

**PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA**  
**MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH**

**Mohamed Seddik Ben Yahia University – Jijel**

**Faculty of Letters and Languages**

**Department of English**



**Investigating the Effect of Pre-editing on Machine Translation's Final Output**

**Dissertation Submitted in Partial Fulfillments of the Requirements for the Master  
Degree in English Didactics**

**Submitted by:**

**Amira IZRI**

**Meriem ZINE**

**Supervised by:**

**Bakir BENHABILAS**

**Board of Examiners:**

**Chairperson: Dr. Fateh BOUNAR**

**Examiner: Pr. Ammar BOUKRIKA**

**Supervisor: Mr. Bakir BENHABILAS**

## Dedication

*I dedicate this work to:*

*The light of my life, my parents, brothers and sisters, who have shared their words of advice and have always been the source of inspiration and encouragement;*

*My dear Fiancé “Zayn”, who continually provided me with moral and emotional support that gave me the strength to never give up;*

*My beloved nephews “Moundir” and “Youcef” and my niece “Noursine”; for being the source of hope and happiness;*

*Finally, my best friend “Meriem”, whom I consider as my sister, for being the kindest, sincerest and most generous friend I have ever had; for her unlimited support and help throughout the whole years of university studies; and for her great efforts that were expended to finish this work.*

*Amira*

## Dedication

*In the name of the almighty God, the Most Gracious, the Most Merciful. All praises to Allah for the strengths and His blessing in completing this thesis.*

*This work is wholeheartedly dedicated to:*

***My beloved mother**, the source of motivation and inspiration, whose love and support have sustained me throughout life, and who prepared me to face the challenges with faith and humility.*

***My late father**, who always had confidence in me and offered me strength and encouragement in all my endeavours.*

*My lovely little sister '**Marwa**' and beloved brothers '**Anis**' and '**Bilal**', who constantly encouraged and supported me throughout my life journey.*

*My favourite cousins, for their endless support.*

*My classmate, my partner, my best supporter, and above all, my true best friend '**Amira**', with whom I shared unforgettable moments in life, who always believed in me and urged me to strive to achieve my goals.*

*Lastly, my friends, who shared their words of advice and encouragement.*

**Meriem**

## Acknowledgement

First of all, our deepest gratitude is expressed to “Allah” for giving us the strength to keep going regardless of the pitfalls we have come across.

We would like to express our special thanks to **Mr. Benhabilas Bakir**, our supervisor and our teacher, for his precious assistance, guidance and support to accomplish this work.

We also owe warm thanks to **Miss. Mezrag Sara** and **Mr. Boukrika Ammar** from the English department, **Mr Benamor Mohammed** from the French department, **Mr. Nouari Razzak**, **Mr. Gueham Toufik** and **Mr. Boulahia Mohammed** from the Arabic department who devoted a portion of their precious time to evaluate the translations we worked on and without whom the completion of our dissertation would not have been possible.

We are extremely grateful to **Mr. Hammouda Sami** from the management department for his valuable contribution in performing the statistical analysis of this research study.

Finally, we very much appreciate the great help and the knowledge that were received from our major and dear friend **Zeraoulia Asya**.

### **Abstract**

In a society with a state of the art technology and a globalised world with hundreds of languages, machine translation, as an interdisciplinary field, has become an integral part of foreign language learning. Its output quality, however, depends greatly on how one uses it. Hence, this research study primarily aims at investigating the effect of pre-editing on machine translation's final product. Accordingly, it is hypothesised that pre-editing has a positive effect on machine translation's final output; more precisely, there is a relationship between pre-editing quality and machine translation quality. To verify the aforementioned hypotheses, the researchers conducted an experimental based study with a set of 21 third year EFL students at the University of Mohamed Seddik Ben Yahia. This study was completed relying on a mixed-approach design, for the data were collected and analysed both qualitatively and quantitatively. The results obtained were very interesting. At first, the findings showed that most of the students were successful in applying the pre-editing rules and provided well-simplified texts. Second, the results corroborated the positive effect of pre-editing approach on machine translation's intelligibility and fidelity. Lastly, the findings revealed that there is a significant relationship between pre-editing quality and machine translation quality. To conclude, the current study provides some useful recommendations and points out the necessity of teaching students the concept of pre-editing for better translation quality.

**Key words:** Machine Translation, Pre-editing Approach

## **List of Abbreviations**

**MT:** Machine Translation

**ALPAC:** Automated Language Processing Advisory Committee

**DBA:** Direct-based Approach

**SL:** Source Language

**TL:** Target Language

**RBA:** Rule-based Approaches

**CBMT:** Corpus-based Machine Translation

**EBA:** Example-based Approach

**SMT:** Statistical Machine Translation

**HMTA:** Hybrid Machine Translation Approach

**HAMT:** Human Aided Machine Translation

**NMT:** Neural Machine Translation

**CLs:** Controlled Languages

**HOCL:** Human Oriented Controlled Languages

**MOCL:** Machine Oriented Controlled Languages

**CFE:** Caterpillar Fundamental English

**PEP:** Plain English Program

**ILSAM:** International Language of Service and Maintenance

**PACE:** Perkin's Approved Clear English

## List of Tables

**Table 1:** Frequency of Use of Pre-editing Rules

**Tables 2:** Descriptive Analysis of the Pre-edited Texts' Translation Produced by Google Translate

**Table 3:** Descriptive Analysis of the Raw Text's Translation Produced by Google Translate

**Table 4:** Descriptive Analysis of the Pre-edited Texts' Translation Produced by Microsoft Translate

**Table 5:** Descriptive Analysis of the Raw Text's Translation Produced by Microsoft Translate

**Table 6:** Descriptive Analysis of the Pre-edited Texts' Translation Produced by Systran

**Table 7:** Descriptive Analysis of the Raw Text's Translation Produced by Systran

**Table 8:** Comparison between Pre-edited and Raw texts' Translations in Terms of Fidelity

**Table 9:** Comparison between Pre-edited and Raw texts' Translations in Terms of Intelligibility

**Table 10:** Pre-editing and Translation Evaluation Scores

**Table 11:** The Pearson Correlation between Pre-editing and Machine Translation

## **List of Figures**

**Figure 1:** Machine Translation Architectures

**Figure 2:** Machine-aided Human Translation Model



## Table of Contents

Dedication.....	2
Dedication.....	3
Acknowledgment.....	4
Abstract.....	5
List of Abbreviations.....	6
List of Tables.....	7
List of Figures.....	8
Table of Contents.....	9
1. Introduction.....	13
2. Statement of the Problem.....	15
3. Research Questions.....	16
4. Hypotheses.....	16
5. Significance of the Study.....	16
6. Research Methodology.....	17
7. Organization of the Study.....	17
<b>Chapter One: Machine Translation and Pre-editing Approach</b>	
<b>1.1. Section One: Machine Translation.....</b>	<b>18</b>
Introduction.....	18
1.1.1. Definition of Machine Translation.....	18
1.1.2. Historical Development of Machine Translation.....	19
1.1.3. Approaches of Machine Translation.....	21
1.1.3.1 Direct Approach.....	21
1.1.3.2 Rule-based approach.....	23

PRE-EDITING AND MACHINE TRANSLATION	10
1.1.3.2.1. Interligua Approach.....	23
1.1.3.2.2 Transfer Approach.....	23
1.1.3.3. Corpus-based Approach.....	24
1.1.3.3.1. Example-based Approach.....	24
1.1.3.3.2. Statistical Approach.....	25
1.1.3.4. Hybrid Approach.....	25
1.1.4. Purpose of using Machine Translation.....	26
1.1.5. Issues of Machine Translation.....	27
1.1.6. Evaluation of Machine Translation.....	29
1.1.7. Machine-aided Human Translation.....	32
1.1.8. Human-aided Machine Translation.....	33
1.1.9 Neural Machine Translation.....	34
Conclusion.....	35
<b>1.2. Section Two: Pre-editing Approach.....</b>	<b>36</b>
Introduction .....	36
1.2.1 Definition of Pre-editing .....	36
1.2.2. Controlled Language.....	37
1.2.2.1. Emergence of Controlled Language.....	37
1.2.3. Pre-editing Approach.....	38
1.2.4. Advantages of Pre-editing on Machine Translation.....	41
Conclusion.....	41
<b>Chapter Two: Research Methodology and Data Analysis and Interpretation</b>	
<b>Section One: Research Methodology.....</b>	<b>43</b>
Introduction.....	43
2.1. Research Methodology .....	43

PRE-EDITING AND MACHINE TRANSLATION	11
2.1.1. Research Paradigm.....	43
2.1.2. Setting.....	44
2.1.3. Population and Sampling.....	45
2.1.4. Data Collection Procedures.....	45
2.1.4.1. Pre-implementation Phase.....	45
2.1.4.2. Implementation Phase.....	46
2.1.5. The Nature of Data Analysis.....	47
2.1.6. Limitations of the Study.....	50
Conclusion.....	51
<b>Section Two: Data Analysis and interpretation.....</b>	<b>52</b>
<b>Part One: Data Analysis.....</b>	<b>52</b>
Introduction.....	52
2.2.1.1. Analysis of pre-editing rules in the participants' texts.....	52
2.2.1.2 Analysis of the Translation Outputs.....	53
2.2.1.3. Correlation Analysis between the Quality of Pre-editing and Machine Translation...59	
Conclusion.....	61
<b>Part Two: Data Discussion and Interpretation.....</b>	<b>62</b>
Introduction.....	62
2.2.2.1. The Frequency of Use of pre-editing rules.....	62
2.2.2.2. The Effectiveness of Pre-editing on Machine Translation's Final Output.....	64
2.2.2.3. The Correlation Between Pre-editing Quality and Machine Translation Quality...64	
Conclusion.....	65
Recommendations.....	65
General Conclusion.....	66
<b>References.....</b>	<b>68</b>

**Appendices**

**Resumé**

**ملخص**

## 1. Introduction

The concept of translation has been neglected for a very long period of time as a result of the negative effect that traditional methods, particularly Grammar Translation method, had on the field of teaching foreign languages. Scholars claimed that it does not have a role in enhancing the learner's language because it does not serve the purpose of producing users of the foreign language with high communicative competence. However, recently, they realized the fact that previously, translation was just used as a means for explaining the meaning of words and sentences, whereas the core of translation is not only to give equivalents of terms and phrases in the target language. It rather plays a crucial role in interrelating and connecting various cultures which speak different languages all around the world. Thus, they revealed the fact that it should be considered as the fifth skill alongside with the four skills of language because it really has an important role in improving the communicative competence. When someone masters his language and the target language, they have the ability to deliver meanings and ideas through translation.

Nowadays, with technological development and huge dependence on electronic devices in different fields of teaching and learning in general and language learning in particular, machine translation appeared to be a very important means that has a vital role in the field of translation. In principle, it refers to the use of computers in performing translation tasks. Since computers are approaching to replace the human in many daily activities, it is undeniable that in the coming years, electronic systems will take the role of human translators. Thus, it is significantly important to teach the use of machine translation for EFL students as an attempt, firstly, to make them know about how to benefit from computer systems; secondly, it has a great effect on improving their language because it leads them to have a more profound linguistic and cultural knowledge through understanding the way the machine interprets the source text.

Generally speaking, computers and translation systems are unable to handle certain types of syntactic ambiguity and semantic analysis. Hence, to reach more satisfactory and adequate translations, many practices and approaches to machine translation have evolved. These approaches, however, are the result of the degree of human-computer collaboration. More specifically, this cooperation has resulted in the classification of machine translation technology. In addition to machine translation as a subtype, machine-aided human translation and human-aided machine translation are two of the approaches that involve the human translator at some point. The former refers to the human translation with the assistance of computer-based translation tools, such as programs for checking spelling, grammar, and online dictionaries. The latter is considered as the most important in which the system itself takes control with the human assistance to help at some point during the process. As stated by Katsburg (2012), “the human role can be compared to that of a consultant or an editor, i.e. that the translator corrects and modifies, what, in the machine’s translation suggestions, is unacceptable for him or her” (p. 42). In human-aided machine translation, the human can, basically, intervene at two major points. The first point is when the human checks and adapts the source text before running the translation process; it is the case of pre-editing. The second point is when the human translator corrects and revises the machine’s translation final output; it is the case of post-editing. While many studies in the field of machine translation have investigated the use of post-editing as an attempt to alleviate the translation problems, very little has been done on pre-editing. However, pre-editing is equally as important as post-editing. As a matter of fact, pre-editing can eliminate the need for post-editing, reduce effort and save time. Frensh (1991) supported its use and encouraged the students to spot what might be potentially ambiguous in the source text to ameliorate any linguistic problems the machine might encounter. Likewise, Shei (2002) sought to investigate the effect of pre-editing on machine translation’s final product. The study involved a group of Chinese students of

English where they should correct a Chinese input, an English input and their English compositions. The results were very interesting. Not only did the pre-editing improve the machine translation, but also it was found that the pre-editing is beneficial for the students. In fact, it encouraged them to think critically, to learn syntactic structures and to be able to spot the machine translation flaws. These advantages, however, are still bleary and still a gap in knowledge in the field of machine translation.

On the basis of what has been discussed above, this study is an attempt to further broaden this knowledge and explore the effect of pre-editing and its relationship to machine translation.

## **2. Statement of the Problem**

Although machine translation, a newly introduced concept to the field of teaching English as a foreign language in higher education, has an important role on EFL students, it still encounters a number of issues which should be highlighted as an attempt to benefit from the use of translation systems. Hence, it is of crucial importance to identify the major role which can a human interference, particularly at the pre-translation phase, has on machine translation's final output. The latter, most of the time, does not reach the quality which is targeted by language users in general and translators in particular. Thus, the pre-editing approach is introduced with the aim of improving the quality of electronic translation systems and making the use of machine translation more helpful and beneficial.

Consequently, since there were not enough studies exploring the effectiveness of pre-editing on the process of machine translation, the current study is an attempt to fill this gap through investigating the major issues that the computer may encounter and exploring the crucial role of pre-editing on bettering machine translation's final output.

### **3. Research Questions**

Following are three main research questions that will guide the current experimental study.

- To what extent do third year EFL learners at Mohamed Seddik Ben Yahia University apply pre-editing rules while simplifying their texts?
- To what extent does pre-editing affect machine translation's intelligibility and fidelity?
- Is there a relationship between pre-editing quality and machine translation quality?

### **Hypotheses**

The above stated research questions led to the formulation of the following hypotheses upon which the researchers will attempt to confirm and verify their validity.

- H1: Pre-editing process has a positive effect on machine translations' final output.
- H2: There is a significant relationship between pre-editing quality and machine translation quality.

### **4. Significance of the Study**

Translation as an independent field plays a significantly important role in social, cultural and academic contexts, mainly in today's globalised world where communication is considered as the most crucial aspect especially among cultures of different languages. This urged the researchers to find more sophisticated techniques for quick and cheap translation. Hence, a new type called "machine translation" has evolved to meet this need. However, even though the technology is well advanced, machine translation was sometimes deemed inaccurate compared to human translation. Thus, the current study is designed to investigate whether pre-editing the source text can improve the output of machine translation. Additionally, this study aims to further broaden the current knowledge of pre-editing in



relation to machine translation and urge EFL teachers to fully introduce and teach their students the concept of pre-editing for better use of machine translation.

## **5. Research Methodology**

In order to address the previously stated research questions, this study will adopt an experimental design entitled “the pre-experimental design: One-shot case study” type. At the same time, as an attempt to fulfill the research aim, a mixed method approach will be followed and data will be gathered and interpreted both quantitatively and qualitatively.

