The Impact of Translation Memory Segmentation on the Translator’s Cognitive activities while Translating Technical Texts from English into Arabic

Amira Mahmoud Muhammad Abdel-maksoud
An instructor at the AUC, Cairo University (CHP) and (CLT)
A PhD Candidate,
English Department, Faculty of Arts
Cairo University

Abstract:
Since the appearance of computer-assisted translation (CAT) tools, Translation Memories (TMs) have been considered a prominent tool that are indispensable in the translation process. Even though Translation Memories (TMs) are known to increase translation productivity, term consistency and quality, many scholars criticize this default feature. They confirm that TM segmentation disrupts the Source Text (ST) linearity and affects the translator’s cognitive activities. Being a novel tool for translators in the Arab world, there is a growing need for an in-depth investigation of this feature. This research explores how the TM segmentation affects the translator's cognitive activities. It is a qualitative research design where participants translate two English user manuals into Arabic, the first one without the software program WordFast Anywhere (WFA) and the second through WFA. All the participants produced two Think Aloud Protocols (TAPs) one in the pre-test and the other in the post test. Three professional translation instructors at the AUC Language Department (AUC LD) analyzed the protocols produced. The evaluation of pre-test protocol, post-test protocol revealed that TM segmentation dominated the translators' cognitive activities. They were working on the segmental level and lost track of the text characterizing features.

Keywords: Translation memory, segmentation, cognitive activity, Think Aloud Protocol, CAT tools, WordFast Anywhere.
تأثر تقسيم ذاكرة الترجمة للنص المترجم إلى وحدات في النشاط الإدراكي للمترجم عند ترجمة النصوص التقنية من الإنجليزية إلى العربية

الباحثة/ أميرة محمود محمد عبد المقصود
مدرس بالجامعة الأمريكية ومحاضر برامج الساعات المعتمدة (قسم لغة إنجليزية – جامعة القاهرة) ومركز اللغات الأجنبية والترجمة التخصصية بجامعة القاهرة، طالبة دكتورة - قسم اللغة الإنجليزية – كلية الآداب – جامعة القاهرة

ملخص البحث باللغة العربية:

منذ بداية ظهور أدوات الترجمة بمساعدة الحاسب الآلي، وبرامج ذاكرة الترجمة تعد أحد الأدوات البارزة التي لا يمكن الاستغناء عنها أثناء القيام بعملية الترجمة. هذا، وعلى الرغم من الشهيرة التي حظيت بها برامج ذاكرة الترجمة استنادًا إلى قدرتها على زيادة الإنتاجية والحفاظ على صحة المصطلحات المستخدمة والإرتباط بالجودة، فإن البعض الباحثين يؤكدون أن هذه الخاصية تكسر السلسلة الخطية للنص المترجم منه وهذا يدوره يؤثر على الأنشطة المعترف بها للترجمة أثناء الترجمة. وإن ذاكرة الترجمة أحد الأدوات التي لا يمكن للمترجم في العصر الحالي الاستعاضة عنها وكونها آداة حديثة الاستخدام في العالم العربي، فإن هناك حاجة متزايدة لتداول هذه الخاصية بمزيد من الدراسة. لذا، فإن هذا البحث يهدف إلى سبر أغوار خاصية تقسيم برامج ذاكرة الترجمة للنص المترجم منه إلى وحدات وأثرا على الأنشطة الإدراكية للترجمة. تستخدم الدراسة الحالية برامج "Wordfast Anywhere " كنموذج لبرامج ذاكرة الترجمة. هذه الدراسة دراسة وصفية يقوم المشاركون فيها بترجمة دلالي استخدام ناقل؛ أدمهما باستخدام الحاسب الآلي فحسب والأخر باستخدام برنامج ذاكرة ترجمة "Wordfast Anywhere " يتولى عدد من مدرسي الترجمة المتخصصين بقسم اللغات بالجامعة الأمريكية تقييم بروتوكولات التفكير بصوت عال عند استخدام الحاسب الآلي في الترجمة عندما يكون النص وحدة واحدة متكاملة، في مقابل تحليل البروتوكولات التي يقومها المشاركين عند الترجمة تحت تأثير برنامج ذاكرة الترجمة وتهديد تحت تأثير خاصية تقسيم النص.
4.1 Introduction

