlets, books and broadcasts in matters affecting public safety or national defence may be imposed by law. The Arabic text was fed into Google Translate to translate into English, and the result was as follows: In the case of martial law or emergency, the law may impose limited censorship on newspapers, publications, literature and radio in matters related to public safety and national defense purposes. By comparing Dickins' translation to Google Translate's translation, we can spot some lexical errors in Google Translate's translation (literature and radio). Despite these lexical errors which make the translation sound literal, the meaning of the translation is quite clear.

Text 8

The following Arabic text is a long, legal text with some political expressions. العمل حق لجميع المواطنين وعلى الدولة أن توفره للأردنيين بتوجيه 1 الاقتصاد الوطني والنهوض به. . . تحمي الدولة العمل وتضع له تشريعا يقوم على المبادئ الآتية: 2 أ - وكيفيته يتناسب مع كمية عمله ً إعطاء العامل أجرا ب ـ تحديد ساعات العمل الأسبوعية ومنح العمال أيام راحة أسبوعية وسنوية مع الأجر.) 2017 Dickens, et al, p. 264) This text was translated into English by Dickins (2017, p. 264) as follows: 23. (i) It is the right of every citizen to work, and the State shall provide opportunities to work to all citizens by directing the national economy and raising its standard. (ii) The State shall protect labour and enact a legislation therefore based on the following principles: (a) Every workman shall receive wages commensurate with the quantity and quality of his work. (b) The number of hours of work per week shall be limited. Workmen shall be given weekly and annual days of rest with wages. The Arabic text was fed into Google Translate to translate into English, and the result was as follows:

Article 23 1. The right of all citizens and the State to provide Jordanians with the direction and promotion of the national economy. 2. The State shall protect the work and shall lay down legislation for it based on the following principles: A - Giving the worker a wage commensurate with the amount of his work and how it works B. Setting weekly working hours and giving workers weekly and annual rest days with pay. By comparing Dickins' translation to Google Translate's translation, we can spot some lexical errors in Google Translate's translation (The right of all citizens and the State to provide Jordanians with the direction, and the amount of his work and how it works) and some syntactic errors (The right of all citizens and the State to provide Jordanians with the direction, and the amount of his work and how it work, and the). These errors make the translation sound very literal and inaccurate.

It is obvious that the errors spotted in the English translations produced by Google Translate are not insignificant ones which could be simply overlooked. In fact, they are very serious mistakes which render the translations as inaccurate and faulty. Anyone who knows English can readily identify these errors by simply looking at the model translations and the translations produced by Google Translate because these mistakes are easy to detect. Thus, one could imagine the number and types of errors Google Translate would make if longer texts or more specialized texts were to be translated by it.

Google Translate cannot be relied on to carry out translation from Arabic into English as it cannot always find the correct lexical word or expression suitable for a given context, not to mention the syntactic errors which result from the literal

translation this tool seems to adopt. Such lexical and syntactic errors are bound to surface in this kind of translation because Arabic and English belong to two different families which have completely different linguistic as well as cultural systems. This means that machine translation cannot replace man-made translation, especially when translation is carried out in different language pairs such as Arabic and English.

Based on the findings of the study, the researcher suggests that a more large-scale quantitative study on the assessment of Google Translate's Arabic-English translation be conducted to either support the findings of the current study or challenge them. Another kind of study may assess the accuracy and effectiveness of another translation tool or application such as Microsoft Translator to see if similar or different findings can be arrived at. Of course, the same language pair or a different language pair can be used and general or technical texts with varying lengths can be used as samples to be compared and examined.

REFERENCES

[1] Abu-Al-Sha'r, M. A., & Zughoul, R. M.)2009(. Arabic-English-Arabic online machine translation: An analytical and evaluative study. Turjuman, 18(2), 33-83.

[2] Al-Dabbagh, U.)2010(. Free Online Machine Translation Services as Providers of Gist Translation: A Fiction or a Reality. Paper presented at the 2nd Jordan International Conference on Translation 30th November. Amman, Jordan: Petra University.

[3] Al-Dabbagh, U.)2013(. Assessing the Quality of Free Online Machine Translation Services. International Journal of Translation-IJT, 25(2).

[4] Al-khresheh, M., H., & Shahab A., A.)2018(. English Proverbs into Arabic through Machine

Electronic

Apply GT

مقتطفات من كليلة و دمنة

قال الفيلسوف: زعموا أن حمامة كانت تفرخ في رأس نخلة طويلة ذاهبة في السماء،

فكانت الحمامة تشرع في نقل العش إلى رأس تلك النخلة، فلا يمكن أن تنقل ما تنقل من

العش وتجعله تحت البيض إلا بعد شدة وتعب ومشقة: لطول النخلة وسحقها، فإذا فرغت

من النقل باضت ثم حضنت بيضها، فإذا فقست وأدرك فراخها جاءها ثعلب قد تعاهد ذلك منها لوقت قد علمه بقدر ما ينهض فراخها، فيقف بأصل النخلة فيصيح بها ويتوعدها أن يرقي إليها فتلقي إليه فراخها. فبينما هي ذات يوم قد أدرك لها فرخان إذ أقبل مالك الحزين فوقع على النخلة. فلما رأى الحمامة كئيبة حزينة شديدة الهم قال لها مالك الحزين: يا حمامة، ما لي أراكي كاسفة اللون سيئة الحال؟ فقالت له: يا مالك الحزين، إن فأفرق منه ، دهيت به كلما كان لي فرخان جاء يهددني ويصيح في أصل النخلة ًثعلبا فأطرح إليه فرخي .قال لها مالك الحزين :إذا أتاك ليفعل ما تقولين فقولي له: لا ألقي إليك طرت عنك ونجوت ، إلي وغرر بنفسك .فإذا فعلت ذلك وأكلت فرخي َفارق ،فرخي بنفسي. فلما علمها مالك الحزين هذه الحيلة طار فوقع على شاطئ نهر. فأقبل الثعلب في الوقت الذي عرف، فوقف تحتها، ثم صاح كما كان يفعل. فأجابته الحمامة بما علمها . مالك الحزين. قال لها الثعلب: أخبريني من علمك هذا؟ قالت: علمني مالك الحزين فقال له الثعلب: يا مالك ً فوجده واقفا ، الحزين على شاطئ النهر ً ه الثعلب إلى مالكا ّفتوج. الحزين: إذا أتتك الريح عن يمينك فأين تجعل رأسك؟ قال: عن شمالي. قال: فإذا أتتك عن شمالك فأين تجعل رأسك. قال: أجعله عن يميني أو خلفي. قال: فإذا أتتك الريح من كل مكان وكل ناحية فأين تجعله؟ قال: أجعله تحت جناحي. قال: وكيف تستطيع أن تجعله تحت جناحك؟ ما أراه يتهيأ لك. قال: بلى: قال: فأرني كيف تصنع؟ فلعمري يا معشر الطير لقد فضلكم الله علينا. إنكن تدرين في ساعة واحدة مثلما ندري في سنة، وتبلغن ما لا نبلغ، وتدخلن رؤسكن تحت اجنحتكن من البرد والريح. فهنيئ لكن فأرني ًا كيف تصنع. فأدخل الطائر رأسه تحت جناحه فوثب عليه الثعلب مكانه فأخذه فهمزه