Firstly, subjects of third year Mohamed Seddik Ben Yahia University, Jijel, will be asked to follow a set of rules in order to pre-edit the source text. Then, they will insert their pre-edited versions in three different translation systems and provide the researchers with the translations. Subsequently, the data gathered after analyzing pre-edited versions along with their translations will be compared to the raw output produced by the machine with the aim of providing a conclusion about the effectiveness of pre-editing on machine translation. Additionally, by means of Statistical Package of Social Sciences (SPSS), the correlation between the scores of both pre-edited texts and their translations will summarize the nature of relationship between the use of pre-editing and machine translation’s final output.

## **6. Organization of the Study**

The current study is composed of two chapters, a theoretical chapter and a practical chapter, in addition to a general introduction and conclusion. The first chapter, which is the theoretical part, will be divided into two sections. The first section provides a general overview about machine translation concept, whilst the second section attempts to give a clear explanation about the pre-editing approach. Then, the second chapter, which is the practical part, will also be divided into two sections. The first section explains the methodology followed in this research, whereas the second one summarizes the data analysis and later discusses and interprets data gathered in this investigation.

## **Chapter One: Machine Translation and Pre-editing**

### **Section One: Machine Translation**

#### **Introduction**

The first chapter of the literature review is exclusively concerned with reviewing the major theoretical aspects related to machine translation and pre-editing approach. Being divided into two sections, the first section deals with some noteworthy notions related to translation, precisely, machine translation (MT). In an attempt to study this phenomenon, it is undeniably deemed essential to first expound this concept. Therefore, this section begins with some definitions of MT brought by some leading figures. Followed by a summary of its historical development. Then, it tackles the distinguished approaches of MT. Afterwards, it sheds light on its purpose, as well as its issues. Subsequently, this section draws some attention to the different types of MT evaluation, namely, human and automatic evaluation. Next, it delves into two main subcategories of MT: machine-aided MT and human-aided MT. Finally, it casts light on the future assumptions regarding MT, namely, the newly proposed approach: the neural approach.

#### **1.1.1. Definition of Machine Translation**

There have been many attempts to define MT. Yet, none of the suggested definitions appears to be fully adequate. This is, perhaps, due to the long history MT has and to the variety of views about it. In the paragraphs below, some detailed definitions attributed to the concept of MT.

According to the European Association of Machine Translation (as cited in Quah, 2006), “MT is the application of computers to the task of translating texts from one natural language into another” (pp.8-9). Similarly, Abdulsalami and Akinsarya (2017) supported this definition and asserted that MT is a computerized system that examines the source language

text and provides a parallel and corresponding text in the target language “ideally without human intervention”.

The limitation of the above two definitions is that neither of them includes any form of human involvement. Researchers such as Hutchins and Somers (1992) put forward that MT is the use of computer systems and software to translate from a source language into a target language, with or without human interference.

However, Archer along with other scholars (as cited in Quah, 2006) disagreed on the definition of MT with respect to the human involvement. Therefore, Hutchins (2002) explained that when the field of MT first emerged, the ultimate goal was to provide a high quality translation for all documents. Nevertheless, that goal turned out to be unattainable. Hence, human revision was required if the translation was to be published. At the same time, it was being recognized that raw MT output could be useful for information gathering purposes and assimilation. Furthermore, Hutchins highlighted the fact that MT output does not always have to be of publishable quality; speed and accessibility may be important as well. From the beginnings, unedited MT output was found to be useful for technical reports, administrative memoranda or any other documents that were to be used by one or two people. Ultimately, with the increasing use of computers and the internet since the 1990s, the use of the crude MT output has remarkably expanded.

### **1.1.2. Historical Development of MT**

The notion of MT pre-exists computers with the philosophers Descartes and Leibniz, whom both came up with the idea of representing languages using universal numerical codes in the seventeenth century. However, it was until the 1930s that the first concrete proposals were made. In 1933, two patents were granted in France and Russia. The French-Armenian, George Astroni, was issued his patent in July 1933 for what he called “Mechanical Brain”, a storage device consisting of a bilingual dictionary and a paper tape, which could be used to

translate from one language into another. Shortly after, a prototype machine was exhibited in 1937. Within the same year, in the proposal of the Russian, Petr Smirnov-Troyanski, he pictured a three-stage process of the mechanical translation: firstly, a native speaking editor who was to analyse the input logically; secondly, the machine to perform the translation; finally, another native speaking editor who was to check and assess the output (Hutchins and Somers, 1992).

After the birth of computers, the actual development of MT can be traced to a conversation between Andrew D. Booth, a British crystallographer, and Warren Weaver, the vice president of the Rockefeller Foundation in 1947. More specifically, to a memorandum written by Weaver in 1947. In the memorandum, he raised four main issues; namely, the problem of multiple meaning, the logical basis of language, the application of communication theory and cryptographic techniques, and lastly, the possibility of language universals (Hutchins, 1986).

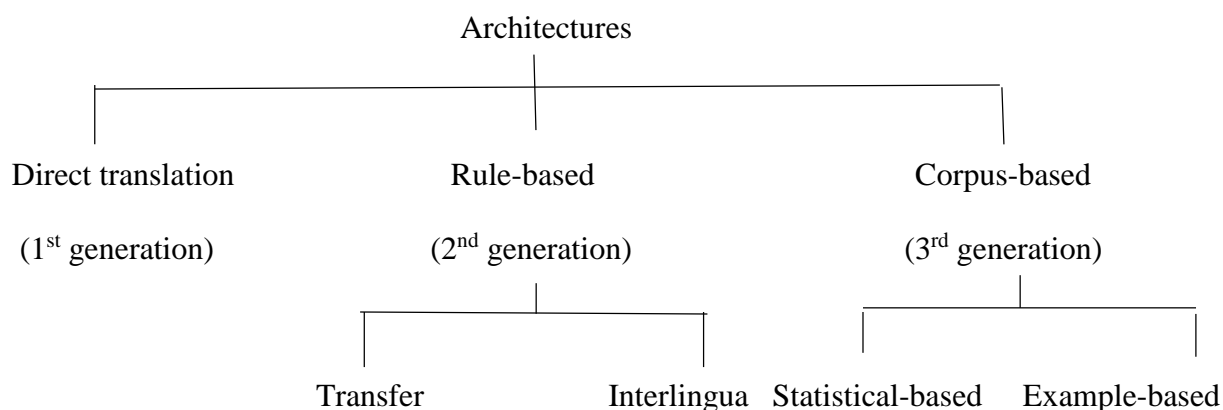
MT took off with the Georgetown-IBM experiment. The first MT system was developed, often referred to as the “first generation”. A bilingual dictionary used for word-for-word translation (Russian-English) in which 250 words and 6 grammar rules were stored in the system and 49 Russian sentences were translated into English. The experiment was a great success and was considered as the starting point of MT era, especially in the USA (Quah, 2006).

Then, followed the infamous ALPAC report and its paralyzing consequences in 1966. A committee known as the Automated Language Processing Advisory Committee was set to evaluate MT research. In its report in 1966, it concluded that MT is slower, less accurate and twice as expensive as the human translators are. Instead, it recommended the development of machine aids for human translators. The ALPAC report brought a virtual end to MT research and funding, especially in the US (Quah, 2006).

It was not until the late 1970s that MT research saw something of a renaissance. The Systran system, replacing the old IBM system, was installed in the US for translating Russian scientific and technical documents. In 1976, the Météo system for translating weather forecasts was installed in Canada. During the 1970s, many other researches in different countries marked the turning point of MT. Notably, in France, with a system for translating Russian Mathematics and Physics texts into French. In Asia, a Hong Kong university developed a Chinese-English MT system called CULT. Then, from the 1970s onwards, MT came to a worldwide availability with the first translation software for personal computers. Subsequently, in 1990 was the appearance of the first translator workstations. Finally, MT has become an online service on the internet in the last five years (Hutchins, 2005).

### 1.1.3. Approaches to Machine Translation

From the beginning of MT field, many MT systems across the globe have evolved. These systems went through three generations, as described by Quah (2006). Figure 01 gives an overview of the architectures of the various MT systems.



**Figure 1. Machine Translation Architectures**

#### 1.1.3.1. Direct Approach

DBA is known as the first generation of MT systems that needs little morphological and grammatical analysis. The bilingual dictionary is the major component of this approach. In direct translation, source language texts are translated without passing through an

intermediary representation. MT systems that use this approach translate the source language text directly to the desired target language (Hutchins and Somers, 1992). The translation process goes through a number of operations. The system starts by replacing the source language words with the equivalent words in the target language. Then, the words are rearranged and generated in the target language text. Quah (2006) stated that this approach does not have the ability to deal with ambiguities and metaphorical expressions or to translate between unrelated language pairs. However, direct MT approach works well with closely related languages that have similar syntactic structure.

Saini and Sahula (2015) summarized the advantages and disadvantages of the direct approach as follows:

- **Advantages of direct MT systems.**

1. The translation is usually easily understood by the reader.
2. The translation is highly effective for languages with similar syntactic structures.
3. Direct systems can be easily carried out.

- **Disadvantages of direct MT systems.**

1. Direct MT involves only morphological analysis without much relationship of the meaning between the words.
2. Direct MT systems can be used only to translate between specific and closely related languages and cannot be adopted for different language pairs.
3. For multilingual translations, direct MT systems can be quite expensive.
4. The source text meaning can be lost in the translation.

### **1.1.3.2. Rule-based Approaches**

Due to the failure and limited quality of the first generation MT system, more accurate linguistic models for translation were developed. As second-generation systems, rule-based approaches were built to make up for the shortcomings of the direct MT approach. RBA require extensive use of linguistic and grammatical rules along with many bilingual dictionaries for language pairs. As noted before, RBA can further be categorized into interlingua and transfer-based translation.

#### **1.1.3.2.1. Interlingua Approach**

In this approach, the source language is converted and translated into an intermediary abstract representation called “the interlingua” which is independent of any languages. Quah (2006) further explained that the abstract representation should include all the necessary syntactic and semantic information for the generation of the target language.

In interlingua systems, the translation from the source language into the target language goes through two independent stages. In the first stage, the source input is analyzed and converted into interlingua representations. In the second stage, the representations are the main sources for generating and synthesizing the target language output (Hutchins, 1986).

The advantage of this approach is that when adding a new language to the system, only an analysis grammar module and a generation grammar module are added (Hutchins and Somers, 1992). However, the main drawback of this approach is defining a universal abstract representation, which preserves the meaning of the text in the original language. Additionally, it is difficult to choose an appropriate neutral representation that can accommodate the possible aspects of syntax and semantic for all languages (Quah, 2006).

#### **1.1.3.2.2. Transfer Approach**

Unlike interlingua approach, transfer approach follows three stages. The first stage is the analysis stage in which the source language text is converted into an abstract source

language representation. The second stage is the transfer of the source language representation into its parallel target language representation. The last stage is the generation of the target text. In this approach, three dictionaries are needed in each stage: a source language dictionary in the analysis stage, a bilingual dictionary in the transfer stage and a target language dictionary in the generation stage (Quah, 2006).

According to Kit, Pan and Webster (2002), “the transfer approach incorporates language analysis and representation at various linguistic levels, but cannot find adequate knowledge to resolve ambiguities involved in the analysis, transfer, and generation” (p. 57). Equally, Quah (2006) further explained this and put forward that the disadvantage of the transfer approach is that it relies on dictionaries, which may not contain enough knowledge to deal with ambiguities included in the three stages.

### **1.1.3.3. Corpus-based Approaches**

In an attempt to overcome the shortcomings of the aforementioned MT systems, two new approaches emerged. Carl and Way (2003) claimed that “CBMT systems assume the existence of a bilingual parallel corpus (usually database of translated sentences) which used and/or consulted the required knowledge for new translations” (p. xvii). Corpus-based approach is furthermore classified into two sub approaches: Example-based MT and statistical MT.

#### **1.1.3.3.1. Example-based Approach**

This approach was first introduced by Nagao in the mid-1980s. According to Sato and Nagao (1990) (as cited in Quah 2006), “the basic idea of an example-based translation is to translate a source sentence by imitating the translation of a similar sentence in the database” (p. 81). In other words, EBA is a memory-based translation that reviews and discovers closely resembling translations of the language pair. In EBA or in the translation by analogy, a bilingual corpus and an algorithm are the crucial components. The algorithm searches for the



closest example to match the source language segment with its target language segment (Arnold, Balkan, Meijer, Humphreys and Sadler, 1994, p. 188).

Kit, Pan and Webster (2002) summarized three main phases involved in the translation process of an example-based system:

- **Matching**

The translation input is segmented and the segments are matched against existing examples in the bilingual corpus.

- **Alignment**

The corresponding fragments are aligned by means of an alignment algorithm.

- **Recombination**

The fragments are then combined to generate the target language text.

#### **1.1.3.3.2. Statistical MT Approach**

The idea of SMT was devised by IBM research group after the success of statistical methods in speech recognition in the late 1980s. MT task was treated as a mathematical and statistical problem using Bayes' theorem in the Candide project (Koehn, 2010). This approach has two essential stages. First, individual words and phrases are aligned in the parallel bilingual corpus. Then, the corpus calculates the probabilities of correspondence of words to other words that are aligned in the other language (Somers, 2008). According to Carl and Way (2003), SMT systems use a translation model from a bilingual parallel corpus and a language model from a monolingual corpus. At the end, the best translation is chosen by maximizing the probabilities of the models.

#### **1.1.3.4 Hybrid MT approach**

When single approaches failed to give a good output quality, hybrid approach was developed. As the name implies, HMTA is the combination of more than one MT approaches, mainly RBA and SMT.

Costa-jussà et al., (2006) claimed that HMTA could be used in different ways. In some cases, translation is first performed using RBA followed by introducing statistics to adjust the output or to enrich resources. In other ways, rules are used to preprocess the input as well as to post process the output of an SMT system.

In short, HMTA emerged to take advantage of the strengths of SMT and RBMT approaches to reach a satisfactory level of accuracy.

#### **1.1.4. Purpose of Using Machine Translation**

Hutchins (2005) summarized four main purposes for using MT:

##### **1. Dissemination**

Translation for dissemination purposes need to be of a high quality in order to be published. Since MT outputs are often regarded as raw outputs, human assistance is necessary.

##### **2. Assimilation**

Translation for this purpose needs not to be of a publishable quality. Unedited output could be sufficient if the text is understood and the user can get the main idea of the content.

##### **3. Interchange**

When individuals who speak different languages communicate, by e-mails, telephones...etc., the translation quality is not an issue as long as it conveys the content of the message and people understand each other.

##### **4. Database**

In the modern world, people use MT to extract information from a database in a foreign language, for searching the internet, and for accessing web pages.

Moreover, Hutchins (2005) pointed out that MT could be very necessary in other cases. First, MT is of great assistance to the human when there is too much to be translated. Second, human translators find technical materials boring, so they seek help from computers

and software. In addition, consistency is a major factor for using MT. Large companies often want terms to be translated in the same way every time. Human translations may change depending on a number of factors, whereas a machine's translation will not since it allows the systems to memorize the terminology. Finally, speed is a key benefit of MT. MT is significantly faster than human translators are. All in all, any one of these factors on its own can be a sufficient reason for using computers and MT systems.

### **1.1.5. Issues of Machine Translation**

Although translations produced by machine translation systems are useful and play a major role in facilitating people's work in many fields, the machine still encounters some problems, which result in having unreliable and inadequate translations. As a result, researchers worked hard on defining those issues as an attempt to find solutions to improve the machine's work.