Due to the integral role technology plays in translation industry and the rapid changes witnessed in this field, translation memories have become an indispensable element in the translators' workstation. Regardless of the pivotal role TMs have, Mor (2018) confirms the need for more empirical research studying translators’ interaction with TMs. Therefore, the present study seeks to unravel the impact of using the software program WordFast Anywhere (WFA) to translate an English user manual into Arabic. It specifically studies the effect of the segmentation feature on the translator's cognitive activities. It traces the effect TM segmentation has on the translator's cognitive activities which are responsible for the TT produced. The present study uses two smart phone user English manuals as the ST and considers the Arabic translation the TT. One manual will be translated as a whole unit without being segmented and the other will be uploaded and translated on WFA in the form of segments.

4.2 Statement and Rationale of the Problem

Based on Teixeira & O'Brien (2017) translation memories are an indispensable element in the translators' workstation worldwide. Although they are extensively investigated by many researchers, this exploration is limited to TM’s positive impact on productivity, and consistency of terminology. As for researches on TM’s drawbacks, they are few and target language pairs rather than English and Arabic. they just explore the detrimental impact TMs have on creativity and translators’ status. Furthermore, the literature investigating TM segmentation is little and does not consider the cognitive load TM segmentation imposes on users.
(Christensen & Schjoldar, 2017). Mor (2018) criticize the literature provided on TM segmentation as it is limited to mere description and speculation rather than research-based results. Alotaibi (2014) and Mahfouz (2018) state that TMs, specifically in the Arab world, are given little attention and therefore more empirical studies on the English/Arabic language pair are needed. Consequently, the current research aims to elaborate on the segmentation feature imposed by TMs and its impact on the translator’s cognitive activities when translating an English user manual user guide into Arabic.

4.3 Research Question

In line with the above studies, the present study seeks to answer the following question:

a) What is the difference between smart phone user manuals translated from English into Arabic in terms of the cognitive activities translators go through when the ST is one unit and when the ST is segmented?

4.4 Literature Review

4.5 Definition of Translation Memory (TM)

According to Reinke (2013), TMs are software programs that support the translator but do not actually perform the task of translation and are mainly used in translation of non-literary texts. They break the ST into segments and force the translator to work in sequential order rather than going back and forth (Mellinger, 2018). TMs, according to Bowker (2002), are databases that can automatically save ST segments and their translations to be reused when translating similar texts. In the current research, the term TM means WordFast Anywhere (WFA) which is the online version of the popular WordFast computer-assisted translation (CAT) program.

4.2 Translation Memory Segmentation (TM segmentation)

Linguistically, segmentation according to Merriam Webster dictionary is the process of dividing into segments. In the current thesis, TM segmentation is the process of breaking the ST down
into smaller units which can be sentences, phrases or sentence-like units such as titles, headings and list items. TM segmentation is a default feature in all TM software programs (Christensen & Schjoldager, 2016)

4.3 TMs in relation to Cognitive activities

The association between the process of translation and cognitive activities was established when Risku (2010) described the TT as a reflection of what goes on in the translators' minds during the translation task. This definition highlights the importance of unveiling the cognitive activities that take place within the translator’s mind during the translation process in order to understand and evaluate the translation product. These cognitive activities based on the cognitive translational theory of Situated Translation propounded by Risku (2002) are attributed to the interaction between the translator and his/her working environment which includes all the tools deployed during the translation process. Thus, tracing the interaction between the translator and his/her environment reflects the cognitive activities taking place. Understanding this explains why TMs are considered a vital artefact that changes the translator’s cognitive activities. It also justifies why Christensen and Schjoldager (2011) and Ehrensberger and Messay (2014) stressed the importance of evaluating the TT through tracing the translators’ cognitive activities which emerge when they interact with their ergonomics (i.e. TM).

