ALANAZI, MOHAMMAD S, Ph.D., December 2019 Modern and Classical Language Studies

THE USE OF COMPUTER-ASSISTED TRANSLATION TOOLS FOR ARABIC TRANSLATION: USER EVALUATION, ISSUES, AND IMPROVEMENTS (368 PP.)

Dissertation Advisor: Sue Ellen Wright

The development of technology since the last quarter of the 20th century has played a momentous role in shaping the translation process for most languages. The Arabic language, however, has encountered challenges and difficulties to catch up with the accelerated changes in computer-assisted translation tools. Those challenges have been examined and investigated extensively during the last decade. However, these tools evaluations made by Arabic language translators have not been adequately taken into consideration in the previous studies. The challenging morphological, syntactic, phonetic, and phonologic characteristics of Arabic language make it one of the most complicated languages for the use of developed translation technology, which can explain a potentially understandable negative assessment among Arabic language translators. This study examined Arabic language translators' evaluation of computer-assisted translation tools and investigated potential problems that can possibly complicate the use of the tools. Finally, the study discussed factors to take into consideration when developing computer-assisted tools to address Arabic language translators' needs. The study hypothesized that Arabic language translators would express concerns regarding language-specific issues during the use of the tools. Complications would occur for Arabic language translators while working with these

applications, e.g. MT suggestions, segmentation, punctuation and script related issues etc. To test the study's hypothesis, a mixed methodological approach was pursued that combines the following: an online survey and an observational experiment. Arabic language translators were recruited to participate in the study. A mixed approach of quantitative and qualitative analysis of the collected data were conducted to demonstrate the responses and evaluation of the participants toward the tools. The results of the study reveal a strong inclination by Arabic language translators in this study to encourage and support the use of CAT tools despite the complications (e.g., segmentation, punctuation and spelling etc.) and suggest that Arabic language translators are more likely to make changes to TM and extensive post-editing to MT suggestions. Triangulation of the survey and experiment findings supports the conclusion that there is no relationship between the complications experienced while using translation tools and the expressed level of satisfaction.

Keywords: Arabic language, MT, CAT tools, evaluation, perspectives, complications

THE USE OF COMPUTER-ASSISTED TRANSLATION TOOLS FOR ARABIC TRANSLATION: USER EVALUATION, ISSUES, AND IMPROVEMENTS

A Dissertation Submitted

to Kent State University in Partial

Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

By

Mohammad S. Alanazi

December 2019

Dissertation written by

Mohammad S. Alanazi

BA, Imam Muhammad Ibn Saud Islamic University, 2009

M.A., University of Florida, 2013

Ph.D., Kent State University, 2019

Approved by

Sue Ellen Wright_____, Chair, Doctoral Dissertation Committee

Said Shiyab_____, Members, Doctoral Dissertation Committee

Erik Angelone_____,

Michael Carl_____,

Yesim Kaptan_____,

Accepted by

Keiran Dunne_____, Chair, Department of Modern and Classical Language Studies

James L. Blank_____, Dean, College of Arts and Science

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

MT	machine translation
TM	translation memory
CAT	computer-assisted translation
TT	translation technology
L2	second language
ST	source text
TL	target language

ARABIC TRANSLITERATION

IPA Representation	Romanization Representation	Arabic Letter
aː	a	أ
b	b	ب
t̪	t	ت
θ	th	ث
dʒ	J	ج
ħ	h	ح
x	kh	خ
d	d	د
ð	th	ذ
r	r	ر
z	z	ز
s	s	س
ʃ	sh	ش
sˤ	S	ص
d̪ˤ	d	ض
t̪ˤ	t	ط
ðˤ	z	ظ
ʕ	3	ع
ɣ	gh	غ
f	F	ف
q	g	ق
k	k	ك
l	l	ل
m	m	م
n	n	ن
h	h	ه
w	w	و
j	y	ي

IPA Representation	Romanization Representation	Arabic Vowels
u	u	ُ
a	a	َ
ɪ	i	ِ
un	un	ٌ
an	an	ً
ɪn	in	ٍ

DEDICATION

إلى والديَّ الحبيبين سعدون وسالمة، إلى أخوتي عزوتي

إلى زوجتي العزيزة نوف وأبنائي الغالين ميس وأسيل وسعدون ومبارك

ما كان هذا الإنجاز ليتم من دون دعمكم المستمر

ACKNOWLDEGMENT

I would like to express my deepest gratitude to my thesis advisor, Sue Ellen Wright – who supported and guided me from the inception of this project until completion, and made every step of writing this dissertation a valuable learning experience. I would also like to extend my appreciation to Said Shiyab and Eric Angelone for their time, constructive feedback, and support as members of my thesis committee. In addition, I want to acknowledge the chair of the department, Keiran Dunne for your continuous support and assistance over the years.

A sincere “thank you” to my friend Abdulaziz Alghanam, who was always there to help me no matter the time of day; Muteb Alqarni and Abdullah Alghamdi thank you for helping me within and outside of the classroom; and Ali Arrabai, Yousef Albudairi, Yazid Alsmail, Hisham Alfayyadh, Mohammad Alzahranl, Sultan Ashlowi, Saif Aloqlan, and Sultan Almahimed thank you for your help and support throughout my doctoral candidacy.

I would also like to extend my gratitude to the professors who supported me at the University of Florida during my master’s degree – Eric Potsdam, Wayland Ratree, and Steve Flocks. I am also indebted to the undergraduate professors who guided and supported me during my formative academic years. Thank you to Caryle Murphy for your support during my graduate studies.

Last but not least, I would like to express my deepest gratitude to the government of Saudi Arabia and Imam Muhammad Ibn Saud Islamic University for sponsoring me to pursue my graduate studies in the United States. This achievement would not have been possible without their unwavering support and belief in my abilities.

CHAPTER I

INTRODUCTION

1.1 Overview

Driven by globalization and technological advances in recent years, the translation field has undergone significant changes. The tremendous developments in computer systems and information technology have led to great enhancements in the outcomes of computer-assisted translation tools, where expectations are increasing in response to current research. However, the Arabic language has “lagged behind” developments in the use of technology due to its challenging characteristics (Abufardeh & Magel, 2008, p. 275). Recent years have witnessed a significant increase in the research studies that examine and seek to develop Arabic automated translation tools in an attempt to keep abreast of development of automated and computer-supported translation for other languages.

Despite this increase in studies, Arabic language translators have been more reluctant to use new computer-assisted translation tools than other translators working with European languages (Al-jarf, 2017; Alotaibi, 2014; Fatani, 2006; Thawabteh, 2009). This dissertation aims to review the literature conducted on computer-assisted translation tools, explore the complications involved in the use of these tools for Arabic, and conduct

an observational study to examine Arabic language translators' evaluation of these tools and what potential developments can be made to improve computer-assisted tools to better meet the translators' needs.

1.2 Statement of the Problem

As stated earlier (Abufardeh & Magel, 2008), Arabic language translation has "lagged behind" developments in the use of technology. There are views from the literature suggesting that Arabic translators have been reluctant to adopt the use of computer-assisted translation tools (Al-jarf, 2017; Alotaibi, 2014; Fatani, 2006; Thawabteh, 2009, 2013). The causes of hesitancies in adopting translation tools in the Arab world as demonstrated in the literature can be due the complications the tools can introduce when they are used for Arabic language (Al-jarf, 2017; Breikaa, 2016; Quaranta, 2007; Thawabteh, 2013). Thus, it would not be feasible to spend money or time on a tool that could cause more problems than potential advantages. The complications of CAT tools for Arabic language use discussed in the literature are language related. These complications are due to the unique characteristics of the Arabic language compared to the characteristics of English and European languages in general.

The morphological, syntactic, phonetic, and phonologic characteristics of Arabic language render it one of the most complicated languages for written and spoken language processing (Boualem, 2003; Soudi, Farghaly, Neumann, & Zbib, 2012; Thawabteh, 2013). Therefore, Arabic is considered a complex language that has a rich morphology system in which words contain complex inflections. Consequently,

morphological analysis of Arabic by computer-assisted translation tools can be difficult and complicated to achieve and is plagued with complications (Attia, 2008; Quaranta, 2007; Sadat, 2013; Soudi et al., 2012; Thawabteh, 2013).

Moreover, Arabic relies on diacritics that determine the long phonemes of word forms. However, these diacritic marks are usually omitted in the majority of written texts. Arabic readers depend on the context and their knowledge of Arabic lexicon in order to overcome the resulting ambiguities. Additionally, the unique Arabic syntactic structure differs from English and European languages in its complexity and varied flexible word order, which makes translation using computer-assisted translation tools more complicated. Although it might be easy for the human mind to achieve the processing requirements, it can be very difficult for the tools to encode the lexical and syntactic ambiguities (Soudi et al., 2012).

Another challenge to using computer-assisted translation tools for the Arabic language is related to compatibility and the difficulty of digitizing Arabic texts. Although Sakhr designed an Arabic-Optical Character recognition program in 1993 to digitize Arabic texts (Zughoul & Abu-Alshaar, 2005), this program has been proven to continuously have difficulties. The accuracy of the text recognition is unreliable and huge misrepresentations occur during the digitizing process. Therefore, it would consume large amounts of time and effort to organize lengthy texts.

Despite the difficulties and complexity of Arabic natural language processing, international technology companies have invested in research for Arabic language

localization. For instance, “Microsoft has invested in worldwide research centers for many years and in this case, our Natural Language Processing researchers in our Advanced Technology Laboratory in Cairo, Egypt took the lead in developing this new language system” (Microsoft Translator, 2016). As a result of companies’ attempts to work on solving Arabic language complications with technology, most of the concerns for general use of Arabic language have been solved. Nevertheless, there is still a desperate need for immense amounts of work to be done on computer-assisted translation tools particularly by Arabic language speakers since native speakers are more likely to determine weak points and identify potential effective solutions.

As has been discussed above, the Arabic language has unique characteristics that can complicate the functions of natural language processing tools. The fact that languages are syntactically, semantically, and phonologically different supports the need to conduct studies addressing the complications and requirements for translation tools to accommodate the characteristics of Arabic language. This dissertation aims to explore the complications involved in using integrated translation memory and machine translation tools in the Arabic language. The research questions involve mixed methods that combine online surveys and an experiment designed to try to determine whether there are aspects of the tools that may not be well coordinated with a number of aspects of Arabic texts that distinguish them from the English language for which most of the tools were originally created. Arabic language translators’ views toward the tools and their evaluation are addressed as well.

1.3 Research Questions and Hypotheses

This dissertation will address the following three main questions that focus on the use of computer-assisted translation tools for Arabic language:

- What are the views of Arabic language translators when evaluating the use of computer-assisted translation tools?
- What are the problems that may complicate the use of computer-assisted translation tools for Arabic language?
- How can the potential complications involved in using computer-assisted translation tools for Arabic language be addressed?

The first question aims to explore the views of Arabic language translators toward computer-assisted translation tools. It will demonstrate their evaluation of the current tools and how satisfied they are with the efficiency of using the tools for Arabic language. This will include discussions on increasing translation productivity due to the use of the tools. This question, eventually, will lead to the second question which focuses on the potential complications that Arabic language translators encounter while using the translation tools. These complications will be explored through the use of an online survey and an experiment to cover most of the concerns that cause hesitation on the part of some Arabic language translators at using computer-assisted translation tools. Finally, the dissertation aims to demonstrate how these potential complications can be addressed to improve the use of computer-assisted translation tools for Arabic language translators.

Given that computer-assisted translation tools were designed originally for languages other than Arabic, the study poses the following the hypotheses:

H1.A: Arabic language translators will express concerns regarding language-specific complications of segmentation when using computer-assisted translation tools.

H1.B: Arabic language translators will express concerns regarding language-specific complications of punctuation when using computer-assisted translation tools.

H1.C: Arabic language translators will express concerns regarding language-specific script-related complications by computer-assisted translation tools.

H2.A: Arabic language translators will express more concerns regarding language-specific complications of segmentation when using computer-assisted translation tools for an Arabic to English translation task than when translating the text from English to Arabic.

H2.B: Arabic language translators will express more concerns regarding language specific complications of punctuation when using computer-assisted translation tools for an Arabic to English translation task than when translating the text from English to Arabic.

H2.C: Arabic language translators will express more concerns regarding language-specific script-related complications when using computer-assisted translating tools for an English to Arabic translation task than when translating the text from Arabic to English.

These hypotheses will be checked through several statistical tests as will be demonstrated in detail in Chapter Four. Moreover, qualitative analysis of the collected data will be conducted. Further details about the methodology used for this study will be presented in Chapter Three.

1.4 Potential Impact and Significance

This study hopes to contribute to the field of computer-assisted translation tools, particularly for the Arabic language. Reading the literature of Arabic translation tools has revealed several studies that examined quality and provided an evaluation of current Arabic machine translation tools (Abuelyaman, Rahmatallah, Mukhtar, & Elagabani, 2015; Alqudsi, Omar, & Shaker, 2012; Boualem, 2003; Farghaly, 2010a; N. Habash, Dorr, & Monz, 2006; Hailat, Al-Kabi, Alsmadi, & Al-Shawakfa, 2013; Izwaini, 2006; Kadhim, Habeeb, Sapar, Hussin, & Abdullah, 2013; Lopez & Post, 2013; Zughoul & Abu-Alshaar, 2005). Other studies proposed suggestions and new potential methods for the current challenges of machine translation for Arabic (Attia, 2008; Khemakhem, Jamoussi, & Ben Hamadou, 2013; Mahmoud, Shquier, & Al-howiti, 2014; Aron Phillips, Cavalli-Sforza, & Brown, 2007; Riesa, Mohit, Knight, & Marcu, 2006; Sadat, 2013; Salem, 2009; Salem, Hensman, & Nolan, 2008a; Shilon, Wintner, Science, & Landman, 2011; Shirko, Omar, Arshad, & Albared, 2010; Soudi et al., 2012).

Some studies have addressed translation memory tools to investigate the limitations and complications for Arabic language (Breikaa, 2016; Quaranta, 2007; Thawabteh, 2013). Another study by Aloitabi (2014) investigated the perspectives of undergraduate

translation students towards studying translation technology. The goal of this study was to examine the impact of teaching CAT tools and the undergraduate students' perspectives toward the translation technology. However, no studies have been found that investigated the views of Arabic language translators toward the computer-assisted translation applications and how these tools can be improved to meet the Arabic language translators' needs.

Moreover, conducting this study may bring about an increased interest by Arabic institutions in the field of computer-assisted translation tools. Despite the proven evidence of time and effort savings through the use of computer-assisted translation tools in the long run, the popularity of using newly developed translation tools is very low in the Arab world among governmental and private sectors (Almutawa & Izwaini, 2015). Investigating the proposed research question would cast light on the complications that concern Arabic language translators and institutions that may have hesitated to invest in those translation tools.

Consequently, an increased awareness of the rapidly growing role of computer-assisted translation tools may pave the way for a change in the core curriculum of translator training in the Arabic universities. Currently, the teaching methods for translation students are following the older, traditional methods where students are taught theoretically about translation studies. These methods simply give students texts to translate without introducing the students to computer-assisted translation tools (Al-jarf, 2017). As a result of the traditional methods adopted in translation schools, "most of the trainee translators, if not all, are not comfortable with the use of state-of-the-art

translation memory tools. It is also unfortunate that the trainee translators are not familiar with subtitling or interpreting software" (Thawabteh, 2009, p. 171). For instance, Aloitabi (2014) illustrates in her paper how students are discouraged from using technology in the translation process except for some electronic dictionaries. This dissertation aims to explore the current developments of the computer-assisted translation tools and to cast light on the Arabic language translators' evaluation of the tools, the complications they encounter, and how the potential complications can be addressed. This, hopefully, will lead to more positive awareness of the importance of translation tools importance in Arabic translation school curriculums.

1.5 Overview of the Dissertation

This dissertation is divided into seven chapters. This present chapter has demonstrated the purpose of the dissertation, the research questions, and hypotheses, and has justified its significance. Chapter Two is divided into three main sections. The first section demonstrates Arabic language characteristics and how these characteristics provide us with challenges to the use of technology. It also highlights the significant contributions by Arab linguists to overcome these obstacles. The second section presents the development of machine translation tools from rule-based to example-based, statistical models, and finally the use of neural networks by highlighting contributions of scholars to Arabic machine translation with different approaches and methods. The third and last section of Chapter Two discusses the shift from machine translation to computer-assisted translation tools. This discussion covers the demonstration of concerns raised by Arabic scholars regarding the use of these tools.

Chapter Three discusses the methodology used for this study. This includes the elaboration of the adopted mixed methodological approach that combines an online survey and an experiment. In this chapter, the materials, procedures, and participants for each method are thoroughly explained. Chapter Four presents the quantitative analysis of the results while Chapter Five demonstrates the qualitative analysis of the data. Chapter Six discusses the findings from both approaches. Chapter Seven will include the conclusion and elaborate on limitations of the study and the suggested future research directions as well.

CHAPTER II

LITERATURE REVIEW

2.1 Introduction

This chapter provides a broad review of the literature with regard to the notion of computer-assisted translation tools for Arabic language. In order to draw a clear picture as to where challenges of Arabic computer-assisted translation tools are embodied, it is imperative to start out by discussing a number of Arabic language characteristics and how these characteristics provide us with challenges to the use of technology. The significant contributions by Arab linguists to overcome these obstacles are highlighted. This discussion will include Arabic morphology, syntax, diacritics, Optical Character Recognition (OCR), social Arabic diglossia and finally the aversion of Arabic governmental and academic institutions toward the use of translation technology. This general view of the Arabic characteristics and its challenges with respect the use of computer-assisted translation tools paves the way for reviewing the development of machine translation in Arabic. The second part of the chapter discusses the development

of machine translation tools from rule-based to example-based, statistical models and finally the use of neural networks. Contributions of scholars to Arabic machine translation with different approaches and methods are highlighted.

Finally, the third part discusses the shift from machine translation to computer-assisted translation tools. This covers the discussions of the evolving translation workstation (e.g. translation memory, terminology management system, OCR and machine translation) and concerns raised by Arabic scholars regarding the use of these tools.

2.2 Challenges for Arabic Language Natural Language Processing

Arabic Natural language processing systems has usually employed both rule-based and machine learning approaches (Farghaly & Shaalan, 2010, p. 3). However, morphological, syntactic, phonetic, and phonological characteristics of Arabic language render it one of the most complicated languages for written and spoken language processing (Boualem, 2003; Soudi et al., 2012; Thawabteh, 2013). This section will both explore some of the challenges encountered by written Arabic natural language processing and show the contributions that have been made by scholars to overcome some of these challenges.

2.2.1 Morphological Analysis in Arabic

Arabic is considered a complex language that has a rich morphological system in which words are subject to intricate inflections. Therefore, “Tokenization is a non-trivial problem as it is closely related to the morphological analysis” (Attia, 2007, p. 65). Arabic has a non-concatenative morphology that renders it harder for natural language

processing system to process texts accurately (Aoun, Benmamoun, & Choueiri, 2010; Attia, 2007; Farghaly, 2010b; Soudi et al., 2012). For instance, one token *wa-sa-ya3lum-oun-ha* 'وسيعلمونها' can be translated into a full English sentence such as '*and they will teach her'*. As can be seen from the previous example, one token yields five strings in English or another Indo-European language. Consequently, the morphological analysis of Arabic by computer-assisted translation tools is overwhelmed with complications that can make processing difficult to achieve (Attia, 2008; Quaranta, 2007; Sadat, 2013; Soudi et al., 2012; Thawabteh, 2013).

Despite these challenges, several scholars (Al-Sughaiyer & Al-Kharashi, 2004; Farghaly & Shaalan, 2010; Soudi, Bosch, & Neumann, 2007) have addressed these complications and proposed potential solutions for the analysis of Arabic Morphology. Other scholars have developed approaches and systems for Arabic morphological analysis. For instance, Cahill (2007) present a syllable-based morphological analysis module for Arabic language. The morphological analysis in this module consider syllables as the primary component in word structures. Although this type of analysis has addressed European languages, the author argues, syllable-based analysis is possible for Arabic language as well. She has presented a comparison analysis of her findings where she demonstrates that output of Arabic syllable-based morphology analysis does not significantly differ from English and German Language. Cahill (2007) concluded that syllable-based morphology analysis is applicable to Semitic languages despite some potential phonological challenges that are encountered in in both English and German.

Furthermore, Daya et al. (2007) have presented a machine learning approach that can extract roots of a Semitic language. In this article, the authors demonstrate results they obtained from Arabic and Hebrew data. Although they argue for the challenge of extracting the roots of Semitic languages even manually by humans, they claim their approach can predict the roots with high accuracy similar to human performance. For instance, the tool as they claim can predict the root of *Katab-na* 'we wrote' as *ktb* 'wrote'.

2.2.2 Syntactical Analysis of Arabic

Arabic has a unique syntactic structure that differs from English and European languages in its complexity and flexible word order. The primary word order in modern standard Arabic is Verb-Subject-Object (VSO). However, it has a free flexible word order due to the existence of case markers (e.g. nominative, accusative and genitive) that can distinguish the subject from objects (Aoun et al., 2010; Farghaly & Shaalan, 2010; Soudi et al., 2012). These case markers are expressed either in diacritics (see Example1) or as a suffix (see Example 2).

1- زيدُ خالدً قابل

Zaid-u, khalid-an qabla

Zaid-NOM Khalid-ACC met.3

'Zaid met Khalid'

2- يدرس المعلمون طلابهم

Yudris	Al-Ma3lm-oun	talab-i-hum
Teach.Pres Singular.M	teachers.Nom.M.Plural	students.Acc.them

'The teachers teach their students'

Moreover, Arabic syntactic structure has a complex and rich agreement system. A modifier agrees in number, gender, and case with its noun phrase (see Example 3). However, quantifiers (numbers) reversely agree with the noun in gender (see Example 4)[1]. Additionally, the verb in Arabic has an agreement system depending on the word order. In VSO structures, the verb must be singular but has to agree only in gender with subjects (see example 2). However, in SVO structures, the verb must agree with subjects in gender, number, and person (See example 5) (Alqarni, 2015; Aoun et al., 2010; Farghaly & Shaalan, 2010).

3- رجال كبار ونساء كبيرات

Rijal	kibar	wa	nisaa	kabirat
Men.Pl.M	Old.Pl.M	and	women.Pl.F	old.Pl.F

'Old men and women'

[1] These rules have exceptions and are too complicated to explain in this context. See (Alqarni, 2015) for further details.

4- ثلاثة رجال وأربع نساء

Thalatha rijal wa arb3 nisaa

Three.F men.M and four.M women.F

'Three men and four women'

5- الطلاب فهموا الدرس

Al-ṯalab fahimu aldars

The.Student.Pl.M understood.Pl.M the.lesson

'The students understood the lesson'

There are other characteristics of the Arabic language that demonstrate the complexity of its structure. Arabic is a pro-drop language (see Example 6) and has no verb copula (see Example 7; Aoun et al., 2010; Farghaly & Shaalan, 2010). Additionally, Arabic syntactic structure, like that of other Semitic languages, has more complex structures that are frequently used like broad subjects (See Example 8) & clitic left dislocation CLLD structures (See Example 9).[2]

6- كتبت الرسالة.

Katab-tu alrisalah

Wrote.M.Singular the.message

'I wrote the message'

[2] See (Alexopoulou, Doron, & Heycock, 2003; Aoun et al., 2010) for further details about broad subjects and CLLD structures in Arabic.

7- أنا طالب.

Ana ṯalib

I student.M.Singular

‘I am student’

8- البيت ألوانه جميلة.

al-bayt-u alwaan-u-hu jamilat-un[3]

the-house-NOM colors-NOM-its beautiful-NOM

‘The house has beautiful colors’

Literally: the house, its colors are beautiful.

9- "وَكُلَّ إِنسَانٍ أَلْزَمْنَاهُ طَائِرَهُ فِي عُنُقِهِ"

Kull-**a** insan-in alzmna-hu ṯair-hu fi 3aniq-h [4]

Every.ACC human.GEN fastened.him deed.his in neck.his

We have fastened every man’s deeds to his neck (Qur’an, 17:13)

Nevertheless, there have been significant contributions that addressed the automatic syntactical and morphological analysis of Arabic. For instance, Diab et al. (2007) have developed an automatic machine learning-based morphosyntactic analysis system. This system, as the authors state, can process standard Arabic texts by analyzing segments and parts of speech. The accuracy rate achieved for tokenizing tested texts is 96.6%,

[3] The broad subject al-bayt 'house' as well as the narrow subject alwaan 'colors' bear a nominative case. The broad subject is assumed to occupy another A-position, spec,TP. See discussion of (Alexopoulou et al., 2003)

[4] kull-a, in Example 9 derives the accusative case from an implicit verb, alzmna 'fasten'. This type of construction is also attested in Modern Standard Arabic by (Owens, 2007).

according to Diab et al. (2007). These findings can enrich the field of Arabic language natural processing.

Al-Taani et al. (2011) have developed a syntactical analysis system that can assess the grammaticality of Arabic sentences. This system was tested on a sample of 170 short sentences from texts taught in K-12 level grades, where sentence length was between 2-6 words. The accuracy of the results was 85.88% according to Al-Taani et al. (2011). The drawback of the system is the difficulty it encounters with the longer sentences. Generally, Arabic language tends to use longer complex sentences in its structure, which affects the accuracy of text analysis.

Boudchiche et al. (2017) develop a morphosyntactic analysis system for Arabic texts. They call the system "AlKhalil Morpho Sys". This system has had two versions. The later version was developed to avoid the errors in the database. The system provides good results for morphologically and syntactically analyzing Arabic texts. According to Boudchiche et al. (2017), the system was able to analyze 99.31% of the words with a speed of 632 words per second. However, it works with fully or partially vowelized texts (i.e. diacritics are available partially or fully). It can't provide morphosyntactic analysis for non-vowelized Arabic texts. This drawback makes Arabic natural language processing harder, since most Arabic texts are non-vowelized. Table 2.1 below shows examples of vowelized and non-vowelized words in Arabic.

Table 2.1 Example of vowelizing Arabic texts

Vowelization Status	Examples	English Representation and Transliteration
Fully vowelized words	الْبَيْتُ أَلْوَانَهُ جَمِيلَةَ	al-bayt-u alwaan-u-h jamilat-un
Partially vowelized words	البيتُ أَلْوانهُ جميلَةَ	the-house-NOM colors-NOM-its beautiful-NOM 'The house has beautiful colors'
Non-vowelized words	البيت ألوانه جميلة	

Most of the earlier studies do not consider the type and complexity of the texts under investigation. The accuracy rate should be tested on complex structures to examine the ability of those systems to process potentially complex Arabic texts. Therefore, the question remains about processing structures that show no similarity to English language and the accuracy of these proposed systems.

2.2.3 Arabic Diacritics

Arabic relies on diacritics that determine the morphosyntactic status of words. However, these redundant diacritical marks are usually removed in the majority of written texts. Arabic readers depend on the contexts and their knowledge of Arabic lexicon in order to overcome complications of lexical and syntactic ambiguity (Soudi et al., 2012).

Dots and diacritics in the current Arabic script were newly introduced into Arabic in the 8th century by an order of the fifth Caliphate of the Umayyad empire (Ibnulyemen, 2017). Old Arabic original manuscripts including the Qur'an did not include these

introduced script forms. This new writing system was mainly introduced to facilitate reading Arabic texts for non-native speakers who were converted to Islam and have inadequate knowledge of Arabic (Ibnulyemen, 2017). According to Farghaly (2010b), Arabic's phonology, morphology, syntax, and dictionaries were established in the 8th century and become the Classical Arabic norm that is still taught in the Arab world today. Due to this time period of the Islamic Empire expansion when non-native Arabic speakers began switching to the Arabic language, an obvious need for a standard Arabic language appeared. This standardization process began with a set of rules and styles for the language to be taught in schools (Farghaly, 2010b, pp. 45–47).

Nevertheless, "Most contemporary texts such as newspapers, academic papers, and modern books….do not show short vowels nor do they have explicit representation of most case markings" (Farghaly & Shaalan, 2010, p. 16). The absence of diacritics from most current Modern Standard Arabic texts makes accurate automatic processing harder for Arabic natural language processing systems. Even Arabic human readers still need some specific diacritics or obvious contexts to process the actual meaning accurately. To demonstrate this challenge, an Arabic word like ('درس') *drs* without diacritics could be a singular noun 'lesson', an active past tense verb 'studied,' a passive past tense verb *'was studied*' or another passive past tense verb '*he was taugh*t' (Farghaly & Shaalan, 2010). It can be argued that the obstacle in the example mentioned earlier can be overcome and processed accurately by natural language processing systems using a syntactic analysis tool that can determine the morphological status of the word (درس) *drs*, depending on where this token can be located in a sentence. However, another example of the same

token with inflectional suffixes can be much harder one to process. To illustrate, the Arabic token 'درسوها' *drs-uha* can be translated in English as a full sentence. However, it can be an ambiguous string to translate without the diacritics or context since this sentence can lead to two possible meanings: '*they studied it*' or '*they taught it*'.

Therefore, natural language processing of Arabic language can be a challenge without an automatic analysis system that can sufficiently restore the diacritics in Arabic texts. Alzand & Ibrahim (2015) have proposed a system that translates Arabic words to English words by adding diacritics to each letter of the word through a morphological model for Arabic language processing. The authors tested the system by using 11 words from the Qur'an and 7 words from Arabic literature. 9 words taken from the Qur'an were diacriticized and translated correctly, while 4 of the seven words from Arabic literature were processed correctly. The authors do not clarify the significance of the difference in results between Arabic literature words and Qur'an words. Nevertheless, the system is a dictionary-based that use morphological analysis to translate single words. The accuracy of the output is still unreliable based on the results noted by the authors.

Furthermore, Zitouni el al. (2006) have proposed a statistical system for diacritizing Arabic words. The experimental results of this system achieve a diacritic error rate of 5.1% and a word error rate of 17.3%. Furthermore, Habash & Rambow (2007) have developed another diacritization system for Arabic texts that is based on lexical resources. They call it the "Mada-D" system. This module succeeds in reducing the word error rate to 17.2 % as compared to Zitouni el al. (2006). However, Habash & Rambow's

system encounters the challenge of unknown words, while Zitouni et al. can perform better with respect to unknown words, since their system follows a statistical approach.

Chennoufi & Mazroui (2017) have proposed a hybrid diacritization system that combines linguistics knowledge with a statistical approach. The system works through several stages; morphological analysis, syntactic analysis, diacritic rules and finally statistical processing. This system is much advanced since it can deal with sentences more accurately. The word error rate of this system, according to Chennoufi & Mazroui (2017), is 6.28 %. The improved results with this system have encouraged researchers to develop tools for better Arabic natural language processing.

Despite these efforts, accuracy of those tools is a long way away achieving efficiency and reliability with large number of texts. Alzand & Ibrahim (2015) claim that the "solution to this dilemma is not difficult but there is a need to use tactics" that are both statistical and morphological analysis can be used to diacritize Arabic words (2015, p. 231). This could be correct from a theoretical perspective, although it would need a huge amount of training data to achieve acceptable accuracy. It might be somehow easier for the human mind to achieve the processing requirements, however, it can be very difficult for the tools to encode the lexical and syntactic ambiguities (Soudi et al., 2012)

2.2.4 Arabic Diglossia

The entire world is dominated by the changes and developments of technologies. The Arabic speaking nations in the Arab world are no different. These nations are receiving advanced developments which bring new terms lacking equivalent words in the Arabic

language. Thus, Arab countries are attempting to localize terminology as much as they can. However, this constitutes a great challenge for some languages, such as Arabic, to incorporate all of these new terminologies, as different languages deal with this complication in different ways. Therefore, Arabic has witnessed great pressure to absorb the huge number of new terms coming from all over the world, especially the Western developed world.

Although there are clear, strict rules for Classical Arabic, variations between the languages exist in different Arabic-speaking countries. These country-specific variants combined with Classical Arabic are referred to as "Modern Standard Arabic" (Farghaly, 2010b, p. 47). Due to these differences, some inconsistencies are noted among the Modern Standard Arabic variants. This can be attributed to the lack of consistent rules among different countries' norms. This inconsistency can be clearly seen in technological terms, as 'mobile phone' is called 'jawal' in Saudi Arabia, 'mahmoul or 'mobile' in Egypt and 'khaliwi' in Syria.

Not only do countries' variants cause potential concerns for the Arabic language, there is an entire new kind of pressure for the Arabic language to handle: adapting to the new terms of the world, particularly the Western countries. This particular concern has divided Arabic-speakers into three groups with the following beliefs: Arabic is a language that requires no reform, Arabic needs to be simplified and to stick to its core roots (represented by the Cairo Academy), and the Arabic language needs to be completely reformed (Maamouri, 1998, p. 54). Due to these competing views, it became obvious that a compromise needs to be made; thus, 'Arabicization', a way to regenerate

the Arabic language while simultaneously adapting the language to be adequate in communicating the terminology of science and technology (Elkhafaifi, 2002, p. 255). Arabicization began by participating groups, which Mahmoud Sieny (1987) classified into the following groups: official academies, research institutes, Arab organizations and associations, colleges and universities, publishers, and individuals (Sieny, 1987, p. 168). The official academies are located in Cairo, Damascus, Jordan, and Iraq. The research institutions are located in Kuwait, Libya, and Morocco. Due to the fact that these agencies are widespread across the Arab world, it is predictable that there are difficulties in finding consistent levels in such an Arabicization process.

As noted by Saraireh (2001), there are three explanations as to why inconsistencies are present in translating technical terms into Arabic: slow, lagging work of the Arabicization agencies, lack of cooperation and coordination among them, and the gap between those who are standardizing the language and those who are using the language (Saraireh, 2001, p. 10). In support of Saraireh, Sieny (1987) notes these important concerns and also presents other causes of inconsistency in standardization among the different variants of MSA (Sieny, 1987, p. 169). Most notably, he discusses linguistic factors that include the excessive number of synonyms in Arabic, which cause translators to use inconsistent terms, opposing views on what Arabicization methods to use among the agencies, and the fact that the Arab world is divided into two groups with two different prevailing second languages (English and French) and sources of science and technology terminology (Sieny, 1987, p. 170). He also emphasizes that there is no hierarchy of standardization,

noting that there is no official body in the standardization process, a lag in production of standardized terms, poor distribution of standardized terms (Ibid).

There are other reasons for the inconsistencies in the standardization of Arabic terminology. Most importantly, there is a gap between the spoken language and the written language in Arabic, which causes the language to be difficult to standardize. This is prominent when it comes to translations of texts in Modern Standard Arabic, as there are no native speakers of this particular language; rather, it is simply learned in schools but not used consistently (Ibrahim & Aharon-Peretz, 2005). Another issue is certain countries that resist standardization because they identify themselves with their particular country's variant. Not only do they identify with their country's variants, sociolinguistics also view standardization as a form of oppression in their culture (Armstrong & Mackenzie, 2013, p. 539). Lastly, conflicts among ideology and intellectual issues can be problematic as well for Arabic standardization.

2.2.5 Arabic Optical Character Recognition (OCR)

Arabic language encounters a difficulty with regard to compatibility with translation tools due to the challenge of digitizing Arabic texts that are not already digitized. The process of digitizing Arabic texts encounters several challenges including:

- the connectivity of Arabic script, (e.g. وسيعلمونها *wa-sa-ya3lum-oun-ha 'they will teach her'*)

- recognizing the dotting (e.g. ب, ت, ث) that distinguishes between the shapes of similar letters
- diacritics (e.g. دُرِسَ)
- and shape changes in Arabic multiple grapheme cases according to their position in the word (e.g. ع, عـ, ـعـ, ـع) (Fakhr, 2011; Farghaly & Shaalan, 2010).

Despite these potential challenges, Sakhr was the first to develop an Arabic-Optical Character recognition program in 1993 to digitize Arabic texts (Zughoul & Abu-Alshaar, 2005). The company of Sakhr OCR program claims that its software has been ranked as the best Arabic OCR software by US government. However, this program has been "found inaccurate and unreliable because the resulted [*sic*] texts required a great deal of cleaning. Some whole pages were not read by the tool at all and were reproduced as images" (Alramadan, 2017, p. 53).

Nevertheless, other tools developers have had their own attempts to introduce OCRs that support Arabic language. Abbyy, a Russian global company, has produced an OCR called Abby Fine-reader (^{TM}) that supports 190 languages including Arabic. Also, RDI has emerged as a result of Ph.D dissertation by El-Mahllawy (2008). This system was adopted by King Abdulaziz University in Jeddah, Saudi Arabia as a research cooperation. Another designed OCR that works with Arabic is Tesseract. This system is maintained by Google and has been released as Apache 2.0. It is a free open source to be developed (Alghamdi & Teahan, 2018).

Furthermore, Arabic scholars have made their attempts to provide better Arabic OCR systems. For instance, Nashwan et al. (2017) have developed an Arabic Optical Character Recognition system that they call 'Holistic Arabic OCR'. This system considers words as single units to avoid segmentation errors. According to the authors, their new approach managed to achieve high accuracy for new font sizes that were not included in the system training data. The authors claim that their Holistic system outperformed Abbyy and Sakhr for computerized texts.

Moreover, Hesham et al. (2016) have presented an Arabic OCR that follows a 'Zone Classification' approach. It can classify scanned documents into text and non-text zones. Then, classified texts zones can be recognized and analyzed by the OCR system. This engine uses a morphological analysis system to be able to classify zones. The authors claim good encouraging accuracy results on multi fonts and sizes as compared to RDI and Sakhr. They state that RDI and Sakhr misclassify more text zones as non-text zones than their system which can have an impact on quality output.

However, RDI, Teseract and Abbyy in addition to Sakhr are among the most well-known Arabic OCRs in the market (Alghamdi & Teahan, 2018, p. 231). Despite these attempts to develop an OCR that is capable of digitizing Arabic texts accurately, Sakhr, RDI, Abbyy and Teseract systems "have low performance accuracy rates, below 75 percent" (Alghamdi & Teahan, 2018, p. 239). OCR is a critical part of the CAT tools where translators cannot get the full benefits of the translation tools without an appropriate OCR. (See some discussions of participants concerns about OCR in Section 5.3.8).

2.2.6 Governmental and Academic Support for Arabic Computer Tools

The complications discussed above demonstrate the Arabic characteristics that pose challenges to the use of computer-assisted translation tools for the Arabic language. Therefore, there is a need for a great deal of work to be done on computer-assisted translation tools by Arabic computational linguists since native speakers are more likely to determine weak points and identify potential effective solutions. However, Arab universities and governmental organizations have shown practically no interest in funding or working on these projects to improve the outputs of Arabic computer-assisted translation tools. There are some attempts to work on Arabic technology that have been carried out by research institutions in Saudi Arabia, UAE and Qatar. Their focus is on developing information technology systems and they have done some work on enriching digital Arabic content, developing voice recognition systems for Arabic and pure MT systems. It would be safe to state there is no research center in the Arabic world focusing on developing computer-assisted translation tools. In fact, the popularity of using newly developed translation tools is still very low in the Arab world among governmental and private sectors (Almutawa & Izwaini, 2015). Fatani (2006) demonstrated her research results about the translation industry in Saudi academic institutions as follows:

> It is clear from our survey of Saudi universities that there is a general reluctance and perhaps even aversion to introducing translation technologies. Students of translation have not been introduced to the most important advances which have been made in software design in recent years nor have they been provided with the ability to deal with modern high tech tools in the workplace of tomorrow such

> as on-line dictionaries, desk-top publishing systems and website automation technologies, and to become familiar with the commercial implications of these tools. (Fatani, 2006, p. 16)

Thawabteh (2013) expresses the same concern regarding the status of translation technology (TT) in the Arab world where he states; “Perhaps it would be safe to assume that TT seems to be of little interest in the Arabic-speaking World where linguistic-oriented approaches to translation are still seen as the academic norm” (Thawabteh, 2013, p. 82). Furthermore, Fatani (2009) compares the state of translation technology as to the development of information technology in Saudi Arabia and concludes that; “Unlike Information Technology (IT), Translation Technology (TT) has not become a strategic tool for many Saudi companies: i.e. it has not as yet become obligatory” (Fatani, 2009, para. 41).

This problem could be linked to the previous unsuccessful outputs of automated translation, which decreased the expectations for the role that translation technology can play (Alotaibi, 2014). The initiative to work in this field has been left to commercial companies or individual Arabs who work in Western universities (Zughoul & Abu-Alshaar, 2005). Consequently, this reluctance of Arabic academies adds another obstacle to the development of Arabic computer-assisted translation tools research. Without a pioneering spirit on the part of Arabic academies, Arabic translation technology will not be able to catch up to the achievement of other languages with respect to better outputs for translation tools. Thus, it is crucial for interested Arabic researchers to publish in the Arabic language and express their thoughts, contributions, and potential progress in

computer-assisted translation tools in order to attract the attention of Arab governmental and educational institutions to the field.

2.3 Development of Arabic Automated Translation Tools

Machine translation or automated translation has been introduced as part of the technological development witnessed in the twentieth century. It is defined comprehensively by Vasconcellos as;

> the technology whereby computers attempt to model the human process of translating between natural languages. The computer, rather than a person, generates the "output"—though it is only a rough draft, not yet fit for most types of consumption. The draft is usually polished into final form by a translator or a bilingual editor, though in some cases it may be used directly by a technical expert who is gathering data for ongoing research (Vasconcellos, 2001, p. 697).

The history of machine translation development can be traced back almost 70 years ago when Weaver proposed his methods for the prospects of machine translation in 1947 (Hutchins, 1986; Koehn, 2010; Zughoul & Abu-Alshaar, 2005). However, it took several decades until the demand for machine translation systems started in the business field. This demand for these systems can be because of the new development in computer systems and information technology (Zughoul & Abu-Alshaar 2005; Koehn 2010). According to Jeffrey Allen (2003), there are many reasons that led to the use of MT. The primary reason, as he illustrates, is globalization. Many corporations started to shift from

being local corporations to being multinational. As a result, business is no longer conducted in one language. In order to be successful, they need to go global. Another reason is the change in expectations in regard to the type and quality of translated material. There is an increasing need for understanding, in their native language, the main idea of a text that only exists in a foreign language with not much emphasis on high quality.

As for work on Arabic language natural processing and machine translation, several scholars claim it did not begin until the 1970s (Ali & Mnasri, 2016, p. 59; Boualem, 2003, p. 1; Elsherif & Soomro, 2017, p. 2317)[5]. However, Yngve (2000) mentions that Arabic was among other languages that have been part of COMIT project at MIT that he was working on with his team in late 1950s (Yngve, 2000, pp. 60–67). Also, Arnold Satterthwait's (1962) investigated parallel computer grammars of Arabic and English (Yngve, 2000, p. 67). Perhaps, the interest in Arabic machine translation research began with financial support from the US government as Vasconcellos (2000) statement shows below;

> Even though the CIA grant was for research on Russian, Dostert believed that insights could be gained from linguists specialized in other languages as well…..Arabic, which had been designated a priority language by the U.S.

[5] All these references have stated the start date of Arabic language natural processing as to be in 1970s without citing a reference. The original source of this information has not been found, so the accuracy of this date ’cannot be verified.

government, was tackled by Nancy Kennedy, a graduate student at the Institute (Vasconcellos, 2000, pp. 92–93).[6]

Also, the Soviet Union had its own attempts with natural language processing and machine translation research for English, German, Arabic, French and other several languages in 1950s (Piotrovskij, 2000, p. 234). Despite conflicting historical dates given with regard to the beginning of Arabic natural language processing, it can be safe to state that the comprehensive research for computer-assisted tools for Arabic translation and natural language processing of Arabic did not increase significantly until the last decade.

2.3.1 Rule-Based Machine Translation

The tremendous development in computer systems and information technology led to great enhancement in the outcomes of machine translation, where expectations are getting higher with current research. In their early days, machine translation systems started with word for word, dictionary-based translations that caused complications to occur, as the system did not account for the syntax, semantics, and pragmatics of different languages. In this context, the rule-based approach emerged to account for linguistic knowledge of languages. The approaches within the rule-based system ranged from the direct method to the transfer method, and interlingua methods (Alqudsi et al., 2012; Hutchins & Somers, 1992; Koehn, 2010). The initial development of the rule-based approach involved using a direct/literal method that analyzed input and output with basic linguistic rules (e.g., word

[6] Nancy Kennedy et al. (1959) have published a report titled "Final Report of Research in Machine Translation from English to Arabic, October 1958 to June 1959" but there was no available access to the content of this published manuscript. This work on Arabic MT was part of famous Georgetown experiment. For further details about the project see (Vasconcellos, 2000).

order). Later, a new method for a rule-based approach referred to as the ‘transfer method’ emerged for refining translation outputs to account for the morphological and syntactic analysis of texts. Another development within rule-based approach appeared with the interlingua method. This method is based on the argument “that MT must go beyond purely linguistic information (syntax and semantics); translation involves ‘understanding’ the content of texts” (Hutchins & Somers, 1992, p. 8). The interlingua method uses two steps of monolingual analysis. First, it analyzes the source language into an abstract universal language representation of meaning and then generates the resulting meaning into the target language through the use of the target-language morphological and syntactic characteristics (Alansary, 2011; Hutchins & Somers, 1992).

However, the rule-based system generally has a major weakness that handicaps its functions, as it is impossible to write down rules that cover all languages (Alqudsi et al., 2012; Charoenpornsawat, Sornlertlamvanich, & Charoenporn, 2002). Additionally, this system lacks language fluency and the ability to adjust for the exceptions that occur in various languages (Systran, 2018). Therefore, it becomes extremely difficult with rule-based systems to account for collocations, idioms, and other seemingly irrational linguistic features (Alansary, 2011; Alqudsi et al., 2012; Hutchins & Somers, 1992; Peng, 2013).

2.3.2 Example-Based Machine Translation

Revolutionary developments in technology supported new operational machine systems that have appeared on the market since the 1980s, covering several previously uninvolved

languages such Arabic. In this regard, the machine translation field attracted more attention for further research from various parts of the world. In 1984, a new approach for a machine translation system was proposed by Nagao. It is referred to as an 'example-based machine translation system' and works with a data-driven approach. In this system, a large amount of data (translated sentences) is collected as a bilingual corpus to allow the system to produce translation by retrieving relevant data from the corpus (Soudi et al., 2012). This approach can produce accurate translations in cases of given texts that match the available corpus (Alqudsi et al., 2012). This system has proved to be promising and demonstrates positive results (Brown, 1996; Furuse & Iida, 1992; Stetina & Nagao, 1997). However, it encounters the challenge of encompassing all texts that need to be translated (Soudi et al., 2012).

2.3.3 Statistical Machine Translation

In 1990, a dramatic change occurred in the field of machine translation when Brown et al. suggested statistical machine translation as a new approach, which has become the most dominant approach in the research field (Soudi et al., 2012). This system does not rely on grammatical rules – instead, the computer begins learning the second language through the probability theory (Alqudsi et al., 2012; Zughoul & Abu-Alshaar, 2005). Therefore, this approach does not need grammar rules, but it requires large-sized training data. The advantage of this approach over the example-based approach is the ability to produce accurate translations in the event that there are no similarities or matches for a given sentence in the corpus (Alqudsi et al., 2012). Although this system does not rely on linguistic knowledge as rule-based systems do, it can produce satisfactory language

fluency and has the ability to catch exceptions to language rules (Systran, 2018). It can be said that the statistical approach has provided a great contribution in the renaissance of the machine translation field because it increased interest and optimism for its outputs. However, it still has not met the requirements for quality (Alqudsi et al., 2012). Cavalli-Sforza & Philips (2012) describe how morphological information can be used in example-based machine translation to produce better quality of translated texts with the use of smaller corpora. They discuss the main differences between statistical and example-based machine translation and express their preference for the use of example-based machine translation over statistical approaches since example-based systems perform better than statistical machine translation on smaller corpora. The weakness of this method is the possibility of overgeneralizing that might occur during the text processing, which can lead to inaccurate translation, as has been discussed by Cavalli-Sforza & Philips (2012).

2.3.4 Hybrid Machine Translation[7]

A new approach that combines methods from different approaches has produced a hybrid machine translation system that incorporates the best from each approach into one single system (Soudi et al., 2012). This approach is meant to avoid the deficiency of machine translation approaches. Peng (2013) recommends the use of hybrid machine approach as he states "multi-strategy machine translation method, combining rules, corpus with

[7] Hybrid machine translation is referred to the use of multiple machine translation approaches (e.g. combination of rule-based and statistical approaches) within a single machine translation system. The same term can refer to the use of integrated translation memory tools with MT system. However, this section discusses the earlier meaning of term. The integrated TM with MT will be discussed in a later subsection.

semantic methods to complete the machine translation system is an effective way to obtain high-quality translations" (2013, p. 7129)

Habash et al. (2006) have demonstrated the challenges in Arabic-English machine translation with an evaluation of different systems and finally suggested the use of a hybrid system to achieve better quality. Additionally, Shaalan & Hosney (2012) proposed a new method for Arabic morphological rule induction through the use of inductive logic programing. This method takes the following four basic steps: word to word alignment, sentence partitioning, Arabic morphological analysis and, finally, transfer-rule induction by identifying each rule and related patterns, and then constructing the link between them. They concluded that this method proved to achieve satisfactory results in terms of translation quality. Furthermore, Khemakhem et al. (2013) developed a hybrid approach for Arabic machine translation that integrates Arabic syntactic knowledge into a statistical machine translation system. This approach, as Khemakhem et al. illustrate, can enhance the poor grammatical texts translated from English into Arabic.