Garvin (1956) claimed that problems of machine translation could be divided into two categories: problems of word selection, which have to do with finding suitable equivalents in the TL for units in the SL; and problems of word arrangement, which deals with the divergences of syntactic systems throughout languages.

First, in word selection, the machine faces no problem when a word of the SL has only one equivalent in the TL. However, when a word has more than one equivalent in the TL, the machine finds it difficult to decide upon which word is more appropriate. This problem mainly occurs because of machine's disability of understanding the context in which that word takes place (lack of contextual and cultural knowledge). For example, the word "immunity" in the English language may have two different meanings in the Arabic language. This depends on whether it occurs in a medical context or a military one.

Second, different languages follow different syntactic rules. Thus, the machine will encounter problems, in order to provide an appropriate final translation, when the word order

in the SL is very different from that of the TL. To illustrate, in the English language, a sentence can never start with a verb; something which is totally correct in the Arabic language.

Hutchins and Somers (1992) provided a more detailed classification of various problems that encounter machine translation. According to them, the problems are caused either by lexical ambiguity or by structural ambiguity.

In lexical ambiguity, they identified two kinds of problems: The first one refers to words which are spelt in the same way and have different meanings, such as the word “pen”, and which could be determined by the context. The latter, in fact, is quite difficult to be understood by the computer. The second one deals with lexical problems that are caused by polysemous words, i.e. words that are spelt the same and have closely related meanings. Hutchins and Somers gave many examples to clarify this point. For instance, the pronoun “I” in English refers to both males and females. In Spanish, this pronoun would create a problem for the machine because they have different pronouns, which are followed by different verb conjugations (different endings). Another example is the verb “to know” in English, in French, this word has two slightly different meanings “savoir” a fact and “connaitre” a thing. Hence, the machine will find it difficult to choose the more suitable word for the given context.

In structural ambiguity, on the other hand, it is mentioned above that Garvin claimed that differences in word order of sentences among languages lead to producing inappropriate and even wrong translation. In Hutchins and Somers’ work, they identified other syntactic problems. The use of some words, such as prepositions or linking items, in long complex sentences results in structural ambiguity. For example, “the police refused the students a permit because they feared violence”, or “the police refused the students a permit because they advocated violence”. In the first example the pronoun “they” refers to the police; in the

second example, changing the verb “fear” with “advocate” switched the meaning of the pronoun “they” from police into students. Another example is “I saw the girl with a telescope”. Here, the use of markers with more than one relation creates a problem. It is not clear whether the prepositional phrase “with a telescope” is attached to the phrase “I saw” or the combination “the girl”. Thus, the syntactic structure affects meaning since these interpretations are, in fact, difficult for the machine.

Arnold et al. (1994) added another point related to syntactic problems. According to them, a word, which may belong to different grammatical classification or different parts of speech (noun, verb, adjective...), creates another kind of problems. The word “use” for example in the sentence, “you must not use abrasive cleaners on the printer” is a verb and its translation in the French language would be “se servir or employer”. Whereas, in the example “the use of abrasive cleaners on the printer casing is not recommended”, the word “use” serves as a noun and its French translation would be “emploi or utilisation”. In this case, it would be a hard and confusing job for the machine to deal with such issue.

However, knowledge about translation procedures and strategies may give an idea for solving potential translation problems as well as improving the machine translation’s quality. It may focus the translator’s attention on the relevant factors that should be taken into consideration before getting through the process of translation. Moreover, the problems are mostly encountered when translating literary texts. In technical, economic and commercial texts, the translation is most of the time successful, especially when the source and target language are close in their morphological, lexical, and syntactic structures. Thus, it would be easy for the translator to adjust the final product and have a coherent text.

### **1.1.6. Evaluation of Machine Translation**

Through time, the concept of evaluating machine translation systems has been a point to discuss. All scholars agreed upon the necessity of evaluating machine translation’s output

in order to decide upon the aspects which need further adjustment and improvement. As a result, two types of evaluation were identified: human evaluation and automatic evaluation. Each of the latter concepts takes into consideration a particular set of characteristics as an attempt to determine the effectiveness of the machine's work.

Firstly, in human evaluation, the scholars Miller, Gates, Magdalen, and Underwood (2001) stated that comprehensibility, readability, fidelity, and coverage represent the main points that a human evaluator should take into account.

- **Comprehensibility**

(Or fluency) refers to whether the output text is understandable. Here, users are going to see whether when reading the translation sentence by sentence, the text is understood; and even if there is a possibility of having some missing items, still the reader would be able to understand the delivered message.

- **Readability**

In contrast to comprehensibility, it refers to the case in which the output text could be read. The two concepts are different because a text can be comprehensible after several close readings even though it is difficult to be read. Whereas, a more easily readable text would save the user's time since he does not need to reread the text many times in order to understand it.

- **Fidelity**

It is the extent to which the information is successfully delivered from the ST into the TL. In other words, the ideas and information of the ST are fully restated in the TL without any changes, neither in meaning nor in the quantity of the information.

- **Coverage**

It refers to the percentage of translated terms, i.e. the successfulness of translation is determined through counting how many terms, of the particular domain of the inserted text, were able to be translated.

However, Howy and Belis (2002) maintained that there are two main characteristics: fluency and adequacy. According to them, fluency deals with the grammatical correctness of the output. The text then will be ranked according to a scale composed of five main degrees: flawless language, good language, non-native language, disfluent language, and incomprehensible language. On the other hand, adequacy is used to evaluate the quantity of information exists in the original text that a translation contains. This type also follows a scale of five degrees, which are: all, most, much, little, and none. In line with this, Koehn and Mouz (2006) assured that the distinction between fluency and adequacy was made because a translation might be disfluent, yet contains all the information from the original text. However, unfortunately, no instructions were given to evaluators in terms of how to quantify meaning or how to count the grammatical errors of the output.

Second, automatic evaluation is the evaluation done by automatic systems as an attempt to define to what extent machine translation's work is correct and appropriate. Some of the famous systems are BLEU, NIST, F-measure, Meteo. In this respect, Nyberg, Mitamura and Carbonell (1994) identified three major criteria that automatic evaluation adopts: completeness, correctness, and stylistics.

- **Completeness**

A system is complete if for each input string there is a corresponding one in the output. Completeness is divided into three types: **lexical** (if the system has source and target language lexicon for every word or phrase in the translation domain), **grammatical** (if the system is able to analyze grammatical structures of the SL and generate them into

grammatical structures in the TL), and **mapping role** (if the system assigns an output structure to every input structure in the translation domain).

- **Correctness**

A system is correct if it assigns a correct output string to every given input string. There are three types: **lexicon** (words in TL should be correctly chosen for the sense in the input), **syntactic** (there should be no grammatical errors) and **semantic** (meaning of the target sentence should be equivalent to that of the source sentence).

- **Stylistics**

This one deals with **syntactic style** (output sentence should not be grammatical but inappropriate for the context), **lexical appropriateness** (words chosen should be appropriate for the context), **usage appropriateness** (the most conventional or natural expression should be chosen) and others such as **formality, level of difficulty of the text**...etc.

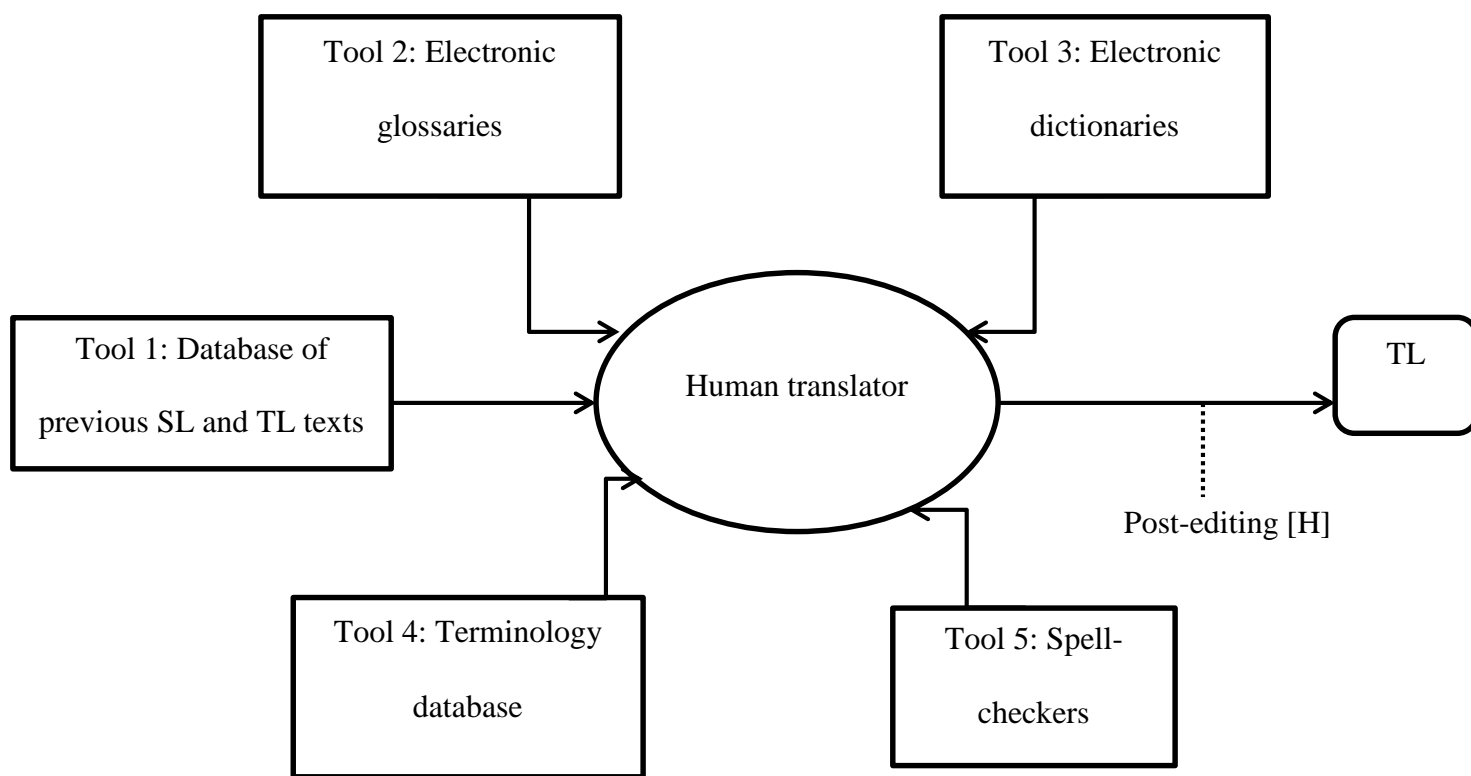
Although scholars defined the characteristics mentioned above for the evaluation process, this latter is not considered as an easy task because of the language's dynamic, ambiguous and complex nature. Thus, it is difficult to estimate whether a translation is correct or how to consider a translation as a correct one (Goutte, 2006).

### **1.1.7. Machine-aided Human Translation**

Machine aided human translation is the translation done by a human translator with the help offered by computer tools. Sager (as cited in Quah, 2006) stated that it is the use of computer software by translators to perform part of the process of translation. Those software or systems are known as "workbenches" or "work-stations" as they combine a number of tools. This type of translation is done by a human translator who uses a number of tools such as spell checkers, electronic glossaries, electronic dictionaries, terminology database, and a collection of previously translated texts and their originals, which refer to translation memory, to support the translation process.



Sager clarified his idea with the following figure



**Figure 2. Machine-aided Human Translation Model**

### 1.1.8. Human –aided Machine Translation

Katberg (2012) identified HAMT as the translation that is done by a machine translation program with human assistance. The program here is designed in a way that requires human’s collaboration. The latter plays the role of an editor and intervenes at three main stages: before, during, and after.

First, when the interference takes place at the beginning, it will be known as “pre-editing”. Here, the human adapts the source text in a way that facilitates the work for the machine and enables the program to decode the text. At this level, the adaptation will include the word order of sentences. The editor will try to re-order the words in sentences of the ST either all in the same way, or in a way similar to the one of the target language. Thus, pre-

editing will lead to avoid word order problems. In addition, the pre-editor could change the language of the ST into a more “controlled language”. He is going to adjust the ST in a way where he excludes expressions which are unconventionalized and which cause stylistic variations.

Second, when the human intervenes in the middle of the translation process, it will be known as “interactive editing”. The human is asked by the program for assistance in solving particular translation problems that a translation program is unable to fulfill. For instance, the program will give suggestions for the editor and he, in return, is supposed to choose the most appropriate and suitable one. This mostly happens with words, which have many equivalents in TL, and the selection depends on the context in which they appear.

Third, when the interference takes place at the end of the process where the machine has already produced its translation, it will be known as “post-editing”. Here, the editor is supposed to correct the machine’s translation and adjust the product in order to provide a more appropriate final output.

### **1.1.9. Neural Machine Translation**

Neural Machine Translation is a recently proposed approach to Machine Translation. It was first introduced by Kalchbrenner and Blunson in 2013 as an attempt to improve the work of machine translation systems and raise the chance of having more adequate and accurate human-like translations, with higher quality, in comparison with the traditional approaches.

According to Koehn, Och and Marcu (2003), NMT is a system that is composed of large artificial neural networks that are trained to read a sentence and provide a correct translation without the need of storing phrase tables and language models that are used in conventional systems. In other words, unlike traditional systems, which contain small sub-components stored separately, the neural system can be trained directly on both source and

target languages. Hermann and Blunson (2014) further explained that it is mainly composed of two major elements: “**encoder and decoder**”. The encoder reads and encodes a source sentence into a fixed-length vector. Then, the decoder outputs the translation from the encoded vector.

Finally, although neural machine translation has managed to provide more appropriate translations, it still encounters a number of problems. Working with fixed-length sentences make it difficult for the machine to deal with long sentences (Cho et al, 2014). Furthermore, it is incapable of translating rare words because it contains a fixed size of vocabulary (Bahdanau, Cho & Benjio 2015) (as cited in Sutskever, Lee, Vinyals & Zaremba ,2015).

## **Conclusion**

Concluding what has been discussed above, this section has dealt with the major points concerning MT. It has firstly defined MT and it spotted light on some key points along its history. Then, different approaches of MT were highlighted. Additionally, it dealt with the purpose and issues of MT. Next, this section set forth the two main ways for evaluating MT. Furthermore, it gave an account of its subcategories, mainly, MAHT and HAMT. Lastly, it ended with the recently introduced approach to MT, the neural approach.

## **Section Two: Pre-editing Approach**

### **Introduction**

After introducing the main aspects pertinent to the field of machine translation in the first section, this section is devoted to shed light on the pre-editing approach as a major phase during machine translation process. First, it provides the definition of pre-editing approach. Next, it elucidates the concept of controlled language and its role on the emergence of pre-editing, as well as the main steps that are followed in order to simplify the original text. Finally, it presents the main advantages that the process may cause on the translation done by electronic systems.

#### **1.2.1. Definition of Pre-editing**

The interference of human translator is vital for the MT systems to provide an appropriate translation output. When this interference takes place before running the MT, it is called “pre-editing”. The latter is significantly important because MT systems are unable to interpret the source text correctly.