O’Brien et al., (2017) explained that being engaged with a complex system (i.e. the TM system) causes a cognitive load on the translators. Based on this, O’Brien et al. (2017) and Teixeira and O’Brien (2017) investigated the cognitive friction resulted from TMs. O’Brien et al., (2017) argued that the ordinary flow of the text was interrupted due to the segmented ST and this adds intrinsic load on the translator. Bundgaard and Christensen (2019) demonstrated that researchers used this theory of cognitive load to examine the resistance and cognitive friction the translators encounter.
4.4 Think Aloud Protocol in Translation Studies

According to Hayes and Flower (1980) Think Aloud Protocol (TAP) is a technique in which participants are asked to verbalize aloud everything that occurs to them while performing a specific task. It is the transcription of the verbalization reports which include the participants’ conscious cognitive activities (Leow, 2015). It is a medium used by researchers who seek getting closer observation of the human mind while it is at work (Zahou, 2012). It is, according to Dastyar (2018), the transcription of every thought a translator has and reports during the translation process. Bernardini (2001) and Christensen and Schjoldager (2011) explained that TAP was borrowed from cognitive psychology for the purpose of studying the internal translation processes and the mental activities which cannot be studied directly.

Being an important tool for obtaining information about the hidden cognitive activities, Ericsson and Simon (1993) classified this verbal reporting method into introspective verbalization and retrospective verbalization. They distinguished introspective verbalization from retrospective verbalization by time of implementation. They illustrated that introspective verbalization takes place while the participants are performing the task to ensure that verbalization is not constrained by memory. Conversely, retrospective verbalization, is usually conducted immediately after some form of processing takes place. Bernardini (2001) divided the TAP into concurrent TAP and Post TAP. She valued the concurrent verbalization rather than the Post (TAP) and regarded it a more effective method for reflecting the cognitive activities and retrieving information from the short term memory during the translation process. It gives a faithful account of the cognitive activities occurring without interfering with the mental processes. Finally, Christensen and Schjoldager (2011) classified TAP into online and offline TAPs based on the data collected.

Based on the above, translation instructors, researchers, curriculum developers and students support the use of TAP and voiced its benefits. Translation instructors can use the strategies
observed in the TAP to find out where problems occur in the translation process specifically if the participants are translation students (Kussmaul and Tirkkonen-Condit, 1995). Specifying translation problems, can help translation instructors see matters more clearly. Translation instructors can reach knowledge-based conclusions when evaluating the TT taking into consideration the indispensable data the TAP reveals. Pietrzak (2015) explicated that TAP can help researchers trace the cognitive activities translators divulge through the simultaneous disclosure of the information stored in the short-term memory (STM). It is described by Bernardini (2002) and Zhou and Lee (2012) as an appropriate verbal-reporting method for revealing the cognitive activities during translation. On pedagogical level, Pietrzak (2015) advocated using the hidden activities the TAP discloses to produce a more appropriate translation curriculum. Leow (2015) elucidated that TAP had the advantage of providing insights into learners' cognitive processes as opposed to simple processing. Araghian et al., (2018) enlisted TAP among the most useful translation assessment tools on basis of the potential pedagogical usefulness it provides those who work in the translation assessment field. Finally, Byrne (2006) demonstrated that the TAP is an ideal tool for formative usability testing as it provides rich qualitative data for understanding the nature of the interaction occurring between translators and TMs.

4.5 Technical Texts

A technical text is any text that deals with a specific knowledge such as instruction guides, installation guides and software user manuals (Byrne, 2006). The present study refers to user manuals, based on Alaoui (2015) as typical examples of technical texts. Such texts are void of any literary tendencies, do not include cultural undertones, subtexts or intricate literary techniques. They are specialized texts that usually tackle a specific topic (Hatim & Mason 1996). User manuals are technical texts that are characterized by using a simple direct language, having repetitive patterns and proper use of vocabulary (Byrne, 2006). Such
characteristics, based on Ehrensberger-Dow, 2019; Garcia, 2012; & Byrne, 2006, make them suitable texts to be translated using TMs software programs.

Using manual user guides as a representative of technical texts in the corpus of the present study is due to the consensus among Ehrensberger-Dow (2019), Garcia (2012), and Byrne (2006) that software user manuals are suitable texts to be translated by TMs. Finally, Bundgaard and Christensen (2019), Mor (2018) and Christensen and Schjøl-dager (2011) considered technical texts and manual user guides the best texts to investigate TM features.