همزة دقت عنقه. ثم قال: يا عدوي نفسه، ترى الرأي للحمامة، وتعلمها الحيلة لنفسها، وتعجز عن ذلك لنفسك، حتى يستمكن منك عدوك، ثم أجهز عليه وأكله. فلما انتهى المنطق للملك والفيلسوف إلى هذا المكان سكت الملك. فقال له الفيلسوف: أيها الملك مع وفور سرورك ً، وأعطيت من كل شيء سببا ، وملكت الأقاليم السبعة ، عشت ألف سنة وقرة عين رعيتك بك، ومساعدة القضاء والقدر لك، فإنه قد كمل فيك الحلم والعلم. وزكا منك العقل والقول والنية، فلا يوجد في رأيك نقص، ولا في قولك سقط ولا عيب. وقد ولا ضيق الصدر عندما ينوبك من ، عند اللقاء ً فلا توجد جبانا ، جمعت النجدة واللين الأشياء. وقد جمعت لك في هذا الكتاب شمل بيان سأ ما لتني ً وشرحت لك جوابا ،الأمور عنه منها فأبلغتك في ذلك غاية نصحي، واجتهدت فيه برأيي ونظري ومبلغ فطنتي، لقضاء حقك وحسن النية منك .بأعمال الفكرة والعقل .فجاء كما وصفت لك من ًالتماسا النصيحة والموعظة مع أنه ليس الآمر بالخير بأسعد من المطيع له فيه، ولا الناصح بأولى بالنصيحة من المنصوح، ولا المعلم للخير بأسعد من متعلمه منه. فافهم ذلك أيها الملك ولا حول ولا قوة إلا بالله العلي العظيم

Analyse GOOGLE TRANSLATION

Philosopher said, claimed that the dove was spawn at the top of a long Palm going in the sky, was the dove initiate the transfer of the nest to the top of that palm tree, it can not convey what is transported out of the nest, and put it under the eggs only after the severity of fatigue and discomfort: the length of the palm tree and crushed, if completed transportation Badt then incubated eggs, if hatched chicks came to her and realized Fox has a compact that which for a time had

taught him as much as her brood rise, to stand in the origin of Palm Faisih out and Atuadha to live up to it thus he received to her brood. While she may some day be realized her Farrakhan as I accept the heron and he fell on the Palm. When he saw the dove bleak sad severe carefree Heron told her: O dove, what I Araki Kasfah bad case color? She said to him: O heron, if a fox Dhat him whenever I had Farrakhan came threaten me and yell at the origin of Palm, Vavriq it shall put him two young. Heron told her: If Attac to do what you say Vcoli him: Do not you were two young, difference to me and tricked yourself. If I did that and I ate two young, I flew about you and I survived myself. When the flag sad this trick owner flew and fell on the river beach. Fox turned to the time who knew, stood underneath, then shouted as he did. Lean dove including flag heron. Fox told her: Tell me your knowledge of this? She said: Teach me Heron. Fox head down to sad owner of the river, and found him standing. Fox said to him: O Heron: If you are faced with your right hand, where the wind makes your head? He said about north. He said: If you are faced with your left hand on where to make your head. He said: I will make him my right hand or rear. He said: If you are faced with the wind from everywhere, all the hand to make it where? He said: Make it under the wings. He said: How can you make it under your wing? What I see is preparing for you. He said: Yes: said: Varney how to make? Vlamra bird O God, please us. You know, nor are in one hour as we know in a year, and what Thblgn not inform, and Tdkhaln Raskn under Ajnankn from the cold and wind. Congratulations But Varney how to make. Bird enter his head under his wing fox leaped upon his place took him Vhmzh connecting rang his neck. Then he said: O my enemy himself, see the opinion of the dove, and learned the trick itself, and is unable to yourself, so you Istmcn your enemy, then finished it and eat it. When he finished the logic of the

king and philosopher to this place minted King. Philosopher said to him: O king, I lived a thousand years, and reigned seven regions, and give everything a reason, and as soon as you have been pleased with the apple of your eye your parish, and help you to fate and destiny, it is fulfilled in you dream and science. And Zacchaeus you mind and speech and faith, there is no shortage in your mind, not in your words fell not defect. Rescue and soft were collected, there is no coward at the meeting, and chest tightness when Anopk of things. You have collected in this book included a statement of things, and I explained to you what you ask him answer them in that very Viblgtk advise me, and it worked hard in my opinion and my opinion and the amount of Aftnta, petitioned to spend your right and good faith from you. The work of the idea and the mind. He came as described to you for advice and counsel with it It's not good Bosaad of obedient to him and there is no mentor FIRST advice advisable, nor the teacher for good Bosaad of educated him. Vavhm so, O king, and no power except God Almighty

Apply GT

والطيرُ كلُّ سَبُعٍ وبَهيمة وهَمَج . والسباعُ من الطير على ضَربَيْن : فمنها العِتاقُ والأحرارُ والجوارحَ ، ومنها البغاث[1] وهو كلُّ ماعظمَ من الطير : سبعاً كان أو بهيمة ، إذا لم يكنْ من ذواتِ السلاحِ والمخالبِ المعقَّفة ، كالنُّسورِ والرَّخَم والغِربان ، وما أشبهها مِنْ لئامِ السباع .

ثم الخَشاش ، وهو ما لطُف جرمُه وصَغُر شخصه ، وكان عديمَ السلاح [ولا يكون[2]] كالزُّرَّقِ[3] واليُؤيُؤ[4] والبادنجار[5] .

فأما الهَمَج فليس من الطير ، ولكنَّه ممَّا يطير . والهمَجَ فيما يطيرُ ، كالحشراتِ فيما يمشى .

والحيّاتُ من الحشرات ، وأىُّ سبعٍ أدخَلُ فى معنى السَّبُعيَّة مِنَ الأفاعى والثعابين ؟ ولكن ليس ذلك من أسمائها ، وإن كانتْ من ذوات الأنيابِ وأكَّالة اللُّحوم وأعداءِ الإنسِ وجَميعِ البهائم ، ولذلك تأكلُها الأوعَال[6] والخنازيرُ والقَنافِذُ والعِقبان[7] والشاهْمُرك[8] والسنانير ، وغير ذلك من البهائم والسباع . فمَنْ جَعَلَ الحيَّاتِ سِباعاً ، وسمَّاها بذلك عندَ بعضِ القولِ والسببِ فقدْ أصابَ ، ومن جَعلَ ذلك لها كالاسمِ الذى هو العلامةُ

Direct Machine Translation Approach:

This is oldest approach and translation is performed at word level. There is no additional intermediary representation be- tween source and target languages. Words of source language text is directly translated into the target language. This is uni-directional bilingual translation system. Direct machine translation approach involves the word by word translation with some modification at grammar level. The translation is not good as it is just the replacement of words from target language into source language text i-e word by word meaning replacement. Interlingual Machine Translation Approach: This approach introduces an intermediary language represen- tation between source and target languages. This intermediary language is called Neutral Language. Neutral language can represent any natural language. It is independent of source and target Languages. It is also useful for multilingual translation machine system.KANT system was developed on interlingual approach in 1992 by Nyberg and Mitamura[5]. Building in- terlingual language is not an easy job. Too much efforts are required to develop truly neutral language. Transfer base Machine Translation Approach In this approach the text of source language is converted into intermediary representation, it is then used to generate the target language text with help bilingual dictionary and grammar rules. Transfer based machine translation process is divided into three phases. Analysis In this phase source language text is analyzed on basis of lin- guistic information and heuristics to parser the text (syntactic representation) Transfer The syntactic representation of source language is converted into the syntactic form of target language. Generation The final text in target language is generate with help of morphological analysis. This approach heavily dependent on the grammar and

structure of sentence and changes to a monolingual component affect all transfer modules for that language. 2) Corpus-based Machine Translation Approach: It is ac- tually data driven machine translation. It was introduced an alternative approach to the rule-based approach. In this ap- proach the bilanguage parallel corpus is used to extract the translation for new sentences. A large amount of raw data is collected in parallel corpora. The raw data is actually the translation between source and target languages and this data is used for translation. The sub-approaches of Corpus-based Machine Translation are Statistical Machine Translation and Example-based Machine Translation. Statistical Machine Translation(SMT) This approach is basis on statistical model. It has two statistical probabilities models: language model and translation model and massive parallel corpora of source and target languages. The advantage of SMT system is that linguistic knowledge is not required for building them. The difficulty in SMT system is creating massive parallel corpus. We have to two models in SMT, one is Word-based and other is phrase-based.