Hutchins and Somers (1992) provided a thorough definition of pre-editing and asserted that “(...) pre-editing involves checking source texts for foreseeable problems for the system and trying to eradicate them. It can include the identification of names, the marking of grammatical categories of homographs, indications of embedded clauses, breaking of coordinate structures, flagging or substitution of unknown words,...etc.”(p.157). Likewise, Quah (2006) put forward that “the main task of pre-editing is to discover any elements such odd phrases or idioms or typographical errors that might create problems for the machine translation system during the translation process. The human translator or editor amends the source language text accordingly” (p.11). Simply put, pre-editing is the process of adjusting and regulating the source text until appropriate output is obtained.

## **1.2.2. Controlled Language**

Natural languages are what humans use in daily communication such as English, Chinese... etc. Due to the variations in their informal rules and their subtle differences in culture and meaning, both humans and machines may experience difficulty in understanding or translating them. Controlled natural language's main goal is to address this issue with restricted rules to eliminate any ambiguity or complexity. Ultimately, seeking to reach a readable and comprehensible text or document, which result in an easily translatable input.

Traditionally, a CL can be defined as “an explicitly defined restrictions of a natural language that specifies constraints on lexicon, grammar, and style” (Nyberg, Mitamura and Huijsen, 2003, p. 245). The resulting effect is comprehensibility improvement, style consistency, post-editing and processing easiness as well as the reusability of texts for translation (Douglas and Hurst, 1996).

CLs can be classified into two broad categories: human-oriented controlled languages and machine-oriented controlled languages. In, HOCL, the ultimate goal is making texts and technical documents legible, comprehensible and clear for humans, especially, for non-native speakers. However, the main objective of MOCL is to improve translatability for machines (Reuther, 2003).

### **1.2.2.1. The Emergence of Controlled Language**

The history of controlled language was summarized by Kaji (1999). He claimed that the concept dates back to 1930s in Odgen's Basic English. The latter was proposed as an international language and a foundation for English learning. It consists of 850 basic words along with some inflection and derivation rules. It aims to provide a simplified English version with more restricted lexicon, syntax, and grammar.

From the mid-1960s to 1970s, the first practical controlled language “Caterpillar Fundamental English” was designed by Caterpillar Inc. They provided simplified versions of

documents in order to make them easily understood by non-native speakers through the use of extremely limited vocabulary and grammar.

Then, the European Association of Aerospace Industries introduced Simplified English (SE); the most widely used controlled language in the industrial field. It is a human-oriented controlled language which aims to enhance readability and consistency of aircraft maintenance. The SE guide consists of about 3.100 words and 57 writing rules.

During the 1980s, the CFE helped in the emergence of other controlled languages including Smart's Plain English Program (PEP), White's International Language of Services and Maintenance (ILSAM), and Perkin's Approved Clear English (PACE). The latter has a major role in the field of machine translation. It successfully applied controlled language on the source text and made the post-editing work three times faster than usual.

In recent years, the development of controlled language has been carried out by a number of organizations. Caterpillar Inc. replaced CFE with "Caterpillar Technical English" in relation with KANT, a machine translation system developed by Carnegie Mellon University. They extended the vocabulary from 1.000 words to 70.000 words. There are other controlled languages such as Scania Swedish for truck maintenance documents at Scania, a Swedish truck manufacturer, Controlled English at Alcatel Telecom of Belgium, and Easy English Language at IBM Corporation.

Finally, unlike the previously existing CL, which was specified for technical domains, new approaches are recently adopted for other domains where descriptive texts can be used. However, studies have shown that the application of controlled language on descriptive texts, which convey various types of information, is harder than procedural and technical ones.

### **1.2.3. Pre-editing Approach**

In many cases, machine translation systems are unable to interpret the source text correctly, which lead them to provide inappropriate and even wrong translations. Thus, this

results in facing huge problems during the post-editing stage of translation. As a result, researchers suggest adapting the pre-editing approach; in other words, they require interpreting the source text at the very beginning as an attempt to eliminate the difficulties a machine system may encounter such as: complex grammatical structures, ambiguous words, semantic problems, misspellings...etc.

In order to solve the problems mentioned above, researchers managed to introduce a number of methods and strategies that the editor needs to follow. They claimed that the source text should be rewritten in a more controlled language; i.e. a more controlled vocabulary and controlled grammar to limit both lexical and syntactic ambiguities and to help the machine to be able to decide upon the suitable translation. In line with this, Mitamura (1999) believed that there are some lexical problems, which need to be fixed before the machine starts the work. She acknowledged that the use of functional words, such as determiners and conjunctions, should be discouraged since they increase syntactic ambiguity; use of practical forms (-ing and -ed forms) should be restricted when used after conjunctions (while driving the vehicle...) or in reduced relative clauses (directional stability caused by wheel lock-up...); both examples, she believed, should be rewritten as follows: while you are driving the vehicle... and the directional stability that is caused by the wheel lock-up. In addition, whenever a word has many synonyms in a particular domain, the editor should use the most frequently used word that has a clear meaning to make it easier for the machine to give the suitable translation. Moreover, Mitamura focused on the importance of orthography checking. The misplacement, addition, or omission of one letter in a word leads to having another word with another possible meaning. For example, “copy the files from the accounts directory to my directory”. In this case, it is obvious for the human reader that the word “files” is the intended word and it was just misspelt; something, which is difficult for the machine to deal with since it does not have a contextual knowledge. Furthermore, the word “flies” is, in fact, a

meaningful word. As for grammatical constraints, Mitamura argued that there are two types of constraints: phrase level constraints and sentence level constraints. In the former type, she included replacing phrasal verbs with single words. For instance, “turn on” should be rewritten as “start”, and repeating prepositions in conjoined constructions to avoid ambiguous readings. For example, the phrase “recorded memory of radio and each control unit” is ambiguous since it can be divided as (recorded memory of radio) and (each control unit), or recorded memory (of radio and each control unit). Therefore, it needs to be rewritten this way: recorded memory of the radio and of each control unit. In sentence level constraints, Mitamura included the coordination of sentences (the two parts of a conjoined sentence should be of the same type), relative clauses should always be introduced by a relative pronoun and elliptical constructions should be avoided altogether.

Wojcik (1998) provided a list of writing rules. These rules specify that:

- a) Determiners such as the, a, an, this, that...etc. should be used when appropriate.
- b) Passive sentences should be avoided in descriptive writing.
- c) Sentences length should be limited to 25 words maximum.
- d) Noun clusters with more than three words should be avoided.
- e) There should not be more than one word modifying a noun or a noun cluster.
- f) –ing forms should be carefully avoided.
- g) The word “that” should not be omitted after verbs.

There are also some explicit rules introduced by MT developers for writing content that needs to be translated using MT. For instance, IBM developers recommended the use of simple short sentences, avoidance of idiomatic and slang expressions, avoidance of ambiguous words (words which have different meanings depending on the context in which



they appear), repeating nouns or phrases instead of using pronouns when possible, use of proper punctuation, and use of complete sentences.

#### **1.2.4. Advantages of Pre-editing on Machine Translation**

Pre-editing the source text during the process of machine translation has a number of advantages on the final output provided by the computer. According to Baker, Franz, Jordan, Nyberg and Mitamura (1994), it improves the translated text in terms of readability, comprehensibility, and reusability. They believed that reducing the use of homonymy, synonymy, and ambiguous structures enhances the readability and intelligibility of the text. In addition, when a text is written in a standard terminology, it is easier to reuse the text elsewhere and thus provide a translation once again. Another advantage is discussed by Mitamura (1999). She claimed that the use of pre-editing process to simplify the source text is helpful and beneficial for both computers as well as humans. For computers, encouraging clear and direct writing along with reducing complexity and ambiguity leads to improve the quality of the machine's translation and speeds up its work. Therefore, this will help the human to easily post-edit the translation output by making slight corrections, or it will totally eliminate the post-editing step altogether. Finally, Nyberg et al., (2003) shed light on the benefit of reducing the risks of misunderstanding. Through using more controlled vocabulary, the editor will decrease semantic ambiguity and will limit the chance of delivering a wrong message that is not meant by the original text.

#### **Conclusion**

This section has been devoted to give an overview about key issues related to pre-editing approach. It first introduced the notion of pre-editing. Then, it clarified the concept of controlled language and its role on the appearance of pre-editing. Additionally, this section explained the major rules that should be followed during the process of simplifying the

original text. Finally, it shed light on the advantages of pre-editing on machine translation process.

## **Chapter Two: Research Methodology and Data Analysis and Interpretation**

### **Section one: Research Methodology**

#### **Introduction**

While the previous chapter cast light on the theoretical aspects of machine translation field and provided a comprehensive review about the pre-editing approach, the present chapter is concerned with the practical framework which aims to investigate the effectiveness of pre-editing on machine translation's final output. It is basically composed of two sections: the research methodology and the data analysis and discussion. The research methodology section includes the research paradigm, the setting, the population and sampling, the research procedures involved in data collection, the methods followed in data analysis, along with research limitations. The second section, on the other hand, is divided into two parts. The first part displays the main results obtained from the implemented experiment; the second part summarizes and discusses the major findings of the current investigation.

#### **2.1. Research Methodology**

The research methodology section is devoted to such issues as the research paradigm, setting, population as well as sampling. It also describes the research procedures involved in data collection and the methods followed in data analysis and lastly the limitations of the study.

##### **2.1.1. Research Paradigm**

The current research study aims to investigate the effectiveness of pre-editing the source text on machine translation's final output. To probe into the topic, the researchers opted for an experimental design as it is viewed to be the best method to identify a cause-effect relationship (Nunan 1992). Grotjham (as cited in Nunan and Baily, 2009) claimed that experimental design includes a collection of research methods with principles and procedures for determining statistical significance as well as criteria for determining the quality of the

study. Therefore, since there are various types of experimental designs, the criteria of this study led the researchers to follow a subtype known as the pre-experimental design: the One-Shot case study. It is a design in which the researcher does not administer a pre-test; instead, he depends on the post-test analysis after the treatment is implemented. It also does not require a control group because it tests the effect of the treatment on the experimental group in comparison with the constant variable. In this investigation, the participants receive a treatment about the use of pre-editing rules. Then, the translation provided after the text is simplified will be compared to the one produced by the machine without human interference as an attempt to answer the research problems posed.

For data collection and analysis, the researchers adopted both quantitative and qualitative approaches. The former, according to Aliaga and Gunderson (2002) (as cited in Apuke, 2017), explains phenomena by collecting numerical data that are analyzed using mathematically based methods (statistics in particular), whereas the latter, as Creswell (2009) claimed, uses non-numerical data and describes the results in words rather than numbers. Accordingly, the reason behind the use of a mixed approach is that the researchers aim not only to explore the effectiveness of pre-editing approach on machine translation output (qualitatively), but also to identify how many rules the participants managed to apply (quantitatively). Simultaneously, the researchers opted for the Statistical Package of Social Sciences (SPSS) to establish the relationship between the pre-editing quality and translation quality.

### **2.1.2. Setting**

This research study was carried out at the English Department of Mohamed Seddik Ben Yahia University in Jijel, precisely at language laboratories, where students were able to use computers with internet to provide the researchers with the appropriate data for the investigation.

### **2.1.3. Population and Sampling**

The participants of this study were a total of twenty-one (21) third year EFL students (nineteen (19) females and two (2) males). Since the aim of generalizability is not intended in this research, the researchers adopted convenient sampling. A third year group was chosen based on their accessibility, availability, and willingness to participate in the study.

The rationale behind working with third year students is that they have translation as a module for their second time. Thus, they are familiar with the different aspects related to the field of translation as well as the problems which are mostly encountered. Furthermore, with the technological development and the great dependence on computers in different fields of learning, it is important, at this level, to raise students' awareness about the methods and approaches which help them to exploit the use of computer systems which are devoted for language learning in general and translation in particular. As a result, the researchers see that third year students have enough linguistic knowledge to apply a set of pre-editing rules on the source text and provide a simplified version as an attempt to facilitate the translation process using electronic systems.

### **2.1.4. Data Collection Procedures**

The experimental design previously described has been used to gather data along two main sequential steps: pre-implementation phase and implementation phase.

#### **2.1.4.1. The Pre-implementation Phase**

As a way of start and striving to find convincing answers to the stated research questions and at the same time test the earlier stated hypotheses, the researchers collected a set of pre-editing rules, which were agreed upon by many researchers, as the most widely used rules to prepare any text for the machine translation. (See Appendix A)

As far as the text is concerned and since it was almost impossible to find a decent one that works with all the selected rules, the researchers opted for a text adopted from the text

book “Between English and Arabic: A practical course in translation” by Bahaa-eddin Abulhassan Hassan (2014). The text is written in general English; hence, suitable for the participants’ level. More importantly, it needs to be interpreted using seven of the selected rules. (See **Appendix B**)

#### **2.1.4.2. The Implementation Phase**

Practically, this phase went through two basic steps. Prior to the first phase, the participants were informed of the purpose of this study. Additionally, the students were ensured that their identities and answers would be kept confidential. This is in line with research ethical principles, to ensure that the participants’ rights are not violated.

In the first phase, the students were taught the concept of the pre-editing approach in relation to machine translation; where they learnt about the major issues a machine system might encounter when interpreting the source text and how to apply the pre-editing rules to eliminate those issues and facilitate the task for the machine. Remarkably, due to constraints beyond the researchers’ control, the treatment was conducted in just one session that lasted approximately one hour. Eventually, provided with a printed version of the rules along with examples for further explanation, the participants were given the text and asked to rewrite it in a simplified version using the taught pre-editing rules.

In the second phase, the students were called to the language laboratory and asked to translate their pre-edited versions of the text using three different translation systems; namely, Google Translate, Microsoft Translate and Systran. The aim behind working on three translation systems is to ensure reliability and to prove that pre-editing rules are effective on more than one system. Essentially, prior to this step, the researchers checked the equipment of the laboratory to avoid any unanticipated obstacles. Afterwards, the students were randomly divided into three groups of seven. Each group was required to perform the translation using one of the aforementioned systems.

### **2.1.5. The Nature of Data Analysis**

After the application of the pre-editing rules by the participants and the translation process was fulfilled, the researchers reached an important phase. That is, data analysis.

To answer the first research question, which aims at investigating the participants' use of the rules, the researchers counted the frequency of occurrence and percentages of the pre-editing rules in the subjects' simplified versions of the text and presented them in a table.

As far as the analysis of the translation outputs is concerned, the researchers first decided to rely on the criteria featured in the infamous ALPAC report, namely, intelligibility and fidelity. As it was mentioned in the previous chapter, the ALPAC report was published in 1966 by a committee that was set to evaluate MT quality. The committee designed an experiment to lay the foundation of a standard procedure for measuring the quality of scientific translation, be it human or mechanical. It concluded that the major characteristics of translation are its intelligibility and its fidelity to the original text (ALPAC, 1966). Before striving to better evaluate and analyze the final output, it is of considerable necessity to first provide a definition of the aforementioned criteria.

#### **1. Intelligibility**

Intelligibility essentially refers to the "subjective evaluation of the degree of comprehensibility and clarity of the translation" (Van Slype, 1979, p.62). To investigate intelligibility, ALPAC report issued a nine-point scale, where the higher the score, the more intelligible the translation. However, the more the number of scales, the more will be errors in human evaluation. Hence, it was modified to a four-point scale by Van Slype, which is widely used. Thus, in the current study, the researchers decided to follow the four-point scale defined by Van Slype.

## **Scale of intelligibility.**

### **3: Very intelligible**

All the content of the message is comprehensible, even if there are errors of style and/or of spelling, and if certain words are missing or badly translated, but close to the target language.

### **2: fairly intelligible**

The major parts of the message passes.

### **1: Barely intelligible**

A part of the content is understandable, representing 50% of the message.