4.6 Previous studies

In an attempt to overcome the dearth of research available on the TM segmentation on the translators, Mor (2018) studied the effect of TMs on the phenomenon of linguistic interference. The study included 90 participants divided into three groups: including professional, in-house and novice ones who worked on an unpopulated TM. He used screen casting software, keystroke and mouse logs of each process and finally direct observation by the researchers present in the classroom. The study lasted for five sessions, each one was structured in three phases of 70 minutes each, with two 20-minute breaks in between. The results reported that TMs may negatively affect cohesion of the TT rendered. However, it had points of weakness which were attributed to two variables. First, the tools used were varied in number and order. Second, the participants who took part in the study had various profiles.

Trying to understand the effect TM segmentation has on the translator's cognitive activities, Christensen and Schjoldager (2011) conducted a small scale pilot study which included 23 MA students. The data of the study was gathered through the retrospective comments participants provided on desktop-based TM. Christensen and Schjoldager (2011) provided the participants with a populated TM created by aligning authentic parallel texts from a company website in English and Danish. The participants filled out an online questionnaire immediately after the translation
process was accomplished. Students mutually reported that the software tended to control the translation process and complained that they were compelled to uncritically accept the propositions they were offered. Furthermore, unveiling the cognitive process translators went through retrospective comments was criticized by Bernardini (2001). She explained that retrospective comments can yield information that may not be consistent with what actually goes on inside the participants’ minds as the risk of distortion proliferates with time.

Bundgaard et al., (2016) studied human-computer interaction (HCI). The study had two objectives: first, to gain a deeper understanding of the interaction between TM and humans during the translation process and second, to document how the segmentation feature imposed by TM software programs restrained the translation process. The study strength was in using a two data-collection method: a semi-structured interview and a quantitative evaluation of the text produced focusing on the effect of integrating MT into TM on syntagmatic cohesion. The results showed that the one translator the research studied was always able to keep track of the text’s cohesive entity and CAT tools do not necessarily have constraints on translators in relation to segmentation. Although Bundgaard et al., (2016) tried to resolve one of the issues that TM segmentation has but the corpus selected hindered achieving this target.

LeBlanc (2013) tried to explore how Canadian translation vendors perceived translation technology in general and TMs in particular. This ethnographic study was conducted in three different translation firms and services located in a minority francophone community in Canada. The study was based mostly on interviews with translators and shadowing sessions of translators at their workstations. The results evinced that TMs had advantages and disadvantages. The advantages included increasing productivity; improving consistency; and eliminating repetitive work. As regards the disadvantages, the participants considered TMs a barrier to creativity due to the effect they have on
translators’ decision-making authority. The study revealed that TMs can contribute to error propagation in case of being polluted by inaccurate TUs. The study also criticized the TUs proposed by the TM due to the potential adverse effect they can have on the translators' satisfaction if misused. The participants labeled segmentation as a mechanical and unnatural feature which could lead to problems in the TT produced.

As for the Arab world, the literature has very limited number of researches (i.e. Mahfouz, 2018; Alotaibi, 2014). Mahfouz (2018) tried to explore Arabs’ attitudes to CAT tools through a survey and semi-structured interviews. The study revealed a favorable attitude among participants towards using CAT tools regardless of the mixed and contradicting opinions it revealed on some aspects. Although this study is considered a contribution, the methodology adopted was limited to surveys and semi-structured interviews. This study would have been more effective in case TMs had been studied from a different perspective. The sampling Mahfouz (2018) used included both translation students and professional translators and this made the sample loses it harmonious nature.

Alotaibi (2014) tried to study TM in the Arab world through using multiple data-collection methods including a pre-research questionnaire, a post-research questionnaire, semi-structured interview and classroom observations. The study targeted measuring students’ perception of CAT tools, the expectations and the level of knowledge students have towards this newly-introduced tool. The results of the study showed that the more familiarity the participants have with CAT tools; the more positive attitude they developed. The study recommended the integration of CAT tools into translation classrooms for enhancing the future translators’ skills and maximizing their opportunities in the translation job market. It was limited, however, to assessing users’ attitudes towards these tools as an external factor; it did not reach conclusions through an empirical study but through pre-research questionnaire, a post-research questionnaire, semi-structured interview and classroom observations.
Based on the literature review presented, it can be concluded that researches on TMs were mainly focusing on exploring their advantages and contributions. Even when researchers tried to study TM’s constraints, they discussed just the negative effect TMs have on the translators’ creative ability, the degrading status translators had because of TMs and the low pay translators receive when TMs are utilized. Accordingly, the current study aims to provide more empirical research on the actual interaction between translators and TM systems in practical contexts.