2.3.5 Neural Network-based Machine Translation

During recent years, research in machine translation has switched from the now traditional statistical approach to deep neural networks systems (Ive, 2017, p. 10). A deep neural network is a powerful machine learning tool (Sutskever, Vinyals, & Le, 2014, p. 3104) that has played a role in other developed technology research in other fields but has been introduced first to the translation field by Kalchbrenner & Blunsom (2013) and Sutskever et al. (2014). This emerging approach has provided more successful outputs. Therefore, it gained the attention of researchers as an attractive alternative

approach (Ive, 2017, p. 17). Although neural network systems are somehow related to the statistical approach in that large-sized data collections are still needed for better performance, neural networks are heuristic in nature in that, on their own and without human intervention, “they learn an intricate computation” (Sutskever et al., 2014, p. 3104). The neural network machine translation model “aims at building a single neural network that can be jointly tuned to maximize the translation performance” (Bahdanau, Cho, & Bengio, 2014, p. 1).

2.4 Shift to Human-Machine Translation

The developments of machine translation reviewed above seem appealing to translators and to translation agencies since they can translate large amounts of texts very quickly and cheaply through the machine translation systems. However, a complete dependency on machine translation is still a dream in the unforeseen future (Bowker, 2002, p. 4). “Unfortunately, raw MT output cannot always meet the end user’s expectations in terms of translation quality, thus making MT plus post-editing a necessary and standard practice” (Jia, Carl, & Wang, 2019). Therefore, most machine-translated texts are still subject to post-editing by humans to meet the requirements of readability and quality (Hutchins, 2003b, pp. 9–10; Jia et al., 2019; Lagarda, Ortiz-Martínez, Alabau, & Casacuberta, 2015, p. 5).

A full dependency on machine translation systems would require highly controlled language structure, e.g., factory manuals designed to produce good readable texts (Bowker, 2002; Hargrave & Savourel, 1997). The research focus of machine translation

systems “has shifted away from the notion that machines should be designed to replace the human translators” to the way that technology “can support human translators” (Bowker, 2002, p. 4). These conclusions were already foreseen by Alan Melby in the 1980s as “Rather than replacing human translators, computers will serve human translators” (Melby, 1981, p. 28) Therefore, the aspect of computer-assisted translation tools should not be recognized as competitors “to human translators, but they are aids to enable them to increase productivity” (Hutchins, 1997, p. 113).

2.4.1 The Translator’s WorkStation (Computer-assisted Translation Tools)

The notion of interactive machine translations can be traced back to the proposals of Alan Melby and Martin Kay in the early 1980s (Kay, 1980; Melby, 1979, 1981, 1982). Seeking to develop tools according to this new shift, researchers started to think of ways to use computer-assisted translation tools to promote the translator’s work (Hargrave & Savourel, 1997). This led to the establishment of a newly developed field referred to as ‘computer-assisted translation’ (CAT tools) (Hutchins, 1998). Bowker (2002) distinguishes CAT tools from machine translation tools by explaining that in CAT tools, translators take the full responsibility to translate the text and use the tools to assist them in increasing their translation productivity.

There are many commercial CAT products available in the market that have become the workstation for translation agencies and freelance translators around the world such as SDL Trados, MemoQ, Multitrans, DéjàVu and Wordfast. However, SDL Trados is the most widely used translation tool among translators (Lagoudaki, 2006, p. 20; Moorkens

& O'Brien, 2013, 2017; Tabor, 2013). The following subsections demonstrate the traditional components of translator workstation tools including translation memory, terminology management and optical character recognition.

2.4.1.1 Translation Memory Tools

The initial developments of translation memory tools can be traced back to the 1980s (Hutchins, 1998, 2003a). Melby was the first who suggested the use of a bilingual concordance as a translation tool (Hutchins, 1998, p. 297). Melby came up with insightful proposals that have led to the use of translation memory as a translation tool (Melby, 1981, 1982, 1984). Nevertheless, these tools first started to become available on the market in the 1990s (Hutchins, 2003a, p. 14). Several commercial companies introduced computer-assisted translation tools to help translators increase their productivity. The core of these tools is a translation memory where the tool stores the texts in the form of bilingual segments (bitext). The purpose of translation memory was to assist translators in identifying identical or similar segments and provide suggestions, so the translators can reuse the translation or repetitive segments during the translation process without having to re-translate all texts (Hutchins, 2003b, p. 14). This approach saves the time and effort of translators, which can lead to increasing productivity. Therefore, translation memory tools among other tools (e.g. terminology management tools and optical character recognition software) have contributed to the workstation for translation agencies and professional translators since they are designed to facilitate the hard work of translators.

With the accelerated development in technology, “Many translators have had no choice but to embrace a technology they were not prepared for.” (García, 2006, p. 98). That is, “Technology is not an option in today’s professional world; it is a necessity.” (Gil & Pym, 2006, p. 18). Thawabteh (2013) discusses the difference between machine translation and translation memory outputs for the Arabic language in the following statement;

> MT is less efficient than TM tools. The former gives rise to many translation problems, especially in the translation of remote languages as is the case with Arabic and English. Unless it is meticulously used by the translators, MT may have disastrous consequences insofar as any translation activity is concerned. The latter, however, offer a gateway to success in translation profession if fastidious attention to technical details is paid. (Thawabteh, 2013, p. 87)

Translation memory system have become “perhaps the most aggressively marketed and widely used CAT tools in the industry” (Mcbride, 2009, p. 162). However, “few attempts to address TMs are made in the Arabic-speaking World” (Thawabteh, 2013, p. 81).

McBride (2009) has investigated the translators’ “perceptions” towards the translation memory tools. She aimed to explore the thoughts and opinions of translators regarding the use of translation memory software. The author has collected her corpus data without considering the experience or employment status of the translators from internet discussion forums (ProZ.com & TranslatorsCafe.com). In her findings, she has demonstrated the discussions of translators that included: the benefits of owning TM

system, prices and new updates, technical problems, and file formats. She concludes that her corpus didn't have any discussion regarding the language of TM interface or other languages complications. However, she asserts that this doesn't essentially mean that there are no complications but assumes that discussions of other languages complications (e.g. Arabic, Hebrew, Japanese etc.) might be posed in non-English forums.

2.4.1.2 Terminology Management tools

Terminology management can be crucial particularly with specialized texts. It reduces inconsistency with terminological usage to ensure the standardization of terms. This important feature can guarantee consistency of terms among group projects where more than one translator is working on the same project (Melby, 1992, pp. 158–159). Furthermore, It also can help "to cut costs, improve linguistic quality, and reduce turnaround times for translation, which is very important in this age of intense time-to-market pressures" (Bowker, 2002, p. 77). Moreover, terminology management can be extremely helpful for translators in some situations such as dealing with ad hoc terminology that occurs in so-called "laundry lists…for which there is no external contextual reference and no internal coherence" (Wright & Wright, 1997, p. 147). Melby (1992) mentions another advantage of the terminology management tools that "the translator can become knowledgeable in a particular domain more easily and more efficiently" (Melby, 1992, p. 160).

The initial development of terminology management tools can be traced back to the 1960s with the use of main-frame term banks by major companies (e.g., Siemens) and

national terminology documentation efforts (e.g. Termium). However, commercial terminology systems have been available in the markets for translators since the 1980s (Bowker, 2003, pp. 50–51). Surprisingly, terminology management tools have not been examined for Arabic language, despite the interest such a system holds for Arabic standardization complications. There have been found no study that discusses the use of this extremely important tool for Arabic, although it would facilitate the work of Arabic language translators by reducing the inconsistency that might occur with non-standardized terms.

2.4.1.3 Optical Character Recognition (OCR) Tool

Optical Character Recognition is "a technique that aims to automatically convert a machine-printed or handwritten text image into an editable text format" (Alghamdi & Teahan, 2018, p. 229). The automatic recognition of scanned texts is an important part of the translator workstation since printed texts can be transferred into editable texts which facilitate the translation process through enabling the use of translation memory and machine translation systems. This turns into a huge advantage for translators by saving the time and cost.

The initial development of OCR system can be associated with the appearance of digital computers in 1940s (El-Mahallawy, 2008, p. 1). However, OCR systems that support Arabic language did not appear until the 1990s. Elaboration on Arabic OCR tools and the challenges these tools encounter have been already discussed in a previous section (2.1.5).

2.4.2 Challenges encountered with CAT tools with Arabic

Despite the witnessed advanced development of translation technology, computer-assisted translation tools have encountered some complications with some languages such as Arabic. Quaranta (2007) has conducted an experimental study to evaluate the use of SDL Trados 2007 for Arabic language. In her study, she translated two sales contracts using the tool to identify potential problems that translators encounter when translating Arabic texts. In her findings, she demonstrated some of the difficulties she had encountered while using the tools. The differences between Arabic and English create several complications, including morphological issues as well as formatting, punctuation, and segmentation difficulties. Concluding her study, the author suggested implementing morphological analysis software with the tools to overcome the problems derived from complicated Arabic morphology.

Also, Breikaa (2016) demonstrated various problems that Arabic language translators encountered while using CAT tools. She summarized the problems into three different categories. First, there are situations preventing translators from using the tools, such as translating scanned documents in Arabic that, unfortunately, cannot be converted to a readable version. Second, there are language-specific complications that make the tools harder to use such as punctuation, segmentation, and sentence structural differences between Arabic and English. Finally, the author demonstrated the technical problems of the tools that include a difference in text directions, alignment tools, and dealing with tags. Dealing with these technical complications, as Breikaa illustrated, would consume the translators' time and efforts.

Thawabteh (2013) has conducted a study to investigate the problems encountered by 10 Arab graduate translation students while using the Translator's Workbench translation memory tool .[8] The author discussed the linguistic and technical. limitations of the tool. The tool had the inability to handle diacritics of Arabic language, which can change the meaning of the text in most cases. Moreover, the morphological analysis, matching and segmentation processes were the core of the limitations and deficiencies of the tool.

To demonstrate which CAT tool has better features for Arabic-English translators, Moujaes (2016) has conducted a comparative study between SDL Trados and MultiTrans in terms of quality features for the user in Arabic and English language translation. He tested the tools according to specific criteria, including reliability, usability, performance, etc. The author concludes that SDL Trados outperforms MultiTrans in quality features, especially the time factor. SDL Trados, according to Moujaes (2016), is much faster in performance than MultiTrans.

Other studies have addressed the pedagogical aspect of teaching computer-assisted translation tools. Al-jarf (2017) discussed the lack of technological knowledge among translation students in King Saud University in Saudi Arabia. She emphasized the importance of training students by specialists in the CAT tools and other translation-related technology to meet the demands of the modern market. Also, Alotaibi (2014) demonstrated the relationship between translation students and the use of computer-

[8] Translator's workbench is an outdated tool that was released in 1992. It was replaced later by SDL Trados Studio in 2007 (SDL Trados, n.d.)

assisted translation tools. In her study, Alotaibi focused on the impact of teaching CAT tools on students' perspectives toward the technology. With more than 100 participants in her classes, Alotaibi conducted a mixed methodological approach that included questionnaires, class observation, and interviews with her students. In her findings, Alotaibi revealed a lack of knowledge among students regarding CAT tools. The majority, as she stated, have shown interest in learning the tools. Also, the author discussed the set of emotions her students have expressed during the classes, which ranged from worry, doubt, anxiety, disappointment, enthusiasm, and excitement. According to Alotaibi, these emotions did not change completely, even at later stages of the class. In her conclusion, she argued for a relationship between increased knowledge of the tools and the change into positive perspectives on the part of the students toward the use of computer-assisted translation tools. Mahfouz (2018) reaches similar findings in her survey results of 114 translation students and professional translators in Egypt. She concludes that participants with better computer skills and more year of experience in CAT tools environment have more positive "attitudes" toward the use of CAT tools in their translation work. Although she finds that participants showed they have some difficulties in the use of CAT tools for their work, the results in general show positive "attitudes" about the use of CAT tools. She also addresses some of the translators' concerns with CAT tools environment as training requirement, problems with coherence of texts due to segmentation style and lower creativity in translation due to the dependence on previous translations.

Despite the challenges and difficulties that have occurred with these tools for specific languages, the demand for CAT tools has increased in the industry. Professional translators and translation agencies have adopted the use of translation memory tools to save costs and increase productivity. Researchers have expressed some concerns about "blind faith" among translators in the content of translation memories as described by Bowker (2005, p. 19). In an experimental study to measure the trust of translation memory content among new users of translation technology, Ford (2016) conducted an experiment that involved Saudi professional translators as well as students. She discovered that all Saudi professional translators performed better jobs in editing the "fuzzy" matches than students. However, all of these translators failed to notice the incorrect 100 percent identical matches that were provided for them as a test of their ability to catch those wrong matches during the given translation task, an indication of full trust in the content of previous human translation.

2.4.3 Integrated Computer-assisted Translation Tools

In recent years, there have been significant developments in computer-assisted translation tools. One system called EURAMIS can integrate a translation memory with machine translation and has been used by the European Commission since 1995 (Hutchins, 2003b, p. 16). However, the first commercial translation memory tool in the market that was integrated with a machine translation tool was DéjàVu X (Lagoudaki, 2008, p. 263). Also, most commercial companies have developed these translation tools integrated with MT translation. SDL Trados studio, as an example, provides a translation memory tool

that can integrate MT suggestions when there is no available translation memory for a given translation segment.

The task of translators then has changed from translation from scratch to post editing the translations provided by the translation tools. There are many definitions for post-editing and all revolve around the same meaning. Jeffrey Allen (2003), borrowing Veal and Way's definition of post-editing, defines it as "the correction of machine translation output by human linguists/editors" (Allen, 2003, p. 297). He indicates that the task of the post-editor is to edit, modify and/or correct pre-translated text rather than to translate "from scratch". Krings (2001) discussed the difference between the human translation and post editing. In human translation, the translator expresses the meaning of a source sentence or text in another language while taking into consideration both source and target culture. However, the post-editor checks the translation against the original source text for unintended omissions, misrepresentation or better lexical choices. The emphasis is on adjusting relatively predictable difficulties rather than on the discovery of unintended lapses or errors. The differences in both types of translations are relevant to frequency, repetitiveness, and error types. In terms of frequency, where a human translator might mistranslate a word once, machine translation will result in consistent mistranslation of the same word throughout the whole text.

Nevertheless, post-editing has never been an easy job for translators. Therefore, several studies have been conducted comparing translation to post-editing based on the cognitive effort involved. For instance, Sekino (2015) conducted a study on Japanese into Portuguese and found that there was not much difference between post-editing and

translation for both cognitive effort and time levels. She ascribed the findings to the huge grammatical differences between Japanese and Portuguese.

Lee and Liao (2011) conducted a study comparing human translation to post-editing of machine translation output. Participants in this study were students. They used time only as measure of effort, which might raise some doubts about the results. They concluded that the results indicated that the MT text was very helpful in reducing errors in some student translations; the use of MT also reduced the gap between students of divergent language proficiency levels. It should be noted that their aim was to see if machine translation could help reduce student errors.

Using keystroke logging and eye tracking, Koglin (2015) conducted a study comparing cognitive effort in translation and post-editing. The study used English into Spanish as a language combination. Data analysis shows that the cognitive effort required to post-edit a MT output is lower in comparison to manual translation.

In a study to evaluate the translators perspectives of the current tools, Moorkens & O'Brien (2013) investigated translators perspectives toward the post editing of the machine translations using translation memory tools. More than a hundred professional translators participated in the survey. Most participants used SDL Trados for their post editing, while others used Microsoft Word. In their findings, the authors demonstrated the dissatisfaction that translators have expressed toward the current tools environment. The participants suggested some improvements to the user interface, which include

providing access to features like dictionaries, Internet search, and better concordance search.

Through a web-based interactive computer-assisted translation tool, Federico et al. (2012) have measured the translators productivity levels and efforts while using SDL Trados that integrate machine translation suggestions. Twelve professional translators (English-German & English-Italian) were given translation tasks where the first half of the text has translation memory matches and the other half depended on Machine translation suggestions and Translation memory. In their findings, the authors show that the efforts of the translators had been decreased significantly when translators had access to both machine translation suggestions in addition to the translation memory. Additionally, translators were able to save time when they moved from translation memory mode to both translation memory and machine translation suggestions.

2.5 Summary

This chapter has reviewed the literature conducted on the Arabic language characteristics and its challenges to natural language processing. Also, the developments of machine translation and the contributions made for Arabic language have been discussed. Then the discussion traces the research focus shift from the notion of complete dependency on machine translation into human integration with machine translation. However, the literature shows a lack of studies focusing on human interactive machine translation for Arabic language. This current study will attempt to fill this gap found in the literature by investigating the views of Arabic language translators toward the computer-assisted

translation applications and how these tools can be improved to meet the Arabic language translators' needs. The next chapter will describe the methodology of the study, explaining in detail its design, procedure, materials and participants.

CHAPTER III

METHODOLOGY

3.1 Introduction

As discussed in Chapter Two, with the exception of anecdotal comments, there has been a lack of experimental research investigating Arabic language translators' responses to and evaluations of computer-assisted translation tools. This chapter explains the methodology used to answer the research questions stated in Chapter One. This research study will examine the views of Arabic language translators when evaluating computer-assisted translation tools used for Arabic translations, the problems that may complicate the use of the tools, and how these complications can be addressed to better meet the needs of Arabic language translators. To answer the research questions and examine the research hypotheses, a mixed methodological approach was adopted that combines an online survey and an experiment. This combination of varied methods will provide the advantage of triangulation analysis, which can increase confidence in the research data, reveal distinctive findings, and provide a better understanding of the problems involved (Thurmond, 2001). In addition, triangulation, in the view of Shreve and Angelone (2010), "is the use of two or more data acquisition methodologies within a single study to

improve the quality, validity, and reliability of research findings" (Shreve & Angelone, 2010, p. 6)

3.2 Research Approach

As noted, this study's research methodology utilizes a mixed approach that combines more than one method, whereby the two components of the study involve both qualitative and quantitative data analysis. "Mixed approach" refers to the use of more than one method for data collection or analysis. However, it is "often understood to mean using both qualitative and quantitative approaches" (Saldanha & O'Brien, 2014, p. 23). Green (2015) argues in favor of mixed-methods research and states that it provides "the possibility of meaningful engagement with and dialogue across not just different types of methods and data but also different logics of inquiry, different ways of knowing, and thus different perspectives on understanding" (2015, p. 608). Although the use of quantitative procedures can provide the generalizability of results to go beyond the level of exploratory studies, Göpferich (2008) emphasizes that "qualitative analysis should not be neglected either as it provides insights that cannot be gleaned through quantitative procedures" (as quoted in Kiraly, 2013, p. 204).

Given the nature of the research questions posed in this study, the use of mixed approaches is appropriate as it "allows us to tackle complex topics (whether related to translations or translators) with a flexibility that surpasses that of the qualitative or quantitative approaches alone" (Meister, 2018, p. 77). The present study is concerned with understanding Arabic language speakers' evaluation regarding the use of computer-

assisted translation tools and how these tools can be improved to meet their needs. Kiraly (2013) illustrates the importance of using qualitative analysis in complex topics:

> … if we are trying to understand extremely abstract and complex processes, it may well be the case that statistical studies are simply not up to the task – or might, at best, supplement qualitative studies, rather than the other way around. (Kiraly, 2013 p. 205).

In addition, mixed methods research, according to Meister (2018), "is well suited to the nature of research in translation studies" (2018, p. 78), as it can promote "internal coherence in research design and implementation by enforcing reflexive and conscious choices in all phases and at all levels of the research process, whereby it also foregrounds ethical aspects and implications of research" (Meister, 2018, p. 77). Furthermore, she adds in her argument in favor of the mixed methods approach, that this type of research is "well worth exploring in greater depth for all those who pursue translation studies research that crosses the traditional boundaries of the qualitative and quantitative approaches" (Meister, 2018, p. 79).

The following sections in this chapter elaborate in detail the methodology used for this study and provide a thorough presentation on how the research was conducted using two separate methods including an online survey and an experiment.

3.3 Survey

3.3.1 Participants

A survey was conducted in order to investigate the participants' previous experience with the evaluation of translation tools. The study aimed to collect up to 100 responses and to encourage as many participants as possible, so drawings for gift cards (10 Amazon or other online vendors with $20 value) were offered as incentives. Nevertheless, only 57 participants started the survey and only 49 have completed all the questions to date. About eight participants have chosen not to complete the whole survey. This is considered a small withdrawal number, given the length of the survey, which might take up to 15 minutes to complete, and the importance of time to translators, who may not be willing to volunteer their time to complete online surveys.

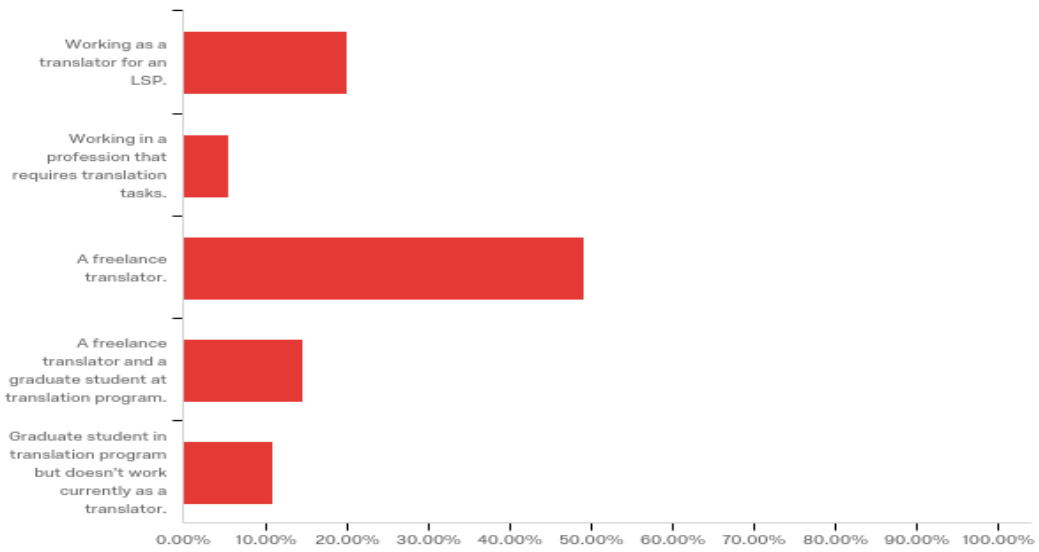

Figure 3.1 Participants' Employment Status

The employment status of participants as provided in Figure 3.1 reveals that 49% of the participants were freelance translators, 20% worked for language service providers and 15% were freelancers and graduate students at the same time. This indicates that the majority of the participants are freelancers. As for participants' years of experience as shown in Figure 3.2 below, 45% of participants have 1-5 years of experience, while 24% have 6-9 years, and 29% of participants have more than ten years of translation experience.

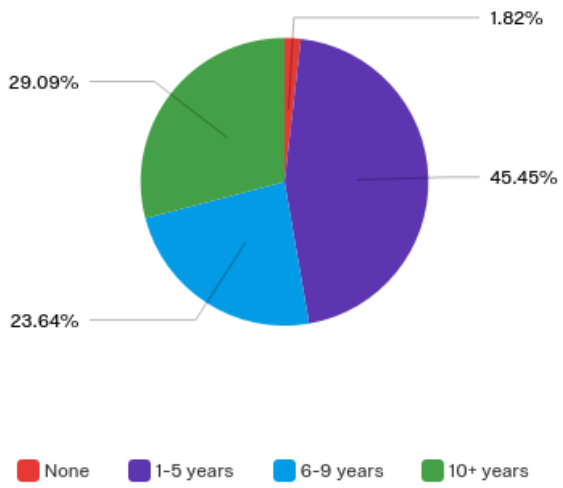

Figure 3.2 Participants' Translation Experience

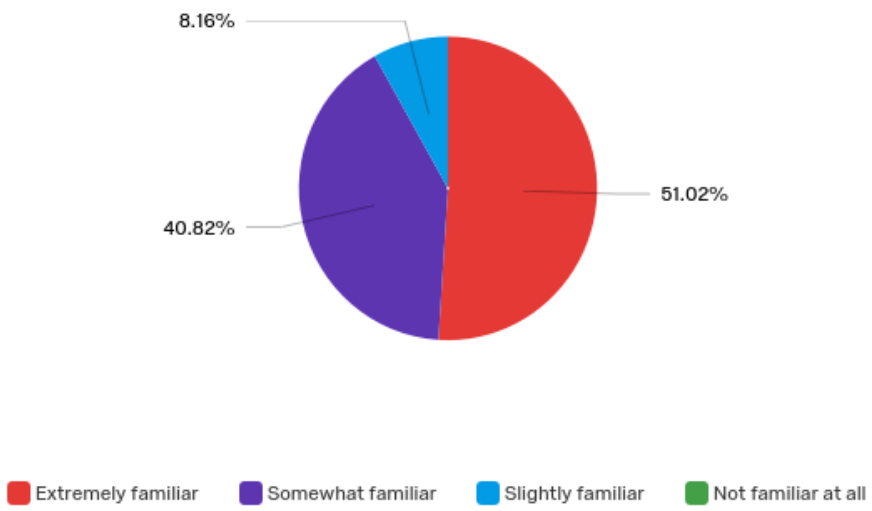

Figure 3.3 Participants' familiarity with CAT tools

Figures 3.3 and 3.4 illustrate the familiarity of participants with computer-assisted translation tools and MT tools. Familiarity refers to how well someone can use and interact with CAT tools to produce translated texts by using the tools. Extremely familiar means that a translator has used CAT tools quite enough to the extent that he/she can use them without much effort and without asking for help from another person. Both figures reveal that the participants have enough knowledge about the tools to participate in the survey. Figure 3.4 below shows that participants' familiarity with MT is less than computer-assisted translation tools in general. Furthermore, participants were asked to rate their knowledge of CAT tools from 1 out of 10. The responses ranged from 5 to 10 and the average was 7.5 out of 10.

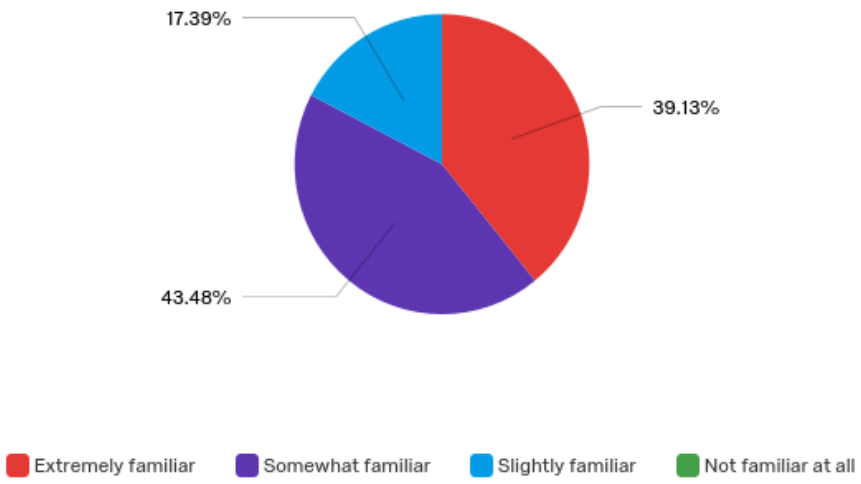

Figure 3.4 Participants' familiarity with MT tools

3.3.2 Materials and Procedure

The survey was created using Kent State Qualtrics and was posted after receiving the Institutional Review Board (IRB) approval #18-211 for a closed group on Facebook called WeArabize [Home of CAT Tools, Dictionaries and Subtitling], which has more than 14,000 members. After two weeks of low participation, the decision was made to post the survey on another closed group on Facebook called Arabic Freelancers. This could explain the significantly high number of freelancing participants in the survey, since the majority of members in those groups are freelancers. The members of both groups should be familiar or interested in the tools since the groups' goal is learning about tools use and complications for Arabic translators. However, to rule out meaningless contributions, the survey included questions about the participants' knowledge, experience, and their actual use of the tools.

The survey contained 18 questions, which included both open-ended and close-ended questions. The questions are designed to reveal the background of participants (e.g., experience, actual use of tools, and education level), their assessment of the tools, the problems they are encountering while using the tools, and, hopefully, their thoughts on how these complications can be addressed to better meet their needs. The survey has taken into consideration both quantitative and qualitative data (See Appendix A). The time period required for data collection took over three months to reach the minimum required numbers. The survey was conducted between January 5th and March 14, 2019. Once it was apparent that there was no new participation anymore despite attempts to keep the post on top on the Facebook groups, a decision was made to end data collection for the survey on March 15, 2019.

3.4 Experiment

The experiment required participants to perform the following tasks: translate two texts in two directions (from Arabic to English and vice versa) using SDL Trados Studio 2019 Then immediately after completing the translation tasks, respondents were required to participate in a short interview to discuss their experience using the tool.

3.4.1 Translation Task

The first phase in the experiment required the participants to perform two translation tasks while being observed by cued retrospective protocol[9] via screen recording to investigate complications that occurred when using computer-assisted translation tools. Keystroke logging was excluded since it does not work currently with SDL Trados tool. The participants worked at a desktop computer with a keyboard and a mouse. The translation tasks were performed using the SDL Trados Studio 2019 translation memory tool integrated with MT and fed with a translation memory that covers 45% of the test passages, allowing for MT suggestions for the missing segments. SDL Trados was chosen since it is the most widely used translation tool among translators (Lagoudaki, 2006, p. 20; Moorkens & O'Brien, 2013, 2017; Tabor, 2013), and the tool is familiar to the participants. The Flashback Express 5 recorder software tool was installed to record the screen during the translation process. There was no video recording of participants during the translation tasks to avoid distraction or discomfort.

The required translation tasks were performed in both directions, English-Arabic, and vice-versa, whereby the text fit the following parameters:

- Both texts reflect typical translation problems
- Both texts consist of five to seven sentences
- Both texts were designed to reflect segmentation, punctuation and script-related problems

[9] 'Cued retrospective protocol' refers to the verbal report produced retrospectively with respect to a cue of a screen recording in this study. For further details about the protocol see (Saldanha & O'Brien, 2014).

The overall word count was estimated to be around 100 words for each task (See Appendix 3). Both translation tasks are from the same genre and have the same difficulty level. They both contain educational medical content that provides basic information about specific diseases. The Arabic text was about oral cancer and the English text was about anorexia. The Arabic-English text was divided into three segments, while the English-Arabic text was divided into four segments. The Arabic text had fewer segments than the English one since the Arabic language tends to have longer, more complex sentences than English. The translation memory coverage for the segments of both translation tasks was purposefully designed to present three different conditions: the segment was provided with a fuzzy match of 70-80 %, the segment was presented with 100% match, and the segment was not covered by the translation memory, in which case SDL Cloud machine translation would provide the suggested translation. In all conditions, participants had the option and ability to edit any segment they thought needed it.

3.4.2 Semi-Structured Interview

In the final stage, each participant was interviewed after completing the translation task in order to explore their experience with the tools. This involved face-to- face interviews to inquire about their experiences and their assessment of the tool, what problems they encountered, and what suggestions they have to improve the quality of the tool. To obtain sufficient information from the participants, a semi-structured interview involving closed-ended and open-ended questions was conducted, which was designed to allow the

interviewees to express their thoughts in depth (See appendix B). The interviews were audio-recorded and transcribed for accurate data capture.

By the end of data collection, a total of 190 minutes of interviews had been recorded with 13 participants. The interviews were transcribed into a MS Word document that is 60 pages and around 15 thousand words. It took more than four weeks of careful and repeated listening to ensure accurate transcription. Although this was time-consuming, "the time spent in transcription is not wasted, as it informs the early stages of analysis, and you will develop a far more thorough understanding of your data through having transcribed it" (Braun & Clarke, 2006, p. 88).

3.4.3 Participants

Thirteen graduate students from the translation program at Kent State University were recruited. The participants were selected based on their experience using computer-assisted translation tools. This requirement for recruiting participants was designed to avoid any problems based on lack of sufficient experience with translation tools. Therefore, the selection process started with second-year master's students and Ph.D. students who have previous experience or familiarity with computer-assisted tools. However, the experience of participants with the tools varied. Their experience level using computer-assisted translation tools ranged from one year up to eight years. The participants' education level was as follows; six PhD translation students and seven masters level students who are in their second year. It is worth mentioning that the population of Kent State University students might not reflect typical Arabic-English

users of CAT tools. These students start with more extensive training in CAT tool utilization and the theory underlying CAT tools. Their perspectives toward CAT tools might not be the same as other Arabic language translators who were not introduced to CAT tools in their Academic programs.

Participating in this research study was voluntary and has had no effect on the subjects' progress in their studies. Participants were offered a $25 Walmart gift card upon their completion of the required tasks. They were given a $10 Walmart gift card if they decided to withdraw from the translation experiment or the interview at any time. Nevertheless, all participants chose to complete their tasks with no signs of discomfort or hesitation. Participants' identities were confidential; therefore, their names or any information that might disclose their identities were hidden. All participants were identified by pseudonyms when direct quotations were used within the demonstration of qualitative data analysis section.

3.4.4 Research Procedure

Once Institutional Review Board (IRB) approval requirements had been resolved and the SDL Trados Studio with cloud machine translation had been installed and was ready to use, an invitation for participation in this experiment was sent through e-mail and personal contact with the students. The experiments were conducted only on one computer, so participants who volunteered to participate were scheduled for appointments after signing the consent form and providing their contact information, thus

each participant was assigned a specific time and date. The experiments were conducted in the period between April 12th and May 01, 2019.

On the scheduled experiment date, each participant was briefed on the procedure for the translation tasks and the computer equipment was tested and prepared for each participant. The participants were informed about the audio-recording of the interview and assured that their confidentiality, anonymity, and privacy would be strictly maintained throughout the study. Then, the translation tasks and the screen recording started. The participants were asked to translate general Arabic and English texts (Arabic-English and English-Arabic) that were covered partially by the translation memory. The translation memory covered around 45% of the texts. For the rest of the non-covered text, the MT tool in SDL Trados Studio provided the participants with translation suggestions and then participants needed to post-edit the suggested translation. The learning and update mode of MT and TM were switched off to ensure that the same translation suggestions were provided for each participant and to prevent the system from incorporating previous solutions into new suggestions during subsequent trials. The translation process took between 12 and 30 minutes to complete depending on the individual differences between participants. The average time was about 16 minutes for all participants to complete both tasks.

After they completed the translation task, the participants were interviewed to ask them about the process they experienced while using the computer-assisted translation tool. The screen recordings of the translation process were reviewed during the interview to

trigger the participants' recollection of the experience they had during the experiment. The length of the interview was between 10 and 20 minutes depending on the participant.

3.5 Data Elicitation

The evaluation of Arabic language translators regarding computer-assisted translation tools was triangulated using the mixed methodological approach of a survey and an experiment. Both quantitative and qualitative approaches were adopted to analyze the data. The quantitative aspect of the data was extracted from the close-ended questions that were included in the online survey and the experiment.

The data analysis started with the quantitative aspects of the data and then qualitative analysis was conducted. This manner of analyzing the data quantitatively and then qualitatively subsequently granted "the potential advantage of exposing some trends that can be further probed via qualitative data" (Saldanha & O'Brien, 2014, p. 23). A qualitative analysis was crucial in this present study since the nature of the research questions was concerned with understanding the Arabic language speakers' evaluation regarding the use of computer-assisted translation tools and how these tools can be improved to meet their needs. Thus, the study results in general were based on the assessment of the tools and what problems participants encounter for Arabic translation. Consequently, the use of mixed approaches that combine both quantitative and qualitative analysis would be more valid and appropriate since it would take into consideration all possible insights that cannot be found through quantitative analysis alone (Kiraly, 2013;

Meister, 2018). Table 3.1 provides the general objectives of both methods used, the online survey and the experiment.

Table 3.1 Demonstration of Survey and Experiment objectives

Survey	Experiment
• Addresses Arabic language translators' perspectives toward the CAT tool environment. • Measures the satisfaction level among Arabic language translators with the current CAT tools. • Explores the concerns and complications regarding the use of CAT tools. • Addresses participants' suggestions for the developments of CAT tool.	• Examines the participants' perspectives toward the use of CAT tools. • Examines the evaluation and satisfaction level among participants toward CAT tools. • Examines the complications of segmentation, punctuation, and script related problems and how these complications can impact the participants' perspectives toward the CAT tools. • Addresses the participants' suggestions for developing CAT tools to meet their needs.

3.5.1 Quantitative Analysis

The responses to the closed-ended questions (quantitative data) from the online survey were analyzed using Qualtrics survey analysis software. The results were reported with charts and graphs to demonstrate the responses of participants.

As for the experiment, data were extracted from the transcribed script based on the answers to the closed-ended questions and then entered in the Qualtrics SPSS statistical analysis tool. The categorical variables were re-coded in SPSS to yield numeric values to allow statistical analysis of the entered nominal variables (see Appendix L). In addition, the screen recordings were reviewed to extract the time each participant spent on each translation task and the total time for both tasks and examine the participants behavior toward the translation tools (e.g. responses to fuzzy matches, TM and MT suggestions; see Appendix L). The research study has four variables: 1) *segmentation*; 2) *isomorphism between English and Arabic source and target texts*; 3) *tools evaluation*; 4) *level of complexity*. The first two variables, *segmentation* and *isomorphism* between English and Arabic translation have been treated as the independent variables and then manipulated to study their effects on the dependent variables: *tools evaluation* and *complications* encountered by participants while using the tools. The time spent on the translation process was used as a measurement of the dependent variable *level of complexity* that participants encountered. A full statistical analysis was conducted to determine if there are significant relationships between the independent and dependent variables, which will be demonstrated in detail in Chapter Four.

3.5.2 Qualitative analysis

The next phase of data analysis involved conducting a qualitative analysis of the collected data from the open-ended questions from both the survey and the interviews. The qualitative analysis of the interview data was based on the adopted thematic analysis approach through the use of the NVivo qualitative analysis tool. The strategies of analysis consisted of six phases adopted from (Braun & Clarke, 2006). The first phase involved familiarization with the data. The familiarizing process started with transcribing the audio recordings of the interviews into a Microsoft Word document, followed by relistening to the recordings and reading the scripts simultaneously to confirm accurate transcription of the data. Next, reading the scripts several times while classifying initial codes and themes using the NVivo tool. The second phase generated initial codes from the data. Braun & Clarke (2006, p. 88) state that "Codes identify a feature of the data that appears interesting to the analyst, and refer to the most basic segment, or element, of the raw data or information that can be assessed in a meaningful way regarding the phenomenon". The coding process followed a theory-driven approach aimed at looking for answers to the research questions. The option of a data-driven approach was excluded since it would lead to irrelevant codes for the addressed research questions. The third phase in this process was searching for themes. This step aims to sort the initial codes into broad themes. The next and fourth phase begins with reviewing the sorted themes. This involves carefully reading the coded data to annotate and classify the presented themes and then to find coherent patterns within the themes. The fifth phase requires defining and naming the found themes. This phase also involves organizing the themes to be

consistent within and related to the research questions. The sixth and final phase is reporting the analysis of results. This will include a demonstration of the results in thematic formats with direct quotations from participants and discussions to clarify the participants' evaluation of the translation tools and descriptions of their thoughts (Braun & Clarke, 2006).

As for the qualitative survey data that were collected from open-ended questions, they were coded manually because of the low number of responses that participants provided. The data collected were compared to the themes extracted from interview questions and then combined to the appropriate pre-defined themes. Other responses that did not fit into the pre-defined themes, which included other complications involving tools use for Arabic language that were not covered in the experiment interviews data, were reported separately and discussed in the context of the qualitative data results (see Section 5.3.8 in Chapter Five).

Moreover, the observed experiment participants' behavior from screen recordings will be correlated with their responses to their views and evaluation to the tools. Following this stage, another round of the analysis involved validation of the complications raised by translators in both the survey and experiment. This stage of examination aimed to identify the complications and to rule out any problems that occurred due to participants' lack of experience with the tools. This also involves careful searching for potential solutions that IT specialists might not have already taken into consideration, which will be demonstrated in the (discussion section) in Chapter Six.

CHAPTER IV

RESULTS (I): Quantitative Data

As discussed in Chapter three, data collection consisted of an online survey and an experiment. This chapter analyzes and demonstrates the quantitative results from both methods to answer the three research questions posed in Chapter one.

4.1 Survey Results

4.1.1 General Observation of the Survey Results

There are a number of international companies that offer computer-assisted translation tools in the market worldwide. Most of these available tools support a variety of languages. This level of flexibility provides translators with more options to find which tool works best for them. In the survey, participants were asked which computer-assisted translation tool they have been using in their translation work. They were able to choose more than one tool if applicable. Around half of the participants chose SDL Trados Studio, 28% chose MemoQ, 14% chose WordFast and around 12% chose other tools, as can be seen in Figure 4.1 below. The results of this question are similar to the findings of (Lagoudaki, 2006, p. 20) that SDL Trados is the most widely used tool among translators.

As revealed by the survey results, SDL Trados seems to be the most popular tool among Arabic language translators as well. Also, the results from this survey question are similar as well to the results of the ProZ survey by (Tabor, 2013) about the translation tools used among translators around the world, which found 43.2% of translators use SDL Trados, 11.5% use Wordfast, and 8.% use MemoQ. Figure 4.1 below demonstrates the CAT tools most frequently used among survey participants.

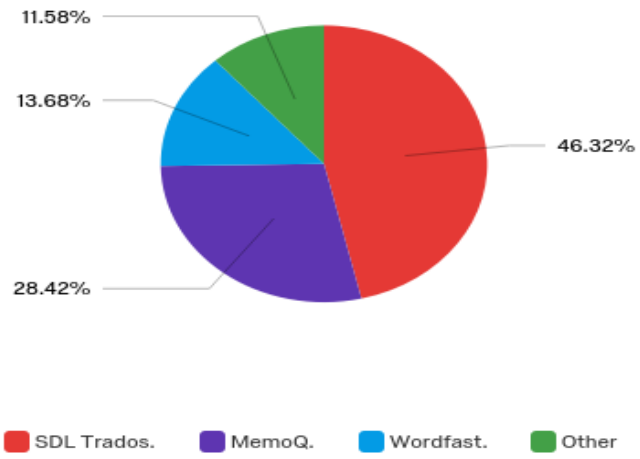

Figure 4.1 The commonly used translation tools among participants

On a topic related to the translation tools, participants were asked whether they think that machine translation is beneficial for Arabic language translators to use during the translation process. Most participants agreed that the use of MT is beneficial for Arabic language translators, while only 22% thought otherwise, as can be seen in Figure 4.2 below.

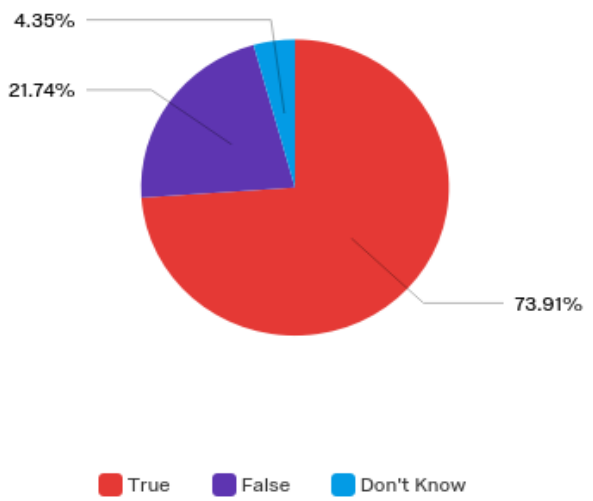

Figure 4.2 Whether participants think MT is beneficial for Arabic Language translators

Furthermore, participants were asked how important the use of machine translation engines is for Arabic language translators. Importance refers to the significant value for using modern tools that incorporate MT with TM. Around half of participants thought it is very important, 26% revealed that it is somewhat important, 9% thought it is slightly important, while 13% expressed that it is not important at all, as can be seen in Figure 4.3 below. The two figures below present the views of survey participants regarding the use of MT in their translation work.

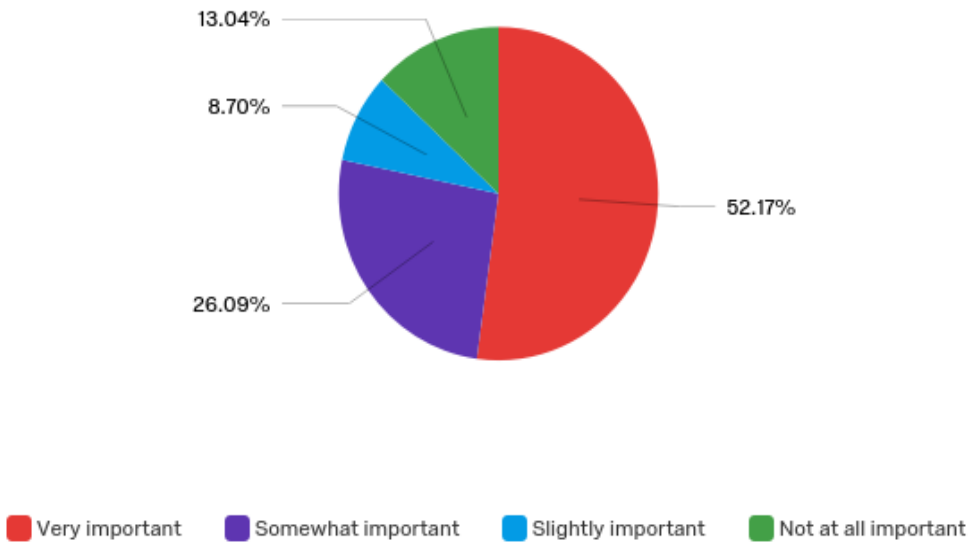

Figure 4.3 The importance of using MT among Arabic language translators

4.1.2 The Quantitative Results of the Survey

As has been discussed in Chapter Two, CAT tools have some complications when used for Arabic language translation (see for further details: Breikaa, 2016; Quaranta, 2007; Thawabteh, 2013). Therefore, participants in the survey were asked if they encountered complications in the use of computer-assisted translation tools for Arabic language. As illustrated in Figure 4.4, 61.2% of participants reported that they had complications while using the translation tools for Arabic language while around 32.6% revealed they had no complications.

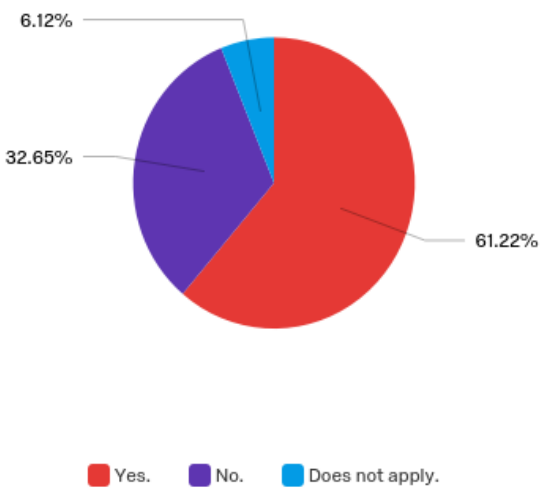

Figure 4.4 Participants' response about whether they encountered complications with the tools or not

Results of the McNemar non-parametric test revealed statistically significant association between online survey participants who reported complications and those who didn't, p = .039. Therefore, a significantly larger percentage of participants reported that they experienced complications while using CAT tools for Arabic language (61.2%) than participants who didn't report complications (32.6%).

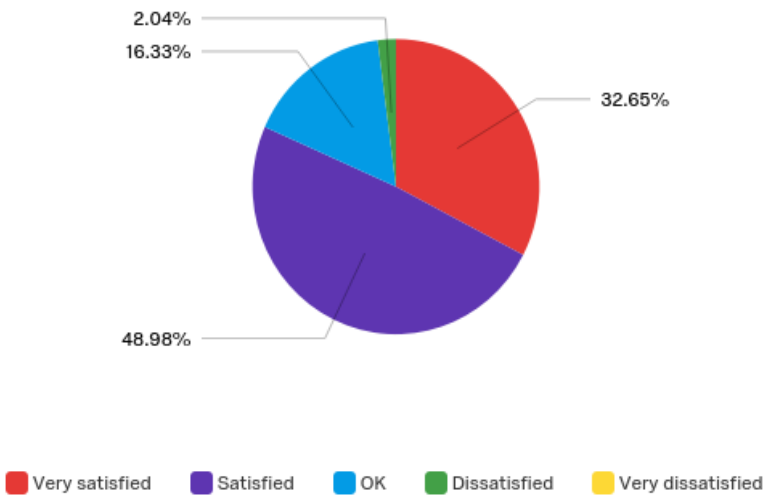

Figure 4.5 The satisfaction level among participants regarding the use of translation tools

Furthermore, participants were asked to rate their satisfaction level for the current computer-assisted translation tools. As shown in Figure 4.5, around 33% reported *high satisfaction* and 49% chose the *satisfied* option. Meanwhile, approximately, 16% were "*OK"* with the tools and only 2% were *dissatisfied*. Table 4.1 below demonstrates the distributions of participants' responses to both questions demonstrated in Figure (4.4 and 4.5).

Table 4.1 Distribution of Participants' Satisfaction Level

	Very satisfied	Satisfied	OK	Dissatisfied
Number of participants who reported no complications encountered	8	5	2	1
Number of participants who reported they encountered complications	6	18	6	0
Does not apply	2	1	0	0

Figures 4.6 below presents the level of satisfaction with the tools related to the participants' responses to the question whether they encountered any complications during their use of the translation tools for Arabic language.