### **0: Unintelligible**

Nothing or almost nothing of the message is comprehensible (Van Slype, 1979, p. 71).

## **2. Fidelity**

Fidelity refers to the extent to which information is successfully retained in the translated text compared to the source text. In line with this, Halliday and Briss (1977) defined fidelity as “measurement of correctness of the information transferred from the source language to the target language” (p. 72). Similarly, Van Slype (1979) maintained that fidelity is a “subjective evaluation of the measure in which information contained in the sentence of the original text reappears without distortion in the translation” (p. 72). To measure fidelity, the researchers opted for the four-point scale introduced by Van Slype.

## **Scale of fidelity**

### **3: Completely or almost completely faithful**

### **2: Fairly faithful**

More than 50% of the original information passes in the translation.

### **1: barely faithful**

Less than 50% of the original information passes in the translation.

### **0: Completely or almost completely unfaithful** (Van Slype, 1979, p. 78).



In order to ensure reliability, six teachers from Mohamed Seddik Ben yahia University, Jijel were asked to evaluate the translations. To measure intelligibility, three monolinguals, Arabic native speakers, were selected to judge twenty-four texts (twenty-one translations of pre-edited texts and three raw translations). The evaluators were provided with the Arabic translation of the intelligibility scale, which is as follows:

<b>Grade 3</b>	مفهوم جدا	
<b>Grade 2</b>	مفهوم الى حد ما	
<b>Grade 1</b>	بالكاد مفهوم	
<b>Grade 0</b>	غير مفهوم تماما	

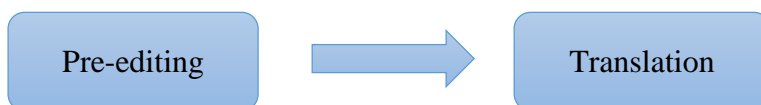
To measure fidelity, three bilinguals (Arabic-English) were asked to assess the informativeness of the translations. The bilingual evaluators were provided with the original text along with the subjects' pre-edited texts. Afterwards, the researchers assigned scores for the texts and reported the findings in tables. The latter were accompanied by commentaries to cast light on important aspects.

After the evaluation of the fidelity and intelligibility of the translations, and in order to further reach rigorous results, the researchers evaluated the correctness of the pre-edited texts in terms of their grammar and sentence structure in accordance with each one's translation. After having evaluated the pre-editing and the translations, the researchers assigned scores for both of the variables based on a five-point scale, which is as follows:

Very good	Good	Average	Bad	Very bad
5	4	3	2	1

The purpose of this step is to determine whether the quality of pre-editing affects the quality of translation. For this, the researchers relied on the Statistical Package of Social Sciences (SPSS) to calculate the Pearson Correlation for it is considered as the most commonly used measure to assess the association between two variables. In line with this, Cramer (2003) has put forward that “Pearson’s  $r$  is the measure of the linear relationship between two interval or ratio variables, and can have a value between -1 and 1” (p. 2). Afterwards, after having tested the relationship between the variables, it is deemed necessary to find out how likely it is that the findings of the research are due to chance. Simply put, it is crucial to determine whether the results are statistically significant. The significance of the result is measured by a probability coefficient ( $p$ ). A result is said to be statistically significant if the  $p$ -value is lower than the significant level of 0.05 or 0.01 ( $*p < 0.05$  or  $**p < 0.01$ ).

This step is further illustrated in the study model below.



### 2.1.6. The Limitations of the Study

Although this research has been carried out with a solid research methodology and it has reached its aim, there were a number of constraints that were beyond the researchers' control.

- One of the major limitations of this study is the lack of relevant and reliable references. Significantly, concerning the second section in the theoretical part, which lacks both resources and prior research studies on the topic.
- Another limitation that is worth of noting is time constraint. Since the approach of pre-editing was not previously introduced to the students, the researchers in fact attempted to teach the participants this approach thoroughly. The students will then practice pre-

editing as a separate step of this study before even running the machine translation. However, it was not possible due to time restrictions.

- Another limitation which was also due to time constraints is the incapability of adding another research tool, which at that time seemed to be necessary for the reliability of this research. Respectively, the researchers intended to conduct a focus group with the participants regarding their attitudes about pre-editing and the difficulties they faced when working on their texts.
- Lastly, it is relevant to point out the technical problems occurred during the data collection procedures. As it was stated before, this research was carried out in the language laboratory. It is fairly to mention that the laboratory hold about twenty computers. However, the participants were able to use only four, mainly, due to the unavailability of the remaining computers and to the fact that half of them were not working, which took a long time to finish the practical part since only four students could use the computers at a time.

## **Conclusion**

All in all, the first section explained briefly the methodology followed in this research. It described the research paradigm, the setting, the population and sampling as well as the methods adopted for data collection, analysis, and interpretation. Finally, this section shed light on a number of limitations of the current research.

## **Section Two: Data Analysis and Interpretation**

### **Part One: Data Analysis**

#### **Introduction**

The first section of the fieldwork chapter was devoted to describe the research design and methodology used for data collection. In the second section, on the other hand, the data gathered from the experiment are carefully analysed and interpreted. The purpose of this section is to answer the questions and to verify the hypotheses which, as previously mentioned, were meant to highlight the efficiency of pre-editing on machine translation's final product.

The first part of the second section is further divided into three subsequent phases; the first stage is mainly concerned with the analysis of the participants' pre-edited versions of the text, the second stage aims at analyzing the translations which, as stated in advance, is based on two main criteria. The last step however, is concerned with the analysis of the relationship between pre-editing and machine translation's final output.

#### **2.2.1.1. Analysis of Pre-editing Rules in the Participants' Texts**

This part of the section aims at counting and presenting the frequency of occurrence of the pre-editing rules in the participants' simplified versions of the text. As it was previously mentioned, the selected text does not contain all the taught rules, instead it includes seven rules, which were seen necessary for the simplification of any text. They are as follows:

1. Sentences should not contain more than twenty (20) words. They should be well structured and they express one idea.
2. Passive voice should be avoided especially in descriptive writing.
3. Avoidance of synonymous words.
4. It is better to avoid using phrasal verbs.
5. The use of simplified terms.

6. Repetition of subject if necessary.
7. The use of proper punctuation.

Table 1

*Frequency of Use of Pre-editing Rules*

<b>Pre-editing rules</b>	<b>Frequency of occurrence</b>	<b>Percentage %</b>
Rule n°1	18	85.71%
Rule n°2	09	42.85%
Rule n°3	19	90.47%
Rule n°4	19	90.47%
Rule n°5	06	28.57%
Rule n°6	02	9.52%
Rule n°7	19	90.47%

As it is plainly shown in table 1, rule n°3, rule n°4 and rule n°7 are the most frequent rules and are equally used by the students with a percentage of 90.47%. Then, followed by rule n°1 which occurred in eighteen (18) of the student's texts. 42.85% of the total number of the students succeeded in applying the second rule, in which the majority of them used it the same way. However, the least frequent rules are the fifth and sixth rules. In which only 28.75% and 9.52% of the students managed to apply them properly.

### **2.2.1.2. Analysis of the Translation Outputs**

This part of the second section displays the results and findings of the translation process. As it was noted before, the participants performed the translation on their pre-edited texts using three different programs.

Human evaluation has been the main source for measuring the translation quality. Teachers of Mohamed Seddik Ben yahia University were asked to rate the translation outputs in terms of how faithful and intelligible they are. Afterwards, the researchers assigned grades

for the texts based on the aforementioned criteria. Remarkably, each criterion was out of three (/3). (See Appendices C, D and E for more details)

Table 2

*Descriptive Analysis of the Pre-edited Texts' Translation Produced by Google Translate*

<b>Texts</b>	<b>Fidelity/3</b>	<b>Percentage %</b>	<b>Intelligibility/3</b>	<b>Percentage %</b>
Text n°01	03	100%	02	66.66%
Text n°02	02	66.66%	03	100%
Text n°03	02	66.66%	01	33.33%
Text n°04	02	66.66%	02	66.66%
Text n°05	02	66.66%	01	33.33%
Text n°06	02	66.66%	02	66.66%
Text n°07	03	100%	03	100%
	<b>Total</b>	76.18%	<b>total</b>	66.66%

As indicated in table 2, only two texts were given the highest fidelity rating “completely faithful” with a percentage of 100%. The rest of the texts received “fairly faithful” with a percentage of 66.66%. In terms of intelligibility, two texts were given the highest rating “very intelligible” and six of them were rated “fairly intelligible” and scored two (2/3). The lowest intelligibility rating is “barely intelligible” and the evaluators gave it only to one text with a percentage of 33.33%. To sum up, the pre-edited texts translated by Google Translate were 76.18% faithful and 66.66% intelligible.

Table 3

*Descriptive Analysis of the Raw Text's Translation Produced by Google Translate*

<b>Text</b>	<b>Fidelity/3</b>	<b>Percentage %</b>	<b>Intelligibility/3</b>	<b>Percentage %</b>
Raw text	02	66.66%	02	66.66

The results from the above table show that the evaluators gave the raw translation of Google Translate the same fidelity and intelligibility ratings. The text scored two (2/3) on both criteria with a percentage of 66.66%.

Table 4

*Descriptive Analysis of the Pre-edited Texts' Translation Produced by Microsoft Translate*

<b>Texts</b>	<b>Fidelity/3</b>	<b>Percentage %</b>	<b>Intelligibility/3</b>	<b>Percentage %</b>
Text n°01	02	66.66%	03	100%
Text n°02	02	66.66%	02	66.66%
Text n°03	02	66.66%	02	66.66%
Text n°04	03	100%	02	66.66%
Text n°05	02	66.66%	02	66.66%
Text n°06	03	100%	02	66.66%
Text n°07	03	100%	02	66.66%
	<b>Total</b>	80.94%	<b>Total</b>	71.42%

As the above table demonstrates, the highest fidelity rate “very faithful” was given to three texts and the rest of them were rated as “fairly faithful”. As far as intelligibility is concerned, the highest score was given to one text with the rating of “very intelligible”, while

the rest of the texts scored two out of three (2/3) as “fairly intelligible”. It is worth noting that none of the texts scored zero (0) or given the lowest rate neither on fidelity nor on intelligibility. Briefly, Microsoft Translate produced seven texts that are 80.94% faithful and 71.42% intelligible.

Table 5

*Descriptive Analysis of the Raw Text's Translation Produced by Microsoft Translate*

<b>Text</b>	<b>Fidelity/3</b>	<b>Percentage %</b>	<b>Intelligibility/3</b>	<b>Percentage %</b>
Raw text	01	33.33%	01	33.33%

The results displayed in table 5, show that the raw text produced by Microsoft Translate scored the same on both criteria. The evaluators rated it as “barely faithful” in terms of fidelity and “barely intelligible” in terms of intelligibility.

Table 6

*Descriptive Analysis of the Pre-edited Texts' Translation Produced by Systran*

<b>Texts</b>	<b>Fidelity/3</b>	<b>Percentage %</b>	<b>Intelligibility/3</b>	<b>Percentage %</b>
Text n°01	02	66.66%	02	66.66%
Text n°02	03	100%	03	100%
Text n°03	02	66.66%	02	66.66%
Text n°04	02	66.66%	02	66.66%
Text n°05	01	33.33%	01	33.33%
Text n°06	01	33.33%	02	66.66%
Text n°07	02	66.66%	02	66.66%
	<b>Total</b>	61.90%	<b>total</b>	66.66%



As indicated in table 6, the majority of texts, four out of seven texts, were evaluated to be “fairly faithful” with a percentage of 66.66%. The “completely faithful” rating was given to one text only with the percentage of 100%. The two texts left were ranked to be “barely faithful” with a percentage of 33.33%. Concerning intelligibility, five texts were scored 2/3 as “fairly intelligible”. However, one of the two texts left was ranked to be “completely intelligible” and the other one was scored 1/3 as “barely intelligible”. Overall, Systran produced the translation of seven pre-edited texts with the percentage of 61.90% in terms of fidelity and 66.66% in terms of intelligibility.

Table 7

*Descriptive Analysis of the Raw Text's Translation Produced by Systran*

<b>Text</b>	<b>Fidelity/3</b>	<b>Percentage %</b>	<b>Intelligibility/3</b>	<b>Percentage %</b>
Raw text	02	66.66%	01	33.33%

Table 7 displays that the raw translation produced by Systran received the score of 2/3 with the percentage of 66.66% in terms of fidelity and the score 1/3 with the percentage of 33.33% in terms of intelligibility.

Table 8

*Comparison between Pre-edited and Raw texts' Translations in Terms of Fidelity*

	<b>Google translate</b>	<b>Microsoft Translate</b>	<b>Systran</b>	<b>Total</b>
Pre-edited texts	76.18%	80.94%	61.90%	73%
Raw Texts	66.66%	33.33%	66.66%	55.55%

The above table indicates that Microsoft Translate produced the best translation for the pre-edited texts in terms of fidelity with a percentage of 80.94%, followed by Google

Translate with a percentage of 76.18%. However, Systran scored the lowest with a percentage of 61.90%. In contrast to simplified texts, Google Translate and Systran's translations for the original text were faithful with a percentage of 66.66%, whereas Microsoft Translate translation received a percentage of 33.33%. Simply put, the pre-edited texts' translations received a higher percentage in terms of fidelity in comparison with the original text's translations.

Table 9

*Comparison between Pre-edited and Raw Texts' Translations in Terms of Intelligibility*

	<b>Google translate</b>	<b>Microsoft Translate</b>	<b>Systran</b>	<b>Total</b>
Pre-edited texts	66.66%	76.18%	62.37%	68.52%
Raw Texts	66.66%	33.33%	33.33%	44.44%

As it is shown in the above table, Microsoft Translate produced the best translations of the pre-edited texts in terms of intelligibility with a percentage of 76.18%, followed by both Google Translate and Systran with close percentages of 66.66% and 62.37%. On the other hand, concerning raw texts, Google Translate produced the best translations with a percentage of 66.66%, whereas the two other systems, Microsoft Translate and Systran, received an evaluation with a percentage of 33.33 %. To sum up, the total intelligibility percentage of the three systems' translations shows that the texts which were simplified through pre-editing process led to provide translations with higher intelligibility in comparison to the non-pre-edited texts.

### 2.2.1.3. Correlation Analysis between the Quality of Pre-editing and Machine Translation

This part of the analysis section aims at investigating the correlation between pre-editing and machine translation. In order to do so, the researchers first evaluated each pre-edited text in accordance with its translation. As stated before, the evaluation was based on a five-point scale. Therefore, the texts were scored from one (1) to five (5). The table below presents the evaluation scores.

Table 10

#### *Pre-editing and Translation Evaluation Scores*

<b>Texts</b>	<b>Pre-editing/5</b>	<b>Translation/5</b>
<b>Text n°1</b>	5	3
<b>Text n°2</b>	5	4
<b>Text n°3</b>	4	2
<b>Text n°4</b>	4	3
<b>Text n°5</b>	3	2
<b>Text n°6</b>	2	2
<b>Text n°7</b>	5	4
<b>Text n°8</b>	4	4
<b>Text n°9</b>	3	2
<b>Text n°10</b>	5	5
<b>Text n°11</b>	3	3
<b>Text n°12</b>	4	3
<b>Text n°13</b>	5	4
<b>Text n°14</b>	5	5
<b>Text n°15</b>	4	3

Table 10

*Continued*

<b>Text n°16</b>	5	4
<b>Text n°17</b>	2	2
<b>Text n°18</b>	4	4
<b>Text n°19</b>	4	2
<b>Text n°20</b>	1	1
<b>Text n°21</b>	2	1

As it is demonstrated in the above table, most of the translations have good scores after a well pre-editing of the text. For example, texts n° 2, 7, 10, 13, 14, and 16, scored the highest, for the pre-editing was very well performed which in turn, resulted in accurate and remarkable translations. The latter scored between four (4) and five (5). In addition, it should be worth of noting that when the quality of pre-editing was poor, the translation quality was poor. To illustrate, the translation of text n°20 rated “very bad” and scored one (1) after a poor performance of pre-editing of the text which also scored one (1).