In word-bases MT sentences are consider as combination of single words and structure relation between the words are ignored while in phrase-based model consider sentences as combination of phrases or chunk. The basic concept in SMT

is probability. The probability score of translations are gen- erated from already available translated data (parallel corpus, translated by human), the translation having high probability is selected as final translation. The probability is calculated with help of language and translation models.

Example-based Machine Translation (EBMT)

Example-based machine translation contains the point to point mapping between the source and target language sentences i-e we have examples data that is translated between the source and target language[6]. This data is used for translation. The basic idea is if already translated sentence occur again it, the same translation is likely to be correct again. Basically, EBMT is memory-based translation and the concept of analogy is used for the translation.

Syntactic structures polysemy and ambiguity

The phenomena of polysemy and ambiguity of language structures require special research and development efforts since they are immanent features of natural

language and manifest themselves at all language levels. We focus here on syntactic structures polysemy and ambiguity because they are far less explored than lexical polysemy but they lead to a great number of mistakes and failures of translation if not treated properly. Our experience shows that syntactic polysemy and syntactic ambiguity are not the same phenomena, which has very important practical impact. Hence, their distinctive features should be stated, and the corresponding processing features developed.

Syntactic structures polysemy

By syntactic polysemy we mean the immediate realization of more than one categorial meaning within the head element of a language structure. The polysemous structures display variable manifestation of their categorial features depending on the functional role in the sentence. Consider such language phenomena as the Gerund, the Participle and the Infinitive. The Gerund comprises the features of both the Verb and the Noun, which affects the translation strategy when the appropriate means are to be chosen for representation of the English Gerund via the Russian language forms. The structures similar in category to the English Gerund are the Russian Verbal Nouns denoting “Activity”, e.g. singing □ penie, reading □ chtenie, and both the English Gerund, and the Russian Verbal Noun allow direct object arguments if derived from transitive verbs. However, the direct transfer of the Gerund into the Russian Verbal Noun is the least probable translation variant of the three possible transfer schemes: □

 Verbal Noun or Gerund forms (Eng) Clauses with the Finite Verb forms (Rus) Verbal Noun or Gerund forms (Eng) Clause with the Infinitive (Rus) Verbal Noun or Gerund forms (Eng) Verbal Noun (Rus). This fact can be accounted for by the mechanisms employed in the Russian language for configuring sentential structures and is to be envisaged in the machine translation engine. Consider the other most productive polysemous language structures which comprise more than one categorial meaning: Participle Verb + Adjective Infinitive Verb + Noun Nominal Phrase as the Nominal Modifier Noun + Adjective Verbal Phrase as the Verbal Modifier Verb + Adverb. Thus we introduce the notion "polysemous syntactic structure" to determine the set of possible transfer schemes for a given language structure. When a polysemous structure is assigned specific categorial attributes realized in this structure, the possible and preferable transfer schemes become predictable for the given structure. The predominant categorial meaning of a polysemous syntactic structure (or syntaxeme) is determined by the syntactic function realized at a given moment. Thus the transfer scheme for a "stone wall" construction will be as follows:

Noun1 + Noun2 [Eng.] Adjective + Noun2 [Rus] The weight for this transformation will be higher than for the transformation: Noun1 + Noun2 [Eng] Noun2 + Noun1 (Genitive) [Rus] if the dictionary contains an Adjective as one of the possible translation equivalents for Noun1, that is the case when the dictionary is composed by various methods including acquisition of lexical units from parallel texts. Judging by the function we establish the transfer field [8] within which the translation procedure will be carried out. The Functional Transfer Fields (FTF) support the possible paraphrasing variants and envisage the synonymous ways of conveying the same functional meaning across languages.

The syntactic polysemy manifests the action of the mechanism of the transposition of meaning stated by the Semiotic Universal Grammar (SUG) [9]. The awareness of the mechanism action gives cues to designing the procedures for polysemous structures parse and transfer, the multiple syntactic-semantic data structures being associated with a given syntactic pattern. Of special interest is the situation of the categorial shift in translating a syntactic pattern. The category of a syntactic pattern, i.e. phrase structure, is determined by the category of the head word of this phrase structure. Thus, when transfer employs conversion, and the category of the head word shifts to another category, the whole structure is assigned the feature of the new category. Thus a Nominal modifier of a Nominal Phrase becomes an Adjective in translation; a Verbal unit acting as a Verbal modifier becomes an Adverbial clause containing the Finite Verbal form. The latter case accords with the SUG principle of the Verb being the Sentence Nucleus [9]. 2.2 Syntactic structures ambiguity

We find it important to differentiate between polysemous and ambiguous syntactic structures.

A polysemous structure implies possible realizations of meanings which are compatible within one language structure and can be transferred to the structures of another language which are isofunctional to the source language structure. An ambiguous syntactic structure presupposes alternative ways of interpretation, the meanings being incompatible within one language structure, thus we deal with ambiguity when we try to discern some Finite and Nonfinite verbal forms: Gerund / Present Participle; Infinitive / Present Simple; Past Participle / Past Simple. Ambiguous structures can be misleading to the parsing

procedures and subsequent machine translation, as for example, the “garden path” is a well-known language phenomenon which may give incorrect parse at the early stage of analysiis, that could be corrected only at the final stage:

e.g. The new control system updated continuously displayed robust performance. A broader context is necessary for correct interpretation of ambiguous syntactic structures.

Cognitive transfer based on functional semantics

For machine transfer the realization of language transformations which must be performed in the course of translation from one language into another, is the most complex problem. The current development stage of machine translation systems is characterized by research in the area of cognitive semantics, probabilistic language models and by development of the semantics- syntactic presentations, which consider polysemy and ambiguity of syntactic structures. The new content to the problem of language transformations are given by the today realia: the need for projecting and developing the training components of the systems of machine transfer and processing of text knowledge on the basis of the already existing and newly created corpora of parallel texts. Translation activity involves the search for equivalence between structures of different languages. However, to establish whether the structures and units are equal or not, we need some general equivalent against which the language phenomena would be matched. In Contrastive Linguistics the notion of tertium comparationis is widely employed to denote this general equivalent, and in [10,11] it was

demonstrated that the approach based on the principle “from the meaning to the form” focusing on Functional Sytax would yield the necessary basis for equivalence search. What differs our approach is the attention to the semantics of configurations, i.e. the study of the way languages tend to arrange structures in order to convey certain meanings. And we focus on the linear patterns of the languages under study, since we assume that linearization is not a random process but it is determined by the cognitive mechanisms of speech production and the way they manifest themselves in syntactic potentials of a given language. The primary object of our contrastive language study was to establish what particular language meanings are represented in the categorial-functional systems of the English and Russian languages.

Computer-assisted translation vs machine translation

As computers have become an integral and crucial part of every aspect of our life,

translation without the use of computers is nowadays impossible. Using a computer makes the translation process easier and faster. Hutchins (2003, p.2) lists some reasons why computers are used in translation:

• To lessen the amount of work for TRLs

• Technical materials can be very monotonous to translate for humans

• Computers provide consistency

• Quicker results

• In the case of machine translation, some translations can be completely performed with a

computer and do not require a human TRL

• Cost reduction

Translation performed by computers can be divided into computer-assisted translation and machine translation. In computer-assisted translation, the translation is primarily performed by a human TRL, with the assistance of computer software that facilitates some aspects of the

translation process. This is the main difference between CAT and MT, in which the translation is done primarily by a computer. Bowker & Fisher (2010, p.60) consider translation as something that should be looked at on a "continuum of translation possibilities", with varying degrees of human or computer assistance. CAT is located in the middle of the continuum, between human

translation on the one end and machine translation on the other end.