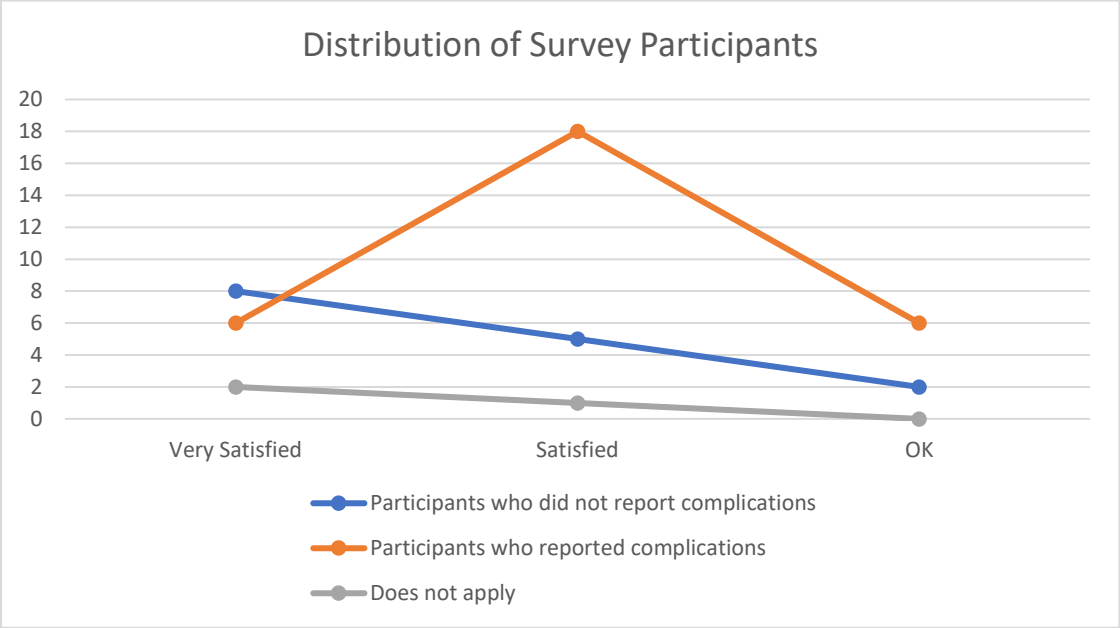

Figure 4.6 Distribution of participants' responses

Table 4.1 and Figure 4.6 shown above, indicate that participants who reported that they have encountered complications while using CAT tools for Arabic language likely

experienced less satisfaction than participants who reported having no complications while using the CAT tools. However, the chi-square independence test was adopted to determine whether there is a significant relationship between the level of satisfaction and the reported complications. The results show no significant relationship between the two variables as p =.18 is greater than .05. This nonsignificant result could be due the sample size, as only 49 participants responded to both questions, which could create some inconsistency in the data. However, it is possible that the current complications of translation tool use for Arabic language do not influence the satisfaction level of translators.

On a related question, participants were asked about the importance of using CAT tools by Arabic language translators. As can be seen in Figure 4.7 below, around 84% think it is very important. 16% think it is somewhat important, while zero participants undervalued the importance of using CAT tools for translation in Arabic.

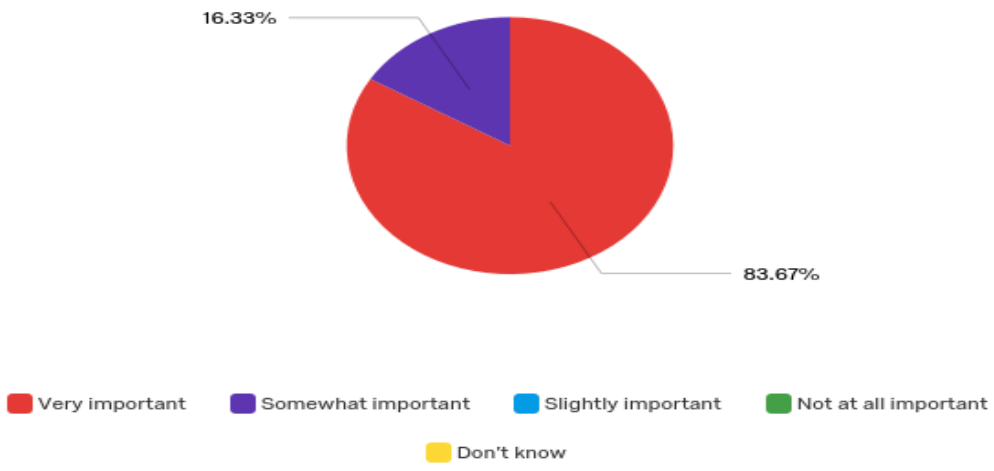

Figure 4.7 Participants' responses regarding the importance of translation tools use

In addition, participants were asked if they agree or disagree that integrated translation memory systems with MT suggestions when there is no TM for a given translation segment will increase Arabic translators' productivity. As shown in Figure 4.8 below, the majority of translators agreed to this statement while only around 10% of participants disagreed.

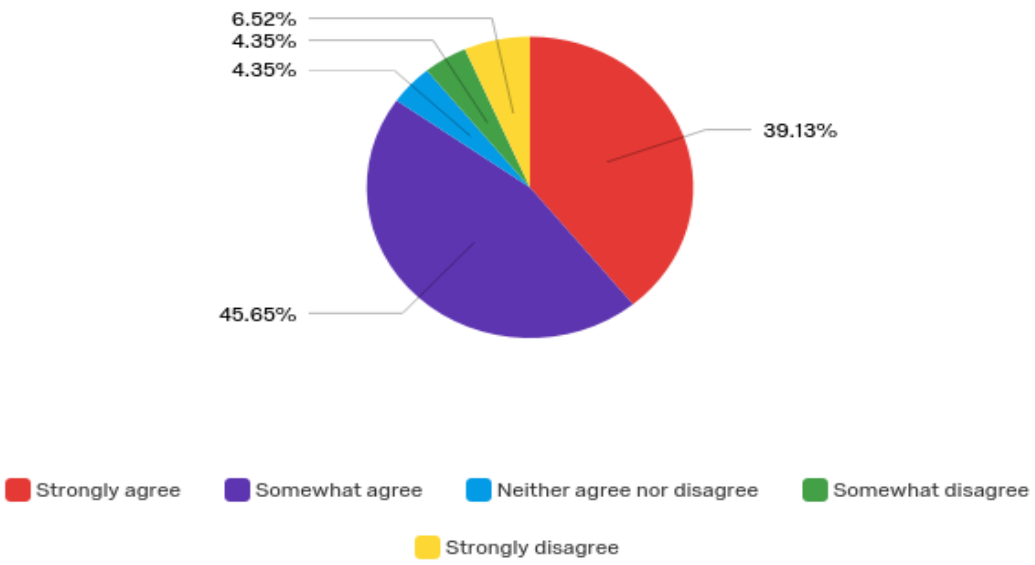

Figure 4.8 Participants' responses to a productivity increase with integrated CAT tools with MT

To sum up, the responses from survey participants in this section reveal the answer to the first research question on how Arabic language translators evaluate computer-assisted translation tools. As has been demonstrated above, despite the fact that the majority of participants encountered complications while using the tools for Arabic language, the

Arabic language translators encourage the use of computer-assisted translation tools. Therefore, results show that the majority of participants are satisfied with the current translation tools. However, there are considerable variations between responses regarding the level of importance given to CAT tools as compared to the importance of MT use among the survey participants. Although this finding cannot be generalized to the community of Arabic language translators, given the small sample of 49 participants, it suggests that Arabic language translators appreciate the use of CAT tools more than they do MT systems. Probably, this reflects the poor quality of MT tools, which still struggle to provide better outcomes for Arabic language translation. Nevertheless, the satisfaction level reported by participants does not eliminate the complications that Arabic language translators encounter while using CAT tools. As has been shown in the results above, more than sixty percent of participants stated they have complications while using the current translation tools for Arabic. These language-related complications will be discussed thoroughly in the following chapter.

4.2 Experiment Results

4.2.1 General Observations of the Experiment Results

All experiment participants completed their required tasks, which included translating two texts in each of two directions: from Arabic to English and vice versa using SDL Trados 2019 Studio, immediately followed by a short interview. Some participants asked if they could use an online search before they started their translation process. They were told they were free to use the online resources as needed with no restrictions. As has been

discussed in Chapter Three, participants were not given a translation brief, so some of their choices while translating may well reflect divergent intention with respect to implicit assumptions. However, quality output is not the concern of this study as it aims to investigate translators' views toward the tools, the complications they encounter and how the tools can be developed to meet their needs.

The mean time for participants to complete both translation tasks was 16.56 minutes (SD 6.75), with times ranging from 9.06 minutes to 30.25 minutes. The mean time for completing the Arabic-English translation task was 9.85 minutes (SD 4.99), with times ranging from 4.39 minutes to 19.29 minutes. Furthermore, the mean time for completing the English-Arabic translation task was 6.58 minutes (SD 2.26), with times ranging from 3.46 minutes to 11.08 minutes. Table 4.2 below reports some general information about the time spent for the translation tasks.

Table 4.2 General Overview of Experimental Tasks: Time (in minutes)

	Arabic-English Task	English-Arabic Task	Total Time for both Tasks
Mean	9.85	6.58	16.56
Std Dev	4.99	2.26	6.75
Min	4.39	3.46	9.06
Max	19.29	11.08	30.25
Median	8.44	6.07	14.51

Regarding which task was harder, 76.9% participants reported having more complications with the Arabic-English translation, while for 23.1% thought the English-Arabic task was harder. All respondents reported that the second segment of the of Arabic-English translation was the hardest to properly translate. On the other hand, the

second segment of the English-Arabic translation was the hardest to translate for 46.2% participants, the third segment was hardest for 46.2%, and the fourth one was hardest for 7.7% participants. The two Figures 4.9 and 4.10 below demonstrate the distribution of time spent by each participant on each task. We can see that the mean time spent for the Arabic-English translation task was more than the English-Arabic translation. Despite the fact that 23.1% of participants stated that they had more difficulties with the English-Arabic translation, most translators spent more time with the Arabic-English translation task except two participants, as can be seen in the Figures (4.9 and 4.10). Epsilon spent about the same time to complete each translation task while Lambda spent considerably more time on the English-Arabic translation task than the Arabic-English task.

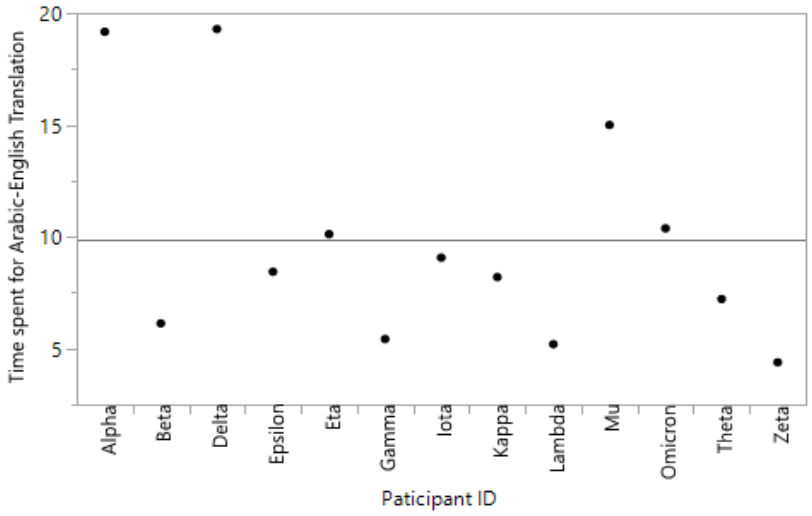

Figure 4.9 Time Spent for Arabic-English Translation Task

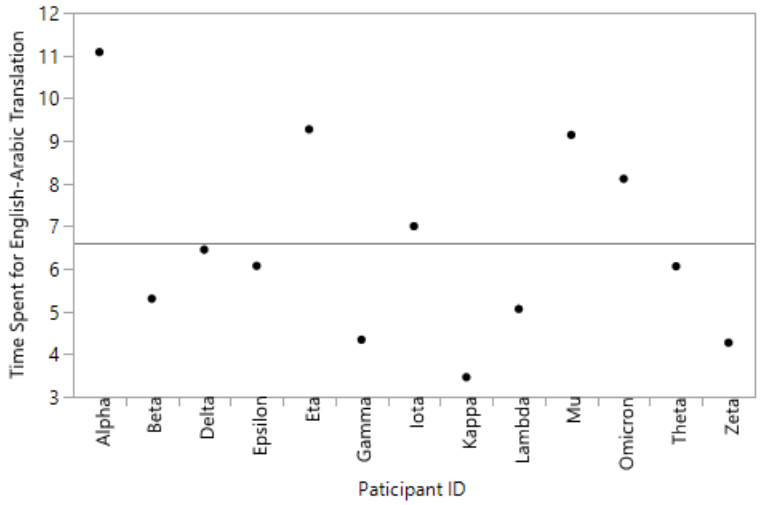

Figure 4.10 Time Spent for English-Arabic Translation Task

Before conducting the paired-samples of the t-test, the assumptions of normality and the absence of outliers were tested. According to the Shapiro-Wilk statistics, the distribution of the time spent on the Arabic-English task varied more than expected (p = .04), while the distribution of the time spent on the English-Arabic task was approximately normal (p = .64). Figures 4.9 and 4.10 demonstrate distributions of both variables. As can be seen in Figure 4.9, there are two outliers regarding the time spent on the Arabic-English task (participants Alpha and Delta), and no outliers in the time spent on the English-Arabic task. Due to the lack of normality and the presence of outliers in the variable of time spent on the Arabic-English task, the results from this analysis will be limited only to this sample and should not be generalized to the whole population of Arabic translators or to a range of texts or text types. It is worth mentioning that Alpha, who spent significantly more time than others, provided the best translation product. See section 4.2.2 for further details about participant responses to MT and TM suggestions.

In order to check if there is significant difference in the mean time spent between the Arabic-English and English-Arabic translation tasks by participants, a descriptive statistical test was conducted. Table 4.3 below shows the difference in mean time and standard variation between both translation tasks.

Table 4.3 Paired Samples Statistics (Time spent in both tasks)

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Time Spent in Arabic-English Task (min)	9.845	13	4.99914	1.38651
	Time Spent in English-Arabic Task (min)	6.5854	13	2.25723	.62604

From Table 4.3 above, the mean time for completing the Arabic-English task (9.849) is greater than the English-Arabic task (6.585) with a lower standard deviation. The correlation between the two variables is a significant high positive linear relationship of .714, as can be seen in Table 4.4.

Table 4.4 Paired Samples Correlations (Time spent in both tasks)

		N	Correlation	Sig.
Pair 1	Time Spent in Arabic-English Task (min) & Time Spent in English-Arabic Task (min)	13	.714	.006

According to the results of the paired-samples t test, there was a statistically significant difference in the time spent for Arabic-English (M = 9.85, SD = 4.999) and English-

Arabic (M = 6.59, SD = 2.257), t (12) = 3.15, p = .008. Participants spent significantly more time in the Arabic-English task than the English-Arabic task when using computer-assisted translation tools.

Table 4.5 Paired Samples T-Test (Time spent in both tasks)

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Time Spent in Arabic-English Task (min) - Time Spent in English-Arabic Task (min)	3.26385	3.73798	1.03673	1.00501	5.52268	3.148	12	.008

Therefore, it can be concluded that the mean time for completing the Arabic-English translation task with mean time 9.85 ± 5 is statistically different from the mean time spent to complete the English-Arabic translation task with mean time 6.59 ± 2.26. This would suggest that Arabic to English translation is harder for Arabic language translators who have Arabic as their mother tongue, as stated by the majority of participants.

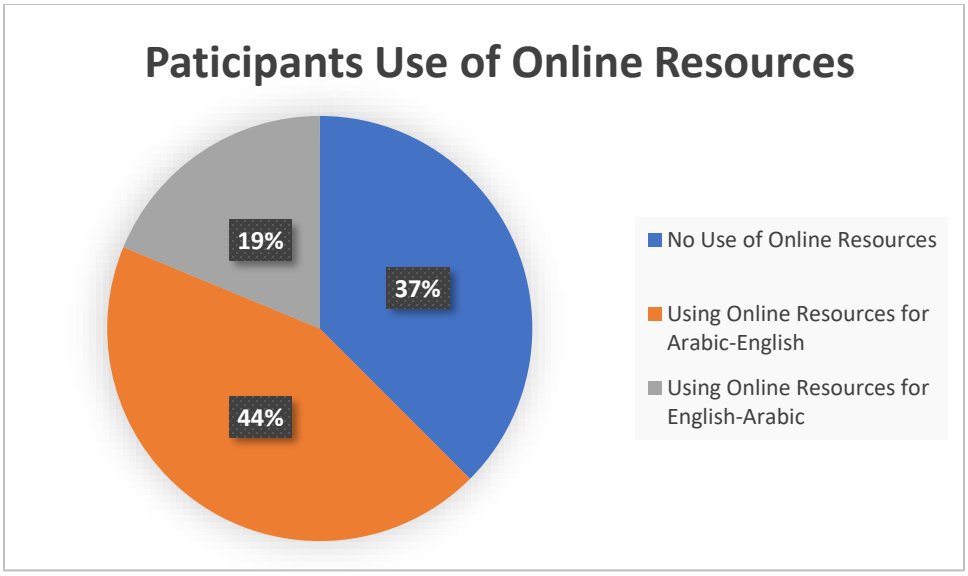

Figure 4.11 Participants Use of Online Resources

Reviewing the screen recordings to observe specific patterns and activities revealed that 7 participants used online resources (e.g., Google search engine, online dictionaries, etc.) for the Arabic to English translation task, while only 3 participants used online resources for the English to Arabic translation tasks. Thus, only 6 participants chose not to use other resources than the MT and TM provided to them for both tasks. The higher number of participants using additional online resources for the Arabic-English translation task could be attributed to the requirements of translation to participants' second language where they needed to validate the terms that they were not familiar with. However, there is no statistically significant difference between using online resources for the Arabic-English and English-Arabic translation tasks, or between participants who used the online resources and the participants who didn't. Figure 4.11 above reports general information of the participants' use of online resources during their required translation tasks.

Two sample t-tests show a significant relationship between time spent by participants who used online resources for the Arabic-English task (M =12.32, SD = 5.51) and those who did not use online resources (M = 6.69, SD = 2.20), t (8.1) = 2.4, p = 0.04. Participants who decided to use online resources spent significantly more time to complete the task than participants who did not. On the other hand, two sample t-test revealed insignificant results between time spent to complete the translation for participants who used online resources for the English-Arabic translation task (M =9.04, SD = 2.88) and those who did not use online resources (M = 6.13, SD = 1.97), t (1.36) = 1.17, p = 0.37. This insignificant result can be attributed to the time spent searching for online resources as they spent less time searching as compared to the first translation task from Arabic to English. However, participants tended to spend more time to complete the translation if they used online resources. This was statistically significant for the first translation task from Arabic to English and insignificant for the second translation task from English to Arabic.

4.2.2 Observations of Participants' Use of the CAT Tool

The participants' behavior during the experiment was captured by screen recorder and analyzed to determine if there are patterns in their use of CAT tools. Examining the translation texts produced by participants and screen recording data allowed to penetrate their activities in order to gain a better, more intimate understanding of how they respond to TM and MT suggestions and mostly to focus on how their behavior toward the tools

reflect on their evaluation and satisfaction of the translation tools. The data analysis aimed to examine participants' behavior regarding TM suggestions, MT suggestions and fuzzy match segments. The following sections demonstrate the responses of participants to the CAT tools.

4.2.2.1 Responses of Translators to TM Suggestions

Translators are highly likely to accept TM suggestions and most studies indicate that translators trust them (Bowker, 2005; Ford, 2016; LeBlanc, 2013). However, there are some participants in the study whose behavior in response to the TM suggestions was the opposite of what was expected. The total number of participants who made changes to the Arabic-English TM were 8 (61.5%) and 8 (61.5%) for the English-Arabic TM. However, four of those participants changed both Arabic-English and English-Arabic TMs. Some of these participants made massive changes and some others made minor changes to TM suggestions. The Table 4.6 below demonstrates samples of the changes made to TM.

Table 4.6 Samples of Terminology TM Changes

Participant ID	Example of change	TM suggestion
Alpha	Recognizing Signs of Oral Cavity Cancer	How to Recognize Signs of Oral cancer
Zeta	How to Recognize Signs of Mouth Cancer	How to Recognize Signs of Oral cancer
Eta	How to identify signs of mouth cancer	How to Recognize Signs of Oral cancer

Lambda	How to identify signs of mouth cancer	How to Recognize Signs of Oral cancer
Kappa	How to Identify Signs of Mouth Cancer	How to Recognize Signs of Oral cancer
Beta	Although your primary care physician and dentist are trained	Although your doctor and dentist are trained
Epsilon	Although your physician and dentist are trained to detect oral cancers	Although your doctor and dentist are trained

The explanation for the changes to Arabic-English TM can be attributed to MT suggestions. Most of these participants preferred 'mouth cancer,' which was suggested by MT, instead of 'oral cancer' that was provided by TM. This could be linked to isomorphism between Arabic and English as the original term in Arabic 'oral cancer' would be literally translated as 'mouth cancer'. This might have caused confusion to some participants. Some of those who made these changes and preferred the term of *mouth cancer* over *oral cancer* are ATA certified translators, so their level of translation experience might not be involved here. The same applies to the change of 'recognize' to 'identify' in TM. The participants preferred the term that has a direct equivalent in Arabic. Although both terms can mean the same in Arabic, *recognize* is more appropriate in this context. Again, isomorphism might have played a role in the decisions made by participants to change the TM suggestions in Arabic-English. Furthermore, there is one participant who changed 'doctor' into 'primary care physician'. Another participant changed it into 'physician'. Those changes reflect the urge of translators to address the English reader with the best terms they think are most appropriate to be used.

In addition, some participants made language-use changes to Arabic-English TMs. These changes were minor and did not influence the meaning of the target text. The Table 4.7 below demonstrates some examples of these changes.

Table 4.7 Sample of Language Use Changes to TM

Participant ID	Example of change	TM suggestion
Alpha	Although your regular doctor and dentist are trained to detect oral cavity cancers	Although your doctor and dentist are trained to detect oral cancers,
Delta	Although your dentist or doctor is trained to detect oral cancers	Although your doctor and dentist are trained
Delta	recognizing the signs yourself may help in an earlier diagnosis	recognizing the signs yourself may facilitate an earlier diagnosis
Zeta	recognizing the signs yourself may help in an earlier diagnosis and treatment	recognizing the signs yourself may facilitate an earlier diagnosis
Lambda	Although your doctor or dentist are trained	Although your doctor and dentist are trained
Mu	Although your doctor and dentist are trained to diagnose oral cancers	Although your doctor and dentist are trained to detect oral cancers
Omicron	recognizing the signs yourself may facilitate (deletion of an article) early diagnosis	recognizing the signs yourself may facilitate an earlier diagnosis

As shown above, some of these changes are grammatical or stylistic in nature. Some participants preferred ‘help in’ instead of ‘facilitate’. Two participants changed ‘and’ into ‘or’ in the TM. Another participant removed the article ‘an’ from ‘early diagnosis’. These changes could be linked to isomorphism between Arabic and English as these suggested changes by participants reflect the influence of their Arabic language as mother tongue.

On the other hand, the changes made to English-Arabic are stylistic preferences where some participants restructure the statements to fit better into the Arabic writing style. The changes in English-Arabic TM were extensive. The Table 4.8 below demonstrates two examples of these changes. For further examples of TM changes, see the translations of participants in (Appendix K).

Table 4.8 Examples of English-Arabic TM Changes

Participant ID	Proposed change in TM	English Representation of TM suggestion	English representation of the participant's translation
Alpha	يعاني الفرد من مرض فقد الشهية عندما يمتنع عن تناول الطعام والشراب اللازمين للحفاظ على وزن صحي	When an individual refuses to consume the amount of food and drink required to maintain a healthy body weight… that person suffers from anorexia.	An individual suffers from anorexia when he/she refuses to consume the amount of food and drink required to maintain a healthy body weight
Theta	عندما يكون للمريض صورة مشوهة عن جسمه، وخوف شديد من الزيادة في الوزن،	When an individual refuses to consume the amount of food	When a patient has a distorted body image, and an intense fear of

ويمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن الجسم الصحي	and drink required to maintain a healthy body weight… that person suffers from anorexia.	gaining weight and refuses to consume the amount of food and drink required to maintain a healthy body weight maintain a healthy body weight

The two examples in the table demonstrate the urge for some participants to change TM to better address Arabic language readers. These changes are stylistic with no terminological changes to the TM structure. Contrary to the changes in Arabic-English TM, the changes in English-Arabic TM were extensive and included major changes in structure. This resulted in a significantly longer time spent to complete the translation task as shown in the figure below.

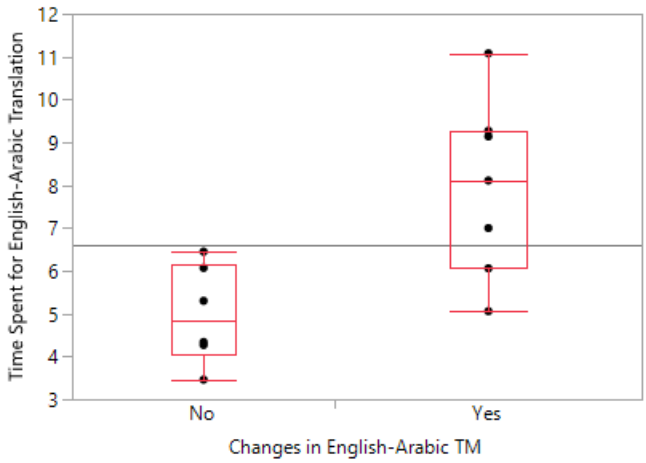

Figure 4.12 Changes in English-Arabic TM and Time spent

According to the results of the paired-samples t-test, there was a statistically significant difference in the time spent by participants who made changes to English-Arabic TM (M = 7.96, SD = 2.07) and who did not make changes to the TM (M = 4.98, SD = 1.155), t (9.24) = 3.25, p = .009. Participants who made changes to English-Arabic TM spent significantly more time to complete their translation task than participants who just accepted TM suggestions. On other hand, the paired-samples t test revealed an insignificant difference in time spent by participants who made changes to the Arabic-English TM (M = 10.42, SD = 6.34) and who just accepted TM suggestions without changes (M = 9.17, SD = 3.27), t (0.4) = 9.24, p = 0.65. This variation can be attributed to the level of changes to TM in each task as illustrated above. Extensive changes to the English-Arabic TM were made by participants, e.g. restructuring of sentences, as compared to the changes made to Arabic-English TM that involved terminology and minor stylistic changes.

The findings suggest that Arabic language translators are more likely to check TM and change suggestions. This does not necessarily contradict other findings in the literature that translators generally are more likely to accept TM suggestions without checking them for several reasons. First, the translation text was short and probably participants took their time to validate TM and change its suggestions. Second, participants decided to change TM suggestions without being instructed whether or not to change TM. At the same token, they haven't been told they will not get paid if they edit 100% match segments. The scenario is anticipated to be totally different in a real-world translation task where translators won't get paid for editing a 100% match segment. However, one of the online survey participants complained about the problems of some clients who are not willing to pay for 100% match segments while these segments still need to be edited. This might suggest that Arabic language translators are more likely to change TM suggestions, which are contrary to the anticipated results of this study.

Through the samples of translations performed by 13 participants, it can be seen that there are patterns of editing TM. These patterns range from changing terms into style preferences. The primary goal of the study is not focusing on quality of translation, since the conditions of providing high quality translation were not provided to translators. Therefore, evaluation of the translations for quality was not conducted, as the goal of the study is to examine the views of Arabic language translators toward TM and MT suggestions. Contrary to the anticipations of the study, 8 out of 13 (61.5%) of participants preferred to make changes to TM suggestions despite the accuracy of the TM content. This is probably not a significant percentage of participants. However, most of the

changes were made in 100% match segments in which it is unexpected to see this number of participants make those changes.

4.2.2.2 Responses of Translators toward Fuzzy Matches

Examining the responses of participants toward the fuzzy matches in Arabic-English revealed variation between participants. 7 participants out of 13 (54%) completed translation of all missed parts in the fuzzy segment. On the other hand, 9 participants (69%) addressed the second missed part of the TM but ignored the first part. 4 participants ignored completion of the fuzzy segment. Table 4.9 below demonstrates the translations of participants who addressed the missed parts in the 75% fuzzy match. Underlined words represent the missing parts of TM.

Table 4.9 Samples of Translating Missed Parts of a Segment in TM

Participant ID	Example of fuzzy match translation	TM suggestion
Alpha	recognizing the signs yourself may facilitate an earlier diagnosis and timely treatment. More awareness is for the better	recognizing the signs yourself may facilitate an earlier diagnosis
Delta	recognizing the signs yourself may help in an earlier diagnosis and treatment. The more awareness you possess the better	recognizing the signs yourself may facilitate an earlier diagnosis
Epsilon	recognizing the signs yourself may facilitate an earlier diagnosis and timely treatment. The more you are aware, the better it is for you	recognizing the signs yourself may facilitate an earlier diagnosis
Zeta	recognizing the signs yourself may help in an earlier diagnosis and treatment. and the more you	recognizing the signs yourself may facilitate an earlier diagnosis

	become aware the better	
Theta	recognizing the signs yourself may facilitate an earlier diagnosis. Also, medication at an early stage is helpful. The more consciousness you have the better you live	recognizing the signs yourself may facilitate an earlier diagnosis
Lambda	recognizing the signs yourself may facilitate an earlier diagnosis and cure in the appropriate time. The more conscious you are, the better your health will be	recognizing the signs yourself may facilitate an earlier diagnosis
Kappa	recognizing the signs yourself may facilitate an earlier diagnosis and timely treatment. The more you are aware, the better	recognizing the signs yourself may facilitate an earlier diagnosis

According to the results of the paired-samples t test, there was an insignificant difference in the time spent by participants who completed translation of fuzzy matches (M = 10.27, SD = 6.29) and who did not complete translation of fuzzy matches (M = 9.35, SD = 3.44), t (0.3) = 9.5, p = 0.74. Moreover, the paired-samples t test revealed insignificant difference between participants who addressed fuzzy matches (M = 9.82, SD = 5.6) and those who ignored fuzzy matches (M = 9.90, SD = 3.94), t (0.02) = 8.3, p = 0.97. It may be assumed that participants who addressed all ST content should spend significantly more time to complete the translation task. This was not the case here as revealed by the data and it can be attributed to variations in level of experience between participants. It is worth mentioning that participants who ignored the fuzzy match segment in the Arabic-English translation task have lower levels of experience in translation than the other participants who addressed the fuzzy match segment. Despite that, the results show no

significant impact on time spent to complete translation by addressing fuzzy matches. This finding cannot be generalized, however, due to the sample size and the difference in experience levels between participants.

On the other hand, all participants addressed and completed translation of the fuzzy matches in the English to Arabic translation task. All missed parts from TM were translated by the 13 participants. There's no explanation for this variation in performance between the two tasks. It could be attributed to the fact that translation into a native language is easier for translators when they have strong command of the language. Thus, they can spot the fuzzy match segments and address them easier than translation tasks into L2.

In total, this is an interesting finding as it is similar to the finding of (Ford, 2016) for the Arabic-English task and completely different for the English to Arabic task. Participants seemed to pay more attention to the fuzzy matches in their translation into Arabic more than they did the other way around. Ford (2016) found out that 76% of professional translators and 53% of translation student participants spotted the fuzzy matches in an English-Arabic translation task and translated them completely. In the findings of this experiment, only 54% spotted the fuzzy matches completely and 69% addressed them partially in Arabic-English. On the other hand, 100% of participants spotted the fuzzy matches in the English-Arabic task. The interesting part of this finding is that there are participants who did not attempt to address fuzzy matches in their Arabic-English task, but they did with their English-Arabic task.

4.2.2.3 Responses of Translators toward MT Suggestions

The behavior of participants in response to the MT suggestions ranged from extensive post-editing to minor post-editing. No participants accepted the suggestions of MT as they were without minor post-editing. 77% of participants made extensive post-editing to the MT suggestions in the Arabic-English translation task while 23% of participants made minor changes to the MT suggestions in the English-Arabic translation task. On other hand, 54% of participants made extensive changes to the MT suggestions in the English-Arabic translation task, while 46% of participants made minor changes. The lower number of participants who made extensive changes to MT suggestions in the English-Arabic translation can be attributed to the better quality of MT suggestions for into Arabic translation. Extensive post-editing had a significant impact on the time spent to complete translation tasks as shown in the Figure 4.13 below.

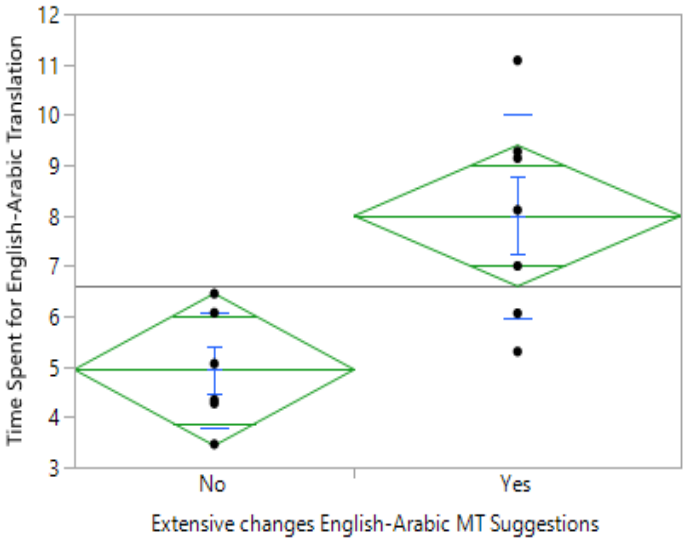

Figure 4.13 Extensive Changes in English-Arabic MT Suggestions and Time Spent

Two sample t-tests reveal significant difference in the time spent between participants who made extensive post-editing changes to the MT suggestions in the Arabic-English translation task (M = 11.3, SD = 4.8) and those who made minor changes (M = 5.06, SD = 0.55), t (9.7) = 4.05 = p, = 0.0024. Moreover, participants who made extensive post-editing changes to the MT suggestions in the English-Arabic translation spent significantly more time (M = 7.9, SD = 2.01) than participants who made minor changes (M = 4.94, SD = 1.14), t (11) = 3.3 = p, = 0.0075. In addition, the group of participants who extensively post-edited MT suggestions in both directions spent significantly more time to complete both tasks than participants who made only minor changes. According to the results of the paired-samples t test, there was a significant difference in the total time spent for both tasks by participants who extensively post-edited the MT suggestions (M =18.57, SD = 6.40) and participants who performed minor post-editing to the MT suggestions in both tasks (M = 9.83, SD = 0.66), t (4.2) = 9.6, p = 0.002.

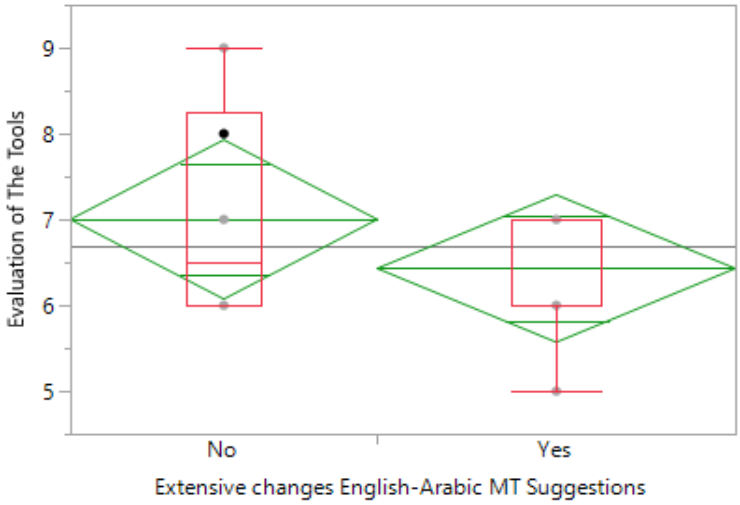

Figure 4.14 Extensive Changes English-Arabic MT Suggestions

The amount of work done in the translation process can have an impact on the evaluation of the tools. It can be seen from the figure above that participants who made extensive post-editing changes gave a lower evaluation for the tools. However, the paired-samples t-test revealed an insignificant relationship between the evaluation and extensive post-editing of the MT suggestions, $t(11) = 1$, $p = 0.3$. This insignificant result could be due to the small sample size of 13 participants. Despite that, there is no statistically significant relationship between the extensive effort of post-editing MT suggestions and evaluation of the tools.

4.2.2.4 Statistical Analysis of Translations Output

Transfer of ST structure into TL structure

The texts produced by participants were analyzed quantitively (average number of words per sentence, number of words and number of sentences) and compared to the MT and TM suggestions. This was conducted to examine if there are patterns of transferring source text (ST) structure into target language structure and how this may relate to the Arabic language translators' evaluation and views toward the CAT tools. MT suggestions tend to transfer ST into TL (Toral, 2019). In the perfect situation, it would be expected that the average number of words per sentence decreases, and the average number of words and sentences increases in the final output produced by participants for Arabic-English translation. The opposite is true for English-Arabic translation where the average number of words per sentence increases, and number of words and sentences decrease. These two opposite expectations are based on the two language's different morphosyntactic, punctuation and general structure rules.[10]

[10] . See Section 2.1 for further details about Arabic language structure.

Table 4.10 Demonstration of TM and MT Suggestion Statistical Information

	Arabic-English Translation			English-Arabic Translation		
	Average number of words per sentence	Number of words	Number of sentences	Average number of words per sentence	Number of words	Number of sentences
Source text	47	94	2	28	84	3
MT and TM suggestions	45	90	2	27.67	83	3

Table 4.11 Demonstration of Participants' Final Translation Output Statistical Information

	Arabic-English Translation			English-Arabic Translation		
Participant ID	Average number of words per sentence	Number of words	Number of sentences	Average number of words per sentence	Number of words	Number of sentences
Alpha	18.17	109	6	39.50	79	2
Beta	23.40	117	5	28	84	5
Gamma	26	104	4	29	87	3
Delta	23.40	117	5	29	87	3
Epsilon	23	115	5	27.67	83	3
Zeta	54	108	2	28	84	3
Theta	22.60	113	5	26.67	80	3
Eta	25.25	101	4	25.67	77	3
Iota	26	104	4	30.67	92	3
Lambda	38.67	116	3	28.67	86	3
Kappa	20.60	103	5	29.33	88	3
Mu	23.75	95	4	28.33	85	3
Omicron	20.80	104	5	30.67	92	3

Arabic language tends to have longer sentence structure. The source text of Arabic-English represented this complication with three sentences that consisted of 90+ words. MT tends to transfer source text structure with its suggestions. An analysis of

participants' response to these suggestions was conducted to examine their effort to avoid transferring ST structure into TL. A measure of the average of words per sentence was used to determine whether or not there is a relationship between extensive post-editing and minor post editing in the Arabic-English translation task. Figure 4.15 below demonstrates the average of words per sentence and the amount of post editing.

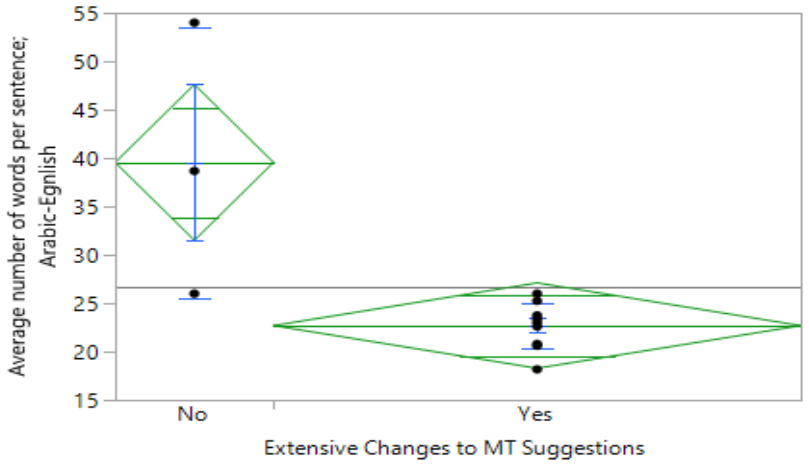

Figure 4.15 Extensive Changes to MT Suggestions and Average Number of Words per Sentence

Two sample t tests reveal a significant relationship between the groups of participants who did extensive post-editing and the average word numbers per sentence, $t(11) = 4.4$, $p = 0.0019$. Participants who performed extensive post-editing in the Arabic-English task have a significantly lower average of words per sentence in their TL.

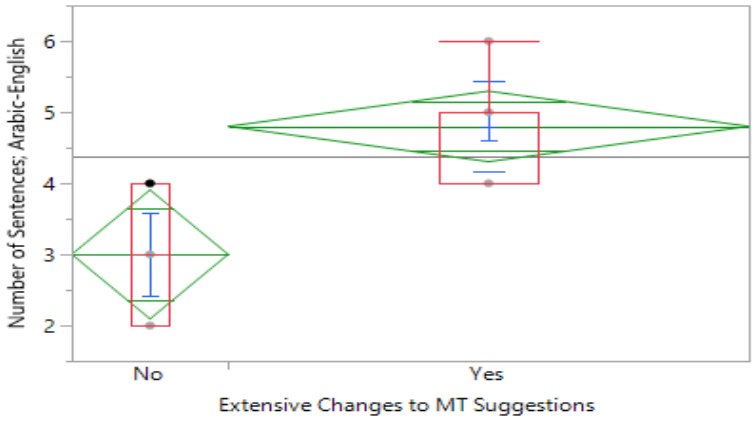

Figure 4.16 Extensive Changes to Arabic-English MT Suggestions and Number of Sentences

In addition, participants who made extensive post editing changes have significantly more sentences in their Arabic-English translation than the ones who made only minor changes as revealed by t test, t (11) = 3.8, p = 0.0028. On the other hand, participants seem to have similar translation outputs in terms of average number of words per sentence, number of words and number of sentences in their English-Arabic task. Both the participants who did extensive and minor changes have a similar average number of words per sentence. No significant change in English-Arabic average number of words per sentence can be due to the flexibility of the structure of the Arabic language, although strict punctuation of Arabic was not followed by most participants. On the other hand, most participants made significant changes in the punctuation of the Arabic-English translation task to target English readers. This explains the significant difference between MT suggested text and the final output in numbers of sentences and other statistical information as shown in Tables 4.10 and 4.11 above.

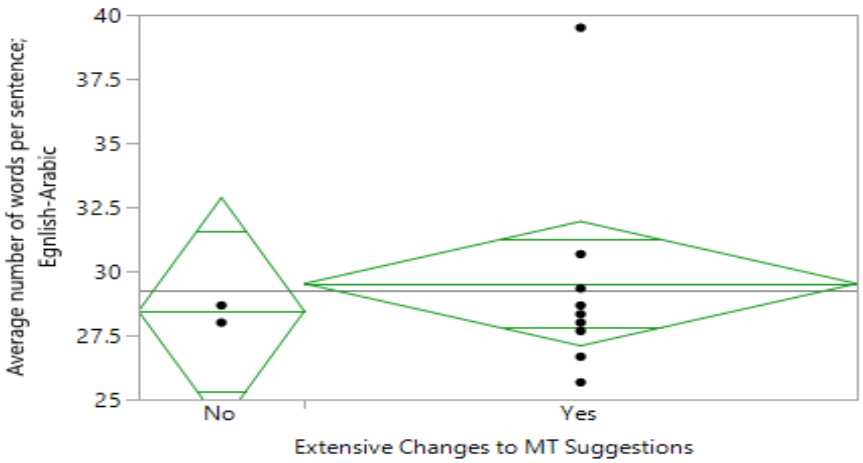

Figure 4.17 Extensive Changes to English-Arabic MT Suggestions and Average Number of Words per Sentence

As a result of that, there's no significant relationship between the amount of effort in post-editing and the structure of the target language. The reasons behind this variation between the two tasks need to be addressed in future research. It would be interesting to find out why only one participant avoided transferring ST into TL in the translation into Arabic language as shown in Figure 4.17 above. This participant had an average of 39.5 words per sentence, which was significantly different from other participants. This participant is an outlier, although he/she has produced the appropriate translation output as the Arabic language structure was addressed accurately. This participant also has more translation experience than any of the other participants.

4.2.2.5 Summary

Several scholars suggest that less experienced translators rely on TM suggestions (Bowker, 2005; Ford, 2016; LeBlanc, 2013). The behavior of the experiment participants toward the TM suggestions reflect this fact in the study. More experienced translators

made changes to both MT and TM suggestions while less experienced translators accepted all TM suggestions and made minor changes to MT suggestions. Those who changed the TM and MT suggestions in the English to Arabic task generally improved it, e.g. by using better wording, flow and clarity to target Arabic readers. However, this was not the case with the Arabic to English task as some translators' decisions to change TM suggestions were appropriate and others were perhaps unwise. This can be attributed to the complications of translating into L2. The isomorphisms between Arabic and English could have a strong impact on the decisions that translators made related to changing Arabic-English TM suggestions. Despite those variations between the outcomes of both tasks, high quality translation is not the primary concern of this study for several reasons. First, participants were not provided with proper conditions to produce high quality translation as they were not given a term base or a translation brief about the topic they were translating.

Second, all participants were native speakers of Arabic, so their translations into English are subject to questions from the outset. Despite the need of job markets in the Arab world that force translators to translate into their L2, translation studies scholars and researchers are known for their strong opposition to translations into L2. Peter Newmark's outspokenness on the problem cannot simply be overlooked. In his *Approaches to Translation*, Newmark speaks against the *"unnaturalness"* of L2 translations. He also warns translators not to engage in L2 translation; because if they do so, they will be "*caught"* every time as if it was a predetermined fact that no matter how

non-native translators try, their translations are unequivocally destined to have errors (1981: 180). [11]

Third, the translation texts were too short, and the participants did not have the time to thoroughly research the topics in order to guarantee appropriate terminology or background knowledge, so they were to an extent just guessing at appropriate usage. Standard metrics like the ATA Framework are designed for use with texts of at least 250 words+. The relative "accuracy" of any metric increases in proportion to the number of words and the relative familiarity of the translator with the text topic

Finally, quality and evaluation of translation texts are out of the scope of the research questions. As a result of raising these concerns, analytical evaluation of the given quality of the performed translation texts was excluded. The study was more concerned with the responses and evaluation of Arabic language translators toward the TM and MT suggestions while they use CAT tools. It can be seen from the participants' attempts to make changes in TM and MT suggestions that translators are not likely to accept suggestions from TM and MT while they are using translation tools. In the following chapter, the views of participants toward the translation tools and their evaluation will be discussed in detail and will be correlated with their responses toward the MT and TM suggestions.[12]

[11] For further discussions about translation into L2, see (Beeby Lonsdale, 2001; Campbell, 1998, 2000).
[12] See Section 5.2 and 6.1 for further discussions about participants evaluation of the translation tools.

4.2.3 Quantitative Results of the Experiment

It is assumed that several factors influenced the participants' experience with the CAT tools, based on the dataset used for this study. This section will provide a statistical analysis of the quantitative data of the experiment. The quantitative dataset comprises secondary data with discrete and continuous variables. As discussed in Chapter Three, the research study has four variables: 1) *segmentation*; 2) *isomorphism between English and Arabic source and target texts*; 3) *tools evaluation*; 4) *level of complexity*. The first two variables, segmentation and isomorphism between English and Arabic source and target texts have been treated as the independent variables and then manipulated to study their effects on the dependent variables: tools evaluation and complications encountered by participants while using the tools. The time spent on the translation process was used as a measurement of the dependent variable level of complexity that participants encountered. Time spent is more complicated but after looking at participants translation output and examining screen recording, it can be found that variations between participants with regard to time spent represent complications. For instance, participants who decided to change TM suggestions have created their own complexities which resulted in more significant time spent than participants who accepted TM suggestions. Thus, time spent can represent the level of complexities encountered or at least perceived by participants although these complexities differs from one participant to another. Participants' behavior during the translation process has been discussed in the previous section and will be correlated with following chapter.

For easier referencing, the hypotheses for the study (as posed in Chapter One) will be restated below:

H1a: Arabic language translators will express concerns regarding language specific complications of segmentation when using translation tools.

H1b: Arabic language translators will express concerns regarding language specific complications of punctuation when using translation tools.

H1c: Arabic language translators will express concerns regarding language specific script-related complications when using translation tools.

H2a: Arabic language translators will express more concerns regarding language specific complications of segmentation when using translating tools for the Arabic to English translation task than vice versa.

H2b: Arabic language translators will express more concerns regarding language specific complications of punctuation when using translation tools for the English to Arabic translation task than vice versa.

H2c: Arabic language translators will express more concerns regarding language specific script-related problems when using translation tools for the English to Arabic translation task than vice versa.

In order to answer the first two research questions posed in Chapter One and to test the hypotheses, various statistical methods were adopted. The quantitative analytical section of the experiment has been divided into two sections. The first section demonstrates the

results of the study's tested hypotheses. The second section illustrates the statistical analysis conducted to test relationships between the independent and dependent variables. The second section includes an exploratory data analysis and inferential statistics. The exploratory analysis section tests the assumptions of normality, including missing values and outliers in the datasets, which invariably indicate the most suitable method to use for the inferential data analysis.

4.2.3.1 Descriptive Analyses of the Hypotheses

The overall evaluation of the computer-assisted translation tools, measured on the 10-point scale, was rather positive, M = 6.69, SD = 1.032, Min. = 5.00, Max. = 9.00, *N* = 13. None of the respondents were dissatisfied with the tools. This finding supports the results of the online survey which suggested high satisfaction level among Arabic language translators regarding the translation tools. In addition, all participants in the experiment reported that the computer-assisted translation tools help increase productivity. Again, this finding supports the results of online survey responses regarding the importance of CAT tools for increasing productivy.