In order to determine whether there is a relationship between the pre-edited texts’ quality and their machine translation’s quality, the researchers opted for a statistical means known as Pearson Product – Moment Correlation Coefficient ( $r$ ) which was calculated using SPSS system.

Table 11

*The Pearson Correlation between Pre-editing and Machine Translation*

Variables	Pre-editing
Machine translation	Pearson correlation ( $r$ )
	Sig(P-value)
	N

*Note.* \* $p < .05$ . \*\* $p < .01$       N = Number of participants      p = Probability coefficient

According to what is displayed in the above table, the Pearson's  $r$  is positive; thus, there is a positive relationship between pre-editing and machine translation. Whenever pre-editing improves, machine translation improves as well. Additionally, the relationship between the two variables is significantly strong since the Pearson's  $r = 0.831$ . When the latter is close to 1, it indicates that the correlation between variables is strong. However, when it is close to 0, it means that there is either weak or no relationship between the variables. As Evans (1996) suggested, the absolute value of  $r$  may belong to five ranges:

When  $r$  ranges between 0.00 and 0.19: the correlation is very weak.

When it ranges between 0.20 and 0.39: the correlation is weak.

When it ranges between 0.40 and 0.59: the correlation is moderate.

When it ranges between 0.60 and 0.79: the correlation is strong.

Finally, when it ranges between 0.80 and 1.0, as it is the case in this research study, the correlation is very strong. Moreover, the significance of correlation is also explained by the probability coefficient ( $p$ ) which equals 0.00 (\*\* $p < 0.01$ ).

## Conclusion

This part has taken as its major concern the analysis of data collected during the experiment. Firstly, the analysis of pre-editing process displayed the rules which the students succeeded to apply as an attempt to simplify the original text. Then, the texts that were translated after the subjects' interference were evaluated based on two main criteria:

intelligibility and fidelity. After comparing the results to the translation of the original text, it was found that pre-edited texts produced a more faithful and comprehensible translation. Additionally, this section has been also devoted to display the results of the statistical analysis of the correlation between the quality of pre-editing and machine translation. First, the quality of both pre-edited texts and their translation was evaluated on a five-scale basis. Then, the analysis of the results revealed that there is a strong and positive relationship between the two investigated variables. Hence, when the text is pre-edited correctly and accurately, it leads to provide a correct and accurate translation.

## **Part Two: Data Discussion and Interpretation**

### **Introduction**

This part focuses on the interpretation and discussion of the overall results drawn from the analysis. Hence, the research questions that were raised at the outset of the current study will be answered. Furthermore, based on the results, the researchers will confirm or reject the hypotheses that were set prior to the experiment.

This research work sought to answer the following questions:

1. To what extent do third year EFL learners at Mohamed Seddik Ben Yahia University apply pre-editing rules while simplifying their texts?
2. To what extent does pre-editing affect machine translation's intelligibility and fidelity?
3. Is there a relationship between pre-editing quality and machine translation quality?

#### **2.2.2.1. The Frequency of Use of Pre-editing Rules**

The analysis of the data unveiled the general image of the use of pre-editing rules by 21 third year EFL students in the process of simplifying the source text. It showed the most frequently used rules that were successfully applied by the subjects as well as the rules which the students did not really focus on. As the data gathered displayed, most of the subjects were

able to follow the suitable rules in order to simplify the original text as much as they could. The majority of students with a percentage of 90, 47% followed the rules of avoiding phrasal verbs and synonymous words. They changed the phrasal verb “made up” with suitable simple verbs such as create, form and establish, and omitted the word “millennia” which has practically the same meaning as the word “centuries”. Thus, it can be said that those rules were easily adopted by the students. In addition to the previously stated rules, the use of accurate punctuation along with the use of simple sentences was also successfully done by most of the subjects. To make it easier for the machine, they replaced the verbs “was and remains” in the first sentence by the simple verb “is”; and they, most of the time, put the verb “should” instead of “can and must be” to avoid wordiness that may be caused in the target language. However, the remaining rules were found to be somehow difficult for most of the subjects. Some of them, almost 43% of the students, were able to change the second sentence which is written in the passive voice into “Over the centuries, the reciprocal understanding and interpenetration have enriched those cultures and so have created the unique mosaic of human civilization”. Others, a number of 6 students, could replace the word “mosaic” by the combination “in an artistic way” or “unique diversity”. Finally, only two students applied the rule that concerns repetition of the subject if necessary. They repeated the subject “project” instead of the preposition “which” in the last sentence of the text. All in all, the results showed that the subjects succeeded in appropriately applying most of the rules which led to provide correct simplified versions of the original text and which positively affected the translations provided by the computer.

### **2.2.2.2. The Effectiveness of Pre-editing on Machine Translation's Final Output**

The effectiveness of pre-editing on the final output of machine translation is analyzed through comparing the translation of pre-edited texts with the translation of the original text on the basis of two main criteria: intelligibility and fidelity

To ensure the effect that pre-editing causes on machine translation, the researchers decided to adopt three electronic systems: Google Translate, Microsoft Translate and Systran. They believed that this would assure the reliability of the work. Thus, the results obtained from evaluating the translations confirmed the first suggested hypothesis. At the beginning, they evaluated the translations done by each system in terms of both pre-edited and non-pre-edited texts separately. The results showed that each time the simplified texts' translation surpassed the raw texts' translation with a significant difference concerning both fidelity and intelligibility. Then, they summarized the results in a general comparison which considered the three systems as a translation done by the computer. Thus, the huge difference in the percentage of both faithfulness and comprehensibility of the texts revealed the fact that pre-editing strongly and positively affects machine translation.

### **2.2.2.3. The Correlation between Pre-editing Quality and Machine Translation Quality**

Investigating the relationship between pre-editing and machine translation is the core of the current research. Pearson's Product Moment was run to assess this relationship, which in fact detected a strong positive correlation between pre-editing quality and machine translation's final output. As it was mentioned in the previous section, the strength of the correlation is determined by the magnitude or the absolute level of the Pearson Correlation Coefficient. Accordingly, the results suggest that the better the pre-editing, the better the translation, and the worse the pre-editing quality, the worse the translation produced by the



machine. More importantly, Pearson's  $r$  indicates that the correlation is statistically significant, for the p-value is less than the significant level 0.01. In a nutshell, there is conclusive evidence about the strength as well as the significance of the relationship between pre-editing and machine translation.

## **Conclusion**

This part aimed to discuss the major findings that sprung up from the data analysis. Primarily, it gave a detailed picture of the students' use of pre-editing rules. Then, it provided a conclusion about the effect of pre-editing on machine translation's final output. Lastly, it explained the nature of the relationship between the pre-editing quality and machine translation's quality.

## **Recommendations**

According to what has been achieved in this research study, the positive effect of pre-editing on machine translation led to state the following recommendations.

Due to the huge dependence on electronic means in the field of teaching and learning, syllabus designers are recommended focusing on the importance of teaching the module of machine translation for its crucial benefits on EFL students as future users of the English language.

Teachers are required to teach the pre-editing process for two main reasons:

- It makes the use of machine translation more helpful and may eliminate the process of post-editing.
- It positively affects the students' language. It makes them able to express their ideas in a more direct and controlled language. More importantly, it improves their critical thinking and helps them acquire how to analyze different types of texts.

## General Conclusion

This study is an attempt to shed light on the effectiveness of pre-editing on machine translation's final output. Thus, it aims to confirm or reject the hypotheses which state that: there is a positive effect of the pre-editing process on machine translation's final output. Also, there is a significant relationship between pre-editing quality and machine translation quality. To achieve the study's aim, the researchers investigated the students' use of pre-editing rules to simplify the original text. Then, they evaluated and compared the translations of the simplified texts and the original text as an attempt to check the effect of pre-editing. Finally, they explored the nature of the relationship between the quality of pre-edited texts and their translation which was done by computer systems.

The thesis is composed of two chapters. Being divided into two sections, the first chapter reviews the major theoretical aspects related to the field of machine translation in the first section, and presents a general overview of the pre-editing approach in the second section. Equally divided into two sections, the second chapter elucidates the practical part of this research. On the one hand, the first section introduces the research methodology followed to pursue the current study. On the other hand, the second section is divided into two parts. The first part provides the statistical and correlation analysis of the collected data, whereas the second part deals with the interpretation and discussion of the results obtained from data analysis.

As regards the main findings of this experimental study, it was found that third year EFL students have successfully applied most of the pre-editing rules, mainly, the avoidance of phrasal verbs and synonymous words, as well as the use of well-structured sentences. Furthermore, the results showed that pre-editing truly improves machine translation's intelligibility and fidelity. Lastly, the findings revealed that when the students produced well-pre-edited texts with correct grammar and sentence structure, the translation was of good

quality. However, when the pre-edited texts lacked coherence and cohesion, the translation was in some way inaccurate and of poor quality. Following this, it can be concluded that there is indeed a relationship between pre-editing quality and machine translation quality.

Narrowly speaking, the findings gathered in this study revealed that the majority of the participants are unaware of the pre-editing approach nor of its rules. Hence, the findings justify the need for recommending the emphasis of teaching pre-editing in relation to machine translation.

In a nutshell, due to the fact that the crux of the current research was to uncover the possible effect of pre-editing on machine translation, and after the findings generated from the analysis, the teachers are highly recommended to expose their students to pre-editing rules so as to achieve better translation quality.

### References

- Abdulsalami, B. A. & Akinsanya, B.J. (2017). Review of different approaches for machine translations. *International Journal of Mathematics Trends and Technology (IJMTT)*, 48(3), 197-202.
- ALPAC. Division of Behavioral Sciences, National Academy of Sciences & National Research Sciences. (1966). *Language and machines: computers in translation and linguistics*. Washington, WA: a report by the Automatic Language Processing Advisory Committee.
- Arnold, D., Balkan, L., Meijer, S., Humphreys, R. L., & Sadler, L. (1994). *Machine translation: An introductory guide*. London, England: NCC Blackwell.
- Arnold, D. (2003). Why translation is difficult for computers. In H. Somer (Ed), *Computers and translation*. Amsterdam, Netherlands: John Bnejamins
- Apuke, O. D. (2017). Quantitative research method: *A synopsis approach*. *Arabian Journal of Business and Management Review*, 6(10), 40-47
- Baker, K. L., Franz, A. M., Jordan, P. W., Mitamura, T. & Nyberg, E. H. (1994). Coping with ambiguity in large-scale machine translation systems. Proceedings of the 15th International Conference on Computational Linguistics, Pittsburgh, PA: Carnegie Mellon University.
- Beldjazia, A., & Alatou, D. (2016). Precipitation variability on the massif forest of Mahouna (North Eastern-Algeria). *International Journal of Management Sciences and Business Research*. 5(3). 21-28
- Carl, M. & Way, A. Eds. (2003). Recent advances in example-based machine translation. 21. Saarbrucken, Germany.
- Chee, D., J. (2015). *Pearson's Product-Moment Correlation: Sample Analysis*. Hawaii: University of Hawaii at Mānoa School of nursing.

- Cho, K., Merriënboer, B. V., Gulcehre, C., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation.
- Costa-Jussà, M. R., Rapp, R., Lambert, P., Eberle, K., Banchs, R. E., & Babych, B. (2006). Hybrid approaches to machine translation. In J. Hirschberg, E. Hovy & M. Johnson (Eds.). *Theory and applications of natural language processing*. Switzerland: Springer International Publishing AG.
- Cramer, D. (2003). *Fundamental statistics for social research: Step by step calculations and computer techniques using SPSS for windows*. London: Routledge.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3<sup>rd</sup> ed.). Oaks, CA: SAGE.
- Douglas, S. & Hurst, M. (1996). Controlled language support for Perkins Approved Clear English (PACE). Conference Proceedings of the First International Workshop on Controlled Language Applications, Leuven, Belgium.
- Garvin, P. L. (1956). Some linguistic problems in machine translation. In M. Halle et al. (Eds), *For Roman Jakobson*. The Hague, Netherlands: Mouton.
- Goutte, C. (2006). *Automatic evaluation of machine translation quality*, January. Canada: National Research Council Canada.
- Halliday, T. C. & Briss, E. A. (1977). The evaluation and systems analysis of the Systran machine translation system. Columbus, OH: Battelle Columbus Laboratories.
- Hassan, B. A. (2014). *Between English and Arabic: A practical course in translation*. Cambridge Scholars Press.
- Hermann, K. M. & Blunsom, P. (2014). Multilingual distributed representations without word alignment. Oxford: Oxford University Press.
- Hutchins, J. (1986). Machine translation: Past, present, future. In B. L. Meek (Ed), *Computers*

- and their applications*. New York, NY: Halsted Press.
- Hutchins, J., & Somers, H. L. (1992). *An introduction to machine translation*. London, England: Academic Press.
- Hutchins, J. (2002). Machine translation today and tomorrow. In G. Willée, B. Schröder, & H. Schmitz (Eds.), *Computational linguistics: Achievements and perspectives* (pp. 159-162). Sankt Augustin: Gardez! Verlag.
- Hutchins, J. (2005). Current commercial machine translation systems and computer-based translation tools: system type and their uses. *The International Journal of Translation*, 2(2), 5-38.
- Kaji, H. (1999). *Controlled languages for machine translation: State of the art*. Tokyo, Japan: Hitachi, Ltd.
- Kalchbrenner, N. & Blunsom, P. (2013). *Recurrent continuous translation models*. Oxford: Oxford University Press.
- Katsberg, P. (2012). Machine translation tools: Tools of the translator's trade. *Communication and Language at Work Journal*, (1), 34-45.
- Kit, C., Pan, H., & Webster, J.J. (2002). Example-based machine translation: A new paradigm. In S.W. Chan (Ed.), *Translation and information technology* (pp. 57-78). Hong Kong, China: Chinese University Press.
- Koehn, P. Och, J. F., & Marcu, D. (2003). Statistical phrase-based translation. Conference Proceedings of HLT-NAACL. California: Southern California University.
- Koehn, P. and Monz, C. (2006). Manual and automatic evaluation of machine translation between European languages. Conference Proceedings of the Workshop on Statistical Machine Translation. Edinburgh, UK.
- Koehn, P. (2010). *Statistical machine translation*. Cambridge: Cambridge University Press.
- Miller, K. J., Gates, D. M., Underwood, N. & Magdalen, J. (2001). *Evaluation of machine*

- translation output for an unknown language source: Report of an ISLE-based investigation.* Washington D.C: The MITRE Corporation.
- Mitamura, T. & Nyberg, E. (1999). Automatic rewriting for controlled language translation. Conference Proceedings of the Sixth Natural language Processing Pacific Rim Symposium Post-Conference Workshop Automatic Paraphrasing: Theories and Applications.
- Nunan, D. (1992). *Research methods in language learning.* Cambridge: Cambridge University Press.
- Nunan, D. & Bailey, K. M. (2009). *Exploring second language classroom research: A Comprehensive guide.* Boston, MA : Hainle, Cengage Learning.
- Nyberg, E., Mitamura, T. & Carbonell, J. G. (1994). *Evaluation metrics for knowledge-based machine translation.* Pittsburgh, PA: Carnegie Mellon University.
- Nyberg, E., Mitamura, T. & Huijsen, W. O. (2003). Controlled language for authoring and Translation. In H. Somers (Ed), *Computers and Translation.* Amsterdam, Netherlands: John Benjamins.
- Quah, C. (2006). *Translation and technology.* New York, NY: Palgrave Macmilan.
- Reuther, U. (2003). Two in one-can it work? Readability and translatability by means of Controlled language. Conference Proceedings of EAMT/CLAW2003, Dublin, Ireland.
- Saini, S. & Sahula, V. (2015). A survey of machine translation techniques and systems of Indian languages. Jaipur, India.
- Somers, H. (2008). Corpora and machine translation. In A. Ludeling, M. Kyto & T. McEnery (Eds), *Corpus linguistics: An international handbook.* Berlin, Germany: Mouton De Gruyter.
- Uzkoreit, H. (2007). Survey of machine translation evaluation. In X. Gros (Ed), *Statistical.*

*and hybrid machine translation between all European languages.* Saarbrücken, Germany.