The terms Computer-assisted Translation and Machine Translation can be confusing for non-specialists, due to being closely related. Both types of translations use a computer during the translation process. However, as Craciunescu (2004) points out, CAT and MT are each the result

of different approaches used in translation; they produce different results and are not used in the same contexts.

Unlike CAT, where the TRL does most of the work with some assistance by a computer,

in MT the translation is done primarily by a computer, with possible human intervention during

the pre- or post-editing process. (Bowker & Fisher 2010).

Optimism for MT was high in the 1950s, with predictions of fully automatic systems within

a few years. However, the optimism faded within a few years, as the importance of the human

factor in translation became clear. In 1966, the National Science Foundation, at the request of sponsors of MT, set up the Automatic Language Processing Advisory Committee (ALPAC) which

reported that compared to human translation, MT was more expensive, slower and less accurate.

Although seen as unreasonably critical, ALPAC's influence brought a virtual end to MT research

in the US for over a decade (Hutchins, 2007). According to Garcia (2015), research into MT

stalled due to lack of computation power but continued when capacities expanded, despite the

aforementioned setbacks.

Today, there is a new type of MT, which uses neural networks to find correspondences between SL and TLs. Neural networks have the advantage of being able to use data to learn complex relationships among natural languages, without having to resort to hard-to-design manual hand features (Srivastava, Shukla & Tiwari, 2018). There have been attempts to augment CAT with automation from MT, starting in the 90s, but the software was not powerful enough then. In 2006, Lingotek launched a web-based CAT integrated with MT. It was followed by Trados and various other CAT tools. The integration of MT into CAT provides CAT tools with the option of either continuing to work the traditional way, accepting, repairing, and rejecting exact, fuzzy and no matches, or using MT solutions to populate no matches, which can then either be accepted, modified, or rejected. However, this technology is still in its infancy, and it is questionable whether it would lead to improvements in time and quality

of translations (Garcia, 2015).

History of CAT Tools

Craciunescu (2004) explains that the main characteristic of computer-assisted translation is the use of a variety of tools that make the work of the TRL quicker and more accurate. These tools are also called CAT tools, although they are known by many names, which is a source of some terminological confusion. Garcia (2015) notes that CAT tools have been referred to by various other names in the industry and literature. Some of the names are CAT software, CAT systems, "TM (translation memory), TM tools / systems / suites, translator workbenches /

workstations, translation support tools, and latterly translation environment tools (TEnTs)" (p. 69). It is worth noting that the term translation memory has another meaning, being one of the main components of a CAT tool,. Also, the term

Translation Environment Tool (TEnT) is often used synonymously with CAT tool, but it should be noted that there is a slight difference, as CAT tools refer to all the tools and software TRLs use when translating, while TEnT refers to specific software that integrates all tools into one (Barois, 2018). To avoid confusion, I will be using the term CAT tools in all instances in this thesis.

According to Bowker & Fisher (2010), the term CAT tools should be applied strictly to software that was specifically designed with translation in mind, to avoid confusion with other tools and software TRLs use when working. This definition excludes tools such as e-mail, spelling checkers, and word processing programs. Following that definition, Microsoft Word, despite being a word processing program and having a spell check feature, is not a CAT tool.

The history of CAT tools is relatively short. Garcia (2015) points out two major periods for CAT tools; the classic era, lasting from 1995 to 2005, and the modern

era, encompassing a period from 2005 till present day. The classic period started when CAT tools were fully developed and became available for commercial use around the mid 1990-s, with the decade that ensued being focused on stability and increasing processing power. The modern era is characterized by the increased number of potential scenarios for CAT usage. According to Garcia, “a more granular approach towards text reuse has emerged” and “the amount of addressable data expanded” (p. 69).

User interfaces have been simplified, and cloud computing has made CAT tools accessible to a broader audience while also making it possible to integrate TMs with MT. Finally, the power of the Web made it possible for translation aficionados around the world to cooperate on projects, significantly reducing translation times.

Although CAT tools were finalized in the mid 1990-s, their development had started much

earlier. Bowker & Fisher (2010) note that the development can be traced to the 1960s, when term banks were created, which allowed storage of large amounts of structured information incomputers. Although structured information could be stored in computers, the technology was not yet advanced enough to be used for the purposes of translation, and human translation was still seen as more efficient. TRLs used typewriters in the 1960s, and kept paper copies of their work,

consulting them when the need arose.

One of the main reasons for the development of CAT tools was the frustration of machine translation developers, who were looking to design a product that could

assist in producing useable, yet faster and cheaper translations. The ALPAC (Automatic Language Processing

Advisory Committee) report of 1966 was critical of machine translation but was supportive of the idea of funding Computational Linguistics, particularly machine-aided human translation, listing reports which showed that using tools such as electronic glossaries can both increase productivity and reduce errors (Garcia, 2015).

Bowker & Fisher (2010) note that the developments in computational linguistics in the late 1970s and early 1980s were crucial for the development of modern CAT tools. These developments made it possible to not only use computers to store, but also search for and retrieve information. Typewriters quickly became a relic of the past with the advent of the personal computer, which was one of the key factors that allowed the emergence of computer-assisted translation. Personal computers made possible the storage of documents in soft copies, while at the same time providing a more convenient way of querying them. In 1984, possibly foreseeing the future importance of computers for translations, two German TRLs, Jochen Hummel and Iko Knyphausen, established Trados (TRAnslation & DOcumentation Software) which was initially

established as a provider of language services (1984-1989) (Garcia, 2005). Trados would later go on to become the most widely used CAT tool.

The first prototype of a CAT tool, known as the Translation Support System (TSS) developed by ALPS (Automated Language Processing Systems) in Utah emerged in

the mid-1980s. TSS contained a multi-word processor and had a terminology management system, allowing access to previously translated segments. However, the technology had not yet been developed enough for ALPS to profit from its software, and TSS was taken off the market by the late 1980s (Garcia, 2005).

By the early 1990s, technology had developed to such a degree to allow for the

commercialization of CAT tools, and technologically proficient TRLs with a penchant for business pursued the opportunity, as competition suddenly became fierce. In 1990, T

rados launched their erminology database named MultiTerm. In 1992, the first version of the Translator's Workbench TM tool was launched. The same year, other programs were launched, such as IBM Deutschland's

Translation Manager 2, and STAR AG's Transit. Over the next few years, many other CAT tools were launched, many of which were shortly discontinued, with the notable exception of Déjà vu, which was launched in 1993 and remains in use until today. Out of all these programs, Trados was

the one that became the industry standard, in large part due to successful tender bids to the European Commission in 1996 and 1997. As previously stated, features that were present by the mid-1990s were standardized on the most advanced tools and would stay relatively the same over the course of the ensuing decade (Garcia, 2015).

In 2005, the modern era of CAT tools began. That same year, SDL (Software and

Documentation Localization), a multinational services company with headquarters in the United Kingdom, acquired Trados. In 2009, SDL Trados Studios 2009 was released, which saw a shift toward integrating all functions into one proprietary interface. Lingotek, the first web-based tool, was launched in 2006. In 2009, Google launched the web-based Translator Toolkit, aimed at non-

professional users. CAT tools of the modern era have recognized the importance of STs and the supply side of translation, and "begun creating authoring tools for precisely the same gains of consistency and reuse" (Garcia, 2015, p.79). According to Garcia, while the classic era of CAT Tools was defined by the amount of computer processing power and connectivity, the modern era has largely been defined by cloud computing and Web 2.0. Cloud computing made local storage

and processing largely irrelevant, as all data is now stored remotely, online. Web. can be defined as a more interactive and collaborative kind of Web, with an emphasis on "social interaction and collective intelligence", and social media websites such as Youtube, Myspace

(which was superseded by Facebook) and others (Murugesan, 2007, p.34). For CAT tools, this

means that users now take on a more active role, and user experience and feedback have gained more importance (Garcia, 2015).