Nevertheless, participants have reported some language-related complications while using CAT tools for Arabic. Table 4.12 below presents percentages of participants who reported having segmentation, punctuation and script-related problems when translating from Arabic to English and from English to Arabic using the translation tools.

Table 4.12 Frequencies and percentages of participants who reported about the Complications with segmentation, punctuation and spelling (script-related) by using computer-assisted translation tools

	Translation task			
	Arabic-English		*English-Arabic*	
Complications	Frequency	Percent (%)	Frequency	Percent (%)
Segmentation	9	69.2	13	0.0
Punctuation	8	61.5	9	69.2
Script-related	13	0.0	11	84.6

N = 13.

The first hypothesis (H1a, H1b and H1c) was evaluated by performing one-sample chi-square test for all six combinations of the specific complications (segmentation, punctuation and spelling) and the translation tasks (Arabic-English and English-Arabic). It was tested whether the number of participants who reported experiencing a specific complication (e.g. segmentation, punctuation and spelling) and those who did not report it significantly differed, with assuming equal probabilities.

The second hypothesis (H2a, H2b and H2c) was evaluated performing the McNemar test for each pair of complication variables, i.e. both segmentation variables (Arabic-English and English-Arabic), both punctuation variables and both spelling variables. The McNemar test is "used to compare two paired samples when the data are nominal and dichotomous" (McCrum-Gardner, 2008, p. 40). The variables tested in this statistical analysis are dichotomous (nominal) that involve agreement level between participants (e.g. yes and no). The paired sample t-test would be appropriate with continuous variables, but since the tested variable are nominal (yes or no) to specific complications

that occurred during the use of translation tools, the McNemar test was appropriate as non-parametric alternative.

4.2.3.2 Language-specific Complications when Using Computer-assisted Translation Tools

This section presents the results for the first hypothesis (H1A, H1B and H1C). Since 9 participants out of 13 reported segmentation complications while translating from Arabic to English, one-sample chi-square test was adopted. The results of the test revealed no significant difference in the proportion of participants who experienced segmentation complications in the Arabic-English task (69.2%) and those who did not (30.8%), χ^2 (1) = 1.92, p = .17. On the other hand, in the English-Arabic translation task, none of the participants experienced segmentation complications (0.0%), and therefore the chi-square statistic test was not computed. Therefore, *hypothesis 1a* was not supported by the data for any of the translation tasks. This indicates that there is not sufficient evidence to conclude that Arabic language translators have significant segmentation complicaitons with CAT tools. This finding could be due to the sample size. Also, examining screen recordings revealed that there were participants who didn't address the punctuation of English while translating from Arabic. This aspect is not strictly speaking a segmentation problem, but manupluation of the data over time might result in segmentation complication, as punctuation and segementation are related complications.

In addition, there was no significant difference in the percentage of participants who experienced punctuation problems in the Arabic-English task (61.5%) and those who did not (38.5%), χ^2 (1) = .69, p = .40. There was also no difference in punctuation problems

in the English-Arabic task between those who did (69.2%) and those who did not experience them (30.8%), χ^2 (1) = 1.92, p = .17. *Hypothesis 1b* was also not supported by the data for any of the two tasks.

Furthermore, in the Arabic-English task none of the participants experienced script-related problems (0.0%), and hence the chi-square test was not conducted. However, in the English-Arabic task there was a statistically significant difference at 1.0% level of significance between proportions of those who experienced script-related problems (84.6%) and those who did not (15.4%), χ^2 (1) = 6.23, p = .01. As presented in Table 4.6, more participants reported having script-related problems than the proportion of those who did not. *Hypothesis 1c* was not supported by the data for the Arabic-English task, but was supported for the English-Arabic translation task. This indicates that Arabic language translators face significant spelling complications when they are using CAT tools for English to Arabic translation. Examining screen recording reveals that majority of participants attempted to correct spelling output of MT. However, there are some participants who ignored these complications but when they were asked about it in the interview, they revealed they had problems with spelling.

4.2.3.3 The Difference in Experiencing Language-specific Complications between Arabic-English and English-Arabic Translation when Using Computer-assisted Translation Tools

This section presents the statistical results for the second hypothesis (H2A, H2B and H2C). Results of the McNemar non-parametric test revealed a statistically significant association between the Arabic-English and English-Arabic translation tasks regarding

reported segmentation complications, p = .004. A significantly larger percentage of participants reported segmentation complications in the Arabic-English translation task (69.2%) than during the English-Arabic task (0.0%). *Hypothesis 2a* was supported by the data. These findings indicate that Arabic language translators report significantly more segmentation complications in Arabic to English translation than English to Arabic.

However, there was no statistically significant association between the two tasks regarding reported punctuation problems, p = 1.00. Therefore, statistically, there was an equal percentage of participants who reported punctuation problems in both the Arabic-English (61.5%) and English-Arabic task (69.2%). Hence, *Hypothesis 2b* was not supported by the data. This indicates that Arabic language translators report equal punctuation problems regarding both translation directions.

Finally, there was a statistically significant difference between the two tasks regarding script-related problems, p = .001. There was a significantly larger percentage of participants who reported script-related problems in the English-Arabic translation task (84.6%) than the Arabic-English task (0.0%). *Hypothesis 2c* was supported by the data. The results of the significant hypothesis 1C and 2C indicate that Arabic language translators encounter spelling problems in English to Arabic translation.

4.2.3.4 Descriptive Analysis of Variables

4.2.3.4.1 Effects of Language-specific Complications when Using Computer-assisted Translation Tools on the Overall Evaluation of the Tools

Before performing multiple regression analysis, the assumptions of multicollinearity/singularity, outliers, normality, linearity, homogeneity of variances and independence of residuals were evaluated. The *segmentation* variable for English-Arabic and *spelling* for Arabic-English translation were excluded from the analysis, as they were constants. No important violation of the assumptions was detected, except for the sample size of $N = 13$. Due to the small sample size, the results are limited only to this particular sample and should not be generalized.

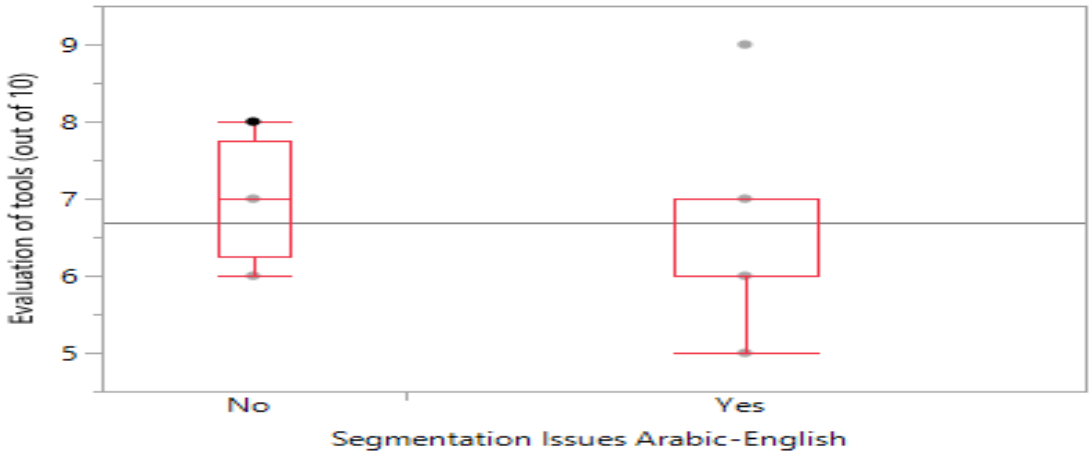

Figure 4.18 Participants' evaluation of the tools related to reporting of segmentation complication

However, there was an observation of one outlier in segmentation complications for the first task. One of the participants, as shown in Figure 4.18, is considered as an outlier for the score of tools evaluation, measured on the 10-point scale. This means that this

participant, who has given a rating for the tool evaluation as 9 out of 10, is considerably different from the others in that group who gave a lower evaluation score. Also, it was depicted that those who reported no complications with segmentation in the Arabic-English translation task have scored the tools evaluation higher as compared to the other group.

However, the results for the multiple linear regression model were non-significant, F (4, 8) = .14, p = .96. Therefore, segmentation, punctuation, and spelling complications for Arabic-English and English-Arabic translation did not affect evaluation of the translation tools. This indicates that the relationship between the independent variables (segmentation, punctuation, and spelling complications*) and the dependent variable* (evaluation of the tools) was not supported by the data (Table 13, $p > .05$). Thus, there is no evidence indicating that the complications of the tools for Arabic language may affect evaluation of the tools. This is similar to the findings of the online survey, where there was no signficant relationship between the complications of the tools and the evalution rates.

Table 4.13 Unstandardized and standardized regression coefficients of predictors of overall evaluation of the tools

Predictor	Unstandardized Coefficients		Standard-ized Coefficients	*t*	*p*	95.0% Confidence Interval for *B*	
	B	Std. Error	β			Lower Bound	Upper Bound
(Constant)	7.175	.943		7.606	.000	4.999	9.350
Segmentation complications Ar-En	-.349	.763	-.163	-.457	.659	-2.109	1.411
Punctuation problems Ar-Eng	.063	.851	.031	.075	.942	-1.899	2.026
Punctuation problems En-Ar	-.365	1.004	-.170	-.364	.726	-2.680	1.950
Spelling En-Ar	-.032	1.292	-.012	-.025	.981	-3.011	2.948

4.2.3.4.2 Effects of Language-specific Complications by Arabic-English Translation when Using Computer-assisted Translation Tools on the Time Spent

None of the assumptions of multicollinearity/singularity, outliers, normality, linearity, homogeneity of variances and independence of residuals was significantly violated, except for the aforementioned limitation of the sample size. The *spelling* variable was excluded, as it was a constant.

The linear regression model was non-significant, F (2, 10) = .08, p = .92. Therefore, the existence of segmentation, punctuation, and spelling problems in the context of Arabic-English translation did not affect the time spent on translation (Table 14, $p > .05$), and hence *Hypothesis 5a* was not supported by data.

Table 4.14 Unstandardized and standardized regression coefficients of predictors of the time spent on translation by Arabic-English translation

Predictor	Unstandardized Coefficients		Standardized Coefficients	*t*	*p*	95.0% Confidence Interval for *B*	
	B	Std. Error	β			Lower Bound	Upper Bound
(Constant)	10.570	3.136		3.370	.007	3.582	17.558
Segmentation complications Ar-En	-1.330	3.306	-.128	-.402	.696	-8.696	6.036
Punctuation problems Ar-En	.325	3.136	.033	.104	.920	-6.663	7.313

These insignificant results can be explained by the fact that participants who reported segmentation complications in this task spent less time than participants who didn't report them as shown in the table below:

Table 4.15 Group Statistics (Time spent vs. Segmentation Complications in Arabic-English task)

	Segmentation complications in the Arabic-English Task	N	Mean	Std. Deviation	Std. Error Mean
Total Time Spent in Both Tasks (min)	Yes	9	15.6489	6.12509	2.04170
	No	4	18.6025	8.61453	4.30726

Consquently, it cannot be assumed that complications occuring during the translation process have a significant influence on the time spent or the level of complexity. Again,

this finding can be due the small sample size. As mentioned earleir, time spent was a problematic and more complicated since there were other factors have impacted the variation between participants (e.g. TM changes, extensive MT post-editing vs minor post-editing). This could have caused inconsistancy in the data.

4.2.3.4.3 Effects of Language-specific Complications by English-Arabic Translation when Using Computer-assisted Translation Tools on the Time Spent

The *segmentation* variable was excluded as a constant. The assumptions of multicollinearity/singularity, outliers, normality, linearity, homogeneity of variances and independence of residuals were not signifcantly violated, however there was the sample size limitation.

The multiple regression model was not significant, F (2, 10) = 2.23, p = .16. The existence of punctuation and spelling problems by English-Arabic translation also did not affect the time spent on translation (Table 16, p > .05). *Hypothesis 5b* was not supported.

Table 4.16 Unstandardized and standardized regression coefficients of predictors of the time spent by English-Arabic translation

Predictor	Unstandardized Coefficients		Standardized Coefficients	*t*	*p*	95.0% Confidence Interval for *B*	
	B	Std. Error	β			Lower Bound	Upper Bound
(Constant)	9.040	1.454		6.217	.000	5.800	12.280
Punctuation complications En-Ar	1.679	1.607	.357	1.045	.321	-1.902	5.261
Spelling En-Ar	-4.275	2.056	-.711	-2.079	.064	-8.857	.307

4.2.3.5 Summary of the Quantative Analysis for the Experiment

In summary, the results suggest that there are no relationships between the complications that occur while using CAT tools for Arabic language and the evaluation rates. Also, there are no relationships between these complications and the level of complexity, or the time spent on processing translation in both directions. However, results of the McNemar tests suggest that participants do encounter complications when they are using translation tools for Arabic language. These complications will be demonstrated within the qualitative analysis in the following chapter.

CHAPTER V

RESULTS (II): Qualitative Data

5.1 Qualitative Analysis

This Chapter will explore the findings of the study to answer the research questions posed in Chapter One. This will include findings of how Arabic language translators evaluate the use of CAT tools, the problems and complications they encounter, and what improvements need to be addressed so the translation tools can better meet their needs and expectations. As the quantitative analysis in Chapter Four revealed some insights into the research questions, the qualitative data will explore the findings of the quantitative data in depth. As Göpferich (2008) emphasizes "qualitative analysis … provides insights that cannot be gleaned through quantitative procedures" (as quoted in Kiraly, 2013, 204).

The results of qualitative data for both the online survey and experiment will be combined in one section to avoid unnecessary redundancies, as both have raised some mutual complication regarding the use of CAT tools for Arabic language. There were several concerns raised by participants when they were asked about the complications

they encounter while using computer assisted translation tools that will be demonstrated in this chapter.

This chapter is divided into three major parts. The first part addresses the participants' views regarding the use of translation tools. The second part focuses on the complications of CAT tools for Arabic language translation. The final section addresses the suggestions of some participants toward developing the CAT tools, to meet the translators' needs. In all parts of this chapter, direct quotations were frequently incorporated to clarify the participants' viewpoints and emphasize their voices.

5.2 The Views and Evaluation of Arabic Language Translators about the Efficiency of Translation Tools:

The hybrid CAT tool is designed to reduce the amount of work that translators have to do. It provides them with MT suggestions in addition to translation memory suggestions. Translators still need to work on post editing, but the amount of work done by the tools reduces the effort and time spent on these tasks. Consequently, translators in the experiment have expressed their appreciation of the efficiency the tools provide to them. The benefits of CAT tools as conveyed by participants are as follows: time saving, solutions and dynamics, MT suggestions with terminology concerns, and the software's fast performance.

5.2.1 Time Saving

Time is valuable for translators, particularly freelancers, as increased efficiency gives them the chance to make more money when they can increase their productivity. The translation tools provided this benefit to translators because they help during the translation process as well as in managing translation projects and processing textual contents significantly faster than without using the tools. This appreciation was expressed by several participants as will be presented below. Beta describes his experience with the time savings when he was asked if the translation tools helped to increase his productivity:

> Definitely that... I finished both tasks in about 15 minutes, I would say. It would have taken me a much longer time to look up terms. At least, when I had to fix them now it was only a few gender issues, conjugations. Just like restructuring the sentence... But if I started from scratch, it would have taken me longer. So, yeah, definitely that they... The tools helped. (Beta)

Theta adds by describing her experience using the tools as compared to her previous translation experience "Yeah, it helped because if I compare what I did to what I used to do before, I think it saves time and effort". Another interview participant, Lambda describes the efficiency of text processing and the way translation memory works: "It saved me a lot of time. That way, I can see the next sentences, if anything is repeated, so I don't have to re-translate it, it will be saved into my TM."

To sum up participants' views regarding time concerns, participants seem to understand the importance of time saving with the use of translation tools. Rather than translating

everything from scratch, all translators need to do is just post-edit which would definitely result in less time than translating the whole texts from scratch without using the tools' help.

5.2.2 Solutions and Dynamics

In addition to the time saving, CAT tools provide translators with features that reduce the needed efforts with translation projects significantly. First, translators don't have to type all translated segments, as previously stored translations are saved automatically and when there's a match for a segment, it will appear in that segment section with the matching percentage. Participants expressed appreciation of the efficiency that tools provide to them as Alpha states: "it provides you with solutions to the translation problems that appear in front of you. So, you have solutions, and then something else you don't have to type so much".

Another important feature that CAT tools provide to their users is the formatting of the target texts. The translation tools will do the formatting, as it copies the formatting of source texts, which saves the energy and time consumed to correct these style problems. Participants emphasized the importance of this feature as expressed by Alpha; "The most important thing is that when I want to produce the target document, I don't have to worry much about the formatting problems, so that helps a lot, and saves a lot of time." (Alpha)

Moreover, the way source and target texts are presented in the tools gives the ability to translators to go through both texts without the need to divide the screen. It is already divided and presented to translators in an easy and helpful way. Some participants

revealed their gratitude for this feature as stated by Lambda: “The software itself, it's very useful and especially with the dynamics that are used in the program, that it’s right to left, so you can really focus on one segment rather than just looking at the whole text.”

5.2.3 SDL Trados Translation Tools Performance

The speed of translation tools and troubleshooting complications are extremely important for translators since it saves time and effort to deal with technical problems. SDL Trados 2019 impressed participants with the speed and dynamics of the software that eliminated most technical complications and troubleshooting during the translation process. One interview participant, Zeta, expressed his appreciation of the tool’ speed in the following quotation “It was efficient. It was super-fast. The translation memory did pretty much everything, I just fixed some sentences.” This statement goes along with findings of a comparative experimental study that examined the performance and quality reliability between MultiTrans and SDL Trados by (Moujaes, 2016). The author concluded that SDL Trados Studio is much faster in performance than MuliTrans when used for Arabic language.

5.2.4 MT and Terminology

On top of these factors, the addition of MT adds a valuable feature for translators. It provides them with a suggested translation and all they need to do is to post edit the

provided translation of the segment as needed. This saves time looking in dictionaries for unfamiliar terms. Several participants appreciated the MT suggestions provided to them with meaning of terms that they were unfamiliar with. Delta describes the addition of the MT feature as a time saver in his statement below;

> I would say automated translation is a bonus, and the CAT tool, it saves time. Sometimes I use Google Translate. I might sometimes just dump a whole sentence or paragraph and then paste it in the tool. And it's a time saver here.

Kappa as well gives credit to MT suggestions for introducing him to the translation of some terms that he was not familiar with as can be seen in his statement below:

> There were several terms that I didn’'t know, but I didn't need to look them up online, because the MT provided me with the Arabic translation. So, I was focusing on the... Just the translation, making it better. So, having the terminology ready, I think made it easier.

As can be revealed from these quotations, participants view the integration of machine translation tool with translation memory system as a helpful feature, especially with the suggestions it provides for scientific terms that translators may know nothing about which results in easier terminology management during the translation process. However, those participants, who embraced the MT suggestions for the terms that they are not aware about, did not take the time to cross-check those suggested terms. Those participants have accepted ‘mouth cancer’ instead of TM suggestion ‘oral cancer’ without

looking up the more appropriate term. As Kappa stated above ‘There were several terms that I didn’t know, but I didn't need to look them up online”, he didn’t try to look up for those unfamiliar terms.

Acceptance of suggested terms can be a negative feature if they aren’t always correct which was the case for this situation. ‘Mouth cancer’ is not the appropriate term for this translation. Despite that, there are seven participants who looked up and search to validate the suggested terms from MT in their Arabic-English translation.

5.2.5 Understanding Limitations

Moreover, some participants who volunteered in the experiment have demonstrated an understanding of the limitations of the tools. For instance, one interview participant, Alpha, reports below his expectation of the MT output as he thinks he will need to post-edit the suggested translation for it to be readable:

> Of course, you have to understand that this is a kind of translation that conveys the meaning to you, but it’s not... It doesn't look like... For example, when I translate from Arabic to English or English to Arabic, sometimes the translation doesn’t look like... It looks like a translation. So, what I try is to make it... So, this is a limitation I understand because this is a machine!

The same interview participant, Alpha, claims that he faced no complications while using CAT tools during the experiment although he stated later on in the interview that he faced minor problems as he states:

> I didn’t face any difficulties because I’m accustomed and used to using Trados as a translation tool, but I can envision that others who are not accustomed to using this tool might face issues in how to deal with the various functions….
>
> Yeah, I had some issues, minor issues. Like S for R or R for S, plural and singular, that's all about.

On the same side, Beta expresses below his understanding of the tool’s limitation due the difficulties of Arabic language differences from English language when he describes his experience with CAT tools:

> I don't think I changed my perspective of the tools because I have been working with the tools and I know that they are beneficial but they... And they have their limitations, especially with language pairs that are like Arabic and English. This is a difficult language pair, and I understand that, if there are real limitations between, let’s say in Spanish and English, there are much, much, much more limitations between Arabic and English because the nature of the language is the European... Like Arabic is a Semitic language. English is a Germanic language, like two different families.

Another interview participant, Epsilon, reported his expectation of the tool that he cannot depend on without the need of post-editing work to be done on the translation text as he states: “This semester I’m also working on Trados. It is just the same. I mean, we cannot rely on Trados. We have to post-edit the texts.” All these points raised by those participants show an understanding of the limitations. This tolerance toward the

complications might suggest that those translators are already familiar with the complications, whether they are solved or not. It is worth mentioning that an understanding of the limitations cannot be associated with or linked to the level of experience or knowledge of the tools, because participants who revealed understanding the limitations of the translation tool have various experience levels with the tools. Some of them have longer years of experience up to ten years while other participants had one year of experience. However, this indication of the results does not contradict the findings (Alotaibi, 2014) "that the more the student translators became familiar with CAT tools, realizing their reasonable potentials and current limitations, the less anxious they were" (2014, p. 65). In fact, the finding of the results could support the idea that trainee translators would show an understanding of the translation tool limitations after they become more familiar with them since all the experiments have the minimum requirement of familiarity of the translation tools.

5.2.6 The Correlation between the Participants' Behavior during Translation Tasks and their Views toward CAT Tools

Examining the positive views mentioned above and comparing them to the participants' behavior during translation tasks reveal that there is no relationship between participants' behaviors and their views toward the translation tools. The Table 5.1 below demonstrates their behavior and their views to the translation tools.

Table 5.1 Demonstration of participants' behavior and their views toward translation tools

Participant ID	Behavior toward translation tools	Views toward translation tools	Quotes from participants
Alpha	**Arabic-English task:** changes in TM, addressing all fuzzy matches, extensive post-editing MT suggestions. **English-Arabic task:** changes in TM, extensive post-editing MT suggestions.	Positive, understanding of limitations, evaluation rate 7 out 10.	You have to understand that this is a kind of translation that conveys the meaning to you, but it's not... It doesn't look like…
Beta	**Arabic-English task:** changes in TM, addressing part of fuzzy matches, extensive post-editing MT suggestions **English-Arabic task:** accepting TM suggestions and extensive post-editing MT suggestions.	Positive, understanding of limitations, evaluation rate 7 out 10.	I have been working with the tools and I know that they are beneficial but they... And they have their limitations, especially with language pairs that are like Arabic and English. This is a difficult language pair, and I understand that.
Gamma	**Arabic-English task:** no changes in TM, minor post-editing MT suggestions. **English-Arabic task:** accepting TM suggestions, ignored fuzzy matches, minor post-editing MT suggestions.	Less positive, criticized translation texts, addressed the issues and complications he encountered, evaluation rate 6 out of 10.	For the Arabic, it was okay. It would need a lot of revision, I guess. I mean the source text, the first one where I translate from Arabic into English, the English output was a bit off
Delta	**Arabic-English task:** changes in TM, extensive post-editing of MT suggestions. **English-Arabic task:** accepting TM suggestions, minor post-editing MT suggestions.	positive, understanding of limitations, evaluation rate 6 out 10.	I would say automated translation is a bonus, and the CAT tool, it saves time.
Epsilon	**Arabic-English task:** Changes in TM, addressing	Positive, understanding of	This semester I'm also working on Trados. It is

	all fuzzy matches, extensive post-editing MT suggestions **English-Arabic task:** accepting TM suggestion, minor post-editing MT suggestions.	limitations. 8	just the same. I mean, we cannot rely on Trados. We have to post-edit the texts
Zeta	**Arabic-English task:** changes in TM, addressing all fuzzy matches, minor post-editing MT suggestions. **English-Arabic task:** accepting TM suggestions, minor post-editing MT suggestions	Overwhelmingly positive, evaluation rate 9 out of 10.	It was efficient. It was super-fast. The translation memory did pretty much everything, I just fixed some sentences
Theta	**Arabic-English task:** accepting TM suggestions, addressing all fuzzy matches, extensive post-editing MT suggestions **English-Arabic task:** changes in TM suggestions, extensive post-editing MT suggestions	Very positive, understanding of limitations, evaluation rate 7 out of 10.	It helped because if I compare what I did to what I used to do before, I think it saves time and effort
Eta	**Arabic-English task:** accepting TM suggestions, ignoring fuzzy matches, extensive post-editing MT suggestions **English-Arabic task:** Changes in TM, extensive post-editing MT suggestions	Less positive, no understanding of limitations, evaluation rate 6 out of 10.	They need to work more for Arabic. They ignore Arabic, we have a lot of issues, Arabic, Japanese, Chinese
Iota	**Arabic-English task:** accepting TM suggestions, ignoring fuzzy matches, extensive post-editing MT suggestions **English-Arabic task:** changes in TM suggestions, extensive post-editing MT suggestions	Less Positive, evaluation rate 6 out 10	It was helpful for some segments. Other segments, no. I needed to delete the suggestion and re-translate the whole thing
Lambda	**Arabic-English task:**	Positive,	It saved me a lot of

	changes in TM suggestions, addressing fuzzy all missed parts of matches, minor-post-editing MT suggestions. **English-Arabic task:** changes in TM suggestions, minor post-editing MT suggestions.	evaluation rate 7 out of 10	time. That way, I can see the next sentences, if anything is repeated, so I don't have to re-translate it, it will be saved into my TM
Kappa	**Arabic-English task:** accepting TM suggestions, addressing all missed parts in fuzzy matches, extensive post-editing **English-Arabic task:** accepting TM suggestions, extensive post-editing MT suggestions.	Less positive, evaluation rate (6) out of 10.	I think the machine translation was not really accurate, so I had to make a lot of amendments, changes to the translation
Mu	**Arabic-English task:** accepting TM suggestions, ignoring fuzzy matches, extensive post-editing MT suggestions. **English-Arabic task:** changes in TM suggestions, extensive post-editing MT suggestions.	Less positive, evaluation rate 6 out of 10.	I would say the tools saved me some time, but the translation needed to be looked at thoroughly to catch all the mistakes and make all the necessary corrections
Omicron	**Arabic-English task:** changes in TM suggestions, addressing one part of fuzzy match, extensive post-editing. **English-Arabic task:** changes in TM suggestions, extensive post-editing MT suggestions.	Less positive evaluation rate 5 out of 10.	The tool is not really smart enough till now, it definitely needs a lot of corpora in order to develop a systematic and a stylistic kind of language in Arabic.

From the Table 5.1 above, it can be seen that some participants who had overwhelming positive views are the ones who made minor post-editing to MT suggestions. For instance, Theta who embraced the tools for time saving has made minor post-editing for

MT suggestion. Lambda and Zeta accepted all TM suggestions and did very minor post-editing for MT suggestions. Alpha, on the other side, has made extensive changes to MT and TM suggestions. This participant has taken the longest time (30.25 mins for both tasks) to provide more accurate translation. This participant and Beta have shown understanding of the limitations, but they didn't express overwhelming positive views as other did. However, there are participants like Gamma, who accepted TM suggestions and made minor post-editing MT suggestions, expressed less positive views, which is contrary to the other groups. Also, Kappa, Mu and Omicron made extensive post-editing MT and made changes to TM suggestions and revealed less positive views.

To sum up, the variation between participants in the Table above reveal that there is no relationship between the behavior of participants and their views toward the tools. The responses of participants to the MT and TM suggestion did not have an impact on their views and evaluation toward the translation tools. This finding might support the previous finding that there is no relationship between complications and evaluation score or satisfaction level. The interesting part is that participants who have less experience level in the use of translation tools have embraced the tools more than the ones who have more experience (Epsilon, Zeta, Theta, Lambda). The ones who have more experience were less enthusiastic with embracing the tools than the ones who have less experience. This would be attributed to experience of the limitations they have encountered during their translation careers than the ones who have just been introduced to the tools just within a year. However, this is not always the case, as Eta who has one year of experience, expressed less positive evaluation than others. This participant can be

considered as an outlier among the others who have the same level of experience. To conclude, participants' responses to TM and MT, using online resources and addressing fuzzy match have no relationships with their views or evaluation of the translation tools.

5.3 Language-Specific Complications When Using Computer-Assisted Translation Tools

5.3.1 Segmentation

Segmentation refers to the coherence of segmentation when Arabic texts are processed. This complication can involve dividing one sentence into several segments or combining several sentences into one segment, which causes alignment complications with the target texts. Despite the fact that SDL Trados provides translators the option to change the segmentation rules into sentence-based or paragraph-based segmentation rules, the segmentation rules of SDL Trados studio and other CAT tools cause difficulties when Arabic language is used, since the tools adopt the English language punctuation rules when it comes to automatic text processing for segmentation. Figure 5.1 below shows an example of segmentation complications for Arabic texts when used with SDL Trados. The sentences are segmented inappropriately into clauses.

Figure 5.1 Inappropriate Segmentation for Arabic Texts

These complications of segmentation occur when the source text is Arabic. Thus, all experiment participants agreed that they faced no segmentation complication in their translation task from English to Arabic. However, 9 participants out of 13, approximately 69%, reported segmentation problems in their translation task from Arabic to English. Interview participants who reported the segmentation complication during Arabic-English translation expressed their dissatisfaction with the segmentation rules used by the CAT tools. In addition, some online survey participants have raised the same concern regarding segmentation rules. One interview participant, Gamma, reported the complication he encountered with segmentation in the following quotation;

> Well, for Arabic, my only issue is that the segmentation rules are a bit difficult because...Yeah. So, you can see they're very long because sometimes there's no full stops. So, you can see that I added full stops in the English version where I felt the sentence was ending.

Another interview participant, Iota, reports the complication he encountered with segmentation where he had to restructure the content of the segment to be readable in English.

> It was the segmentation here. It was a prompt for me, it was where to put the 83%. How to restructure the whole sentence because it doesn't make sense if I translated the sentence as is.

Moreover, Theta expresses her concern that it was time consuming to deal with segmentation complications in the Arabic-English translation task as stated below:

> I have a comment on the segmentation rule, I guess the one who used this tool, he has to put the segmentation rule built on smaller segments. Because this is the major problem. It took much of my time to just break sentences. So, if it were just a small segment or chunks it would be better.

Gamma proposes a suggestion to deal with segmentation complications by providing special segmentation rules for Arabic language, as he states in his own words:

> I'd say they need to work on some kind of segmentation rule, like a preset segmentation rule for Arabic. I know that you can play around with it and change the segmentation rules by yourself, but I think if there was a preset segmentation rule that is set for Arabic, I don't know how they would do it but it would make it much easier, translating from Arabic into English.

Another interview participant, Delta, shares the same concern regarding segmentation and he proposed the idea of complementing special segmentation rules for Arabic language;

> The segmentation in the first task was problematic, I think, because the second segment was too long to follow. Well, I would suggest having more advanced segmentation rules. Yeah, so to make the sentence... The segments smaller for me.

Also, another survey participant reports "segmentation, not compatible with Arabic's complex sentence structure" as a limitation of the tool. It is worth mentioning that those participants who reported segmentation problems in the experiment did not attempt to split the long segment in question. SDL Trados provides splitting and merging of segments features within the tools. Examination of the screen recordings show that no participant addressed the complications with those provided features. Perhaps they are not aware of these features.

From the quotations of interview participants demonstrated above and other participants from the online survey who raised the same concerns with segmentation rules for Arabic language, it is apparent that CAT tools have not taken into consideration the structure of different languages apart from the European languages. It is worth mentioning that the experimental translation texts contained a low translation load. In real translation situations, translators could face more complicated problems with segmentation, as the texts for translation would be longer. However, there are 4 participants who didn't report they had segmentation complications, probably because of the short text task. Despite that, segmentation complications with Arabic texts causes difficulties for translators, especially the ones who are new users to translation tools and are not adapted to the current limitations. All these complications of segmentation can be solved manually through the features provided by SDL Trados. However, the time-consuming process of

going from manual work to change settings and reworking on segmentations complications make it harder for translators. If participants were given a 50-page Arabic-English project and asked to translate it, they all would probably agree there would be segmentation problems because it would be time consuming to fix all the segmentation complications manually for the large number of segments. Thus, proper automatic segmentation is a necessity to facilitate the translation process.

5.3.2 Punctuation Complications

Arabic language uses similar punctuation marks to those in the English language. However, it differs slightly based on the direction of Arabic as a Semitic language using right-to-left. For instance, the comma is written (،) instead of (,) and question mark is written as (؟) instead of (?) (N. Y. Habash, 2010, p. 14). Despite the similarity in punctuation marks between Arabic and English, Arabic punctuation rules are completely different from the English ones (AlQinai, 2008). SDL Trados 2019 adopts the punctuation rules of English in texts processing while dividing segments, where such a strategy causes difficulties for Arabic language translators to deal with during the use of the translation tool. Punctuation complications refer to the transfer of English/Arabic punctuation rules into the target language. In the experiment, 8 participants reported they had punctuation problems in their Arabic-English translation while 5 stated they didn't. As for the English-Arabic translation task, 9 participants reported they had punctuation problemss while 4 participants stated they didn't. Despite the disagreement between

participants, the majority of participants from the experiment and some online survey participants have expressed their concerns about the punctuation complications while using CAT tools.

An interview participant, Beta, demonstrated the complications of the differences between Arabic and English punctuation:

> In Arabic you can have, for instance, you can have a comma and then start a new sentence, and it wouldn't be a fragment. But in English, you cannot just write a sentence and then... Like a complete sentence, and then like a full stop, and then another complete sentence. That'll be a fragment... A run-on sentence like, so... But in Arabic, comma splice is very common. It's... It's okay. That's why I had to make these... So many changes with the punctuation. So, whenever a sentence ended, I had to change the punctuation in English from a comma into a full stop and then capitalize the letter or...

Another interview participant, Omicron, reports the difference in using 'wa' [and], and the comma use in English. Arabic language tends to use 'and' somehow equivalently as a period (full stop). Most new sentences in the same paragraph start with 'and' without using the period. This is a major difference in punctuation rules from English language as reported in Omicron's statement below:

> The punctuation in Arabic kind of copied the one in English, where I had to add things like wa it's not only commas that we have in Arabic, but we have the wa instead of the commas.

Theta expresses her concerns that the tool has adopted the English punctuation rules as she states: “I guess the tool follows the English rules of punctuation, so I need to change and of course because I break sentences, so I change the punctuations.” Moreover, Mu has reported the amount work that needed to be done regarding the punctuation complications as he states:

> I had to change the punctuation because when translated from Arabic to English, the punctuation was copied. And in Arabic, sentences tend to be longer and segments are separated by comma, where in English sentences are shorter and there is a period at the end of the sentence, and then you start a new sentence. So, I had to change the punctuation and make sure to change the letters or to capitalize the first letter of each sentence to an upper case.

Other participants from the survey and the experiment have shared the same concern regarding the punctuation problems while using the CAT tools for Arabic language translation. They expressed that it was time consuming while they needed at the same time to work on editing the translation to convey the right meaning. Some other participants who didn’t agree they had complications with punctuation revealed that they think it is part of the post-editing job where they should expect such type of work while using the translation tools. However, examining the screen recordings and the translation outputs reveal that some participants didn’t address the punctuation problems. Some of those participants did report punctuations problems while others reported the

complications when they were asked about punctuation. There is only one participant who addressed the errors but didn't reported the punctuation problems as a complication.

5.3.3 Arabic Language-Script Related Complications

Arabic language script relies on diacritics that determine the syntactic and morphological status as well the pronunciation of words. These diacritical marks are usually removed in the majority of written texts. However, 'Hamza' can be a diacritical mark and letter form depending on its use in a given word. Its use within Arabic texts is vital since missing it could result in a misunderstanding of the intended meaning (N. Y. Habash, 2010). The default spell checker in SDL Trados is not compatible with Arabic as it doesn't recognize the errors in spelling. However, the tool gives the user an option to use Microsoft word spell checker instead. The editing tool in Microsoft office is much better, although it still has some problems with Arabic particularly with detecting the appropriate use of 'Hamza' and diacritical marks for Arabic. Therefore, since the translation tool does not detect the correct use of diacritics or Hamza, translators need be careful reviewing the translation texts.

All participants from the experiment revealed they had no complication with spelling in their Arabic to English task, while 12 out of 13 participants reported they had complications with the spelling in their English to Arabic task. Probably, translators have more agreements with the spelling checker for the Arabic language than they do with other complications of the tools. This could be explained in that spelling problems cause a headache to translators since they need to detect the spelling errors for each Arabic

word in their translation, which consumes more time and effort as compared to the other complications. In the interviews, participants have expressed their concerns with the translation tool's spelling output. One interview participant, Beta, reports the complication of Hamza recognition in the following quotation:

> It did not have a Hamza under it. But this is something very important. It's like the accent in Spanish, for example, or in French. It changed the pronunciation of the word. The fact that the machine did not trans... Did not recognize that error... And if I take that for granted that the machine recognizes all the errors, like the spell checker is on, and I just like to ignore... Like I don't... If I didn't pay enough attention, I would just make a mistake and I wouldn't know that there was a mistake. So, yeah. This is an example; the spell checker is not really that accurate.

Other participants share the same concerns as Lambda stated: "I faced some issues with the diacritical marks with the Hamza." Kappa also reported the problem of MT output of the missing Hamza through the text: "the problem sometimes is the output coming from the MT doesn't have the Hamza trait, so you need to fix it." Moreover, Iota explains the complexities of Arabic script by describing the special characteristics of Arabic that represent a challenge for the tool to deal with "Arabic is also a special case with the translation tools. Because of the diacritic, we call it the diacritic system in Arabic, the tool cannot sometimes differentiate between words and cannot read the texts very well because of that."

Another interview participant, Omicron, compares the spell checker output of English with the Arabic one where he states:

> It would do a perfect job for English but not really with... Let me check. So, this one works fine. It does not always provide the correct one. It is good to have it in there for the Arabic, but it's not as perfect as with the English, with the English it was way better.

The previous discussions with participants reveal the problems of spelling outputs the translation tools provide for processing Arabic language texts. Arabic translators have more effort to deal with post editing and detecting the spelling errors of the Arabic texts which could likely lead to more time doing post translation work. It is worth mentioning that some participants didn't address the spelling problems in Arabic language text, but they reported the complications when they were asked about it. Time spent has no relationship with this complication because there were other factors that impacted the time spent variation between participants, as has been discussed in Chapter Four (e.g. changing TM suggestion and extensive MT postediting).

5.3.4 Language-specific output

The evaluation of participants regarding language-specific output is not as positive as it is with efficiency of the tools. There are some dissatisfactions and concerns raised by participants. One interview participant, Eta stated "They need to work more for Arabic. They ignore Arabic, we have a lot of issues, Arabic, Japanese, Chinese." Moreover, a survey participant reports "Arabic, I believe, is not given much importance as other languages within CAT tools." This participant's statement is probably inaccurate as there

are factors that have strong impact on Arabic language MT quality and CAT tools in general. First, isometric parallelism is greater between closely related languages, like English and European languages, which makes MT potentially more effective. Arabic language has unique characteristics that would complicate the functions of natural language processing tools. Second, there are greater data resources backing up the neural network-based MT when working between some of the other languages. The Arabic language content on the internet, for instance, is less than 1% (Mishal, 2015). However, this type of disappointment revealed by some participants is understandable as it is caused by several complications that Arabic language translators face while using CAT tools. These include several concerns that will be discussed in the following sections. An interview participant, Delta, reported his experience with language-specific complications in this quotation below:

> The tools helped speed up the process, but it also introduced some problems with re-arranging things. It caused editing problems. It took me a while to re-arrange some syntactic structures.

Another interview participant, Omicron, describes the complications caused by the two languages differences in structure which adds more work to be done while using CAT tools for translation between Arabic-English.

> They were kind of efficient to find the translation in there, but it required a lot of post-editing, especially the text that were really kind of consistent sometimes in

> terms of translation, especially I mean regarding the structure, the way English is structured is definitely different from Arabic, and that's why I had to do some sort of foregrounding and backgrounding for certain information. This is... and I had to change different things like this is totally perfect in Arabic to have a kind of complete sentence of three lines, it's quite long one but it's totally perfect in Arabic. Which wouldn't be the case in English. So, I had to really decide where each sentence can stop, where I can start another new sentence.

Omicron adds in his demonstration of his experience with the tools that the CAT tools still have challenges to recognize the systematics and stylistic characteristics of Arabic language, as can be seen in his statement below:

> The tool is not really smart enough till now, it definitely needs a lot of corpora in order to develop a systematic and a stylistic kind of language in Arabic. And I know this is kind of hard... Yeah, and that's why it requires a lot of post-editing.

Moreover, Delta agrees with Omicron's statement, as he suggests that there is a need for a large amount of parallel corpora data, so the performance of the tools can be improved syntactically and semantically, as he states in his own words: "I would say, there needs to be more context added to these tools because, obviously, the structure has some issues, the syntactic structure into English and into Arabic."

5.3.5 MT Output Accuracy

The accuracy of MT suggestions had poor quality for both directions, although participants have given the English-Arabic direction better ratings. They thought the suggestions of English to Arabic were much better than the translation suggestions from Arabic-English. An interview participant, Mu, expresses below his opinion about MT translation output between the two translation tasks:

> Yeah, I think the Arabic to English was not as accurate as the English to Arabic and needed more modifications. I think that might be because of the sentences' structure, the Arabic structure is kind of like way different from the English, where if you translate English to Arabic, it's easier to like to get the meaning.

Another interview participant, Kappa, expresses his dissatisfaction with the quality of MT output as he states: "I think the machine translation was not really accurate, so I had to make a lot of amendments, changes to the translation."

Furthermore, another interview participant, Lambda, emphasizes below the condition of MT that it is still not smart enough to be able to depend on its suggestions:

> Okay, the tools, they're very useful, but I would say not... It doesn't function as a human brain. Sometimes it just gives you a bunch of options and neither of them is correct. So, you need to read the context again, and make sure this is the right term to use.

Therefore, MT suggestion was a sort of help for translators although it introduced some problems as the quality was poor for the Arabic-English translation task. Despite that, it helped some participants with some terms that they were unfamiliar with, as has been discussed in the previous section.

5.3.6 Limited Beneficial Use of CAT Tools for Arabic

One interview participant pointed out some limitations with respect to the use of CAT tools in translation. As Iota argues, the translation memory systems rely completely on the quality of the inserted TM and without such quality, the CAT tools become useless in translation, as he states in his own words:

> Well, first of all, if you're depending on the tool itself without machine translation, it means that if your input is good, then the output will be good. If it's bad, then the output will be bad. Well, you know, they say "Garbage in, garbage out." So that's one of the limitations.

In addition, Iota presents another limitation of CAT tool use for Arabic language. He argues that CAT tools can be only beneficial for controlled language texts (e.g. technical, legal etc.) while it is hard to be used for literary texts as he states in the following quotation:

> I think that the CAT tools cannot really overcome the type of language they deal with. So, for example, as I said, if you're trying to translate a technical text, a medical text, legal, anything like that, yeah, the tool would be of great help. But if

> you're translating, if I can say like a loose-language text or something like analysis or a political article or news or something like that... Yeah, the tool wouldn't help that much. Even the machine translation.

Moreover, one survey participant agrees with Iota statements as he/she reports about CAT tools limitation, they "cannot be used in artistic or creative approaches". Moreover, Akhrameev (2016), a certified Russian-English translator, agrees with both statements regarding the feasibility of using CAT tools with literary texts as he states in his online report about the CAT tools: "it is not feasible to buy an expensive CAT tool if you are translating literature with 0% of repetitive text segments". Akhrameev adds in his argument "when it comes to technical, financial and legal texts, CAT tool will certainly make translation processes easier and faster". The opinions of the two participants and Akhrameev (2016) are similar in that both think the use of CAT tool is still limited to specific types of texts, so CAT tools cannot help all translators, especially the ones who are working on literary translations, as they argue.

5.3.7 Bi-directionality complications

Bi-directionality refers to the use of both directions (left to right and right to left) in the same translation segment in any CAT tool. It is a complication that affects translators who translate texts that include English scientific acronyms or for instance companies' names in Latin alphabets. The problem occurs when the translator needs to keep the English original term within the Arabic language text as it is in its Latin alphabet form. Although this complication was not designed to be tested in the experiment, it was raised

by two of the participants in the experiment as a general complication they encounter in their frequent use of translation tools. Kappa, as he states below, describes the complication he encounters as a translator when he has to include English terms within the Arabic segments:

> "There are a lot of problems that happen when you have English texts, and you need to use it in the translation with the Arabic text. Usually, there are problems that happen, and usually when you clean up the files, you would need to do a lot of formatting. bi-directionality problem, yeah. If you have to use like let's say a company's name in English, you wanna keep it in English. Sometimes there is problems with the directionality and then it appears in a formatting. If it's a Word document, you would see problems that you need to fix. And if let's say in a company, a project manager doesn't know Arabic, he wouldn't notice a problem. So, needs to go back to the Arabic translator to fix it. It is a disturbing for Arabic language users."

Another interview participant, Omicron, discusses the same concerns in this statement;

> I have comment about bilingual texts in the same segment, there is an issue with Arabic as right-left language in CAT tools, in case you get English and Arabic words in the same segment, you will get issues with formatting.

Moreover, there are 11 participants from the online survey who have raised the same problem regarding the bi-directionality (RTL and LTR) complications as can be reported in the following quotations: “Direction (RTL). All translation tools have room for improvement in terms of direction”, “not compatible with Arabic right to left layout”, “LTR issues and final output”, “Being different in direction, Arabic creates all the problems”, “Direction of the text when we found English and Arabic in the same sentence” and there are other comments by online survey participants.

It would be assumed that the problem is encountered in all language pairs that have different directionality from English (e.g. Hebrew). Several complaints about the complication were discussed online for other languages with a writing system that starts from the right such as Hebrew, Urdu, and Persian. This study focuses on the complications for Arabic language. Therefore, it can’t be proved or disapproved that right-left languages may encounter the same problem. The problem described by participants shows that segments get messed up when another English name or acronym or Latin alphabets in general is used in the Arabic segment. This leads to complications in formatting after generating target texts in a Microsoft Word file as they describe. Translators would need to start from scratch fixing all formatting that include tables and columns if applicable. This problem becomes worse if bi-directionality occurred more frequently in a given Arabic translated text. However, after validating the complication and searching for potential solutions for the problem, it was found that the complications can be resolved with a few clicks in 2019 SDL Trados as can be demonstrated below.

If a translator has bilingual texts (e.g. Arabic and English in the same segment), some steps need to be taken need, 1) switch language (Alt + Shift) and 2) and switch direction (Control+ Shift) each time switching languages is needed while typing. Also, the users will need to verify language settings while they are working on the project: File > Options > Editor > Languages and check the settings. This is needed because the translation tool might get confused with bilingual texts since it detects the languages used automatically. Figure (4.15) demonstrates the process of verifying language settings in SDL Trados 2019.

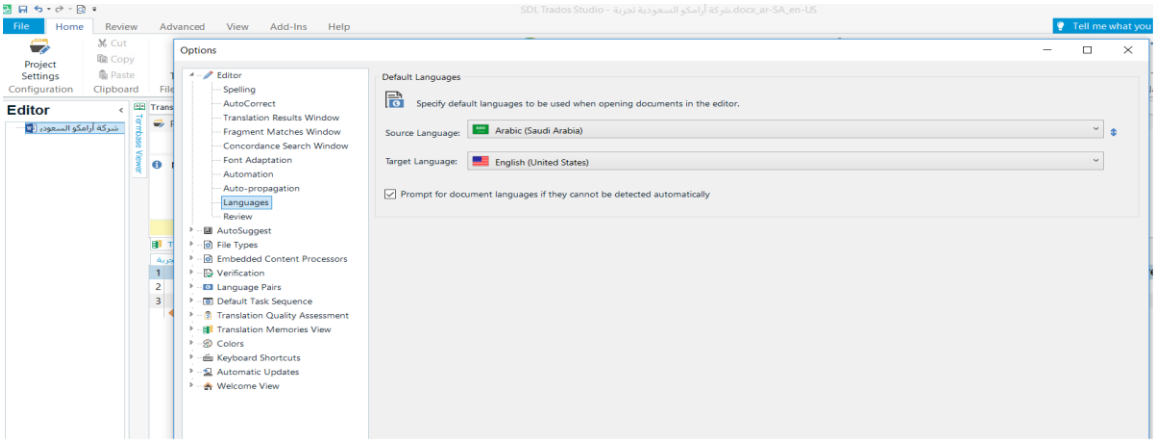

Figure 5.2 Verifying Language Settings in 2019 SDL Trados

Furthermore, the bi-directionality was tested on MemoQ.9 and revealed no complications with having both Arabic and English texts in the same segment. Figure (4.16) demonstrates a screenshot of experimental work having both Arabic and English in the same target segments. The problem can be avoided if language and direction are switched appropriately as demonstrated above for SDL Trados use.