4.7 Methodology

The current study adopted a one group pre-test post-test design with convenience sampling through recruiting all the Egyptian male and female students enrolled in a 6-week translation course at the AUC where the researcher works. The participants received and translated two texts from English into Arabic, one as a pre-test and the second as a post-test. Convenience sampling is effective when investigating claims about products or seeking more insight about attitudes or behavior of translators. The two texts belong to the same genre to avoid the influence of any extraneous factors like memory and to boost the internal validity (Laukkanen, 1996). The current research collected quantitative from analyzing the TAP presented in the zoom video recording. The current study used non-probability sampling because it is concerned with the unique characteristics of the sample itself aside from any larger population. Convenience sampling technique was also used in the present study because of being a well-established sampling technique various translation researchers have been using. For instance, Bundgaard et al., (2016) used convenience sampling to investigate computer-translator interaction. It was also adopted by Šanca (2018) to investigate the use of CAT Tools.

4.8 Data Collection Tools

The present study obtained data through comparing the TAPs every participant produced in relation to the segmentation problem
when the participants translated without a TM and when they translated with a TM.

The first procedure for data collection took place in session (6) when the participants rendered the pre-test and was followed by rendering the post-test in session (7). Before the pre-test was administered, all the participants received a clear explanation about the purpose of the research to feel at ease and avoid any unnecessary anxiety that can hamper the study. Accordingly, the researcher just instructed the participants to perform the translation task and say out loud what comes to their mind while working on the ST. They were asked to verbalize their thought rather than interpreting them to avoid any change that may affect the sequence of the gathered information. The participants were free to verbalize in English, Arabic or both because they all were bilinguals and were already engaged in a bilingual processing task. The translations provided through the pre-test and the post-test were collected and evaluated by three different translation instructors. The verbalizations produced by every participant in the pre-test and the post-test were compared.

4.9 Results

To answer this research question, the current study analyzed the TAP produced by a corpus of 60 Egyptian translation students in the pre-test and the post-test. In the pre-test where the text is presented as a whole unit, all the participants showed a comprehensive view of the text. All the participants scrolled the screen up and down to see the layout of the text. They were aware of the beginning and end of each section which the TAP showed clearly. They were aware of the number of sub items each section includes and the way they are presented as reflected by the screen recording. All the participants did not work on statement by statement. They read the whole section first, take some time to think or look up words in dictionaries, then start the translation of the targeted section. None of the participants translated a title without reading the whole section. None of the participants overlooked the boundaries between sections as none of them
moved from one section to the other directly. They either take rest, read the section under translation as a whole or whispered being ready for the following part. This reflects their understanding of the text unity. All the participants showed awareness of term/word repetition. They all were conscious when a word/term is repeated in the same section or another different section. This repetitive action shown through screen recording while reviewing the TAP confirmed the panoramic view the participants had. It saved their time and achieved term consistency in all the sections even if the term chosen is not the most appropriate one.

The TAP of the post-test where the TM is utilized and the ST is segmented shows that all the participants are controlled by the segmentation feature. They all deal with the segment as a whole unit. The screen recording does not reflect any scrolling up or down for the segments. None of the participants shows realization of the relationship between segments. None of the participants voices the beginning of any section or the end. For the participants the segment was the unit of translation. None of the participants realizes the reference to a pronoun if it refers to a previous segment. The participants show acceptance of the segments provided.

4.10 Limitations and Recommendations for Future Research

The current study has some limitations as by relying on the qualitative data. A deeper analysis using a larger sample with a variety of dimension could confirm the results or produce more insight.

The current study investigated the effect of TM segmentation on the translator's cognitive activities when translating under the effect of TM segmentation. Another study can explore the same topic using professional translators. Also, the current study focuses on the performance of novice Egyptian translators using English as the SL and Arabic as the TL. More studies can tackle the same language pair but have Arabic as the SL and English as TL. Other
studies can investigate the effect of TM segmentation on other text types rather than user manuals.

References


The Impact of Translation Memory Segmentation on the Translator’s Cognitive activities while Translating Technical Texts from English into Arabic

Amira Mahmoud Muhammad Abdel-maksoud


[113]