As can be seen from this overview, CAT tools have developed quickly over the last few decades, alongside the rapid evolution of technology. Esselink 2000; Lagoudaki 2006 (as cited in Bowker & Fisher, 2010) conclude that CAT tools have become increasingly more accessible, popular, and affordable, and have become

a necessity for TRLs in today's globalized information age, in order to translate large amounts of text more quickly. Along with that, the process of computer-assisted translation has become increasingly more streamlined, as modern CAT tools have integrated various components such as termbases and TMs and have made their interfaces more user-friendly.

Main components of CAT tools

According to Craciunescu (2004), the two most important components of CAT tools are terminology databases and translation memories (see section 2.2), and they provide the foundation of all CAT tools.

A terminology database (also called a glossary, termbase, or lexicon) can be defined as a standardized dictionary, in which TRLs can check meanings, translations and spellings of words. (Soluling, accessed on August 22, 2021). In other words, it is a centralized database containing various technical or industry specific words. It combines images, reference materials and terms that are pre-approved, and allows TRLs to share them in real-time, which makes the translation process faster (especially if it is also automated), as well as more consistent, as the same term can always be translated in the same way. The primary benefit of using a terminology database it that

it saves time translating, especially with texts that contain large amounts of specialized content, as the TRL does not have to spend a lot of time researching specialized terminology. Instead, they

can quickly find correct translations of key terms. Other benefits include the elimination of ambiguity within an abundance of specialized content and help with the management of key terminology (Language Scientific, 2015).

Terminology bases are usually integrated into CAT tools. However, there are also

termbases that can be accessed online without using a CAT tool, such as the Croatian Struna (struna.ihjj.hr) or IATE (Interactive Terminology for Europe), as illustrated below.

Although the terms termbase and glossary are often used interchangeably, it should be noted that there is a slight difference in meaning, as the term changes depending on the format.

For example, a termbase is usually in a format that can be incorporated into translation software automatically. A glossary is usually delivered in a format like Excel or Word. The advantage of using these formats is that they can easily be converted into termbase formats, so it can be used as a standard format in cases where TRLs do not know which software they will use beforehand (Proz, 2013). In the rest of this thesis, I will be using the term termbase to refer to both glossaries and termbases, unless it is necessary to stress this difference.

Unlike termbases, a TM is used to monitor the process of translation in real-time,

memorizing each translated passage. When the TRL encounters a paragraph or a sentence that has already been translated, the TM tool notifies them and allows them to either insert the previously translated text or modify it. TMs are created when a translation or a localization project is started and evolves over the course of it, as new content is translated and therefore stored into memory, meaning that the memory expands with each subsequent translation. An expanded TM is a valuable resource for future translations. A major benefit of using a TM is consistency: the same phrases can always be translated in the same way; therefore, multiple TRLs can work on a project and use consistent phrasing and terminology. TMs also reduce the cost of translation, time spent translating, and lower the turnaround time of the translation and localization processes (Language Scientific, 2015).

Bowker & Fisher, (2010) offer a more technical clarification of how TMs work. They explain that TMs store both ST and TTs in its database as bitexts (bitexts are a collection of aligned texts, source and target, that are considered equivalents of each other). The texts are divided into segments, which are usually sentences,

although Garcia (2015) notes that they can also be “a title, caption, or the content of a table cell” (p. 71). Each segment from the ST is then linked to the corresponding segment in the translated TT. When translating a new text, the TM again divides the new text into segments and compares them with the segments stored in its database. Pattern-matching is used to identify whether any portion of the text has already been translated as part of a text from the database. When a match is found, the TM presents it to the TRL, and they can decide whether to accept, modify, or reject the displayed match. It should be noted that depending

on how similar the two segments are, a different type of match is displayed. Garcia (2015, p.72) describes three main types of matches:

• Exact match (or 100 per cent match): a source segment from the database precisely matches the active segment that needs to be translated. The TRL still needs to check whether the translation can be reused, or if some minor modifications are necessary.

• Fuzzy match: a source segment matches the active one to a degree. The degree is displayed as a percentage and is calculated on the Levenshtein distance, looking at how many insertions, substitutions or deletions are required to get an exact match. To avoid distractions, only segments with a matching percentage of 70 and above are usually offered, although Bowker & Fisher (2010) note that thresholds usually range from 60% to 70%.

Depending on the segment offered, the TRL can decide whether to use the proposed segment or start from scratch.

• No match: the TM has failed to find source segments that exceed the match threshold (usually 70%) and no match is offered.

Along with these three main types of matches, some authors add other types. For example, Bowker & Fisher, (2010) for instance, make a more detailed division and describe six types of matches. Along with exact, fuzzy and no matches, they also define full match, sub-segment match, and term match.

There are many other components that make up a standard CAT tool, other than the three mentioned. Some of the more common ones are: concordancer, document analysis module, machine translation system, project management module, quality control module, term extractor etc. (Bowker & Fisher, 2010). A brief description of each is provided below.

It should be noted that editors, TM, termbases and the other tools are intertwined. For

example, CAT tools usually use a Terminology Management System (TMS), to store and retrieve terminology information from the termbase. Additionally, TMs and termbases can be integrated to make the translation process more automated (Bowker & Fisher, 2010).

Current trends and future developments

ProZ, the most widely used website dedicated to translation, conducted a survey in 2013 to assess TRLs' attitudes towards CAT tools. Among other things, TRLs

were asked which CAT tools they used to determine the market share. According to the results, SDL Trados had the biggest share, with over 80% of TRLs using it. It was followed by Wordfast and memQ, a CAT tool developed in Hungary, which has been gaining popularity over the last decade.

The translation market is steadily growing, and CAT is growing with it. According to Kučiš (2010), in the EU translation market only, in 1997, there were 1,125,709 pages translated. In 2007, the number reached 1,762,773, and the number has probably grown even more since then, with more and more freelance TRLs being hired to translate, as EU's TRLs find themselves overwhelmed with work. Such a situation suggests CAT is here to stay.

Bowker & Fisher (2010) see the future of CAT as one of emerging possibilities, with CAT continuing to rapidly develop along with technology. Some of the possibilities are improvements in TMs, with the addition of linguistic analysis and the ability to consider the context of matching segments, a feature that is already present in some TMs. Current TMs identify the differences in fuzzy matches, but in the future, they might also be able to indicate which elements of the target segment should be modified or preserved. Along with that, the standardization of products for sharing translations will make sharing translations easier. Finally, the Internet opens up various possibilities for CAT, such as allowing for the crowdsourcing of translations and collaborative translations, connecting TRLs from all over the world, and allowing them to share their knowledge, which will undoubtedly lead to further improvement of CAT tools.

Ranking CAT TOOLS

When it comes to ranking which features are the most important when it comes to CAT tools, most (29.6%) of the participants chose speed, which is an understandable answer, given that TRLs are expected to translate texts quickly, and this is sometimes even more important than giving completely accurate translations. The next highest option was fullness of features (23.5%), followed by ease of use (21%), and compatibility (16%). These four options seem to be the most important when discussing CAT features. The other options picked were customization (4.9%), tech support (1.2%) price (1.2%) and two features reported by the participants, TMs and quality assurance & control, both at 1.2%. Of note is that the price of a CAT tool does not seem to be very important to professional TRLs, provided the other features are present. For the sake of comparison, as listed on their site, the newest full version of SDL Trados, Trados Studio 2021, is

priced at 385€ (down from 695€), while the price of the version that supports two PCs simultaneously is 475€ (down from 855€). The price of an annual subscription (One-year license to use Trados) is 295€. The price of Memsource, as listed on their site, starts from $27 a month for

the most basic edition (containing the essential CAT tool features, such as TMs, termbase, integration with machine translation etc.), all the way to $350 a month for the most advanced edition.