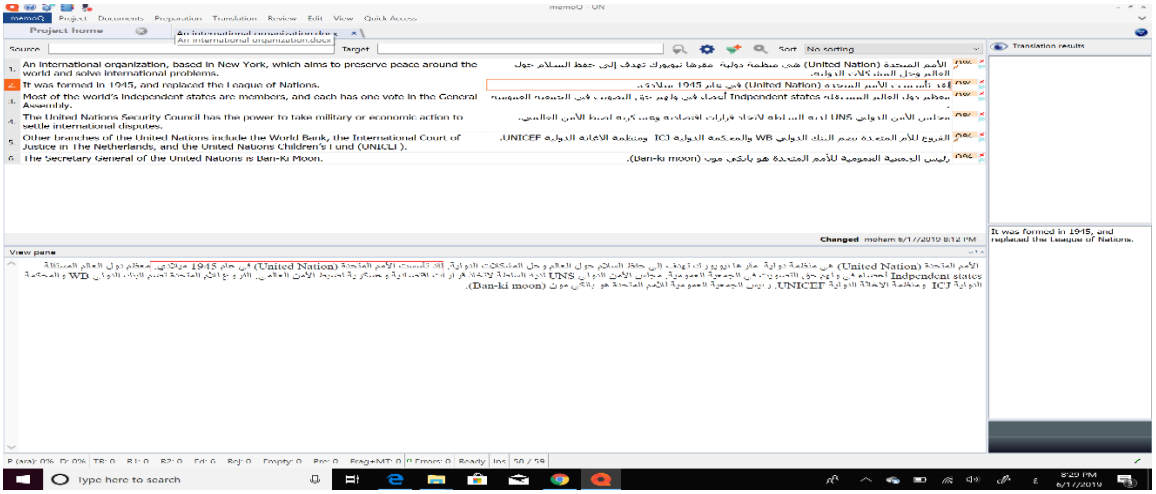

Figure 5.3 Bi-directional Texts in an Experimental Work on MemoQ9

An online search about the complication was conducted to examine the other complaints and how SDL Trados responded to these complaints. It was found out this problem was a bug in SDL Trados 2017 and an update was provided to solve the complication for that version. In 2019 SDL Trados, some experimental work was done to check if the problem still exists or not. It was found that there was no problem if languages were switched appropriately as explained above. Therefore, the participants, who reported their concerns regarding bi-directionality complications in bilingual segments, are either unfamiliar with the solution for this problem in 2019 SDL Trados or they are still using older versions of SDL software which still have complications with bi-directionality. A suggestion was made that SDL Trados offer updates that can fix the bugs occurring with bi-directionality. Also, it would be beneficial to produce tutorial videos about translating bilingual texts that involve RTL and LTR languages in the same segments. Such tutorial videos would

help translators who are unfamiliar with the problem which will lead to higher satisfaction levels and fewer complaints with the CAT software.

5.3.8 Other Arabic Language Related Complications with CAT Tools

There are other responses from the online survey that didn't fit into the complications discussed above regarding the CAT tools, and which also included other complications of the tools use for Arabic language that were not covered in the interview data. This section reports separately those complications raised by the survey participants. Some of these concerns are not relevant to the research questions addressed here or Arabic language-related complications. For instance, some participants reported the problems with clients who are not willing to pay for the 100% match segments, while this type of match still needs to be edited due to contextual differences. This type of complication is general and is not exclusive to Arabic language or has nothing to do specifically with the translation tools. Therefore, there are some exclusions for some points raised in the survey responses. However, there are two Arabic language-related complications that were raised that are worth discussing in this section.

First, participants have raised their unpleasant experience with OCR tools for Arabic. As one survey participant reported "Working on the OCR for Arabic Language is a great step towards improvement". This statement shows the need to work on developing a more accurate OCR tool. Most translation tools don't provide accurate results for Arabic scanned documents that are not digitalized. As has been discussed in Chapter Two, the process of recognizing Arabic texts automatically encounters several challenges

including: the connectivity of Arabic script, recognizing the dotting that distinguishes between similar letters shape, diacritical marks, and Arabic multiple grapheme cases change shape according to their position in the word. Consequently, most up to date available OCR tools in the market "have low performance accuracy rates, below 75 percent" (Alghamdi & Teahan, 2018, p. 239). Therefore, one participant reports "most of the translators do not recommend using OCR if the source document is in Arabic". Translators have to digitize scanned documents by retyping all intended translation contents which is time consuming and prevents translators from the potential advantages the translation tools provide them with.

Second, translators raised the problems of terminology management tools as they don't recognize the plural forms of Arabic terms, for instance. This difficulty of the tools to recognize those forms can be attributed to the characteristics of Arabic language terms which tend to have prefixes and suffixes that could change their morph-syntactic status.

Terminology management systems can be crucial, particularly with specialized texts. They reduce inconsistency with terminological usage to ensure the standardization of terms. This important feature can guarantee consistency of terms among group projects where more than one translator is working on the same project (Melby, 1992, pp. 158–159). Furthermore, it also can help "to cut costs, improve linguistic quality, and reduce turnaround times for translation, which is very important in this age of intense time-to market pressures" (Bowker, 2002, p. 77). Despite all the benefits that terminology management systems can provide to translators, some participants in the survey complained that available terminology management systems don't recognize Arabic

plural forms or any other prefix or suffix of a given term which forces translators to enter the plural forms for each term in their term base. This complication reduces the potential benefits of terminology management tools as the complication of adding only the plural forms for each term could consume a lot of time and effort for translators. Solving such a problem in the next version of Multi-term tool would facilitate the work of Arabic language translators by reducing the inconsistency that might occur with non-standardized terms.

5.4 Needed Improvements in CAT Tools

In previous sections, the complications that need to be addressed so translation tools can meet the needs of Arabic language translators have been discussed. There are other general suggestions that were proposed by online survey participants and the experiment participants that would improve the use of the CAT tools for Arabic language translators. However, those suggestions are general and could be used for other languages as well.

Participants from both the online survey and experiment have suggested some features to be added to the CAT tools to help in increasing their productivity and saving time and effort. The suggestions are as following: search of concordance, dictionaries, online search. They would like to see these features included within the CAT tool software itself instead of the need to exit and use other resources to do their own search.

For instance, Delta, has suggested more access features e.g. search of concordance within the tools itself without the need to leave the tool every time to do an online search to verify specific a term as he states below:

> I'm looking forward to seeing a CAT tool that is interactive enough to let me do the research about the vocabs or terms or any other aspect, cultural aspects, that needs to be researched. I would like to see a feature that allows me to do that within the tool itself. I don't have to exit the tool and come back.

Another interview participant, Zeta, shared the same concern where he suggested the search of concordance within the tool in the following quotation:

> What I'm saying is, for some terms, if you want to see what other suggestions are other than the machine translation that's already given to you. Sometimes you just want to see if this term is actually used in English, then you validate it... Just going through various texts, but you have to go to use Chrome or Firefox.

According to participants, reliable search of the concordance within the translation tools can help them in validating specific translated terms which could result in saving time and effort. Some participants spent considerable time searching online for specific terms while they were performing the translation tasks. Therefore, their need for reliable search features within the tools is reasonable. The current developments in translation tools are seeking to provide products that give translators access to all the features and services they need while translating.

Furthermore, a full editing environment similar to the Microsoft office editing interface was suggested by some participants. Participants are looking for translation tools that can

provide them with more access to features, so they don't have to use other applications to finish their translation work e.g. Microsoft Word. SDL Trados provides some editing features, but the translators need to go to the settings in order to change font size, for instance. It would be great for translators, if editing features are provided within the user interface like Microsoft Word does.

Moreover, the current integrated MT with SDL Trados Studio provides translation suggestions when there is no TM for a given translation segment. However, it does not provide suggestions to fuzzy segments despite the percentage of matching. Some participants complained about this shortcoming where they thought MT should give suggestions even if there is a TM for a given translation segment. The availability of TMs for some segments does not necessarily mean there is no need for MT suggestions. Therefore, it would be great if MT can provide suggestions for all type of segments.

Moreover, some participants suggested adding a feature to process the image files to translate the texts in those images. This feature would save time and efforts for translators if such features can be accessible within the tools. However, adding this feature will require the use of OCR systems to process and recognize texts. This will result in poor outcomes as discussed previously. The current OCR systems are still having complications in accuracy with Arabic language. Although Arabic language translators might still get the benefits of processing image texts for non-Arabic texts (e.g. English to

Arabic translations), this feature might add a greater value for other translators than Arabic.

Other participants complained about bugs and slow responses from the tools. This concern probably can be attributed to their personal computer processor as the translation tools require fast and high-quality computer processors. Thus, one participant complained that there is no 64bit version of SDL Trados. The available version has only one option as 32bit. The participant argues that this type of version can be cumbersome for the regular processor and would need high RAM and fast processor. After searching online about this complication, several translators from several language pairs were complaining about the problem in SDL Trados community website. The response from the company was suggesting the use of a fast-solid-state drive and fast processor to see the benefits of the tool rather than asking for an upgrade to 64-bit. Their response indicates that they are keeping the 32-bit version for faster performance. However, not all freelance translators can afford renewing their devices periodically to keep abreast of the developments in technology. For instance, the translation rates in the Arab world are very low. Experienced professional translators in Egypt don't make more than a thousand dollars monthly. Thus, providing a 64-bit version would facilitate the translation process for the majority of Arabic translators who don't have access to fast computer processors.

These suggestions from the experiment responses were mentioned as well in other studies (Moorkens & O'Brien, 2013, 2017) where authors reported similar suggestions of their participants who wanted to see some improvements to the user interface that include

providing access to features, e.g., dictionaries, Internet search, and better concordance search. Despite the fact that the results of this dissertation suggest that participants seem to be satisfied with the current translation tools, they are looking for some better features that give them more confidence to make translation decisions without the need to use outside resources. A quotation of one of the interview participants, Epsilon, summarizes the translators' hopes for a better tool as he sates: "The tool is good. Trados is good. But I think it can be better."

CHAPTER VI

DISCUSSIONS OF THE RESULTS

This chapter presents the findings of the study from both the online survey and the experiments. It discusses the responses to the three research questions posed in Chapter One. This chapter is divided into four parts. The first part is an overview of the evaluation and views of Arabic Language translators toward the CAT tools. The second part demonstrates behavior of participants toward translation tools. The third part discusses the current complications of the CAT tools for Arabic language and potential solutions. The fourth part identifies the needed improvements for the CAT tools to meet the needs of Arabic language translators.

6.1 CAT tools from the viewpoint of Arabic language translators

As demonstrated in Chapter Four, the overall evaluation of the computer-assisted translation tools and user satisfaction level was rather positive from both participant groups in the study. Therefore, there were no significant relationships between their

responses and any complications they may have anticipated or encountered with the translation tools for Arabic language and the evaluation of the tools.

As for the online survey, around 61.2% have encountered complications; however, the evaluation of the tools was not affected statistically as most of the participants reported considerable high satisfaction level toward the current translation tools. The same applied to the participants in the experiment, as the majority of the respondents reported segmentation, punctuation, and Arabic script related complications during their translation tasks. However, their evaluation or their views toward the translation tools were not significantly impacted by the complications they encountered.

Contrary to the anticipated results, participants from both the online survey and experiment provided positive feedback affirming the importance of translation tools use for Arabic language translation. The findings from both conducted methods are contrary to the views in the literature, which suggest Arabic translators have been reluctant to adopt the use of computer-assisted translation tools (Al-jarf, 2017; Alotaibi, 2014; Fatani, 2006; Thawabteh, 2009, 2013). Participants in the study showed an understanding of the tools' limitations and an enthusiasm for more future improvements in the tools to better meet their needs. LeBlanc (2013) established similar findings regarding Canadian translators, where he states "Unlike what is sometimes claimed or assumed, the translators who took part in this study are certainly not opposed to the use of new technologies. On the contrary, translators have welcomed the introduction of new tools over the years" (LeBlanc, 2013, p. 10). There's no link between Arabic language translators and the Canadian translators since those translators are dealing with different

language set (e.g. English-French translation) but this finding by LeBlanc is similar to the results of this study where translators have welcomed the use of translation tools, as contrary to the anticipations.

Nevertheless, the findings of the study do not necessarily contradict other studies discussed in the literature about the reluctance to adopt translation tools among Arabic language translators. In fact, this study targeted only translators who currently use the tools in order to guarantee an objective evaluation of tools use and the complications they are encountering. Non-users of the tools were excluded from study to avoid subjective evaluation that might not add valuable contributions to the literature. Therefore, discussions of Arabic language translator's reluctance toward CAT tools might not represent the full community of translators. There are possible explanations for the assumptions or claimed reluctance cited in literature. First, participants in those studies might not be actual users of the tools, so their thoughts and opinions might not reflect the whole community of Arabic language translators or be grounded in actual personal experience. Second, those studies that claim reluctance can be considered a little bit old (Alotaibi, 2014; Fatani, 2006; Thawabteh, 2009). There have been revolutionary developments and progress made in the CAT tools within the last four years. Thus, it is possible that Arabic language translators have changed their thoughts and views toward the use of CAT tools in translation environments. The initial questions and approach taken in this study did not take into account the potential for shifts in perspectives over time. This question needs to be addressed in future research in order to gauge a better

generalized data that can represent the overall community of Arabic language translators assessed at a given point in time.

Tabor (2013) reports his results from an online survey that involved over three thousand translators from around the world. In his report of the findings, 32% of participants revealed they have never used CAT tools in their translation work. This can be considered a high percentage in an era of rapid technological advancement. Therefore, the findings of Tabor's report suggest that it would not be unthinkable that Arabic language translators who have never used CAT tools could be more than 32% of the Arabic language translator population. However, testing this hypothesis would involve collecting a large number of surveys from Arabic language translators to determine the percentage of Arabic language translators who are not using CAT tools in their translation work. It might be anticipated that the results would reveal more than 32%.

This estimate can be attributed to the reluctance of translation schools as well as public and private translation centers to adopt translation technology in their educational curricula. For instance, in current practice, several professional translation units in the public sectors of Saudi Arabia still have not adopted the use of the translation tools in their translation work, even though these translation units translate books and thousands of words each day (e.g. the Institute of Public Administration, the Translation Department in the Royal Court, and other translation centers in public sectors). This information was revealed during data collection attempts with several translation centers. When contact was initiated to invite them to participate in the online survey, they responded that they don't use the computer assisted translation tools at all. King Abdullah

Institute for Translation and Arabization was the only translation center which had limited use of the tools in its work environment. Only four translators among the many other translators have access to the tools. The majority of the translators still use only Microsoft Word documents to translate with some access to online dictionaries. This is a surprising finding since translation tools can facilitate the work of translation tasks by significantly reducing the required time and effort. Financial factors cannot be an issue in the decisions against adopting those tools in these translation centers, because these departments have enough financial support from the government to purchase licenses for the most updated tools for each of their staff translators.

The benefits of using CAT tools in translation work are evident despite some complications for Arabic language. Therefore, when the potential advantages of using CAT tools are compared to the current complications for Arabic language translation, the advantages clearly outweigh all the complications and challenges that might occur. Most translators probably would agree to this assessment, as they seek increased productivity and better efficiency. However, this does not undervalue the necessity to solve those complications identified by Arabic language translators, which some consider minor concerns while others consider them major concerns. Microsoft Word has been successful in providing great service for Arabic language writers and translators with almost no complications except for the spelling and grammar checker, which is under continuous development and improvement.

This dissertation has aimed to explore the importance of the translation tools' use for the translation process by discussing the professional translators' evaluation of the tools and

the complications they encounter and how these complications could be addressed to better meet the translators' needs. Hopefully, these research findings will encourage the public and private sectors to adopt the use of technology in translation to increase their productivity and save the time and efforts of their translators.

6.2 Participants Behavior toward the Translation Tools

The observation of participants' behavior while completing the required translation task through screen recordings revealed interesting findings. Isomorphism between Arabic and English source texts had an impact on some participants' choices and translation decisions. As has been discussed in Chapter Four, some participants made changes to the TM for terms (e.g. 'mouth cancer' instead of 'oral cancer', 'identify' instead of 'recognize'). These changes could be linked to isomorphism between Arabic and English as these suggested changes by participants reflect the influence of their Arabic language as their mother tongue. However, other cases of these changes also might reflect some participants' failure to confirm language usage by looking up terms in parallel texts instead of relying on MT suggestions.

In addition, those changes and other changes in the TM reveal that there is a strong inclination by Arabic language translators to change TM suggestions perhaps without strong theoretical or usage-oriented justification. It has been found that 61.5% of participants have changed TM suggestions in both directions. This suggests that Arabic language translators are more likely to check and edit TM suggestions, which is contrary to what is assumed in literature. Bowker (2005) has warned of 'blind faith' in TM. Also,

Ford (2016) concluded in her experimental study that the majority of translators are more likely to trust TM suggestions. The participants in this experiment showed unexpected behavior when they decided to edit even correct translation suggestions, for instance, from English-Arabic TM to improve it, e.g. by using better wording, flow and clarity to target Arabic readers.

Furthermore, the majority of participants in the experimental study have made extensive post-editing in both directions. It was found that 84.6% of participants made extensive changes to MT suggestions in Arabic-English translation while 53.8% made extensive post-editing in English-Arabic translation. The variation between participants' efforts in post-editing between the two tasks can be attributed to the better quality for MT suggestions in English-Arabic. Participants expressed their satisfaction with MT output for the into Arabic direction. Thus, results suggest that Arabic language translators are more likely to make extensive post-editing for MT suggestions in both directions.

The variations between participants' behavior toward TM suggestions and MT suggestions had a significant impact on time spent to complete the required translations. It was found that participants who made changes to TM or extensive post-editing to MT suggestions have spent significantly more time than others. Thus, there was a drawback for the use of time spent to measure other complexities while using the CAT tools e.g. segmentation, punctuation and spelling problems.

6.3 Complications with using translation tools for Arabic language

Although there is a considerable range of opinions from Arabic language translators regarding the translation tools, the results of the online survey and interviews from the experiment, show that Arabic language translators still have some complications regarding the use of the translation tools for Arabic language. About 61% of the online survey responses indicated that they encounter complications while using the tools. Results of the McNemar non-parametric test revealed a statistically significant association between online survey participants who reported complications and those who didn't, $p = .039$.

Also, the majority of the participants of the experiment revealed complications with segmentation, punctuation, and spelling problems that occur while using the translation tools. Despite the fact that results of the chi-square test of the experiment data only revealed significant results for spelling problems and failed to show significant results for segmentation and punctuation, the majority of the experiment participants reported they have had complications with segmentation, punctuation, and spelling while completing the required translation tasks. McNemar non-parametric tests revealed that a significantly larger percentage of participants reported segmentation complications for the Arabic-English translation task than the English-Arabic task. Also, it revealed that an equal percentage of participations reported punctuation complications in both the Arabic-English and English-Arabic translation tasks.

In addition, the statistical model's regression analysis revealed no significant relationship between the complications of the tools for Arabic language and the time spent to process

and complete translation tasks. The insignificant results might be attributed to the small sample size of 13 participants, which could lead to an inconsistency in the data. Also, there were other factors that impacted the time spent as has been discussed earlier. Participants who made changes to TM and extensive post-editing to MT suggestions spent significantly more time than others. One would assume that overcoming the complications might be time consuming, although the data point toward no significant difference among participants who reported problems and those who didn't. The denial of facing complications while using CAT tools could be due to the pride of participants as Alpha's statement could explain:

> I didn't face any difficulties because I'm accustomed and used to using Trados as a translation tool, but I can envision that others who are not accustomed to using this tool might face issues in how to deal with the various functions.

Another explanation of the few participants who did not report encountering problems while using the CAT tools is that they might think those complications are part of the post editing task as stated by Lambda in the following quotation "I think it's part of my post editing job to fix these minor issues, nothing more."

Moreover, one interview participant, Iota, stated he didn't pay attention to the complications because he didn't use the MT suggestions and instead, restructured the source translation texts. Participants were not informed in advance about the investigated complications, because the study was seeking objective judgements which would be tainted if participants were informed about the study's areas of focus in advance. Another factor can be attributed to the short translation texts required for the experiment task and

participant were not under time pressure which could made the task easier than actual translation work.

However, statistical analysis of both conducted methods, the experiment and online survey, revealed only partial support for the hypotheses of the study. Despite that, there are evident concerns that Arabic language translators still encounter when using CAT tools for Arabic language translation. Segmentation (e.g. several sentences in one segment or one sentence divided into more than one segment. etc.), punctuation (e.g. suggesting English comma (,) instead of Arabic comma (،) etc.), and Arabic script related problems (e.g. spelling errors) were reported by the majority of the participants in the experiment. Moreover, seven online survey participants reported segmentation complications, four reported punctuation problems, and four reported Arabic script related complications as well. Therefore, these concerns are valid and cannot be undervalued even though there are participants who chose not to report them.

The segmentation and punctuation complications are related to each other since the punctuation rules for Arabic are different from English. On the other hand, the translation tools are based on English language segmentation rules and punctuation rules, where each segment is divided based on the use of the period which is considered the end of English sentences. In Arabic, the rule is completely different as a period is used when a complete idea is expressed, so it is totally normal for a paragraph to be one long sentence as it conveys the meaning of one idea. This major language difference complicates the use of CAT tools for Arabic language. Also, the complicated rules of Arabic language punctuation make it harder for a natural language processing system to assist translators

to process the texts appropriately. For instance, the comma in Arabic can be used for ending independent sentences that have the same idea as the previous sentence and can be used as well for listing things and clauses. SDL Trados studio offers options to customize the segmentation rules and allows users to create advanced special rules in addition to the available sentence-based and paragraph-based segmentation rules. However, creating special segmentation rules will require preprocessing of the Arabic texts to meet the new customized segmentation rules. If the 'comma' is set as the marker in the segmentation control system as the end of each segment so that it will be treated as equivalent to the English period, this will create complications as has been explained since Arabic uses commas for lists and clauses as well.

Thus, enforcing sentence-based segmentation as the default option would be a logical and useful way of dividing up the texts for easier processing for most languages. However, the similarity and complicated rules of Arabic punctuation don't help translators to formulate their own segmentation rules. In order to improve the tools to better meet the needs of Arabic language translators, tool developers need to consider making special segmentation rules for Arabic language based on the Arabic language punctuation rules. Improving the competency of CAT Tools in understanding the structure of Arabic sentences, phrases, expressions, and words is vital to developing the tools for Arabic language users. This can be accomplished by adopting and developing one of the available morpho-syntactic analysis tools designed for the Arabic language. More information on this topic can be reviewed in previous research articles (Al-Taani, Al-

Awad, & Abu-Salem, 2011; Boudchiche, Mazroui, Ould Abdallahi Ould Bebah, Lakhouaja, & Boudlal, 2017; Boudlal et al., 2010; Diab, Hacioglu, & Jurafsky, 2007). However, there is no magical solution that solves the problem in the meantime until automated translation tools are revised and improved to process Arabic texts more accurately.

Thus, there are two possible solutions for this matter meanwhile. First, the Arabic language needs to reform its punctuation rules to be more compatible with CAT tools. However, this solution might not be practical for several reasons. First, Arabic linguists are divided into three schools of thought on the matter: 1) Arabic is a language that requires no reform, 2) Arabic needs to be simplified and stick to its core roots as represented by the Cairo Academy, and 3) the Arabic language needs to be completely reformed (Maamouri, 1998). Second, the official Arabic academies are located in Egypt, Syria, Jordan, and Iraq. The research institutions are located in Kuwait, Libya, and Morocco. Due to the fact that these agencies are far-flung across the Arabic world, it is highly likely that there would be significant difficulties in finding consistent methods for reforming Arabic punctuation rules or the Arabic language in general.

Previously nationalistically-driven attempts, aimed at the creation of pan-Arab standardized language reforms, have been superseded by national language planning at the level of nation states. It would be possible in the foreseen future that Arabic States standardize their reforms individually with current wind of changes from Arabic nationalism into state patriotism. Despite the difficulties of reforming Arabic language, there is a need to address the current concerns e.g. comma use. It would be needed to

differentiate between the used comma in compound sentences and the used one for short clauses and listings. This reform of punctuation might facilitate the work with Arabic natural language processing.

The second solution, which is more practical than reforming Arabic language, involves preprocessing Arabic texts to modify punctuation to follow English punctuation rules since translation into English will already require changing the structure of the translation to be acceptable for English language readers. Arabic language translators could eliminate the complication of segmentation through the preprocessing of Arabic texts to facilitate the use of CAT tools. It wouldn't be difficult to change the punctuation rules to be adaptable with the target texts. Although preprocessing is not convenient for translators, as it might create more work for them, it would help increase their productivity and save their time while translating using CAT tools. In addition, such preprocessing of texts might help in improving the performance of MT suggestion output. MT seems to provide better translation suggestions for shorter sentences. Therefore, given that the two languages differ syntactically, morphologically, and semantically, the punctuation complications will remain a problem so that Arabic translators would still need to complete post editing for the target texts to meet the language requirements for each translation direction.

Moreover, using the spell checker tool was a problem for most translators. Statistically, it was found that translators encounter spelling problems while using the spell checker for English to Arabic translation, as the translation tools do not provide an appropriate system that functions the same way it functions with English. Nevertheless, statistically

there was no significant relationship between the time spent on the English to Arabic translation task and the spelling problems. Moreover, there was no significant relationship between evaluation of the tools and the spell check complications. The insignificant results could be attributed to the sample size and the other factors that had a significant impact on time spent. However, the results of the study suggest that spelling problems have no influence on time spent or evaluation of the tools by Arabic language translators.

SDL Trados relies on MS Word and Hunspell spell checkers. Some blame Microsoft for the spelling problems. However, SDL Trados could invest in improving the quality of the output of the spell checker for Arabic language. The company has implemented impressive improvements for the Arabic language translators during the last decade including multiple improvements for quality and performance for Arabic. Compared to the segmentation problems, the spell checker tool is not a hard problem to fix because the tool just needs to be equipped with the appropriate syntactical and morphological analysis data specific to the Arabic language. Most of the errors that spell checker in MS Word or the Hunspell spell checker in SDL Trados don't recognize are easy to detect. For instance, neither spell checker system recognizes the need of Hamza in this word "الى" which should be "إلى". This word means 'to or into' in English and the word without the Hamza does not exist in Arabic. There are many other instances that show the dysfunction of the tools regarding spell checking requirements for the Arabic language. Microsoft Word spell checker can detect spelling errors with the letter forms, but when it comes to diacritical forms or Hamza, it fails to provide any suggestion. The spelling tools

need to be improved by developing the efficiency of Arabic natural language processing to improve accuracy and the quality of the output.

However, the tools currently do not have the capability to recognize the diacritics of Arabic language scripts. The major complication is that most Arabic language texts don't include the diacritics. As has been discussed in Chapter Two, Arabic speakers depend mostly on their cognitive effort to process the texts. There are developed systems that can install diacritics in the Arabic scripts, but these systems have not been tested independently. For further information and details about these systems, consult previous research (Alzand & Ibrahim, 2015; Chennoufi & Mazroui, 2017; Nizar Habash & Rambow, 2007; Zitouni, Sorensen, & Sarikaya, 2006). CAT tools developers are not likely to want to take the risk of adopting systems that might provide poor quality output to their users. This adds a greater complication to the current concerns. However, CAT tools developers should consider investing in the development of a better natural language processing system that is capable of recognizing and processing Arabic texts sufficiently.

Furthermore, participants in the experiment found the translation task from Arabic to English harder than the translation task from English to Arabic, as has been demonstrated in the results above. This led to a statistically significant difference in the time spent between the two tasks. In addition to translating into their native language, English to Arabic is easier because there are better suggestions from MT with minimal post edit requirements as was stated by some interview participants, Gamma for instance reports; "the second one where I was translating into Arabic, the output was much nicer, in my

opinion". The output problems in the MT translation direction from Arabic to English can be attributed to the long structure of Arabic sentences because it provides better results with shorter sentences. For longer Arabic sentences, the MT produces a much lower accuracy output. This problem causes difficulties for translators when translating from Arabic to English. Therefore, some translators prefer to translate from scratch than to use the MT suggestions as has been stated by Iota, "I deleted the segment and re-translated". There are other complications with MT regarding the recognition of the Arabic language script and punctuation rules. The SDL Cloud MT system doesn't recognize Arabic script "diacritics" and punctuation rules properly. However, some translators considered these to be minor concerns compared to struggling with MT to provide an accurate translation that can be at least understood by readers.

On the other hand, as an experiment, the same translation texts for the experiment were copied into Google translate. It provided better terminological options as well as better spelling and punctuation output. As an example of the different output between Google translate and SDL Cloud MT suggestions, Table 6.1 below demonstrates the differences in terms of average number of words per sentence, number of words and number of sentences for the Arabic to English translation.

Table 6.1 Statistical Information from SDL Cloud MT and Google Translate Suggestions

Arabic-English Translation- MT and TM suggestions			Arabic-English Google translate Suggestions		
Average number of words per sentence	Number of words	Number of sentences	Average number of words per sentence	Number of words	Number of sentences
45	90	2	23	94	4

As can be seen from the table above, Google translate produced better output in terms of English punctuation rules. On the other hand, SDL Cloud MT copied the punctuation rules from Arabic text into the English one. Thus, there is a considerable difference between 45 words per sentence and 23 words per sentence for the English texts. Also, there's a difference in the number of sentences, where Google suggested 4 sentences, while SDL cloud MT suggested 2 sentences. Table 6.2 below demonstrates an example of both engines' suggestions for the second segment in Arabic-English translation.

Table 6.2 An example of Google Translate and SDL Cloud MT Suggestions

Google Translate Arabic-English	SDL Cloud MT Arabic-English
Approximately 2% of all cancers diagnosed each year are in the United States. Early detection and treatment of oral cancers is important to raise the risk of survival. The five-year life expectancy of oral cancer patients who do not have the disease is 83% While it is only 32% after cancer spread to other body parts.	Mouth and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America, the early detection of mouth cancer diseases in the appropriate time is important to raise the prospects of survival, life rate for five years infected oral cancers who did not have the disease spread is 83%, for example, while only 32% after the cancer is spread to other parts of the body.

From the example shown in Table 6.2, it can be seen that Google Translate accounted for the punctuation requirements for the target text, while SDL Cloud MT did not. Also, Google Translate has suggested ‘oral cancer’ while SDL Cloud suggested ‘mouth cancer’. However, Google translate suggested ‘risk of survival’ while SDL Cloud suggested ‘prospects of survival’. Neither translation is perfect, but Google translate tends to have fewer language specific complications in general than SDL Cloud MT, although it produces error in the first sentence of the segment. Therefore, it is recommended that Arabic translators use Google translate through SDL as it is an option that SDL Trados offers to its users if they have a subscription with Google Translate.

As another complication, 11 participants in the online survey and 2 participants in the experiments reported bi-directionality complications (right to left vs. left to right language), it was found that the problem can be solved with a few clicks as was explained in Chapter Four. This concern raises an interesting point that not all complaints and problems reported by participants are actual software problems as they can be due to lack of knowledge or experience in dealing with the tool appropriately. Therefore, all mentioned complications were tested and validated to avoid claiming complications that have nothing to do with the translation tools’ shortcomings. Therefore, adequate training for Arabic language translators in the use the CAT tools is needed to eliminate the relevant problems and concerns that are related to user errors more than software complications.

6.4 Improvements required for CAT tools

The CAT tools have been improved through impressive developments in the last decade. They have been developed from a basic translation memory system into more interactive translation tools that integrate with machine translation. Better features as well were added for the sake of facilitating the translation process. Moreover, there has also been much improved performance and support for right-to-left languages such as Arabic and Hebrew in the latest versions of SDL Trados (Shannon, 2015). Furthermore, the participants of the online survey and experiment have expressed satisfaction about the current CAT tools.

Despite these impressive developments and the high satisfaction reported by translators, there are still some complications that need to be addressed as has been discussed earlier (e.g. segmentation, punctuation and Arabic script related problems). Beyond that, some participants shared their thoughts about the possible improvements that they think are needed to improve the current translation tools. Those suggestions are general and could be used for other languages as well. The proposed suggestions included improving the editing environment, having more access to features like reliable concordance search , and online search etc. as has been discussed in Chapter Five. Similar suggestions (e.g. editing environment, online search within the user interface etc.) were raised by other studies (Moorkens & O'Brien, 2013, 2017) and more were discussed by translators in online forums (see SDL community forums). Thus, CAT developers need to take these proposed ideas into consideration in their upcoming versions, so the tools better meet the expectations and needs of translators.

CHAPTER VII

CONCLUSION, LIMITATIONS, AND FUTURE RESEARCH

7.1 Conclusion

Looking back at the results of this study, it is apparent that there is a strong inclination by Arabic language translators in this study to encourage and support the use of CAT tools. Also, screen recordings and translation outputs suggest that Arabic language translators are more likely to make changes to TM and extensive post-editing to MT suggestions. In addition, triangulation of the survey and experiment findings supports the conclusion that there is no relationship between the complications experienced while using translation tools and the evaluation of the tools and expressed level of satisfaction. The isomorphism between Arabic and English source texts had an impact on some participants' choices for Arabic to English translation. Although sample size for both data sets are considerably small, which means the results cannot be generalized to all Arabic language translators, the findings suggest that Arabic language translators are satisfied with the translation tools despite the complications they encounter.

Furthermore, despite the fact that there are 33 % of the online survey participants as well as some experiment participants who reported they encountered no problems with the translation tools, the majority of both groups reported some complications, as has been discussed in Chapters Four, Five and Six. However, looking at the improvements and developments of SDL Trados, for instance, from the 2007 version up to the newest 2019 version, it is noticeable that most of the Arabic language complications have been solved. However, there are a few complications that have been left unsolved with the latest version which are; segmentation, punctuation, Arabic script related problems, and poor MT output. Otherwise, the newest version of SDL Trados is efficient enough with Arabic language. This indicates that the tool developers are trying hard to provide a better product to their Arabic language users. Due to the unavailability of access to other translation tools and the unfamiliarity of most participants with the other translation tools, this study did not investigate the complications of other available translation tools in the market (e.g. MemoQ, Wordfast etc.).

In conclusion, the results of the research showed partial support for the hypotheses but introduced interesting findings at the same time. The study aimed to find enough evidence which either supported the hypotheses and the relationship between variables or contradicted them. Even though some of the hypotheses are "not supported (or fully supported) this is still a valuable research outcome" (Saldanha & O'Brien, 2014, p. 18). The valuable outcome of the dissertation is that the complications Arabic language translators encounter while using the tools available for Arabic translation have not impacted their evaluation or satisfaction level toward the tools. It seems Arabic language

translators are adapting to the complications including Arabic script related problems which were statistically significant. However, there is still a need to improve the quality and eliminate the current complications, so the tools can attract more Arabic language users.

Furthermore, the tools developers would need to do more client education to perhaps dispel retained fears resulting from previous problems that have resulted in negative evaluation among translators in previous studies. Some of the complaints have already been solved (e.g. Arabic language directionality). Thus, tool developers and translation educational programs need to address those concerns and introduce the current limitation and how can deal with them to translators who are having trouble to adapt with the translation tools.

7.2 Limitations

One limitation of this study is the sample size of Arabic language translators. The nature of the study required that participants spend their time and energy to participate in the survey, the translation task, and an interview. This led to a smaller number of relevant participants who were willing to volunteer some of their time to the study. Overall, there were 49 participants for the online survey and 13 participants for the experiment. It was a challenge to get this minimum required number of participants although valuable incentives, in the form of gift cards, were offered to participants in both conducted methods to encourage as many participants for the study as possible. Furthermore, there was a limited population of individuals for the experiment who are qualified to

participate. Thus, the results could give an understanding of the complications and challenges that Arabic language translators face and what potential improvements should or could be adopted to meet their needs. Nevertheless, it is difficult to argue that the results demonstrated for the collected sample size, in this study, will hold true for all Arabic language translators, at all levels of experience.

Moreover, although all the participants in the experiment are graduate students in the translation program, some of them are considered trainee translators since they have only one year of experience in the translation field. Therefore, their responses to some specific questions regarding the complications of the tools and the time they spent on the translation process might have caused some inconsistencies in the data.

Furthermore, passages of translation tasks were unnaturally short, so participants didn't encounter a full range of segmentation problems or potential translation problems. A much longer passage might give more meaningful data, but it would also become even harder to find participants who are willing to spend more time translating long passages. Thus, the current experiment does not represent actual translation work. A long-term evaluation with in-house translators using computer-assisted translation tools would be needed to address the full range of complications and how translators deal with them.

Another limitation is the unfamiliarity of the new version of 2019 SDL Trados translation tool to some participants as they have been using older versions of the tool, which caused some complications at the beginning of the translation task as they started to get used to

the new functions of the tool. However, participants were able to adapt easily as they were instructed how to locate the needed features during the translation process.

In addition, time spent to complete translation tasks was used as a measurement of the level of complexity that participants encountered (e.g. segmentation, punctuation and spelling problems). It was assumed that time spent would represent these complications. However, the findings of the studies revealed that there were other factors that impacted the time spent significantly, including using online resources, changing TM suggestions, and making extensive post-editing to MT suggestions. Thus, time spent was problematic since it could not be an ultimate measurement of the complications of spelling, punctuation and segmentation. The results revealed no significant time differences spent between participants who reported the complications and those who didn't. Therefore, this is considered one of the limitations of the data.

Moreover, the responses of participants regarding complications they encountered during the use of CAT tools were subjective, because some participants see these complications as part of their job as post editors, other participants see them as time consuming. The goal of the study, however, was to represent the various views of Arabic language translators regardless of their differences. The study attempted to gather objective evaluations as much as possible by recruiting only translators who were experienced with the CAT tools and excluding translators who were not familiar with the CAT tools to avoid a poor evaluation of the tools due the participants' lack of experience.

7.3 Future Research Directions

It is necessary to investigate the views of Arabic language translators, the complications they encounter, and needed improvements to be addressed in future versions of the CAT tools, given the reluctance of translation centers and schools to adopt the technology in classes and programs in Arabic language translation. This dissertation has addressed these concerns and focused on the Arabic language translators' evaluation of the translation tools, the complications they encounter, and the improvements they need to meet their requirements. However, the limitations outlined above suggest several possible avenues of future research which include expanding participant pools that would provide the generalizability of the present study findings.

In addition, this study has targeted only Arabic language translators who are currently using the CAT tools. Thus, it would be recommended for future research to conduct a longitudinal study that examines the views of Arabic language translators who do not use CAT tools in their translation work and are unfamiliar with them. This will require introducing those translators to the developed translation tools through designed workshops. The study should aim to examine how participants respond to the effectiveness of the translation tools before and after they have learned about the tools and had hands-on experience using them in their translation work. It would be interesting to check if those participants' reluctance to use the tools would have been impacted after they have been introduced to the advantages and limitations of current translation tools or not. Nevertheless, recruiting a minimum acceptable number of participants of such a longitudinal study would be a challenge and would probably take several years, perhaps

working with a translators' professional organization or tracking successive cohorts of advanced students.

Furthermore, another interesting research direction would be investigating the cognitive efforts of Arabic language translators in comparison with translators from other languages while using computer-assisted translation tools. It would be interesting to examine if there is a significant difference in cognitive efforts made by Arabic language translators while using CAT tools as compared to the effort made by other translators from different language pairs e.g. English-Spanish. The findings of such research might provide further insight into the complications of the translation tools use for Arabic language. It would be anticipated that Arabic language translators would have to use more cognitive effort than translators of other languages. A result of such prospective research would be a valuable contribution in the field.

In addition, although translation to non-native languages can be a hard task for the majority of translators, most language service providers in the Arab world expect their hired translators to be able to translate in both directions. Therefore, translation directionality is a real concern in the translation job market. This study indicated that there is significant difference in level of difficulty in translation direction as stated by the majority of translators and as the time consumed during the study has shown. However, it would be interesting to cast light on the concern by conducting a study to examine the cognitive effort of more experienced and professional Arabic language translators when they translate in both directions and to see if there is a significant difference in cognitive efforts during both parts of the translation process. The current study was limited because

some of the participants were not professional translators, which further indicates that the findings of the study cannot be generalized.

Finally, investigating the accuracy differences in machine translation outputs regarding the translation directions from Arabic to English and vice versa would add a great contribution to the field. Even though more studies focused on solving MT complications for the Arabic to English direction, the MT output of English to Arabic seemed to be much better in accuracy and clarity. Although Arabic is more complex, it is probably in many cases perfectly capable of accepting more simple English syntax without sounding odd, but a careful Arabic stylist / human translator might in some cases merge sentences and change the style. But translating the complex Arabic structures directly into English, which doesn't support this much flexibility, is likely to end up sounding weird if it isn't carefully edited. These concerns could be due to the nature of Arabic language structure where it tends to use longer sentences which complicate the natural language processing for the MT systems when translating into English. Indicating the reasons behind the poor quality of MT translation output from Arabic to English and providing potential solutions would increase the number of machine translation users, particularly in controlled language environments (e.g. legal, medical reports, manuals etc.).

References

Abuelyaman, E., Rahmatallah, L., Mukhtar, W., & Elagabani, M. (2015). Machine Translation of Arabic language: Challenges and keys. *Proceedings - International Conference on Intelligent Systems, Modelling and Simulation, ISMS*, *2015-Septe*, 111–116. https://doi.org/10.1109/ISMS.2014.25

Abufardeh, S., & Magel, K. (2008). Software Localization : The Challenging Aspects of Arabic to the Localization Process (Arabization). *Proceedings of the IASTED International Conference on Software Engineering*, (September 2014), 275–279.

Akhrameev, S. (2016). *Large number of CAT tools – is it a problem or a solution?* Retrieved from https://theopenmic.co/large-number-of-cat-tools-is-it-a-problem-or-a-solution/

Al-jarf, P. R. (2017). Technology Integration in Translator Training in Saudi Arabia. *International Journal of Research in Engineering and Social Sciences*, *07*(3), 1–7.

Al-Sughaiyer, I. A., & Al-Kharashi, I. A. (2004). Arabic Morphological Analysis Techniques: A Comprehensive Survey. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*. Retrieved from http://www.abdelali.net/ref/AMA_Survey.pdf

Al-Taani, A. T., Al-Awad, N., & Abu-Salem, H. (2011). An Adaptive Parser for Arabic Language Processing. *Article in International Journal of Computer Processing Of Languages*, *23*(1), 67–80. https://doi.org/10.1142/S1793840611002218

Alansary, S. (2011). Interlingua-based Machine Translation Systems: UNL versus Other Interlinguas. In *11th International Conference on Language Engineering, Ain Shams*

University. Retrieved from https://www.bibalex.org/isis/UploadedFiles/Publications/Cairo2011a_1.pdf

Alexopoulou, T., Doron, E., & Heycock, C. (2003). *Broad Subjects and Clitic Left Dislocation*. Retrieved from http://pluto.huji.ac.il/~edit/edit/peri.pdf

Alghamdi, M., & Teahan, W. (2018). Experimental evaluation of Arabic OCR systems. *PSU Research Review*, *1*(3), 229–241. https://doi.org/10.1108/PRR-05-2017-0026

Ali, I. H., & Mnasri, Z. (2016). Statistical Analysis of Arabic Prosody. *Statistical Language and Speech Processing: 4th International Conference*. Retrieved from https://books.google.com/books?id=1YggDQAAQBAJ&pg=PA59&lpg=PA59&dq=Research+on+written+Arabic+language+processing+started+in+the+1970s&source=bl&ots=VoEY4ua_3Q&sig=HVJRaN6xr6AqRrjWyTTnBq4j-cA&hl=en&sa=X&ved=0ahUKEwjq0PaTmsXZAhURy1kKHeUhBCcQ6AEIMzAD#v=onep

Allen, J. (2003). Post-editing. In H. Somers (Ed.), *Computers and Trnslation: A translators guide*.

Almahairi, A., Cho, K., Habash, N., & Courville, A. (2016). First Result on Arabic Neural Machine Translation. *Arxiv*. Retrieved from http://arxiv.org/abs/1606.02680

Almutawa, F., & Izwaini, S. (2015). Machine Translation in the Arab World : Saudi Arabia as a Case Study. *Trans-Kom. Wissenschaftliche Zeitschrift Für Translation Und Kommunikation*, *8*(2), 382–414.

Alotaibi, H. (2014). Teaching CAT Tools to Translation Students: an Examination of Their Expectations and Attitudes. *Arab World English Journal*, (3), 75–87.

Alqarni, M. (2015). *THE MORPHOSYNTAX OF NUMERAL-NOUN CONSTRUCTIONS IN MODERN STANDARD ARABIC By MUTEB ALQARNI A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY*. University of Florida.

AlQinai, J. (2008). Mediating punctuation in English Arabic translation. *Journal of Applied Linguistics*, *5*(1), 5–29. https://doi.org/10.1558/japl.v5i1.5

Alqudsi, A., Omar, N., & Shaker, K. (2012). Arabic machine translation: a survey. *Artificial Intelligence Review*, *42*(4), 549–572. https://doi.org/10.1007/s10462-012-9351-1

Alramadan, R. (2017). *Arab Women's Representation in Arab Women's Writing and Their Translation* (Kent State University). Retrieved from https://etd.ohiolink.edu/

Alsaket, A. J., & ◌Ab Aziz, M. J. (2014). ARABIC-MALAY MACHINE TRANSLATION USING RULE-BASED APPROACH. *Journal of Computer Science*, *10*(6), 1062–1068. https://doi.org/10.3844/jcssp.2014.1062.1068

Alzand, A., & Ibrahim, R. (2015). Diacritics of Arabic Natural Language Processing (ANLP) and its quality assessment. *IEOM 2015 - 5th International Conference on Industrial Engineering and Operations Management, Proceeding*, 227–231. https://doi.org/10.1109/IEOM.2015.7093716

Aoun, J., Benmamoun, E., & Choueiri, L. (2010). *The Syntax of Arabic*. Retrieved from https://books.google.com/books?hl=en&lr=&id=Q4MhAwAAQBAJ&oi=fnd&pg=PR9&dq=the+syntax+of+arabic+aoun&ots=R3UW0pk7jC&sig=RqZ6EJanKWVtrcsf

8Yr1aGckDoQ#v=onepage&q=the syntax of arabic aoun&f=false

Armstrong, N., & Mackenzie, L. (2013). *Standardization, Ideology and Linguistics* (New York). Retrieved from https://books.google.com/books?hl=en&lr=&id=rryQdSGEb7oC&oi=fnd&pg=PP2&dq=Standardization,+ideology+and+linguistics&ots=_YGf3yi6KU&sig=L0FrzO6JWLxHWMDMc9CIGRkkcHU#v=onepage&q=Standardization%2C ideology and linguistics&f=false

Attia, M. A. (2007). *Arabic Tokenization System*. Retrieved from http://www.aclweb.org/anthology/W07-0809

Attia, M. A. (2008). *Handling Arabic Morphological and Syntactic Ambiguity within the LFG Framework with a View to Machine Translation*. Retrieved from http://attiaspace.com/Publications/Attia-PhD-Thesis.pdf

Austin, P. K. (2001). *Lexical functional grammar*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.12.3985&rep=rep1&type=pdf

Bahdanau, D., Cho, K., & Bengio, Y. (2014). *Neural Machine Translation by Jointly Learning to Align and Translate*. 1–15. https://doi.org/10.1146/annurev.neuro.26.041002.131047

Beeby Lonsdale, A. (2001). Directionality. In M. Baker & G. Saldanha (Eds.), *Routledge Encyclopedia of Translation Studies* (Second, pp. 63–67). Routledge.

Boualem, M. (2003). *Session on Arabic Language Processing*. Retrieved from https://www.mail-archive.com/mt-list@eamt.org/msg00612.html

Boudchiche, M., Mazroui, A., Ould Abdallahi Ould Bebah, M., Lakhouaja, A., & Boudlal, A. (2017). AlKhalil Morpho Sys 2: A robust Arabic morpho-syntactic analyzer. *Journal of King Saud University - Computer and Information Sciences*, *29*(2), 141–146. https://doi.org/10.1016/j.jksuci.2016.05.002

Boudlal, A., Lakhouaja, A., Mazroui, A., Meziane, A., Ould Abdallahi Ould Bebah, M., & Shoul, M. (2010). Alkhalil Morpho SYS1: A Morphosyntactic Analysis System for Arabic Texts. *International Arab Conference on Information Technology*, (June), 1–6. Retrieved from http://www.itpapers.info/acit10/Papers/f653.pdf

Bowker, L. (2002). *Computer-aided Translation Technology: A Practical Introduction - Lynne Bowker - Google Books*. Retrieved from https://books.google.com/books/about/Computer_aided_Translation_Technology.html?id=ly29-mc6dO0C&printsec=frontcover&source=kp_read_button#v=onepage&q&f=false

Bowker, L. (2003). Terminology tools for translators. *Computers and Translation: A Translator's Guide*, 49–65.

Bowker, L. (2005). Productivity vs Quality? A Pilot Study on the Impact of Translation Memory Systems. *Localization Focus*, 13–20. Retrieved from https://www.localisation.ie/sites/default/files/publications/Vol4_1Bowker.pdf

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Breikaa, Y. (2016). *The Major Problems that Face English – Arabic Translators while Using CAT Tools*. Retrieved from https://www.academia.edu

Brown, R. D. (1996). *Example-Based Machine Translation in the Pangloss System*. Retrieved from http://delivery.acm.org/10.1145/1000000/992660/p169-brown.pdf?ip=173.88.250.91&id=992660&acc=OPEN&key=4D4702B0C3E38B35.4D4702B0C3E38B35.4D4702B0C3E38B35.6D218144511F3437&__acm__=1520054191_757e0997c5e3a96d3793a95136aa80b3

Campbell, S. (1998). *Translation into the second language*. Retrieved from https://books.google.com/books/about/Translation_Into_the_Second_Language.html?id=SExiAAAAMAAJ

Campbell, S. (2000). Critical Structures in the Evaluation of Translations from Arabic into English as a Second Language. *The Translator*, *6*(2), 211–229. https://doi.org/10.1080/13556509.2000.10799066

Cavalli-Sforza, V., & Phillips, A. B. (2012). Using Morphology to Improve Example-Based Machine Translation: The Case of Arabic-to-English Translation. In A. Soudi, A. Farghaly, G. Neumann, & R. Zbib (Eds.), *Challenges for Arabic Machine Translation* (pp. 23–48). John Benjamins Publishing Company.

