MERIT RESEARCH JOURNALS
www.meritresearchjournals.org

Merit Research Journal of Education and Review (ISSN: 2350-2282) Vol. 2(9) pp. 173-177, September, 2014
Available online http://www.meritresearchjournals.org/er/index.htm

Review

Achieving Translation Quality through Translation Memory

Siddig Ahmed Ali

College of Arts and Science at Arras, Qassim University, Saudi Arabia

Email: sidigahmed45@gmail.com

Abstract

This article aims to exhibit the significance of translation memory for professional human translators. It also encourages them to build their comprehensive TM programs. Moreover, human translator can use such memory by utilizing many software programs to conduct their jobs effectively and effortlessly. Translation memory is one of these programs, which aided human translator to save time; develop consistency and high quality of translation. Furthermore, the study enlightens the students of translation and novice translator the translators with ways of establishing translation memory. It illustrates the nature of TM by providing historical background, building TM and types of software. It reveals TM pros, cons and challenges faced by human translators with the nature of its systems. The research followed the descriptive method due to its appropriacy for the study. Finally, the study recommends the use of the TM especially for translation technical documents and volume documents as it enhance the quality of translation.

Keywords: Machine translation, Translation memory, fuzzy matches, Tados, competence, translation quality

INTRODUCTION

In the era of globalization, there is global expansion in the field of technology. The greatest way to access information and knowledge is through using technology to save time and money and reducing the manual access. For instance, when human has accumulated documents to translate, they resort to machine translation to save time and reduce the cost of human translation expenses. Computer-Assisted Translation (CAT) is one type of translation involving automation. The process using technology for translated is termed as Machine translation or (MT). Arnold, Balkan, Humphreys, Meijer, and Sadler (1994:2) defined MT as the attempt to automate all, or part of the process of translating from one human language to another. Translation of natural languages by machine has become a reality in the late twentieth century. In that era, the translation was devoted to technical manuals, scientific documents, commercial prospectuses, administrative memoranda, medical reports. In this study, the researcher attempts to shed lights on translation memory as a significant component of machine translation.

The history of MT translation

Machine Translation (MT) can be traced back from the 1950s and 1960s around the globe. The impact of ALPAC report in the United States (Hutchins, 2003) is the evidence for the emergence of Machine Translation. However, Machine Translation on the web started with SYSTRAN which Offer free translation for short texts (1996), followed by AltaVista Babelfish, which racked up 500,000 requests a day (1997). Franz-Josef Och (the future head of Translation Development AT Google) won DARPA's speed MT competition (2003). Availability of greater computing power has made access to and usage of MT more straightforward. Machine translation has also gained wider exposure to the public through several

dedicated services, typically available through search engine services. Most internet users will be familiar with at least one of Babel Fish, Google Language Tools, Babylon, or Windows Live Translator. Most of these services used to be powered by the rule-based system developed by Systran, Cancedda, Dymetman Foster and Goutte (2009:15).

Translation memory (TM)

Translation memory (TM) is one of the crucial computer-based aids for human translators. It is a particular type of translation supporting tools that maintain a database of a source and target language sentence pairs, and automatically retrieves the translation of those sentences in a new text that occur to the database, Macklovitch and Russell (2000:137. In addition, Wright and Budin (2001:683) defined TM as a database of bilingual sentences or sentence units. These units are recorded in memory from the time when a translator completes translating each segment of text. The Previously translated segments are automatically shown to translators subsequent to their coming across the same or similar segment. It has been estimated that more than 20 percent of technical text is repeated. Somers (2003:31) introduces MT as the idea of consulting a database of previous translations, usually a sentence-by-sentence basis, looking for something similar enough to the current sentence to be translated and can then use the retrieved example as a model. When an exact match is found, it can be simply cut and pasted into the target text. The first development for TM in the 1970s but not generally available until the mid-1990s. Since then the database is stored in segments formats including phrases, sentences, sentence-like units and paragraphs that have previously been translated. TM includes both the source text (ST) and target text (TT) in language pairs which are stored in so-called "translation units". Translation memory manager (TMM) handles these translation units. Translation memory managers (TMM) are Software programs and tools consisting of a database of text segments in a source language and their translations in one or more target languages.

Building TM database

A prerequisite for a TM is of course a database of translation examples. Known to computation linguists as "aligned parallel corpus", there are principally three ways of building a MT database: building it up as you go along, importing it from elsewhere, or creating it from a parallel text.

Building it as you go along

Perhaps the simplest method is to build it up as you go along. Each sentence you translate is added to the database. Somers: (2003: 33) says if you are working on a text that is similar to one you worked on before, you can load onto the database that you created last time and continue to add to it this time. Conversely, if you are working on different projects and want to develop separate databases for each of them, this can be done. Unfortunately, this method of developing the database is painfully slow, and there will take a long time before the translator really feels the benefits of software.

Importing someone else's

In this method, the translator imports the database from elsewhere. With the proliferation of TM products and increasing numbers of translators using them, it makes sense to share their assets. Fortunately, and despite the variety of software products, developers have agreed a common interchange format, which the TM database developed using one product can be "imported" into another. (ibid: 34)

Aligning a parallel text

This is more complex method is to take an existing translation together with the original text and have software build a TM from it automatically. This involves alignment above all else, though as the previous paragraph indicated, once aligned there will be an amount of indexing and other database manipulations that need not to concern us here.