When the participants that use CAT tools were asked to rate their computer usage skills on a scale from 1 to 5, most (50.6%) rated them as 5. Their responses averaged out to 4.43. In contrast, the participants that do not use CAT tools rated

their computer usage skills slightly lower, with most (53.8%) rating them as 4, and the mean score being 4.31. Although this is not statistically

Significant (partly because only 13 participants that do not use CAT tools responded to the questionnaire), it might imply that users of CAT-tools perceive themselves as slightly more skilled in the use of computers compared to participants that do not use CAT tools.

When the participants that do not use CAT tools were asked which CAT tools they have heard of, most (76.9%) reported hearing of SDL Trados, while other CAT tools were not as widely known to them, with memoQ (46.2%), Wordfast Pro (46.2%) and Memsource (38.5%) being among the slightly more well-known ones. 23.08% (N=3) reported not having heard of any CAT tools.

Most of the comments were positive in their

appraisal of CAT tools, describing them as extremely helpful when translating. One participant stated that CAT tools keep improving, while also recommending the use of modern machine translation, based on neural MT technology, for

Trados. Another participant made a similar comment, stating that what makes CAT tools so effective is their integration with machine translation. Yet another participant suggested a study that examines how much TRLs use the machine translation options offered by many CAT tools (including Trados); this is something that could be examined in future studies. As mentioned previously during the study, one participant commented that CAT tools are a requirement for TRLs who work with foreign clients, while another participant lamented the fact that higher-education institutions in Croatia do not teach their students how to use CAT tools, as many clients and translation agencies require the TRL to use

CAT tools. Several participants noted that one of the main strengths of CAT tools is not only the option to store terms once translated and save them for future use, but also their ability to help users to be consistent in terminology use across one or multiple translations. One noted that CAT tools not only create personal terminology bases for each user but are also capable of storing separate terminologies for each client, which enables users to have consistent translations for each client. Another participant stressed the importance of CAT tools for the creation, management and utilization of terminology. However, even though terminology bases and TMs make things faster

and more efficient, one participant pointed out that creating a solid database is a time-consuming process, meaning that it will take a TRL many translations and many hours of work before they notice a significant increase in efficiency in their translation process.

Translation today is much more reliant on computers, as computer technology has rapidly grown in the last few decades. Computers can significantly facilitate the translation process. Translation using computers can be divided into computer-assisted translation and machine translation. Both translation types serve different purposes. The potential today lies in integrating these two types of computer translation.

CAT tools, i.e. software that is used for computer-assisted translation, traces its roots to the 1960s, although CAT tools in their finalized form first appeared in the mid-1990s. The history of CAT tools can be divided into two eras, each with its own characteristics. They are the classic era, from 1995 to 2005, and the modern era, from 2005 to the present. The classic era was focused on stability and computing power, while modern CAT tools have shifted from those aspects and focus on using the Internet and cloud computing to make themselves accessible to a broader audience, as well as to provide the possibilities of integrating CAT and MT.

Many TRLs today (86.2% of participants surveyed in this study) use CAT tools, not only because they make the translation process more efficient, but also because many clients require TRLs to use CAT tools. In Croatia, this is especially common with foreign clients. The study conducted in this thesis has shown that the

majority of freelance and in-house TRLs in Croatia use CAT tools. It has also confirmed that the majority of those TRLs have largely positive opinions on CAT tools and consider them to be helpful when translating. The most used CAT tool among TRLs in Croatia is SDL Trados, although Memsource also has a significant share of the market and has increased it quickly, which in part may be due to its focus on cloud computing.

When it comes to how much CAT tools increase the efficiency of a translation, it is understandable that the opinions of TRLs are varied. Most of this comes down to just how extensive a user's database (consisting of their termbases and TMs) is. For a new user, who is starting out with a clean database, translations using CAT tools will not be significantly faster compared to not using a CAT tool, although CAT tools, such as SDL Trados, provide an interface, such as using a split screen for STs and TTs, which facilitates the translation process. One of the limitations of this study was that it was not able to objectively measure exactly how much a TRL's efficiency is increased, and instead relied on the subjective reports of the participants.

Several participants mentioned that what makes CAT tools especially efficient is their integration with MT, mentioning they would like to know just how many TRLs take advantage of the MT option offered in many CAT tools. This is something that should be examined in a future study.

As this study has confirmed, the main drawback, if it could be called that, for CAT tools, is that in order to achieve their fullest potential, a user will need to have many translations already performed and stored into a TM, meaning that their full potential cannot be realized from the start. Nevertheless, CAT tools are very

helpful for TRLs and have become a necessity nowadays, and their full potential has still not been reached. As technology continues to develop, so will CAT tools, and they should become even better in the coming years.

References

Barois, M. (2018). What are Translation Environment Tools (TEnTs)? Accessed July 26, 2021, from https://mastertcloc.unistra.fr/2018/01/22/translation-environment-tools-computer-aided- translation-software/

Bassnett, S. (2013). Translation. Routledge.

Bowker, L., & Fisher, D. (2010). Computer-aided translation. Handbook of translation studies, 1, 60-65.

Craciunescu, O., Gerding-Salas, C., & Stringer-O'Keeffe, S. (2004). Machine translation and computer-assisted translation: a new way of translating. Translation journal, 8(3), 1-16.

Esselink, B. (2000). A practical guide to localization (Vol. 4). John Benjamins Publishing.

Garcia, I. (2005). Long term memories: Trados and TM turn 20. The Journal of Specialized Translation, 4, 18-31.

Garcia, I. (2015). Computer-aided translation: systems. Routledge Encyclopedia of Translation Technology, 68-87.

Hutchins, J. (2003). Machine translation and computer-based translation aids.

Hutchins, J. (2007). Machine translation: A concise history. Computer aided translation: Theory and practice, 13(29-70), 11.

IATE, European Union Terminology (2021). Accessed 19 July 2021, from https://iate.europa.eu/search/result/1627225521071/1

Kučiš, V. (2010). Prevodilački alati u funkciji kvalitete prijevoda. Informatologia, 43(1), 19-33.

Lagoudaki, E. (2006, November). Translation memories survey 2006: Users' perceptions around TM use. In proceedings of the ASLIB International Conference Translating & the Computer (Vol. 28, No. 1, pp. 1-29).

Language Scientific (2015). What's The Difference Between Translation Memory (TM) and A Glossary? Accessed 18 July 2021, from: https://www.languagescientific.com/whats-the-difference-between-translation-memory-tm-and- a-glossary/

Merriam-Webster. (2021). Translation. In Merriam-Webster.com dictionary. Accessed July 19, 2021, from https://www.merriam-webster.com/dictionary/translation

Murugesan, S. (2007). Understanding Web 2.0. IT professional, 9(4), 34-41.

Ministry of Science and Education. Accessed September 15, 2021, from mzo.gov.hr

Proz (2013). Glossary vs termbase. Accessed 18 July 2021, from: https://www.proz.com/forum/localization/244198-glossary_vs_termbase.html

TRANSLATION TECHNOLOGIES With the growing amount of translation, technology is becoming more and more important in the translation profession. Not only it helps the translator to be efficient, but it is almost vital for the translator to use technology if he or she wants to remain competitive in the 21st century's global market.

Types of Translation Technologies These terms are very closely linked together and sometimes their distinction might be un- clear. For instance, according to Quah (2006, 6), the term Computer-Aided/Assisted Trans- lation (CAT) is frequently used in the professional translating community and in Translation Studies, whilst people developing software for this type of tools often use the term Machine- Aided Translation (MAT).

1 Computer-Aided Translation (CAT)

According to Bowker, Computer-Aided Translation is a process in which human translators use a computer software to assist them in completing their translations (Bowker 2002, 144). In comparison to Bowker's statement, Austermühl refers to CAT as translation whose pro- cess involves any assistance of a computer in general (Austermühl 2014, 178).