Van Slype, G. (1979). Critical study of methods for evaluating the quality of machine translation. Final Report, Bureau Marcel van Dijk, Brussels.

Wojcik, R. (1998). Boeing technical English: An extension of AECMA SE beyond the aircraft maintenance domain. Proceedings of the Controlled Language Application Workshop.



## **Appendix A**

### **Pre-editing Rules**

1. Sentences should not contain more than twenty (20) words. They should be well structured and they express one idea.
2. Passive voice should be avoided especially in descriptive writing.
3. Slangs and automatic expressions should be totally avoided.
4. Avoidance of synonymous words.
5. It is better to avoid using phrasal verbs.
6. –ing and –ed forms should be carefully avoided after conjunctions.
7. Relative pronouns should be used to introduce relative clauses.
8. Avoidance of ambiguous words and use of simplified terms.
9. Repetition of subject when necessary.
10. Use of proper punctuation.
11. Spell checking.

## **Appendix B**

### **The original Text**

Dialogue between cultures was and remains the main road for the development of human civilization. Through the reciprocal understanding and interpenetration of cultures over the centuries and millennia, those cultures have been mutually enriched, and so have made up the unique mosaic of human civilization. It is this dialogue that can and must be the answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. This is the essence of the “Dialogue between Cultures” project which was proposed in UNESCO and has the support of cultural figures in various countries.

## Appendix C

### Fidelity and Intelligibility Evaluation (Google Translate)

#### Pre-edited text n°01

Dialogue between cultures is the key for the development of human civilization. Through the mutual understanding and contradictions between cultures over time, those cultures have been mutually enriched, and so have created the unique combination of human civilization. This dialogue should be the answer to the growing danger of intolerance and violence. This is the purpose of "Dialogue between Cultures" which was presented in the United Nations educational scientific and cultural organization and was supported by many cultural figures worldwide.

#### Translation n°01

الحوار بين الثقافات هو مفتاح تطور الحضارة الإنسانية. من خلال التفاهم المتبادل والتناقضات بين الثقافات مع مرور الوقت، تم إثراء تلك الثقافات بشكل متبادل، وخلق بذلك مزيج فريد من الحضارة الإنسانية. يجب أن يكون هذا الحوار هو الرد على الخطر المتزايد المتمثل في عدم التسامح والعنف. هذا هو الغرض من "الحوار بين الثقافات" الذي تم تقديمه في منظمة الأمم المتحدة للتربية والعلوم والثقافة والذي كان مدعومًا من قبل العديد من الشخصيات الثقافية في جميع أنحاء العالم.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Pre-edited text n°02

Dialogue between cultures is the main road for the development of human civilization. Through time, cultures have been mutually enriched through the exchanged understanding and interpenetration. Therefore, they created the unique mosaic of human civilization. The aim behind "Dialogue between Cultures» project, which was proposed in UNESCO and has the support of cultural figures in various countries, is to stop the growing danger of various manifestations of intolerance, violence, and aggressive nationalism.

## Translation n°02

الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الإنسانية. عبر الزمن، تم إثراء الثقافات بشكل متبادل من خلال التفاهم المتبادل والاختراق. لذلك، أنشأوا الفيسفساء الفريدة للحضارة الإنسانية. الهدف من وراء مشروع "الحوار بين الثقافات"، الذي تم اقتراحه في اليونسكو ويحظى بدعم الشخصيات الثقافية في مختلف البلدان، هو وقف الخطر المتزايد لمختلف مظاهر التعصب والعنف والقومية العدوانية.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Pre-edited text n°03

Dialogue between cultures was and remains the main way for the development of human civilization. Cultures have been mutually enriched and they established the unique mosaic of human civilization through the reciprocal understanding and interpenetration over centuries. This dialogue between cultures should be the answer to the growing danger of various manifestations of intolerance, violence, and aggressive nationalism. This is the core of the project "Dialogue between Cultures" which was proposed in UNESCO and has the support of cultural figures worldwide.

### Translation n°03

كان الحوار بين الثقافات ولا يزال هو الطريق الرئيسي لتنمية الحضارة الإنسانية. تم إثراء الثقافات بشكل متبادل وأنشأوا الفسيفساء الفريدة للحضارة الإنسانية من خلال التفاهم المتبادل والاختراق على مر القرون. يجب أن يكون هذا الحوار بين الثقافات هو الرد على الخطر المتزايد للعديد من مظاهر التعصب والعنف والقومية العدوانية. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي تم اقتراحه في اليونسكو ويحظى بدعم الشخصيات الثقافية في جميع أنحاء العالم.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

#### Pre-edited text n°04

Dialogue between cultures remains the main road for the development of human civilization. Those cultures have been mutually enriched through the reciprocal understanding and interpenetration over the centuries. Thus, they created the unique mosaic of human civilization. This dialogue should be the answer to the growing danger of various manifestations of intolerance, violence, and aggressive nationalism. This is the essence of the “Dialogue between cultures” project which was proposed in UNESCO and has the support of cultural figures in various countries.

#### Translation n°04

يبقى الحوار بين الثقافات الطريق الرئيسي لتنمية الحضارة الإنسانية. وقد تم إثراء تلك الثقافات بشكل متبادل من خلال التفاهم المتبادل والتداخل على مر القرون. وهكذا، خلقوا الفسيفساء الفريدة للحضارة الإنسانية. يجب أن يكون هذا الحوار هو الرد على الخطر المتزايد للعديد من مظاهر التعصب والعنف والقومية العدوانية. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي تم اقتراحه في اليونسكو ويحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Pre-edited text n°05

Dialogue between cultures is the main road for the development of human civilization. The mutual understanding of cultures over centuries made them to mutually enriched and created the unique mosaic of human civilization. It is this dialogue between cultures which can and must be the answer to the growing danger of various manifestations of intolerance and violence. This is the essence of the “Dialogue between cultures” project, which was proposed in UNESCO and has the support of cultural figures in multiple countries.

## Translation n°05

الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الإنسانية. جعلهم التفاهم المتبادل للثقافات على مر القرون من إثراء متبادل وخلق فسيفساء فريدة من الحضارة الإنسانية. إن هذا الحوار بين الثقافات هو الذي يمكن ويجب أن يكون هو الحل لخطر متزايد من مختلف مظاهر التعصب والعنف. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي تم اقتراحه في اليونسكو ويحظى بدعم الشخصيات الثقافية في العديد من البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Pre-edited text n°06

Dialogue between cultures was and remains the main way for the development of human civilization. The reciprocal understanding and interpenetration of cultures over the centuries and millennia led those cultures to be mutually enriched. They created the unique mosaic of human civilization. It is this dialogue that must be the answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. This is the essence of the "Dialogue between Cultures" project which was proposed in UNESCO and has the support of cultural figures in various countries.

### Translation n°06

كان الحوار بين الثقافات ولا يزال هو الطريق الرئيسي لتنمية الحضارة الإنسانية. أدى التفاهم المتبادل والاختراق بين الثقافات على مر القرون والآلاف السنين إلى إثراء تلك الثقافات. لقد صنعوا الفسيفساء الفريدة للحضارة الإنسانية. هذا الحوار هو الذي يجب أن يكون الجواب على الخطر المتزايد للعديد من مظاهر التعصب والعنف، بما في ذلك القومية العدوانية. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي تم اقتراحه في اليونسكو ويحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	



## Pre-edited text n°07

Dialogue between cultures is the main road for the development of human civilization. The reciprocal understanding and interpenetration of cultures over the centuries have enriched those cultures. consequently, they created the unique mosaic of human civilization. It is this dialogue that must be the answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. This is the main purpose of the “Dialogue between Cultures” project which was proposed in UNESCO and has the support of cultural figures in various countries.

## Translation n°07

الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الإنسانية. التفاهم المتبادل والاختراق بين الثقافات على مر القرون قد أثرى تلك الثقافات. وبالتالي، خلقوا الفسيفساء الفريدة للحضارة الإنسانية. هذا الحوار هو الذي يجب أن يكون الجواب على الخطر المتزايد للعديد من مظاهر التعصب والعنف، بما في ذلك القومية العدوانية. هذا هو الهدف الرئيسي لمشروع "الحوار بين الثقافات" الذي تم اقتراحه في اليونسكو ويحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Translation of the original text (without pre-editing)

كان الحوار بين الثقافات ولا يزال الطريق الرئيسي لتنمية الحضارة الإنسانية. من خلال التفاهم المتبادل والاختراق بين الثقافات على مر القرون والآلاف السنين، تم إثراء تلك الثقافات بشكل متبادل، وبالتالي شكلت الفسيفساء الفريدة للحضارة الإنسانية. هذا الحوار هو الذي يمكن ويجب أن يكون هو الحل للخطر المتزايد المتمثل في مختلف مظاهر التعصب والعنف، بما في ذلك القومية العدوانية. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي تم اقتراحه في اليونسكو ويحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Appendix D

### Fidelity and Intelligibility Evaluation (Microsoft Translate)

#### Pre-edited text n°01

Dialogue between cultures is the main road for the development of human civilization. The reciprocal understanding and interpenetration of cultures over centuries has enriched those cultures, and this has formed the unique mosaic of human civilization. This dialogue between cultures can answer the growing danger of various manifestations of intolerance and violence. Aggressive nationalism is included. This is the basis of the "Dialogue between Cultures" project, which was proposed in UNESCO. The project, which has the support of cultural figures in various countries.

#### Translation n°01

إن الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الإنسانية. وقد أثرى التفاهم المتبادل والتفاعل بين الثقافات على مدى القرون تلك الثقافات، وشكل الفسيفساء الفريدة للحضارة الإنسانية. هذا الحوار بين الثقافات يمكن ان يجيب على الخطر المتزايد لمختلف مظاهر التعصب والعنف. بما في ذلك القومية العدوانية. وهذا هو أساس مشروع "الحوار بين الثقافات"، الذي اقترح في اليونسكو. المشروع، الذي دعمته الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Pre-edited text n°02

The exchange of cultures through dialogue is the main road for the development of human civilization. The reciprocal understanding and interpenetration of cultures over the centuries has enriched those cultures and has built the unique mosaic of human civilization. The dialogue between cultures is the answer to the growing danger of various manifestations of intolerance, violence and aggressive nationalism. UNESCO proposed the “Dialogue between Cultures” project and has the support of important people in various countries.

## Translation n°02

ان تبادل الثقافات من خلال الحوار هو الطريق الرئيسي لتنمية الحضارة الإنسانية. وقد أثرى الفهم المتبادل وتداخل الثقافات علي مر القرون تلك الثقافات وبني السيفساء الفريدة للحضارة الإنسانية. والحوار بين الثقافات هو الرد على الخطر المتزايد لمختلف مظاهر التعصب والعنف والنزعة القومية العدوانية. واقتрحت اليونسكو مشروع "الحوار بين الثقافات" ودعم من أشخاص مهمين في بلدان مختلفة.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Pre-edited text n °03

Dialogue between cultures is the main road for the development of human civilization. Through the reciprocal understanding and interpenetration of cultures over centuries, those cultures have been mutually enriched, and so have built the unique diversity of human civilization. This dialogue between cultures can be the answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. This is the essence of the "Dialogue between Cultures" project, which was proposed in UNESCO and has the support of cultural figures in various countries.

### Translation n°03

ان الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الإنسانية. ومن خلال التفاهم المتبادل والتفاعل بين الثقافات علي مر القرون، تم إثراء تلك الثقافات بصوره متبادله، وهكذا بنت التنوع الفريد للحضارة الإنسانية. وهذا الحوار بين الثقافات يمكن ان يكون الرد على الخطر المتزايد لمختلف مظاهر التعصب والعنف، بما في ذلك النزعة القومية العدوانية. هذا هو جوهر مشروع "الحوار بين الثقافات"، الذي اقترح في اليونسكو والذي يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	مفهوم جدا	
<b>Grade 2</b>	مفهوم الى حد ما	
<b>Grade 1</b>	بالكاد مفهوم	
<b>Grade 0</b>	غير مفهوم تماما	

#### Pre-edited text n°04

Dialogue between cultures was and remains the main road for the development of human civilization. Through the reciprocal understanding and interpenetration of cultures over the centuries, those cultures have been mutually enriched, made up the unique mosaic of human civilization. The dialogue between cultures must be the answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. The essence of the project entitled the “Dialogue between Cultures” was proposed in UNESCO and it has the support of cultural figures in various countries.

#### Translation n°04

وكان الحوار بين الثقافات ولا يزال الطريق الرئيسي لتنمية الحضارة الإنسانية. ومن خلال التفاهم المتبادل والتفاعل بين الثقافات علي مر القرون، تم إثراء تلك الثقافات بشكل متبادل، وشكلت فسيفساء فريدة من نوعها للحضارة الإنسانية. ويجب ان يكون الحوار بين الثقافات هو الرد على الخطر المتزايد لمختلف مظاهر التعصب والعنف، بما في ذلك النزعة القومية العدوانية. تم اقتراح جوهر المشروع المعنون "الحوار بين الثقافات" في اليونسكو، وهو يدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Pre-edited text n°05

Dialogue between cultures is the main road for the development of human civilization. The cultures have been mutually enriched through the reciprocal understanding and interpenetration over the centuries. Thus, they created the unique mosaic of human civilization. It is this dialogue between cultures, which should be the answer to the growing danger of various manifestations that include intolerance, violence and aggressive nationalism. This is the essence of the "Dialogue between Cultures" project, which was proposed in UNESCO and has the support of cultural figures in various countries.

### Translation n°05

ان الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الإنسانية. وقد تم إثراء الثقافات بشكل متبادل من خلال التفاهم المتبادل والتداخل على مر القرون. وهكذا، فإنها خلقت فسيفساء فريدة من الحضارة الإنسانية. هذا الحوار بين الثقافات هو الذي ينبغي ان يكون الحل للخطر المتزايد لمختلف المظاهر التي تشمل التعصب والعنف والنزعة القومية العدوانية. وهذا هو جوهر مشروع "الحوار بين الثقافات"، الذي اقترح في اليونسكو والذي يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Pre-edited text n°06

Dialogue between cultures was and remains the main reason for the development of human civilization. The reciprocal understanding and interpenetration between cultures over the centuries has enriched those cultures and has created a unique mosaic of human civilization. This dialogue between cultures must be the answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. This is the content of the "Dialogue between Cultures" project, which was proposed in UNESCO and was supported by cultural figures in various countries.