Charoenpornsawat, P., Sornlertlamvanich, V., & Charoenporn, T. (2002). Improving translation quality of rule-based machine translation. *COLING-02 on Machine Translation in Asia -*, *16*, 1–6. https://doi.org/10.3115/1118794.1118799

Chennoufi, A., & Mazroui, A. (2017). *Morphological, syntactic and diacritics rules for automatic diacritization of Arabic sentences*. https://doi.org/10.1016/j.jksuci.2016.06.004

Craciunescu, O., Salas, C. G., & Keeffe, S. S. O. (2004). Machine Translation and

Computer - Assisted Translation : a New Way of Translating ? *Translation Journal*, *8*(3), 1–10.

Devlin, J., Zbib, R., Huang, Z., Lamar, T., Schwartz, R., & Makhoul, J. (2014). Fast and Robust Neural Network Joint Models for Statistical Machine Translation. *Association for Computational Linguistics Conference (ACL)*, *17*, 1370–1380. Retrieved from http://acl2014.org/acl2014/P14-1/pdf/P14-1129.pdf

Diab, M., Hacioglu, K., & Jurafsky, D. (2007). Automatic Processing of Modern Standard Arabic Text. In A. Soudi, A. Bosch, & G. Neumann (Eds.), *Arabic Computational Morphology: Knowledge-based and Methods* (pp. 159–180). Dordrecht, The Netherlands: Springer.

El-Mahallawy, M. (2008). *A LARGE SCALE HMM-BASED OMNI FONT-WRITTEN OCR SYSTEM FOR CURSIVE SCRIPTS* (Cairo University). Retrieved from http://www.medar.info/Archive/PhD-Mahallaway_Apr2008.pdf

Elkhafaifi, H. M. (2002). Arabic language planning in the age of globalization. *Language Problems and Language Planning*, *26*(3), 253–269. https://doi.org/10.1075/lplp.26.3.03elk

Elsherif, H. M., & Soomro, T. R. (2017). PERSPECTIVES OF ARABIC MACHINE TRANSLATION. *Journal of Engineering Science and Technology*, *12*(9), 2315–2332. Retrieved from http://jestec.taylors.edu.my/Vol 12 issue 9 September 2017/12_9_2.pdf

Fakhr, M. W. (2011). *Arabic Optical Character Recognition (OCR) Systems Overview*. https://doi.org/10.13140/2.1.3898.3682

Farghaly, A. (2010a). Arabic Machine Translation : A Developmental Perspective. *International Journal on Information and Communication Technologies*, *3*(3), 3–10.

Farghaly, A. (2010b). *The Arabic Language, Arabic Linguistics and Arabic Computational Linguistics*. Retrieved from https://www.researchgate.net/publication/235478761

Farghaly, A., & Shaalan, K. (2010). Arabic Natural Language Processing: Challenges and Solutions. In *ACM Trans. Asian Language Information Processing* (Vol. 8). Retrieved from http://doi.acm.org/10.1145/

Fatani, A. (2006). *Globalization and the Translation Industry in Saudi Arabia: Factors Inhibiting the Use and Integration of Translation Technology into Mainstream Coursework*. Jeddah.

Fatani, A. (2009). The State of the Translation Industry in Saudi Arabia. *Translation Journal*, *13*(4). Retrieved from http://translationjournal.net/journal/50saudi.htm

Federico, M., Cattelan, A., & Trombetti, M. (2012). Measuring user productivity in machine translation enhanced computer assisted translation. *… for Machine Translation …*. Retrieved from http://amta2012.amtaweb.org/AMTA2012Files/papers/123.pdf

Ford, D. (2016). *Can You Trust a TM ? Results of An Experiment Conducted in November 2015 and August 2016 with Students and Professional Translators* . (November 2015), 69–80.

Furuse, O., & Iida, H. (1992). *An Example-Based Method for Transfer-Driven Machine Translation*. Retrieved from http://mt-archive.info/TMI-1992-Furuse.pdf

García, I. (2006). *Translators on translation memories: a blessing or a curse?* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.543.6222&rep=rep1&type=pdf

Gil, J. R., & Pym, A. (2006). Technology and translation (a pedagogical overview). In *Translation Technology and its Teaching (with much mention of localization)* . Retrieved from http://www.intercultural.urv.cat/media/upload/domain_317/arxius/Technology/translationtechnology.pdf#page=13

Greene, J. (2015). Preserving Distinctions Within the Multimethod and Mixed Methods Research Merger. In S. N. Hesse-Biber & R. B. Johnson (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. https://doi.org/10.1093/oxfordhb/9780199933624.013.37

Habash, N., Dorr, B., & Monz, C. (2006). Challenges in building an Arabic-English GHMT system with SMT components. *Proceedings of the 11th Annual Conference of the European Association for Machine Translation (EAMT-2006)*, 56–65.

Habash, N. Y. (2010). Introduction to Arabic Natural Language Processing. *Synthesis Lectures on Human Language Technologies*, *3*(1), 1–187. https://doi.org/10.2200/S00277ED1V01Y201008HLT010

Habash, Nizar, & Rambow, O. (2007). *Arabic Diacritization through Full Morphological Tagging*. 53–56. Retrieved from http://www.aclweb.org/anthology/N07-2014

Habash, Nizar, & Sadat, F. (2012). Arabic preprocessing for Statistical Machine

Translation: Schemes, techniques and combinations. In A. Soudi, A. Farghaly, G. Neumann, & R. Zbib (Eds.), *Challenges for Arabic Machine Translation* (pp. 73–94). John Benjamins Publishing Company.

Hailat, T., Al-Kabi, M. N., Alsmadi, I. M., & Al-Shawakfa, E. (2013). Evaluating English to Arabic machine translators. *2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, AEECT 2013*, *4*(1), 66–73. https://doi.org/10.1109/AEECT.2013.6716439

Hargrave, J., & Savourel, Y. (1997). *Machine assisted translation tools*. Retrieved from https://patents.google.com/patent/US5724593A/en

Hassan, H. (2012). Lexical syntax for Arabic SMT. In A. Soudi, A. Farghaly, G. Neumann, & R. Zbib (Eds.), *Challenges for Arabic Machine Translation* (pp. 109–134). John Benjamins Publishing Company.

Hesham, A. M., Abdou, S., Badr, A. A., Rashwn, M. A., & Al-barhamtoshy, H. M. (2016). A Zone Classification Approach for Arabic Documents using Hybrid Features. *(IJACSA) International Journal of Advanced Computer Science and Applications*, *7*(7), 158–162. https://doi.org/10.14569/IJACSA.2016.070722

Hutchins, J. (1986). *Machine Translation: Past, Present, Future*. Chichester: Ellis Horwood.

Hutchins, J. (1996). ALPAC: the (In)famous Report. *MT News International*, (14), 9–12.

Hutchins, J. (1997). *Translation Technology and the Translator*. 113–120. Retrieved from http://hutchinsweb.me.uk/ITI-1997.pdf

Hutchins, J. (1998). The Origins of the Translator ’ s Workstation. *Machine Translation*,

13(4 (1998)), 287–307. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.14.7374&rep=rep1&type=pdf

Hutchins, J. (2003a). Machine Translation and Computer-based Translation Tools: What's Available and How It's Used. *A New Spectrum of Translation Studies*, (March), 20.

Hutchins, J. (2003b). The Development and Use of Machine Translation Systems and Computer-based Translation Tools. *INTERNATIONAL JOURNAL OF TRANSLATION*, *15*(1). Retrieved from http://ourworld.compuserve.com/homepages/WJHutchins

Hutchins, J., & Somers, H. L. (1992). An Introduction to Machine Translation. In *psychotransling.ucoz.com* (Vol. 362). https://doi.org/10.2307/3721978

Ibnulyemen. (2017). history of Arabic diacritics and dotting. Retrieved August 2, 2018, from https://blogs.transparent.com/arabic/the-beginning-of-dotting-and-diacritics-in-arabic/

Ibrahim, R., & Aharon-Peretz, J. (2005). Is Literary Arabic a Second Language for Native Arab Speakers?: Evidence from Semantic Priming Study. *Journal of Psycholinguistic Research*, *34*(1). https://doi.org/10.1007/s10936-005-3631-8

Ive, J. (2017). *Towards a Better Human-Machine Collaboration in Statistical Translation : Example of Systematic Medical Reviews*.

Izwaini, S. (2006). Problems of Arabic Machine Translation : Evaluation of Three Systems. *Proceedings of the International Conference on the Challenge of Arabic*

for NLP/MT, 118–148.

Jia, Y., Carl, M., & Wang, X. (2019). How does the post-editing of neural machine translation compare with from-scratch translation? A product and process study. *The Journal of Specialised Translation*, (31). Retrieved from http://www.jostrans.org/issue31/art_jia.pdf

Kadhim, K. A., Habeeb, L. S., Sapar, A. A., Hussin, Z., & Abdullah, M. (2013). An Evaluation of Online Machine Translation of Arabic into English News Headlines: Implications on Students' Learning Purposes. *Turkish Online Journal of Educational Technology*, *12*(2), 39–50.

Kalchbrenner, N., & Blunsom, P. (2013). Recurrent Continuous Translation Models. *Emnlp*, (October), 1700–1709. https://doi.org/10.1146/annurev.neuro.26.041002.131047

Kay, M. (1980). The Proper Place of Men and Machines in Language Translation. *Machine Translation*, *11*(1–2), 1–20. https://doi.org/10.1023/A:1007911416676

Khemakhem, T., Jamoussi, S., & Ben Hamadou, A. (2013). Integrating morpho-syntactic features in English-Arabic statistical machine translation. *Proceedings of the Second Workshop on Hybrid Approaches to Translation, ACL 2013*, 74–81.

Kiraly, D. (2013). Towards a View of Translator Competence as an Emergent Phenomenon: Thinking Outside the Box(es) in Translator Education | University of Westminster. In D. Kiraly, S. Hansen-Schirra, & K. Maksymski (Eds.), *New Prospects and Perspectives for Educating Language Mediators* (pp. 197–224). Narr Verlag.

Koehn, P. (2010). Statistical Machine Translation. *Cambridge: Cambridge University Press*, 433.

Krings, H. (2001). *Repairing texts: investigations of machine translation post-editing processes*. Kent State University Press.

Lagarda, A. L., Ortiz-Martínez, D., Alabau, V., & Casacuberta, F. (2015). Translating without In-domain Corpus: Machine Translation Post-Editing with Online Learning Techniques. *Computer Speech and Language*, *32*(1), 109–134. https://doi.org/10.1016/j.csl.2014.10.004

Lagoudaki, E. (2006). Translation Memories Survey 2006: Users' perceptions around TM use. In *Translating and the Computer* (Vol. 28). Retrieved from http://mt-archive.info/Aslib-2006-Lagoudaki.pdf

Lagoudaki, E. (2008). The Value of Machine Translation for the Professional Translator. *8th AMTA Conference*, (October), 21–25.

LeBlanc, M. (2013). Translators on translation memory (TM). Results of an ethnographic study in three translation services and agencies. *Translation & Interpreting*, *5*(2), 1–13. https://doi.org/10.12807/T&I.V5I2.228

Lee, J., & Liao, P. (2011). A Comparative Study of Human Translation and Machine Translation with Post-editing. *Compilation & Translation Review, 4.2*.

Lopez, A., & Post, M. (2013). Beyond bitext : Five open problems in machine translation. *20 Years of Bitext Workshop in EMNLP-13*, 2–4.

Maamouri, M. (1998). Language Education and Human Development Arabic diglossia and its impact on the quality of education in the Arab region PRELIMINARY

COPY. *The Mediterranean Development Forum*. Retrieved from www.literacyonline.org

Mahfouz, I. (2018). *Attitudes to CAT Tools: Application on Egyptian Translation Students and Professionals*. (4), 69–83.

Mahmoud, M., Shquier, A., & Al-howiti, K. M. (2014). *Fully Automated Arabic to English Machine Translation System : Transfer-based approach of AE-TBMT* .

Mcbride, C. (2009). *Translation Memory System: An Analysis of Translator's Attitudes and Opinions* (University of Ottawa). Retrieved from https://ruor.uottawa.ca/bitstream/10393/28404/1/MR61311.PDF

McCrum-Gardner, E. (2008). Which is the correct statistical test to use? *British Journal of Oral and Maxillofacial Surgery*, *46*(1), 38–41. https://doi.org/10.1016/j.bjoms.2007.09.002

Meister, L. (2018). On methodology: How mixed methods research can contribute to translation studies. *Translation Studies*, *11*(1), 66–83. https://doi.org/10.1080/14781700.2017.1374206

Melby, A. (1979). *ITS : Interactive Translation System*. *5*(1), 234–241. Retrieved from file:///Users/SK/Documents/MyResearch/Library/articles/C80-1064.pdf

Melby, A. (1981). Translators and Machines - Can they Cooperate? *Meta: Journal Des Traducteurs*, *26*(1), 23. https://doi.org/10.7202/003619ar

Melby, A. (1982). Multi-Level Translation Aids in a Distributed system. *Proceedings of the Ninth International Conference on Computational Linguistics*, 215–220. Retrieved from http://ww2.cs.mu.oz.au/acl/C/C82/C82-1034.pdf

Melby, A. (1984). Machine Translation with Post Editing versus a Three-Level Integrated Translator Aid System. In I.D.K. Kelly (Ed.), *Progress in Machine Translation: Natural Language and Personal Computers* (pp. 119–125). Wilmslow, UK: Sigma Press.

Melby, A. (1992). The Translator Workstation. In John Newton (Ed.), *Computers in translation A practical appraisal* (pp. 147–165). https://doi.org/10.1075/ata.vii.14mel

Microsoft Translator. (2016). Microsoft helps translate your Arabic conversations face-to-face or across the globe - Microsoft Translator Blog. Retrieved June 5, 2019, from Microsoft Translator Blog website: https://www.microsoft.com/en-us/translator/blog/2016/03/08/microsoft-helps-translate-your-arabic-conversations-face-to-face-or-across-the-globe/

Mishal, A. (2015). Importance of Arabic Language Content. Retrieved July 20, 2019, from https://www.ccjk.com/importance-arabic-language-content/

Moorkens, J., & O'Brien, S. (2013). User attitudes to the post-editing interface. *Proceedings of Machine Translation Summit XIV: Second Workshop on Post-Editing Technology and Practice, Nice, France*, 19–25.

Moorkens, J., & O'Brien, S. (2017). *Assessing User Interface Needs of Post-Editors of Machine Translation Assessing User Interface Needs of Post-Editors of Machine Translation Joss Moorkens & Sharon O ' Brien , ADAPT / Dublin City University*. (January).

Moujaes, M. (2016). *A comparative study of the Quality Assurance feature in SDL*

Trados and Multitrans in the context of the World Intellectual Property Organization A Comparative Study of the Quality Assurance Feature in SDL Trados and Multitrans in the Context of the World.

Nashwan, F. M. A., Rashwan, M. A. A., Al-barhamtoshy, H. M., Abdou, S. M., & Moussa, A. M. (2017). *A Holistic Technique for an Arabic OCR System*. 1–11. https://doi.org/10.3390/jimaging4010006

Oettinger, A. G. (2000). Machine translation at harvard. In J. Hutchins (Ed.), *Early Years in Machine Translation* (pp. 73–86). John Benjamins Publishing Company.

Owens, J. (2007). A Linguistic History of Arabic. In *A Linguistic History of Arabic*. https://doi.org/10.1093/acprof:oso/9780199290826.001.0001

Peng, L. (2013). A Survey of Machine Translation Methods. *TELKOMNIKA*, *11*(12). Retrieved from https://pdfs.semanticscholar.org/aded/ce21e5f7015a9a189a406f04ea50575c0c6a.pdf

Phillips, Aaron, Cavalli-sforza, V., & Brown, R. (2006). Improving Example Based Machine Translation Through Morphological Generalization and Adaptation. *Machine Translation*. Retrieved from http://www.mt-archive.info/MTS-2007-Phillips.pdf

Phillips, Aron, Cavalli-Sforza, V., & Brown, R. D. (2007). Improving example based machine translation through morphological generalization and adaptation. *... 9th Machine Translation ...*, 369–375. Retrieved from http://mt-archive.info/MTS-2007-Phillips.pdf

Piotrovskij, R. G. (2000). MT In The Former USSR And In The Newly Independent

States (NIS). Prehistory, Romantic Era, Prosaic Time. In J. Hutchins (Ed.), *Early Years in Machine Translation* (pp. 233–242).

Quaranta, B. (2007). *Arabic and Computer-Aided Translation : an integrated approach*.

Riesa, J., Mohit, B., Knight, K., & Marcu, D. (2006). Building an English-Iraqi Arabic Machine Translation System for Spoken Utterances with Limited Resources. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2*, 745–748.

Sadat, F. (2013). Towards a Hybrid Rule-based and Statistical Arabic-French Machine Translation System. *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, (September), 579–583. Retrieved from http://aclweb.org/anthology/R13-1076

Saldanha, G., & O'Brien, S. (2014). *Research Methodologies in Translation Studies*. https://doi.org/10.4324/9781315760100

Salem, Y. (2009). A Generic Framework for Arabic to English Machine Translation of Simplex Sentences Using the Role and Reference Grammar linguistic model. *Unpublished Master" s Thesis, Institute of Technology Blanchardstown, Dublin, Ireland*. Retrieved from http://www.linguistics.buffalo.edu/people/faculty/vanvalin/rrg/Yasser_Salem_MSc_thesis.pdf

Salem, Y., Hensman, A., & Nolan, B. (2008a). Implementing Arabic-to-English Machine Translation using the Role and Reference Grammar Linguistic Model. *Machine Translation*.

Salem, Y., Hensman, A., & Nolan, B. (2008b). Towards Arabic to English Machine Translation. *ITB Journal*, (17), 20–31. Retrieved from http://arrow.dit.ie/itbinfoart/1/

Saraireh, M. A. T. (2001). Inconsistency in Technical Terminology: A Problem for Standardization in Arabic. *Fédération Internationale Des Traducteurs*. Retrieved from https://www.ingentaconnect.com/content/jbp/bab/2001/00000047/00000001/art00002?crawler=true&mimetype=application/pdf&casa_token=ozzkpBh6XnkAAAAA:P4mpmyZRpE54gljDRvQ0GThNRicjRUuvp2gV4BJc5_Luv80w7Vkcvz59C5g2vne1E7zoAYYwchIN3fM

SDL Trados. (n.d.). The history of SDL's translation software. Retrieved March 19, 2018, from https://www.sdltrados.com/about/history.html

Sekino, K. (2015). An investigation of the relevance-theoretic approach to cognitive effort in translation and the post-editing process. In *Translation & Interpreting* (Vol. 7). Retrieved from http://trans-int.org/index.php/transint/article/viewFile/375/189

Shaalan, K., & Hossny, A. H. (2012). Automatic rule induction in Arabic to English machine translation framework. In *Challenges for Arabic Machine Translation* (pp. 135–154). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=mzh&AN=2012932939&site=ehost-live

Shannon, E. (2015). *Better Support for MS Word documents in SDL Trados Studio 2015*. Retrieved from https://www.sdltrados.com/blog/better-support-for-ms-word-documents-in-sdl-trados-studio-2015.html

Shilon, R., Wintner, S., Science, C., & Landman, F. (2011). Transfer-based Machine Translation between morphologically-rich and resource-poor languages : The case of Hebrew and Arabic MA thesis submitted by. *Computer*, (March).

Shirko, O., Omar, N., Arshad, H., & Albared, M. (2010). Machine translation of noun phrases from Arabic to english using transfer-based approach. *Journal of Computer Science*. https://doi.org/10.3844/jcssp.2010.350.356

Shreve, G. M., & Angelone, E. (2010). *Translation and Cognition* (G. M. Shreve & E. Angelone, Eds.). https://doi.org/10.1075/ata.xv

Sieny, M. E. (1987). Pan-Arab and international cooperation in technical terminology. *[Translating and the Computer 9.* Retrieved from http://www.mt-archive.info/Aslib-1987-Sieny.pdf

Soudi, A., Bosch, A., & Neumann, G. (2007). *Arabic Computational Morphology: Knowledge-based and Methods*. https://doi.org/10.1007/978-1-4020-6046-5_10

Soudi, A., Farghaly, A., Neumann, G., & Zbib, R. (2012). *Challenges for Arabic Machine Translation.* Amsterdam / Philadelphia: John Benjamins Publishing Company.

Stetina, J., & Nagao, M. (1997). *Corpus Based PP Attachment Ambiguity Resolution 1 with a Semantic Dictionary II*. Retrieved from http://www.aclweb.org/anthology/W97-0109

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems (NIPS)*, 3104–3112. https://doi.org/10.1007/s10107-014-0839-0

Systran. (2018). What is Machine Translation? Rule Based vs. Statistical |. Retrieved March 3, 2018, from http://www.systransoft.com/systran/translation-technology/what-is-machine-translation/

Tabor, J. (2013). *CAT tool use by translators: what are they using?* Retrieved from https://prozcomblog.com/2013/03/28/cat-tool-use-by-translators-what-are-they-using/

Thawabteh, M. (2009). *Apropos translator training aggro: A case study of the Centre for Continuing Education*. Retrieved from https://www.jostrans.org/issue12/art_thawabteh.pdf

Thawabteh, M. (2013). The Intricacies of Translation Memory Tools: With Particular Reference to Arabic-English Translation. *Localisation Focus The …*, *12*(1), 79–90. Retrieved from http://works.bepress.com/mohammad_thawabteh/16/

Thurmond, V. A. (2001). The Point of Triangulation. *Journal of Nursing Scholarship*, *33*(3), 253–258. https://doi.org/10.1111/j.1547-5069.2001.00253.x

Toral, A. (2019). *Post-editese: an Exacerbated Translationese*. Retrieved from http://arxiv.org/abs/1907.00900

Toral, A., & Sánchez-Cartagena, V. M. (2017). *A Multifaceted Evaluation of Neural versus Phrase-Based Machine Translation for 9 Language Directions*. (i). Retrieved from http://arxiv.org/abs/1701.02901

Vasconcellos, M. (2000). The Georgetown Project and Léon Dostert: Recollection of Young Assistant. In J. Hutchins (Ed.), *Early Years in Machine Translation* (pp. 87–96).

Vasconcellos, M. (2001). Terminology and Machine Translation. In S. E. Wright & G. Budin (Eds.), *Handbook of Terminology Management : Application-Oriented Terminology Management* (Vol. 2, pp. 697–724). John Benjamins Publishing Company.

Wright, S. E., & Wright, L. D. (1997). Terminology Management for Technical Translation. In S. E. Wright & G. Budin (Eds.), *Handbook of Terminology Management*. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Yngve, V. H. (2000). EARLY RESEARCH AT MIT. In J. Hutchins (Ed.), *Early Years in Machine Translation* (pp. 39–72).

Zbib, R., & Badr, I. (2012). Preprocessing for English-to-Arabic Statistical Machine Translation. In A. Soudi, A. Farghaly, G. Neumann, & R. Zbib (Eds.), *Challenges for Arabic Machine Translation* (pp. 95–108). John Benjamins Publishing Company.

Zitouni, I., Sorensen, J. S., & Sarikaya, R. (2006). *Maximum Entropy Based Restoration of Arabic Diacritics*. 577–584. Retrieved from https://aclanthology.info/pdf/P/P06/P06-1073.pdf

Zughoul, M. R., & Abu-Alshaar, A. M. (2005). English/Arabic/English Machine Translation: A Historical Perspective. *Meta: Journal Des Traducteurs, 50*(3), 1022. https://doi.org/10.7202/011612ar

Appendices

Appendix A: Online Survey

Please answer the following questions:

General Background Information

1) Please choose one: I'm currently _____

□ Working as a translator for a LSP.

□ Working in a profession that requires translation tasks.

□ A freelance translator.

□ A freelance translator and a graduate student at translation program.

□ Graduate student in translation program but don't work currently as a translator.

2) How many years of translation experience do you have?

□ None □ 1-5 years □ 6-9 years □ 10+ years

3) What are your academic qualifications, if any?

□ PhD completed/ongoing □ MA completed/ongoing □ BA completed/ongoing

4) If your studies are ongoing, which year are you in?

□ 1st year □ 2nd year □ 3rd year □ 4th year □ 5th year

□ Other, please specify: _____

Background Question about CAT tools:

5) How familiar are you with the current computer-assisted translation tools?

□ Extremely familiar □ Somewhat familiar

□ Slightly familiar □ Not familiar at all

6) From you experience, have you encountered issues and challenges in the use of computer-assisted translation tools for Arabic language?

□ Yes □ No □ Does not apply

7) If yes, what are the issue you faced while using the tools with Arabic?

……

8) Which computer-assisted translation tool you have been using in your translation work;

□ SDL Trados □ MemoQ □ Wordfast □ Other…….

9) How do you rate your knowledge of any computer-assisted translation tool you have used? Rate from one to ten, where 1 is very poor and 10 is excellent:

…….

10) Based on your experience with the tool, how would you rate your satisfaction level of the tools?

□ Very Satisfied □ Satisfied □ OK □ Dissatisfied

□ Very dissatisfied

11) Do you have specific suggestions for improving the features of any tool you have not been fully satisfied with?

……
……
……
……

12) In your opinion, how important is the use of CAT tools by Arabic translators?

□ Very important □ Somewhat important

□ Slightly important □ Not at all important □ Don't know

Background Question about machine translation tools:

13) How familiar are you with machine translation tools for Arabic?

□ Extremely familiar □ Somewhat familiar

□ Slightly familiar □ Not familiar at all

14) In your opinion, how important is the use of machine translation tools by Arabic translators?

□ Very important □ Somewhat important

□ Slightly important □ Not at all important □ Don't know

15) Machine translation in general is useful for Arabic language translators;

□ True □ False □ Don't know

16) Some translation memory systems integrate MT suggestions when there is no TM for a given translation segment. Do you anticipate that the MT suggestions will increase Arabic translators' productivity?

□ Strongly agree □ Somewhat agree □ Neither agree or disagree

□ Somewhat disagree □ Strongly disagree □ Don't know

17) My evaluation of machine translation for Arabic language is:

□ Positive □ Neutral □ Negative □ Don't know

18) What limitation do you think the new computer-assisted translation tools have for Arabic language?

……

Appendix B: Interview Questions

1) **How was your experience with the tools during the task?**

……
……
……
……
……
……
……

2) **How can you rate the outputs of the tools during your translation task? Out of 10?**

……
……
……
……
……
……

3) **Do you think the tools helped your productivity, and how?**

……
……
……
……
……

4) **What limitations do you think the tools have?**

……
……
……
……
……
……

5) **What difficulties did you face during your translation task?**

……
……
……
……
……
……
……

6) **Have you changed your general assessment of the tools after this experiment?**

……
……
……
……
……
……

7) **Which translation task was harder and why?**

……
……
……
……
……
……

8) **Which segment in each translation task was difficult for you? And why?**

……
……
……
……
……
……

9) **Which segment in each translation task was easier for you? And why?**

……
……

……
……
……
……

10) Did you have difficulties with segmentation or punctuations in both Arabic and English translations?

……
……
……
……
……
……

11) What suggestions do you have to improve the tools use?

……
……
……
……
……
……

12) Is there any other concern you like to share about your experience with the tools?

……
……
……
……
……
……

Appendix C: Translation Texts

Arabic to English Translation text (task 1)

كيفية التعرف على علامات سرطان الفم

تمثل سرطانات الفم والحنجرة 2% تقريبًا من جميع السرطانات التي تشخص كل عام في الولايات المتحدة الأمريكية، فإن الاكتشاف المبكر لسرطانات الفم وعلاجها في وقت مناسب أمرٌ مهمٌ لرفع احتمالات النجاة، فمعدل الحياة لخمس سنوات للمصابين بسرطان الفم الذين لم ينتشر لديهم المرض يبلغ 83% مثلًا بينما هو 32% فقط بعد انتشار السرطان إلى أجزاء الجسم الأخرى .رغم أن طبيب الأسنان أو طبيبك المعتاد مدربين على اكتشاف سرطانات الفم إلا أن التعرف على الأعراض بنفسك قد يسهل التشخيص المبكر، والعلاج في الوقت المناسب، وكلما زاد وعيك كان أفضل.

The suggested translation:

How to Recognize Signs of Oral Cancer

Oral cancers of the mouth and throat account for about 2% of all cancers diagnosed each year in the U.S. Early detection and timely treatment of oral cancers is important because it greatly increases the chances of survival. For example, the five-year survival rate for those with oral cancer that hasn't spread is 83%, whereas it's only 32% once the cancer spreads to other parts of the body. Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis and more timely treatment. The more aware you are, the better.

English to Arabic Translation text (task 2)

How to Beat Anorexia

When an individual refuses to consume the amount of food and drink required to maintain a healthy body weight, has a distorted body image, and an intense fear of gaining weight, that person suffers from anorexia. Anorexia is an extremely dangerous eating disorder that can lead to severe dehydration, lowered blood pressure, bone density loss, and fainting among other consequences. Fortunately, most people who suffer from anorexia can beat it with the right combination of physical, psychological, and social therapy.

The suggested translation:

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي، ويكون له صورة مشوهة عن جسمه، وعنده خوف حاد من الزيادة في الوزن، هذا الشخص يعاني من مرض فقدان الشهية. مرض فقدان الشهية هو اضطراب أكل شديد الخطورة يمكن أن يؤدي إلى الجفاف الحاد، وانخفاض ضغط الدم، وفقدان كثافة العظام، وفقدان الوعي وسط العديد من العواقب الأخرى. لحسن الحظ، أغلب من يعانون من فقدان الشهية يمكن أن يتغلبوا عليها بالمزيج الصحيح من العلاج الجسدي، والنفسي، والاجتماعي.

Appendix D: Designed TM for Experiment

Designed TM for Arabic-English Translation Task

1-	exact	كيفية التعرف على علامات سرطان الفم	How to Recognize Signs of Oral Cancer
2-	Fuzzy	تمثل سرطانات الفم والحنجرة 2% تقريبًا من جميع السرطانات التي تشخص كل عام في الولايات المتحدة الأمريكية، فإن الاكتشاف المبكر لسرطانات الفم وعلاجها في وقت مناسب أمرٌ مهمٌ لرفع احتمالات النجاة، فمعدل الحياة لخمس سنوات للمصابين بسرطان الفم الذين لم ينتشر لديهم المرض يبلغ 83% مثلًا بينما هو 32% فقط بعد انتشار السرطان إلى أجزاء الجسم الأخرى	~~Oral cancers of the mouth and throat for about 2% of all cancers account diagnosed each year in the U.S. Early detection and timely treatment of oral cancers is important because it greatly increases the chances of survival. For example, the five-year survival rate for those with oral cancer that hasn't spread is 83%, whereas it's only 32% once the cancer spreads to other parts of the body.~~
3-	Fuzzy	رغم أن طبيب الأسنان أو طبيبك المعتاد مدربين على اكتشاف سرطانات الفم إلا أن التعرف على الأعراض بنفسك قد يسهل التشخيص المبكر ~~والعلاج في الوقت المناسب وكلما زاد وعيك كان أفضل.~~	Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis ~~and more timely treatment. The more aware you are, the better.~~

Designed TM for English-Arabic Translation Task

1	Exact	How to Beat Anorexia	كيفية التغلب على مرض فقدان الشهية
2	Fuzzy	When an individual refuses to consume the amount of food and drink required to maintain a healthy body weight, has a distorted body image, and an intense fear of gaining weight, that person suffers from anorexia.	عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي، ويكون له صورة مشوهة عن جسمه، وعنده خوف حاد من الزيادة في الوزن، ~~هذا الشخص يعاني من مرض فقدان الشهية.~~
3	Fuzzy	Anorexia is an extremely dangerous eating disorder that can lead to severe dehydration, lowered blood pressure, bone density loss, and fainting among other consequences.	مرض فقدان الشهية هو ~~اضطراب أكل شديد الخطورة يمكن أن يؤدي إلى الجفاف الحاد، وانخفاض ضغط الدم، وفقدان كثافة العظام، وفقدان الوعي وسط العديد من العواقب الأخرى~~
4	MT	Fortunately, most people who suffer from anorexia can beat it with the right combination of physical, psychological, and social therapy.	~~لحسن الحظ، أغلب من يعانون من فقدان الشهية يمكن أن يتغلبوا عليها بالمزيج الصحيح من العلاج البدني، والنفسي، والاجتماعي.~~

Appendix E: Consent Form for the Experiment

Informed Consent to Participate in a Research Study

Study Title: *(The Use of Computer-assisted translation Tools for Arabic Translation: User Evaluation, Issues, and Improvements.)*

Principal Investigator: *Dr. Sue Ellen Wright and Mohammad Alanazi*

You are being invited to participate in a research study. This consent form will provide you with information on the research project, what you will need to do, and the associated risks and benefits of the research. Your participation is voluntary. Please read this form carefully. It is important that you ask questions and fully understand the research in order to make an informed decision. You will receive a copy of this document to take with you.

Purpose:

The research project aims to explore the issues involved in the use of computer-assisted translation tools for Arabic and conduct an study to examine Arabic language translators' perspectives of these tools and what potential developments can be made to computer-assisted tools to meet the translators' needs.

Procedures

The research study will require the participants to perform the following two tasks:

1. Translate two texts in two direction: from Arabic to English and vice versa using SDL Trados 2017 translation. The translation task should not take more than 30 minutes to complete.

2. Participate in a short interview to discuss the task experience with the tool. The interview should not take more than 20-30 minutes.

Audio and Video Recording and Photography

The interview of this study will involve audio recording which will be transcribed for analysis. The audio recordings will be used exclusively for the academic purpose of this study and will be disposed of upon the completion of analysis.

Benefits

This research will not benefit you directly. However, your participation in this study will help us to better understand the issues and complications that concern the Arabic language translators while using computer-assisted translation tools, so recommendations can be made to develop the tools for Arabic language translators.

Risks and Discomforts

There are no anticipated risks beyond those encountered in everyday life.

Privacy and Confidentiality

Your study related information will be kept confidential within the limits of the law. Any identifying information will be kept in a secure location and only the researchers will have access to the data. Research participants will not be identified in any publication or presentation of research results; only aggregate data will be used. Any reference to a specific participant will be made using an anonymous name.

Compensation

For participating in this research, you will receive a 25-dollar gift card for Walmart or some other vendor. If you decide to stop participating before completing all required tasks, you are to be given 10-dollar gift card for Walmart or some other vendor.

Voluntary Participation

Taking part in this research study is entirely up to you. You may choose not to participate or you may discontinue your participation at any time without penalty or loss of benefits to which you are otherwise entitled.

Contact Information

If you have any questions or concerns about this research, you may contact *Mohammad Alanazi* at *malanaz6@kent.edu* or *Dr. Sue Ellen Wright* at *330-672-2150* or email at *swright@kent.edu*. This project has been approved by the Kent State University Institutional Review Board. If you have any questions about your rights as a research participant or complaints about the research, you may call the IRB at 330.672.2704.

Consent Statement and Signature

I have read this consent form and have had the opportunity to have my questions answered to my satisfaction. I voluntarily agree to participate in this study. I understand that a copy of this consent will be provided to me for future reference.

_____ _____

Participant Signature **Date**

Appendix F: Consent form for Online Survey

Informed Consent to Participate in a Research Study

Study Title: *(The Use of Computer-assisted translation Tools for Arabic Translation: User Evaluation, Issues, and Improvements.)*

Principal Investigator: *Dr. Sue Ellen Wright and Mohammad Alanazi*

You are being invited to participate in a research study. This consent form will provide you with information on the research project, what you will need to do, and the associated risks and benefits of the research. Your participation is voluntary. Please read this form carefully. It is important that you ask questions and fully understand the research in order to make an informed decision. You will receive a copy of this document to take with you.

Purpose:
The research project aims to explore the issues involved in the use of computer-assisted translation tools for Arabic and conduct an experimental study to examine Arabic language translators' perspectives of these tools and what potential developments can be made to computer-assisted tools to meet the translators' needs.

Procedures
The research study will require the participants to complete this online survey that contains 18 questions. The survey should not take more than 20-30 minutes to complete.

Benefits
This research will not benefit you directly. However, your participation in this study will help us to better understand the issues and complications that concern the Arabic language translators while using computer-assisted translation tools, so recommendations can be made to develop the tools for Arabic language translators.

Risks and Discomforts
There are no anticipated risks beyond those encountered in everyday life.

Privacy and Confidentiality
Your study related information will be kept confidential within the limits of the law. Any identifying information will be kept in a secure location and only the researchers will have access to the data. Research participants will not be identified in any publication or

presentation of research results; only aggregate data will be used. Any reference to a specific participant will be made using an anonymous name.

Compensation

For participating in this survey, you will be offered to enter a drawing of a 20-dollar ten gift cards for Amazon or some other online vendor.

Voluntary Participation

Taking part in this research study is entirely up to you. You may choose not to participate or you may discontinue your participation at any time without penalty or loss of benefits to which you are otherwise entitled.

Contact Information

If you have any questions or concerns about this research, you may contact *Mohammad Alanazi* at *malanaz6@kent.edu* or *Dr. Sue Ellen Wright* at *330-672-2150* or email at *swright@kent.edu*. This project has been approved by the Kent State University Institutional Review Board. If you have any questions about your rights as a research participant or complaints about the research, you may call the IRB at 330.672.2704.

Consent Statement and Signature

I have read this consent form and have had the opportunity to have my questions answered to my satisfaction. I voluntarily agree to participate in this study. I understand that a copy of this consent will be provided to me for future reference.

_____ _____

Participant Signature **Date**

Appendix G : Flyer

The Flyer

The Use of Computer-assisted Tools for Arabic Translation: User evaluation, Issues, and Improvements

Principal Investigator: Prof. Sue Ellen Wright

Co-Investigator: Mohammad Alanazi

Background:

The research project aims to explore the issues involved in the use of computer-assisted translation tools for Arabic and to conduct an experimental study to examine Arabic language translators' evaluation toward these tools and what potential developments can be made to computer-assisted tools to meet translators' needs.

Procedure:

You will be asked to perform the following tasks:

1. Translate two texts in two direction: from Arabic to English and vice versa using SDL Trados 2017 translation tool.

2. Participate in a short interview to discuss the task experience with the tool. The interview should not take more than 20 minutes.

Risks and Benefits:

There will be no risks for you as the participant since the participation will be voluntary. You can discontinue your participation at any time without any penalty. Any identifying information will be kept in a secure location and only the researchers will have access to the data. Participation in the project is in no way linked to your course participation or your grade in any course.

This research will not benefit you directly. However, your participation in this study will help us to better understand the issues and complications that concern the Arabic

language translators while using computer-assisted translation tools, so recommendations can be made to develop the tools for Arabic language translators.

Appendix H: Experiment Recruitment Script

Hello,

My name is Mohammad Alanazi and I am a Ph.D. candidate in translation studies at Kent State University. I am conducting a study that will help me to collect data for my dissertation. This study is being conducted by Prof. Sue Ellen Wright, principal investigator, and co-investigator, Mohammad Alanazi, and it has been approved by Kent State University Institutional Review Board.

I am currently looking for Arabic language translators to participate in this study. The study will require the participants to perform two tasks: Translate two texts in two direction: from Arabic to English and vice versa using SDL Trados 2017 translation, and participate in a short interview to discuss their task experience with the tool. All the tasks should take no longer than 60 minutes to complete.

The study aims to explore the issues involved in the use of computer-assisted translation tools for Arabic and conduct an experimental study to examine Arabic language translators' evaluation of these tools and what potential developments can be made to computer-assisted tools to meet the translators' needs. It will help to cast light on the issues and complications that concern the Arabic language translators while using computer-assisted translation tools, so recommendations can be made to develop the tools for Arabic language translators.

Your participation is extremely important to us, and we greatly appreciate you taking the time to share your experience by participating in the study. For participating in this research, you will receive a 25-dollar gift card for Walmart or some other vendor. If you

decide to stop participating before completing all required tasks, you are to be given 10-dollar gift card for Walmart or some other vendor.

If you are interested in participation, please read this consent form which will provide you with information on the research project, what you will need to do, and the associated risks and benefits of the research.

Appendix I: Survey Recruitment script

Hello,

My name is Mohammad Alanazi and I am a Ph.D. candidate in translation studies at Kent State University. I am conducting a study that will help me to collect data for my dissertation. This study is being conducted by Prof. Sue Ellen Wright, principal investigator, and co-investigator, Mohammad Alanazi, and it has been approved by Kent State University Institutional Review Board.

I am currently looking for Arabic language translators to participate in this study. The study will require the participants to complete the survey below. The survey should not take more than 20-30 minutes to complete.

The study aims to explore the issues involved in the use of computer-assisted translation tools for Arabic and to examine Arabic language translators' evaluation of these tools and what potential developments can be made to computer-assisted tools to meet the translators' needs. It will help to cast light on the issues and complications that concern the Arabic language translators while using computer-assisted translation tools, so recommendations can be made to develop the tools for Arabic language translators.

Your participation is extremely important to us, and we greatly appreciate you taking the time to share your experience by participating in the study. For participating in this research, you will be offered to enter a drawing of a 20-dollar ten gift cards for Amazon or some other vendor.

If you are interested in participation, please read the consent form in the beginning of the survey which will provide you with information on the research project, what you will need to do, and the associated risks and benefits of the research.

Appendix J: Interviews Transcription

Interview with Alpha

00:01 Speaker 1: Okay, let's start the interview with Alpha.

How was your experience with the tools during the task?

00:07 Alpha: Well, I enjoyed using Trados machine translation, embedded into Trados. This is the way I do my translation. So yeah.

How can you rate the outcomes of the tools during your translation task? Out of 10?

00:28 Alpha: It's not exemplary but it's helpful. I would say 7.

00:31 S1: Okay.

00:32 Alpha: Yeah.

Do you think the tools helped your productivity? And how?

00:37 Alpha: Sure.

00:39 S1: How?

00:41 Alpha: First of all, it provides you with solutions to the translation problems that appear in front of you. So you have solutions, and then something else you don't have to type so much. What was the question again?

01:03 S1: Did it help your productivity?

01:06 Alpha: Yes, the most important thing is that when I want to produce the target document, I don't have to worry much about the formatting issue, so that helps a lot, and saves a lot of time. I also like the dictionary suggestions that come with automatic machine translations, yes.

What limitations do you think the tools have?

01:37 Alpha: Of course, you have to understand that this is a kind of translation that conveys the meaning to you, but it's not... It doesn't look like... For example, when I translate from Arabic to English or English to Arabic, sometimes the translation doesn't look like... It looks like a translation. So, what I try is that to make it... So this is a limitation I understand because this is a machine. Other things, for example, in Arabic, it doesn't provide you with the accents. Sometimes, sometimes I feel it's still the machine translation misses parts of the translations, it's like it deletes some parts of it.

02:26 Alpha: I don't know where does that come from, when the paragraph is so long, or longer than... So, it provides you with a translation, but parts of it is like it's being cut off. And if you don't pay attention and you rely much on machine translation, you discover that you have left out parts of the original, of the source text not translated in the target text.

What difficulties did you face during your translation task?

03:01 Alpha: I didn't face any difficulties because I'm accustomed and used to using Trados as a translation tool, but I can envision that others who are not accustomed to using this tool might face issues in how to deal with the various functions.

Have you changed your general assessment of the tools after this experiment?

03:26 Alpha: About what? In comparison with what?

03:28 S1: With the tools in general. Have you changed it, after this experiment or...

03:35 Alpha: For me, no, it hasn't changed at all because I'm a user of these tools. And I actually call for the use of these tools.

Which translation task was harder and why?

03:57 Alpha: I don't have an answer for this question. None of them was harder than the other, but I would envision that maybe if I want to choose, I would say from Arabic into English was more difficult than from English into Arabic.

Which segment in each translation task was difficult for you? And why?

04:17 Alpha: Segment-wise, yes, there are segments. For example, in the Arabic to English, [04:24] "Surtan Alfm", the oral cavity cancer, the segment which was saying... It was talking about the rate of survival.

04:35 S1: Okay.

04:35 Alpha: Yeah.

04:36 S1: Segment number two.

04:37 Alpha: Yeah.

04:38 S1: Okay.

04:38 Alpha: Yeah, this was a little bit...

04:41 S1: And for the English to Arabic?

04:47 Alpha: For the English to Arabic? For the English to Arabic... Yeah, I think also segment number three.

04:58 S1: Okay.

05:00 Alpha: Yeah.

05:00 S1: For the segment number two, why do you think it's harder? Is it because it's a long segment, it's a long sentence or any other thing that you think why it's the hardest one? Do you think it's because it has a long segment?

05:14 Alpha: Yeah, you see, in Arabic, it's a very long sentence so you have to divide it into chunks, smaller sentences in English.

Which segment in each translation task was easier for you? And why?

05:35 Alpha: Yeah, the one with the least words, let's say, yes. Number one.

Did you have difficulties with segmentation or punctuations in both Arabic and English translations?

05:50 Alpha: I didn't see these, no.

05:54 S1: You discussed about the segment... Segmenting the Arabic...

05:57 Alpha: Segmentation, yes, I had to... From Arabic to English, I had to adjust and to make segments shorter for the English reader.

06:10 S1: But you didn't have it with the English to Arabic?

06:16 Alpha: From the English to Arabic...

06:16 S1: No problem with the segmentation.

06:18 Alpha: There was no issues, no.

06:21 S1: Okay. Punctuations, did you have any problem with punctuation?

06:26 Alpha: Punctuation? No.

06:27 S1: Comma?

06:29 Alpha: No.

06:29 S1: Spelling?

06:34 Alpha: Yeah, I had some issues, minor issues. Like S for R or R for S, plural and singular, that's all about...

What suggestions do you have to improve the tools use?

07:00 AA: No, I don't have any suggestions.

Is there any other concern you like to share about your experience with the tools?

07:04 AA: No, thank you.

07:05 S1: Thank you so much.

Interview with Beta

00:01 Speaker 1: Okay, how was your experience with the tools during these two tasks?

00:13 BETA: I think it was pretty easy. Most of the text was translated correctly. It was only a few minor mistakes.

00:21 S1: Okay.

00:21 BETA: That were easy to fix.

00:23 S1: Okay. And how can you rate the output of the tools during your translation tasks?

00:28 BETA: Like out of... Like a percentage?

00:30 S1: Yeah, the quality... The output, out of 10?

00:32 BETA: I would say it was like 70%.

00:36 S1: Good. Okay. Do you think the tools helped your productivity?

00:42 BETA: Definitely that... I finished both tasks in about 15 minutes, I would say.

00:48 S1: Yeah.

00:48 BETA: It would have taken me a much longer time looking up terms. At least, when I had to fix them now it was only a few gender issues, conjugations. Just like restructure the sentence... But if I started from scratch, it would have taken me longer. So, yeah, definitely that they... The tools helped.

01:08 S1: Okay. What limitations do you think the tools have from your experience and during the use of it...

01:14 BETA: I would say maybe the issue of consistency with terms because at some point one of the terms was translated in a certain way and then the same term in another sentence was translated differently. So unless I'm aware of like... Unless I manage my terms, I know that this is...

01:33 S1: This was in the same segment?

01:34 BETA: Not in the same segment.

01:35 S1: Okay.

01:36 BETA: In a different segment.

01:37 S1: Okay.

01:38 BETA: So, it was translated in a way... In the text it was like, "Oral cancer." But then it was like, "Mouth cancer." So, unless it is... Unless I know that... Unless...

01:53 S1: But you know that some segments are from translation memory and some segments are from MT?

01:58 BETA: Yes.

01:58 S1: So... So, maybe this is the difference...

02:02 BETA: Maybe. Maybe that's why. I guess, I guess that's why. But yeah, in... But for the most part, I think the translation was fine.

02:11 S1: Okay, what difficulties did you face during the translation task?

02:16 BETA: What difficulties I faced? It was not... It was not very difficult but it was sometimes... Whether I should change that or not, whether that the machine translation would be... Because I don't want to over-edit. I just... If it's... If it's legible, if people can understand this translation, I should not be making any... Like any modifications. So, yeah. But sometimes when I read it and I thought, "Okay, it's... Ignore the source text." And I would read it and I would think, "Okay, would I understand that if I read it for the first time?" And if I did, then I wouldn't change it. If I did not, then I had to make changes to make it clear.

02:56 S1: Okay. Have you changed your general assessments of the tools through this experiment after using that machine translation and the CAT tools at the same time?

03:06 BETA: I don't think I changed my perspective of the tools because I have been working with the tools and I know that they are beneficial but they... And they have their limitations, especially with language pairs that are like Arabic and English. This is a difficult language pair, and I understand that, if there are real limitations between, let's say in Spanish and English, there are much, much, much more limitations between Arabic and English because the nature of the language is the European... Like Arabic is a Semitic language. English is a Germanic language, like two different families.

03:37 S1: Okay. Which translation task was harder? And why?

03:41 BETA: Which one was harder? I think the one from... They were both not difficult... Not very difficult, but I would think that the more challenging one would be the one from Arabic into English.