.2 Machine Translation (MT) Initially, the term was used to describe only automatic systems without any human interven- tion. Due to the fact that researchers and scholars cannot agree on the exact definition of Machine Translation regarding to the human involvement, this term still persists to

describe both automated systems and systems involving human intervention (Quah 2006, 9).

According to Bowker (2002, 149), this term describes the process in which computer is the primary former of the text's translation and human is just an assistant for pre- or post- editing.

3 Machine-Aided Translation (MAT) As it was previously stated, Machine-Aided Translation is a term used in a community of people who develop this kind of computer tools (Quah 2006, 6). Otherwise, it has the same definition as Computer-Aided Translation that was described above.

Since this is a thesis directed towards the translation field, the term Computer-Aided Translation will be used for both CAT and MAT.

4 Human-Aided Machine Translation (HAMT) According to Lehrberger and Bourbeau (1988, 7), this translation is completed by a computer with the necessary help of a human translator. Assistance of the human translator can be required before the text is forwarded to machine translation, throughout the whole process, or during the post-editing phase when the final output is revised.

The human intervention is necessitated mainly because particular linguistic structures are difficult to automatically analyse, and also because of a semantic aspect (Lehrberger and Bourbeau 1988, 7).

5 Machine-Aided Human Translation (MAHT) Machine-Aided Human Translation is defined as a translation where computer software completes a particular part in the process of the whole translation and the human translator may use a variety

of tools or their combination to complete it. Therefore, this type of translation is primarily put together by the translator (Quah 2006, 13)

6 Fully Automatic Machine Translation (FAMT) The Fully Automatic Machine Translation is described as translation in which no human assistance is used between entering the text into the machine for processing and the definitive crude output of the machine translation. Nonetheless, this machine output still requires some revision, similarly as the human translation would, although the machine itself decides which parts of the text should be checked (Lehrberger and Bourbeau 1988, 8).

2 History of Translation Technologies The transformation of translation from craftsmanship to a mechanical process is inextricably linked to and strongly influenced by technological advancements. Over the last 40 years, the translator's work has changed from translating paper documents without using any tools to translating electronic documents inserted in a variety of software, tools, and applications.

All the aforementioned technologies have a large range of different formats and serve diverse purposes (Pastor et al. 2018, 205).

The concept of using a machine to translate dates back to the 1950s, the age of the Cold War, when U.S. scientists tried to use the help of technology to translate Russian doc- uments (House 2018, 18). According to Kay (1973, 217), for the first time in history, a con- siderable amount of money was invested in projects that could be eventually inauspicious. The U. S. government's departments invested

money to obtain a computer program or ma- chine that would be able to produce fully automated translations.

Later in the 1960s, the Automatic Language Processing Advisory Committee (ALPAC), a board of seven scientists, was established to evaluate the national sponsorship of research on machine translation. After some research, the committee produced a highly critical report that claimed machine translation was inefficient, expensive, and unsatisfac- tory. They also recommended to discontinue financing the research in this field. This resulted in suspension of development of machine translation in the U. S. for more than a decade (Kay 1973, 217).

In the early 2000s, when researchers moved from a ruled-based approach to a statis- tic-based approach, the quality of machine translation improved significantly. Due to statis- tical probabilities and new computer advancements, which made computers capable of pro- cessing and storing massive volumes of data, translation technologies could use these data for continuous improvement. As an example of system, which is supported by the statistic- based approach, is globally well-known Google Translate (House 2018, 18-9).

Despite the fact that many advances are being made in the field of translation technologies almost every day, a fully automated high quality computerized translation is still not available.

.3 Approaches to Machine Translation In spite of the fact that machine translation is unfavourably famous for its errors, significant advances continue to be made and it has improved considerably in the past few decades. The error reduction is due to numerous factors and their combination. One of the factors being the possibility of building larger databases of grammatical rules and lexicons. Another one is the new linguistic knowledge that has been acquired (Bowker 2002, 3).

With a better understanding of linguistic theory, it is easier to interpret related rules and create controlled languages (CLs) to eliminate ambiguity. A crucial factor is improvements and utilisation of the computers themselves at their full potential. This has resulted into new approaches to machine translation, such as statistical-based and example-based ap- proaches (Bowker 2002, 3).

Various methodologies have been adopted to achieve the automated translation. Ac- cording to Tripathi and Sarkhel (2010, 388-9), the translation process is divided into two levels – metaphrase and paraphrase. They refer to metaphrase as the word-for-word translation, which produces the result of 'literal' translation for every word in source text, and therefore, this level may cause loss of meaning of the original text. Paraphrase, on the contrary, would carry the meaning of the original text without necessarily producing the word-for- word translation (9). The typical metaphrase level would be seen, for instance, in the direct translation approach that is heavily based on word-for-word translation (Quah 2006, 177). These two levels are the cores of approaches to machine translation and every translated text in general.

Sinwai (2017) claims that from the beginnings of machine translation research to the year 2012, a total of 22 approaches have been formed (217). Chosen methodologies are further described based on their distinct importance.

Rule-Based Machine Translation Methodology

The rule-based machine translation methodology is focused on transferring structures of source language to structures of target language (Tripathi and Sarkhel 2010, 389) and its development is dated from the 1960s to the late 1980s (Somers 2003, 177). Sin-wai (2017, 219-20) describes this machine translation as "a relatively traditional machine translation method which depends on the preparation and maintenance of a large number of rules and lexical information in the form of dictionaries, both general and specialized". This method- ology involves various approaches or sub-types. Two of the most important are interlingua and transfer translations (Tripathi and Sarkhel 2010, 389). Quah (2006) mentions that the highest usage of linguistic theory is implemented in the rule-based methodology due to the utilisation of formal grammar (177).

Interlingua translation transforms the source language into an independent interme- diary language, from which, subsequently, several target languages can be potentially de- rived (Tripathi and Sarkhel 2010, Sin-wai 2017). The term interlingua is presented as "allsentences that mean the "same" thing in the same way, regardless of thelanguage they hap- pen to be in" (Jurafsky and Martin 2000 cit. from Quah 2006)

Transfer translation converts the source language text into an abstract representation, while the characteristics of the source language stay unchanged. Afterwards, a correspond- ing representation is produced for the target language, which now bears the characteristics of the target language. This representation is created by using grammatical rules and diction- aries (Tripathi and Sarkhel 2010, Sin-wai 2017).

Quah (2006, 73-4) claims that this approach is not as ambitious as the interlingua translation approach. As this translation depends on dictionaries, it may cause difficulties in resolving language ambiguities. On the contrary to the interlingua translation, this approach employs multiple transfer models for every translation unit.

.2 Corpus-Based Machine Translation Methodology

This approach uses a vast pre-existing database of already translated texts to generate a whole machine translation system (Sin-wai 2017, 218). Due to its exceptional level of accu- racy is the corpus-based methodology one of the most popular (Tripathi and Sarkhel 2010, 390), and its development is dated to the 1990s. Therefore, it is more modern in the field of machine translation (Somers 2003, 163).

Quah (2006, 177) introduces two of the main approaches/sub-types – prominent ex- ample-based translation and statistical-based translation. Sin-wai (2017, 218) presents an additional sub-type of corpus-based approach – a memory-based machine translation, alt- hough both Tripathi and Sarkhel (2010, 391) and Quah (2006, 81) include this translation in the example-based translation approach.

The example-based translation uses an immense amount of corpora to analyse equiv- alent, previously translated examples of language pairs to, subsequently, generate similar text in the target language (Sin-wai 2017, 218). Quah (2006, 81-3) refers to these examples as "a matched pair of segments", which should preferably be at the length of a sentence. If no similar matches are identified in the bilingual corpus or if the STs are metaphorical, the example-based approach will probably fail to be effective.