## Translation n°06

كان الحوار بين الثقافات ولا يزال السبب الرئيسي لتطور الحضارة الإنسانية. وقد أثرى التفاهم المتبادل والتداخل بين الثقافات علي مر القرون تلك الثقافات وشكل فسيفساء فريدة من نوعها للحضارة الإنسانية. يجب ان يكون هذا الحوار بين الثقافات هو الرد على الخطر المتزايد لمختلف مظاهر التعصب والعنف، بما في ذلك النزعة القومية العدوانية. وهذا هو مضمون مشروع "الحوار بين الثقافات"، الذي اقترحه اليونسكو وأيدته شخصيات ثقافية في بلدان مختلفة.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	



### Pre-edited text n°07

Dialogue between cultures is still the main way for the development of human civilization. The reciprocal interpenetration of cultures over the centuries has enriched those cultures, and so has created the unique mosaic of human civilization. This dialogue between cultures can be the answer to the growing danger of various manifestations of violence, including aggressive nationalism. This is the essence of the "Dialogue between Cultures" project, which was proposed in UNESCO and was supported by cultural figures in various countries.

### Translation n°07

لا يزال الحوار بين الثقافات هو السبيل الرئيسي لتنمية الحضارة الإنسانية. وقد أثري التفاعل المتبادل بين الثقافات علي مر القرون تلك الثقافات، وهكذا خلق فسيفساء فريدة من الحضارة الإنسانية. هذا الحوار بين الثقافات يمكن ان يكون الرد على الخطر المتزايد لمختلف مظاهر العنف، بما في ذلك النزعة القومية العدوانية. وهذا هو جوهر مشروع "الحوار بين الثقافات"، الذي اقترح في اليونسكو وأيدته شخصيات ثقافية في بلدان مختلف.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	مفهوم جدا	
<b>Grade 2</b>	مفهوم الى حد ما	
<b>Grade 1</b>	بالكاد مفهوم	
<b>Grade 0</b>	غير مفهوم تماما	

### Translation of the original text (without pre-editing)

وكان الحوار بين الثقافات ولا يزال الطريق الرئيسي لتنمية الحضارة الإنسانية. ومن خلال التقاهم المتبادل والتفاعل بين الثقافات علي مر القرون وألفيات، تم إثراء تلك الثقافات بشكل متبادل، التالي فقد شكلت الفسيفساء الفريدة للحضارة الإنسانية. وهذا الحوار هو الذي يمكن ويجب ان يكون الرد على الخطر المتزايد لمختلف مظاهر التعصب والعنف، بما في ذلك النزعة القومية العدوانية. وهذا هو جوهر مشروع "الحوار بين الثقافات" الذي اقترح في اليونسكو والذي يدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Appendix E

### Fidelity Evaluation (Systran)

#### Pre-edited text n°01

Dialogue between cultures is the main way for the development of human civilization. Those cultures have been mutually enriched through reciprocal understanding and interpenetration over the centuries and so have created the unique mosaic of human civilization. It is this dialogue that should be the answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. This is the essence of the “Dialogue between Cultures” project, which was proposed in UNESCO and has the support of cultural figures in various countries.

#### Translation n°01

ان الحوار بين الثقافات هو السبيل الرئيسي لتنمية الحضارة الانسانية. وقد تم اثراء تلك الثقافات من خلال التفاهم المتبادل والتداخل على مر القرون، كما خلقت الفسيفساء الفريدة للحضارة الانسانية. هذا هو الحوار الذي ينبغي ان يكون الرد على الخطر المتنامي لمختلف مظاهر التعصب والعنف، بما في ذلك القومية العدوانية. - وهذا هو جوهر مشروع "الحوار بين الثقافات" الذي اقترح في اليونسكو، وهو يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Pre-edited text n°02

Dialogue between cultures remains the main road for the development of human civilization. The reciprocal understanding of cultures over the centuries has enriched those cultures and made up the unique mosaic of human civilization. This dialogue can be the answer to the growing danger of various manifestations of intolerance, including aggressive nationalism. This is the essence of the project "Dialogue between Cultures" which was proposed in UNESCO. The project, which has the support of cultural figures in various countries.

## Translation n°02

لا يزال الحوار بين الثقافات الطريق الرئيسي لتنمية الحضارة الإنسانية. وقد أثرى الفهم المتبادل للثقافات على مر القرون تلك الثقافات وشكل الفسيفساء الفريدة للحضارة الإنسانية. قد يكون هذا الحوار هو الرد على الخطر المتنامي لمختلف مظاهر التعصب، بما في ذلك القومية العدوانية. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي اقترح في اليونسكو، المشروع الذي يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Pre-edited text n°03

Dialogue between cultures remains the main device for the development of human civilization. Through the reciprocal understanding and interpenetration of cultures over the centuries, those cultures have been enriched and created the unique mosaic of human civilization. This dialogue must be the answer to the growing danger of social phenomena. "Dialogue between Cultures" was proposed in UNESCO and has the support of cultural figures in various countries.

### Translation n°03

لا يزال الحوار بين الثقافات الاداة الرئيسية لتنمية الحضارة الانسانية. فمن خلال التفاهم المتبادل والتداخل بين الثقافات على مر القرون، تم اثراء تلك الثقافات وشكلت الفسيفساء الفريد للحضارة الانسانية. هذا الحوار يجب ان يكون الاجابة على الخطر المتنامي للظواهر الاجتماعية. اقترح "الحوار بين الثقافات" في اليونسكو، وهو يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	مفهوم جدا	
<b>Grade 2</b>	مفهوم الى حد ما	
<b>Grade 1</b>	بالكاد مفهوم	
<b>Grade 0</b>	غير مفهوم تماما	

#### Pre-edited text n°04

Dialogue between cultures is the main road for the development of human civilization. The reciprocal understanding and interpenetration between cultures over the centuries has enriched those cultures. Thus, they created the unique mosaic of human civilization. It is this dialogue between cultures, which should be the answer to the growing danger of various manifestations that include intolerance, violence and aggressive nationalism. This is the essence of the "Dialogue between Cultures" project, which was proposed in UNESCO and has the support of cultural figures in various countries.

#### Translation n°04

ان الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الانسانية. وقد أثرى التفاهم المتبادل والتداخل بين الثقافات على مر القرون تلك الثقافات. وبالتالي، فقد خلقوا الفسيفساء الفريدة للحضارة الانسانية. وهذا هو الحوار بين الثقافات، الذي ينبغي ان يكون الرد على الخطر المتزايد لمختلف المظاهر التي تشمل التعصب والعنف والقومية العدوانية. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي اقترح في اليونسكو، وهو يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	مفهوم جدا	
<b>Grade 2</b>	مفهوم الى حد ما	
<b>Grade 1</b>	بالكاد مفهوم	
<b>Grade 0</b>	غير مفهوم تماما	

### Pre-edited text n°05

Dialogue between cultures is the main solution for the development of human civilization. Over centuries, cultures have been mutually enriched through reciprocal understanding and interpenetration. Those mutually enriched cultures formed human civilization in an artistic way. This dialogue between cultures can be the only answer to the growing danger of various manifestations of intolerance and violence, including aggressive nationalism. This is the essence of the “Dialogue between Cultures “ project, which was proposed in the UNESCO and has the support of cultural figures in various countries.

### Translation n°05

يشكل الحوار بين الثقافات الحل الرئيسي لتنمية الحضارة الانسانية. وعلى مدى قرون، اصبحت الثقافات تنعم بالاثراء من خلال التفاهم المتبادل والتداخل المتبادل. وقد شكلت تلك الثقافات ذات الاثراء الحضارة الانسانية بطريقة فنية. وقد يكون هذا الحوار بين الثقافات هو الرد الوحيد على الخطر المتنامي لمختلف مظاهر التعصب والعنف، بما في ذلك القومية النزعية. هذا هو جوهر مشروع "الحوار بين الثقافات" الذي تم تشكيله في منظمة اليونسكو، وهو يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Pre-edited text n°06

Dialogue between different cultures was and remains the primary road for the development of human civilization through time. Through the mutual understanding and interpenetration of cultures over centuries, those cultures have been mutually enriched. Thus, they created the unique mosaic of human civilization. It is this dialogue between cultures which must be the answer to the problem of the growing danger of multiple manifestations of intolerance, violence, and aggressive nationalism. This is the core of the "Dialogue between Cultures" project which was proposed to UNESCO and has the support of multiple cultural figures in different countries.

### Translation n°06

وكان الحوار بين الثقافات المختلفة هو الطريق الرئيسي لتنمية الحضارة الانسانية عبر الزمن. فمن خلال التفاهم المتبادل والتداخل بين الثقافات عبر القرون، اصبحت تلك الثقافات متأثرة ببعضها البعض. وبالتالي، فقد خلقوا الفسيفساء الفريدة للحضارة الانسانية. هذا هو الحوار بين الثقافات الذي يجب ان يكون الاجابة على مشكلة الخطر المتزايد من مظاهر التعصب والعنف والقومية العدوانية المتعددة. وهذا هو جوهر مشروع "بين الحوار الثقافات" الذي اقترح على اليونسكو، وهو يحظى بدعم شخصيات ثقافية متعددة في بلدان مختلفة.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	مفهوم جدا	
<b>Grade 2</b>	مفهوم الى حد ما	
<b>Grade 1</b>	بالكاد مفهوم	
<b>Grade 0</b>	غير مفهوم تماما	



### Pre-edited text n°7

Dialogue between cultures is the most important way of human civilisation. Through the mutual understanding of cultures over time, those cultures have been simultaneously enriched and have formed the human civilisation unique. It is this discourse between cultures which should be the answer to the growing danger of various phenomena of intolerance and violence. This is the essence of the “ Dialogue between Cultures” project which was proposed in UNESCO and has the support of cultural figures in different countries.

### Translation n°07

الحوار بين الثقافات هو اهم طريقة للحضارة البشرية. وبفضل الفهم المتبادل للثقافات مع مرور الزمن، تم اثراء تلك الثقافات في نفس الوقت، وشكلت الحضارة البشرية الفريدة. فهذا الخطاب بين الثقافات هو الذي ينبغي ان يكون الرد على الخطر المتزايد لمختلف مظاهر التعصب والعنف. - وهذا هو جوهر مشروع " الحوار بين الثقافات" الذي اقترح في اليونسكو، وهو يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

### Translation of the original text (without pre-editing)

وكان الحوار بين الثقافات هو الطريق الرئيسي لتنمية الحضارة الانسانية ولا يزال. فمن خلال التفاهم المتبادل والتداخل بين الثقافات على مر القرون والالاف، كانت تلك الثقافات تتكاثر بشكل متبادل، كما انها تشكل الفسيفساء الفريدة للحضارة الانسانية. هذا هو الحوار الذي يمكن ويجب ان يكون الرد على الخطر المتنامي لمختلف مظاهر التعصب والعنف، بما في ذلك القومية العدوانية. وهذا هو جوهر مشروع "الحوار بين الثقافات" الذي اقترح في اليونسكو، وهو يحظى بدعم الشخصيات الثقافية في مختلف البلدان.

<b>Grade 3</b>	<b>Completely faithful</b>	
<b>Grade 2</b>	<b>Fairly faithful</b>	
<b>Grade 1</b>	<b>Barely faithful</b>	
<b>Grade 0</b>	<b>Completely unfaithful</b>	

<b>Grade 3</b>	<b>مفهوم جدا</b>	
<b>Grade 2</b>	<b>مفهوم الى حد ما</b>	
<b>Grade 1</b>	<b>بالكاد مفهوم</b>	
<b>Grade 0</b>	<b>غير مفهوم تماما</b>	

## Appendix F

### Results of Pearson Product-Moment Correlation for Pre-editing and Machine Translation

Correlations

		preedited	Translation
preedited	Pearson Correlation	1	,831**
	Sig. (2-tailed)		,000
	N	21	21
translation	Pearson Correlation	,831**	1
	Sig. (2-tailed)	,000	
	N	21	21

\*\* . Correlation is significant at the 0.01 level (2-tailed).

## Résumé

Dans une société à la fine pointe de la technologie et à la mondialisation avec des centaines de langues, il y a la traduction automatique, cette dernière, en tant que domaine interdisciplinaire, est devenue une partie intégrante de l'apprentissage des langues étrangères. Sa qualité, cependant, dépend largement de la façon dont on l'utilise. Par conséquent, cette étude vise principalement à étudier l'effet de la pré-édition sur le résultat final de la traduction automatique. Nous supposons que la pré-édition a un effet positif sur le résultat final de la traduction automatique ; plus précisément, il existe une relation corrélationnelle entre la qualité de pré-édition et la qualité de traduction automatique. Pour vérifier les hypothèses susmentionnées, les chercheurs ont mené une étude expérimentale avec un ensemble de 21 étudiants de troisième année à l'Université de Mohamed Seddik Ben Yahia. Cette étude a été fondée sur une approche mixte, car les données ont été recueillies et analysées qualitativement et quantitativement. Les résultats obtenus étaient très intéressants : ils ont montré que la plupart des étudiants ont réussi à appliquer les règles de pré-édition et ont fourni des textes bien simplifiés. Puis, ils corroboraient l'effet positif de l'approche de pré-édition sur l'intelligibilité et la fidélité de la traduction automatique. Enfin, ils ont révélé qu'il existe une relation significative entre la qualité de pré-édition et la qualité de traduction automatique. Pour conclure, la présente étude fournit quelques recommandations utiles et souligne la nécessité d'enseigner aux élèves le concept de pré-édition pour une meilleure qualité de traduction.

## ملخص

في مجتمع يتمتع بأحدث التقنيات وعالم يتحدث مئات اللغات، أصبحت الترجمة الآلية، كمجال متعدد التخصصات، جزءاً لا يتجزأ من تعلم اللغة الأجنبية. ومع ذلك، فإن درجة جودتها تعتمد إلى حد كبير على كيفية استخدامها. انطلاقاً مما سبق، تهدف هذه الدراسة العلمية في المقام الأول إلى دراسة تأثير "التعديل المسبق" للنص المراد ترجمته على النسخة النهائية للترجمة الآلية. وفقاً لذلك، يُفترض أن "التعديل المسبق" له تأثير إيجابي على النتائج النهائية للترجمة الآلية؛ بتعبير أدق، هناك علاقة بين جودة التعديل المسبق للنص وجودة الترجمة الآلية. للتحقق من الفرضيات المذكورة أعلاه، أجرت الباحثان دراسة تجريبية على عينة تتكون من 21 طالباً في السنة الثالثة لغة إنجليزية في جامعة محمد صديق بن يحيى. اعتمدت هذه الدراسة على تصميم مقارنة مدمجة، إذ تم جمع البيانات وتحليلها من حيث النوعية والكمية. كانت النتائج المتحصل عليها مثيرة للاهتمام. فقد أظهرت النتائج أن معظم الطلاب كانوا موفقين في تطبيق القواعد الخاصة بتعديل النص الأصلي وقدموا نصوص مبسطة بشكل جيد. ثم أثبتت النتائج التأثير الإيجابي لتعديل النص على وضوح الترجمة الآلية ودقتها. أخيراً، كشفت النتائج أن هناك علاقة طردية بين جودة التعديل المسبق للنص وجودة الترجمة الآلية. في الختام، تقدم الدراسة الحالية بعض التوصيات المفيدة وتشير إلى ضرورة تعليم الطلاب كيفية تبسيط النص الأصلي للحصول على ترجمة ذات جودة أفضل.