Alignment involves matching up the source text and the translation segment of a segment of a translation pairs. Segments are usually understood to corresponding to sentences or other more or less easily distinguishable text portions, such as titles. If the translation is straightforward, then so are the alignment. However, there are three factors can make alignment more difficult than it at first seems. First is the difficulty of accurately recognizing where sentences begin and end; the second is the fact that depending on the language pair a single sentence in one language may not necessarily correspond a single sentence in the other language. The third factor is that translators may more or less freely to change the relative order for sentences in translation. (ibid: 34).

Translation Memory systems

The idea of creating a TM system was first proposed in 1970s; however, these systems were not commercially

available until 1990s (Somers, 2003:31). Then in short period, they were quickly accepted by users (Ibid, 33). The number and variety of TM systems available to translators have grown remarkably, which has resulted in a high level of competition among software vendors. TM systems are programs that create databases of source text (ST) and target text (TT) segments in such a way that the paired segments can be re-used (Gil and Pym, 2006: 8). TM systems provide various benefits of translators, translation vendors and clients. Firstly, these systems save a great deal of time of translating the texts that include a high degree of repeated terms and rephrases. The translator translates the segment only once and re-uses it whenever he/she encounters repetitions. Especially in the case of digital texts such as websites, software, and online manuals, the translator is sometimes called on to render no more than updates or adaptations (Gil and Pym, 2006: 6).

Matching

Obviously, the most important function for a TM system is its ability to match the sentence to be translated against the database. Where there is an exact match, the system will normally take the corresponding target language phrase and paste it directly into the target text, though the user will always have the option of rejecting it. Where there is not an exact match the system presents one or more close matches. One is the decision of how to deal with cases where no exact matches can be found. Developers generally opt to search for similar matches and to calculate a ranking of identified 'fuzzy' matched++s which are then offered to the translator as a possible base for the translation +of the new segment.

SDL Trados Studio: SDL Trados, Catalyst or WordFast

Trados is a translation memory tool made up of separate but integrated components, including WinAlign, an alignment tool, MultiTerm, a terminology management tool, and translator's workbench (TWB) a tool that allows the user to search the translation memory database and retrieve segments for insertion into a new text. Trados translator's workbench is a translation memory application developed by Trados. It can be used to translate any document that can be opened by Microsoft Word. It creates a statistical overview of numbers of internal repetitions, fuzzy or extracts matches found in translation memory. The translate command in tools menu of translator's workbench enables the user to divide into segments and pre-translate files with translations retrieved from a translation memory database. Translator's workbench is closely integrated with MultiTerm, a terminology management application that interacts with workbenches. Another recent tool has been added to translators' workbenches are the TagEditor. It provides direct support for HTML and other markup languages Esselink (2000: 368).

Limitations and Challenges of TM Systems

Like every technology, TM systems have some limitations and challenges. First, there is a learning curve associated with using translation memory systems. Using full versions of these systems might require a considerable amount of investment. Besides, updates are generally not free of charge. There are free demo versions of TM systems that are available online to get potential customers, that is, translators, to try such systems. However, these generally do not include all of the features that full versions have.

TM systems bring some quality-related problems. First, errors contained in TMs might also be recycled. Mostly, clients do not make corrections on TM, but on final documents. If the contents of TM are not updated in order to reflect corrections made by the client to the final document, the translator will reproduce the same errors when s/he is asked to translate a document, which is highly similar to the version translated by him/her previously. In that case, the client will be irritated because the corrections made previously need to be corrected in the newer version again and this is not the kind of added value that the client is looking for (Bower and Barlow, 2008: 15-16). The use of TM systems has caused the clients to ask translators to reduce translation fees. According to Gil and Pym (2006), this encourages translators to work fast and often uncritically with the previously translated segments, with a corresponding decline in quality. When higher quality work is required, Gil and Pym's suggestion is to put special emphasis on revising the outputs of TM systems. This means more time checking and editing, yet entails mandatory price reductions - the infamous 'Trados discounts' (Garcia, 2009:202). Another limitation of these systems is that most translators who use these systems no longer focus on translating texts, but segments (Garcia, 2009:201). Communicative context of target text is of no concern to them. During a translation process when TM system finds full matches, a translator who works under time pressure tend to accept these matches without reading, checking and/or editing, and just continues with the rest of the translation. Furthermore, in order to maximize the "recyclability" of a text, the translator may choose to structure the sentences in the target text to match those in the source text and avoid using pronouns, conjunctions, transitions and various references (Bowker and Barlow, 2008:15). According to Heyn (1998:135) and Mogensen (2000: 28), the result may be a text that is inherently less coherent or readable, and of a lesser

overall quality (qtd. in ibid). Mossop (2006: 790) calls it “collage translation”.