03:52 S1: Why?

03:53 BETA: As I told you, like because maybe because there were some inconsistency, sometimes. Sometimes they were awkward structures. Like if... Like when I read the translation it didn't make sense, but then when I got back to the original, then I understood why it did that. Maybe because the machine is not fed with enough text translated from Arabic into English. I'm not sure.

04:18 S1: Do you think it's because of the segmentation, the long segments?

04:24 BETA: I think, yes because in Arabic you can have, for instance, you can have a comma and then start a new sentence, and it wouldn't be a fragment. But in English, you cannot just write a sentence and then... Like a complete sentence, and then like a full stop, and then another complete sentence. That'll be a fragment... A run-on sentence like, so... But in Arabic, comma splice is very common. It's... It's okay. That's why I had to make these... So many changes with the punctuation. So whenever a sentence ended I had to change the punctuation in English from a comma into a full stop and then capitalize the letter or...

05:00 S1: So, which segment in the each translation task was hard or difficult? And why? Both, if you want to do it...

05:07 BETA: Let me... Yeah, yeah because I will not...

05:10 S1: You can go back.

05:11 BETA: If I can just have a quick look at it. Yeah. I would say the third segment in the English to Arabic text was a bit difficult because there was... Because I had to make so many changes. There were... It was not fine because the translation... If you read the translation it wouldn't make sense because the names of diseases or the names of symptoms are not... Are not translated correctly. It was... It's literal. So I had to make some changes in order to make that legible translation.

05:55 S1: Do you think it's because of the segmentation?

06:00 BETA: I don't think so. I think it's a problem with terms.

06:02 S1: Okay.

06:02 BETA: Because here like... There's like, "Severe hydration." For example, "Lowered blood pressure, bone density loss." And, "Fainting." So all of these, or most of them at least, were translated in a... Or were not translated correctly.

06:17 S1: Okay.

06:18 BETA: So I had to change them. And they were even messed up. Like for example, the word, "Lowered." For some reason, I don't know why, what the reason is, it was describing... In the Arabic translation, it was describing dehydration instead of describing the blood pressure.

06:35 S1: Okay.

06:36 BETA: So...

06:36 S1: Okay, let's go now to the other translation...

06:41 BETA: Yeah, the other translation.

06:41 S1: Yes.

06:41 BETA: What was the most challenging one? I think... Yeah, number two was the most challenging one. The first reason, as you mentioned, is the punctuation: The commas, the run-on sentences. Also because the structure was very complex. It was... It was a very long sentence. Like we start with a subject, and then the verb is almost by the end of the sentence. So I think that caused the problem for the machine because it didn't recognize where is the verb, the main verb of the sentence. So it made like fragments... Like a lot of fragment sentences. Yeah. And I think that... It was the longest, also. It was the longest segment. So, I think that's why.

07:31 S1: Did you face any problem with the functionality? Other stuff? Bi-directionality, numbers, symbols?

07:44 BETA: Not really. Well, no. I don't think there was... No, there were not major issues. No.

07:54 S1: Okay. Which segment was easier in each?

07:55 BETA: The easiest? I think, for both, the title was easy. The title was the easiest. It was automatically translated and surprisingly, it was translated right. [chuckle]

08:04 S1: Okay. And now let's go... Did you have any difficulties with segmentation, punctuation or... In both Arabic and English translation?

08:17 BETA: So the punctuation, as I told you, like the commas and the run-on sentences in Arabic, which are fine, but they are not in English.

08:29 S1: Do you mean the produced commas?

08:32 BETA: I'm sorry?

08:32 S1: The output commas is in the English one?

08:35 BETA: Yes, yes because they are fine in Arabic, in the source text, but they are not... They are not in English. That's why if you... If you can... If you can have a look at what I did here. So whenever a sentence ended, like here it was, in Arabic... In the source text, the American...

[amerkia]

08:49 BETA: There was a comma after it. But in Arabic, like after the word "American," I had to put a full stop.

08:55 S1: Okay.

08:55 BETA: I could not start a new sentence with just a comma. And that's what I did with the rest of that specific segment.

09:04 S1: Do you mean the difference between... There are both English and Arabic punctuation rules?

09:07 BETA: Yes, yeah. This is like, I would say, this is like arbitrary changes. I cannot just not do them, I have to do them.

09:15 S1: Okay. What suggestions do you have to improve the tools use? Any suggestions?

09:21 BETA: I would say, if there was a way of like that the machine translation would recognize the terms that were already translated and keep those terms consistent, that would've... That would've been better. I mean, like in the case of "Oral cancer" and "Mouth cancer," for example. I know that there was a translation memory and a machine translation, but if there was a way the machine translation would have recognized, "Oh, this term was translated this way." So whenever it occurs again, it'll be trans... It'll be translated consistently. So, it would be... If it would... That would have been much better. It's... It's kind of interactive in a way, but I'm not sure if that would be a possible thing.

10:04 BETA: Also, again, punctuation. [chuckle] I don't know. There are not like so many issues that I could talk about, but I have suggestions for the things that I think were important, like an important changes. So whenever there is a comma, especially in languages like Arabic, I think the machine should... I don't know... Kind of figure out a way, if this is a new sentence or if this is related to the part of the first sentence or phrase.

10:34 S1: Yeah.

10:36 BETA: That would... That would be better. Yeah.

10:38 S1: Do you mean for the machine to recognize the parts of the punctuation?

10:42 BETA: Yes, yes, if... Well, the machine will not really recognize they are... I think if we input like some rules for the machine to recognize how... For example, if this sentence makes perfect sense and it's just, it's over, and then there is comma and then this is a new sentence. Then the machine automatically could place a comma instead or like a semicolon or something. I don't know what rules we can give to the machine to do that, but I think whoever created the machine translation tool, can figure out this way.

11:22 S1: Sure, sure. Is there any other concerns that you'd like to share about your experience with the tools?

11:34 BETA: I would say there are some minor mistakes that the machine translation has with translation into Arabic, from English into Arabic. For example, there are like some diacritical marks like Hamza. That's... There are rules for that and when it's translated... The machine is not even recognizing that there was a mistake with the word...

[إلى]

12:00 BETA: Because it did not have a Hamza under it. But this is something very important. It's like the accent in Spanish, for example, or in French. It changed the pronunciation of the word. The fact that the machine did not trans... Did not recognize that error... And if I take that for granted that the machine recognizes all the errors, like the spell checker is on, and I just like ignore... Like I don't... If I didn't pay enough attention, I would just make a mistake and I wouldn't know that there was a mistake. So, yeah. This is an example, the spell checker is not really that accurate. Also with some spaces, like for example, with the "Wow." Which is...

[foreign language]

12:43 BETA: In Arabic, it's usually attached to the word, okay? The machine did not recognize... When I tried... When I attached it to the word like, "And something." It would just underlined it as if it was an error. So again, the spell checker. I don't know why for some reason... But yeah, that would be one thing that... That this would make me worry because I will have to go through every little thing to make sure that the punctuation is right. But, for example, in English, that was not a problem. That was not a concern.

13:17 S1: So it's time consuming to look at...

13:19 BETA: Exactly, yeah. And it defeats the purpose of machine translation because

this should help me not... I should not be like worried that everything would be wrong.

13:29 S1: Okay. Any other concerns?

13:32 BETA: No, I think that's everything.

13:33 S1: Okay. Thank you so much!

Interview with Gamma

00:02 Speaker 1: Okay. Let's start the interview, How was your experience with the tools during the task, both tasks?

00:14 Gamma: Okay. So, using SDL Trados made it easier since there is a term memory, translation memory and it gives you some nice suggestions. However, I'm not quite sure but the Arabic, the source Arabic, the one that you translate from Arabic into English, the source text seems a bit... It doesn't seem like it's actually written in Arabic, it seems like it's machine translation. A lot of the sentences seemed kind of awkward, I didn't like it. It made it harder to translate into English. And as for the English, it was personally, it was much easier to translate. I'm guessing since it was actually written in English. I might be wrong for the Arabic, but that's what I felt, so it made it much easier. The only thing, the issue is just with whenever there's a listing, when they're counting several symptoms, it was just difficult with adding, sometimes you need to add the letter wa in some situations, that's all.

01:39 S1: Okay. How can you rate the output of the tools during the translation tasks? Out of 10?

01:45 Gamma: I think 6 in general. For the Arabic, it was okay. It would need a lot of revision, I guess. I mean the source text, the first one where I translate from Arabic into English, the English output was a bit off, I'm guessing because the source text was a bit... It wasn't perfect. But the second one where I was translating into Arabic the output was much nicer, in my opinion.

02:12 S1: Okay. Do you think the tools helped your productivity and how?

02:18 Gamma: Yes. Actually, it made it much easier. If you go back to the recording, you can see that I noticed the word oral was used once and then the other time it was used mouth, and those were both from the term memory. So, I guess that there was some kind of...

02:36 S1: One of them was MT, one of the...

02:39 Gamma: Yes. And so I went back to just to make sure, I Googled it really quick. I found that the first was... The first thing that showed up was oral so I changed them all to keep it consistent, as oral. I'm guessing the translation memory will save these kind of...

02:56 S1: Okay. What difficulties... Sorry. What limitation do you think the tools have?

03:03 Gamma: Well, for Arabic, my only issue is that the segmentation rules is a bit difficult because...

03:13 S1: The longer one?

03:13 Gamma: Yeah. So you can see they're very long because sometimes there's no full stops. So you can see that I added full stops in the English version where I felt the sentence was ending.

03:29 S1: Okay. What difficulties did you face during both translation tasks other than the segmentation? Did you have any problem with punctuation, spelling, grammar?

03:41 Gamma: Mainly, as I just said the consistency. I'm guessing if there was a... There's no term base, correct?

03:51 S1: No.

03:52 Gamma: Yeah. I'm guessing if there was a term base for this text, since it's medical, I think it would be much easier to translate when you have the terms already there, you don't necessarily need to go and look for them.

04:06 S1: Okay. Have you changed your general assessment of the tools after this experiment or it's the same?

04:13 Gamma: I'd say it's kind of the same because I've used them before so I already have a general idea.

04:19 S1: Have used before the both TM and MT, together?

04:25 Gamma: No. I've only used TM and a term base, but no MT.

04:31 S1: How was your experience with the use of MT? The interaction with TM and MT, do you think it's helpful?

04:38 Gamma: I think it is helpful in some situations, yes, it makes it much easier.

04:43 S1: Okay. Which translation task was harder and why?

04:50 Gamma: The first one where I translated from Arabic into English, and I've mentioned this at the beginning...

04:55 S1: You mentioned the… what about the segmentation, do you think it's...

04:58 Gamma: That's also an issue because of the segmentation where there's no full stops and no paragraphs, so there's no breakage of the segments.

05:07 S1: More complex sentences.

05:08 Gamma: Mm-hmm.

05:09 S1: Okay. Which segment in each translation was harder and why?

05:15 Gamma: I'd say the second one in the Arabic.

05:17 S1: And of the English one? And the Arabic to English... Sorry, the English to Arabic?

05:24 Gamma: It's probably the second one.

05:29 S1: The second one?

05:30 Gamma: Mm-hmm.

05:31 S1: Okay. Which segment was the easiest in both?

05:38 Gamma: The easiest for me? Well, the first one I didn't really, didn't do any changes to it, so I'd say that was the easiest for the first one, for the Arabic to English.

05:51 S1: The first one.

05:53 Gamma: And I'd say the same for the English as well.

05:57 S1: Okay. Did you have any difficulties with the segmentation, punctuation in both Arabic and English translations?

06:04 Gamma: Mainly from the Arabic to English, I had some issues.

06:09 S1: Okay. Did you have any problem with the punctuations, you said the segmentation. What about the punctuation?

06:17 Gamma: I'd say in Arabic, that's a bit difficult since we don't usually use this way of listing things, the symptoms. We usually just use wa, wa, wa and so it was kind of confusing. I didn't wanna change much since I'm not sure what the client would really want here.

06:39 S1: Okay. What suggestions do you have to improve the tools use here?

06:48 Gamma: I'd say they need to work on some kind of segmentation rule, like a preset segmentation rule for Arabic. I know that you can play around with it and change the segmentation rules by yourself, but I think if there was a preset segmentation rule that is set for Arabic, I don't know how they would do it but it would make it much easier, translating from Arabic into English.

07:12 S1: Okay. Is there any other concern that you'd like to share about your experience?

07:18 Gamma: No.

07:18 S1: Thank you so much,

Interview with Delta

00:02 Speaker 1: Okay, let's start the start the interview with Delta. How was your experience Ali, with using the tools during the translation task?

00:14 Delta: For Task 1, Arabic into English, using the automated translation was less helpful because it created multiple syntactic errors as well as vocab errors.

00:42 S1: Okay. How can you rate out both of the tools during your translation tasks?

00:50 Delta: On what scale?

00:51 S1: Out of 10, if you want.

00:53 Delta: Out of 10. I would say six out of 10.

00:58 S1: Do you feel these tools helped your productivity and why did it help?

01:05 Delta: The tools helped speed up the process, but it also introduced some problems with re-arranging things. It caused editing problems. It took me a while to re-arrange some syntactic structures. And also, the fact that the translation is from Arabic into English, which is my second language. I had to look up some vocabs and make sure that they are at the right terms.

01:42 S1: What limitation do you think the tools have?

01:45 Delta: Can I add just one thing for the other question?

01:47 S1: Yeah. Sure.

01:48 Delta: So, from English into Arabic, I thought that the tool, especially the automated translation, was more helpful to me. The result was more accurate.

01:58 S1: The English from... Into English?

02:01 Delta: English into Arabic.

02:02 S1: Yeah.

02:02 Delta: Yeah. And more accurate so I just did minor changes, some minor changes.

02:11 S1: What difficulties do you think... What limitations do you think the tools have? Yeah. What limitation do you think the tools have?

02:21 Delta: Well the limitation is, I would say, there need to be more context to be fed to these tools because, obviously, the structure has some issues, the syntactic structure into

English and into Arabic.

02:52 S1: What difficulties did you face in the translation tasks? Any problems [02:56] you faced.

03:00 Delta: During the translation task, I had to look up some terms. So, the major difficulty I faced was looking up some terms. And the other difficulty was re-arranging the second segment of the first paragraph from Arabic into English. Yeah.

03:29 S1: Have you changed your general assessment of the tools after this experiment?

03:36 Delta: I would say automated translation is a bonus, and the CAT tool, it saves the time. Sometimes I use Google Translate. I might sometimes just dump a whole sentence or paragraph and then paste it in the tool. And it's a time saver here.

03:57 S1: Okay, which translation task was harder?

04:01 Delta: From Arabic into English was harder.

04:04 S1: Why?

04:05 Delta: I think it involved a longer segment. And then, this is the first part. One long sentence, Arabic sentence, that needed to be translated into English to a multiple... To multiple sentences. And the other factor is that English is my second language.

04:31 S1: Which segment in each translation do you think was the hardest one?

04:39 Delta: The second segment in the first paragraph was the hardest in all the tasks. And the second paragraph, I didn't see any difficulty in any of the segments but I would say 2 too.

04:50 S1: Okay. Which segment in each translation task was easier for you?

04:58 Delta: I would say the segments in the second translation were very easy to me. They were of the same difficulty.

05:11 S1: And the first task?

05:14 Delta: The first task, the last segment was easier than the other.

05:17 S1: Did you have difficulties with segmentation, punctuations in both Arabic and English?

05:24 Delta: Yes, yes.

05:25 S1: In both or in just one side?

05:28 Delta: Both of them, there were some punctuation issues.

05:31 S1: What about segmentation?

05:35 Delta: The segmentation in the first task was problematic, I think, because the second segment was too long to follow.

05:48 S1: What suggestions do you have for the tool, to improve the tools, if you have any?

05:54 Delta: Well, I would suggest having more advanced segmentation rules. Yeah, so to make the sentence... The segments smaller for me.

06:09 S1: For Arabic ?

06:11 Delta: Arabic into English. The problem that I faced at least was Segment 2, was due to the fact that the segment was too long.

06:21 S1: Is there any other concern you would like to share about your experience today with the tools?

06:28 Delta: Well, I'm looking forward to see a CAT tool that is interactive enough to let me do the research about the vocabs or terms or any other aspect, cultural aspects, that needs to be researched. I would like to see a feature that allows me to do that within the tool itself. I don't have to exit the tool and come back.

06:57 S1: Okay. Any other concerns, suggestions, adjustments?

07:03 Delta: No.

07:03 S1: Thank you so much.

07:04 Delta: Thank you.

Interview with Epsilon

00:00 Speaker 1: Okay. Let's start the interview with Epsilon. How was your experience with the tools in both translation tasks?

00:12 EPSILON: The experience was good. Sometimes the translation was good, but sometimes it is just bad. You have to edit it... You have post-edit it. I think the punctuation is not placed in the correct place.

00:30 S1: Okay. How can you rate the output of the tools? Out of 10?

00:35 EPSILON: I would say it's good. I would give it like eight out of ten.

00:40 S1: Okay. Do you think the tools helped your productivity during the task, you have been more productive during the...

00:49 EPSILON: Yes, I think it helps a lot. It helps a lot in translation. A tool like SDL Trados restores translations that can help you in the future if you translate texts or translate similar texts in a particular field.

01:13 S1: What limitation do you think the tools have?

01:17 EPSILON: A lot. For example, the punctuation, grammatical mistakes, and that's it. Other than that, it is good.

01:28 S1: Did you face difficulties with the spelling?

01:32 EPSILON: Sometimes yes. Sometimes you have to post-edit the spelling.

01:36 S1: Any missing segments, translation?

01:40 EPSILON: Yes, there was a missing a part of a sentence.

01:45 S1: Okay. Any problems with the plural. The plural?

01:52 EPSILON: Yes. Sometimes yes, I had problems with plural.

01:57 S1: Okay. So these are the difficulties that you faced. Okay. Have you changed your general assessment of the tools after this experiment or it's the same?

02:06 EPSILON: It's probably the same because this year... This semester I'm also working on Trados. It is just the same. I mean, we cannot rely on Trados. We have to post-edit the texts.

02:22 S1: Yeah. Is it your first time to use MT with the translation only or?

02:27 EPSILON: No, it's not the first time. I've tried it before.

02:29 S1: Okay. Which translation task was harder and why? Arabic to English or English to Arabic?

02:37 EPSILON: I think... Which one was harder?

02:43 S1: Yeah.

02:44 EPSILON: The first one.

02:45 S1: The Arabic into English?

02:46 EPSILON: Yes, the first one was really difficult from... The Arabic to English. You have to post-edit a lot. But the other one was easier.

02:57 S1: Why?

03:00 EPSILON: Because the Arabic to English translation was... Had a lot of mistakes that you have to post-edit. But the English to Arabic had little mistakes.

03:19 S1: Okay. Did you face any segmentation issues?

03:25 EPSILON: No, not really. Not really.

03:27 S1: With the Arabic?

03:29 EPSILON: With the Arabic to English?

03:30 S1: Yeah.

03:32 EPSILON: Yes, there are a lot of... There are a lot of problems. For example:

[Surtanat alfm]

03:41 EPSILON: Sometimes they say, mouth cancer. The correct is oral and throat cancer, yes.

03:49 S1: Which segment in each translation was the hardest one?

03:54 EPSILON: Probably the second one of this one, from the Arabic to English. The second one.

04:02 S1: Okay, and the second task? The segment was...

04:12 EPSILON: The second one also.

04:14 S1: Okay. Okay. And which segment was the easiest one?

04:22 EPSILON: The first one. The first one of each. [chuckle]

04:25 S1: Okay. Did you have difficulties with the segmentation, punctuations in both Arabic and English translation tasks?

04:33 EPSILON: Yeah, punctuation.

04:34 S1: In both or in just one of them? Which one is...

04:38 EPSILON: I think punctuation is still an issue in both texts, in my opinion. You have to post-edit it. And this segment was not translated completely, so I translated it. But the punctuation, sometimes you have to post-edit it because it's not correct.

05:01 S1: Okay, what suggestions do you have to improve the tools use?

05:09 EPSILON: To work on grammatical mistakes, punctuation, spelling, and that's it.

05:23 S1: Okay. Is there any other concern that you would like to share about your experience?

05:29 EPSILON: The tool is good. Trados is good. But I think it can be better. I don't have any further suggestions.

05:37 S1: Okay, thank you so much.

05:38 EPSILON: Thank you.

Interview with Zeta

00:00 Speaker 1: Okay, let's start the interview with Zeta. First question, how was your experience with the tools during both translation tasks?

00:13 Zeta: It was efficient. It was super fast. The translation memory did pretty much everything, I just fixed some sentences.

00:24 S1: What about the machine translation? Is it the same?

00:26 Zeta: The machine translation was really helpful, especially with the medical terms because I'm not familiar with medical terms. There were maybe two mistakes. One of them with the cancers, they just translated it as [00:48] "Surtanat" in the Arabic segment. The other one, I actually can't remember the other one but... I don't call this a mistake, but there is inconsistency. For some segments, the translation of mouth cancer or the [01:11] "Surtan Alfm" and some segments, the translation is oral cancer. So, I just used the mouth cancer throughout the translation.

01:21 S1: Okay. How can you rate the output of the tools during the translation task? Output in general. Out of 10?

01:28 Zeta: On a scale out of 10, I would give it a nine.

01:31 S1: Okay. Why?

01:33 Zeta: Because it facilitated the translation process and also helped me when... As I said, the medical terms. And also the translation itself, I just post-edited some sentences.

01:48 S1: So do you think the tools helped your productivity during the task?

01:52 Zeta: Absolutely, yes.

01:54 S1: What limitations do you think the tools have?

02:00 Zeta: Limitations.

02:00 S1: Punctuations, spelling... Dictionary.

02:06 Zeta: Probably... Yeah, the spelling in the Arabic sentences, it has the red underline for some words that are correctly spelled. But I won't call this as a limitation, but everything is perfect.

02:27 S1: Okay, what difficulties did you face during the translation tasks? Terminology, punctuation, segmentations?

02:37 Amer: Yeah, as I said, the medical terminology because... Actually, I'm not familiar

with too many terms.

02:47 S1: Have you changed your general assessment about the MT or the CAT tools in general after this experiment?

02:55 Zeta: Well, I'm actually pro-machine translation... And also CAT tools in general. I think they did a pretty good job for us.

03:06 S1: So you haven't changed your... Okay. Now, which translation task was harder and why? Is it the English to Arabic or the Arabic to English?

03:17 Zeta: For me, the Arabic to English.

03:21 S1: Was harder.

03:22 Zeta: Yeah, It's always easier to translate to your native language.

03:26 S1: Which segment in each translation was difficult and the most difficult and why?

03:34 Zeta: In the other text...

03:36 S1: Let's start with the Arabic...

03:39 Zeta: Arabic-English. Yeah, I think there is a grammatical mistake with the Arabic one. [03:47] I think there is something wrong with this sentence.

03:50 S1: Okay.

03:51 Zeta: It has to be changed.

03:53 S1: So you have this... Number two as the difficult segment?

03:56 Zeta: Yeah.

03:57 S1: Okay.

03:58 Zeta: It's not as difficult as... It's not...

03:58 S1: I mean the most difficult one, the hardest one. So yeah.

04:01 Zeta: Yeah.

04:01 S1: And for the segment, the other task... Which segment do you think it's the

hardest one? There's...

04:10 Zeta: I would say the segment number three.

04:13 S1: Okay, number three.

04:14 Zeta: Yeah, because the English... As you can see, too many words divided with the comma but you can't follow that in Arabic. You have to re-arrange the sentence structure.

04:32 S1: Okay. Which segment was easier in each task? The easiest one.

04:41 Zeta: I would say... Yeah, this one... Segment number one in this task.

04:49 S1: And, in the other one, the same?

04:51 Zeta: No, because I'm not familiar with anorexia... Anorexia.

05:00 S1: Yeah, it's depression... So which one is the easiest?

05:02 Zeta: Last one, number four.

05:09 S1: Okay. Did you have any difficulties with punctuations or segmentation?

05:18 Zeta: In the Arabic one, yes. Because...

05:21 S1: This one?

05:23 Zeta: No, no, no. This one... Number three.

05:26 S1: Okay.

05:27 Khaed: Because as you can see, I have a comma after [05:31] and a comma after [05:33]. And then, I just left the other two without a comma.

05:38 S1: Okay. What about segmentation for both Arabic to English and English to Arabic?

05:42 Zeta: If it is up to me, I would divide the segment number two into two segments.

05:49 S1: Okay.

05:50 S1: Like after the [05:53] and then I will divide it to...

05:56 Zeta: Because of the Arabic length. It's one sentence in Arabic language...

05:58 Zeta: Yeah, it's too long. It's too long.

06:04 S1: Okay, what suggestions do you have... Do you have any suggestions to improve the tools use?

06:13 Zeta: I would say the terms should be consistent, like the mouth cancer or oral cancer. Other than that, I think everything is pretty much good.

06:31 S1: And is there any other concerns that you would like to share about your experience with the tools?

06:40 Zeta: With the tool?

06:40 S1: Yeah.

06:42 Zeta: I would really like if it has a Google Translate or some...

06:52 S1: It has a SDL Trados machine translation, by the way.

06:56 Zeta: No, no. What I'm saying is, for some terms, if you want to see what are other suggestions other than the machine translation that's already given to you. Sometimes you just want to see if this term is actually used in English, then you validate it... Just going through various texts, but you have to go to use Chrome or Firefox.

07:23 S1: So you are talking about the corpus, if there's a corpus or something?

07:27 Zeta: Probably, Yeah.

07:28 S1: Okay, thank you so much, Zeta.

07:31 Zeta: You're more than...

Interview with Theta

00:01 Speaker 1: Okay, let's just start the interview with Theta. How was your experience, Theta, with the translation tasks?

00:09 Theta: I guess the tool was very helpful because, especially with the English into Arabic, because the... The Arabic output was pretty good but the English... But the Arabic into English wasn't that good, because there are some missing segments and some word, only one word appeared in the English section as an Arabic word. So it wasn't translated.

00:43 S1: Okay, how can you rate the output of the tools? Out of 10?

00:48 Theta: If we say out of ten, I'll say, maybe seven...

00:55 S1: Do you think the tools help your productivity during the translation?

01:00 Theta: Yeah, because if I compare what I did to what I usually do before, I think it saves time and effort.

01:13 S1: What limitation do you think the tools have?

01:17 Theta: Maybe missing some segments. Maybe if I was provided with a missed-translation memory, maybe I won't come up with this translation. Also, the nature of the text, the text was well written in the source text, then of course I would have a good output.

01:40 S1: Okay what difficulties did you face during the translation tasks?

01:48 Theta: Maybe the structure of the Arabic sentence.

01:53 S1: The long sentence, units.

01:54 Theta: Yeah, the long, the second one was, yeah, was long. So I need to break it into two sentences in the English segment. Also the third. The third one, the last sentence wasn't really clear in Arabic. So, I need to render it into English in a different way. Yeah, that's it.

02:24 S1: Have you changed your general assessment of the tools after this experiment or is it the same?

02:30 Theta: Sorry, can you repeat that?

02:31 S1: Have you changed your general assessment, your evaluation of the tools after this experiment or it would be the same evaluation or same assessment?

02:40 Theta: I've tried this tool with other texts, they were technical texts, but there are

problem with consistency, but I guess here we have only three segments. So it was really good, the experience was good.

02:53 S1: Which translation task was harder and why? I think you said the English to... Arabic to English.

03:02 Theta: The Arabic into English was harder because there are many... I told you that I need to break the sentences. I need to change, the word "Surtan" wasn't translated. Also, the output I get from the tool wasn't exactly the same of the source, so I need to change it.

03:26 S1: Which segment in each translation task was harder and why?

03:32 Theta: I guess two, yeah.

03:35 S1: That second one which...

03:37 Theta: So from Arabic into English was the second segment and then from English into Arabic I'd say, maybe the second one 'cause I changed it.

03:52 S1: Okay, did you have difficulties with segmentation and punctuations, spelling?

04:01 Theta: I guess the tool following the English rule of punctuation, so I need to change and of course because I break sentences, so I change the punctuations. Other difficulties, I guess that's all.

04:21 S1: Okay, what suggestions do you have for the tools, to improve the tools you use.

04:32 Theta: Maybe we need to work on the... I don't know, do we, are we using a machine translation or translation memory?

04:41 S1: Both.

04:43 Theta: So we need to choose, or to work with using a good translation memory and a good machine translation. Because the tool itself is very helpful, but when you use the missed out translation memory or machine translation...

05:02 S1: These two are where the machine translation and these two were translation memory. So when CM which means complete match from translation memory, this is one from machine translation. So let's still try this and the same for the... Where is it? Here, this is following the translation memory, but this is from machine translation. That's a problem, that is... This is a segment, [05:35].

05:42 Theta: I have a comment on the segmentation rule, I guess the one who used this tool, he has to put the segmentation rule built on smaller segments. Because this is the

major problem. It took much of my time to just break sentences. So if it were just a small segment or chunks it would be better.

06:06 S1: Are there any other concerns that you'd like to share about your experience?

06:16 Theta: No that's all.

06:17 S1: Okay, thank you so much Theta. Thank you.

06:20 Theta: You're welcome.

Interview with Eta

00:00 Speaker 1: Okay, start interview with Eta. You can go back to your... Work, So you can review it if you need it. How was your experience with the tools during work tasks?

00:18 Eta: I found it really helpful, especially with the English. But I need it to look for [00:25] para texts.

00:28 S1: Okay, How can you read the outputs during you translation tasks? Out of 10?

00:36 Eta: The outputs?

00:36 S1: Outputs of the tools yeah.

00:40 Eta: The outputs, was...I would say 6

00:41 S1: English to Arabic and English to Arabic.

00:42 Eta: Yes the Arabic to English? Yes, I think it was better compared to the one from an English to Arabic.

00:51 S1: Why do you think so?

00:54 Eta: Because of the terminology, the medical terminology is the big issue. Especially I don't have a background about the medical terminology.

01:03 S1: Okay.

01:05 Eta: So I was trying to make it sound natural 'cause Arabic is my first language.

01:11 S1: Do you think that will help your productivity during translation tasks and how?

01:19 Eta: Yes, of course.

01:20 S1: How?

01:20 Eta: It makes it easier, faster. But I guess I was lacking experience because I haven't translated for a long time.

01:28 S1: Okay.

01:29 Eta: Maybe I could be faster if I was translating...

01:31 S1: Yeah but for the MT suggestion, was it okay to?

01:34 Eta: Yes, yeah, it was so good.

01:36 S1: What limitations do you think the tools have? Any issues, you encounter while you are translating?

01:44 Eta: Limitation? In this case, I couldn't find any limitations because I have access for the whole text, the context, I can read it.

01:55 S1: Did you search online for the...

01:57 Eta: About the texts? Yes. This is the only thing and the machine translation is not connected to the...

02:03 S1: To the context you mean?

02:08 Eta: Yes.

02:09 S1: What difficulties did you face during the translation task? Any difficulties, terminology, punctuation...

02:15 Eta: Punctuation and terminology is the big issue, with Arabic.

02:21 S1: And the plural in Arabic, do you have a problem with the plural?

02:25 Eta: Yes. My gosh I think even in the English the text from Arabic into English, there was [02:34] "Surtanat".

02:36 S1: Okay. It wasn't translated?

02:38 Eta: No It wasn't.

02:40 S1: Okay, so it doesn't apply to that. Have you changed your general work system to the tools after of this experiment? I mean, have you changed your assessment, general assessment after experiencing this?

02:55 Eta: No because I know that CAT tools are very helpful. Especially with these kind of text, technical, specialized text.

03:07 S1: Okay, which translation task was harder? And why?

03:10 Eta: English.

03:11 S1: I think you said Arabic to English.

03:13 Eta: Yes, Arabic to English.

03:15 S1: And you say why? Because of the terminology.

03:18 Eta: Yes.

03:18 S1: Which segment of the translation task was easier? Harder sorry.

03:26 Eta: Harder?

03:27 S1: Difficult.

03:28 Eta: Arabic to English.

03:30 S1: Yeah.

03:31 Eta: The second one.

03:31 S1: Okay.

03:33 Eta: English to Arabic... I think the third one.

03:48 S1: Okay... Which segment was easier?

03:53 Eta: The first one, the title.

03:56 S1: Okay, do you have difficulties with the segmentation or punctuations?

03:58 Eta: Punctuations yes. But the...

03:58 S1: You said, segmentation or the Arabic into English.

04:05 Eta: No. It was good, the segmentation was good. It was clear like the chunks are complete.

04:11 S1: What suggestion do you have to work the tools? Do you have any suggestions? That we'd like to see while working on translation?

04:23 Eta: Especially for this tool for [04:24] For SDL Trados you mean? Or generally?

04:26 S1: In general.

04:28 Eta: They need to work more for Arabic. They ignore Arabic, we have a lot of issue, Arabic, Japanese, Chinese.

04:34 S1: What do you think they should [04:35] do about Arabic is it just punctuations and spelling?

04:39 Eta: In this case yes but generally sometimes you cannot convert the text. Or insert the text in the tool.

04:47 S1: Is there any other concern you would like to share?

04:50 Eta: Any other concern that you would like to share?

04:52 S1: Have you had any problem with the Hamza?

04:55 Eta: Yes. Yes I did. I didn't mention that.

05:00 S1: Do you think any of it was time consuming for editing, compared to translation from scratch?

05:08 Eta: Because I'm not experienced, I don't have experience in editing. So it's time consuming For me, it takes a while, and as I told you I didn't translate for a while.

05:18 S1: Okay thank you so much Eta.

05:20 Eta: You're welcome.

Interview with Iota

00:01 Speaker 1: Okay, now let's start the interview with Iota. Okay, Iota, so how was your experience with the tools during both tasks?

00:12 Iota: It was helpful for some segments. Other segments, no. I needed to delete the suggestion and re-translate the whole thing.

00:25 S1: How can you rate the output of the tools during both tasks?

00:32 Iota: The output? I can't complain about it. I would give it 6 out of 10. In general, it was good. And it seems that tools, the CAT tools, are always helpful when you're dealing with like scientific, legal, economic, you know, texts with controlled language. Yeah, they can be very helpful.

00:57 S1: And you think the tools helped your productivity here in both tasks?

01:02 Iota: It did, yeah. It did.

01:03 S1: And how?

01:04 Iota: For some segments, I didn't need to... Editing was very... Post-editing was very minimal. I didn't do much. Yeah, so, it saves time.

01:19 S1: Okay, what limitations do you think the tools have in general?

01:22 Iota: In general?

01:23 S1: Yeah.

01:23 Iota: Well, the first thing... So when you say the tools, do you include the machine translation or without...

01:34 S1: Both, both.

01:35 Iota: Okay. Well, first of all, if you're depending on the tool itself without machine translation, it means that if your input is good, then the output will be good. If it's bad, then the output will be bad. Well, you know, they say "Garbage in, garbage out." So that's one of the limitations. The other limitation, I think that the CAT tools cannot really overcome the type of language they deal with. So, for example, as I said, if you're trying to translate a technical text, a medical text, legal, anything like that, yeah, the tool would be of great help. But if you're translating, if I can say like a loose-language text or something like analysis or a political article or news or something like that... Yeah, the tool wouldn't help that much. Even the machine translation. And Arabic is also a special case with the translation tools. Because of the diacritic, we call it... The diacritic system in Arabic, the tool cannot sometimes differentiate between words and cannot read the

texts very well because of that.

03:13 S1: The CAT tools or the MT?

03:14 Iota: The MT. Here, I'm talking about the MT. Yeah, the MT cannot really differentiate between, cannot distinguish Arabic words...

03:22 S1: Or recognize...

03:23 Iota: Yeah, recognize Arabic words because of the diacritic system in Arabic.

03:26 S1: Okay. What difficulties did you face during your translation tasks?

03:32 Iota: Now?

03:33 S1: You know, like segmentation or punctuations or spelling.

03:46 Iota: I don't know. Maybe the only problem is that it will underline some correct words, it would underline it as...

03:58 S1: What about Hamza?

04:00 Iota: The Hamza, yeah...

[foreign language]

04:06 Iota: Yeah, here. Okay, yeah, Hamza. There is no Hamza, so it doesn't suggest that.

[foreign language]

04:15 Iota: Yeah, so the underlining, underlined words. For example, like this one is underlined, but it's correct. This one is underlined, it's correct. Correct, correct. This one is incorrect because there must be a Hamza that's underlined, so yeah.

04:30 S1: What about the plural in Arabic? Did you face... Recognize the plural system in Arabic?

04:42 Iota: Like for me as a translator or for the machine?

04:45 S1: No, for the machine.

04:47 Iota: For the machine... Recognizing the plural?

04:49 S1: Yeah, did you have a problem with the Arabic terms?

[foreign language]

05:00 Iota: Honestly, I'm not sure.

05:04 Iota: I'm not sure to be honest with you, because if there was a problem here, I deleted the segment and re-translated, so I didn't pay attention.

05:11 S1: Have you changed your general assessment of the tools after this experiment or it's the same?

05:16 Iota: Actually this exercise, I think the tool did a good job for this task. I can say that this experiment made my experience with the tool now is better than previous experiences. So yeah.

05:36 S1: Is it because of the MT interaction?

05:40 Iota: MT here, for sure is...

05:42 S1: This is SDL cloud machine translation.

05:44 Iota: Okay. Yeah, for sure, the MT here did... Is this the google MT or...

05:49 S1: No. This is SDL cloud...

05:51 Iota: Oh, SDL Cloud?

05:52 S1: Yeah.

05:52 Iota: Okay. So, okay. One time I worked with the SDL cloud machine translation and it was so bad, it was so bad. This time I don't know if they fixed something, but this time it seems very good.

06:05 S1: This is not one [06:08] Did itself, so yeah.

06:11 Iota: Adaptive-like?

06:12 S1: Yeah.

06:12 Iota: Yeah, it's much better than the one I used in the past, like one year ago. Yeah, It's much better.

06:17 S1: So we have seen some improvements in the MT?

06:19 Iota: Yeah, yeah, if this... Yeah. Because I'm sure the one I used one year ago was SDL cloud or something.

06:24 S1: Is it the same difficulty of the text or plus? Do you know... Was it the same... That one you used? Do you know?

06:34 Iota: When I used it last time the major problem was the word order. It was doing a literal translation. So, for example, when an individual refuses...

[foreign language]

06:50 Iota: So it was like a literal translation. I didn't account for the different syntactic structures between Arabic and English, but now it looks more natural.

07:00 S1: Okay. Which translation task was harder for you and why? The Arabic to English or the English to Arabic?

07:09 Iota: I would always say Arabic to English is harder. A little bit harder because Arabic is my native tongue so it would be easier to deal with Arabic.

07:20 S1: Which segmented in each translation task was the hardest one?

07:26 Iota: In the first task... It was the second one. In the first task, Arabic to English, it was segment two. In the English to Arabic, it was segment three.

07:57 S1: Which segment was the easiest in both tasks?

[background conversation]

08:22 Iota: I'm so sorry.

08:23 S1: That's okay.

08:26 Iota: So the easiest?

08:30 S1: Yeah, in both texts.

08:33 Iota: So in task number one, segment number one, it was already translated. It was easy.

08:40 S1: The other task?

08:44 Iota: Oh, and now I see that task three was easy, but the tool here... There is a small

missing part that the tool didn't catch.

[foreign language]

09:08 S1: Okay, the second one?

09:09 Iota: The second one... The first segment was already translated. The last segment, number four, it was also easy.

09:21 S1: Okay. Did you have difficulties with segmentation or punctuations in both tasks?

09:29 Iota: Segmentation, no.

09:32 S1: With your second segment, did you have any in the first task?

09:36 Iota: The first, the second segment? First task...

09:40 S1: Did you have segmentation issues?

09:43 Iota: Oh, definitely, yeah. Definitely. It was the segmentation here. It was a prompt for me, it was where to put the 83%. How to restructure the whole sentence because it doesn't make sense if I translated the sentence as is.

[foreign language]

10:10 Iota: That was also... It took me a few seconds, some time I mean, to recognize what they mean.

10:21 S1: Did you face any punctuation issue in both tasks?

10:29 Iota: Punctuation, no. I wouldn't say...

10:29 S1: You have any suggestions for both the tools' use?

10:35 Iota: The tools' use in general?

10:37 S1: Yeah, any suggestions?

10:47 Iota: Maybe this is not something that can be done by translators, but something that can be done by authors. It's just to write for machine translation. If you write for machine translation then you would make the tools life much easier, if you know what I mean. Like if I want to write a text to be translated, if I'm taking into account that it will be machine translated, I can change the way I write. Like I use shorter sentences, be more

consistent with using verbs, starting sentences with verbs.

11:45 Iota: Yeah, before you put it in the machine translator. So for the translator, here's the thing. For a translator, if you have a text and you can change it in a way that wouldn't change the meaning, but would make it easier for the the machine translation to translate it... Yeah, definitely.

12:01 S1: Is there any concerns that you would like to share about your experience?

12:09 Iota: It was a wonderful experience.

12:13 S1: Thank you so much.

Interview with Lambda

00:01 Speaker 1: Okay, let's just start the interview with Lambda. Okay, your first question. How was your experience with the tools during the task, both tasks?

00:12 Lambda: Okay, with the tools...

00:14 Lambda: Okay, so I think it was a bit challenging in the beginning because I am not used to the keyboard. [chuckle] But other things about the software itself, it's very useful and especially with the dynamics that are used in the program, that it's right to left, so you can really focus on one segment rather than just looking at the whole text.

00:42 S1: Okay. How can you rate the output of the tools during your translation task?

00:47 Lambda: Out of 10? And...

00:49 S1: Yeah.

00:49 Lambda: The reason? Okay, so the output... Okay, you mean this output, right?

00:55 S1: In both directions.

00:56 Lambda: In both directions, okay. I think I would give it more rate if I see it exported in a Word document, that way I can really give a specific rate. But for what I see now, I think I would give it like a seven because there are some issues, if you see here, it says there are some spelling mistakes even though there aren't, so these words that are underlined in red indicating that the spelling is wrong, but it's correct.

01:27 S1: Okay, do you think the tools help your productivity during and how?

01:35 Lambda: Yes, it does. It saved me a lot of time. That way, I can see the next sentences, if anything is repeated, so I don't have to re-translate it, it will be saved into my TM.

01:50 S1: Okay. What limitation do you think the tools have?

01:56 Lambda: Okay, the tools, they're very useful, but I would say not... It doesn't function as a human brain. Sometimes it just gives you a bunch of options and neither of them is correct. So you need to read the context again, and make sure this is the right term to use. But I think it’s part of my post editing job to fix these minor issues, nothing more.

02:16 S1: You're referring to the MT?

02:18 Lambda: MT, yeah.

02:19 S1: Yeah. What difficulties did you face during your translation task?

02:24 Lambda: Difficulties?

02:25 S1: Yeah.

02:26 Lambda: I would say difficulties at the terminology level. So there are some scientific terms that you need to know your locale first before you translate them, because a target audience use different terminology.

02:45 S1: Have you faced issues with Hamza in Arabic?

02:49 Lambda: Oh, correct, yeah, I faced some issues with the diacritical marks with the Hamza yeah.

02:55 S1: And the plural?

02:56 Lambda: And plural as well. You're right, yeah.

03:00 S1: Punctuations?

03:01 Lambda: Punctuations. I see in the second one, it's more punctuated. There are so many commas.

03:09 S1: Have you ever changed your general assessments of the tools after this experiment? Have you thought of changing your assessments?

03:18 Lambda: No, it's the same as I thought, yeah.

03:20 S1: Okay. Which translation task was harder? The Arabic to English, or English to Arabic do you think?

03:29 Lambda: Okay. For me, of course, the most challenging is, more challenging is from English to Arabic, yeah.

03:38 S1: Why?

03:40 Lambda: Maybe because most of my duties, like my job, I have been doing this for four years now, I've been translating from Arabic to English, so I find it easier to translate rather than English to Arabic, which is different from what others said maybe. [chuckle]

03:57 S1: Yeah.

03:57 Lambda: Yeah.

03:58 S1: Which segment in each translation task was difficult for you? Which segment that was the hardest one in each task and why?

04:07 Lambda: Okay, let's see. So in the first one, task one...

04:12 S1: Yeah, Arabic to English which...

04:15 Lambda: I would say number two, from Arabic to English, yeah, not because of the length, but I had... It's a long sentence, that first of all, a number thing is, the second thing, the percentage that you have here, with the...

04:32 S1: Did you have a problem with the percentage?

04:35 Lambda: No, these numbers. So because in Libya we don't use these numbers.

04:39 S1: Okay.

04:40 Lambda: So, [chuckle] it sometimes takes me a while to differentiate between the two and six.

04:45 S1: Okay.

04:46 Lambda: But I know the two is the same direction of three and six is the other direction. I didn't get used to writing these numbers.

04:57 S1: Okay. And task number two?

05:00 Lambda: The second one. Yeah, I think number four here, number four, took me a while to do that.

05:09 S1: Number four?

05:09 Lambda: Yes.

05:10 S1: Okay.

05:11 Lambda: Can you even do this one.

05:15 S1: Which segment was the easiest one?

05:17 Lambda: The easiest would be number one. [chuckle]

05:19 S1: Yeah. Titles.

05:20 Lambda: Yeah, the titles.

05:22 S1: Okay. Okay. Did you have difficulties with segmentation or punctuation in both Arabic and English translation?

05:33 Lambda: I think they are very... They are properly segmented, you see. They end in the same sentence. Sometimes you have the rest of the sentence here, when you segment something, but here they are properly segmented. Did you segment them with a period?

05:49 S1: Yeah.

05:50 Lambda: Okay. I can see. All this is very good.

05:53 S1: Did you have any problem with segmentations?

05:55 Lambda: No.

05:55 S1: You know longer, because this one is a very long one?

06:00 Lambda: Oh, this one, number two?

06:01 S1: Yeah. Because also it is, it's more than one sentence in English.

06:06 Lambda: Yeah, I think... Yeah, but segmentation if you do it like from Arabic to English, the period segmentation would work better. Yeah, I have experienced this before in my Masters. But when you do English to Arabic, you can do with the comma, but again, period is always the better. And it depends on the text, really.

06:27 S1: And punctuations?

06:31 Lambda: Punctuations. On the first task, yeah. There isn't much punctuations as the second task you see?

06:43 S1: Yeah.

06:43 Lambda: English uses a lot of punctuations.

06:46 S1: Okay. What suggestions do you have to improve the tools use?

06:52 Lambda: Okay, to improve the tools? Maybe this auto-correction, I think this has... Is this connected to Microsoft or I don't know?

07:02 S1: Yeah, you can.

07:08 Lambda: The Arabic auto correction.

07:09 S1: Spelling checker?

07:11 Lambda: Yeah, it's spelling checker.

07:13 S1: What else? Anything, edit environment, dictionaries, punctuation, segmentation?

07:26 Lambda: No.

07:28 S1: Okay.

07:28 Lambda: Everything looks good.

07:30 S1: Did you have any time consuming issues with editing?

07:35 Lambda: No.

07:35 S1: [07:35] changes?

07:36 Lambda: When I made the changes? No, it was easy.

07:39 S1: Okay. Is there any other concern that you would like to share with your... About your experience?

07:47 Lambda: Okay. Maybe about when the source text has bold, how can you bold the target? It automatically, it does that automatically, but what if you want to do something else in the target because sometimes you have underlined English sentence and then you wanna bold it in Arabic. It's like more...

08:14 S1: Editing, yeah. Editing differences between them.

08:18 Lambda: Yeah, yes.

08:19 S1: Okay.

08:20 Lambda: Or parenthesis and square brackets, curly brackets.

08:23 S1: Commas and...

08:25 Lambda: Yeah.

08:26 S1: Okay.

08:28 Lambda: Here, I don't see it is...

08:31 Lambda: It's not bold, yeah.

08:31 Lambda: Yeah, it's not even flexible to do. But I know you can just control, hit control and then B, yeah. And then control is to underline it, this is underline, control I.

08:53 S1: Any other concerns?

08:55 Lambda: No.

08:56 S1: Okay. Thank you so much.

08:58 Lambda: Thank you very much.

Interview with Kappa

00:00 Speaker 1: Okay, let's start the interview with Kappa. Okay Kappa, how was your experience with the tools during both translation tasks?

00:11 Kappa: I think that the experience was very good.

00:15 S1: How can you rate the output of the tools during your translation tasks?

00:21 Kappa: You mean the machine translation output?

00:23 S1: The output in general. You can give a rate for the machine translation and rate for the tools in general, as you like.

00:31 Kappa: I think the machine translation was not really accurate, so I had to make a lot of amendments, changes to the translation. But at the same time, it was helpful with the terminology. So I didn't need to look up for terminology online. I would give it 6 out of 10.

00:47 S1: Do you think that the tools helped your productivity during the translation tasks?

00:51 Kappa: I think yes.

00:53 S1: How?

00:54 Kappa: Again, there were several terms that I didn't know, but I didn't need to look up online, but the MT provided me with the Arabic translation. So I was focusing on the... Just the translation, making it better. So having the terminology ready, I think made it easier.

01:19 S1: What limitations do you think the tools have?

01:24 Kappa: I'm not sure about the limitation in this particular task, but in general, there are a lot of problems that happen when you have English texts, and you need to use it in the translation with the Arabic text. Usually, there are problems that happen, and usually when you clean up the files, you would need to do a lot of formatting. So it's not necessarily that, a lot of people would say Trados provides...

01:49 S1: Are you referring to bi-directionality?