Lastly, the statistical-based translation is based on comparing source text segments to a bilingual corpus. Statistical methods are then applied to acquire new segments of target text. This approach is also not faultless as it needs an extensive bilingual corpus for the translation to be successful (Quah 2006, 77-80).

According to Quah (2006, 91), these two approaches are the most practical for professional translators, as their key advantage is the possibility of using already translated content

COMPUTER-AIDED TRANSLATION TOOLS

As computer-aided translation tools are becoming more and more popular in the professional translators' community, it is crucial for the modern-day translators to be acquainted with different types of computer tools (Bowker 2002, 6).

As it was mentioned in the second chapter, CAT technology, in its largest context, refers to any kind of tool that is computerized and translators use its assistance in their work. Tools such as e-mail, grammar checker, word processors, and the

World Wide Web are great examples. These tools are used not only by translators, but also by people from other professions. Furthermore, these tools are also used outside the professional world for personal correspondence and research as this CAT technology is widely available (Bowker 2002, 6).

Despite the fact that professional translation is mainly technical, there are many innovations, developments, and overall changes in various fields as information technology, business, medicine and others, that it is beyond the bounds of possibility for a technical translator to be familiar with all these fields. Different types of computer tools are suitable for different specialised fields. Each translator has to determine which tool will be the most appropriate for his/her specific translation purposes (Austermühl 2014, 102).

Translation Memory Systems (TMS)

Translation memory represents a collection of linguistic data, which saves the source text and its translated equivalent. The text which is being translated is divided into segments that are frequently in the form of sentences (see Table 1). Translation memories' most crucial feature is that it enables translators to use already translated units again. "Reusing a previous translation in a new text is sometimes referred to as "leveraging." Although language is dy- namic, it is quite repetitive, and people often use the same or similar expressions when com- municating similar ideas" (Bowker 2002, 92-3).

Free and Open-source Tools

Open-source system or software is one that is freely accessible, and it enables its users to analyse it, make changes and improvements of their translation process. System is collec- tively evolved by peers, meaning all public users.

Lately, these systems are rapidly gaining in popularity not only in translation indus- try, but also generally in diverse companies, which are not oriented in translation field. Busi- nesses are supposedly using these systems because they are flexible and have valuable fea- tures, which is only adding to the fact, that they are for free. (Bowker et al. 2008, 27).

1 OmegaT OmegaT is a good example of a free, standalone, open-source Computer-Aided Translation tool originally created by Keith Godfrey but as an open-source project, it was developed by many volunteers that are under the leadership of Aaron Madlon-Kay (OmegaT, 2021).

It is one of the most used computer-aided translation tools intended for professionals but used also by academicians. It includes traditional features for example fuzzy matching, match propagation, and simultaneous processing of multiple-file projects, can be combined with another software and since it is based on Java it can be used basically on any platform such as MS Windows, Linux, Mac OS (Sin-wai 2017, 194). Thanks to a survey that was conducted in 2010 among 458 professional translators, it was found that OmegaT was used by 6 % of

CAT tools users when compared to the market leader SDL Trados, which was used by 50 % of all users questioned in the survey (Translation Tribulations, 2021).

.3 Paid Tools

Nowadays, many commercial translation memory systems are available on the market. Ac- cording to Hutchins (2005, 13), “The oldest and best established are the systems from Tra- dos, STAR (Transit), and Atril (Déjà Vu). Others include the systems from SDL, Multilizer, Terminotix (LogiTerm), Champollion (WordFast), MultiCorpora (MultiTrans), MetaTexis, and so forth”. These systems have very similar features, and therefore with an increasing demand, the competition is very high, but it can be easily said that SDL-Trados, which was created by merging two already leading softwares – SDLX and Trados – is a leader in the industry (Hutchins 2005, 13).

These paid systems can be divided into standalone systems and network-based sys- tems. Standalone systems usually do not require internet connection and are suitable for working offline, have fast response and are compatible with popular operating systems which makes them great for individual users. Network-based systems require internet con- nection since they are operated in a network environment. This means that they can be ac- cessed from anywhere in the world with access to internet connection. Moreover, network- based systems support team collaboration and management as the translation memories and terminology can be shared within whole team (Sin-wai 2017, 169-70).

MemoQ

MemoQ was developed by the translation technology provider Kilgray Translation Technol- ogies, which is based in Hungary. The company's name was generated from the names of the founders – Kis Balázs (KI). Lengyel István (L), and Ugray Gábor (GRAY). This tech- nology was firstly introduced in 2009, after four years of its development (Sin-wai, 2017, 187). MemoQ is further described in the practical part of this thesis.

SDL Trados Studio

This computer-aided translation tool joined the translation industry in 1984 and became broadly popular since then. It was founded by Iko Knyphausen and Jochen Hummel in Ger- many. SDL Trados has more than 270,000 customers worldwide. Among those customers are both self-employed translators and huge translating companies. The name of this trans- lation memory software is created of initial letters from the words TRAnslation, DOcumen- tation, Software (RWS, 2022)

SDL Trados Studio 2021

SDL Trados Studio is the first computer-aided translation tool that is analysed. According to the tool's webpage, SDL Trados is the leader of CAT tools available on the market. With its long history of development, it is expected for SDL Trados to satisfy all the needs of its users (RWS, 2022).

SDL Trados Studio is a paid software, which offers a 30-day trial of its latest version. The trial version is restricted only by not having access to the cloud functions of Trados Live, and the access to products that are not native for the software is also not included. To continue using the tool after the trial period has ended, the user has to purchase full li- cense. There are three products available for selection – Trados Studio Starter, Trados Studio Freelance and Trados Studio Professional. The price for perpetual licence for this tool ranges from 519 EUR for Freelance (currently at discount from 695 EUR) to 2.495 EUR for Pro- fessional licence (Trados, 2022).

BIBLIOGRAPHY Published Sources:

Austermühl, Frank. 2014. Electronic tools for translators. New York: Routledge.

Bahaa-eddin, A. Hassan. 2011. Literary Translation: Aspects of Pragmatic Meaning. New- castle upon Tyne: Cambridge Scholars Publishing.

Bassnett, Susan. 2014. Translation. London: Routledge.

Bourbeau, Laurent, and John Lehrberger. 1988. Machine Translation: Linguistic character- istics of MT systems and general methodology of evaluation. Vol. 15. John Benjamins Publishing Company.

Bowker, Lynne. 2002. Computer-Aided Translation Technology: A Practical Introduction. Ottawa: University of Ottawa Press.

Chan, Sin-wai. 2017. The Future of Translation Technology: Towards a World without Ba- bel. London: Routledge.

House, Juliane. 2018. Translation: The Basics. New York: Routledge.

Kufnerová, Zlata, Milena Poláčková, Jaromír Povejšil, Zdena Skoumalová, a Vlasta Stran- ková. 1994. Překládání a čeština. Praha: H & H

Pastor, Gloria C., and Isabel Durán-Muñoz. 2018. Trends in E-tools and Resources for Translators and Interpreters. Leiden: Brill/Rodopi.

Quah, Chiew K. 2006. Translation and Technology. New York: Palgrave Macmillan.

Somers, Harold L. 2003. Computers and Translation: A Translator's Guide. Amsterdam: John Benjamins Publishing Company.

Venuti, Lawrence. 2018. The Translator's Invisibility: A History of Translation. 3rd ed. Lon- don: Routledge.

Wright, Sue E., and Leland D. Wright, Jr. 1993. Scientific and Technical Translation, Vol. 6. Amsterdam: John Benjamins Publishing Company.

Online Sources: Bowker, Lynne, Cheryl McBride, and Elizabeth Marshman. 2008. “Getting More than You Paid for?: Considerations in Integrating Free and Low-cost Technologies into Trans- lator Training Programs.” Researchgate, January 2008. Accessed May 3, 2021.