Pros and cons of TM

As translation work done by TM is much faster, the resulting profitability of the TM system is measured through lower costs, shorter time, terminology consistency and saved layout of the page, and one of the key advantages is that translation units in TM can be organized as pairs among number of languages. Despite of mentioned savings, the hidden costs should be also mentioned, such as costs of software, maintenance, education for work with TM system, building of glossaries and TM revision. The biggest lacks of TMs are non-existence of knowledge of the language to be translated and context insensitivity. Among obstacles in use of TMs, several ones could appear Heuberger (2006):

- It is possible that TM tools do not fit into existing translation or localization process regarding approval of translation changes
- Customization of the TM systems and sufficient training
- Significant investment (purchase of software, importing of past translations into TM database (i.e. alignment process), training, if necessary additional terminology tools, maintenance costs (upgrading with memory, fast network card, hard disk)
- Protection of TM investment by developing proper strategy for maintaining TM database (data on frequency of updated, regular distribution, backup), ownership of the TM (regulated by agreement), confidentiality, support of the TMX format
- Legalities – whether the intellectual property of the TM belongs to freelancer, vendor or end client.

According to Esselink (2000: 367), TM has the following drawbacks:

Most TM tools do not display the format and layout of the document, which makes it difficult to see how translated text will be displayed in the final layout. A proofreading check is therefore required once the translated files have been converted back to their original format.

TMM often poses problems to vendors, especially when working with several teams of translators in different locations.

Since many last minute language or translation changes are implemented in the translated files once they have been reconverted to their original format, changes are not always inserted in TM; this makes it difficult to keep the TM database up-to-date.

Preparing sources files for translation in TM tools and converting files back adds additional, often time-consuming, steps to the translation process.

TM filters have not always been undated to support new versions of the files formats they process. As the result, translatable text is either not recognized, or markup is presented as translatable text.

Translators do not have the opportunity to change the overall structure of the texts, i.e. to change the sentences within a paragraph. Creating filters for customs i.e. File format that are not supported by default, can be a complicated and time-consuming programming task.

Improving translation quality

Human translators are often loaded with large translation volumes for translation; therefore, translation memories can support them in such projects. Not only these volumes, translation quality is required as well. TM allows the translators to store their translation in database and recycle them in new versions of translation by automatic retrieving matched segments usually sentence for reuse. More specifically, TM has been generally associated with to the benefits in terms of quality, consistency, speed, improvement (O'brien, 1998:119). TM increases translators' productivity and enhance translation quality by ensuring terminologies and expressions that are used consistently. Thus, Quality is achieved when TM is applied.

Human translation competence through TM

The term translation competence has been developing in the previous two decades. It is considered paramount in translation as it reflects the translators experience, training on using translation memory tools. (Schäffner 2000: 146) classified the term into convergent connected types. One of these types is transfer competence – the ability to produce TTs that satisfy the demands of the translation task. Šeböková (2010, p. 56-57) focused purely on translation competence. She classified the competence into Tools competence – the ability to use various tools that will help trainees facilitate translation tasks (e.g. word processor to translation memories or CAT tools).

ACKNOWLEDGEMENTS

I wish to express our gratitude to those who inspired me by the ideas and thoughts throughout writing this article. I am obliged to extend my thanks, colleagues, and students at Qassim University, KSA.

CONCLUDING REMARKS

The paper addressed the topic of translation memory as a system for translation quality. It reveals the signify-cance and the way of establishing powerful TM to

develop translation accuracy. Quality is also achieved when TM is adopted especially for technical documents.

REFERENCES

Arnold B, Humphreys, Meijer, Sadler (1994). Machine Translation, an Introductory Guide. Published in the USA.

Cancedda N, Dymetman M, Foster and Goutte C (2009). Learning Machine Translation. The MIT Press Cambridge, Massachusetts London, England.

Esselink B (2000). A Practical Guide to Localization. John Benjamin's Publishing

Esselink B (2000). A Practical Guide to Localization. John Benjamin's Publishing, Jan 1.

Heuberger Andres (2006). What you Need to Know about Translation Memories? Multilingual webmaster.com http://www.multilingual webmaster.com/ library/trmemories.html [12/2/2006].

John H (2003). ALPAC: The (in) famous report. In Sergei Nirenburg, Harold Somers, and YorickWilks, editors, Readings in Machine Translation, pages 131–135. Cambridge, MA, MIT Press,

Macklovitch E, Russell G (2000). Envisioning Machine Translation in the Information Future: 4th Conference of the Association for Machine Translation in the Americas, AMTA 2000, Cuernavaca, Mexico, October 10-14, 2000 Proceedings

Rodrigo E (2008). Translation and localization. John Benjamin's Publishing Co.

Schäffner C, Beverly A (2000). (eds). Developing translation competence. John Benjamin's Library. (38). Amsterdam: John Benjamin's Publishing Company.

Šeböková S (2010). Comparing translation competence. M.A. Brno: Masaryk University

Somers (2003, 33). Computers and Translation: A Translator's Guide. H. L. Somers John Benjamin's publishing house.

Somers H (2003). Computers and Translation: A Translator's Guide. John Benjamins Publishing co.

Wright S, Budin G (2001). The Handbook of Terminology Management. John Benjamin's Publishing.