01:53 Kappa: Yeah, bi-directionality problem, yeah. If you have to use like let's say a company's name in English, you wanna keep it in English. Sometimes there is problems with the directionality and then it appears in a formatting. If it's a Word document, you would see problems that you need to fix. And if let's say in a company, a project manager doesn't know Arabic, he wouldn't notice a problem. So needs to go back to the Arabic

translator to fix it.

02:23 S1: Okay. What difficulties did you face during the translation tasks? Have you had any difficulties?

02:32 Kappa: Are we talking about the text or the tool?

02:36 S1: The tools.

02:37 Kappa: No, I didn't find any problems with the tool because I've been using it for for six, seven years.

02:42 S1: Okay. Have you changed your general assessment after this experiment or it's the same?

02:47 Kappa: It's the same.

02:48 S1: Which translation task was harder and why?

02:52 Kappa: Well, the Arabic into English was harder of course.

02:54 S1: Why?

02:55 Kappa: Because it's easier to translate into your native language.

02:58 S1: Okay. Which segment in each translation task was harder for you?

[pause]

03:14 Kappa: Say number two in the Arabic into English task.

03:17 S1: Okay, for the other task, which segment do you think...

03:25 Kappa: I didn't find any. It was easy.

03:27 S1: Okay, good. Do you know why it's harder, the segment number two in the Arabic...

03:34 Kappa: It was long.

03:35 S1: Long one. Okay. Which segment in each translation was easier? The easiest one?

03:50 Kappa: The last one.

03:51 S1: The last one.

03:51 Kappa: Well, the first one or the last one, but not the one in the middle.

03:56 S1: Okay. For the English into Arabic?

04:01 Kappa: English into Arabic... The easiest was probably the last one too.

04:07 S1: Okay. Did you have difficulties with the segmentations or punctuations in both Arabic and English translations?

04:19 Kappa: with segmentation yes.

04:26 S1: No punctuation problem with number two in say, Arabic into English translation? Do you think it was too long for a segment?

04:33 Kappa: Yeah, I had to... No, when I translated it, I had to use more segments than the original one. Yeah, so I divided into more sentences in English 'cause it's very long in Arabic.

04:49 S1: That was not a problem for you?

04:51 Kappa: Didn't cause a problem. This is a translation problem. I don't think it has anything to do with the tool.

04:56 S1: Okay. What suggestions do you have to improve the tools use?

05:03 Kappa: Again, this idea of bi-directionality, it happens a lot with translation... Translating marketing texts usually. A lot of style guides in companies would want a lot of names to remain in English. So I think the developer needs to look at this problem and solve it because a lot of people complain about it even on online forums.

05:25 S1: Okay, is there any other concern you will like to share about your experience with the tools?

05:30 Kappa: No. I think the tools are very helpful, but some issues that need to be fixed with regards to... Regarding directionality. That's all.

05:40 S1: Did you have any problems with the... For the spelling, the accents, the diacritics... Hamza...

05:51 Kappa: Let me look at the Arabic...

05:53 S1: Punctuations?

05:57 Kappa: I don't think I had problems with punctuation, but the problem sometimes is the output coming from the MT doesn't have the Hamza trait, so you need to fix it. Like hi, for example.

06:09 S1: Okay. Thank you so much, Kappa.

06:12 Kappa: Thank you.

Interview with Mu

00:02 Speaker 1: Okay, let's just start the interview with Mu. How was your experience with the tools during the translation tasks?

00:12 Mu: I would say the tools saved me some time, but the translation needed to be looked at thoroughly to catch all the mistakes and make all the necessary corrections.

00:29 S1: Okay, how can you rate the output of the tools during both translation tasks? Out of 10?

00:35 Mu: I would say, I would rate, like if I'm to make an evaluation out of 100%, I would say it was 50% for the first text, the English one.

00:46 S1: Arabic to English?

00:48 Mu: Than the English to... Yeah, the Arabic to English was like 50, and the English to Arabic was a little better, I would say, like 60. In general I would say 50%

00:58 S1: Okay, do you think the tools helped your productivity and how?

01:04 Mu: Yes, I think the tool helped me with the translation.

01:08 S1: Okay. How it helped? With the translation you mean the suggestions? The suggested translation in MT or the translation memory, which one?

01:21 Mu: Yes, the suggested translation helped me because the translation was there, some of the segments were either correct or needed few changes or modifications. However, some of the segments were not accurate and needed to be re-translated.

01:42 S1: Okay, what limitation do you think the tools have?

01:46 Mu: Like limitation on my translation?

01:50 S1: No, in general, with the tools, for the tools used, what limitations do you think in general that they have? The tools have?

02:00 Mu: I think the limitation that I could think about here is that I think the tool would suggest a translation based on the input, and sometimes it's not like the equivalence of the translation that we need for this text. In this case we will have to re-translate because it just doesn't convey the meaning.

02:32 S1: What difficulties did you face during your translation tasks? Did you face any difficulties?

02:40 Mu: I don't think I faced any difficulties, I just had to make the necessary changes,

sometimes look up a term to make sure that I'm choosing the right term or like the most common term, for example, when the tool translated oral cancer in two different ways, it was inconsistent, oral cancer and then mouth cancer. I had to look that up and oral cancer is the most common one. So I had to change all the other translation to match.

03:22 S1: Have you changed your general assessment of the tools after this experiment, or is it the same?

03:29 Mu: My general assessment like before and after?

03:31 S1: Yeah, after this experiment. Before and after this experiment. Have you changed any evaluation in your mind, assessment of the tools?

03:42 Mu: No, I mean, I used this tool before, so I kinda like knew what to expect.

03:46 S1: Okay. Which translation task was harder and why?

03:51 Mu: The Arabic to English was harder.

03:54 S1: And why?

03:55 Mu: Than the English to Arabic. I think the...

04:05 Mu: Yeah, I think the Arabic to English was not as accurate as the English to Arabic and needed more modifications. I think that might be because of the sentences structure, the Arabic structure is kind of like way different from the English, where if you translate English to Arabic, it's easier to like get the meaning.

04:31 S1: Okay. Which translation task was the... Which segment in each translation task was harder and why?

04:42 Mu: In the Arabic to English, the second segment was the harder.

04:47 S1: Okay. Why?

04:49 Mu: I believe because of the sentence structure, and specifically the part where it says 83% of those who had oral cancer, but they didn't experience metastases, had a life expectancy of five years. And then it talks about that 32% of the patients had the same life expectancy when they didn't experience metastases. I think that segment was the harder because the suggested translation was inaccurate and kinda needed to be re-written or re-translated.

05:34 S1: Okay. And for the second task, do you have any segment?

05:38 Mu: The second task was easier than the first one. The segments were not hard, I would say they were equally, like the level of difficulty was equal between the three segments or the four segments.

06:00 S1: Okay, which segment in each task was the easiest one?

06:03 Mu: The first one was, I believe, was accurate and correct.

06:07 S1: Okay. Did you have difficulties with segmentation, punctuations, in both tasks or just one or in neither?

06:20 Mu: Yes, I had to change the punctuation because when translated from Arabic to English, the punctuation was copied. And in Arabic, sentences tend to be longer and segments are separated by comma, where in English sentences are shorter and there is a period at the end of the sentence, and then you start a new sentence. So I had to change the punctuation and make sure to change the letters or to capitalize the first letter of each sentence to an upper case, but...

07:02 S1: Okay. What suggestions do you have to improve the tools used? Do you have any suggestions?

07:13 Mu: One of the suggestions is consistency when entering the terms or the segments, because as I could see, Surtan Alfm had two different translation.

07:26 S1: Yeah, like one from TM, one from MT, so the Arabic translation has two sources.

07:30 Mu: Yeah, there is oral...

07:31 S1: And mouth.

07:32 Mu: Cancer and mouth cancer, so...

07:34 S1: The mouth is from their machine translation.

07:37 Mu: Yeah.

07:37 S1: So it was a literal source translation.

07:40 Mu: Yeah.

07:41 S1: Okay. So you are suggesting that to be a very, some interacting between the MT and TM more interacting.

07:50 Mu: Yeah, test your consistency, otherwise you will have more than one term referring to the same thing.

07:57 S1: Okay. Is there any other concern you would like to share about your experience today about the tools?

08:05 Mu: I think the tool was very helpful with this type of text. I am not sure about whether it will be at the same level of effectiveness with other type of texts.

08:21 S1: What do you mean other types?

08:23 Mu: Like this is a medical, [08:25]

08:26 S1: This is medical, but it's similar, it's general, it's not...

08:28 Mu: Yeah, general.

08:29 S1: Specific scientific...

08:30 Mu: Yeah, I don't think it would be like at the same, the translation would be at the same level if the text was literally.

08:45 S1: Okay. Okay, thank you so much.

08:47 Mu: You're welcome.

08:48 S1: Thank you.

08:48 Mu: Thank you.

Interview with Omicron

00:01 Speaker 1: Okay, let's start with the interview with Omicron, okay how was your experience with the tools during both tasks?

00:11 Omicron: They were kind of efficient to find the translation in there, but it required a lot of post-editing, especially the text where we really kind of consistent sometimes in terms of translation, especially I mean regarding the structure, the way English is structured is definitely different from Arabic, and that's why I had to do some sort of fore grounding and back grounding for certain information. This is... And I had to change different things like this is totally perfect in Arabic to have a kind of complete sentence of three lines, it's quite long one but it's totally perfect in Arabic. Which wouldn't be the case in English. So I had to really decide where each sentence can stop, where I can start another new sentence. And the text in Arabic seems to be kind of like back-translated from English, and that's why it was kind of really difficult to understand certain things. Like...

[foreign language]

01:32 Omicron: There's definitely a better way to structure this in Arabic, and that's why it caused kind of difficulty translating them, translating this phrase into English. And the one that I had to post edit required me kind of long time to really first get the meaning from the source text and translated into the target text.

[foreign language]

01:58 Omicron: I had to resort to deletion for something that would be redundant. Like.

[foreign language]

01:58 Omicron: It's basically the dentist for the oral mouth health, this is totally perfect to have the repetition of doctor while I can of resorted to taking that out from the target text. And from the TM there was part that was omitted from the translation which is...

[foreign language]

02:34 Omicron: So I had to translate that since it wasn't there. Overall for the English, it was kind of easier to really have some sort of like terminology in there it made it easier rather than...

02:53 S1: English to Arabic or from Arabic into English?

02:56 Omicron: From the Arabic into English, made it kind of easier doing the post editing that translating this myself, but I had to actually read the Arabic text at least twice in order to get the image because it wasn't, it seems to be translated from English, and it was not really a perfect translation, so I had to read it like more than once to get the

meaning and then check the translation with the source text. The Arabic, because this is machine translation, it basically kept a kind of formal translation of the source text which is the English text so I had to kind of re-write because it wasn't clear in Arabic, and even this title in Arabic that I got from the TM wasn't really smooth, did not really have a flow when you read in Arabic.

03:58 Omicron: So, I had to change, and I had to basically take some parts fore ground it, make it at the beginning rather than having it at the end. What else? Yeah, especially the first sentence in Arabic I had to do a lot of like addition to it in order to have a kind of [04:30] Parallel style. I kind of don't remember the translation that was in there otherwise I could have been like this one.

[foreign language]

04:42 Omicron: This is the intended meaning I guess from the English text.

[foreign language]

04:49 Omicron: And that's it has a distorted body image which was literally translated.

[foreign language]

04:56 Omicron: Something like that.

05:00 S1: So you are talking about the literary translation, the MT.

05:04 S1: Yeah.

05:04 Omicron: Okay. How do you... How can you rate the output of the tools [05:08] both tasks, the output that you got.

05:13 Omicron: The output, efficiency-wise, I would give it around 80% since the translation is there and I had to only post edit which would take only 10 minutes max, rather than translating the whole thing. style-wise, language-wise, it's not really that much efficient.

05:45 S1: How can you rate the tools in general?

05:48 Omicron: I would give it around 50% because I had to do a lot of post editing basically in there.

06:01 S1: Okay, do you think the tools helped your productivity?

06:07 Omicron: Yes, it did.

06:09 S1: And why? How?

06:12 Omicron: Because the translation was there like, in terms of the disease name for example, I wouldn't really have to think a lot about the meaning and I wouldn't really go to dictionary if I don't have the meaning. So, since it's there, it just makes perfect sense for me.

06:33 S1: Okay. What limitations do you think the tools have from [06:38] and this experience?

06:42 Omicron: The tool is not really smart enough till now, it definitely needs a lot of [06:50] coropora in order to develop a systematic and a stylistic kind of language in Arabic. And I know this is kind of hard... Yeah, and that's why it requires a lot of post-editing. And there is one thing that I noted, from Arabic into English, it was easier and it required kind of less post-editing, which wasn't the case from English into Arabic. From English into Arabic, it had to be restructured a lot.

07:30 S1: Okay, what difficulties did you face during the translation tasks? Have you had any difficulties... Like punctuations, segments...

07:48 Omicron: Yes, I had to check for...

07:50 S1: Spelling?

07:51 Omicron: Yes, I had to check for missing segments as I noted that in the English translation, from English into Arabic, there was basically a segment that was not translated. So this required me kind of like more time to read and compare and contrast between the two texts. The punctuation in Arabic kind of copied the one in English, where I had to add things like wa it's not only commas that we have in Arabic, but we have the wa instead of the commas.

08:31 S1: Hamza?

08:32 Omicron: Yes, this is one thing that I had to edit all across the text, like [08:37] they did not all have Hamza [08:43] which I had to also had edit myself. And even on terms of grammar, I got the translation, [09:01]

09:02 S1: This is from English to Arabic?

09:06 Omicron: From English to Arabic yeah, which is not correct grammatically. It would be [09:09] instead of [09:10]

09:12 S1: So it doesn't recognize the plural? Arabic?

09:17 Omicron: [09:17] exactly. Even [09:20] referring to the group instead of [09:24] and this is why I added the plural morpheme [09:26] here. And [09:38] Yeah.

09:46 S1: Okay, did you... Where is it? Have you changed your general assessment after this experience or it's the same? After experience the [09:58] MT and TM at the same time.

10:03 Omicron: Since I had access to the TM before, I feel the translation memory is almost the same, having the same problems, it did not really develop a lot but it's still of much help in terms of time-wise. Rather than translating everything from scratch, all you have to do is just to kind of post-edit which would definitely result in less time than translating the whole thing from zero, from scratch.

10:41 Omicron: It depends on the purpose of my translation if my purpose is to produce a kind of perfect translation, I wouldn't honestly use a translation memory. But if I have lots of clients with lots of translation tasks, I would definitely resort to a translation memory, because it kind of facilitates the translation.

11:03 S1: [11:03] What about MT?

11:06 Omicron: With a post edit, I will use it definitely, especially that it's getting kind of smarter and smarter, better than before.

11:16 S1: But you have not to change your perspective?

11:19 Omicron: No, yeah.

11:21 S1: Okay, which task was harder and why? I think you answered this.

11:25 Majd: Yes, it was translating from English into Arabic.

11:28 S1: Was harder?

11:28 Omicron: Yes.

11:28 S1: Okay. And which segment each translation task was the hardest one?

11:35 Omicron: It was from Arabic into English. In the first paragraph in particular, I had this segment.

11:47 S1: This is from Arabic to English?

11:49 Omicron: Yeah, from Arabic into English.

11:52 S1: From Arabic to English it's segment two?

11:54 Omicron: Yeah, segment two.

11:57 S1: And from English to Arabic?

12:00 Omicron: From English into Arabic it was segment two as well.

12:11 S1: Okay, not three?

12:17 Omicron: No, not three. Because I had to post-edit. But the second segment in this translation from English into Arabic, the second one the translation was not really good translation, the Arabic translation and that's why it took some time.

12:36 S1: Which segment was the easiest in both?

12:44 Omicron: For both, it was the title, the first one.

12:53 S1: Did you have any difficulties with segmentation or punctuation in both Arabic and English? You don't have to answer this but, if you want...

13:03 Omicron: Yeah, I had a difficulty with... This is part of post-editing, I had to take out the commas and add the wa to reflect the addition in Arabic and the one from Arabic into English, I had to break the Arabic sentence into shorter sentences in order to have a better style that is consistent to the English language, yeah.

13:35 S1: Okay, what adjustments do you have for the tools that you...

13:41 Omicron: To include [13:42] Corpora, a lot of [13:48] Corpora in order to be able to detect at least a closed-in structure to the Arabic language in order to be smarter. And it depends on the text type as well. We need definitely lots of medical texts in order to produce something that is considered at least to be efficient.

14:12 S1: Okay, and is there any other concerns you like to share, that you experienced with the tools today?

14:23 Omicron: My only concern is to really look at both translations and not really depend on the translation that you get from the TM or the MT because it seems to be that there are parts which are missing, and you have to have a careful eye to look at these parts, because it's not sometimes a phrase but it's a kind of word inside the text. So you would have to really compare the both of the...

14:53 S1: Isn't that time-consuming or...

14:58 Omicron: It's within the process of translation and post-editing, but it takes time but it's still better than translating the whole thing from scratch.

15:07 S1: What about the spaces and punctuation? Do you have any problem with the spaces?

15:15 Omicron: Yes, because I had to go and edit the one that I had in the Arabic at least and I had to change all those without Hamza into Hamza.

15:27 S1: What about the spelling checker do you think it was doing a good job of...

15:36 Omicron: It would do a perfect job for English but not really with... Let me check. So this one works fine. It does not always provide the correct one. It is good to have it in there for the Arabic, but it's not as perfect as with the English, with the English it was way better. Yeah, also I have comment about bilingual texts in the same segment, there is an issues with Arabic as right-left language in CAT tools, in case you get English and Arabic words in the same segments, you will get issues with formatting.

16:02 S1: Okay, thank you so much.

16:03 Omicron: You're welcome.

Appendix K: Participants' Translations

Arabic-English MT and TM suggestions

How to Recognize Signs of Oral cancer

Mouth and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America, the early detection of mouth cancer diseases in the appropriate time is important to raise the prospects of survival, life rate for five years infected oral cancers who did not have the disease spread is 83%, for example, while only 32% after the cancer is spread to other parts of the body. Although your doctor or dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis.

English-Arabic MT and TM Suggestions

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, ويكون له صورة مشوهة عن جسمه وعنده خوف حاد من الزيادة في الوزن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية خطير للغاية الاضطراب الغذائي قد يؤدي إلى خفض حدة الجفاف، وضغط الدم، وخسارة كثافة العظام وإغماء بين النتائج الأخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكن التغلب عليها مع المزيج المثالي المادي والنفسي والاجتماعي.

Alpha

Arabic to English Translation

Recognizing Signs of Oral Cavity Cancer

Mouth and throat cancers represent 2% of all cancers diagnosed each year in the United States. Early detection and treatment of oral cavity cancers are important in increasing prospects of survival. For example, the survival rate, after five years, for patients suffering from oral cavity cancer which did not spread to other parts of the body is 83%. While only 32% survived after the cancer spread to other parts of the body. Although your regular doctor and dentist are trained to detect oral cavity cancers, recognizing the signs yourself may facilitate an earlier diagnosis and timely treatment. More awareness is for the better.

Average number of words per sentence:	Number of words:	Number of sentences:
18.17	109	6

English to Arabic Translation

التغلب على مرض فقد الشهية

يعاني الفرد من مرض فقد الشهية عندما يمتنع عن تناول الطعام والشراب اللازمين للحفاظ على وزن صحي، وعندما يكون لديه صورة مشوهة عن جسده، وخوف حاد من زيادة الوزن، يعتبر فقد الشهية من الاضطرابات الخطيرة للغاية والتي قد تؤدي إلى الجفاف الحاد، وانخفاض ضغط الدم، وانحسار في كثافة العظام، والإغماء، وغيرها من العواقب. لحسن الحظ يمكن للذين يعانون من فقدان الشهية التغلب على هذا الإضطراب بالحصول على مزيج متوازن من العلاجات البدنية والنفسية والاجتماعية.

Average number of words per sentence:	Number of words:	Number of sentences:
39.50	79	2

Beta

Arabic to English Translation

How to Recognize Signs of Oral cancer

Oral and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America. The early detection and treatment of oral diseases is important to raise the prospects of survival. The rate of extending life for five years for those infected with oral cancer who were early diagnosed is 83%, for example, while it is only 32% for those were not diagnosed early on, and therefore cancer spread to the other parts of the body. Although your primary care physician and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis. The more aware you are, the better.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
23.40	117	5

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, ويكون له صورة مشوهة عن جسمه، وعنده خوف حاد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية اضطراب غذائي خطير للغاية قد يؤدي إلى جفاف حاد، وانخفاض ضغط الدم وهشاشة العظام وإغماء ضمن أعراض أخرى. ومن حسن الحظ فإن معظم الأشخاص الذين يعانو من مرض فقدان الشهية يمكنهم التغلب عليه باتباع نظام سليم جسدياً ونفسياً واجتماعياً.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
28	84	5

Gamma

Arabic to English translation

How to Recognize Signs or Oral Cancer

Cancers of the oral and throat represent 2% of all cancers that are diagnosed each year in the United States of America. The early detection of oral cancers diseases and treating it in time is important to raise the prospects of survival. Life rate is five years for people infected by oral cancers who did not have the disease spread is 83%, for example, while only 32% after the cancer spread to other parts of the body. Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis.

Average number of words per sentence:	Number of words:	Number of sentences:
26	104	4

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي، ويكون له صورة مشوهة عن جسمه، و عنده خوف حاد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية هو إضطراب غذائي خطير للغاية قد يؤدي إلى خفض حدة الجفاف، وضغط الدم، خسارة كثافة العظام والإغماء من بين النتائج الأخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليها مع المزيج المثالي البدني والنفسي والاجتماعي.

Average number of words per sentence:	Number of words:	Number of sentences:
29	87	3

Delta

Arabic to English translation

How to Recognize Signs of Oral Cancer

Oral (Oropharyngeal) cancers represent approximately 2% of all cancers that are diagnosed each year in the United States of America. The early detection of Oral Cancers in time is important to raise the chances of survival knowing that persons with oral cancers that did not spread have 83% five-year longer life expectancy while it goes down to only 32% after the cancer has spread to the other parts of the body. Although your dentist or doctor is trained to detect oral cancers, recognizing the signs yourself may help in an earlier diagnosis and treatment. The more awareness you possess the better.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
23.40	117	5

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي، ويكون له صورة مشوهة عن جسمه، و عنده خوف حاد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية هو إضطراب غذائي خطير للغاية قد يؤدي إلى جفاف حاد، وانخفاض ضغط الدم, ونقص كثافة العظام وإغماء وأعراض أخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليها مع المزيج المثالي من العلاج البدني والنفسي والاجتماعي.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
28.67	86	3

Epsilon

Arabic to English translation

How to Recognize Signs of Oral cancer

Oral and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America. The early detection and treatment of oral diseases is important to raise the chances of survival. Life expectancy of five years for people infected with oral cancer who did not have the disease is 83%, for example, while only 32% after the cancer spread to other parts of the body. Although your physician and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis and timely treatment. The more you are aware, the better it is for you.

Average number of words per sentence:	Number of words:	Number of sentences:
23	115	5

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, ويكون له صورة مشوهة عن جسمه، وعنده خوف حاد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية هو اضطراب غذائي خطير للغاية وقد يؤدي إلى جفاف شديد، انخفاض ضغط الدم، خسارة كثافة العظام وإغماء ونتائج أخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليه بمساعدة العلاج البدني والنفسي والاجتماعي.

Average number of words per sentence:	Number of words:	Number of sentences:
27.67	83	3

Zeta

Arabic to English Translation

How to Recognize Signs of Mouth Cancer

Mouth cancers and throat represent approximately 2% of all cancers that are diagnosed each year in the United States of America, the early detection of mouth cancers in time is important to raise the prospects of survival, life rate for five years infected mouth cancer who did not have the disease spread is 83%, for example, while only 32% after the cancer has spread to other parts of the body. Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may help in an earlier diagnosis and treatment, and the more you become aware the better.

Average number of words per sentence:	Number of words:	Number of sentences:
54	108	2

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, ويكون له صورة مشوهة عن جسمه، وعنده خوف حاد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية خطير للغاية الاضطراب الغذائي قد يؤدي إلى انخفاض حاد من الجفاف، وضغط الدم، وخسارة كثافة العظام وإغماء وعواقب أخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليها مع العلاج البدني والنفسي والاجتماعي المناسب.

Average number of words per sentence:	Number of words:	Number of sentences:
28	84	3

Theta

Arabic to English Translation

How to Recognize Signs of Oral cancer

Oral and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America. The early detection and treatment of oral diseases is important to raise the prospects of survival that life rate for five years infected oral cancers before the disease spread is 83%, for example, while it is only 32% after the cancer spread to the other parts of the body. Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis. Also, medication at an early stage is helpful. The more consciousness you have the better you live.

Average number of words per sentence:	Number of words:	Number of sentences:
22.60	113	5

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يكون للمريض صورة مشوهة عن جسمه، وخوف شديد من الزيادة في الوزن، ويمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن الجسم الصحي, هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية اضطراب غذائي خطير للغاية قد يؤدي إلى جفاف حاد، وضغط الدم، وخسارة كثافة العظام وإغماء والعديد من العواقب أخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكن التغلب عليها بالعلاجات البدنية والنفسية والاجتماعية.

Average number of words per sentence:	Number of words:	Number of sentences:
26.67	80	3

Eta

Arabic to English Translation

How to identify signs of mouth cancer

Mouth and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America. Early detection of mouth cancer diseases is important to provide the necessary treatment in time, which will raise the prospects of survival. Life rate for five years infected oral cancers who did not have the disease spread is 83%, for example, while only 32% after the cancer spread to other parts of the body. Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
25.25	101	4

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على صحة، ويكون له عقله صورة مشوهة عن جسمه، ولديه خوف من الزيادة في الوزن، فهذا مؤشر على أن الشخص يعاني من مرض فقدان الشهية. مرض فقدان الشهية خطير للغاية وله أضرار عدة كالجفاف، وانخفاض في ضغط الدم، هشاشة العظام والإغماء. ولحسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية بإمكانهم التغلب عليها بالعلاج الطبي والدعم النفسي والاجتماعي.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
25.67	77	3

Iota

Arabic to English Translation

How to Recognize Signs of Oral cancer

Oral and laryngeal cancers make up for about 2% of all cancers diagnosed in the United States of America. The detection of oral cancers in the right time is important to raise the chances of survival. Those with oral cancers have 83% chance to live for an average of five years if the disease has not spread to other parts of their body, whereas this percentage drops to 32% if the disease has spread to other parts. Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
26	104	4

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, ويكون لديه صورة مشوهة عن جسمه، ويتملكه خوف حاد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية هو أحد اضطرابات الأكل الخطيرة للغاية والتي قد تؤدي إلى جفاف حاد، وانخفاض في ضغط الدم، ونقص كثافة العظام والإغماء على سبيل المثال لا الحصر. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليها مع مزيج مثالي من العلاج الجسدي والنفسي والاجتماعي.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
30.67	92	3

Lambda

Arabic to English Translation

How to identify signs of mouth cancer

Mouth and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America, the early detection of mouth cancer diseases in the appropriate time is important to raise the prospects of survival, life rate for five years infected oral cancers who did not have the disease spread is 83%, for example, while only 32% after the cancer is spread to other parts of the body. Although your doctor or dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis and cure in the appropriate time. The more conscious you are, the better your health will be.

Average number of words per sentence:	Number of words:	Number of sentences:
38.67	116	3

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, ويكون له صورة مشوهة عن جسمه مما يؤدي إلى خوف حاد من الزيادة في الوزن ومما يجعله يعاني من مرض فقدان الشهية. فقدان الشهية خطير للغاية الاضطراب الغذائي قد يؤدي إلى خفض حدة الجفاف، وضغط الدم، وخسارة كثافة العظام وإغماء بين النتائج الأخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكن التغلب عليها ممن خلال من العلاج الجسدي والنفسي والاجتماعي.

Average number of words per sentence:	Number of words:	Number of sentences:
28.67	86	3

Kappa

Arabic to English Translation

How to identify signs of mouth cancer

Mouth and throat cancer account for 2% of the all cancers that are diagnosed every year in the US. The timely detection and treatment of mouth cancer is important to increase the prospects of survival. Expectancy rate for five years for those with mouth cancer whose disease did not spread is 83% while it is 32% after the disease spreads to other parts of the body. Although your doctor and dentist are trained to detect oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis and timely treatment. The more you are aware, the better.

Average number of words per sentence:	Number of words:	Number of sentences:
20.60	103	5

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, وتكون صورة جسمه مشوهة، وعنده خوف حاد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية هو اضطراب غذائي خطير للغاية قد يؤدي إلى جفاف حاد، وانخفاض ضغط الدم، وفقد في كثافة العظام والإغماء من بين أشياء أخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليها عن طريق المزيج المثالي من العلاج البدني والنفسي والاجتماعي.

Average number of words per sentence:	Number of words:	Number of sentences:
29.33	88	3

Mu

Arabic to English Translation

How to Recognize Signs of Oral cancer

Oral and throat cancers represent 2% of all cancers that are diagnosed each year in the United States of America. The early detection and treatment of oral diseases is important to raise the prospects of survival. 83% of oral cancer patients who have not experienced metastasis have five years life expectancy, where only 32% of oral cancer patients who experienced metastasis have the same life expectancy. Although your doctor and dentist are trained to diagnose oral cancers, recognizing the signs yourself may facilitate an earlier diagnosis and treatment.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
23.75	95	4

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

عندما يمتنع شخص ما عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن صحي الجسم, وتكون له صورة مشوهة عن جسمه، وعنده خوف شديد من الزيادة في الوزن، فإن هذا الشخص يعاني من مرض فقدان الشهية. فقدان الشهية هو اضطراب غذائي خطير للغاية قد يؤدي إلى الجفاف الحاد، وانخفاض ضغط الدم، وخسارة كثافة العظام ومضاعفات أخرى مثل الإغماء. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليها عن طريق العلاج الفيزيائي والنفسي والاجتماعي.

Average number of words per sentence:	**Number of words:**	**Number of sentences:**
28.33	85	3

Omicron

Arabic to English Translation

How to Recognize Signs or Oral Cancer

Cancers of the mouth and throat represent 2% of all cancers that are diagnosed each year in the United States of America. The early detection of oral cancer and its treatment in time is important to raise the probability of survival rate. The expected life rate is five years for those infected with 83% of oral cancer and only 32% after the cancer spread to other parts of the body. Although your dentist is trained to detect oral cancers, recognizing the signs yourself may facilitate early diagnosis. The more you are aware, the better you will be.

Average number of words per sentence:	Number of words:	Number of sentences:
20.80	104	5

English to Arabic translation

كيفية التغلب على مرض فقدان الشهية

يعاني الشخص من مرض فقدان الشهية عندما يمتنع عن تناول كميات من الطعام والشراب اللازمة للحفاظ على وزن جسم صحي, من أجل أن لا تكون له صورة مشوهة عن جسمه، مما يسبب لديه خوف حاد من الزيادة في الوزن. يعتبر فقدان الشهية مرض خطير للغاية قد يؤدي إلى الإصابة بجفاف حاد، وانخفاض في ضغط الدم، وخسارة في كثافة العظام وإغماء إضافة إلى أمراض أخرى. ومن حسن الحظ ان معظم الأشخاص الذين يعانون من فقدان الشهية يمكنهم التغلب عليه من خلال مجموعة من العلاجات الجسمانية والنفسية والاجتماعية.

Average number of words per sentence:	Number of words:	Number of sentences:
30.67	92	3

Appendix L: Interviews and Screen Recordings Quantitative Data

Interviews Quantitative Data

Participants ID	Evaluation of tools (out of 10)	Segmentation Issues Ar-Ee	Segmentation Issues En-Ar	Punctuation Issues Ar-En	Punctuation Issues En-Ar	Spelling Ar-En	Spelling En-Ar	Hardest task	Hardest segment Ar-EN
Omicron	5	Yes	No	Yes	Yes	No	Yes	En-Ar	Second
Mu	6	Yes	No	Yes	Yes	No	Yes	Ar-En	Second
Kappa	6	Yes	No	No	No	No	Yes	Ar-En	Second
Lambda	7	No	No	No	Yes	No	Yes	En-Ar	Second
Iota	7	Yes	No	No	No	No	No	Ar-En	Second
Eta	6	No	No	Yes	Yes	No	Yes	En-Ar	Second
Theta	7	Yes	No	No	Yes	No	Yes	Ar-En	Second
Zeta	9	Yes	No	Yes	Yes	No	Yes	Ar-En	Second
Epsilon	8	No	No	Yes	No	No	Yes	Ar-En	Second
Delta	6	Yes	No	Yes	Yes	No	Yes	Ar-En	Second
Gamma	6	Yes	No	Yes	Yes	No	Yes	Ar-En	Second
Beta	7	Yes	No	Yes	Yes	No	Yes	Ar-En	Second
Alpha	7	No	No	No	No	No	Yes	Ar-En	Second

Hardest segment En-Ar	Did tools Help your Productivity ?
Second	Yes
Second	Yes
Third	Yes
Fourth	Yes
Third	Yes
Third	Yes
Second	Yes
Third	Yes
Second	Yes
Second	Yes
Second	Yes
Third	Yes
Third	Yes

Screen Recordings Quantitative Data

Participant ID	Time spent for Arabic-English Translation	Time Spent for English-Arabic Translation	Total Time for both Tasks	Using Online Resources for Arabic-English	Using Online Resources for English-Arabic	Changes in Arabic-English TM	Changes in English-Arabic TM	Addressing Fuzzy Match Arabic-English	Complete Translation of Fuzzy Match Arabic-English
Omicron	10.38	8.11	18.49	No	No	Yes	Yes	Yes	No
Mu	15	9.14	24.14	Yes	No	No	Yes	No	No
Kappa	8.2	3.46	12.06	Yes	No	No	No	Yes	Yes
Lambda	5.2	5.06	10.26	No	No	Yes	Yes	Yes	Yes
Iota	9.07	7	16.07	Yes	Yes	No	Yes	No	No
Eta	10.12	9.27	19.39	Yes	No	No	Yes	No	No
Theta	7.22	6.06	13.28	No	No	No	Yes	Yes	Yes
Zeta	4.39	4.27	9.06	No	No	Yes	No	Yes	Yes
Epsilon	8.44	6.07	14.51	No	No	Yes	No	Yes	Yes
Delta	19.29	6.45	26.14	Yes	No	Yes	No	Yes	Yes
Gamma	5.43	4.34	10.17	Yes	No	No	No	No	No
Beta	6.13	5.3	11.43	No	No	Yes	No	Yes	No
Alpha	19.17	11.08	30.25	Yes	Yes	Yes	Yes	Yes	Yes

Extensive Changes to Arabic-English MT Suggestions	Extensive changes English-Arabic MT Suggestions	Average number of words per sentence; Arabic-English	Number of Words; Arabic-English	Number of Sentences; Arabic-English	Average number of words per sentence; English-Arabic	Number of Words; English-Arabic	Number of Sentences; English-Arabic	CAT tools suggestions: Average number of words per sentence; Arabic-	CAT tools suggestions: Average number of words per sentence; English-
Yes	Yes	20.8	104	5	30.67	92	3	45	27.76
Yes	Yes	23.75	95	4	28.33	85	3	45	27.76
Yes	No	20.6	103	5	29.33	88	3	45	27.76
No	No	38.67	116	3	28.67	86	3	45	27.76
Yes	Yes	26	104	4	30.67	92	3	45	27.76
Yes	Yes	25.25	101	4	25.67	77	3	45	27.76
Yes	Yes	22.6	113	5	26.67	80	3	45	27.76
No	No	54	108	2	28	84	3	45	27.76
Yes	No	23	115	5	27.67	83	3	45	27.76
Yes	No	23.4	117	5	28.67	86	3	45	27.76
No	No	26	104	4	29	87	3	45	27.76
Yes	Yes	23.4	117	5	28	84	3	45	27.76
Yes	Yes	18.17	109	6	39.5	79	2	45	27.76

Appendix M: Report of Participants Responses in the Online Survey

Detailed Report of Participants Responses in the Online Survey

The Use of CAT Tools for Arabic Translation: User Evaluation, Issues, and Improvements

Note: The written responses of participants are demonstrated as they are without editing or correcting errors.

Q1 - Please choose one: I'm currently

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Please choose one: I'm currently	1.00	5.00	2.91	1.20	1.43	55

#	Answer	%	Count
1	Working as a translator for a LSP.	20.00%	11
2	Working in a profession that requires translation tasks.	5.45%	3
3	A freelance translator.	49.09%	27
4	A freelance translator and a graduate student at translation program.	14.55%	8
5	Graduate student in translation program but don't work currently as a translator.	10.91%	6
	Total	100%	55

Q2 - How many years of translation experience do you have?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How many years of translation experience do you have?	1.00	4.00	2.80	0.88	0.78	55

#	Answer	%	Count
1	None	1.82%	1
2	1-5 years	45.45%	25
3	6-9 years	23.64%	13
4	10+ years	29.09%	16
	Total	100%	55

Q3 - What are your academic qualifications, if any?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	What are your academic qualifications, if any?	1.00	3.00	2.56	0.65	0.43	55

#	Answer	%	Count
1	PhD completed/ongoing	9.09%	5
2	MA completed/ongoing	25.45%	14
3	BA completed/ongoing	65.45%	36
	Total	100%	55

Q4 - If your studies are ongoing, which year are you in?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	If your studies are ongoing, which year are you in? - Selected Choice	1.00	6.00	5.31	1.49	2.21	55

#	Answer	%	Count
1	1st year	5.45%	3
2	2nd year	7.27%	4
3	3rd year	0.00%	0
4	4th year	1.82%	1
5	5th year	9.09%	5
6	Other, please specify:	76.36%	42
	Total	100%	55

Data source misconfigured for this visualization

#	Answer	%	Count
Graduated	Graduated	3.33%	1
A Graduate	A Graduate	3.33%	1
Competed	Competed	3.33%	1
Completed	Completed	30.00%	9
Finished	Finished	6.67%	2
Finished BA	Finished BA	3.33%	1
grad.	grad.	3.33%	1
graduated	graduated	20.00%	6
I do not study now.	I do not study now.	3.33%	1
I have finished my study.	I have finished my study.	3.33%	1
MA research	MA research	3.33%	1
n/a	n/a	6.67%	2
None	None	6.67%	2
Not clear	Not clear	3.33%	1
	Total	100%	30

Q5 - How familiar are you with the current computer-assisted translation tools?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How familiar are you with the current computer-assisted translation tools?	1.00	3.00	1.57	0.64	0.41	49

#	Answer	%	Count
1	Extremely familiar	51.02%	25
2	Somewhat familiar	40.82%	20
3	Slightly familiar	8.16%	4
4	Not familiar at all	0.00%	0
	Total	100%	49

Q6 - From you experience, have you encountered issues and challenges in the use of computer-assisted translation tools for Arabic language?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	From you experience, have you encountered issues and challenges in the use of computer-assisted translation tools for Arabic language?	1.00	3.00	1.45	0.61	0.37	49

#	Answer	%	Count
1	Yes.	61.22%	30
2	No.	32.65%	16
3	Does not apply.	6.12%	3
	Total	100%	49

Q7 - If yes, what are the issue you faced while using the tools with Arabic?

If yes, what are the issue you faced while using the tools with Arabic?
If you work on cloud-based tools for more than 4 hours, the computer starts to freeze sometimes.
Segmentations
Most of these softwares don't support Arabic Pdf files
In Trados, there is always a red line under Arabic translation in some versions. Many professional translators have not known the reason for that. However, I found away to solve this. The problem is that we need every single minute because time is money.
I do not have any challenges.
no
I have problems with search in TM
No issues are found.
NA
Numbers, quality assurance
RTL issues, unfamiliarity with AR grammar and punctuation
There are many problems faced me when replacing the tags to the target Arabic. Another problem is preserving the same length of line as it is in the source English Text.
No
some tools have format issues. some do not offer checkspeller
Tags
Timing assistance
Term base don't recognize plural in Arabic
Numbers difference in memoQ frustrates very much.
Honestly, I didn't face any problems.
Tags
Type of file exchange

Formatting of target files.
Not a technical issue but I'd rather say that I lost some of my creativity due to CAT tools.
Punctuation, Numbers
If you change the case (upper, lower)of the first letter of a word in English, you must enter the word in the TB again. I think there are right in that as cases make distinction betwwen proper ans common nouns. However, I suggest that they enter it as an option in the TM or TB.
My answer is No
Technical problems
May be lost
In general, it took me some time to get use on setting separated segments while keeping there cohesion together.
I do not face any difficult tissue while using the tool
Fonts and tags But technical issues like inability to select all cells in Trados
Nothing
There are no issues
LRT Issues
the Arabic pdf files could not be transferred to any other editable format.
1- Being different in direction, Arabic creates all the problems. For ex. OCR problem is still existing especially when the tool converts a scanned document into a readable Arabic, the result is not 100% of good quality. So, most of the translators do not recommend to use OCR if the source document is in Arabic.
i'm not familiar with Arabic tools it's easier to be English
No issues
Contextual errors due to lack of context.
RTL/LTR direction Preview issue Segmentation issue
Locale issues.
Bi clolumnal Word pages, format corrupted, PDF files issues.
Non
N/A
Arabic right to left layout after cleaning Suggesting inaccurate translation from the TM dealing with client who are not willing to pay for the 100% match, while this type of match needs edits

- Lack of terminology glossaries; - Lack of effective tools that can maximize benefit from existing translation memories; - Misunderstanding of CAT tools and considering them as machine translation software; and - Misunderstanding of CAT tools as software translate themselves rather than software that help translators in translation process.
Direction of the text when we found English and Arabic in the same sentence
incompatibility with Arabic, word order, bad translation

Q8 - Which computer-assisted translation tool you have been using in your translation work; You can choose more than one.

#	Answer	%	Count
1	SDL Trados.	46.32%	44
2	MemoQ.	28.42%	27
3	Wordfast.	13.68%	13
4	Other	11.58%	11
	Total	100%	95

Other

Other - Text
Memsource
MemoQ
Passolo, Smartling, Translation Workspace
Memsource, Smartling
TWS - Memosource
wordpad for movie
Déja vu

Memosource - Matecat
Memsource
Wordbee
Memsource

Q9 - How do you rate your knowledge of any computer-assisted translation tool you have used? Rate from one to ten, where 1 is very poor and 10 is excellent:

How do you rate your knowledge of any computer-assisted translation tool you have used? Rate from one to ten, where 1 is very poor and 10 is excellent:
8
8
7
10
8
7
8
10
8
8
10
8
5
7
8
1
10
7
6
6
5

7
8
6
6
6
9
6
3
5
8
7
8
9
9
8
8
10
7
7
8
9
10
7
8
9
9

10

6

Q10 - Based on your experience with the tool, how would you rate your satisfaction level of the tools?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Based on your experience with the tool, how would you rate your satisfaction level of the tools?	1.00	4.00	1.88	0.75	0.56	49

#	Answer	%	Count
1	Very satisfied	32.65%	16
2	Satisfied	48.98%	24
3	OK	16.33%	8
4	Dissatisfied	2.04%	1
5	Very dissatisfied	0.00%	0
	Total	100%	49

Q11 - Do you have specific suggestions for improving the features of any tool you have not been fully satisfied with?

Do you have specific suggestions for improving the features of any tool you have not been fully satisfied with?
MemQ; It should recognize Arabic language. Sometimes, when I upload a file, it shows the letters as symbols, not as Arabic letters.
Better segmentation, full editing ability to the original text on the editor's interface.
Just more compliance with the nature of arabic text
Trados is a good CAT tool but I think its Term Base need more improvement to be like MemoQ's TB. Generally, there we can not do without all CAT tools because different clients need different CAT tools.
No
it must have
Facilitate research and address problems of language pair
No
No
No
No, I don't
I want to make all cat tools desktop tools.
Yes
None, unfortunately.
Mo
Still need help for assistance
Pop up window in memq
Using voice in Trados like memoQ has. It would save time and effort but shall not be limited for lphone users.
Yeah, what is bad about these tools is that they are slow as they require high RAMs. Even though my computer is excellent, most of these tool still crash.
No, I don't
Probability to edit image inside the tools to translate words

More improving to the formatting of target files.
An auto correcting system for my saved terms depends on data provided by me to the system every now and then.
Arabize it.
I didn't use other tools. But I tried to deal with Trados. It needs to omit many steps in creating TMs. Tags are the often problem the face my fellow colleague.
No
No
No not yet
Trying to improve machine translation to help translators and reading PDF files.
I already use SDL trados and this tool excellnt tool
In Trados, when filtering anything, the select all option is not applicable which waste a lot of time
No
Link to babylon Dictionary
NA
No
Working on the OCR for Arabic Language is a great step towards improvement.
Nope
Dont have any suggestions now
Translation memory search features can be improved in Wordbee.
The more simple, the more effective Instance preview RTL for Arabic No apparent segmentation
Improving translation of PDF format.
TM Teaking, fuzzy match repair by MT. Multiple search and replace.
No
N/A
No
- Improve CAT Tools' compatibility with Arabic Language to be more smart and responsive. - Improve terms extraction process to improve translators' productivity. - Improving the competency of CAT Tools in understanding the structure of Arabic sentences, phrases, expressions and words.

yes fixing bugs in memoQ and better support for Arabic in Fragments assembling feature

Improve quality for arabic

Q12 - In your opinion, how important is the use of CAT tools by Arabic translators?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	In your opinion, how important is the use of CAT tools by Arabic translators?	1.00	2.00	1.16	0.37	0.14	49

#	Answer	%	Count
1	Very important	83.67%	41
2	Somewhat important	16.33%	8
3	Slightly important	0.00%	0
4	Not at all important	0.00%	0
5	Don’t know	0.00%	0
	Total	100%	49

Q13 - How familiar are you with machine translation tools for Arabic?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How familiar are you with machine translation tools for Arabic?	1.00	3.00	1.78	0.72	0.52	46

#	Answer	%	Count
1	Extremely familiar	39.13%	18
2	Somewhat familiar	43.48%	20
3	Slightly familiar	17.39%	8
4	Not familiar at all	0.00%	0
	Total	100%	46

Q14 - In your opinion, how important is the use of machine translation tools by Arabic translators?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	In your opinion, how important is the use of machine translation tools by Arabic translators?	1.00	4.00	1.83	1.05	1.10	46

#	Answer	%	Count
1	Very important	52.17%	24
2	Somewhat important	26.09%	12
3	Slightly important	8.70%	4
4	Not at all important	13.04%	6
	Total	100%	46

Q15 - Machine translation in general is useful for Arabic language translators;

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Machine translation in general is useful for Arabic language translators;	1.00	3.00	1.30	0.55	0.30	46

#	Answer	%	Count
1	True	73.91%	34
2	False	21.74%	10
3	Don't Know	4.35%	2
	Total	100%	46

Q16 - Some translation memory systems integrate MT suggestions when there is no TM for a given translation segment. Do you anticipate that the MT suggestions will increase Arabic translators' productivity?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Some translation memory systems integrate MT suggestions when there is no TM for a given translation segment. Do you anticipate that the MT suggestions will increase Arabic translators' productivity?	1.00	5.00	1.93	1.09	1.19	46

#	Answer	%	Count
1	Strongly agree	39.13%	18
2	Somewhat agree	45.65%	21
3	Neither agree nor disagree	4.35%	2
4	Somewhat disagree	4.35%	2
5	Strongly disagree	6.52%	3
	Total	100%	46

Q17 - My perspectives of machine translation for Arabic language is:

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	My perspectives of machine translation for Arabic language is:	1.00	2.00	1.26	0.44	0.19	42

#	Answer	%	Count
1	Positive	73.81%	31
2	Negative	26.19%	11
	Total	100%	42

Q18 - What limitation do you think the new computer-assisted translation tools have for Arabic language?

What limitation do you think the new computer-assisted translation tools have for Arabic language?
Directionality, segmentation,
Most of CAT tools depend on English-oriented spell checkers and proofing, so they don't provide suitable and accurate correction alternatives
Some translations are good but others aren't because they sometimes translate words not phrases or sentences. I think they need more improvement and time. I believe that they really THREATEN our career as human translators.
Format
lack of insert non-transferable files
Many
Don't know
Literature and marketing text
I guess: Arabic has its specific morphological features.
not acpetabel
only format issues.
التشكيل diacritics
Perhaps tashkeel “diacritics”
Short sentences

The major limitations are related to Arabic itself, Arabic, I believe, is not given much importance as other languages within CAT tools.
Differentiating punctuation marks, quotation marks, help with proper names, identify somewhat similar segment and suggest better translations based on these differences.
Multi meaning
The meaning of words which differ according to context.
The diversity of Arabic semantics compared with English ones.
Many

Not very much. The cases matters. The construction of the sentence as the tool dissect the English sentences and of course overlooks the links between them. And this affects Arabic.
I have no idea.
I don't know
I don't know
Reading PDF files
Can not be used in artistic or creative approaches
Don’t know
No limitations
There is no limitation
LRT issues and final output
it is very good
1- OCR 2- Punctuation of Arabic (in the quality assurance test). 3- Segmentation of the sentence (sentence order). 4- Tags problems (due to differences between Arabic & English). 5- Text Expansion (especially when translating into Arabic: UAE المتحدة العربية الإمارات). post- clean edit process and a lot more...
i think slow program
I don’t understand the question well.
Direction (RTL). All translation tools have a room for improvement in terms of direction.
Do no know
I think limitations will be minimum
Poor terms For some domains Poor phrases Poor idioms
Mo
N/A
segmentation, not compatible with Arabic complex sentence structure not compatible with Arabic right to left layout
- Data-related limitations. - resources-related limitations. - limitations related to the nature and structure of Arabic Language.
no limitation just need further customization for the Arabic text
there are a lot of limitations